FORECASTING SEASONAL STREAMFLOW/PRECIPITATION VARIABILITY ON THE PEACE RIVER AT ARCADIA, FL, CONDITIONED UPON PACIFIC-ATLANTIC SEA SURFACE TEMPERATURES

By

DAVID COLEY

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2003

ACKNOWLEDGMENTS

I would first and foremost like to thank my committee members Dr. Peter Waylen, Dr. Tim Fik and Dr. Paul Zwick for all their support throughout the last four years. Their insight, advice and encouragement have been invaluable to me throughout the writing process, while also supporting my academic and career decisions. I am further grateful to Dr. Bon Dewitt of the Geomatics Department for his insightful comments.

I would also like to thank my wife Cheryl, as without her support and encouragement I would not be where I am today. Finally, my continued and former colleagues Mat Marsik, Jose Torres, Jim Penn, Jim Sloan and David Jenkins all provided ideas, encouragement and insight over the last few years.

Finally, I am grateful to my mother for her support, both financial and emotional.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

FORECASTING SEASONAL STREAMFLOW/PRECIPITATION VARIABILITY ON THE PEACE RIVER AT ARCADIA, FL, CONDITIONED UPON PACIFIC-ATLANTIC SEA SURFACE TEMPERATURES

By

David Coley

August 2003

Chair: Peter Waylen
Major Department: Geography

The Peace River at Arcadia, FL, is in use as a supplemental municipal water supply option for southwest Florida counties in the 21st century. As such, probabilities of encountering very low flows/precipitation levels during the dry season are of critical importance. The association between sea surface temperatures (SSTs) in the eastern Pacific and seasonal streamflow/precipitation variability in the southeast United States is well documented. Less well known is the effect that North Atlantic sea surface temperature variability may have in modifying this response as previously reported in fluctuations of Lake Okeechobee.

Four forecast models are developed to more accurately assess the probability of streamflow/precipitation characteristics during this season conditioned upon SSTs two-three seasons previously. Forecast models presented include lognormal/normal probability distributions, an ordinary least squares multiple regression, three-way contingency tables and a loglinear regression analysis. Seasonal variables examined

include mean, minimum and maximum flows, mean basin-wide precipitation and the number of low-flow days during the dry seasons.

Seasonal flow/precipitation extracted from the historic record is sorted into various subpopulations based on above/below median SSTs from the previous summer. The mean and variance of each subpopulation are calculated and fitted to separate two-parameter lognormal/normal distributions. Probability estimates and return periods for various flows/precipitation levels are then calculated.

The seasonal mean streamflow variable response to antecedent seasonal flow conditions and sea-surface temperature anomalies (SSTAs) is modeled with the ordinary least squares approach. SSTAs are defined as one standard deviation above/below long-term mean. Partial contingency tables are used to construct odds ratios for the dry seasons where there are more or less than the long-term mean of low flow days, conditioned upon sea-surface temperatures. Finally, the number of low-flow days is extracted from the historic record and modeled according to SST categories using a loglinear regression model.

Results indicate that the probabilities of flows at various return periods during the low flow season are strongly conditioned upon Niño 3.4 SSTs during the previous summer. The Tropical North Atlantic appears influential during the spring season. The multiple regression models essentially confirm the streamflow probability distribution results. Finally, the contingency tables and loglinear analysis suggest that the odds ratios and number of low-flow days per winter and spring season are independent of Tropical North Atlantic sea-surface temperatures.

CHAPTER 1
INTRODUCTION

Global Water Issues

Water is an essential ingredient for life. As we continue through the new millennium, water resources continue to occupy a great area of concern for scientists, water resource professionals and policy makers. Primary global issues include water shortages, supply, floods, quality, and allocation. Many articles and news reports over the last several years have portrayed water shortages as perhaps the most pressing issue facing the world today, with many countries facing severe economic and health problems resulting in part from too little water. Other countries face similar problems caused by severe flooding. Because water is not evenly distributed over space, water and water resource issues can be uniquely studied from a geographic perspective.

Water can no longer be taken for granted or regarded as a free commodity. It must be regarded as a resource to be carefully managed, where long term planning is essential to effective management of a dwindling resource.

Florida Water Concerns

In the state of Florida, we are blessed with a favorable geography that creates a water rich environment. The state receives an average annual rainfall of 53 inches (Fernald and Purdum, 1995). As such, water supply has not been a problem except perhaps during times of drought throughout most of Florida history. Florida's population grew by 3 million in the 1990's (23.5%), and is expected to increase to 19 million by 2010 (Dzurik, 2003). Statewide use of water has increased from 2.9 billion gallons a day

(bgd) in 1950 to 7.5 bgd in 1995 (Fernald and Purdum, 1998). Since the state will continue to see such tremendous population growth, water supply and growth management issues are perhaps the number one long-term issues facing the state.

To address supply and other water issues, in 1972 the state legislature enacted the Water Resources Act (Dzurik, 2003). Among its may broad areas, the Act created regulatory bodies known as Water Management Districts to oversee water supply, demand and quality issues. Among many duties that include consumptive use permitting, the districts also establish minimum flow levels for ground and surface water. Consumptive use permits are generally granted to agricultural concerns and to public supply utilities. Usually, they are permitted to withdraw groundwater from well fields. However, as groundwater levels continue to decline in some areas of the state, many public supply utilities are forced to search for alternative supply options, including the use of surface water (Thomsen, cited in Treat, 1998).

Problem Statement and Research Justification

The Peace River at Arcadia, FL is in use as a supplemental municipal water supply option for southwest Florida counties in the 21st century. Known as the Peace River Option (PRO), this plan allows utilities, permitted through the Southwest Florida Water Management District (SWFWMD), to withdraw water from the river and deliver that water to potential customers through an aquifer storage and recovery process (Thomsen, cited in Treat, 1998). The PRO also defines annual average, peak monthly and maximum daily withdrawal amounts and has set a threshold of 3.68 m^3s^{-1} (130 ft^3s^{-1}) below which no surface water withdraw is permitted (Thomsen, cited in Treat, 1998). This level was set following studies by SWFWMD biologists and other scientists to help protect water

quality in the stream and maintain healthy riverine/estuarine ecosystems (Thomsen, cited in Treat, 1998).

Estimation of the risks of flows dropping below various levels during the dry seasons is of critical importance. Because mean and maximum seasonal flows result primarily from seasonal precipitation, it is of equal interest to estimate the probability of this variable during the same or preceding seasons.

As all hydro-meteorological processes such as streamflow and precipitation are inherently random phenomena, obeying probabilistic rather than deterministic laws, it is often not possible or desirable to formulate a deterministic model that can forecast these variables. As a result, the water resource analyst will rely on probabilistic models in order to produce a forecast of the hydrologic variable of interest that lies within a statistically acceptable or significant range of values. Such forecasts usually rely on the historic record, so it is of critical importance to capture the underlying probabilistic properties of that record. Once correctly identified, the analyst may proceed with (among other procedures) generating probability estimates of future flow events thereby allowing water resource professionals to implement future water use policies. The justification for the development of streamflow forecast models for the Peace River should be clear, especially given the population increases that are expected for the state and region through the 21st century.

Thesis Hypothesis and Research Goals

Based on evidence from the literature, there is the potential to incorporate raw sea-surface temperature data from the Equatorial Pacific and tropical North Atlantic as predictors of streamflow and precipitation in southwest Florida. Many studies of these variables use a variety of sea-surface temperature/atmospheric anomaly indices, which

leaves room for subjective interpretation of results. This study will develop streamflow and precipitation forecast models using readily accessible sea-surface temperature data with an appropriate temporal lag. The models will be probabilistic in nature to incorporate the inherent uncertainties involved in the ocean/atmosphere interactions, and to provide measures of risk (and uncertainties) associated with low streamflow and precipitation during the dry seasons. Upon completion, the methods and models will then be available for evaluation by interests such as the agricultural industry, municipal water supply planners, and other potential users. Specific goals of this research include the following:

- Determine the statistical characteristics of seasonal streamflow and precipitation on the Peace River at Arcadia, conditioned upon sea-surface temperature anomalies from the Niño 3.4 region of the Eastern Pacific for the period of study, 1951-2000 (n = 50).
- Streamflow characteristics to be examined include minimum, maximum and mean seasonal discharge rates. Minimum flow is defined as the lowest flow each season, mean flow as the average flow each season, and maximum flow as the largest flow each season.
- Precipitation characteristics to be examined include mean seasonal rainfall.
- Determine whether SSTs in the tropical North Atlantic have a significant effect in either enhancing or diminishing the seasonal impacts that ENSO has on precipitation and streamflow for the same stream gage and region.
- Develop seasonal forecast models for streamflow at Arcadia, FL and basin precipitation for the contributing drainage area.
- Forecast the number of low-flow days expected for the dry seasons conditioned upon Niño 3.4 and tropical North Atlantic SSTs.

Study Area Description

The Peace River at Arcadia, FL, drains a total area of 3,372km^2 in west-central Florida, with a stream length of approximating 166km (fig 1.1).

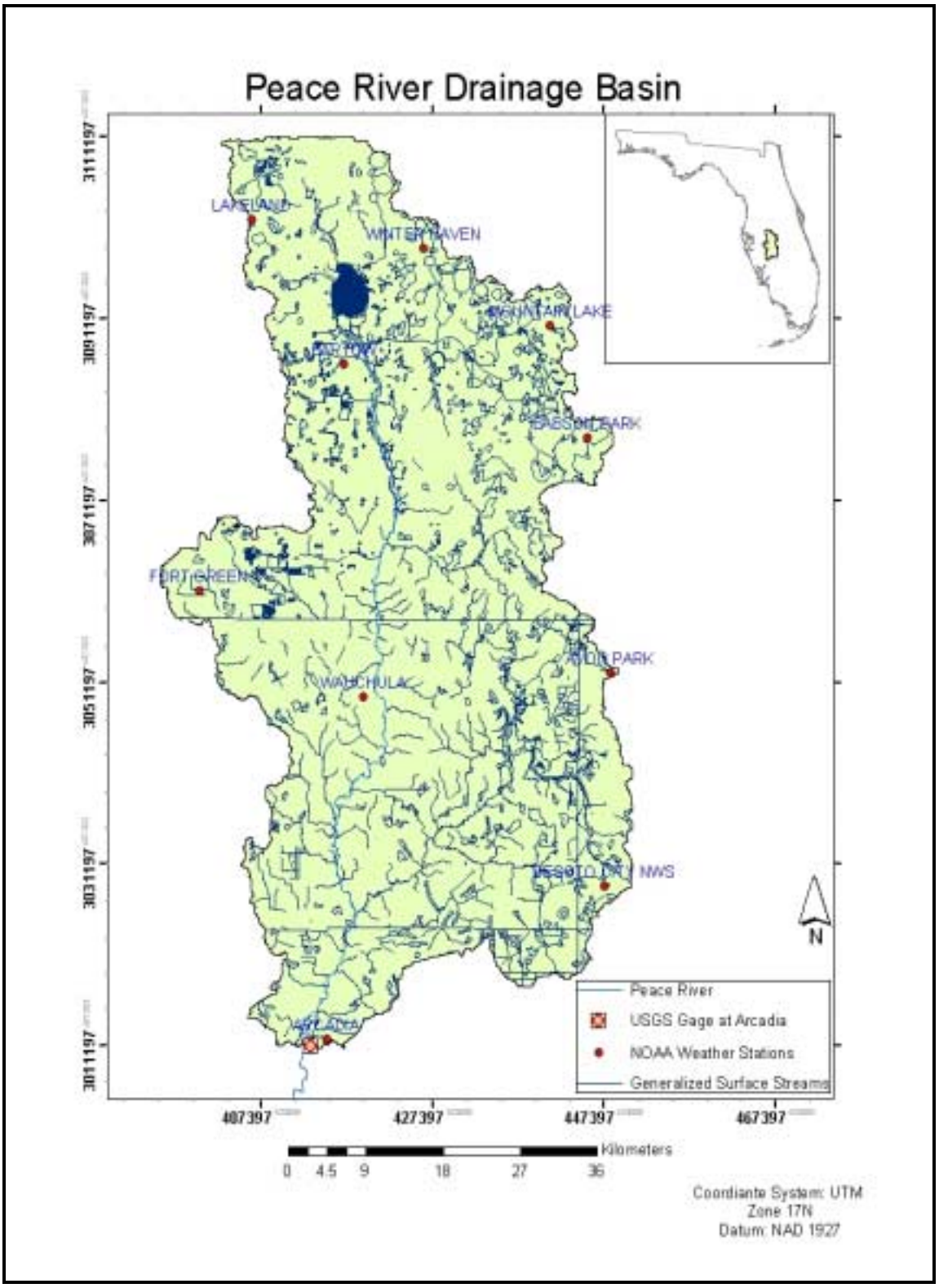

Figure 1.1. Study area of the Peace River drainage basin.

The headwaters begin in Polk County among a group of lakes at the confluence of Saddle Creek and the Peace Creek drainage canal (Fernald and Purdum, 1998), with an average channel slope is of .4m/km.

The river passes through three distinct physiographic provinces, the Polk Upland, DeSoto Plain and Gulf Coastal Lowlands before emptying south into Charlotte Harbor (Randazzo and Jones, 1997). These three provinces define the upper, middle, and lower Peace River basins, respectively. The drainage in the upper basin is internal, characterized by wetlands and depression storage. The upper basin is also an area of negative groundwater pressure head; as such the upper Peace is at times a losing stream, especially during the dry season. The middle and lower portions exhibit a well-defined surface drainage with many incised tributaries, where the stream receives baseflow from the surficial and intermediate aquifers (Lewelling, Tihansky, and Kindinger, 1998). Overall flow regime is characterized in figure 2.1.

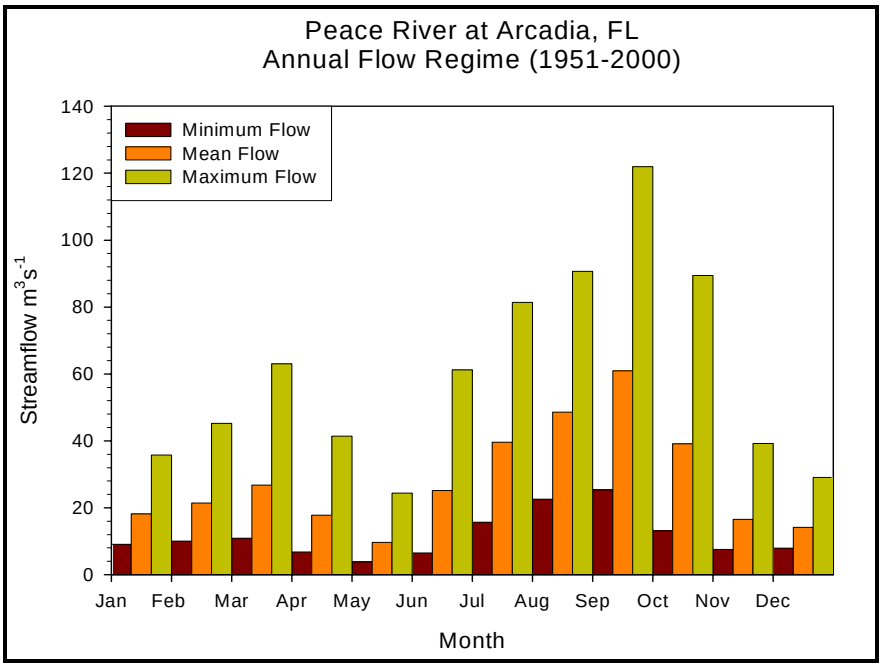

Figure 1.2. Flow regime of the Peace River at Arcadia, FL.

The Peace River is characterized by a wet (June-October) and dry (November-May) season flow regime similar to many streams located in the Florida peninsula. Peak wet season discharge in September is generated by frequent summertime convective activity and the infrequent passage of tropical systems. The winter/spring dry season peak in March is almost exclusively generated by the passage of mid-latitude cyclones (Fernald and Purdum, 1998). Overall, drainage basin climate is considered subtropical and humid. Soils in the basin can be classified as well to moderately drained Entisols and Spodosols (Fernald and Purdum, 1995). Because most of the land use in the basin is rural, and natural surface sediment is porous and well sorted, average annual runoff is generally low, about 304mm yr^{-1} (11.96in yr^{-1}) (Fernald and Purdum, 1998).

Other Applications

Other interests that can utilize reduced uncertainty in streamflow forecasts include recreation and tourism, since higher winter flows are associated with cool, cloudy and rainy conditions, which in turn can dampen the winter tourist season. Ecologists would be interested in both high and low-flows (and their timing) to the Charlotte Harbor estuary, since the magnitude and timing of flows can impact fisheries, for example. Additionally, the insurance industry and emergency management services could use the reduced uncertainty in flood forecasts during winter and spring seasons when such events are infrequent as compared to summer season floods.

Thesis Organization

The thesis is organized according to university guidelines. The introductory chapter is followed by a review of pertinent literature, and a description of the data and relevant statistical models used. The results chapter follows specific methods used to achieve thesis objectives. Finally, a chapter discussing result implications, variable

selection and models justification is presented. A summary and conclusions chapter concludes the thesis.

CHAPTER 2
LITERATURE REVIEW

Hydro-meteorologic Forecasting Concepts

Hydro-meteorological processes such as streamflow and precipitation are inherently random phenomena that obey probabilistic rather than deterministic laws, and can therefore only be explained in a probabilistic sense (Viessman et al., 1989). Where a deterministic forecast provides a single, a probabilistic forecast specifies a probability distribution function for all values of the variable of concern (Krzysztofowicz, 2001). Concepts of probability and probability distributions can be found in most comprehensive statistical and hydrology texts (i.e., Viessman et al., 1989; Maidment, 1993; Burt and Barber, 1996), and will be addressed in the next chapter. Forecasting hydro-meteorologic events such as floods and droughts can thus be described as exercises in defining uncertainties about the underlying physical processes that create these events.

The goal of all forecasts is to reduce uncertainty of future events. Krzysztofowicz (2001) suggests that forecasting hydrologic variables and extreme events should be probabilistic to incorporate the random nature of the variable to be estimated. Perhaps more importantly, a probabilistic forecast also integrates the limits of our own knowledge about the variable of interest. As an illustration of the potential dangers on a deterministic forecast made from operational hydrology, he cites a 1997 flood event on the Red River in Grand Forks, North Dakota. In this instance, a crest of 49 feet (ft) above datum (i.e., a single estimate) was predicted. When the river crested at 54 ft, the town was unprepared for an event of that magnitude, as the forecast included no assessment of accuracy or

uncertainty. Krzysztofowicz (2001) advocates that a deterministic forecast creates an illusion of certainty in the mind of the user, and forces the forecaster to suppress information about uncertainty.

Hirsch et al. (cited in Maidment, 1993) illustrate three potential inhibitors to correctly predicting a hydrological variable in a probabilistic sense. The first of these is the inherent randomness in the hydrologic system itself: including precipitation, topography, soil characteristics, and underlying aquifers. The second obstacle is sampling error. This involves both the spatial and temporal characteristics of the data itself, where hydrologists are usually working with small samples that are part of a potentially infinite population. That sampling error exists reflects a lack of information. For example, precipitation is only collected at certain locations within a watershed for a limited time. As the breadth and length of sampling increases, so should forecast accuracy. Other obstacles include assuming that the variable of interest is from a linear system when the system is non-linear. Finally, an incorrect understanding of the physical processes involved can further inhibit forecast accuracy.

Once the analyst has correctly addressed these concerns, he/she can then rely on probabilistic models in order to produce an operational forecast of the hydrologic variable of interest. Because probabilistic forecasting of streamflow usually relies on the historic streamflow record, forecast accuracy is dependent upon the underlying probabilistic properties. Upon identification, the analyst may proceed with generating probability estimates (among other procedures) of future flow events. It then falls to the user as to the best way in which to utilize this information, such as implementing water use

restrictions during times of drought to the allocation of emergency management services preceding flood events.

All historic streamflow records contain a wealth of information about the climatic variability that gave rise to them. If the forecaster can identify other physical variables (such as oceanic or atmospheric processes) that exhibit a relationship with the variable of interest, that information can provide a basis for a useful forecast (Sharma, 2000). For Florida streams, sea surface temperatures in the eastern Pacific have been identified as a primary environmental factor associated with flow characteristics. This link will be addressed later in this chapter.

ENSO Concepts, Teleconnections and Forecast Potential

ENSO Concepts

The Southern Oscillation (SO) is usually defined as the seesaw in atmospheric pressures at sea level (SLP) between Tahiti and Darwin, while an El Niño/La Niña event is the accompanying oceanic response to the SO The acronym ENSO thus describes both the atmospheric and oceanic components of the phenomenon (Caviedes, 2001). An El Niño (or warm ENSO) event is associated with the negative phase of the SO, (lower than normal pressure at Tahiti, higher than normal pressure at Darwin), which through a relaxation of the Trade winds along the equator allows the large pool of warm water surrounding Indonesia to advect eastward, creating positive sea surface temperature anomalies (SSTAs) across vast regions of the equatorial Pacific (Kahya and Dracup, 1993). The opposite atmospheric-oceanic condition describes a positive, or cold ENSO event. Once in place, both events can produce significant seasonal to annual precipitation anomalies, depending on the strength and duration of the event, for many regions

worldwide. A typical event spans some 18 months with a frequency of 3-7 years, although a 2-10 year frequency is not uncommon (Kahya and Dracup, 1993).

Teleconnections

The concept that distant regions are affected by the ENSO phenomenon is defined as atmospheric teleconnections, where teleconnections are disruptions in large-scale atmospheric circulation patterns. These large-scale disruptions in turn affect regional and local surface climate patterns (Kahya and Dracup, 1993). During the mature phase of an El Niño event, the Northern Hemisphere sub-tropical jet stream is displaced equatorward over the eastern Pacific. The jet then advects anomalously moist air masses over several regions of the United States (including the Florida peninsula) during winter, creating a greater probability of experiencing increased precipitation over much of peninsular Florida. However, the likelihood of increased precipitation diminishes in the southern portion of the state (Kahya and Dracup, 1993; Gershunov and Barnett, 1998). During a positive ENSO phase, both the polar and subtropical jet streams are displaced poleward. As a result, the dry, stable atmosphere in place over the eastern Pacific is advected north of the southeast United States, in turn creating dryer than normal winter conditions over the Florida peninsula and southeast United States (Gershunov and Barnett, 1998).

Forecast Potential

Rasmusson and Carpenter (1982) and Ropelewski and Halpert (1987) provide a thorough account of the atmospheric-oceanic component of ENSO. Rasmusson and Carpenter (1982) present a detailed description of the evolution of sea surface temperatures, sea level pressures, wind field anomalies and accompanying precipitation anomalies across the entire Pacific basin over the duration of an El Niño event. This work also increased the available knowledge of the amplitude, duration, and periodicity

of the phenomenon, where the authors describe the fluctuations in all three components (Rasmusson and Carpenter 1982).

Ropelewski and Halpert (1987) provide an improved analysis of that by Rasmusson and Carpenter (1982). Employing first harmonic vectors extracted from a 24-month composite of precipitation values, Ropelewski and Halpert (1987) identify the seasons during which ENSO-related precipitation anomalies occur across the globe, again including the southeast United States. Most importantly Ropelewski and Halpert (1987) define the strength, or consistency/coherence, of the precipitation relationship with ENSO. The authors find that Florida is affected during the mature phase, which corresponds to late winter-early spring. The strength of this relationship is further quantified in a later study (Ropelewski and Halpert, 1996). Thus, these studies made it possible for other researchers to employ SSTAs and other variables (e.g. SLP or an index created with the SLP differences at Tahiti and Darwin, otherwise known as the Southern Oscillation Index, or SOI) from the eastern Pacific as a predictor for hydrologic variables in various regions around the world.

Other studies quantify precipitation/streamflow and ENSO relationships at interseasonal to interannual timescales for various regions. Simpson and Colodner (1999) detail ENSO influences on seasonal precipitation in Arizona. Gutierrez and Dracup (2001) explore the use of ENSO as a forecast tool in Colombia in support of the hydroelectric power industry. Chiew et al. (1988, 2003) use ENSO to forecast spring and summer season precipitation/streamflow in Australia. Harshburger et al. (2002) examine the ENSO relationship between winter snowfall/spring runoff in the state of Idaho. While employing different methods, these studies all broadly agree that by correctly

identifying the timing of ENSO influence on seasonal precipitation/streamflow for a specific region, it is possible to incorporate that influence in a lagged fashion, thereby increasing (decreasing) forecast accuracy (uncertainty).

The association between ENSO and seasonal precipitation/streamflow variability in the southeast United States is also well reported. In an analysis of temperature and precipitation anomalies across the U.S., Gershunov and Barnett (1998) find that the frequency of winter (DJF) season heavy rainfall events increase by 15-30% during the negative ENSO phase for states that border the Gulf of Mexico. They also report a decrease of 30-50% for extreme rainfall events during a positive ENSO phase for the same region (Gershunov and Barnett, 1998).

Kahya and Dracup (1993) identify the Gulf of Mexico region as one that exhibits a consistent ENSO signature in seasonal streamflow anomalies. The use of streamflow to assess ENSO influence has advantages over precipitation in that streamflow integrates all the hydrologic processes for a given region. By way of first harmonic vectors and an aggregate streamflow composite, Kahya and Dracup (1993) report up to a 50% increase in seasonal (December0-April+1, where (0, +1) indicates a mature phase) streamflow magnitude for the Gulf region during a warm ENSO event. The timing and coherence of their findings are in good agreement with Ropelewski and Halpert (1987) and Ropelewski and Halpert (1996).

Forecast potential for Florida streams has been reported in the literature as well. Employing a lognormal distribution function, Zorn and Waylen (1997) show a negative correlation between winter monthly discharge on the Santa Fe River when monthly flows are conditioned upon the annual ENSO phase (i.e., seasonal flow conditioned upon an

annual warm or cold phase as defined by the Japanese Meteorologic Administration). Sun and Furbish (1997) calculated cross-correlations between SSTs in the eastern Pacific and streamflow/precipitation gauges in Florida. Their results suggest that ENSO is responsible for 30% of the annual streamflow variation and 40% of annual precipitation variation. Schmidt, Lipp, Rose and Luther (2001) detail similar ENSO influences on other Florida streams, all suggesting a strong localized correlation between ENSO and winter seasonal precipitation/discharge.

Tropical North Atlantic Sea Surface Temperatures: Linkages to the Northern Hemisphere Oscillations and ENSO Forcing

Less well known is the effect that tropical North Atlantic (TNA) sea-surface temperatures (SSTs) may have on enhancing or diminishing ENSO relationships with Florida streamflow/precipitation.

Tropical North Atlantic Sea Surface Temperature Fluctuation

A suggested causal mechanism responsible for anomalously warm (cold) tropical North Atlantic SSTs is the North Atlantic Oscillation (NAO) (Hurrell, 1995; Marshall et al. 2001; Wanner et al. 2001), and with ENSO forced atmospheric disturbances (Covey and Hastenrath, 1978; Rogers, 1984; Enfield, 1996; Enfield and Mayer, 1997; Mo and Hakkinen, 2000) which in turn affect regional season precipitation patterns.

The NAO is described as a large-scale meridional oscillation of atmospheric air mass between the Azores high and Icelandic low (Wanner et al., 2001). During the high (low) phase of the NAO, the Azores high is stronger (weaker) than usual, creating stronger (weaker) winds about the center of circulation. This in turn generates stronger (weaker) trades and westerlies, generating a greater (lesser) heat flux at the surface, creating cooler (warmer) than usual SSTs in the tropical North Atlantic (Giannini et al.,

2001). The high phase of the NAO during winter is further associated with anomalous southerly flow and decreased precipitation across the eastern United States (Hurrell 1995; Bucha and Bucha, 2002). While it is the only oscillation present in all seasons, it is most pronounced in amplitude and areal extent during winter (Marshall et al., 2001). Hereafter, NAO fluctuation will imply winter season change.

The duration of the high/low NAO phases exhibits interdecadal fluctuations (Marshall et al., 2001). From about 1900 to 1930, the NAO was high. From the early 1940s to the early 1970s, the NAO exhibited a downward trend, where the amplitude of the pressure differences between the Azores and Iceland was less pronounced (Marshall et al., 2001). Since the 1980s, the NAO has again returned to a positive phase, contributing to warmer surface conditions in the Northern Hemisphere in the last 20 years (Wanner et al., 2001). The positive phase also leads to overall dryer conditions via stronger westerlies creating a more zonal track of winter cyclone activity (Marshall et al., 2001).

It is therefore likely that the ENSO influence in winter over Florida and the southeast United States is moderated by the high/low winter phase of the NAO at the interdecadal scale, as reported by Enfield, Mestas-Nunez and Trimble (2001). These authors found that most of the country experiences less than normal rainfall during the warm phase of the Atlantic Multidecadal Oscillation (AMO), regardless of ENSO phase. The AMO is described as a long-term oceanic phenomenon with a 65-80 year cycle exhibiting a 0.4^{o} C range over the entire North Atlantic (Enfield et al., 2001). Because the duration of the AMO is so lengthy, it may have little value as a predictor of seasonal streamflow/precipitation.

Nevertheless, it is important to grasp the atmospheric effects of the AMO as described by Enfield et al. (2001). They employed a ten-year running mean of the Kaplan et al. (cited in Enfield et al., 2001) global SSTAs and correlated those values with rainfall over the Mississippi basin and south-central Florida. The south-central Florida peninsula was wetter during El Niño years regardless of the AMO phase, during a 30-year period from 1930-1959 in which the entire North Atlantic exhibited above-normal SSTs (Enfield et al., 2001). As explanation, the authors suggest that this period is related to a disruption of the atmospheric ridge-trough pressures associated with the Pacific North American (PNA) pattern (Coleman and Rogers, 2002) across the entire southern tier of the United States, resulting in less winter cyclonic activity (Enfield et al., 2001).

It should be further noted that the use of TNA SSTs as a streamflow/precipitation predictor might be merely a proxy of ENSO. That is, many studies (Enfield, 1996; Enfield and Mayer, 1997; Mo and Hakkinen, 2000; Marshall et al., 2001; Czaja and Frankignoul, 2002; Czaja et al., 2002) have found that ENSO forces TNA SSTAs mainly through modifications in Trade wind strength. Most agree that the TNA SSTs respond to the eastern Pacific SSTs with a lag of 4-5 months. Therefore, any suspected impact of summer TNA SSTs on Florida precipitation/streamflow may reflect an extended ENSO condition from the previous winter.

In considering the isolated influence of the tropical North Atlantic, Enfield (1996) demonstrated that an abnormally warm TNA is associated with significant rainfall departures over the southeast United States, excluding Florida. Although this study suggests a strong link to annual precipitation anomalies in the southeast, the coarseness of the data make it unclear if Florida precipitation is affected. Conversely, Chen and Taylor

(2002) show that lagged warm (cold) ENSO induced TNA SST anomalies favor (inhibit) early season (AMJ) easterly wave development for the Caribbean region, enhancing (diminishing) early season rainfall south of 20^0 north latitude. The authors conclude that although the atmosphere is favorable for easterly wave development over the Caribbean basin including Florida, their model simulations produced positive rainfall departures only south of 20^0 north latitude. This study suggests that spring season TNA SSTAs have no significant effect on spring season rainfall over Florida.

Finally, the physical mechanisms that would create a seasonally lagged relationship (similar to ENSO) between summer season TNA SSTs and winter/spring streamflow/precipitation anomalies in the southeast United States have not been clearly identified in the literature (Wanner et al., 2001).

Summary

In summary, long-term forecasts must incorporate the stochastic nature of hydro-meteorologic variables. Secondly, long term seasonal forecasting should be probabilistic in nature, because it captures the random nature of hydrologic events while allowing the forecaster to provide an error assessment. Third, it is clear that ENSO produces significant departures in seasonal precipitation/streamflow in the southeast U.S. and Florida. As such, summer season SSTAs in various Niño equatorial regions (Niño 4, Niño 3, or Niño 3.4) have shown considerable reliability as predictors because they reflect large-scale atmospheric disturbances that in turn drive regional climate behavior. Fourth, it has not been unambiguously resolved whether summer TNA SSTs, either in isolation or in concurrence with ENSO, possess the same reliability. It may be likely that any influence of the TNA on Florida streamflow/precipitation may be felt simultaneously, or at considerably shorter lags. Lastly, because the strength of the

relationship between Florida precipitation/streamflow is in doubt, this variable will be included as a forecast predictor for the Peace River at Arcadia, FL.

CHAPTER 3
DATA AND RELEVANT MODELS DESCRIPTIONS

Data Description

The data requirement for this study is as follows. Continuous daily mean streamflow data from the United States Geological Survey, USGS Station 02296750 Peace River at Arcadia, FL for the period 1951-2000. Mean monthly precipitation data from the National Oceanic and Atmospheric Administration (NOAA) for the following stations, 1957-1997:

Table 3.1. NOAA weather stations for the Peace River drainage basin.

NOAA ID #	Site Name	Lat	Lon
9707	Winter Haven	28^0 00' 55.06"	81^0 44' 56.26"
4797	Lakeland	28^0 02' 34.06"	81^0 57'12.27"
478	Bartow	27^0 53' 59.08"	81^0 50' 34.27"
N/A	Babson Park	27^0 49' 39.08"	81^0 33' 14.25"
369	Avon Park	27^0 35' 40.11"	81^0 31' 33.25"
228	Arcadia	27^0 13' 44.17"	81^0 51' 27.28"
5973	Mountain Lake	27^0 56' 19.07"	81^0 35' 57.25"
2288	DeSoto City	27^0 22' 56.14"	81^0 31' 52.25"
N/A	Fort Green	27^0 40' 23.11"	82^0 00' 42.29"
9401	Wauchula	27^0 34' 08.12"	81^0 49' 02.28"

Mean monthly SSTs from the Niño3.4 region of the Eastern Pacific (5^o N-S, 90^o – 150^o W) and the tropical North Atlantic (5^o- 20^o N, 30^o – 60^o W) acquired from the National Climate Prediction Center (NCPC) databases, 1951-2000. Because the data used are lengthy, appendix A provides hyperlinks to the raw data sources for examination.

Relevant Models Description

Lognormal Probability Distribution

The lognormal probability distribution function is a special case of the normal distribution, represented by the following pdf:

$$p_x(x) = (2\pi\, x^2\sigma_y^2)^{-1/2} \exp[-(\ln x - \mu_y)^2 / 2\sigma_y^2] \,,\; x > 0$$

Since many hydrologic processes are positively skewed, the natural logarithm of the continuous random variable x is well described by the normal distribution (Zorn and Waylen, 1996). The two moments of its distribution are the natural logarithms of the mean and variance, μ_x and σ^2, equal to:

$$\mu_x = \exp(\mu_y + \sigma^2_y / 2),\; y = \ln(x)$$

$$\sigma^2 = \mu^2_x[\exp(\sigma^2_y) - 1],\; y = \ln(x)$$

Normality of the historic streamflow record is tested with the Kolmogorov-Smirnoff *D* goodness of fit statistic.

Ordinary Least Squares Linear Regression

Ordinary least squares (OLS) regression model descriptions can be found in many texts on statistics (Berry and Feldman, 1985; Graybill and Iyer, 1994; Burt and Barber, 1996). Multiple linear regression goes beyond describing the strength of the association between two or more variables as with a correlation function. In multiple linear regression, the goal is to explain the variation in the dependent variable Y with our knowledge about the behavior of one or more independent variables X (Burt and Barber, 1996). This is accomplished via a population regression function:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k + \varepsilon$$

where α is the Y intercept term, β's are partial slope estimators that describe the contribution of each independent variable to the overall regression function. ε is the error term (Berry and Feldman, 1985). In practice, the true population parameters are unknown, thus the population parameters are estimated from a sample (of size n) from the population. The sample regression function is usually written as:

$$\hat{y}_j = \alpha + b_1x_{1j} + b_2x_{2j} + \ldots b_kx_{kj} + \varepsilon$$

To minimize the error of the sample regression function and to estimate the parameters α and b, we apply the least squares criterion:

$$\sum_{j=1}^{n} (Y_j - \hat{Y}_j)^2,$$

the result is ordinary least squares regression. The significance test about the contributing power of each independent variable b is the Students t-test. That is, large t statistics and small p-values provide evidence against the H_o: b = 0 (Burt and Barber, 1996). The overall explanatory power of the model is provided by the coefficient of determination r^2 and the F-ratio or F-statistic. The r^2 statistic is the square of the correlation coefficient, *r*:

$$r = \sum (X_j - \bar{X})(Y_j - \bar{Y}) / \sqrt{\sum(X_j - \bar{X})^2} \sqrt{\sum(Y_j - \bar{Y})^2}, \text{ from j = 1 for all n.}$$

Squaring the *r* coefficient creates: $0 < r^2 > 1$. Values above .5 indicate that more of the variation in Y is being explained by the predictor variables, whereas values below .5 mean there is less of the variation in Y is being explained by the regression function (Burt and Barber, 1996). The F-ratio is another measure related to r^2, where the F-ratio is the ratio of the regression sum of squares to the error sum of squares, known to readers familiar with ANOVA. Large F statistics indicate more of the variation in Y is

explained by the model or systematic component rather than the error or random component (Berry and Feldman, 1985). Finally, the standard error of the estimate (SEE) or the root mean square error (RMSE) provide an overall goodness of fit test. Large values indicate a poor model fit.

It should be obvious from the functional forms that the relationship between E(Y) and each independent variable Xj is linear and that the effects of each independent variable is additive. Further assumptions of OLS regression are (Berry and Feldman, 1985):

1. All variables must be at the interval level.
2. The mean value of the error term is not significantly different from zero.
3. The variance of the error term is constant.
4. No autocorrelation among the error terms.
5. Each independent variable is uncorrelated with the error term.
6. No perfect collinearity among the independent variables.
7. The error term is normally distributed.

When these assumptions are met, OLS regression is an effective tool for estimating and even predicting the variance in the dependent variable Y.

Two-Way Contingency Tables

Categorical data consist of frequency counts of observations, and can be ordinal or nominal. The number of days with measurable precipitation or the number of days of streamflow above or below some threshold are examples of hydro-meteorologic variables that might be analyzed in a two or three way contingency table. Typically, in a two-way contingency table the column variable is the response (or dependent) variable and the row variable is the explanatory (or independent) variable (Agresti, 1996).

The probability distributions for contingency tables relate to the sampling method. For example, in a two x two table, observation frequencies are chosen and classified by response variables X and Y. Where

$$p_{ij} = P(X = i, Y = j)$$

is the probability that (X,Y) falls in row *i* and column *j*. The probabilites (p_{ij}) form the joint distribution of X and Y and satisfy the condition that the probabilites in each cell sum to one (Agresti, 1996). The row and column totals form the marginal probabilites and do not differentiate the probability distribution of Y for a given level of X. It is usually more helpful to construct a separate probability distribution for Y at each level of X. Two variables are considered independent if the conditional distributions of Y are the same at each level of X (Agresti, 1996).

Descriptive measures include the difference of proportions and relative risk assessments. The difference of proportions ($p_1 - p_2$) compares the success probabilites in the two rows and is not significantly different from zero if the response variable is independent of group classification (Agresti, 1996). The relative risk then is the ratio of the success probabilities, (p_1 / p_2). Another important measure of association is given by the odds ratio, where the odds of success in row 1 are:

$$odds_1 = p_1 / (1 - p_1)$$

and the odds of success in row 2 are:

$$odds_2 = p_2 / (1 - p_2).$$

In either row, the probability of success is a function of the odds,

$$p = odds / odds + 1.$$

An example may be helpful. If $p_1 = .75$, then the odds of success is $.75 / .25 = 3$ for row one, indicating that the odds of success are 3 times greater than the odds of failure. If the odds of success in row two are the same as in row one, then the odds ratio

will equal one and the two response variables are said to be independent (Agresti, 1996). Tests of independence are the Pearson chi-square statistic (X^2):

$$X^2 = \sum (n_{ij} - \mu_{ij}) / \mu_{ij},$$

where n_{ij} are cell counts and the likelihood-ratio chi-square statistic (G^2):

$$G^2 = 2 \sum n_{ij} \ln(n_{ij} / \mu_{ij}).$$

Both are left-tail tests, that is large values suggest evidence against the H_o: Independence.

Three-Way Contingency Tables

The same concepts and tests of association detailed above also apply in a three-way table. The difference lies in measuring association controlling for a third variable. In this manner, while the association between X and Y may prove to be marginally significant, the association may not prove to be conditionally significant. Most statistical software packages such as SAS allow the investigator to construct partial tables that enable one to control for a possibly confounding third variable (Agresti, 1996).

Log-linear Regression Analysis

Log-linear regression is a form of a generalized linear model, or GLM. The OLS model described in the previous section is a special case of the GLM (Agresti, 1996). As such, all GLMs possess three components: the random and systematic components and the link (Agresti, 1996). The random component is the response variable Y, and assumes a probability distribution for the variable. If the response variable is a frequency count of some hydrologic event, we might assume a Poisson or negative binomial distribution for it. The systematic component is the independent variable(s) X, which as a linear combination of predictor variables will take on the form of:

$$\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k$$

as in the OLS model. The link specifies how the two variables will be modeled (Agresti, 1996).

For frequency count data, it is appropriate to employ the log-link in a GLM to model the log of the mean of the response variable. This model has the form of

$$\ln(\mu) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k$$

and is known as a loglinear model. The mean is then:

$$\mu = \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k).$$

Significance tests of the hypothesis H_o: $\beta = 0$ are conducted with the Wald statistic and the likelihood-ratio test. The Wald test statistic is

$$z = \hat{\beta} / \text{ASE}$$

where the ASE is the asymptotic standard error about $\hat{\beta}$. The test statistic has an approximate standard normal distribution when $\hat{\beta} = 0$ (Agresti, 1996). Large z values and small p-values suggest evidence against the null hypothesis. The likelihood-ratio statistic is

$$-2 \ln(l_0 / l_1) = -2 [\ln(l_0) - \ln(l_1) = -2 (L_0 - L_1)$$

where L_0 and L_1 are the maximized log-likelihood functions and yield a chi-square statistic (Agresti, 1996). Large chi-square statistics and small p-values indicate evidence against the null hypothesis.

Gross model fit is described by the summary goodness of fit statistics X^2 and G^2:

$$X^2 = \sum (y_i - \mu_i) / \mu_i,$$

where μ_i is an estimate, or hat-value. The G^2 statistic is:

$$G^2 = 2 \sum y_i \ln(y_i / \mu_i)$$

where μ_i is again a sample estimate. When fitted counts exceed 5, and N is fixed, then both are approximately chi-square distributed (Agresti, 1996). An adequate model fit is indicated by small test statistics and large p-values.

Spatial Interpolation Models

The following section describes relevant models necessary to perform spatial interpolation of seasonal rainfall over the Peace River drainage basin. These models make it possible to create area-averaged, or weighted seasonal precipitation values for the period 1957-1997. They include trend surface analysis and kriging models.

Trend Surface Fitting

To plot some value (z) at location (x,y), we can apply the method of ordinary least squares regression described in a previous section. Generally, a trend surface has the form of

$$Z_{obs(i)} = f(x_i , y_i) + \varepsilon_i$$

Where $Z_{obs(i)}$ = the observed value of the surface at location i,

x_i = the x axis coordinate,

y_i = the y axis coordinate, and

ε_i = the residual at the i^{th} data point.

f is some function that describes the trend of the variable Z over space (Unwin, 1975). Nearly any equation can represent this function. Two more common trend surfaces are a first-order linear trend surface and a second-order, or polynomial trend surface (Unwin, 1975). A first order trend describes a plane, and can be expressed as

$$z_i = a_0 + a_1x_i + a_2y_i + \varepsilon_i,$$

and a second order trend as

$$z_i = a_0 + a_1x_i + a_2y_i + a_3x_i^2 + a_4y_i^2 + a_5x_iy_i + \varepsilon_i,$$

Functional forms that are not linear in their parameters must be transformed so that the function that forms the relationship between $E(Z_i)$ at point (i) is linear and that the effects of each independent variable is additive. All other OLS assumptions apply. Quality of surface fit is gained by minimizing the error term above and below the surface, as in ordinary regression. Overall model fit is described by the RMSE, where small values indicate an adequate fit. Finally, trend surface analysis can be global, to capture an overall trend in the spatial variability in Z, or local to capture local influences. To capture the influence of local trends on a surface, a moving least squares approach is appropriate. It should be noted that trend surface analysis is considered deterministic, and does not account for any spatial autocorrelation that may be present.

Kriging

Kriging methods seek to make a prediction of some value Z at location (i) by quantifying the spatial structure of the data (Johnson et al., 2001), such that

$$Z_i = \mu_i + \varepsilon_i,$$

where μ_i represents a trend and ε_i is a spatially autocorrelated error term. Kriging methods attempt to remove the influence of any global or local trends in the surface by first removing the trend, modeling the surface variation, and then replacing the trend back into the data (Johnston et al., 2001). In effect, kriging assumes that the residuals from the trend surface are autocorrelated to a certain distance and/or direction, and quantifies the degree of correlation.

Quantifying the spatial structure is known as variography, where the modeler fits a spatial-dependent model to the data. Once the model is fit, the surface fit has accounted for the spatial autocorrelation for both direction (anisotropy) and direction. All

variograms are composed of the following parameters: variance, range, sill and nugget effect. Variance of the data (γ_h) over space is described by the y-axis where (h) represents the range between data points on the x-axis. The sill represents the range point (a) at which any autocorrelation can be attributed to random forces. Lastly, a nugget effect describes the error resulting from two variance measurements taken at the same location (McBratney and Webster, 1986). Two of the more common models are known as the spherical and exponential variograms, and are detailed below.

Spherical semivariogram

Each theoretical equation seeks to best fit the empirical data. The spherical model equation is based on the "volume of overlap of two spheres of diameter *h*" (McBratney and Webster, 1986), where *h* represents the range of the data. This model is part of what are considered transitive models, where the property of interest measured is a realization of the random function that relates the semi-variance to the lag (McBratney and Webster, 1986). So, the equation:

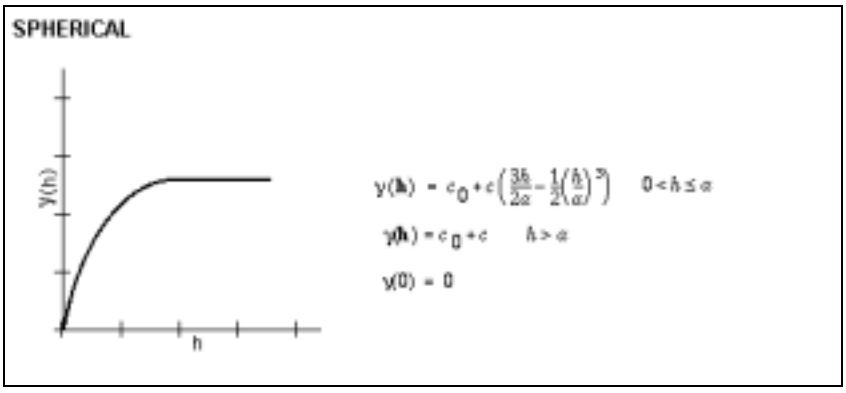

Figure 3.1. The spherical semivariogram equation and functional form.

can be described as: the semi-variance is equal to the nugget effect + the full sill times the parameter(s) h/a, which describe the gradient or rate at which the function reaches the sill

for a distance and direction between the partial sill and the limit of the(McBratney and Webster, 1986).

Exponential semivariogram

In contrast, an exponential model reaches its sill asymptotically, where there is no theoretical limit to the variance of the data (Johnston et al., 2001). This equation takes the form of:

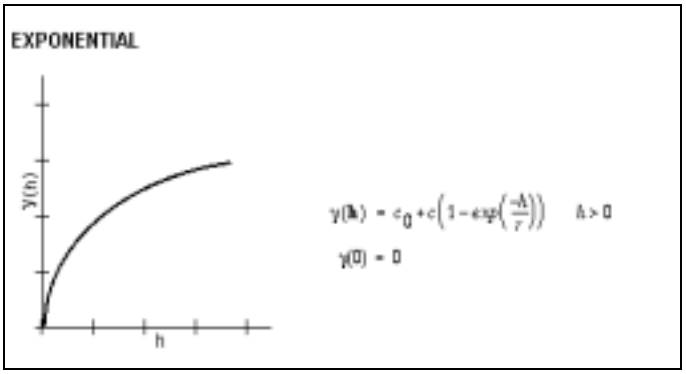

Figure 3.2. The exponential semivariogram equation and functional form.

McBratney and Webster (1986) state that the use of this model should be cautioned as this functional form can result from a variety of causes, although Oliver and Webster (1990) found an exponential best fit due to apparent unbounded increase in variance in all directions of their soils data. It should be noted though that in any case a large nugget effect would produce the same result as a large range in that the distance to which the fitted model approaches the sill is minimized. With a large range and small nugget, the rate at which the theoretical best-fit line approaches the sill is slowed (Dr. Bon Dewitt, pers. comm., 2002).

Summary

This section has presented a rudimentary introduction to the various modeling techniques used in this study. This chapter was not intended to fully educate the reader in

these methods. For more comprehensive model descriptions and assumptions, the interested reader is referred to the sources cited in this chapter.

CHAPTER 4
METHODS

Chapter Overview

The methods needed to achieve the objectives of this thesis are detailed in the present chapter. This chapter is organized by methods used for each model detailed in chapter 3 and differentiated for the streamflow and precipitation portions. Throughout this chapter, each method presented will employ a two to three season lag of SSTs to streamflow/precipitation. First, based on the literature reviewed and the flow regime, winter and spring streamflow/precipitation seasons are defined as:

- Winter = December, January, February (DJF)
- Spring = April, May, June (AMJ).

As such, monthly mean SSTs from both the eastern tropical Pacific Niño 3.4 (simply Niño 3.4) region and the tropical North Atlantic (TNA) are each averaged for the summer season, defined as:

- Summer = July, August, September (JAS).

Hereafter, seasonally lagged SSTs will imply the lag times defined above.

Streamflow Record Differentiation

The streamflow record is first converted from a flow rate of cubic feet per second (ft^3s^{-1}) to cubic meters per second (m^3s^{-1}). Seasonal minimum, mean and maximum streamflow series are created. Mean seasonal flow is the average of all daily flow in the season. Figure 4.1 illustrates of the flow variables during a typical winter season.

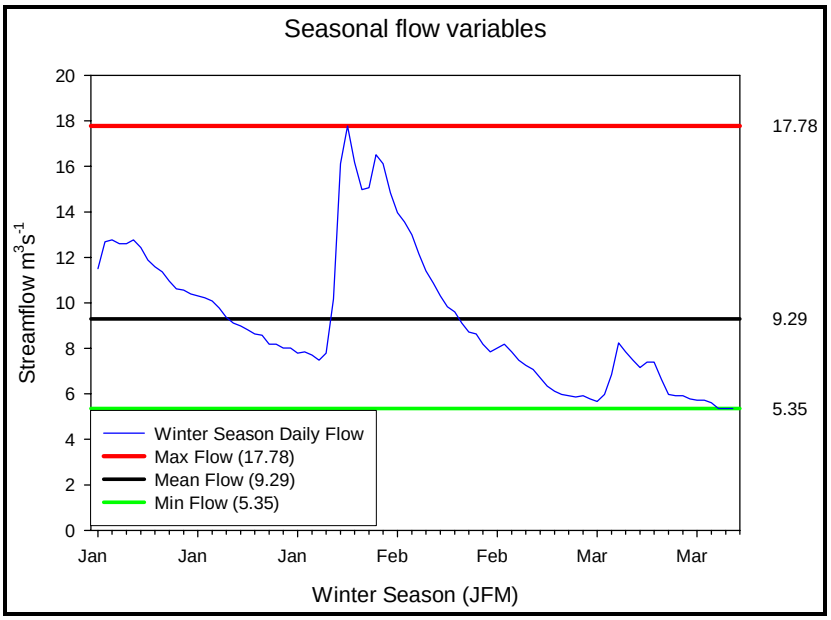

Figure 4.1. Seasonal flow variables.

In order to estimate the influence of Niño 3.4 SSTs, all streamflow records are first differentiated according to the following simple bipartite scheme:

- Warm Pacific (Wp, n = 25)
- Cold Pacific (Cp, n = 25).

The two categories are either above or below the median Niño 3.4 SSTs, and were selected to maximize sample size for each category, while still maintaining some separation of oceanic conditions. To identify any influence of TNA SSTs, the records are further differentiated into four subpopulations:

- Warm Pacific-Cold Atlantic (Wp-Ca, n = 11)
- Warm Pacific-Warm Atlantic (Wp-Wa, n = 14)
- Cold Pacific-Cold Atlantic (Cp-Ca, n = 14)
- Cold Pacific-Warm Atlantic (Cp-Wa, n = 11)

where the TNA SST categories are also either above or below the median value. In this manner the separate or combined influences of the TNA and Niño 3.4 can be estimated and compared.

Determination of Lognormality

A probability distribution is then fit to each subpopulation. Because the two parameter lognormal distribution has proven to be an appropriate distribution for positively skewed hydro-meteorologic variables (Stedinger et al., cited in Maidment, 1993), the natural log of the flow variable is taken. Each subpopulation is then ranked and the Weibull formula estimates a plotting position:

$$i / (n + 1)$$

where $\{i\}$is the rank, and $\{n\}$is the sample size. Selection of a plotting formula is somewhat arbitrary. However, the Weibull provides unbiased exceedence probabilites for all distributions including the lognormal (Stedinger et al., cited in Maidment, 1993). Theoretical exceedence probabilities are then generated for each subpopulation using the mean and variance of the log-flow series. The largest difference (D_{max}) between the observed and fitted exceedence probabilities are calculated and compared to the Kolmogorov-Smirnoff (K-S) *D* values for the appropriate sample size $\{n\}$ and alpha (α) level.

This common goodness of fit test states that the null hypothesis (H_o: reject log normality) is not rejected if $D_{max} < D$. Because the K-S is a right tail test, the alpha level should be generous. In this thesis an alpha (α) of .20 is selected.

Parametric Statistical Tests

An F-test for variance equality, a Students t-test (either pooled or non-pooled, depending on the F-test result) and ANOVA procedures are applied to determine statistical significance between each subpopulation (Burt and Barber, 1996). All tests are performed at the 95% confidence level.

Return Period Generation

An estimate of flows of various return periods can be from the two-parameter lognormal distribution. Return periods represent the probability that a flow of a given level will be exceeded in any one year or season. The return period T, in years, is defined as

$$T = 1 / P(F_s) = 1 / 1 - P(F_f)$$

where $P(F_s)$ is the probability of success and $P(F_f)$ is the probability of failure (Viessman et al., 1989). For example, the 100-year return period for a flood estimate represents a flow level that has an exceedence probability of 0.01 in any given year or season. Conversely, the 100-year return period for a drought estimate has an exceedence probability of 0.99 for a given season

Return periods generated for this study are:

- 100 year
- 50 year
- 20 year
- 10 year
- 5 year
- 2 year

for both floods and droughts.

Tercile Probability Estimates

Estimating the value of the variable corresponding to the upper and lower terciles is gained by using the mean and standard deviation of each season's undifferentiated minimum, mean and maximum populations. Conceptually, a flow level of 25.76 m^3s^{-1} represents the 66.6 percentile of the historic record for mean winter season flow. Tercile flows based on the undifferentiated series establishes an "above average" (66.6 percentile) flow, a "below average" (33.3 percentile) and an "average flow" (between the

upper and lower terciles). If SSTs affect flow then the probabilities of falling in each tercile will depend upon oceanic conditions. This procedure is duplicated for each flow variable, season and ocean surface condition.

By using the seasonally lagged SSTs from both oceans, the probability of experiencing above average, average and below average historic flow levels can be forecast at considerable lead times conditioned upon changing SST conditions in the Niño 3.4 region and in the TNA.

Monte Carlo Simulations

Due to the small sample sizes of each differentiated series, some measure of confidence about the empirical tercile probability forecast is needed. This is accomplished by generating a synthetic streamflow series under a Monte Carlo simulation for each seasonally differentiated Cp-Wp subpopulation based upon the mean and standard deviation of each series and for each flow variable. This method enables us to determine if the observed tercile probability forecast generated about each historically observed Wp-Cp differentiated series are significantly different from one another, or if the observed differences could simply have occurred at random.

As such, a Monte Carlo simulation of 1000 records (1000 realizations of X) each of length (n = 25) is generated with the aid of a random (0-1) number generator in an Excel spreadsheet. The random number is converted to a log normally distributed probability value by specifying a normal distribution and using the log mean and standard deviation of each differentiated realization. A simulation is performed for each flow variable for each season. From here the tercile probability estimates, again based on the historic 66.6 and 33.3 percentile flow levels of each flow variable, are generated for each record of length n = 25, as detailed above. The synthetic lower, middle and upper tercile

probability estimates are then plotted in a box and whiskers format. This allows us to determine if the observed tercile probability estimates lie in the tails of the distribution or in the fail to reject region of the distribution curve.

Further, a synthetic streamflow series is created with a Monte Carlo simulation (1000 trials, each of length 14 or 11 to correspond to the same lengths of the four Pacific-Atlantic categories) using the mean and standard deviation from the two previous Wp-Cp differentiated series. If the observed tercile probability estimates from the various Pacific-Atlantic categories are significantly different from the Cp-Wp tercile probability estimates, those estimates should lie in the tails of the distribution of the Cp-Wp differentiated series. In this manner, the significance of not only the tercile probability estimates of the Pacific differentiated series can be tested, so to can the estimates produced by the various differentiated series outlined above be tested.

Ordinary Least Squares Multiple Regression

As detailed in the previous chapter, OLS regression methods were employed to determine the influence of Niño 3.4 and TNA summer SSTs on winter/spring flow variable response. The purpose is to apply an additional forecast method and to support the log normal probability distribution findings

Independent Variable Definitions

Independent variables to forecast winter and spring streamflow include primary variables, dummy variables and interaction terms. Primary variables are SST data and antecedent seasonal streamflow. Dummy variables are surrogate variables created to assess the influence of extreme Niño 3.4 and TNS SSTs on the forecast model. Dummy variables are also interacted to assess any combined effects of extreme SSTs.

Interaction terms are the various primary and dummy variables multiplicatively interacted with one another.

For the winter season regression model, the primary independent variables include *log transformed*:

- ONDQ = seasonally averaged mean discharge from October, November, December.

- JASPAC = summer season (July, August and September) Niño 3.4 SSTs
- JASATL = summer season TNA SSTs.

Dummy variables include:

- PACPL1 = 1 if Niño 3.4 SSTs > 1σ above mean, 0 otherwise.
- PACMIN1 = 1 if Niño 3.4 SSTs < 1σ below mean, 0 otherwise.
- ATLPL1 = 1 if TNA SSTs > 1σ above mean, 0 otherwise.
- ATLMIN1 = 1 if TNA SSTs < 1σ below mean, 0 otherwise

Interaction variables included as dummy variables are:

- PACATL_PL1 = 1 if both Niño 3.4 *and* TNA SSTs > 1σ above mean, 0 otherwise.
- PACATL_MIN1 = 1 if both Niño 3.4 *and* TNA SSTs < 1σ below mean, 0 otherwise.

- PACPL1_ATMIN1 = 1 if Niño 3.4 SSTs are > 1σ above mean *and* TNA SSTs are < 1σ below mean, 0 otherwise.

- PACMIN1_ATPL1 = 1 if Niño 3.4 SSTs are < 1σ below mean *and* TNA SSTs are >1σ above mean, 0 otherwise.

For the spring season regression model, the only difference in defined terms is the primary independent variable for the antecedent season stream discharge:

- JFMQ = *log transformed*, seasonally averaged mean discharge from January, February, March.

All other variables remain the same. It should be noted that summer Niño 3.4 and TNA SSTs were not interacted with fall or winter streamflow, as it is likely that the SSTs have influence upon those season discharge levels. As such, those interaction terms would

introduce redundant information in the models. All variables listed above are entered into a forward stepwise model described below.

Model Procedures

A forward selection stepwise procedure is a common method to eliminate variables in multiple regression models that are related. It is a safeguard against introducing redundant information into the models and helps to eliminate multi-collinearity among the independent variables. Most statistical software has the capabilities to run both stepwise and complete model procedures.

Here, the significance level for admittance into the models is 0.10 and 0.15 for the winter and spring seasons, respectively. These are commonly used criterion values to allow a predictor into the full model, dependent upon a calculated F value for that variable (Graybill and Iyer, 1994). The stepwise procedure then simply iterates through the variables, successively adding one variable at a time. For a complete description see Graybill and Iyer (1994).

The complete model is then run with the variables that have significant explanatory power. Because both models are log-log models, the anti-logs of the variables and their coefficients are taken to recover model influence.

Three-Way Contingency Tables

The following two sections detail the methods necessary to analyze frequency count data. The first procedure is a three-way contingency table designed to describe the conditional (on summer SSTs) likelihood of experiencing a season during which the number of low-flow days is above or below a threshold value of 3.68 $m^3 s^{-1}$ (130 $ft^3 s^{-1}$). The latter is the response variable, and this procedure allows us to control for a third variable. That is, once the streamflow data are classified by region and SST condition,

we control for region rather than condition. The regions of influence on the response variable are again the Niño 3.4 and TNA. The partial association we seek to identify is whether it is the SST conditions in the Niño 3.4 or in the TNA that proves to have the significant influence. An above or below median level for the SSTs was again chosen to maximize frequency counts.

First, the number of days per season (record) below 3.68 $m^3 s^{-1}$ is extracted from the daily streamflow record. The long-term mean number of days for winter and spring are 8 and 23 days, respectively. The frequency counts are then inserted into a contingency table, cross-classified by ocean and surface condition. The following table is illustrative of the winter season:

Table 4.1 Number of days per season above/below flow threshold by ocean and SST condition. See text for explanation.

Ocean	Condition	# Seasons > 8 days per season	# Seasons < 8 days per season
Pacific	Cold	10	15
	Warm	2	23
Atlantic	Cold	6	19
	Warm	6	19

The next table demonstrates the same procedure for the spring season.

Table 4.2 Number of days per season above/below flow threshold by ocean and SST condition. See text for explanation

Ocean	Condition	# Seasons > 23 days per season	# Seasons < 23 days per season
Pacific	Cold	16	9
	Warm	5	20
Atlantic	Cold	11	14
	Warm	10	15

The frequency counts are entered into a SAS statistical software package, where a PROC FREQ (for procedure frequency) is run, which generates descriptive measures

such as conditional probabilites for cell counts as measured by a chi-square statistic, and the relative risk and odds ratios. This allows us to forecast the likelihood of experiencing more or less than the expected number of low flow days per season for the upcoming winter and spring seasons.

Loglinear Regression Analysis

This procedure permits the forecast of the expected number of low-flow days per season as detailed in the previous section. However, many of the winter and spring seasons on record do not have any days during which low flows occur. To counter this, and to satisfy the large sample chi-square requirement of $\{n_{ij} \geq 5\}$, the data were grouped into categories defined by SST conditions in the Niño 3.4 region, sorted from coolest to warmest. The following categories were defined by quintiles:

- Very Cold
- Cold
- Neutral
- Warm
- Very Warm

where the Very Cold category represents the coldest 20 percentile. Each category after represents the next warmest 20 percent, etc. Table 4.3 is illustrative of the winter season.

Table 4.3. Total number of days per season categorized by SST condition in the Niño 3.4 region for JFM.

Category	Number Cases	Number Days
Very Cold	10	100
Cold	10	83
Neutral	10	154
Warm	10	32
Very Warm	10	29

Table 4.4 illustrates the same procedure for spring season. These data are then entered into a statistical package such as SAS and a log-linear regression analysis is run. Explanatory variables are defined as:

- The mean Niño 3.4 SST for each class
- The mean TNA SST for each class
- The multiplicative interaction term of each explanatory variable for each class.

Table 4.4. Total number of days per season categorized by SST condition in the Niño 3.4 region for AMJ

Category	Number Cases	Number Days
Very Cold	10	429
Cold	10	210
Neutral	10	281
Warm	10	161
Very Warm	10	66

The mean TNA SST for each class requires further explanation. When the frequency counts of low flow days are sorted according to the Niño 3.4 SSTs, the condition of the TNA during the same year and season is simply sorted along side. That is, no stratification of TNA SSTs is applied. If the condition of the TNA for each grouping of Niño 3.4 categories is significant, the software will detect that and enter it into the model.

Next, the distribution of the response variable is defined. Typically, a Poisson distribution is specified for the response variable in loglinear models. However, a Poisson distribution requires that the mean and variance of the data are equivalent (Burt and Barber, 1996). In practice this is rarely the case. If a Poisson distribution is specified for the response variable in a GLM and the distribution is other than Poisson, a phenomenon known as overdispersion occurs (Agresti, 1996). One way to counter this is to specify a negative binomial distribution for the response variable. By specifying a

negative binomial distribution in a GLM, the software applies an adjustment factor to the ASE. As the mean and variance of the grouped count data are not identical, a negative binomial distribution is specified with a log link. The software then generates the regression equation and all diagnostic statistics as detailed in the previous chapter.

GIS and Seasonal Precipitation Modeling Methods

The preceding sections described all methods necessary to generate various forecasts of winter and spring season streamflow as measured at Arcadia, FL. The next sections detail the methods necessary to perform a basin-wide seasonal precipitation forecast, using just the conditional lognormal probability forecast methods applied to the winter season precipitation data.

Drainage Basin Delineation

Delineating the contributing drainage area for the Peace River at Arcadia, FL is conducted using merged 7.5 minute, $30m^2$ spatial resolution Digital Elevation Models (DEM) in a Geographic Information System (GIS). The GIS uses a set of "stream burning" (Maidment, 2002) algorithms that iterates through hydrologically processed grids to create a stream network, allowing the user to select any point along the network from which the drainage area is calculated (Doan, cited in Maidment and Djokic, 2000). The final contributing area to the Peace River at Arcadia is depicted in figure 1.2.

Upon drainage area is delineation, the data from the weather stations within the basin boundary are gathered from the NOAA data website.

Precipitation Interpolation

Total monthly rainfall from the contributing rain gauges is averaged for each winter season and year over the period of precipitation record, 1957-1997. A table field heading corresponding to each year of winter season precipitation is added to the NOAA weather

station database in the GIS. However, because the data contain missing monthly records, a seasonal average cannot be calculated for some years. To overcome this obstacle, various interpolation methods were investigated in order to obtain an interpolated surface that displays the best fit in terms of minimizing the root mean square error (RMSE) and the error associated with each residual.

To accomplish this, an average precipitation year (winter season 1988) is selected among the complete seasonal records in order to create a calibration surface with which to compare interpolation models that are used to fill the missing records. An average year is considered one that displays a mean amount of precipitation variation among the 10 contributing gages. This year was examined for any trend to be accounted for or removed before applying a "best fit" surface. Once a best fit surface has been identified, the interpolation method used to create that surface is applied to years with missing data, thereby providing a complete data set for the period of record.

The results of the various deterministic interpolation methods for the winter season of 1988 are shown below in table 4.5, while the geostatistical models are shown below in table 4.6. Each entry in the deterministic table indicates the type of model, RMSE, and the resulting mean prediction error associated with the surface. The number of entries for each method corresponds to differences in the size and shape of the neighborhood used. The entries in the kriging table correspond to the following:

8. Ord = ordinary kriging model
9. Trend = order trend removal
10. Semivar = the semivariogram model used.
11. Lag size = the size of bin used by the model in m^2.
12. Lag num = the number of bins
13. Aniso = use of anisotropic or directional spatial autocorrelation component.
14. Pred errors = surface error diagnostic statistics.
15. Mean = raw residuals.

16. RMSE = root mean square error of entire surface fit
17. Avg Std Error = the standardized residuals.

Table 4.5. Deterministic interpolation models showing the various goodness of fit statistic and error terms. See text for explanation.

Deterministic Models		
Method	RMSE	Mean predict error
IDW	157.7	-13.24
IDW	159.2	-10.87
Global_poly	156.4	-4.55E-14
Global_poly	181.1	-12.62
Global_poly	357.2	-40.39
Local_poly	154.8	-9.691
Local_poly	154.8	-8.349
Local_poly	182.5	-11.98
Local_poly	182.5	-11.98
Local_poly	273.6	-29.69

Table 4.6. Geostatistical interpolation models, various semi-variograms, and optimal surface statistics.

Kriging Models						Prediction Errors				
Method	trend remove	semivar	lag size	Lag #	Aniso-tropy	Mean	RMSE	Avg std	Mean stdizd	RMSE stdizd
ORD1	none	spher	9000	9	y	-24.85	154.3	124.6	-0.126	1.088
ORD2	none	expon	9000	9	y	-23.24	156.3	127.4	-0.117	1.095
ORD3	none	gauss	9000	9	y	$2.73^{(-13)}$	156.4	138	0.000	1.134
ORD4	const	spher	9000	9	y	-20.84	170.3	127.3	-0.095	1.198
ORD5	const	expon	9000	9	y	-18.93	153.9	132.5	-0.084	1.046
ORD6	const	gauss	9000	9	y	$-2.27^{(-14)}$	156.4	132	0.000	1.185
ORD7	first	spher	9000	9	y	-22.81	166.1	110.6	-0.143	1.341
ORD8	first	expon	9000	9	y	-25.94	170.3	113.2	-0.156	1.332
ORD9	first	gauss	9000	9	y	-12.62	181.1	121.3	-0.104	1.494
ORD10	2nd	spher	8000	12	y	-40.41	355	86.02	-0.447	4.073
ORD11	2nd	expon	8000	12	y	-40.39	357.2	85.65	-0.472	4.171
ORD12	2nd	gauss	8000	12	y	-40.39	357.2	85.65	-0.472	4.171

Of the various models investigated, the ordinary kriging model number 5 produced the most optimal surface for the calibration year of 1988. This model was also the only

one attempted that passed all tests for normality in the raw residuals, and in the standardized residuals. The prediction error and standardized residuals for this model also tested not significantly different from zero, and are normally distributed (see tables B1, B2 and B3 in appendix B). An example of the surface created by the kriging model number 5 is shown in figure 4.6.

Once the residuals have been tested, the same model was applied to all years and stations that have missing rainfall seasonal totals. Because the yearly (seasonal) fields in the NOAA weather station GIS layer have to be defined as numeric (integer, etc), the missing seasonal totals for various rainfall stations are considered as zeros

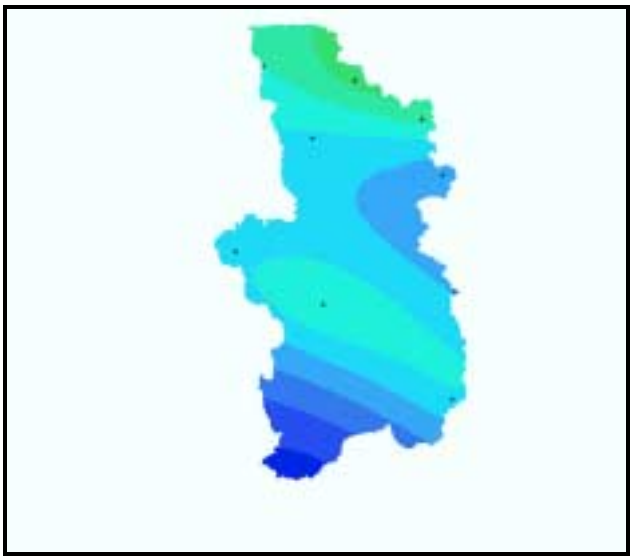

Figure 4.2. Surface produced by ordinary kriging model number 5 for 1988.

instead of no data. It is possible to omit those years with zero values in a deterministic model, but the kriging models require a minimum of 10 points. So, while it is anticipated that the years with the missing data would produce large rmse errors, the decision to use the model to predict totals for the missing data locations was made based on the performance of the model for the calibration year of 1988.

Once all the missing seasonal totals have been predicted by the model, seasonal mean precipitation can be calculated using a simple arithmetic average. As mentioned above, this is an acceptable method for estimating basin precipitation when the basin is of very low relief and the rain gages are evenly distributed throughout the basin (Viessman et al., 1989). Still, this assumption has to be tested, so the mean seasonal precipitation for each year was modeled using the best-fit surface criteria above in order to obtain area-weighted precipitation grids. The model selection criteria to produce each grid were based upon the same parameters as the calibration model.

Upon evaluation of each surface, the GIS then converts them to a raster grid format in order to extract the area-weighted mean precipitation values. Finally, both area-weighted and area-averaged mean seasonal precipitation is tested for normality, and then the entire record is differentiated and tested for normality and differences in means as for winter/spring season streamflow.

Summary

This chapter has presented the methods necessary to complete the data analysis needed to produce seasonal streamflow/precipitation forecasts. Chapter sections detail conditional lognormal probability distributions based on lagged sea-surface temperatures (SSTs) in the Niño 3.4 region of the eastern Pacific and tropical North Atlantic. An ordinary least squares multiple regression method uses precedent seasonal streamflow and interacted dummy variables that account for extreme SST methods to produce a forecast. Three-way contingency tables examine the likelihood of experiencing seasons during which the number of low-flow days exceeds the long-term mean. Fourth, the necessary techniques are presented to perform a loglinear analysis based upon lagged SSTs, providing a measure of low flow days per season.

Lastly, this chapter has extensively detailed necessary procedures to forecast basin precipitation using a GIS.

CHAPTER 5
RESULTS

Chapter Overview

Model and statistical results from the analysis of the various streamflow variables are presented. Differences of means and variances between the differentiated subpopulations are described first, followed by Warm Pacific (Wp)-Cold Pacific (Cp) conditional return period differences in the winter and spring mean discharge. The probability distributions of the undifferentiated historic subpopulations of each flow variable will be followed by Wp-Cp differentiated subpopulations. Next, the results of each historic flow variable further differentiated on the conditions in the Niño 3.4 *and* tropical North Atlantic (TNA) sea-surface temperatures (SSTs) are assessed for each season. The results of the Monte Carlo simulated synthetic streamflow subpopulations will follow each observed probability distribution and season. The OLS multiple regression results are detailed next, to be followed by the categorical data analysis. The latter includes the loglinear regression analysis of low flow days and the results of the three-way contingency tables. Mean winter precipitation probability distributions are presented last. As described previously, the SSTs are lagged behind the flow variables in all models by 2-3 seasons.

Statistical Assessment of Differentiated Flow Variables

Table 5.1 displays the mean and standard deviation of each differentiated streamflow subpopulations.

Table 5.1. Mean and variance in log terms for each differentiated subpopulations. Asterisks denote significant differences at the 95% confidence level between Warm/Cold Pacific differentiated subpopulations. All SST categories are defined in chapter 3.

SST	Winter Season Flow Variable			Spring Season Flow Variable		
	Max	Mean	Min	Max	Mean	Min
Wp	4.29*, 0.84	3.14*, 0.74	2.01*, 0.44	4.25, 0.65*	2.77*, 0.53*	1.20*, 0.37*
Wp-Wa	4.35, 1.04	3.19, 0.87	2.19, 0.56	4.27, 0.80	2.81, 0.73	1.36, 0.45
Wp-Ca	4.22, 0.78	3.07, 0.63	1.77, 0.20	4.22, 0.51	2.71, 0.30	1.01, 0.21
Cp	3.29*, 1.14	2.25*, 0.59	1.30*, 0.34	3.81, 1.39*	2.29*, 1.04*	0.58*, 0.79*
Cp-Ca	3.24, 1.31	2.24, 0.62	1.39, 0.27	4.14, 0.74	2.55, 0.73	0.90, 0.55
Cp-Wa	3.35, 1.04	2.26, 0.61	1.19, 0.44	3.39, 2.02	1.97, 1.33	0.17, 0.85

During the winter season, each flow variable's subpopulation exhibits a significant difference in means (48df), between the Warm Pacific (Wp)-Cold Pacific (Cp) subpopulations. During the spring season, there are significant differences in both means and variances between Wp and Cp subpopulations for the mean and minimum flow variables, while maximum flows exhibit a significant difference only in variance.

To test for a significant effect of warm and cold TNA SSTs (on the Wp and Cp differentiated flow subpopulations), the Warm Pacific-Warm Atlantic (Wp-Wa) and Warm Pacific-Cold Atlantic (Wp-Ca) differentiated subpopulations are compared to the Wp. Similar comparisons are made for the Cold Pacific subpopulation. Sub-division based upon the Atlantic produced no significant differences in the mean or variances of the subpopulation statistics.

These results suggest that streamflow variables, when differentiated conditioned upon a Wp and Cp belong to different statistical populations, and that the TNA SSTs do not significantly change the first two moments of the Wp and Cp subpopulations.

Conditional Lognormal Probability Distributions

Winter Season Return Period Estimates

Figures 5.1 and 5.2 display mean winter flows associated with return periods of floods and droughts. Streamflow magnitude is on the y-axis and return period in years is on the y-axis.

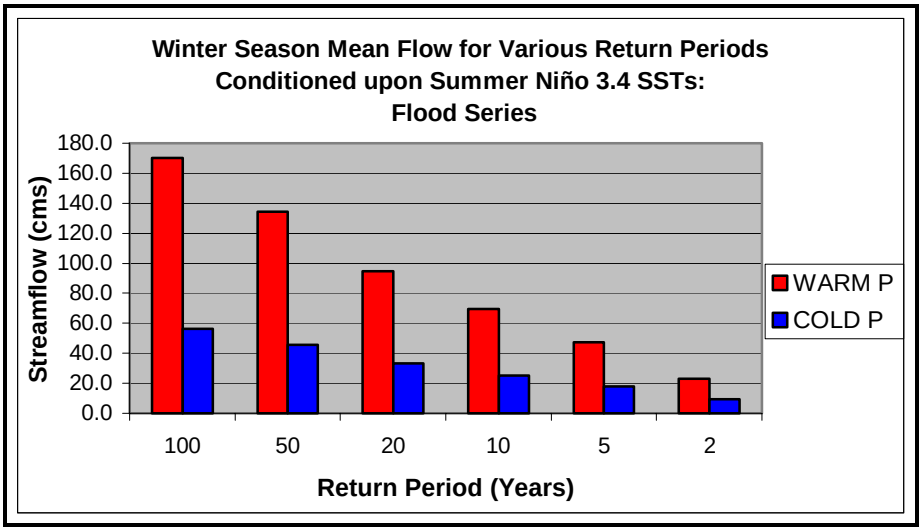

Figure 5.1. Return period estimates for winter season mean flow variable, flood series.

The estimates correspond to exceedence probabilites of 0.99, 0.98, 0.95, 0.9, 0.8 and 0.5 for the 100, 50, . . . 2-year return periods. The graph indicates that higher flows can be expected under a Wp condition at all periods. There is also a greater variability in flows under the same condition. Figure 5.2 indicates that higher "drought" flows can be expected under a Wp condition at all periods, but the variability in flows is reduced as compared to the flood series from the 100-year to the 2-year return period.

Winter Season Minimum Flow Variable

Figure 5.3 displays the probability distribution of winter season historic minima and the historic lower and upper terciles.

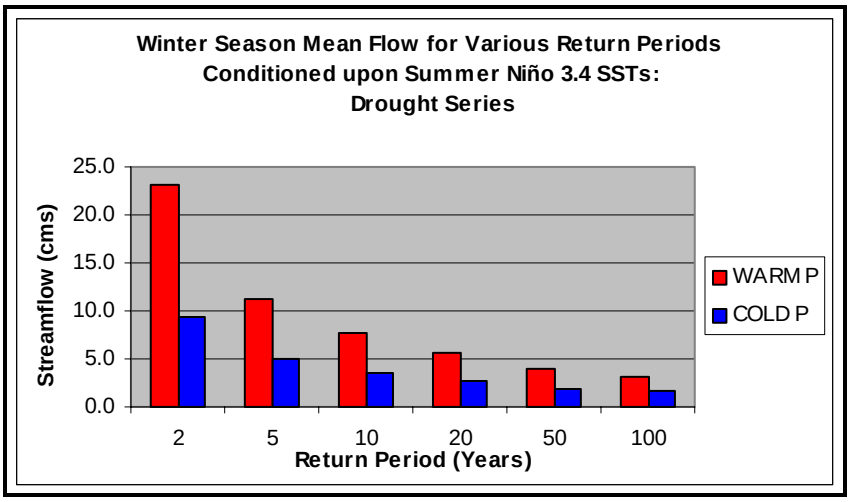

Figure 5.2. Return period estimates for winter season mean flow variable, drought series.

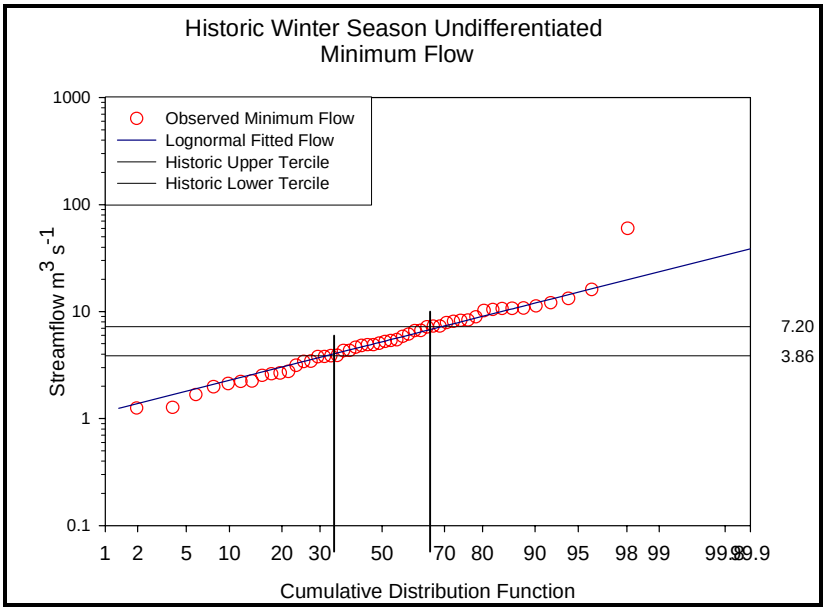

Figure 5.3. Winter season minimum flow lognormal probability distribution.

The vertical drop lines indicate the exceedence probabilites for the upper and lower terciles. For winter season minimum flows, the flow magnitude that corresponds to a 0.333 exceedence probability is 3.86 m^3s^{-1}, while the (1- 0.666) exceedence probability flow is 7.20 m^3s^{-1}. Average minimum seasonal flow is defined as a flow magnitude between these two levels and exceedence probabilities. This establishes the "above average", and "below average" winter season flow categories used throughout the thesis.

The two-parameter lognormal probability distribution (fitted function) provides a reasonable fit for the historic minimum flow subpopulations and passes the K-S *D* at the 80% confidence level, as do all probability distributions of fitted and observed flow variables.

Figure 5.4 depicts the winter season minimum flows conditioned upon Niño 3.4 SSTs two seasons prior, and indicates the probabilities of experiencing "above" and

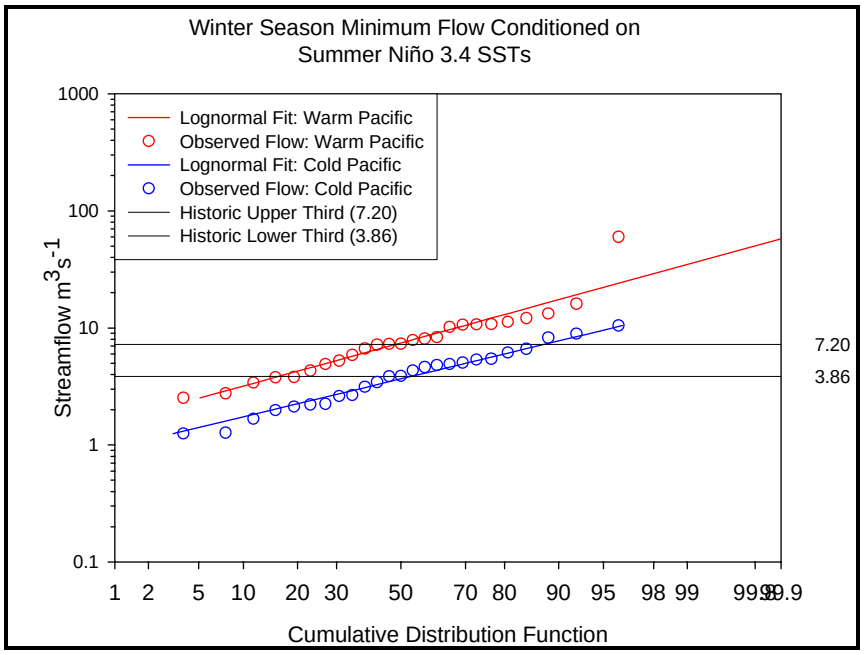

Figure 5.4. Winter season minimum flow differentiated on the preceding summer Niño 3.4 SSTs.

"below" average flows under each oceanic condition. The probability of experiencing below average flow under a CP condition has increased from 0.333 to 0.533. Conversely, under a Wp condition, the probabilities of experiencing the same flow level decrease to 0.160. Above average minimum flow is more likely under a Wp condition (0.520) and are less likely (0.124) under a Cp condition. Average flow probabilities remain fairly constant at 0.343 under a Cp condition and 0.320 under a Wp condition.

Further differentiation into Wp-Wa and Wp-Ca categories (figure 5.5),

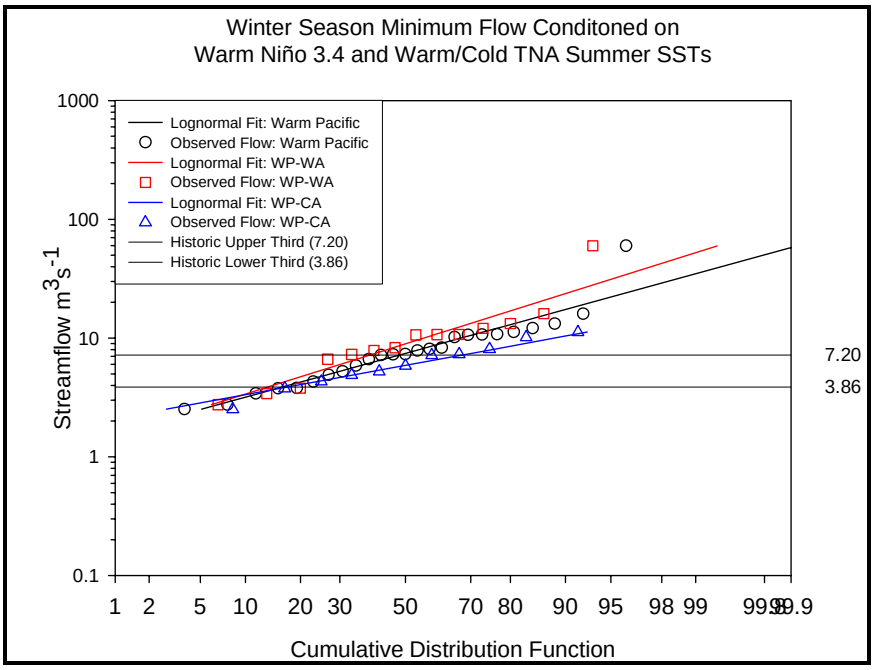

Figure 5.5. Winter season minimum flow differentiated into Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

suggests that the probability of "below" average flow has decreased to 0.134 and increased to 0.169 under Wp-Wa and Wp-Ca conditions, respectively. A very small change associated with changes in Atlantic SST. Somewhat larger changes in probabilities are associated with "above" average minimum flows under a Wp condition. They rise to 0.612 under a Wp-Wa combination and decrease to 0.325 under a Wp-Ca condition. This implies that the combination of warmer Atlantic SSTs and warmer Niño 3.4 may increase the likelihood of average and above average flows.

Figure 5.6 repeats the procedure for a Cold Pacific. The previously established probability of experiencing "below" average flow under an isolated Cp condition remains at 0.533. The probability of experiencing below average flow is 0.471 and 0.595 for Cp-Ca and Cp-Wa conditions, respectively.

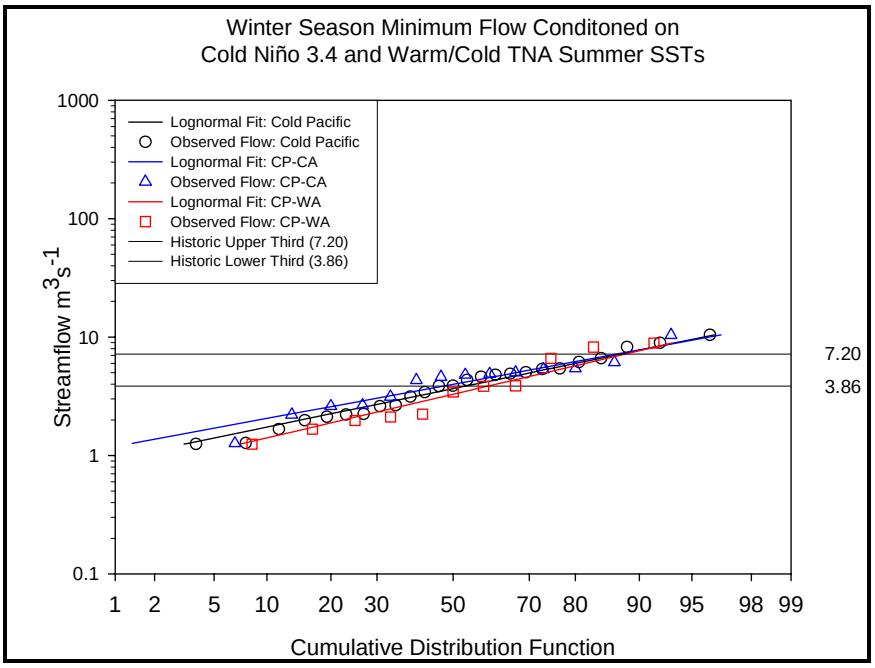

Figure 5.6. Winter season minimum flow differentiated into Cp-Ca/Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

The likelihood of "above" average minimum flow under a Cp condition of 0.124 changes very little under a Cp-Ca condition (0.129) or a Cp-Wa combination (0.118). The likelihood of experiencing average flows is 0.343, 0.278 and 0.399 under Cp, Cp-Wa and Cp-Ca conditions respectively.

Winter Season Mean Flow Variable

Figure 5.7 displays the winter season historic winter mean flow data, fitted lognormal distribution and terciles. Average mean flow lies between 8.24 m^3s^{-1} and 25.76 m^3s^{-1}. Figure 5.8 depicts the probabilities of mean flows when the historic series is differentiated upon a warm or cold Niño 3.4 condition. The probability of experiencing below average flow under a Cp condition increases to 0.428 and decreases to 0.115 for a Wp condition.

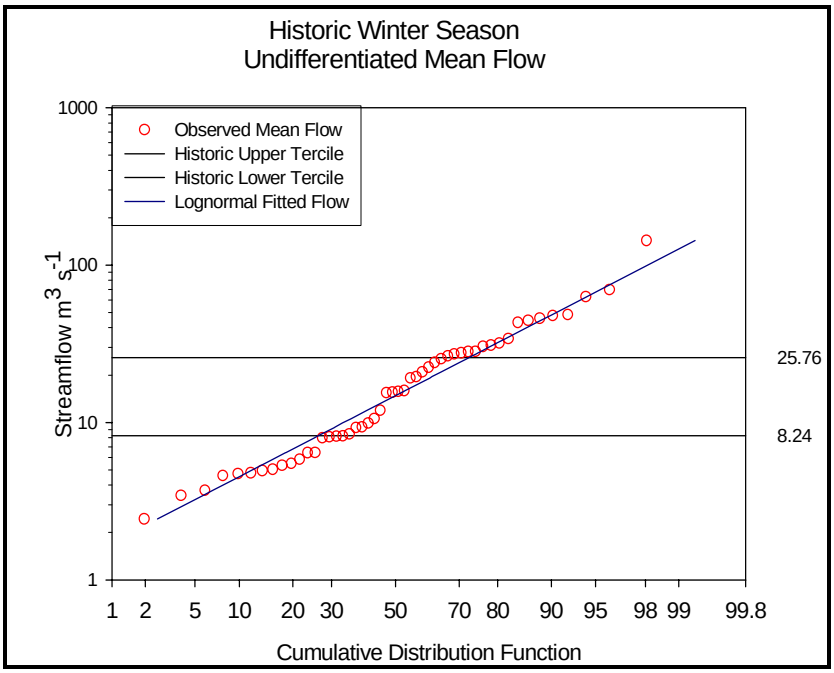

Figure 5.7. Winter season mean flow lognormal probability distribution.

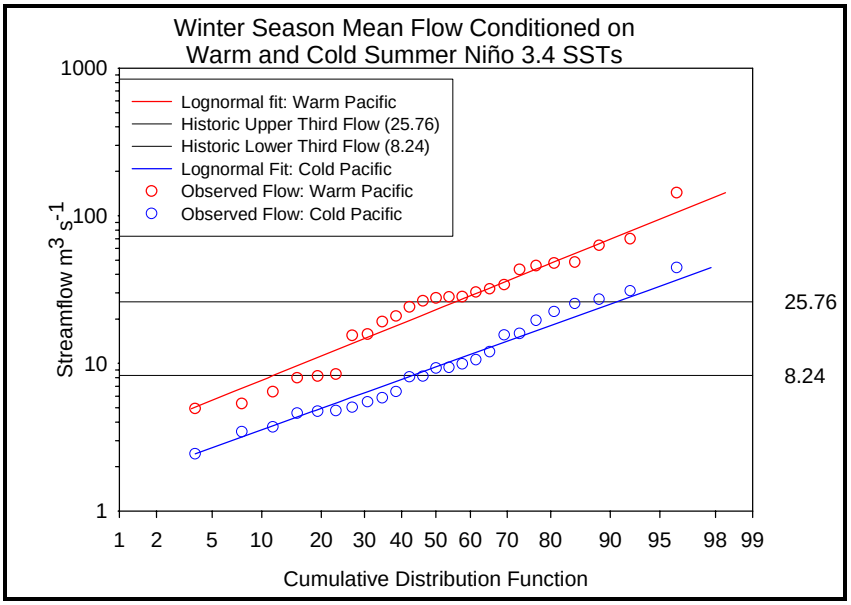

Figure 5.8. Winter season mean flow differentiated on summer Niño 3.4 SSTs.

The probabilities of experiencing above average mean flow increase to 0.449 under a Wp condition, and decrease to only 0.095 under a Cp condition. The likelihood of experiencing average flows is 0.436, and 0.476 a Wp and Cp, respectively.

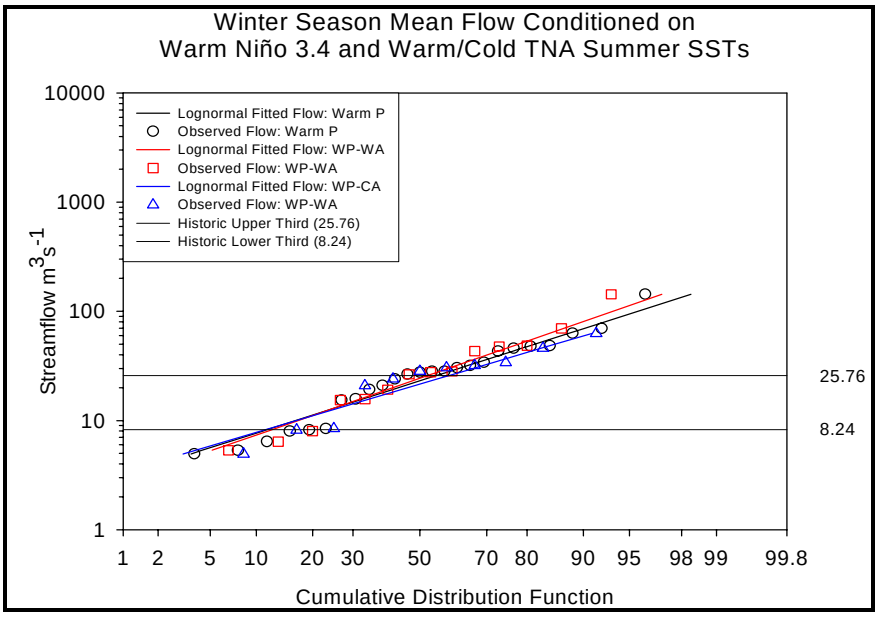

Figure 5.9. Winter season mean flow differentiated into Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

Figures 5.9 and 5.10 show that these probabilities only change slightly when further sub-division based on Atlantic SSTs is carried out.

Winter Season Maximum Flow Variable

The distribution the maximum flow variable and terciles are shown (figure 5.11). The average maximum seasonal flow is defined as being between 28.11 m^3s^{-1} and 77.93. Once again, a clear separation in the distributions and probabilities of flow levels is observed when the data are sub-divided by Pacific SSTs (figure 5.12). The likelihood of experiencing below average flow under a Cp condition increases from to 0.518, while the probability of a similar event under a Wp condition decreases to 0.156. The risks of above average maximum flows have increased to 0.473 under a Wp condition, but decrease to 0.160 under a Cp condition. The likelihood of experiencing average flows is 0.371, and 0.323 Wp and Cp conditions, respectively.

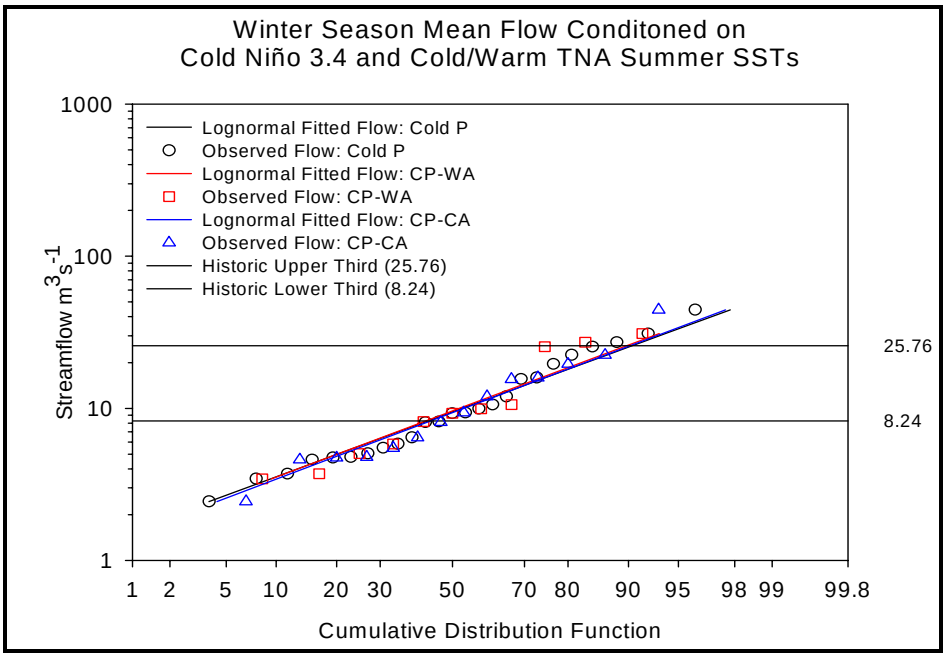

Figure 5.10. Winter season mean flow differentiated into Cp-Ca /Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

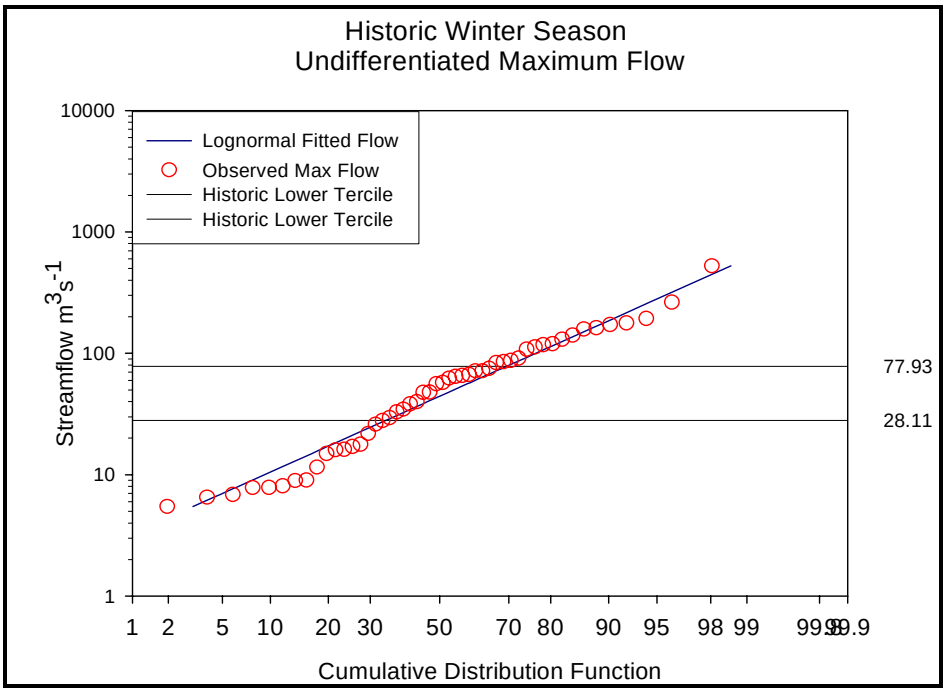

Figure 5.11. Winter season maximum flow lognormal probability distribution.

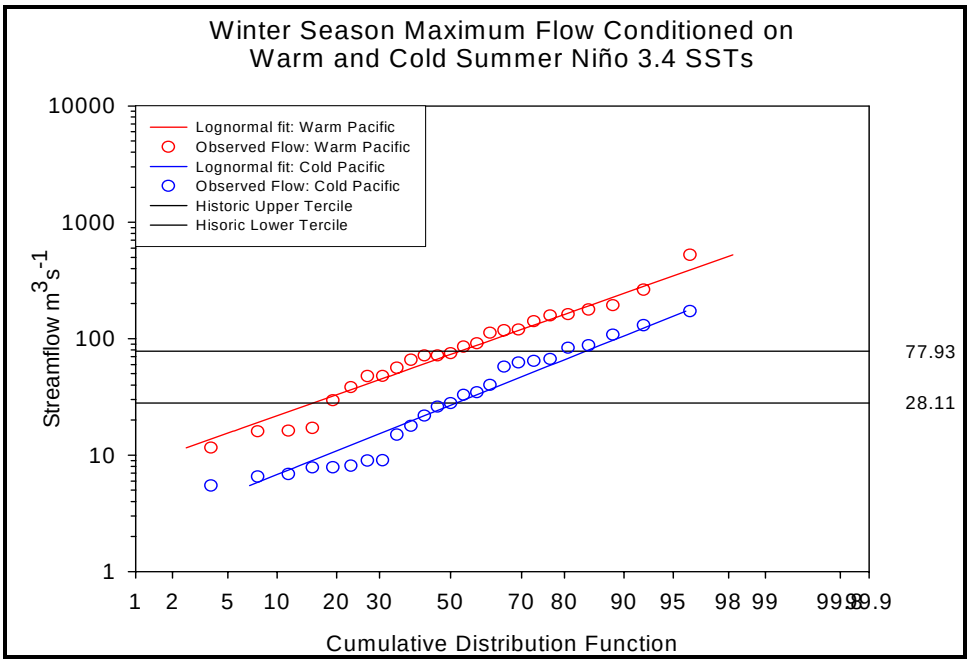

Figure 5.12. Winter season maximum flow differentiated on summer Niño 3.4 SSTs.

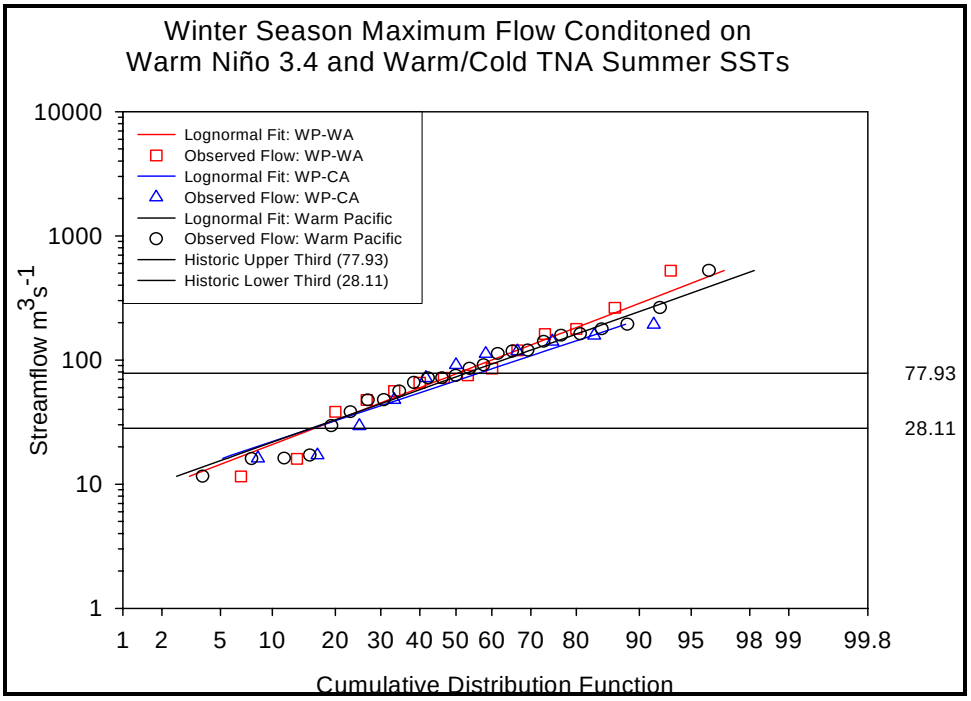

Figure 5.13. Winter season maximum flow differentiated into Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

Figures 5.13 and 5.14 do not indicate any strong sensitivities to the temperature of the TNA.

Winter Season Monte Carlo Simulated Results

The preceding qualitative and simple numerical observations need to be tested quantitatively through the Monte Carlo simulation procedures.

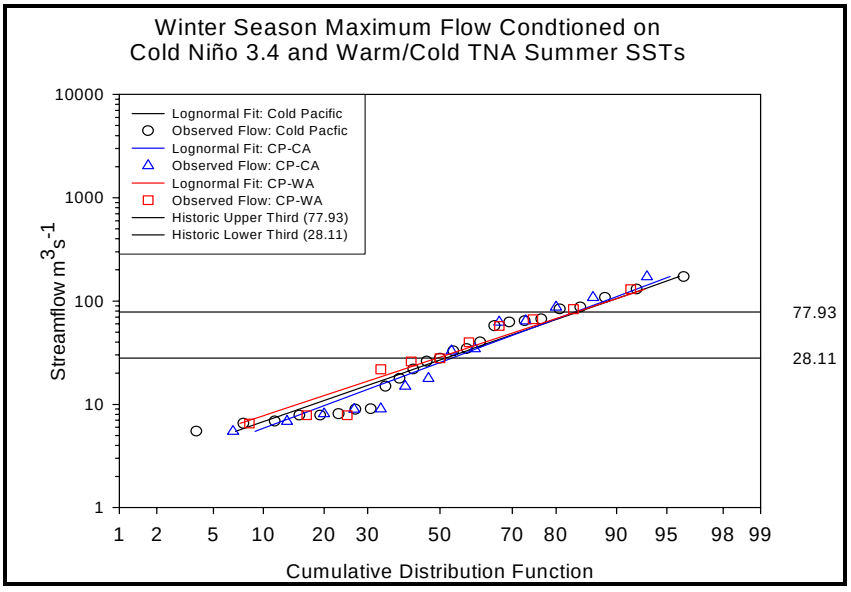

Figure 5.14. Winter season maximum flow differentiated into Cp-Ca/Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

Figure 5.15 displays the results of the Monte Carlo simulations of winter season minimum flows variable conditioned upon a Cp *and* warm or cold conditions in the TNA. The box and whisker plots represent the range of probabilities produced by the simulations (n = 1000) in each flow category. The boxes represent 50% of the simulated probabilites about the median probability value. The upper and lower whiskers represent 75% of the exceedence probability values, while the upper and lower open black circles represent 90% of the values. An outlier can be considered to lie beyond the 75th percentile boundaries represented by the whiskers. Both subpopulations are generated

using the appropriate log mean and log standard deviation of the original Cp subpopulation, but using the record lengths for the Cp-Ca and Cp-Wa subpopulations.

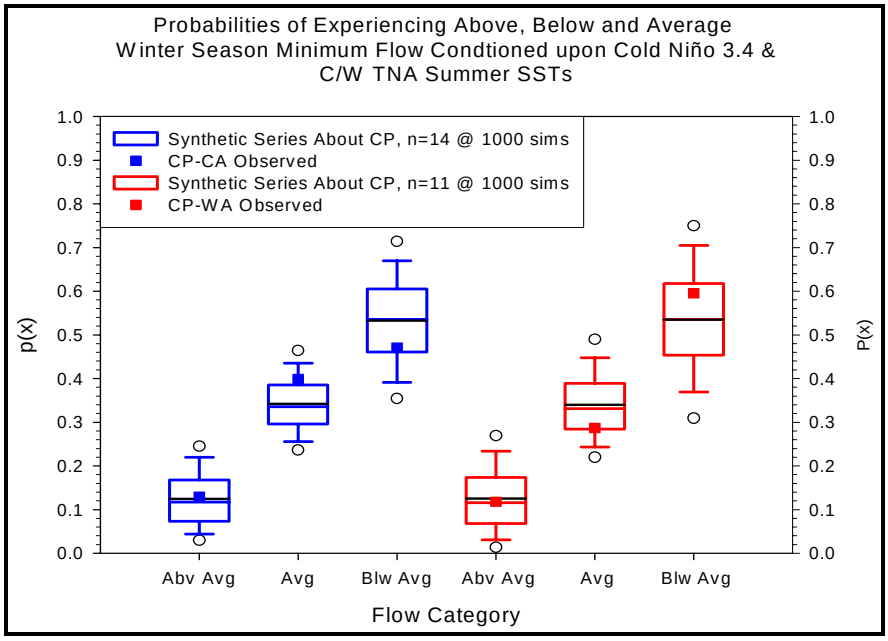

Figure 5.15. Monte Carlo results about the winter season minimum flow variable, Cp. See text for explanation.

The Cp observed probabilities forecast each flow category are 0.124, 0.343, and 0.533 for above average, average, and below average categories, respectively, and are clearly seen in the box and whisker plots. Differences between the red and blue series reflect random chance and differences in sample sizes. The observed Cp-Ca (small squares) and Cp-Wa probabilites all lie within the lower and upper 75th percentile boundaries of their respective simulated flow category, indicating insufficient evidence to suggest that the condition of the TNA affects the risks of flow variables.

Figure 5.16 displays similar results of the Monte Carlo simulation of the same variable conditioned upon a Wp *and* warm or cold conditions in the TNA. The positions of the box and whisker diagrams with respect to the y-axes have changed, reflecting the

different probabilities of flow classes during a warm Pacific, but the interpretation remains the same.

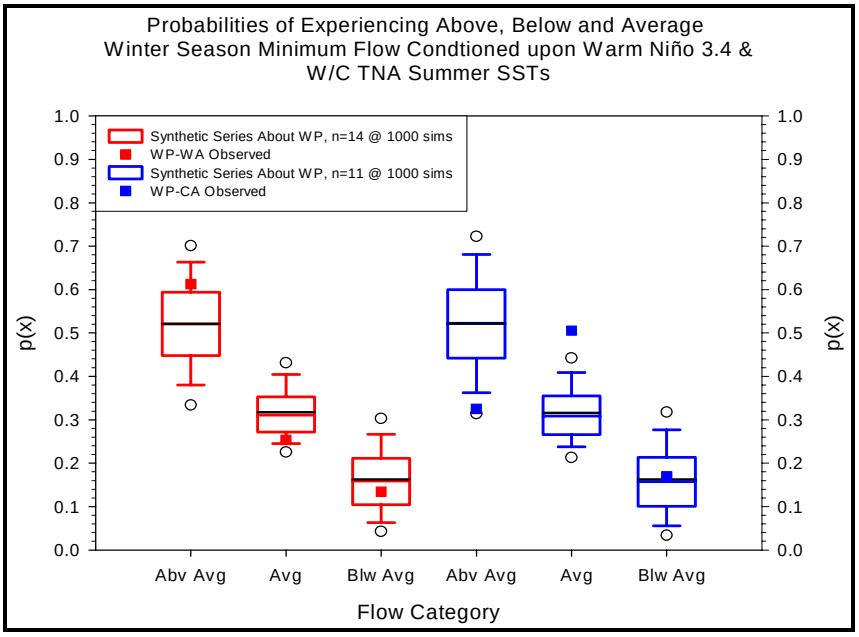

Figure 5.16. Monte Carlo results about the winter season minimum flow variable, Wp. See text for explanation.

The observed probabilites forecast for above average (0.325) and average (0.505) minimum flows under a Wp-Ca condition are beyond the 75th percentile boundaries, implying that the state of the TNA may exert some influence upon the likelihood of experiencing these two conditions.

Figures 5.17 through 5.20 indicate that there is insufficient evidence to believe that TNA SSTs affect either mean or maximum flows during the winter season.

Spring Season Return Period Estimates

Figures 5.21 and 5.22 indicate that the spring drought series appear to be more sensitive to the state of the Pacific than do the flood series. Drought flows of a specific return period are always larger during Wp conditions than Cp conditions. The differentiation is not as clear between flood series.

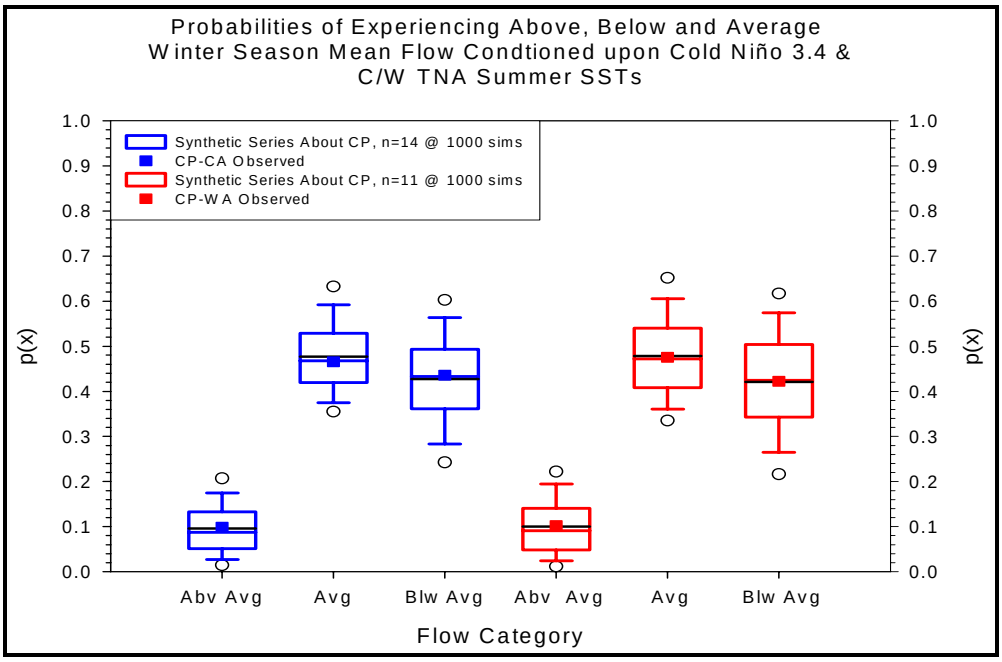

Figure 5.17.. Monte Carlo results about the winter season mean flow variable, Cp. See text for explanation.

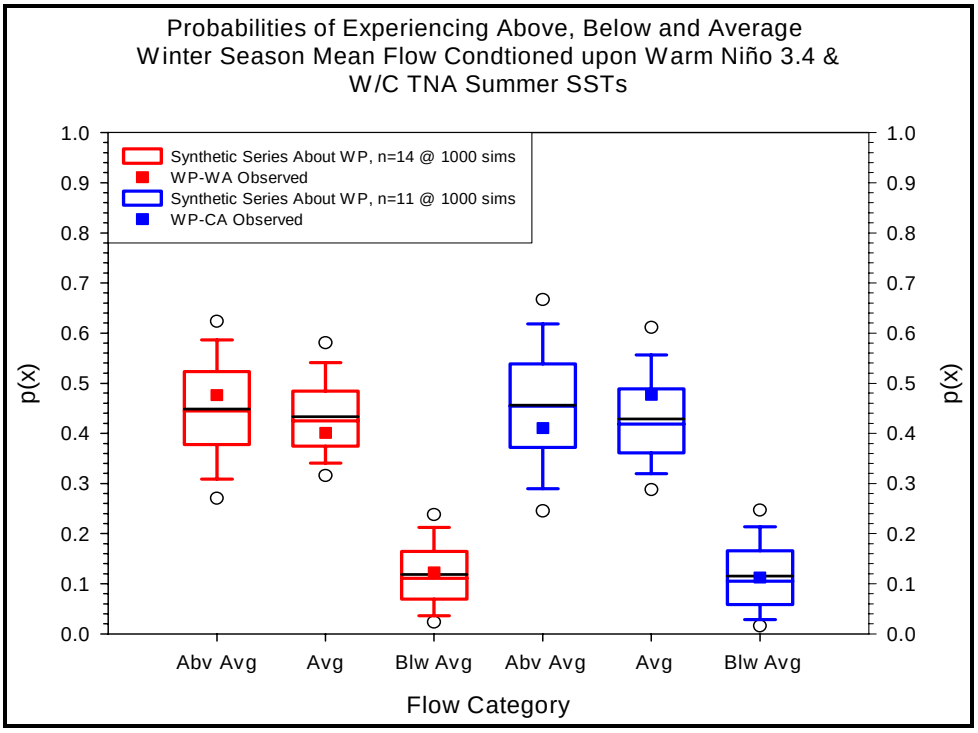

Figure 5.18. Monte Carlo results about the winter season mean flow variable, Wp. See text for explanation.

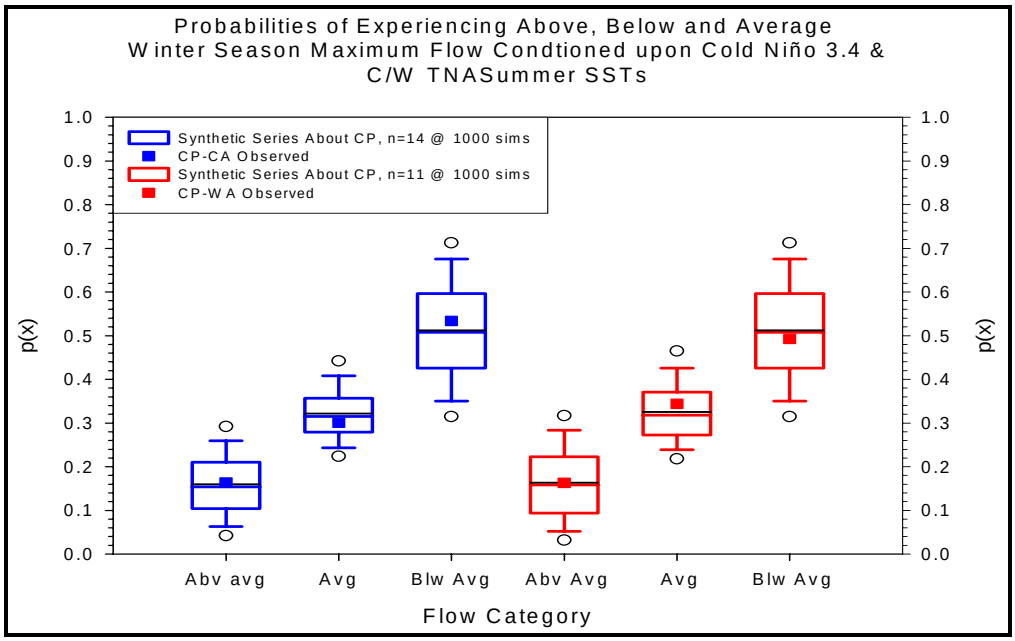

Figure 5.19. Monte Carlo results about the winter season maximum flow variable, Cp. See text for explanation.

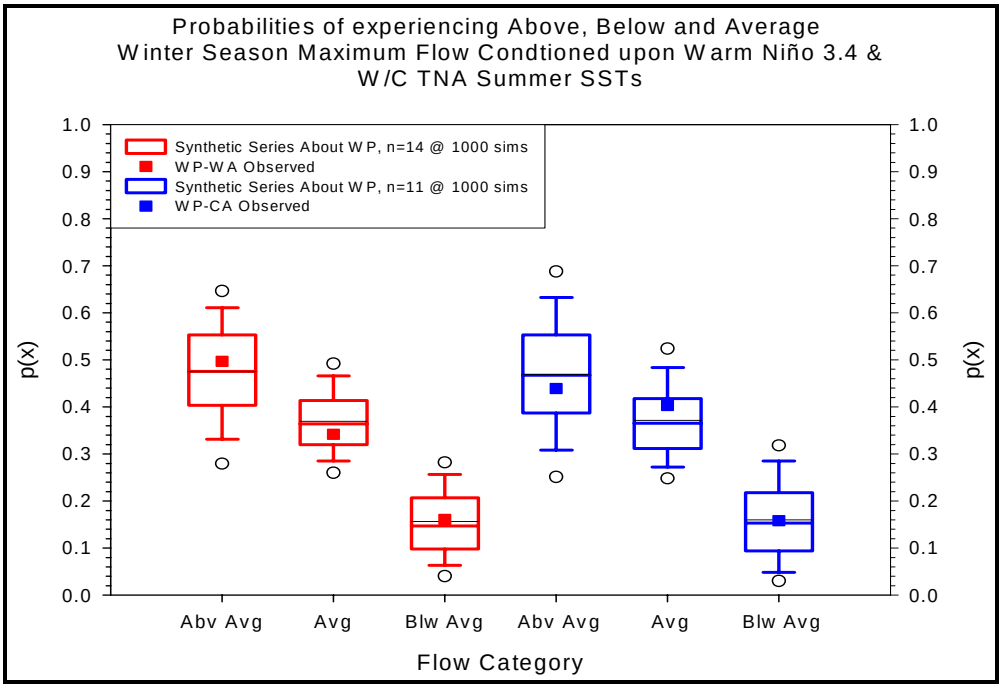

Figure 5.20. Monte Carlo results about the winter season maximum flow variable, Wp. See text for explanation.

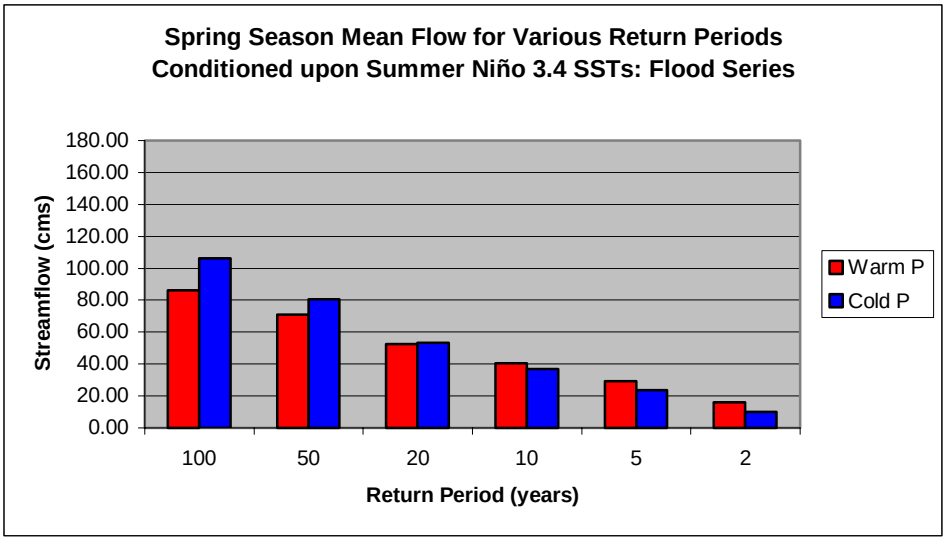

Figure 5.21. Return period estimates for spring season mean flow variable, flood series.

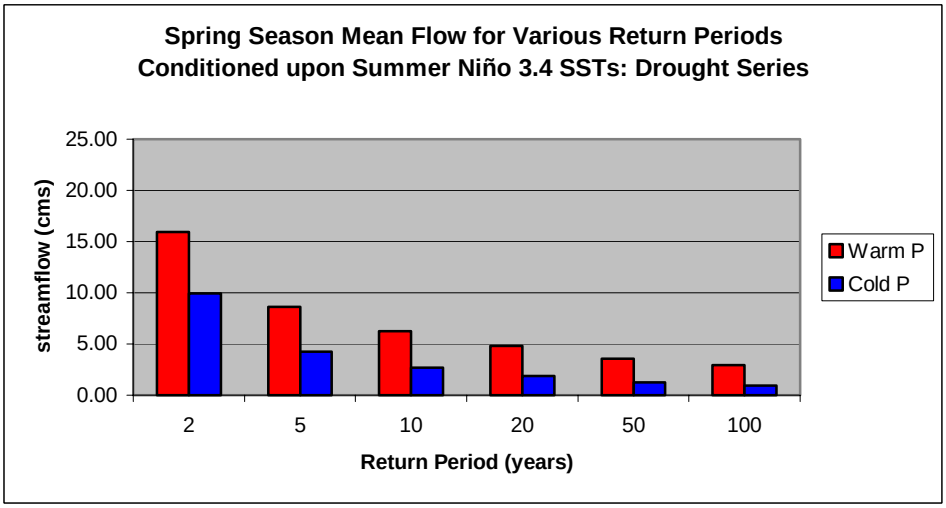

Figure 5.22. Return period estimates for spring season mean flow variable, drought series.

Spring Season Minimum Flow Variable

As for the winter season, the following graphics indicate the exceedence probability distributions of undifferentiated (figure 5.23) and differentiated (figure 5.24) flow variables for the spring season. The probability distributions of all spring season flow variables probability distributions are not significantly different from lognormal at the 80% confidence level.

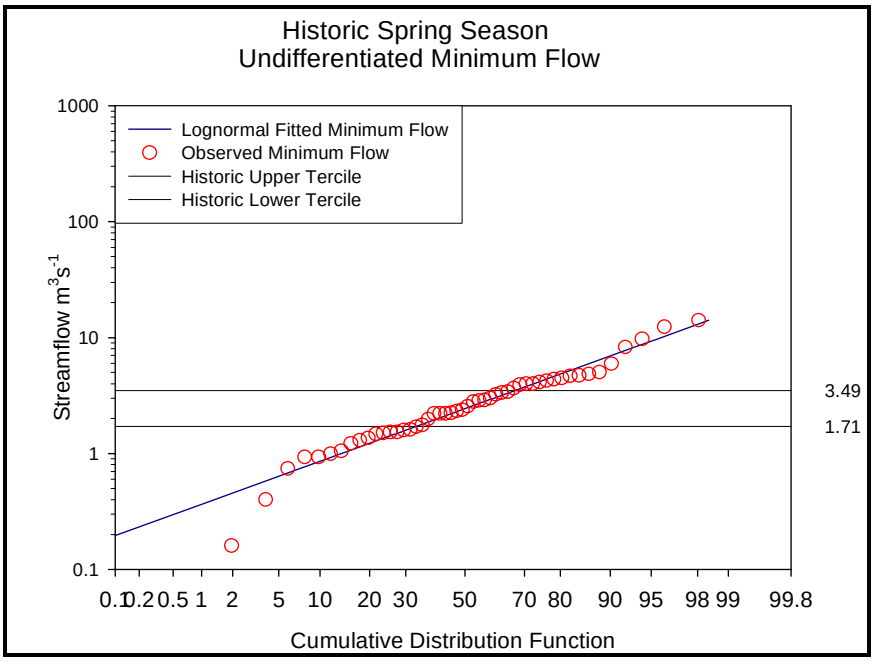

Figure 5.23. Spring season minimum flow lognormal probability distribution.

For spring season minimum flows, "average" flows are defined as being between 1.71 m^3s^{-1} and 3.49 m^3s^{-1}. Figure 5.24 depicts spring season minimum flow conditioned upon Niño 3.4 SSTs from the previous summer.

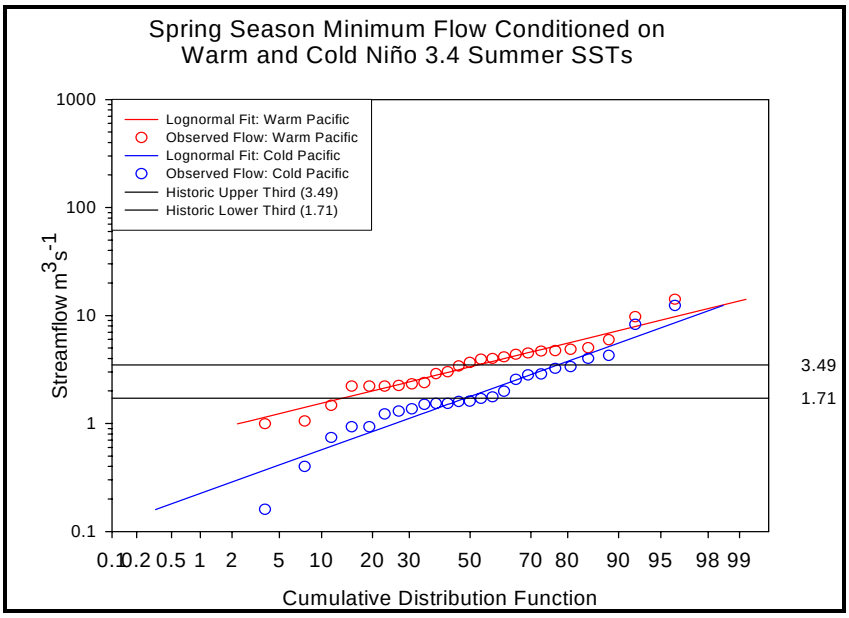

Figure 5.24. Spring season minimum flow differentiated on summer Niño 3.4 SSTs.

The likelihood of experiencing "below" average flow under a Cp condition has increased from 0.333 to 0.482, while under a Wp condition, the probability of experiencing such low flows is only 0.135. The probabilities of experiencing "above" average minimum flows increase to 0.471 under a Wp condition, and decrease to 0.224 under a Cp condition.

Figure 5.25 depicts the exceedence probability distributions of the Wp, Wp-Wa, and Wp-Ca spring minimum flow subpopulations. The probability of experiencing "below" average minimum flow events further conditioned upon TNA SSTs changes very little (0.155 for both Wp-Wa and Wp-Ca conditions, respectively).

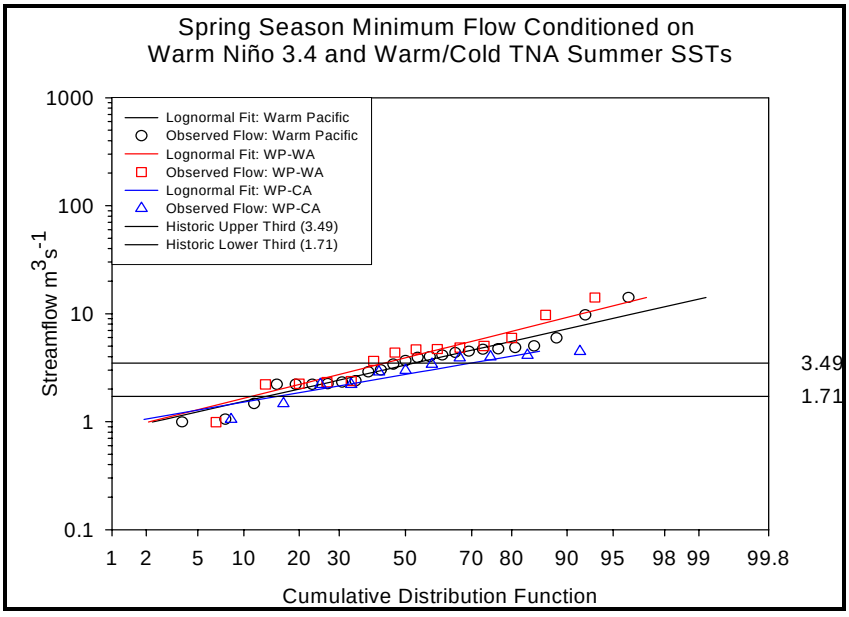

Figure 5.25. Spring season minimum flow differentiated into Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

However, the likelihood of above average minimum flows diverges depending upon the condition of the TNA (Wp-Wa increases to 0.566 and Wp-Ca decreases to 0.300).

Figure 5.26 depicts the probability distributions of minimum flows during a cold Pacific conditioned upon the Atlantic. The probability of experiencing "below" average

flow under an isolated Cp condition remains at 0.482. In combination with a Cold Atlantic this falls to 0.314 and rises to 0.655 in combination with a Warm Atlantic.

The probability of experiencing "above" average flow levels increases slightly to 0.318 with the presence of a Cold Atlantic, and decreases to 0.121 under a warm one. The probabilities of "average" flows are 0.224 and 0.368 under Cp-Wa and Cp-Ca conditions, respectively.

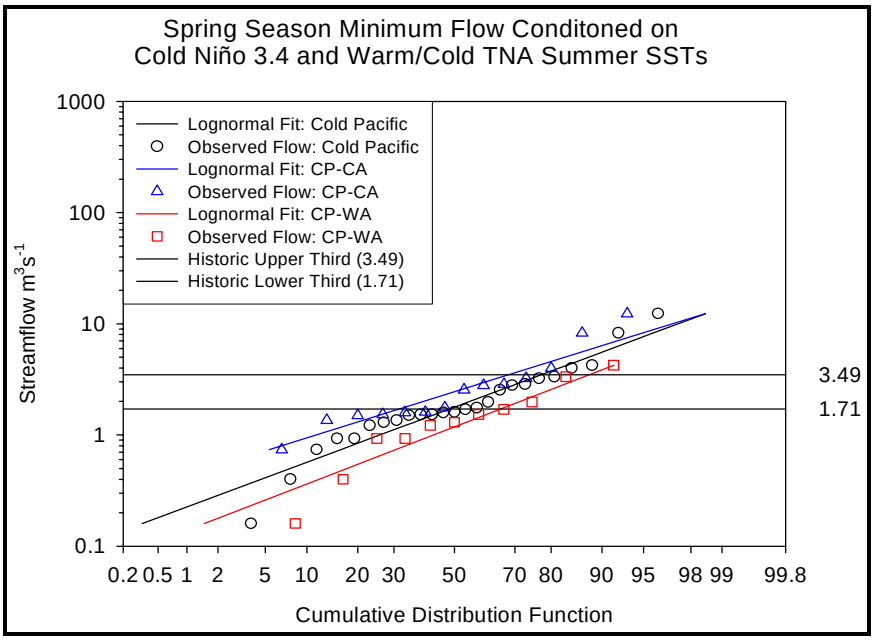

Figure 5.26. Spring season minimum flow differentiated into Cp-Ca/Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

Spring Season Mean Flow Variable

Figure 5.27 displays the distribution of spring season historic mean flows. "Average" mean spring seasonal flow is defined as a flow magnitude between 8.62 m^3s^{-1} and 18.91 m^3s^{-1}. Figure 5.28 depicts spring season mean flow again conditioned upon Niño 3.4 SSTs from the previous summer. The likelihood of experiencing "below" average flow under a Cp condition has increased from 0.333 to 0.444, while under a Wp condition, the probability of experiencing the same flow is just 0.199. The probabilities

of experiencing “above” average mean flows increase to 0.407 under a Wp condition, and decrease to 0.264 under a Cp condition. “Average” flow probabilities are 0.292 under a Cp condition and 0.394 under a Wp condition.

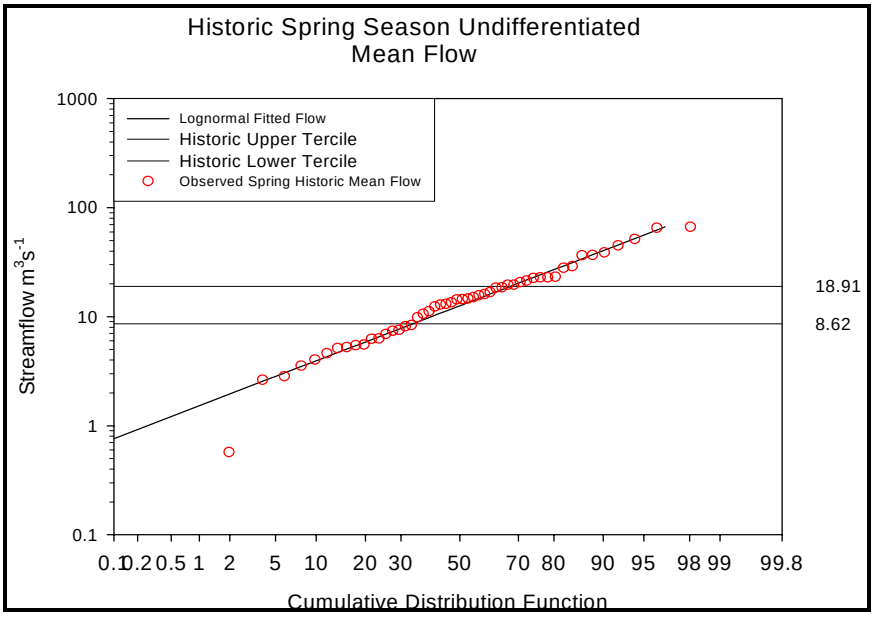

Figure 5.27. Spring season undifferentiated mean flow and probability distribution.

Figure 5.29 shows that the state of the TNA has only a minimal effect on “below” average mean flow probabilities when the Pacific is warm (0.221 under Wp-Wa and 0.157 for Wp-Ca). For the “above” average flow forecast, the effect of the TNA is similarly small, increasing slightly to 0.441 with a warm Atlantic. A cold Atlantic decreases the probability of above average flow to 0.339.

The probability distributions in figure 5.30 show that the effect of the Atlantic on a cold Pacific is similar to the minimum flow variable for “below” average flows. the probability distributions of the Cp-Ca, and Cp-Wa mean flow subpopulations. Here, the probability slightly decreases to 0.320 (Cp-Ca), but increases to 0.562 for Cp-Wa conditions. “Above” average flows display slightly more convergence, decreasing to

0.201 (Cp-Wa) and increasing to 0.326 for a warm Pacific, cold Atlantic condition.

"Average" mean flows remain fairly constant under either condition.

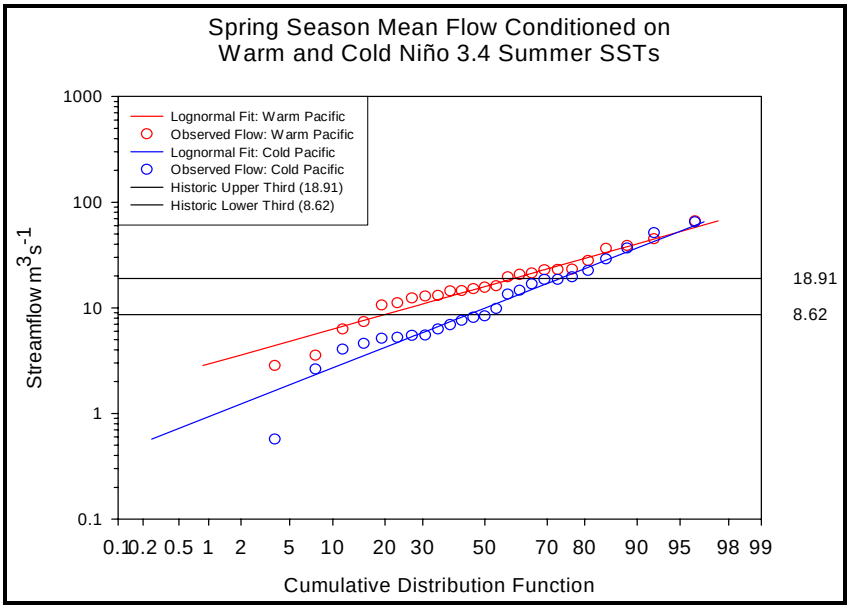

Figure 5.28. Spring season mean flow differentiated on summer Niño 3.4 SSTs.

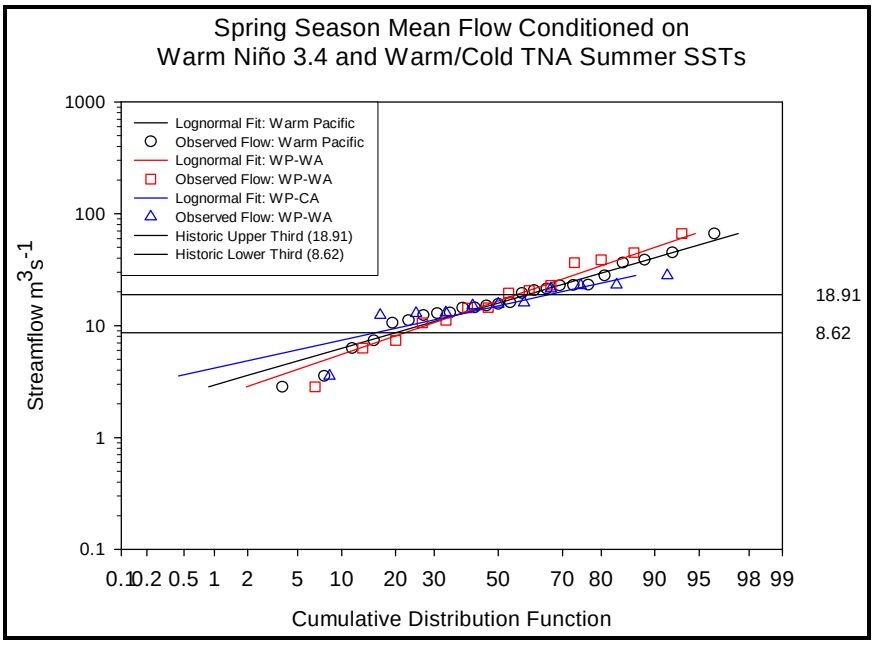

Figure 5.29. Spring season mean flow differentiated into Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

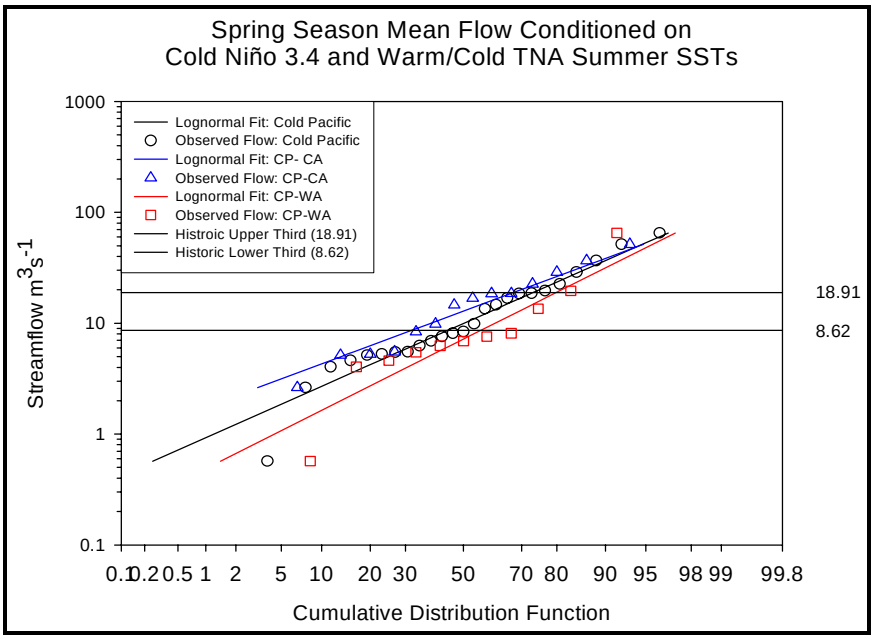

Figure 5.30. Spring season mean flow differentiated into Cp-Ca/Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

Spring Season Maximum Flow Variable

The distribution of the maximum flow variable and terciles are shown (figure 5.31). Average maximum seasonal flow is defined as being between 39.87 m^3s^{-1} and 96.75 m^3s^{-1}. Once again, a similar separation in the distributions and probabilities of flow levels is observed when the data are sub-divided by Niño 3.4 SSTs (figure 5.12), although the convergence in the fitted functions is more pronounced with " above" average exceedence probabilities. Here, "above" average flow events decrease to only 0.259 under a cold Pacific and increase just slightly to 0.344 when the Pacific is warm. The probability of "below" average flow events under a Cp increases to 0.454, while under a Wp the likelihood of experiencing below average events decreases to 0.238.

Figure 5.33 shows that the inclusion of TNA SSTs on a warm Pacific only slightly changes the probabilities of "above" and "below" maximum flow events.

Conversely, there is considerable divergence in “below” average flow probabilities when the state of the Atlantic is considered in combination with a cold Pacific (figure 5.34).

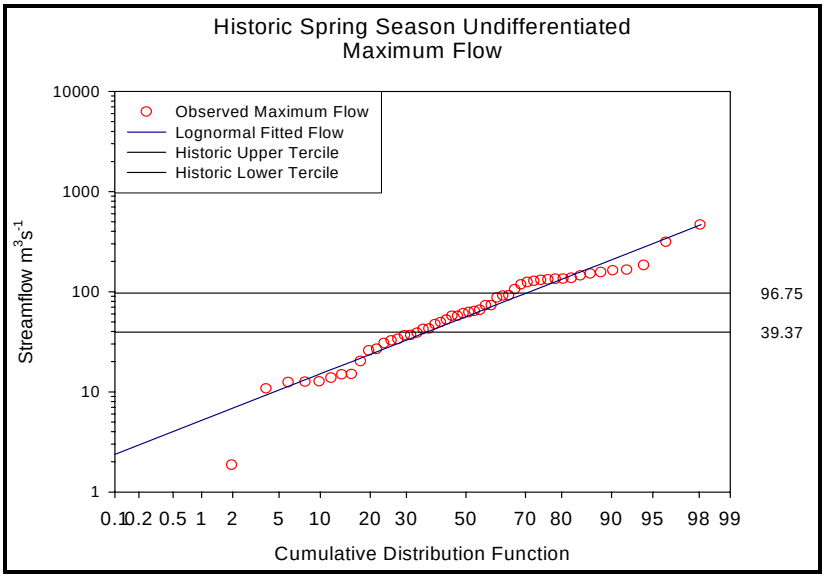

Figure 5.31. Spring season maximum flow lognormal probability distribution.

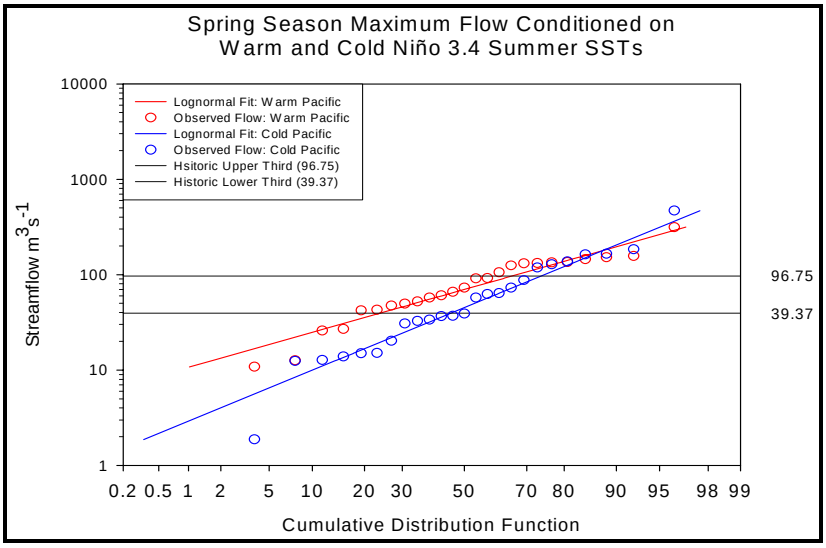

Figure 5.32. Spring season maximum flow differentiated on summer Niño 3.4 SSTs.

When there is a cold Pacific warm Atlantic combination, the probability of “below” average flow events increases to 0.580, while decreasing to 0.293 with a cold Pacific and Atlantic. Neither combination of Atlantic SSTs produces favorable conditions for the

likelihood of "above" average flow, where a Cp-Ca decreases the probability to 0.308 and a Cp-Wa further decreases flow likelihood to 0.202.

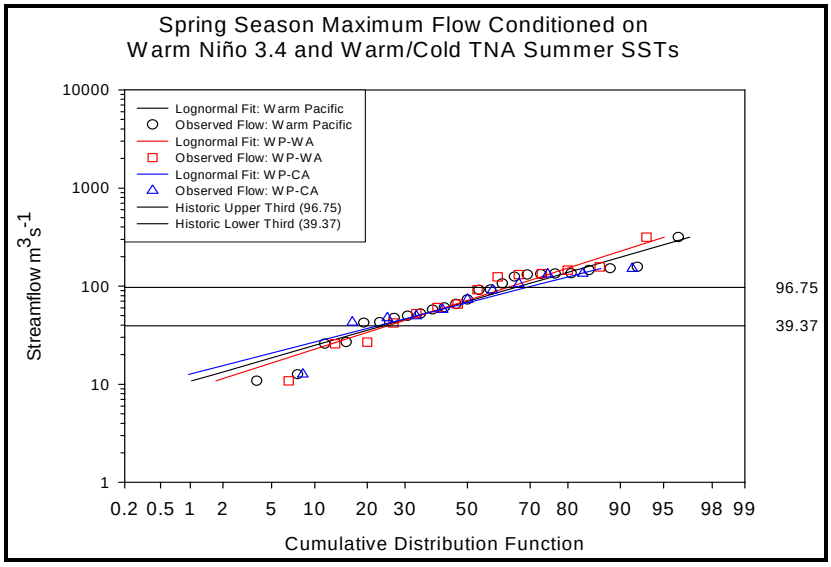

Figure 5.33. Spring season maximum flow differentiated by Wp-Wa/Wp-Ca SST conditions. These distributions are compared to the Wp probability distributions.

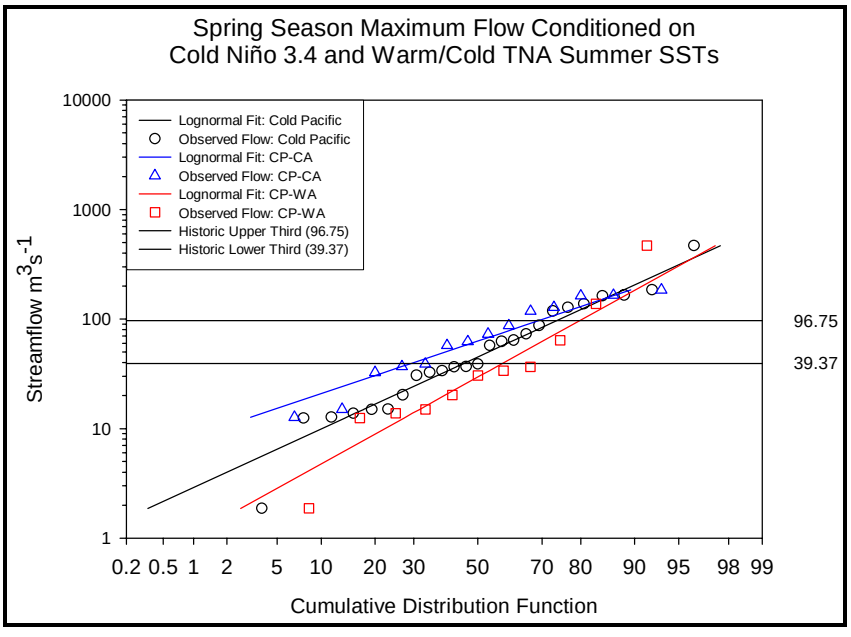

Figure 5.34. Spring season maximum flow differentiated into Cp-Ca/Cp-Wa SST conditions. These distributions are compared to the Cp probability distributions.

Spring Season Monte Carlo Simulated Results

All color schemes, record lengths, and definitions detailed for the winter season remain the same for the spring season Monte Carlo simulations.

Figure 5.35 displays the results of the Monte Carlo simulations of spring season minimum flows conditioned upon a Cp *and* warm or cold conditions in the TNA.

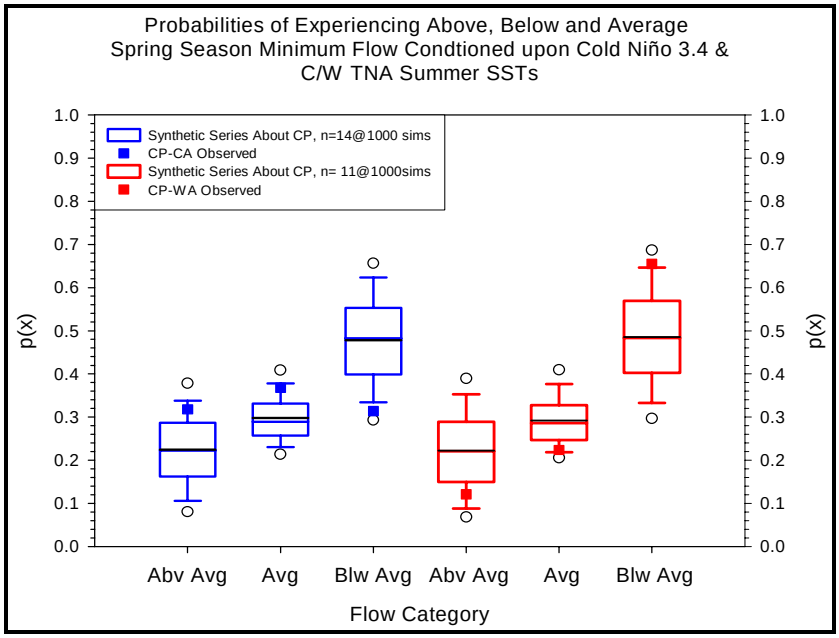

Figure 5.35. Monte Carlo results about the spring season minimum flow variable, Cp. See text for explanation.

The observed probabilities for "below" average flow in both simulated Cp-Ca and Cp-Wa SST combinations lie beyond the lower and upper 75th percentile boundaries, suggesting that the state of the TNA during the previous summer significantly modifies flow response for that category. "Above" average and average flows remain unaffected.

Figure 5.36 displays the results of minimum flows conditioned upon a Wp in the Niño 3.4 *and* warm or cold conditions in the TNA. Only "above" average and average flows appear sensitive to a cold TNA. When both basins are warm there is no significant change in flow forecasts.

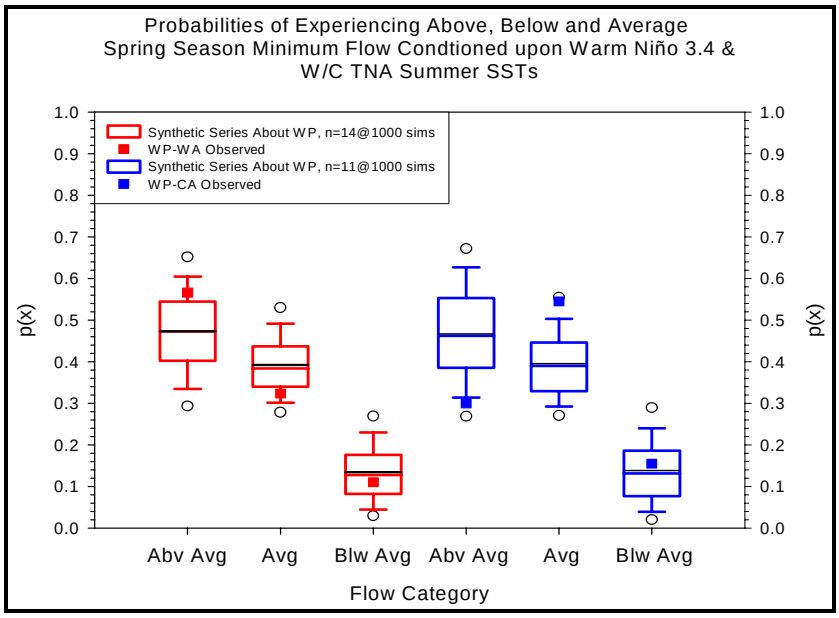

Figure 5.36. Monte Carlo results about the spring season minimum flow variable, Wp. See text for explanation.

With the exception of the average flow category under a Wp, spring season mean flows appear unaffected by the state of the Atlantic, regardless of conditions in the Pacific (figures 5.37 and 5.38).

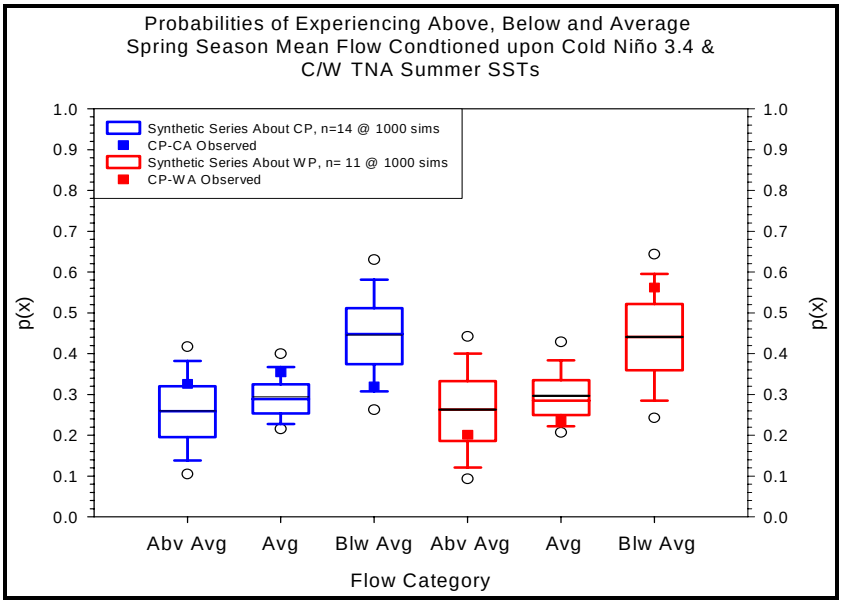

Figure 5.37. Monte Carlo results about the spring season mean flow variable, Cp. See text for explanation.

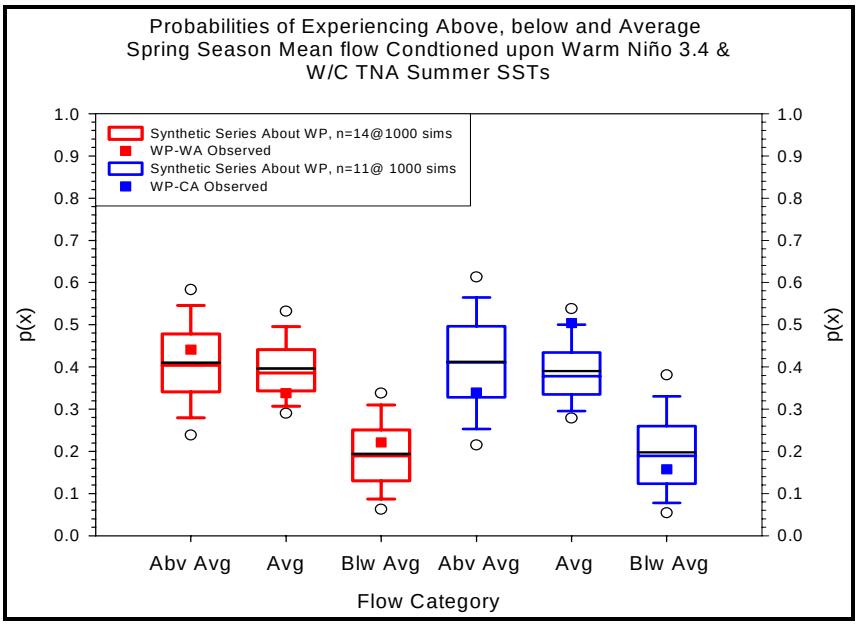

Figure 5.38. Monte Carlo results about the spring season mean flow variable, Wp. See text for explanation.

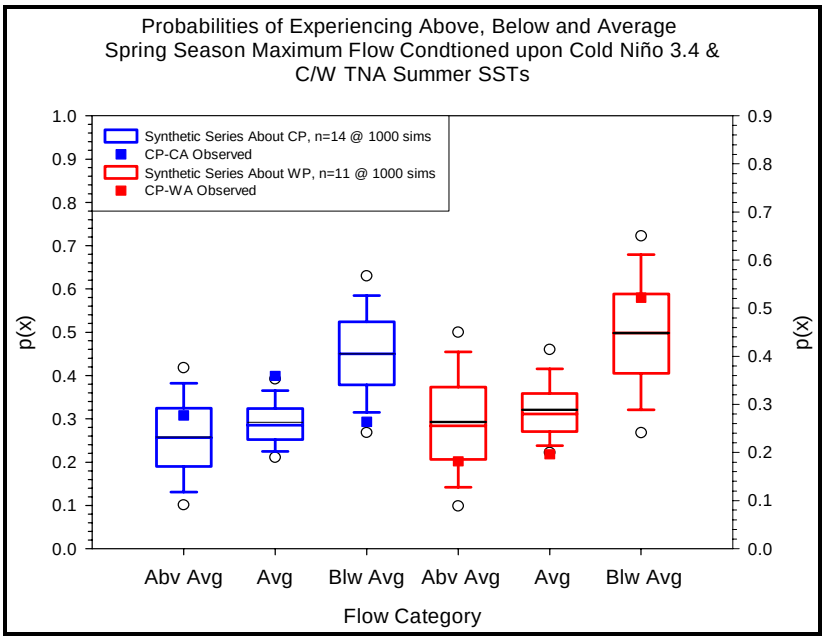

Figure 5.39. Monte Carlo results about the spring season maximum flow variable, Cp. See text for explanation.

As with mean flow, just one category (above average) of the maximum flow variable appears affected by the state of the Atlantic. All other categories of spring season maximum flow remain significantly unchanged (figures 5.39 and 5.40).

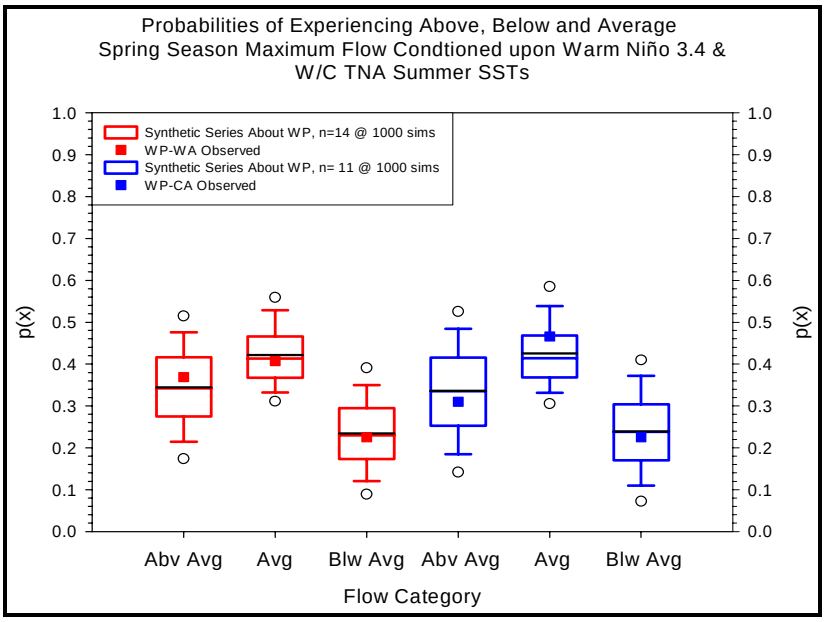

Figure 5.40. Monte Carlo results about the spring season maximum flow variable, Wp. See text for explanation.

Ordinary Least Squares Regression Results

Results of the multiple regression analysis for winter and spring season mean flow are presented.

Winter Season Regression Model Interpretation

As detailed in chapter 3, a forward stepwise procedure is used to eliminate redundant and/or non-influential independent variables. The final regression model is:

- ln(jfmq) = 0.810 + 0.637(ln(ondq)) + 1.104(pacpl1) - 1.283(pacatl_min1) - 1.203(pacmin1_atpl1) +/- 0.556.

This model produces the sample regression function in figure 5.41. The left and right sides are in log and semi-log terms. The OLS sample regression function is in blue and the red asymptotic lines represent the 95% confidence interval about the mean of the dependent variable. Green lines represent the 95% prediction interval about individual observations. Just one observed variable lies outside the 95% prediction interval. Over

half of the observed flows lie outside the 95% confidence interval of the sample regression function.

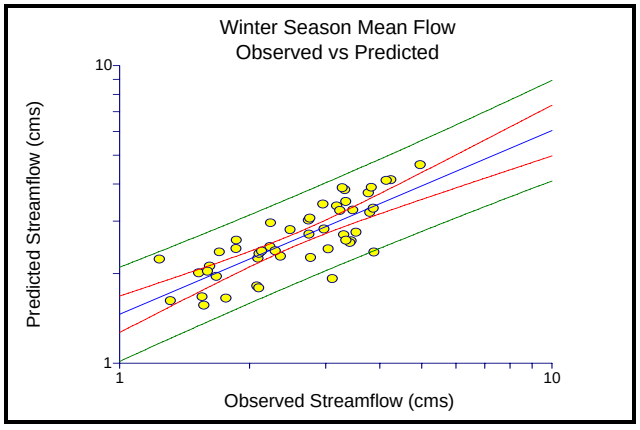

Figure 5.41. Regression output for winter season mean flow. See text for explanation.

This is an indication that the model is likely missing one or more explanatory variables. Table 5.2 shows the beta coefficient estimates, standard errors, and hypothesis tests.

Table 5.2. Winter season multiple regression results.

Effect	Coeff	Std Error	Std Coef	Toler	t	P(2 Tail)
Constant	0.810	0.286	0.000	.	2.831	0.007
ondq	0.637	0.100	0.567	0.941	6.396	0.000
PacPl1	1.104	0.232	0.419	0.949	4.750	0.000
PacAtl_Min1	-1.283	0.564	-0.197	0.989	-2.273	0.028
PacMn1_AtPl1	-1.203	0.406	-0.258	0.973	-2.959	0.005

Multiple R: 0.817 Squared multiple R: 0.667
Adjusted squared multiple R: 0.637 Standard error of estimate: 0.556

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	27.807	4	6.952	22.522	0.000
Residual	13.890	45	0.309		

Case 26 has large leverage (Leverage = 1.000)
Case 49 has large leverage (Leverage = 0.501)
Case 50 has large leverage (Leverage = 0.501)
Durbin-Watson D Statistic 1.815
First Order Autocorrelation 0.070

The constant term (b_o) represents a positive y-intercept of the regression line. The b_1 coefficient represents the amount of change in winter season mean for a one-unit change in fall seasonal flow (ondq). Because x_1 is in log space, the b_1 coefficient is recovered by taking its anti-log. Therefore, a change of 1 m^3s^{-1} in fall season mean discharge will effect a 1.9 m^3s^{-1} increase in winter season mean discharge. The b_2 coefficient may be interpreted as: A greater than 1σ above mean condition in the Niño 3.4 will increase winter mean season mean flow by 1.104 m^3s^{-1}. Both the b_3 and b_4 coefficients are negative and will effect a decrease in winter season streamflow if the they occur during the previous summer.

The adjusted coefficient of determination R^2, indicates that 63.7% of the variation in winter season discharge is explained by model variables, while 36.3% is due to non-systematic forces and/or missing information. The computed F-test value of 22.52 is in the tail of the rejection region, providing ample evidence against the Ho: R^2 not significant (table 5.2).

Winter Season Regression Model Diagnostics

Model diagnostics are provided by several statistics that indicate possible problems in overall model fit or with individual observations. The most important of these is the assumption that model residuals are normally distributed and are of equal variance, shown in figures 5.42 and 5.43.

Both figures indicate no problems with model residuals. The residuals are normally distributed at the 95% confidence level. The residuals vs. predicted plot indicates no problems with functional form or heteroscedasity (an indication of unequal variance). First order auto correlation among the residuals is not a problem in this model as

indicated by the computed vs critical Durbin-Watson D statistic in table 5.2. Observations from 1972, 1999 and 2000 are leveraging the model (i.e. unduly influential), but are not outliers. A leverage value above 0.5 indicates an influential variable.

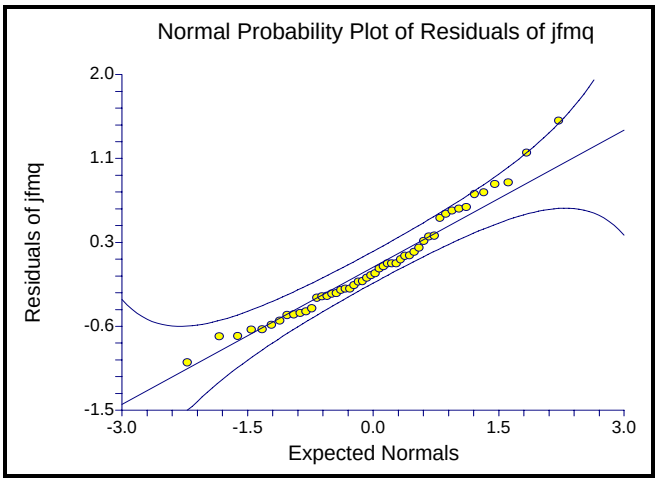

Figure 5.42. Normality plot of winter season OLS model residuals.

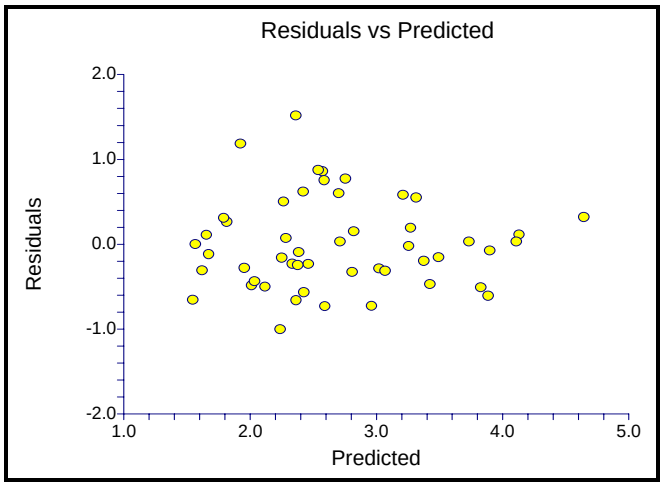

Figure 5.43. Winter season OLS model residuals vs predicted values.

Winter Season Prediction Estimates and Intervals

Figure 5.44 displays the full model prediction estimates and 95% prediction intervals about each estimate. With the exception of the 1959 winter season, all of the

observed seasonal mean flows are contained within the individual prediction intervals, not unexpected as the full model serves as the calibration model.

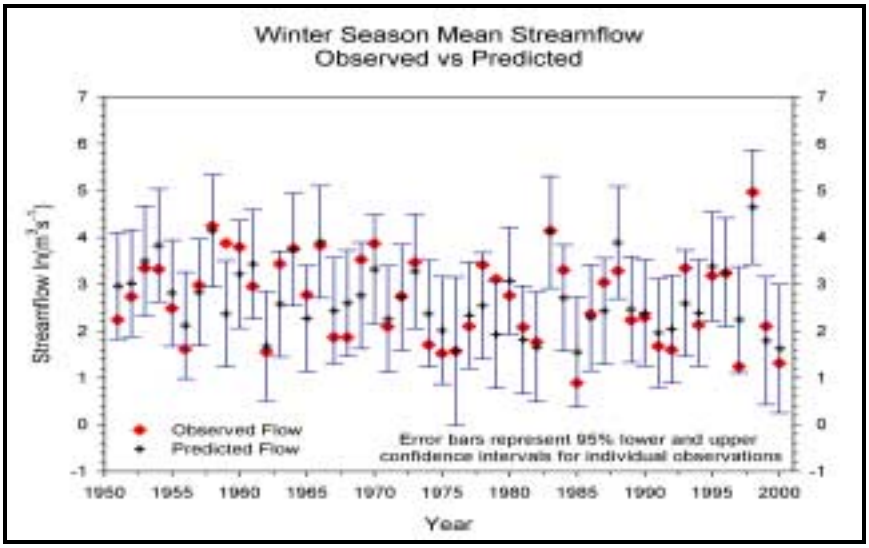

Figure 5.44. Winter season individual 95% prediction intervals.

The spread of the individual prediction intervals about each prediction is a function of the standard error of each prediction, with large standard errors and intervals indicating less model precision. For winter season mean flow, the recovered minimum and maximum prediction interval widths are 9.4 m^3s^{-1} and 23.7m^3s^{-1}.

Spring Season Regression Model Interpretation

The full spring season regression model is:

- ln(amjq) = 1.348 + 0.460(ln(jfmq)) - 1.366(pacmin1_atpl1) +/- 0.735

and produces the sample regression function in figure 5.45. The same color scheme defined for winter applies to the spring season sample regression function (figure 5.45). As with winter season mean flow, only one observed variable lies outside the 95% prediction interval. Several observations lie outside the 95% confidence interval of the sample regression function, again indicating a lack of information in the independent variables.

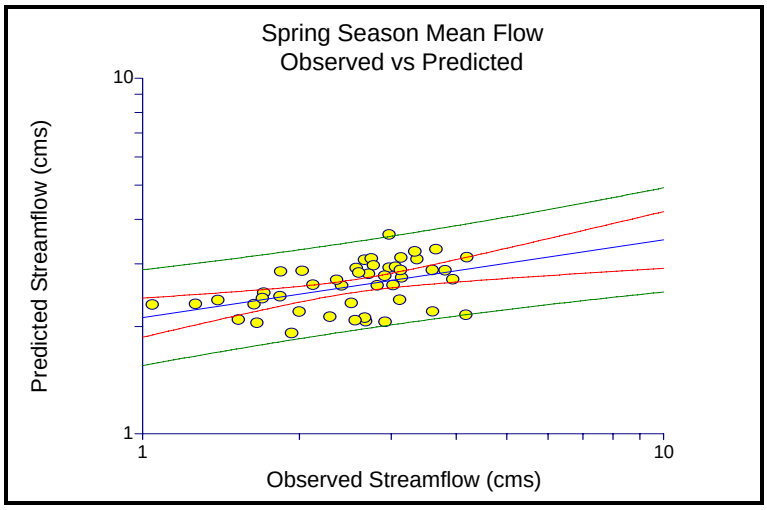

Figure 5.45. Regression output for spring season mean flow. See text for explanation.

Table 5.3 reports the beta coefficients estimates, standard errors, and hypothesis tests. The constant term (b_o) again indicates a positive the y-intercept of the regression line. The partial slope estimator b_1 indicates that a change of 1 m^3s^{-1} in winter season mean discharge will produce a 1.6 m^3s^{-1} increase in spring season mean discharge. The b_2 coefficient will effect a decrease in spring season mean discharge of 1.366 m^3s^{-1}, if present.

The adjusted coefficient of determination R^2, indicates that 34.3% of the variation in spring season discharge is explained by model variables, while the remaining 65.7% is accounted for by non-systematic forces and/or missing information. The computed F-test value of 13.80 is in the tail of the rejection region, providing evidence against the Ho: R^2 not significant (table 5.3).

Spring Season Regression Model Diagnostics

Model diagnostics are again provided by several statistics that indicate possible problems in overall model fit or with individual observations. Figures 5.46 and 5.47 again indicate no concerns with model residuals or assumption violations.

Table 5.3. Spring season multiple regression results.

Effect	Coefficient	Std Error	Std Coef	Toler	t	P(2 Tail)
Constant	1.348	0.336	0.000	.	4.010	0.000
jfmq	0.460	0.117	0.468	0.951	3.946	0.000
Pmin1_Apl1	-1.366	0.544	-0.298	0.951	-2.513	0.015

Multiple R: 0.608 Squared multiple R: 0.370
Adjusted squared multiple R: 0.343 Standard error of estimate: 0.735

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	14.906	2	7.453	13.802	0.000
Residual	25.379	47	0.540		

Case 49 has large leverage (Leverage = 0.504)
Case 50 has large leverage (Leverage = 0.504)
Durbin-Watson D Statistic 2.201
First Order Autocorrelation -0.137

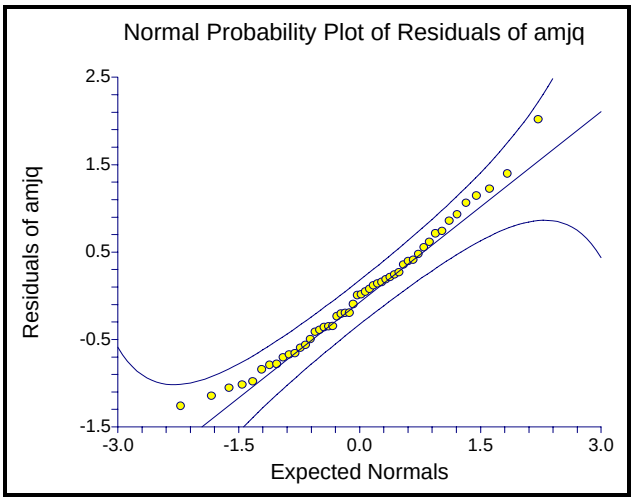

Figure 5.46. Normality plot of spring season OLS model residuals.

The residuals are normally distributed at the 95% confidence level, while the residuals vs. predicted plot indicates no concerns with improper functional form or heteroscedasity. First order auto correlation among the residuals is not a problem in this model as

indicated by the computed vs critical Durbin-Watson D statistic in table 5.3. Observations from 1999 and 2000 are leveraging the model, but are not outliers.

Spring Season Prediction Estimates and Intervals

Figure 5.48 displays the full model prediction estimates and 95% prediction intervals about each estimate. As above, all of the observed seasonal mean flows are contained within the individual prediction intervals excepting the 1982 seasonal flow. The intervals about each individual estimate of spring season mean discharge are

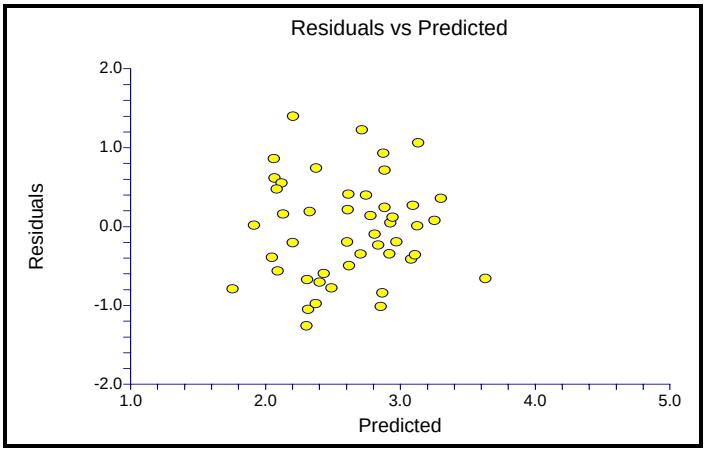

Figure 5.47. Spring season OLS model residuals vs predicted values.

larger than the intervals produced by the model for winter season mean discharge, reflecting less predictive power. For spring season mean flow, the recovered minimum, mean, and maximum prediction interval widths are:

- Min 95% PI = 19.8 m^3s^{-1}
- Max 95% PI = 37.5 m^3s^{-1}

It should be noted that the empirical prediction estimates and intervals do not provide information about scenarios where conditions are changed in the independent variables. For example, what is the winter season mean flow estimate when fall season mean flow is at maximum *and* the Niño 3.4 region is greater than 1σ above its long-term mean? Or, if none of the conditions specified by the models are present in the Niño 3.4

or TNA during the previous summer, then the previous season's flow is the best predictor to estimate the next seasons' flow. Scenarios such as these will be included in the discussion section.

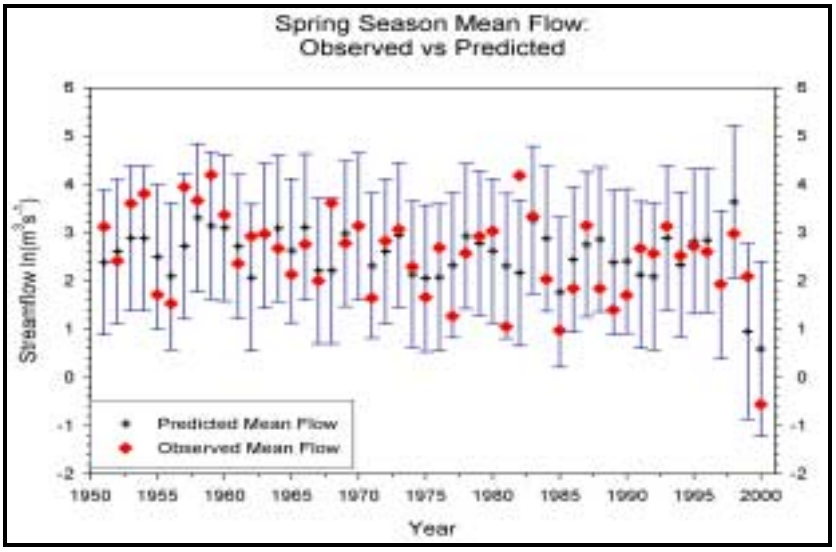

Figure 5.48. Spring season individual 95% prediction intervals.

Three-Way Contingency Table Results

Winter Season

A portion of this thesis seeks to determine the conditional likelihood of experiencing a season during which the number of low-flow days is below a threshold value of 3.68 $m^3 s^{-1}$ (130 $ft^3 s^{-1}$). The partial associations we seek to identify are: Do SST conditions in the Niño 3.4 or in the TNA have the greater influence on winter/spring numbers of low flow days? Table 5.4 provides the observed frequency counts of numbers of seasons during which greater and less than the expected number (8) of low-flow days occurred.

The probability and/or odds of experiencing a season during which the number of low flow days is *greater than* 8 (23 for spring), controlling for the TNA, will indicate a success. The probability and/or odds of experiencing a season during which the number

of low flow days *is less than* 8 (23 for spring), controlling for the TNA, will indicate a failure.

In row one (Ca), the probability of success is equal to: (6/25) = 0.24. In row 2 (Wa), the probability of success is equal to (6/25) = 0.24. In rows 1 and 2, the probability of failure is equal to: (19/25) = 0.76. The *relative risk* then is equal to (0.24 / 0.24) = 1. That is, the chance of success is the same under either SST condition (table 5.5).

Table 5.4. Conditional response of low flow days, controlling for the TNA

	Frequency,# seas w > 8 days flow < 130	# seas w < 8 days flow ,< 130	Total
Cold			
	6	19	25
	6	19	
	12.00	38.00	50.00
	24.00	76.00	
	50.00	50.00	
Warm ^			
	6	19	25
	6	19	
	12.00	38.00	50.00
	24.00	76.00	
	50.00	50.00	
Total	12	38	50
	24.00	76.00	100.00

In rows 1 and 2, the *odds* of success are (0.24 / 0.76) = 0.316 or (1 / 0.316) = 3.165. This means that the odds of a failure in either row are 3.2 times the odds of a success. Since the odds of a success and failure are the same in both rows, the odds ratio is one. Therefore, the odds of a success are no greater than the odds of a failure for a given SST condition, controlling for the TNA. The goodness of fit statistics X^2 and G^2 in table 5.5 indicates that the probability, odds, relative risk, and odd ratios of success are independent of SST condition in the TNA at the 95% confidence level.

Table 5.6 tests the same association, *controlling for the Niño 3.4 SST condition*.

Table 5.5. Goodness of fit statistics and estimates of relative risk for winter season low flow days.

Goodness of Fit Statistics	DF	Value	Prob
Chi-Square	1	0.0000	1.0000
Likelihood Ratio Chi-Square	1	0.0000	1.0000

Odds Ratio and Relative Risk

	Value	95% Confidence Limits	
Odds Ratio	1.0000	0.2731	3.6620
Relative Risk	1.0000	0.3729	2.6818

The probability of success, *controlling for Niño 3.4,* is equal to 0.40, while the probability of failure is 0.60 conditioned on a Cp (row 1). Under a Wp condition (row 2), the probability of success is 0.08, while the probability of failure is 0.92. The relative risk ($1 - p_1 / 1 - p_2$) of 5.00 indicates that the sample proportion of success is 5 times (or 500%) higher under a Cp than under a Wp condition, suggesting that the risk of experiencing a season with more than the long-term mean of low flow days is much lower. In row 1 (Cp) the odds of success are 0.667 times the odds of a failure, or the odds of a failure are 1.5 times the odds of success. Under a Wp, the odds of a success are.087 times the odds of failure, or the odds of failure are 11.5 times the odds of success. The odds ratio then is 7.67, indicating that the odds of experiencing a season during which there are more than 8 days of flow less than 3.68 $m^3 s^{-1}$ under a Cp is 7.67 times the odds of the opposite conditions, controlling for the Niño 3.4.

The goodness of fit statistics indicate that all probabilities, relative risks, odds and odds ratios are statistically significant. The wide 95% confidence intervals about both the relative risk and odds ratios are a function of small sample size; smaller samples create wider intervals (tables 5.7 and 5.8).

Table 5.6. Conditional response of low flow days, controlling for Niño 3.4

Frequency, Expected, Percent, Row Pct, Col Pct	# seas w > 8 days flow < 130	# seas w < 8 days flow < 130	Total
C	10	15	25
	6	19	
	20.00	30.00	50.00
	40.00	60.00	
	83.33	39.47	
W	2	23	25
	6	19	
	4.00	46.00	50.00
	8.00	92.00	
	16.67	60.53	
Total	12	38	50
	24.00	76.00	100.00

Table 5.7. Goodness of fit statistics and estimates of relative risk for winter season low flow days, controlling for Niño 3.4.

Statistic	DF	Value	Prob
Chi-Square	1	7.0175	0.0081
Likelihood Ratio Chi-Square	1	7.5189	0.0061

Table 5.8. Odds ratio and relative risk, winter low-flow days, controlling for Niño 3.4.

Odds Ratios and Relative Risk

	Value	95% Confidence Limits	
Odds Ratio	7.6667	1.4699	39.9866
Relative Risk	5.0000	1.2166	20.5485

Table 5.9 displays the frequencies of spring seasons during which the number of low flow days was greater than or less than the long-term mean.

Spring Season

In row one (Ca), the probability of success is 0.44, while the probability of failure is equal to 0.56. In row 2 (Wa), the probability of success is 0.40, while probability of failure is 0.60. The relative risk then is 1.10. The chance of success (greater than 23 days of low flow) is just 10% greater under a Ca.

In row 1 the odds of success are 0.786 times the odds of failure, or the odds of failure are 1.27 times the odds of success.

Table 5.9. Conditional response of low flow days, controlling for the TNA.

Frequency, Expected , Percent , Row Pct , Col Pct ,	# seasons w/> 23 days flow < 130	# seasons w/< 23 days flow < 130	Total
C	11 10.5 22.00 44.00 52.38	14 14.5 28.00 56.00 48.28	25 50.00
W	10 10.5 20.00 40.00 47.62	15 14.5 30.00 60.00 51.72	25 50.00
Total	21 42.00	29 58.00	50 100.00

In row 2, the odds of success or failure are 0.667 and 1.5, respectively. The odds of a success are not significantly greater than the odds of a failure for a given SST condition, controlling for the TNA. The goodness of fit statistics indicates that the probability, odds, relative risk, and odd ratios of success are independent of SST condition in the TNA (table 5.10).

Table 5.10. Goodness of fit statistics and estimates of relative risk and odds ratios for spring season low flow days, controlling for TNA.

Statistic	DF	Value	Prob
Chi-Square	1	0.0821	0.7745
Likelihood Ratio Chi-Square	1	0.0821	0.7744

	Value	95% Confidence Limits	
Odds Ratio	1.1786	0.3829	3.6274
Relative Risk	1.1000	0.5727	2.1128

Table 5.11 indicates that the probability of success, *controlling for Niño 3.4,* is equal to 0.64. The probability of failure is equal to 0.36 conditioned on a Cp (row 1).

Under a Wp condition (row 2), the probability of success is 0.20, while the probability of failure is 0.80.

Table 5.11. Conditional response of low flow days, controlling for Niño 3.4.

Frequency, Expected , Percent , Row Pct , Col Pct ,	# seasons w/> 23 days flow < 130	# seasons w/< 23 days flow < 130	Total
C	16 10.5 32.00 64.00 76.19	9 14.5 18.00 36.00 31.03	25 50.00
W	5 10.5 10.00 20.00 23.81	20 14.5 40.00 80.00 68.97	25 50.00
Total	21 42.00	29 58.00	50 100.00

The relative risk of 3.2 indicates that the sample proportion of success is 3.2 times (or 320%) higher under a Cp than under a Wp condition. In row 1 (Cp) the odds of success are 1.78 times the odds of a failure. Under a Wp, the odds of a success are 0.25. times the odds of failure. The odds ratio then is 7.12, suggesting that the odds of experiencing a season during which there are more than 23 days of flow less than 3.68 m^3 s^{-1} under a Cp is 7.12 times the odds under opposite conditions, controlling for the Niño 3.4. The goodness of fit statistics indicate that all probabilities, relative risks, odds and odds ratios are statistically significant (table 5.12).

Table 5.12. Goodness of fit statistics and estimates of relative risk for spring season low flow days, controlling for Niño 3.4.

Statistic	DF	Value	Prob
Chi-Square	1	9.9343	0.0016
Likelihood Ratio Chi-Square	1	10.3382	0.0013

	Value	95% Confidence Limits	
Odds Ratio	7.1111	1.9858	25.4651
Relative Risk	3.2000	1.3852	7.3924

Loglinear Regression Results

Winter Season

For winter season low flow days, the final model is:

$$\ln(\hat{u}) = 22.663 - 0.686(X_1), \text{ where } X_1 = \text{mean SST for each class,}$$

and the mean number of low flow days per class ($\hat{u}$) is: $\hat{u} = \exp^{\{(22.663 - 0.686(SST)\}}$.

Tables 5.13 and 5.14 provide details on model parameters:

Table 5.13. Overall winter season log-linear model fit, all observations.

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance (G^2)	3	5.0691	1.6897
Pearson Chi-Square (X^2)	3	6.4788	2.1596
Log Likelihood		1404.2317	

Table 5.14. Winter season loglinear model parameters.

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	22.6632	9.0462	4.9329	40.3935	6.28	0.0122
pac	1	-0.6863	0.3380	-1.3488	-0.0237	4.12	0.0423
Dispersion	1	0.2033	0.1325	0.0567	0.7296		

LR Statistic

	DF	Chi-Square	P-Value
pac	1	3.01	0.0829

Overall model fit is provided by the X^2 and G^2 statistics, and since this is a right-tail test, large values indicate poor fit. Model fit is adequate under the 95% confidence level with 3 degrees of freedom (df). The dispersion parameter is the model generated scaling factor for the asymptotic standard error (ASE), which adjusts the ASE such that variance of the data is now proportional to the mean. The log likelihood value is related to the likelihood ratio statistic, where the value reported here is the maximum value of the likelihood function when β is not forced to zero. The value is arbitrary; the importance of this measure lies in the likelihood ratio statistic for β shown in table 5.13.

Model parameters may be interpreted as follows. The β_1 coefficient has a negative multiplicative effect, indicating an inverse relationship between the number of low flow

days and SST for each class. The Wald-approximate chi-square statistic for the null hypothesis Ho: $\beta = 0$ proves weak evidence against the Ho. Additionally, the likelihood ratio statistic provides evidence against the Ha: $\beta \neq 0$, suggesting that the predictions and prediction intervals in table 5.15 may not be the best estimates.

Table 5.15. Number of expected low-flow days per Niño 3.4 SST class.

Category	Numdays	Pactemp	Pred	Lower	Upper	StReschi
C	83	26.36	96.845	59.616	157.323	0.255
N	154	26.72	75.645	50.202	113.983	1.804*
VC	100	25.79	143.206	67.101	305.630	-0.615
VW	29	27.87	34.358	14.747	80.049	0.323
W	32	27.12	57.486	35.708	92.548	-0.782

Table 5.16 Scaled number of low-flow days per season (winter)

Category	Observed	Predicted	Interval Width
VC	10.0	14.3	6.7 – 30.5
C	8.3	9.6	5.9 – 15.7
N*	15.4	7.5	5.0 – 11.3
W	3.2	5.7	3.5 – 9.2
VW	2.9	3.4	1.4 – 8.0

When scaling the number of observed and predicted days per season for a given Niño 3.4 SST category, the following number of low-flow days is recovered for each season in table 5.16.

There are problems with the neutral SST category. The standardized chi-square residual of 1.8 indicates that there may be a season in the historical record where an extraordinary number of low flow days occur. After close inspection of the historic record, the year 1985 experienced 85 days of flow below 3.68 $m^3 s^{-1}$. After removing this observation, the model was reapplied. The adjusted model is:

$\ln(\hat{u}) = 21.767 - 0.663(X_1)$, and the mean number of low flow days per class is:

$$\hat{u} = \exp^{((21.767 - 0.663(SST))}.$$

The new model estimates for the number of low flow days given SST category are listed in table 5.18

Table 5.17. Goodness of fit and parameter estimates for the 1985 case-omitted winter season low-flow model.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	3	5.5197	1.8399
Pearson Chi-Square	3	5.2346	1.7449
Log Likelihood		1007.2116	

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr> ChiSq
Intercept	1	21.7667	2.8016	16.2757	27.2577	60.36	<.0001
pac	1	-0.6625	0.1055	-0.8692	-0.4557	39.43	<.0001
Dispersion	1	0.0061	0.0162	0.0000	1.1012		

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
pac	1	10.96	0.0009

Overall model fit has improved to under the 90% confidence level at 3 df, (table 5.17). The negative multiplicative effect of β_1 again indicates an inverse relationship between SST and low-flow days. Both the Wald and likelihood ratio statistics provide much stronger evidence against the null hypothesis Ho: $\beta = 0$, and the ASE of the β coefficient has decreased.

Table 5.18. Adjusted number of expected low-flow days per Niño 3.4 SST class.

Category	Numdays	Pactemp	Pred	Lower	Upper	StReschi
C	83	26.36	74.032459	64.45253	85.036304	0.2588977
N	68	26.72	58.32446	50.825726	66.929545	0.3516195
VC	100	25.79	107.99665	87.923625	132.65236	-0.162061
VW	29	27.87	27.226857	20.092248	36.894913	0.1378219
W	32	27.12	44.747765	37.455651	53.459556	-0.601233

When scaling the number of observed and predicted days per season for a given Niño 3.4 SST category, the following number of low-flow days is recovered for each season (table 5.19):

Table 5.19. Adjusted scaled number of low-flow days per season.

Category	Observed	Predicted	Interval Width
VC	10.0	10.7	8.7 - 13.2
C	7.4	7.4	6.4 - 8.5
N	6.8	5.8	5.0 - 6.6
W	3.2	4.4	3.7 - 5.3
VW	2.9	2.7	2.0 - 3.6

The adjusted model with observation year 1985 deleted shows a much-improved fit, although the observed number of low-flow days in the N and W categories is not contained in the 95% prediction interval. The outlier within the neutral SST class has disappeared, and the prediction interval width has narrowed.

Spring Season

As with forecasting winter season low flow days, a forward stepwise procedure is used to estimate the final model:

$$\ln(\hat{u}) = 21.767 - 0.663(X_1).$$

Inclusion of TNA SSTs as a main effect predictor or as an interactive term provided no help in model improvement. Table 5.21 details spring model parameters:

Table 5.21. Goodness of fit and parameter estimates for the TNA predictor-omitted spring season low-flow model.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	3	4.9516	1.6505
Pearson Chi-Square	3	5.1567	1.7189
Log Likelihood		5247.3597	

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr> ChiSq
Intercept	1	27.5366	3.8353	20.0195	35.0537	51.55	<.0001
pac	1	-0.8310	0.1435	-1.1122	-0.5499	33.55	<.0001
Dispersion	1	0.0389	0.0277	0.0096	0.1575		

LR Statistics

Source	DF	Chi-Square	Pr > ChiSq
pac	1	10.46	0.0012

Table 5.22. Spring season low flow day model estimates.

		Observation Statistics				
Category	Numdays	Pactemp	Pred	Lower	Upper	StReschi
C	210	26.36	278.74773	225.60556	344.40773	-0.440912
N	281	26.72	206.67168	171.77141	248.66294	0.6391982
VC	429	25.79	447.64572	324.58995	617.35334	-0.07634
VW	66	27.87	79.474099	54.822518	115.21055	-0.309839

The overall model holds at under the 90% confidence level (table 5.21). The negative multiplicative effect of β_1 again suggests an inverse relationship between SST and low-flow days. Both the Wald and likelihood ratio statistics provide strong evidence against the null hypothesis Ho: $\beta = 0$. Model estimates for the number of low flow days given SST category are listed in table 5.22.

Model residuals indicate no serious outliers or problems in any SST category. The worst fit is in the neutral category, although the standardized chi-square residual is well within the fail to reject region. In the cold and neutral categories, the observed number of low flow days is not contained within the 95% prediction interval (table 5.23), suggesting that cool and cool-to warm SSTs are not as influential in driving long-term variability.

Table 5.23. Number of low-flow days per season (spring).

Category	Observed	Predicted	Interval Width
VC	42.9	44.7	32.4 – 61.7
C	21.0	27.8	22.5 – 34.4
N	28.1	20.6	17.1 – 24.8
W	16.1	14.8	11.9 – 18.3
VW	6.6	7.9	5.4 – 11.5

Winter Season Precipitation Forecast Results

This section presents two examples of area-weighted, winter season precipitation. The first precipitation map is created via an ordinary kriging method, while the second is created with a local polynomial, least-squares trend surface. Overall model(s) fit,

prediction error tables and normality tests of the interpolation surface residuals will be available in Appendix B for inspection, as are normality tests and differences of means results between arithmetically averaged and area weighted winter season precipitation. From here, the analysis mirrors the lognormal probability distribution analysis of seasonal streamflow.

Precipitation Surface Interpolation

Figure 5.49 presents the precipitation surface interpolated by an ordinary kriging method with an exponential variogram. Error associated with this surface is presented in table 5.24. The RMSE indicates an overall surface error of 43.32 mm, while the mean of the prediction errors is unbiased and normally distributed (see table B4 in Appendix B). Finally, the RMSE standardized value suggests that the model is slightly underestimating the variability in the predictions (Johnston, 2001). Figure 5.50 presents area-weighted precipitation for 1974, using a local polynomial, trend surface. The RMSE indicates an overall surface error of 68.50 mm, while the mean of the prediction errors is again unbiased and normally distributed (table B5 in Appendix B).

As previously detailed, a gridded precipitation surface was created for each winter season from 1957-1997. Each surface was examined for error minimization and unbiased residuals according to the criteria described in chapter 4. Both arithmetic-averaged and area-weighted winter season precipitation are normally distributed at less than 80% confidence (see figures B1 and B2 in Appendix B). Table B6 (Appendix B) indicates no significant differences between the two averaging methods.

Niño 3.4 Differentiated Precipitation

Basin-wide seasonal precipitation means between the two sub-divisions are significantly different (table B7, Appendix B), suggesting that a warm Pacific is associated with higher seasonal precipitation.

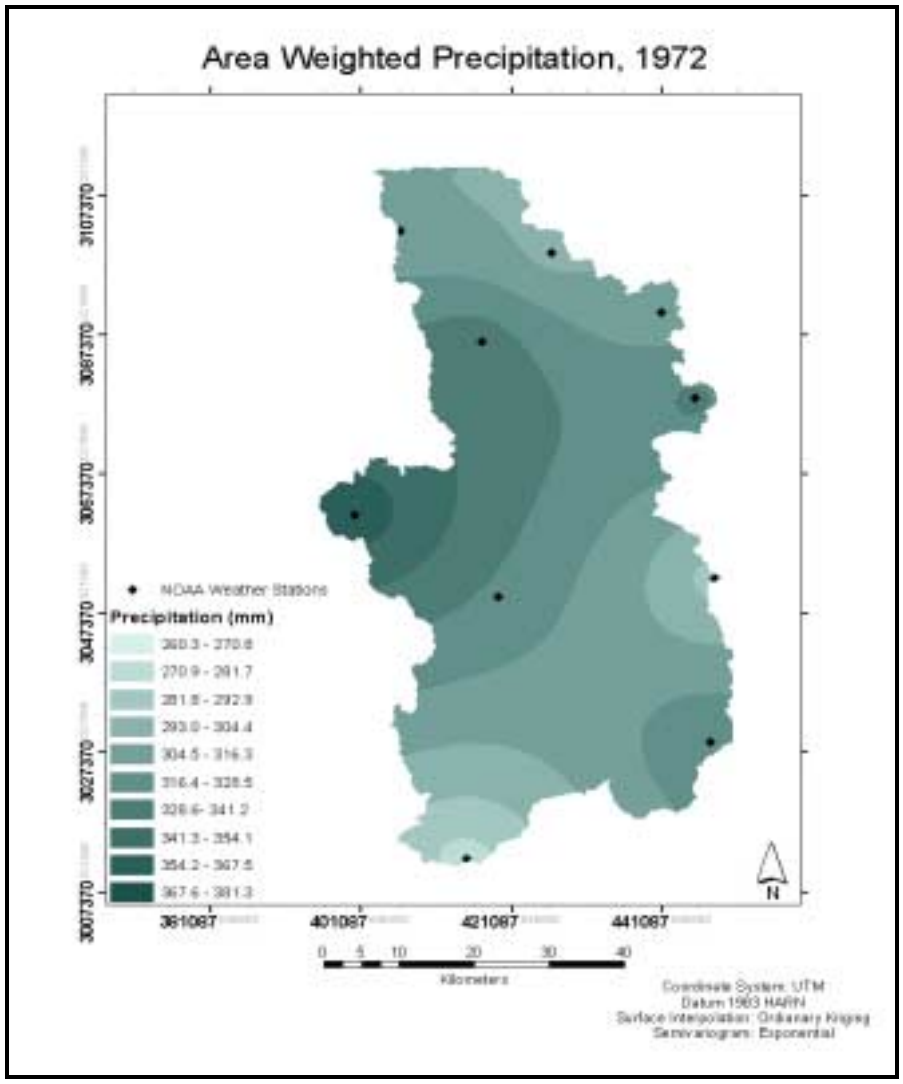

Figure 5.49. Area-weighted precipitation, 1972.

Table 5.24. Surface prediction errors, kriging model.

Surface Prediction Errors			
Ordinary Kriging Model			
Mean	RMSE	Mean Error Std	RMSE Std
0.747	43.32	0.009	1.09

Precipitation Return Period Estimates

Return period estimates (figures 5.51 and 5.52) suggest that higher precipitation can be expected at every level (both floods and droughts) under warm Pacific conditions.

Conditional Normal Probability Distributions

Figure 5.53 displays the winter season historic mean precipitation data, fitted normal distribution and terciles. "Average" precipitation lies between 679 and 1076mm.

Figure 5.54 depicts the probabilities of mean precipitation when the historic series is differentiated upon a warm and cold Pacific. Once again a clear separation in the distributions and probabilities of precipitation levels is observed. The likelihood "below" average precipitation under a Cp condition has increased to 0.382. As with the winter mean flow variable, a Wp condition decreases the likelihood of experiencing the same precipitation level to 0.146. The probabilities of experiencing "above" average mean precipitation increase to 0.449 under a Wp condition, while decreasing to 0.146 under a Cp condition.

Summary

This chapter presented results of winter and spring season streamflow and winter season precipitation forecast analyses. Hydro-meteorological variables studied include minimum, mean and maximum streamflow and mean precipitation. Differences in means, variances and return period estimates between differentiated subpopulations of streamflow and precipitation all suggest the Pacific is influential in above/below average streamflow/precipitation. The lognormal (normal) probability distribution establishes empirical probabilities of experiencing above average, average and below average streamflow (precipitation) conditioned upon SSTs in both the Niño 3.4 and TNA regions. The significance of those results was quantified using Monte Carlo simulations for each

flow variable and season, suggesting the TNA may only be influential upon spring season minimum flows.

An ordinary least-squares regression analysis on winter and spring season mean flow variables supports the findings of the empirical and synthetic forecast probabilities.

Three-way contingency tables quantified probabilities, relative risks, odds and odds-ratios of experiencing a season during which greater or fewer than the mean number of low-flow days occur, suggesting the TNA is independent of low flow day numbers.

Loglinear regression analyses provided a direct measure of the number of days per season below the low-flow threshold that could be expected per SST category, showing that very cold Pacific temperatures are associated with large numbers of low flow days.

Finally, a GIS was used to interpolate missing seasonal precipitation records for NOAA weather stations within the Peace River drainage basin, providing a complete record. The results suggest that a cold Pacific is associated with dry conditions over the Peace River drainage basin.

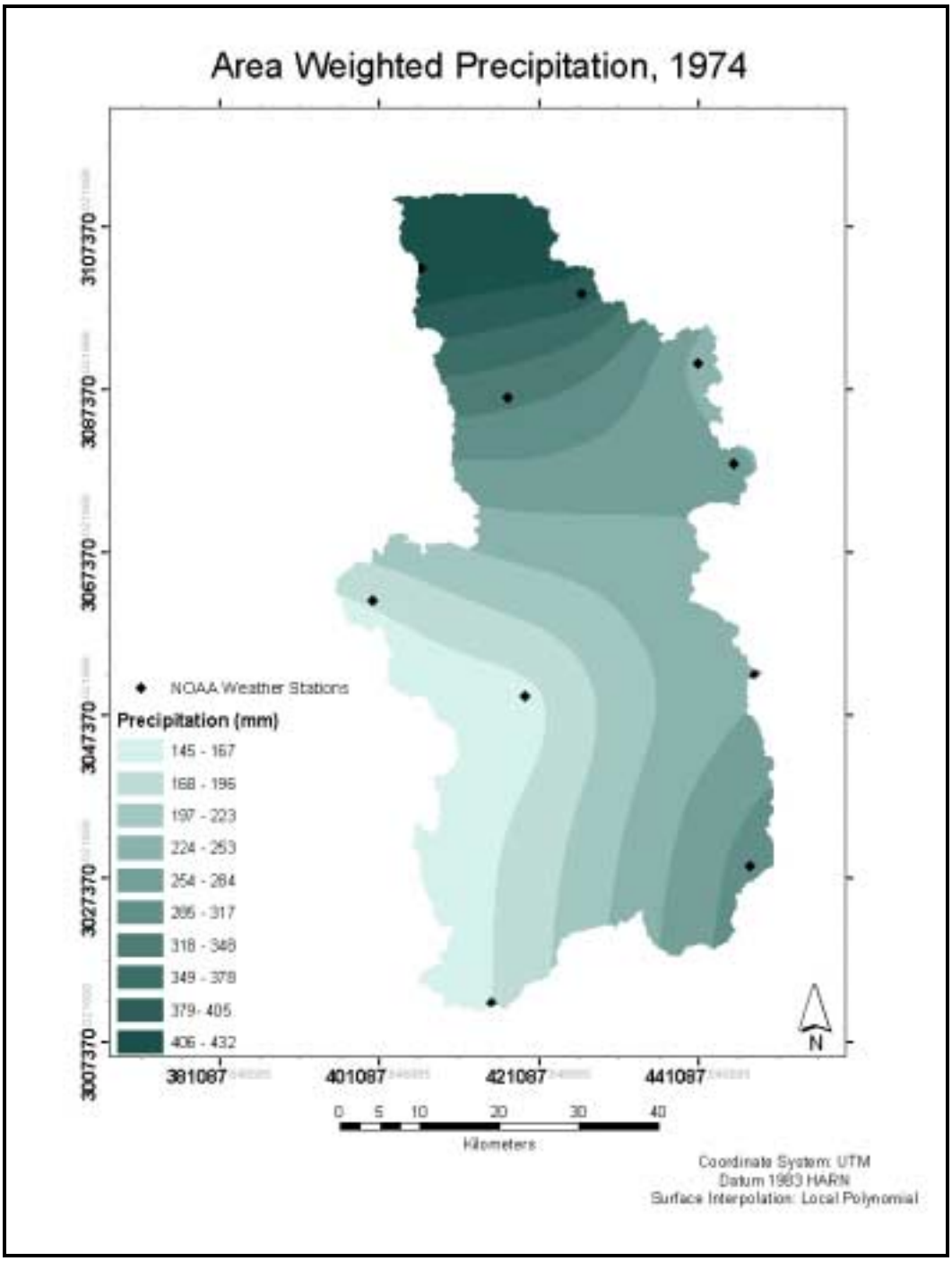

Figure 5.50. Area-weighted precipitation, 1974.

Table 5.26. Surface prediction errors, local polynomial model.

Surface Prediction Errors		
Local Polynomial Surface		
Mean	RMSE	Mean Error Std
3.38	68.5	-8.66E-16

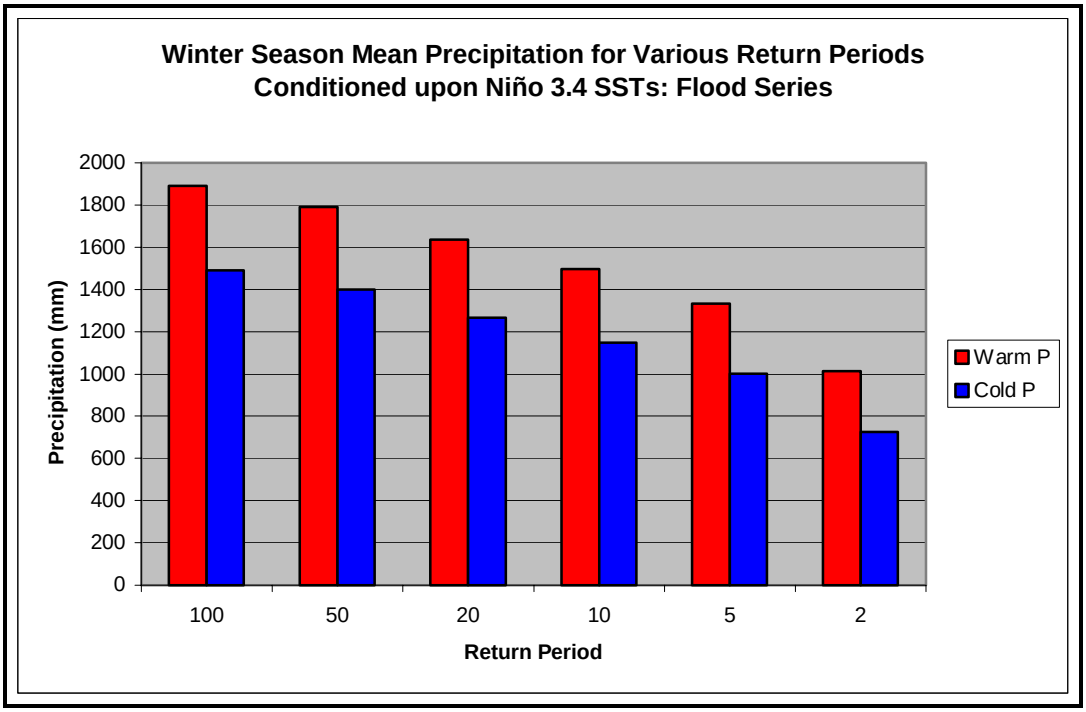

Figure 5.51. Mean winter season precipitation, flood series.

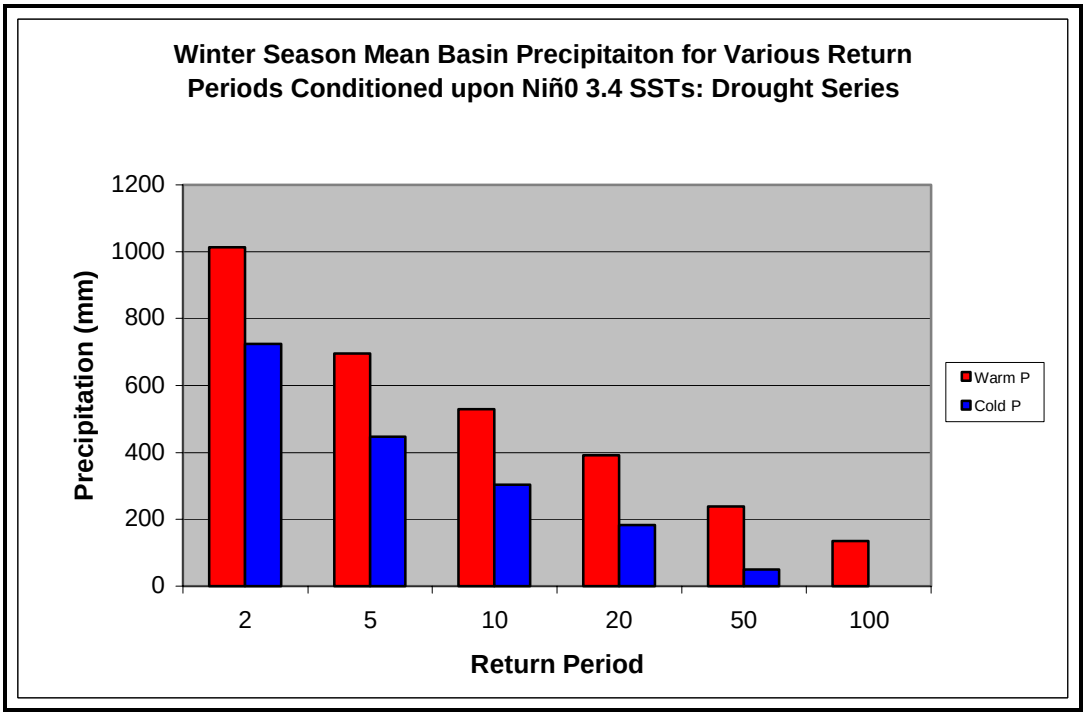

Figure 5.52. Mean winter season precipitation, drought series.

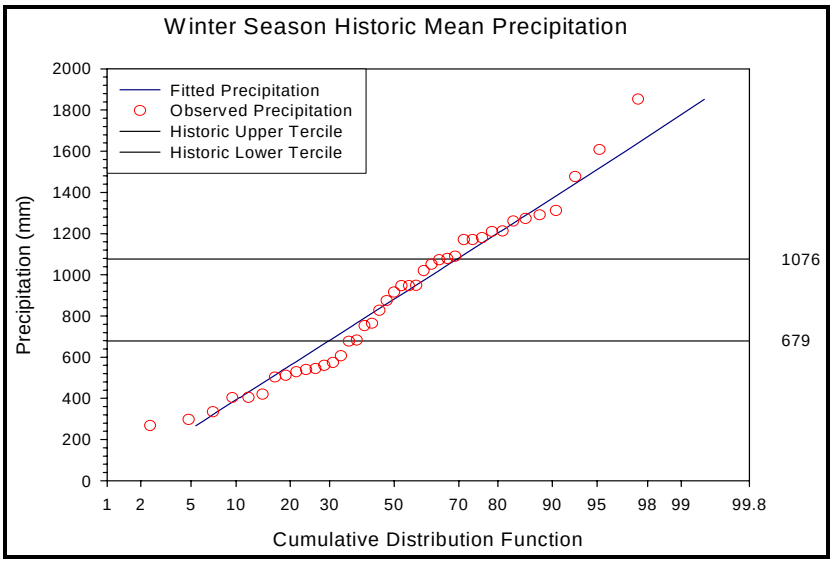

Figure 5.53. Winter season undifferentiated mean precipitation.

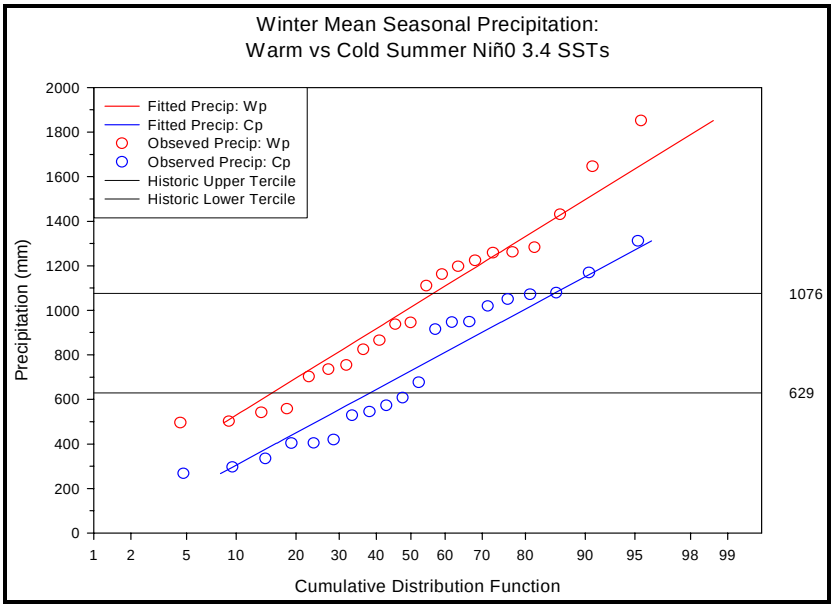

Figure 5.54. Winter season mean flow differentiated on summer Niño 3.4 SSTs.

CHAPTER 6
DISCUSSION

Chapter Overview

The following chapter addresses the implication of study findings and is organized as follows. The lognormal probability distribution is discussed in terms of normalizing the streamflow data to permit the comparison of means and variances and return period estimates. Secondly, the observed probability assessments for above, below and average streamflow under each differentiated SST condition are compared to the Monte Carlo simulated streamflow responses. The implications of the OLS regression results are presented, along with "what-if" scenarios. The three-way contingency table results are discussed, followed by the implications of the loglinear results. An assessment of the methods and results used for precipitation record interpolation and differentiation concludes the chapter.

Lognormal Probability Distributions: Winter and Spring Seasonal Streamflow Goodness of Fit and Differences in Moments

All winter and spring season streamflow variables (minimum, mean and maximum) are well characterized by the lognormal probability distribution as determined by the scaled Kolmogorov-Smirnoff (K-S) *D* goodness of fit test at: $\alpha = 0.20$ (Crutcher 1975). The large significance value ensures a good fit and reduces the possibility of a Type II error. Because confidence is high that the streamflow variables examined for each season follow the lognormal probability distribution, we can be similarly confident that the

parametric procedures used in estimating differences of means and variances between all flow variables are appropriate.

Winter and spring season flow variables exhibit significant differences in means and/or variances when differentiated on Pacific Ocean warm/cold conditions from the previous summer. As such, mean winter streamflow response is nearly 2.5 times greater under Wp conditions (23.10 m^3s^{-1}) than Cp conditions (9.48 m^3s^{-1}). Mean spring season streamflow response under Wp conditions (15.96 m^3s^{-1}) is 1.6 times that of streamflow response under Cp conditions (9.87 m^3s^{-1}). When winter/spring streamflow variables are further differentiated using Niño 3.4 *and* TNA SSTs, no further separation of subpopulations occurs, suggesting that the interactive effect of summer TNA SSTs has no significant effect upon mean streamflow magnitude or variability.

Winter/Spring Return Periods

Winter Cp return period flow magnitudes average just 41% of Wp return period magnitudes. The greatest difference is between the 100-year floods, where a 100-year Cp flood magnitude is 33% of the 100-year Wp flood. A less abstract way of assessing return period differences is in terms of water supply. As indicated in chapter 1, the Southwest Florida Water Management District (SWFMD) has declared a low-flow threshold of 3.68 m^3s^{-1}, and if flow drops below this no surface water withdraws are permitted. Under a Cp condition, the 10, 20, 50 and 100 year return period droughts are all under this limit, while under a Wp condition, only the 100-year drought (3.14 m^3s^{-1}) would lead to a cessation of withdrawals.

Spring Cp return period flow magnitudes average 28% higher than do winter season Cp return periods when compared to Wp return periods for all flow levels. The significantly greater variance of spring season mean flow from table 5.1 is reflected in the

upper end of the return flood probability distribution. The 100, 50 and 20 year return floods are 1.23 to 1.01 times greater than the same return floods under a Wp condition (figure 5.21). At the lower end of the distribution, a Wp condition produces greater flows, with only the 50 and 100-year droughts below the low-flow threshold (figure 5.22). In contrast, the Cp 10, 20, 50 and 100-year droughts are all under the low-flow threshold. Similar to winter mean flow, a Wp condition is favorable for water supply needs. In short, if a Wp condition is present during the summer months, there is a greater likelihood that uninterrupted surface withdraws will be permitted. On the other hand, a Cp condition will be of concern for water supply managers.

Although all the streamflow variables under study do not significantly respond under various Niño 3.4-TNA phases as compared to Niño 3.4 alone, each seasonal flow variable and subpopulation is described by the unique statistical moments (μ, σ^2) of their probability distributions. As such, it is prudent to investigate how the distributions change when conditioned on interactive ocean basin SSTs in terms of experiencing above, below and average flows.

Winter Season Lognormal Probability Distributions

Minimum Flow Variable

A warm Pacific (Wp) condition creates a diminished (greater) likelihood of experiencing below (above) average minimum flows. As such, there is a 37.3% greater chance of encountering below average minimum flow under a Cp condition, while under a Wp condition there is a 39.6% greater chance of above average flow. The observed response upon further division shows that cold Pacific-cold Atlantic (Cp-Ca) and cold Pacific-warm Atlantic (Cp-Wa) conditions both decrease and increase the chances of below average flow by 6.2%. For above average flow, a Wp-Wa or Wp-Ca condition

both increases and decreases the chances of experiencing this flow level by 9.2 % and 19.5% respectively.

However, there is still uncertainty about the significance of the responses because each observed probability distribution represents just one realization of the flow variable under study. Thus, it is not known if these differences are due to sample size or represent the influence of large-scale oceanic-atmospheric processes. This hypothesis is visited via Monte Carlo simulations.

For the minimum flow variable, the effect of warm or cold TNA SSTs on the Cp forecast is negligible, although it was shown that there is less than a 10% chance that the 19.5% observed decrease in probability of above average flow likelihood is due to sample size. It appears that a cold Atlantic in combination with a warm Pacific essentially reduces above average flow likelihood to that of an undifferentiated streamflow forecast probability (i.e. 33.3%), while at the same time enhancing average flow likelihood well above what is forecast under a Wp differentiated condition alone.

Mean Flow Variable

The change in likelihood of experiencing below average mean flow under a Cp to Wp condition is 31.3% greater, while the likelihood of experiencing above average flow is 35.4% greater under a Wp than a Cp. The addition of TNA SSTs does not significantly change any flow category forecast. In brief, conditions in the TNA appear to have no effect in the likelihood of experiencing above average, average and below average flow upon what is already forecast.

Maximum Flow Variable

The forecast results for winter season maximum flow variable essentially mirrors that of the mean flow variable. The occurrence of below average maximum flow is

36.2% greater under a Cp than a Wp condition, while under a Wp condition the chance of experiencing above average flow is 31.3% higher (vs. a Cp condition). Empirical probability differences are virtually unchanged and not statistically significant with the inclusion of warm/cold TNA SSTs.

Spring Season Lognormal Probability Distributions

Spring season flow exhibits greater variability than winter season flow at all levels, reflecting intra-season variability of precipitation generating mechanisms. These range from mid-latitude cyclonic activity in early season to local convection and infrequent tropical storm/easterly wave activity in late season.

Minimum Flow Variable

Warm Pacific conditions again create a diminished (greater) likelihood of experiencing below (above) average minimum flows than do Cp conditions, with just 4 observed flows out of 25 that were above average under a Cp condition. For water supply matters, a Cp forecast condition is cause for concern, especially since above average minimum flow is below the 3.68 m^3s^{-1} threshold. Similar to winter season minimum flow, there is a 34.7% greater chance of encountering a below average flow response under Cp conditions, while under Wp conditions there is a 24.7% greater chance of an above average flow response (vs. Cp conditions). Under Cp-Ca and Cp-Wa conditions, there appears to be some differences in the probability distributions of minimum flows as compared to a Cp conditional distribution (figure 5.35). Every minimum flow magnitude under Cp-Wa conditions has a greater likelihood of occurrence than under a Cp condition alone, while under Cp-Ca conditions flows are less likely to be as large.

Additionally, it appears that the observed forecast probability of 0.655 for below average flow under a Cp-Wa condition is significant above the 90% confidence level. Therefore, it appears that warm (cold) TNA SSTs increase (diminish) the Cp below average flow forecast, although above average and average minimum flow empirical forecasts are not significantly changed by conditions in the tropical Atlantic three seasons prior.

Mean Flow Variable

Spring season mean flow probability distributions are similar to those of minimum flow, where the chance of experiencing below average flow is 24.5% higher under Cp conditions than Wp conditions. The forecast for above average flow is just 14.3% higher under Wp than Cp conditions, reflecting the significantly greater variability in mean flow under a Cp condition. No evidence suggests that the inclusion of Atlantic SSTs significantly alters this forecast.

Maximum Flow Variable

The probability distributions are similar to those of the other spring season flow variables. Chance of below average flow occurrence is 21.6% higher under Cp than Wp conditions, but the chance of above average flow is only 8.5% higher under a Wp than a Cp condition. Thus, the smaller variability of Cp maximum flows is reflected, especially in the upper end of the probability distribution.

Only with average flows does there appear to be sufficient evidence to indicate that Atlantic SSTs may be a factor in the forecast.

Ordinary Least Squares Regression

Winter Season Model

The lognormal probability distributions and tripartite forecasts do not isolate the effects of extremely warm or cold SSTs on streamflow variables responses however the creation of main effect dummy and/or interactive dummy variables in OLS addresses this issue. As reported previously, Wp, Cp-Ca and Cp-Wa variables are represented binomially. The main and interactive effects all test significant against the null hypothesis: Ho: $\beta = 0$. The model may be used to answer such questions as:

- What is the expected winter season mean flow response if none of the conditions represented by the dummy variables are present, *and* fall season mean streamflow is at its *lowest* level?
- What is the expected winter season mean flow response if none of the conditions represented by the dummy variables are present, *and* fall season mean streamflow is at its *highest* level?
- What if fall season mean flow is at its highest level, *and* there is an extremely warm Niño 3.4?
- What if fall season mean flow is at its lowest level, *and* there is an extremely warm Niño 3.4?
- What if fall season is high or low and the Cp-Ca/Cp-Wa interactive conditions are present?

The regression model provides these answers by inserting the presented conditions and recovering their values. Table 6.1 illustrates the estimates and 95% prediction intervals, of flow variables in the absence of any information about SSTs in the preceding summer, and answers the first two questions above.

Here, the best estimate of winter season streamflow is the precedent season streamflow. As such, the best linear unbiased estimate (BLUE) for winter season

streamflow when fall season streamflow is at it minimum mean historic level is 4.70 m^3s^{-1}.

Table 6.1. Expected winter season mean flow if SST conditions are not present.

SST Model Conditions: Not Present			
Fall Season Flow Level	Estimate	LCL	UCL
Mean	13.826	5.435	35.176
Min	4.700	1.761	12.540
Max	46.043	16.881	125.587

BLUEs for mean and maximum fall flow produce 13.83 and 46.04 m^3s^{-1} of winter season flow, respectively. The relatively wide prediction intervals about the BLUE coefficients reflect the relatively large random component explaining streamflow responses.

Tables 6.2, 6.3 and 6.4 answer the remaining questions.

Table 6.2. Expected winter season mean flow when Niño 3.4 SSTs are >1σ above mean *and* various fall flow levels.

Condition: Wp			
Fall Season Flow Level	Estimate	LCL	UCL
Mean	41.702	16.391	106.097
Min	14.175	1.450	15.235
Max	138.873	13.834	153.241

Under these conditions (table 6.2), a summer Wp creates an even higher winter season mean flow response. For example, the BLUE for winter season flow under a Wp condition and maximum fall flow is 92.83 m^3s^{-1} greater than what maximum fall season mean flow will produce alone. The minimum and mean fall season flows *and* Wp conditions produce similar increases in mean flow response.

When conditions in the Niño 3.4 are 1σ below the long-term *and* the TNA is either +/- 1σ away from its long-term mean, a smaller winter season flow response is observed at every level of fall season flow, illustrated in tables 6.3 and 6.4.

Table 6.3. Winter season flow estimates for Cp-Wa conditions at each fall season flow level.

Condition: Cp-Wa			
Fall Season Flow Level	Estimate	LCL	UCL
Mean	4.152	1.632	10.565
Min	1.411	0.529	3.766
Max	13.828	5.070	37.717

Table 6.4. Winter season flow estimates for Cp-Ca conditions at each fall season flow level.

Condition: Cp-Ca			
Fall Season Flow Level	Estimate	LCL	UCL
Mean	3.832	1.506	9.749
Min	1.303	0.488	3.475
Max	12.761	4.679	34.807

The mean flow responses reflect the winter season mean flow lognormal probability distributions when the historic record is further subdivided to include Atlantic conditions. That is, the various TNA SST conditions made virtually no impact upon what is already forecast under Cp conditions. However, it is important to note from a water supply standpoint that under either scenario, winter season mean flow response is lower at all levels (at the given levels of fall season flow) than if the Cp conditions do not exist, as in table 6.1.

Spring Season Model

In the spring season OLS model, the extreme SST effects in the Niño 3.4 and TNA are represented by an interactive dummy variable, with main and interactive effects all testing significant.

The regression model provides insight into streamflow response to SST conditions and precedent seasonal flow levels. Only flow under extreme Cp conditions interacted with a Wa (Wa > 1σ above mean). Other model forms may capture the Wp influence, and the effect of a Wp on spring season flow may not be truly linear. Further, at the upper end of the spring mean flow probability distribution, a Cp produced larger flows than a Wp, as revealed in the return flood estimates. However, the OLS model for spring season did capture the negative influence of the Cp on spring season flow response, the primary concern of water supply management Table 6.5 illustrates various scenarios for the effect of winter season flow level on spring season mean flow response.

Table 6.5. Effects of various winter season flows on spring season mean flow levels.

Condition: Not Present			
Winter Season Flow Level	Estimate	LCL	UCL
Mean	13.302	3.872	45.696
Max	37.832	2.574	36.163
Min	5.804	1.585	21.257

The relatively wide prediction intervals about the BLUE coefficients reflect the large random component of the model. In water supply issues the lower confidence levels (LCL) of the prediction based on winter mean flow is of concern. Both minimum and maximum mean flows fall under the critical low-flow threshold. Table 6.6 illustrates the effects of a Cp-Wa condition on spring season mean flow levels and winter season Cp *conditional* flows. Since Cp conditions affect winter season flow response, only the mean, maximum and minimum winter flows observed under *Cp conditions* are used.

Table 6.6 illustrates considerable concern for water supply under these conditions. Each LCL is well below the low-flow threshold. Only when the is maximum mean flow

occurs during the winter season does the flow estimate exceed that required for permitted surface withdraw.

Table 6.6. Effects of various winter season flows and Cp-Ca conditions on spring season mean flow levels.

Condition: Cp-Wa			
Winter Season Flow Level	Estimate	LCL	UCL
Mean	2.763	0.804	9.490
Max	5.631	1.503	21.106
Min	1.480	0.404	5.421

The spring season model has an adjusted R^2 value of just 0.343. One way to reduce unexplained variance would be to include coincident seasonal precipitation, but then the model ceases to function as a forecast tool. Other possibilities might include using a Box-Cox transformation on the dependent and independent continuous streamflow variables or experimenting with different TNA SST lag intervals, but may be attempted in a future study.

Regardless of model shortcomings, the OLS results provide insight into the effects of extreme SST conditions not captured under the lognormal probability distribution forecasts, while confirming the results of the earlier approach. During winter and spring a Cp condition has a negative effect on streamflow response at every level, and while a Wp condition did not figure significantly in the spring season model, that may simply reflect the fact that the season is less dominated by a single precipitation generating mechanism.

Three-Way Contingency Tables

Winter Season

The goal of the conditional contingency tables is to determine the conditional likelihood of experiencing a season during which the number of low flow days is above

or below the long-term mean for each season. For the winter season, the average number of low flow days per season is 8 days, where low flow is again defined as less than 3.68 m^3s^{-1} (130 ft^3s^{-1}).

When summer Niño 3.4 SSTs are below median, the odds of experiencing a season with more than 8 days of low-flow are over 7 times greater than experiencing less than 8 days. This odds ratio is independent of TNA SSTs.

From a water supply standpoint, a Wp during the previous summer will increase the likelihood of a below-mean number of low flow days in accordance with previous findings.

Spring Season

The average number of low-flow days in spring has increases to 23. When summer Niño 3.4 SSTs are below median, the odds of experiencing more than 23 days of low-flow are over 7 times greater than the odds of a season with less than 23 days. A Cp during the previous summer is therefore, once again, an unfavorable condition from a water supply standpoint. Atlantic SSTs proved to be independent of the odds ratios.

Loglinear Regression Analysis

The final model employed to assess associations of TNA and Niño 3.4 SSTs is loglinear regression. This yields estimates and prediction intervals of low-flow days per season conditioned upon a mean SST per class partition (quintiles) of Niño 3.4 SSTs. TNA SSTs did not test significant as either a main effect or in an interactive sense for either season.

During winter, the mean Niño 3.4 SSTs of the two uppermost percentile classes produces the smallest estimate of low-flow days. According to the model, warm and

warmest Niño 3.4 SSTs in the preceding summer would only produce 2.9 and 3.2 days when withdrawals would cease.

The neutral class (half above-half below median SST) did produce the largest estimate of low-flow days (15.4). That the neutral SST category experienced the largest number of low-flow days is not unexpected. Under these "neutral" conditions, there is little anomalous advection of moist, or dry, air over Florida. The sub-tropical jet maintains a zonal track, and normal winter season precipitation is more variable. One anomalous season, 1985, during which 85 of out of 91 days were below threshold, violates the loglinear model's assumption of a linear systematic component. The model is still valid as the beta parameter for the independent variable is additive, and should prompt the analyst to investigate the anomaly for any incongruity.. However, the cause, climatic or anthropogenic, remains unknown. Removal of the anomaly leads to linearity in the model and a better overall fit as manifested through the narrower prediction intervals.

Similar problems with the linear assumption are encountered in the spring season. In addition to the possible explanation already noted above, spring season flows are produced through mixed precipitation processes, including winter frontal and convective activities. However, no single season within the neutral category displays abnormally high numbers of low flow days. As such, the final model is the best estimate under the stated hypotheses, SST lag intervals and physical assumptions.

Still, the mean Niño 3.4 SST in the coolest category produced nearly twice the number of expected low-flow days (from 23 to 43), while the mean Niño 3.4 SST in the

warmest category produced less than 1/3 of the expected number (from 23 to 7). In effect, low-flow days in both seasons are inversely related to Niño 3.4 SSTs.

Peace River Drainage Basin Winter Precipitation

Basin Delineation and Precipitation Surface Interpolations

The ArcHydro GIS tools delineated well the sub-watershed and permitted the determination of the NOAA weather stations within the contributing drainage area. This is of primary concern considering the physical attributes of the soils and the low relief of the area. No gross errors in basin delineation are present in comparison to the USGS delineated watershed. Successful drainage basin demarcation diminishes the need to rely on secondary data sources and reduces the chances of including unnecessary information, while allowing us to calculate the contributing drainage area(s) at any un-gauged location on the river.

The initial ordinary kriging model (model five) performance is satisfactory (table 4.6). In the absence of primary rainfall data, predictions for missing point data are physically reasonable. It neither grossly over-or-under predicted seasonal rainfall for missing data.

Various area-weighted interpolation methods produced differences from the arithmetic average of only 2mm over the 41 years of data. These may have been larger had there been significant changes in physical characteristics across the basin such as severe elevation changes that could produce an orographic effect.

Winter Season Precipitation

Computed winter season precipitation totals are adequately characterized by the normal probability distribution, as are the conditional Wp and Cp subpopulations. The

differences in means (but not variances) reveals greater winter season precipitation when SST conditions in the Niño 3.4 region are above normal.

The estimated flood and drought quantities reveal that winter season precipitation under Cp conditions is just 58.9% of the Wp. The largest separation is found at the 100-year drought return period, where precipitation under a Cp is 0.007% of Wp precipitation. The average difference between Wp and Cp 2 to 100-year floods is 355 mm. The average difference at the 2-100 year droughts is less at 216 mm.

Significant differences exist in the probability distributions of Cp and Wp precipitation, especially in the tails. Although the Weibull plotting position used to plot the observed precipitation values can bias the extremes of the distributions, it is likely that the observed separation is real. Using the 0.50 exceedence probability as an example, some 289 mm more seasonal precipitation can be expected under Wp conditions. This precipitation forecast would be of value to many, especially agricultural concerns within the Peace River drainage basin.

A Note on the Effect of Ocean-Atmosphere Interactions

The inclusion of TNA SSTs appears to have a significant effect only on the winter season minimum flow variable when the Niño 3.4 region is above median level, and only for the above average and average flow forecasts. From a water management standpoint, the most important forecast category is below average flow, for which the inclusion of TNA SSTs does not appear to have a significant influence.

If the TNA is independent of ENSO, the phases and magnitudes of the NAO are responsible for warm/cool summer TNA SSTs. When the Azores High is anomalously strong, a large pressure gradient exists between it and the Inter-tropical Convergence Zone. The literature suggests (e.g., Wanner et al., 2001; Marshall et al., 2001) that intra-

basin high pressure subdues easterly wave and tropical cyclone activity, and dryer summer conditions are experienced over the Florida peninsula (Wanner et al., 2001; Marshall et al., 2001). Dry summer conditions mean less groundwater storage, and lower flows during the winter and spring dry seasons.

However, if the TNA responds to ENSO forcings at a 4-6 month lag interval as suggested by Enfield and Mayer (1997) and Chen and Taylor (2002), then summer TNA conditions are further modified by conditions in the Eastern Pacific in the previous late winter-early spring. A cool Eastern Pacific over the previous winter (-1) diminishes the likelihood of substantial precipitation over the Florida peninsula during that season. As a cool Eastern Pacific then modifies TNA SSTs, the accompanying higher pressure over the subtropical North Atlantic diminishes convective and tropical precipitation sources. In short, a dry winter (-1) is followed by a dry summer (-1), and would be represented as a Cp-Ca summer (-1) condition. Less flow during winter/spring (0) would likely result, not so much as a result of lower rainfall in those seasons, but in the preceding seasons.

However, if a rapid overturn of SST conditions from a cool winter (-1) to a warm summer (-1) occurs in the Eastern Pacific, then the Eastern Pacific forcing of TNA SSTs is not noted in the data as a warm TNA, and would be represented in the data as Wp-Ca (-1) condition. A portion of the increased precipitation from a warm ENSO event during winter (0) will then serve as groundwater recharge over the Florida peninsula, including the Peace River drainage basin. Since the upper basin is a region of negative groundwater pressure head, the upper Peace is at times a losing stream. Therefore, while the likelihood of below average flow remains unchanged, the likelihood of above average flow would be low.

Spring season flow is normally more variable than winter flow. As such, understanding the large scale ocean-atmosphere interactions that would create a greater (lesser) likelihood of below average spring flow response is more complex. Under an assumption of TNA dependence to ENSO forcings, a warm summer (-1) TNA is associated with the teleconnection pattern produced under a warm wintertime (-1) Eastern Pacific. Warmer conditions in the TNA during summer are favorable for increased easterly wave and tropical cyclone development, effects of which would be felt during summer-fall (-1). Since a warm TNA is also associated with a weakened Azores High, that would create lower sea-level pressure in the subtropical North Atlantic, again favorable conditions for convective activity across the peninsula. These effects would be felt simultaneously.

However, this does not explain how a Cp-Wa increases the likelihood of below average spring flow. For this to happen, a rapid overturn of the Eastern Pacific SSTs from warm winter (-1) to cool summer (-1) is required, similar to that proposed for the winter season low-flow forecast.

The question of TNA dependence on ENSO forcings is not entirely resolved (Enfield and Mayer, 2001). The SST data reveal that Cp-Wa extreme (+/- 1σ) conditions occur just twice in the past 50 years, suggesting that extreme opposing conditions are a relatively rare event. As such, it is these conditions that the spring season OLS model reveals as highly leveraged values, and are antecedent of the lowest spring flows on record. Since it appears that the TNA is, at the least, associated with conditions in the Eastern Pacific, Wp-Wa and Cp-Ca would have to represent extended El Niño and La

Niña conditions. Opposite conditions would then have to reflect a rapid overturn of conditions in the Eastern Pacific and would suggest short-lived ENSO episodes.

For the Peace River summer (-1) Cp-Ca and Cp-Wa conditions both create a negative winter season flow response, suggesting that the ENSO signal dominates local climate response. In spring, the negative association between the Cp-Wa and streamflow suggest that it is the timing and strength of the ENSO event, rather than its duration that is responsible for the observed response. On the other hand, the categorical response of low-flow days analyzed in the three-way contingency tables and in the loglinear regression suggests that low flows are independent of TNA SSTs. This of course does not mean that the TNA, as a proxy for large scale ocean-atmosphere conditions in the subtropical North Atlantic, does not modify local climate and streamflow response. Rather, it suggests that the timing of the response is less than the 2-3 season lag time explored here.

Chapter Summary

This chapter has justified the use of parametric statistical techniques, probability distributions, ordinary least squares, contingency tables and loglinear regression as appropriate methods and models to forecast characteristics of seasonal streamflow/precipitation on the Peace River at Arcadia, FL. While the lognormal distribution adequately captures the probability distribution of all streamflow variables and produces likelihood measures of above, below and average flow, it does not differentiate between moderate and extreme SST events. However, the OLS models do allow the modeler to describe streamflow response to extreme SST events and antecedent seasonal flow scenarios. Finally, categorical data analysis techniques and models allow

the modeler to assess numbers and durations of low-flow days in response to SST categories.

Finally, it does appear that above/below median Niño 3.4 SSTs in summer can be effectively used as forecasters of precipitation and streamflow characteristics for the coming low flow seasons of winter and spring, although results gained through the use of *summertime* TNA SSTs remain ambiguous.

CHAPTER 7
SUMMARY AND CONCLUSIONS

Summary

This thesis addresses several statistical methods to provide winter and spring seasonal forecasts of streamflow/precipitation on the Peace River at Arcadia, FL and surrounding drainage basin. Because the state of Florida is confronted with significant population increases through the first half of the 21st century, accurate long-term forecasts of water supplies will continue to be one of the most critical aspects of growth management faced by planners and policy makers. As a result, proper management strategies and long-term forecasting of water supply is paramount.

Advantages and disadvantages of probabilistic methods to forecast streamflow/precipitation have been addressed. While operational hydrologists mostly favor deterministic methods, probabilistic methods are arguably a more appropriate method for long-term forecasting of hydro-meteorological variables, as the uncertainties are also quantified and presented for the use of planners and decision makers.

It has also shown that, through proper research design and the careful application of statistical methods and models, it is desirous for long-term forecasting to use a probabilistic, rather than deterministic approach. The quasi-stationary behavior of ENSO related SSTs (and SST anomalies) have shown to possess power in as predictors/fluctuations in long-term, basin wide hydro-meteorological variables.

The methods employed for this study are sound, and can be easily exported to other regions. Changes needed to successfully employ these methods would be in research

design, which includes the selection of independent environmental variables, and identifying appropriate lag times in the relationship between the predictor and response variables. Many researchers have used some form of ENSO-related variables, in many Western Hemisphere locations to successfully predict responses in precipitation and streamflow. Others have employed atmospheric-oceanic variables, such as the North Atlantic Oscillation or the Pacific Decadal Oscillation, with varying degrees of success to improve the of forecast hydro-meteorological variables.

Conclusions

This study reveals the following important conclusions. The lognormal distribution adequately captures the probability distribution of all streamflow variables (minimum, mean and maximum). Differences in means and variances between winter/spring subpopulations indicate that the response of winter/spring seasonal streamflow is greater in magnitude and variability under Warm Pacific conditions. Less streamflow/precipitation can be expected under Cold Pacific conditions.

It appears that above/below median Niño 3.4 SSTs in summer can be effectively used as forecasters of precipitation and streamflow characteristics for the coming low flow seasons of winter and spring.

Tropical North Atlantic SSTs for the most part appear to have no significant effect (upon what is already been forecast) except during the spring when the Niño 3.4 SSTs have been below median. This last point is of particular importance from a management standpoint, as below normal summer SSTs in the Niño 3.4 region are associated with drought, seemingly deepened by above normal tropical North Atlantic SSTs.

The OLS results indicated that although there is no concern with functional form for either season, more information is needed to decrease model prediction variability.

Still, OLS provides a method to examine the influence of extreme SST events not accounted for by the probability distribution forecasts.

The three-way contingency tables and loglinear regression analyses indicate that categorical data methods can be used with some success in forecasting numbers of low-flow days during the dry seasons, or a season during which there will be more or less than the average of low-flow days. The results from these models do not necessarily contradict the lognormal probability distribution results, but do indicate that the association of low-flows with summer season TNA SSTs is ambiguous, either as a main effect or as an interactive one with Niño 3.4 SSTs.

The normal distribution satisfactorily captures the probability distribution of winter season precipitation, with similar differences in response conditioned upon the Niño 3.4, as for winter season mean streamflow. Results of the winter season precipitation probability distribution model essentially mirrors that for winter season mean streamflow.

Finally, forecasting basin-wide precipitation during the dry season is of vital importance for agricultural interests within the region, while providing a reliable forecast for seasonal streamflow is vital to ensure an adequate domestic water supply for the communities of southwest Florida into the 21st century.

APPENDIX A
LINKS TO DATA SOURCES

USGS daily streamflow data for the Peace River at Arcadia, FL (Station 02296750) can be assessed at the USGS website: http://waterdata.usgs.gov/nwis/discharge

Monthly precipitation data for the Peace River drainage basin can be accessed at the NOAA website:http://www.ncdc.noaa.gov/oa/climate/online/coop-precip.html

Monthly sea surface temperature data can also be found at NOAA's NCPC website: http://www.cpc.ncep.noaa.gov/data/indices/index.html

Digital elevation data can be accessed from the GIS datadepot at:

http://data.geocomm.com/dem/demdownload.html

APPENDIX B
STATISTICAL TABLES AND FIGURES

Table B1. Normality test for standard prediction errors from kriging model 5. See text for explanation.

Normality Test Section of Standard Error			
Test	Obs Value	10% Critical Decision	5% Critical Decision
Kolmogorov-Smirnov	0.197	0.240	Accept Normality

Table B2. Normality test for standardized prediction errors from kriging model 5. See text for explanation.

Normality Test Section of Standardized error			
Test	Obs Value	10% Critical Decision	5% Critical Decision
Kolmogorov-Smirnoff	0.171	0.241	0.262 Accept Ho

Table B3 Significance test for standardize residuals.

Student's t-test for difference between mean and zero				
Alternative Hypothesis	T-Value	Prob Level	Decision (5%)	p-value
Error<>0	-0.372	0.718	Accept Ho	0.063
Stdd_Error<>0	-0.243	0.814	Accept Ho	0.056

Table B4. Bias and normality tests for kriged surface individual prediction errors.

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (5%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
Error<>0	0.0517	0.959889	Accept Ho	0.050247	0.010070
Assumption			**Value**	**Probability**	**Decision(5%)**
Skewness			0.0019	0.998477	Cannot reject normality
Kurtosis			-1.1555	0.247897	Cannot reject normality

Table B5. Bias and normality tests; least squares trend surface individual prediction errors.

Alternative Hypothesis	T-Value	Prob Level	Decision (5%)
Error<>0	0.4335	0.674871	Accept Ho

Assumption	Value	Probability	Decision(5%)
Skewness	0.1419	0.887139	Cannot reject normality
Kurtosis	-0.1017	0.918983	Cannot reject normality

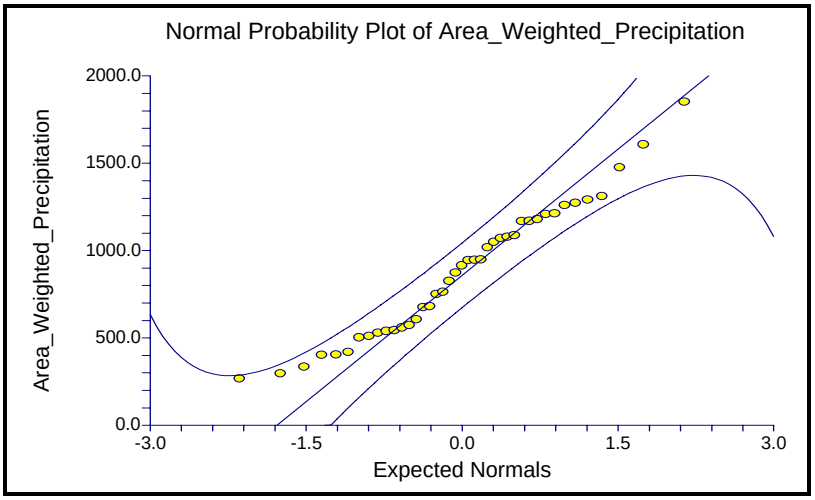

Figure B1. Normal distribution plot of area-weighted precipitation.

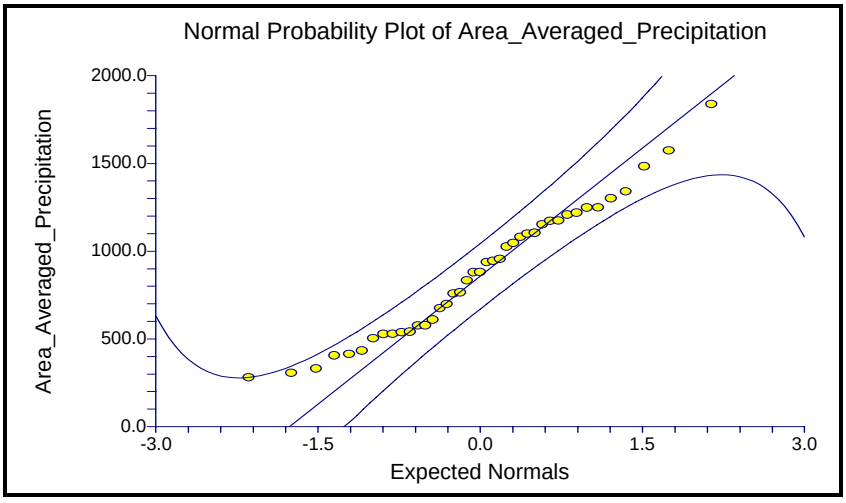

Figure B2. Normal distribution plot of area-averaged precipitation.

Table B6. Area-averaged vs. area-weighted basin precipitation.

	Area Averaged	*Area Weighted*
Mean	883.290	881.424
Variance	142605.988	145844.265
Observations	41.000	41.000
Pooled Variance	144225.126	
df	80.000	
t Stat	0.022	
P(T<=t) one-tail	0.491	
t Critical one-tail	1.664	

Table B7. Niño 3.4 differentiated subpopulation t-test results.

	Warm	*Cold*
Mean	1013.607	724.701
Std. Dev.	378	329
Df	39.000	
t Stat	2.607	
P(T<=t) one-tail	0.006	
t Critical one-tail	1.685	

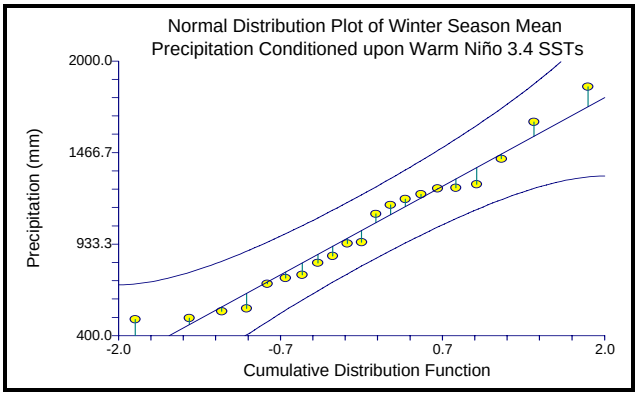

Figure B3. Normality test of winter season precipitation, warm subpopulation.

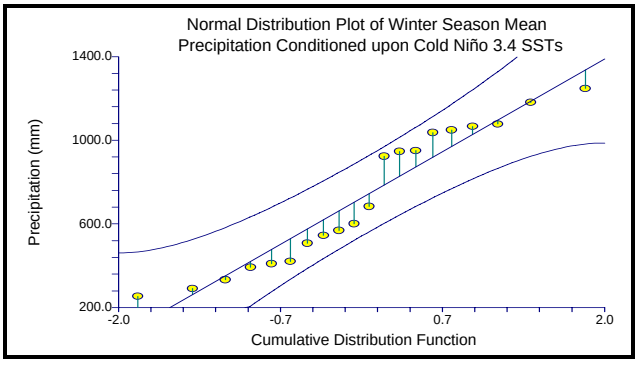

Figure B4. Normality test of winter season precipitation, cold subpopulation.

LITERATURE CITED

Agresti, Alan (1996). An Introduction to Categorical Data Analysis. New York: John Wiley and Sons, Inc.

Berry, William D. and Stanley Feldman (1985). Multiple Regression in Practice. Newbury Park: Sage Publications.

Bucha, Vaclav and Vaclav Bucha Jr. (2002). Geomagnetic forcing and climatic variations in Europe, North America and in the Pacific Ocean. Quaternary International. 91, 5-15.

Burt, James E. and Barber, Gerald M. (1996). Elementary Statistics for Geographers (2nd ed.). New York: The Guilford Press.

Caviedes, Cesar N. (2001). El Niño in History: Storming Through the Ages. Gainesville, FL: University Press of Florida.

Chen, Anthony A. and Michael A. Taylor (2002). Investigating the link between early season Caribbean rainfall and the El Nino+1 year. International Journal of Climatology. 22, 87-106.

Chiew, F.H.S., T.C. Piechota, J.A. Dracup, and T.A. McMahon (1998). El Nino/Southern Oscillation and Australian rainfall, streamflow, and drought: Links and potential for forecasting. Journal of Hydrology. 204, 138-149.

Chiew, S.L. Zhou, and T.A. McMahon (2003). Use of seasonal streamflow forecasts in water resources management. Journal of Hydrology. 270, 135-144.

Coleman, Jill S.M. and Jeffery C. Rogers (2003). Ohio River Valley Winter Moisture Conditions Associated with the Pacific-North American Teleconnection Pattern. Journal of Climate. 16, 969-981.

Covey, Daryl L. and Stefan Hastenrath (1978). The Pacific El Nino Phenomenon and the Atlantic Circulation. Monthly Weather Review. 106, 1280-1287.

Crutcher, Harold L. (1975). A Note on the Possible Misuse of the Kolmogorov-Smirnov Test. Journal of Applied Meteorology. 14, 1600-1603.

Czaja, Arnauld and Claude Frankignoul (2002). Observed Impact of Atlantic SST Anomalies on the North Atlantic Oscillation. Journal of Climate. 15, 606-623.

Czaja, A., P. Van Der Vaart and J. Marshall (2002). A Diagnostic Study of the Role of Remote Forcing in Tropical Atlantic Variability. Journal of Climate. 15, 3280-3290.

Dzurik, Andrew A. (2003). Water Resources Planning (3rd ed.). (chps. 5 & 8). Lanham: Rowman and Littlefield Publishers, Inc.

Enfield, David B., (1996). Relationships of inter-American rainfall to tropical Atlantic and Pacific SST variability. Geophysical Research Letters. 23(23), 3305-3308.

Enfield, David B. and Dennis A. Mayer (1997). Tropical Atlantic sea surface temperature variability and its relation to the El Nino-Southern Oscillation. Journal of Geophysical Research. 102(C1), 929-945.

Enfield, David B., Alberto M. Mestas-Nuñez and Paul J. Trimble, (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. Geophysical Research Letters. 28(10), 2077-2080.

Fernald, Edward A. and Purdum, Elizabeth D. (1995). (Eds.). Atlas of Florida. Gainesville, FL: University Press of Florida.

_____. (1998). (Eds.). Water Resources Atlas of Florida. Tallahassee, FL: Institute of Science and Public Affairs, Florida State University.

Gershunov, Alexander and Tim P. Barnett (1998). ENSO Influence on Interseasonal Extreme Rainfall and Temperature Frequencies in the Contiguous Unites States: Observations and Model Results. Journal of Climate. 11, 1575-1586.

Giannini, Alessandra, Mark A. Cane and Yochanan Kushner, (2001). Interdecadal Changes in the ENSO Teleconnection to the Caribbean Region and the North Atlantic Oscillation. Journal of Climate. 14, 2867-2879.

Giannini, Alessandra, John C. Chiang, Mark A. Cane, Yochanan Kushner and Richard Seager (2001). The ENSO Teleconnection to the Tropical Atlantic Ocean: Contributions of the Remote and Local SSTs to Rainfall Variability in the Tropical Americas. Journal of Climate. 14, 4530-4543.

Graybill, Franklin A. and Hariharan K. Iyer (1994). Regression Analysis: Concepts and Applications. Belmont, CA: Duxbury Press.

Gutierrez, F. and J.A. Dracup (2001). An analysis of the feasibility of long-range streamflow forecasting for Columbia using El Nino-Southern Oscillation indicators. Journal of Hydrology. 246, 181-196.

Harshburger, Brian, Ye Hengchun and John Dzailoski (2002). Observational evidence of the influence of Pacific SSTs on winter precipitation and spring stream discharge in Idaho. Journal of Hydrology. 264, 157-169.

Hurrell, James W., (1995). Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. Science. 269, 676-679.

Johnston, Kevin, Jay M. Ver Hoef, Konstantin Krivoruchko and Neil Lucas. (2001). Using ArcGIS Geostatistical Analyst. Redlands, CA: ESRI Press.

Kahya, Ercan and John A. Dracup (1993). U.S. Streamflow Patterns in Relation to the El Nino/Southern Oscillation. Water Resources Research. 29(8), 2491-2503.

Krzysztofowicz, Roman (2001). The case for probabilistic forecasting in hydrology. Journal of Hydrology. 249, 2-9.

Lewelling, B.R, A.B. Tihansky, and J.L. Kindinger. (1998). Assessment of the hydraulic connection between ground water and the Peace River, west-central Florida. Water-resources investigations report; 97-4211. pp vi-96.

Maidment, David R. (Ed.). (1993). Handbook of Hydrology. (pp. 17.1-19.72). New York: McGraw-Hill, Inc.

Maidment, David R and Dean Djokic. (Eds.). (2000). Hydrologic and Hydraulic Modeling Support. Redlands, CA: ESRI Press.

Maidment, David R (2002). (Ed.). ArcHydro: GIS for Water Resources. Redlands, CA: ESRI Press.

McBratney, A.B., and Webster, R. (1986). Choosing Functions for Semi-variograms of Soil Properties and Fitting Them to Sampling Estimates. Journal of Soil Science. 37, 617-639.

Marshall, John, Yochanan Kushnir, David Battisti, Ping Chang, Arnaud Czaja, Robert Dickson, James Hurrell, Michael McCartney, R. Saravanan, and Martin Visbeck (2001). North Atlantic Climate Variability: Phenomena, Impacts and Mechanisms. International Journal of Climatology. 21, 1863-1898.

Mo, Kingste C. and Sirpa Hakkinen (2001). Interannual Variability in the Tropical Atlantic and Linkages to the Pacific. Journal of Climate. 14, 2740-2762.

Oliver, M. A. and R. Webster. (1990). Kriging: a method of interpretation for geographical information systems. International Journal of Geographical Information Systems. 4(3), 313-332.

Randazzo, Anthony F. and Douglas S. Jones. (1997). (Eds.). Geology of Florida. Gainesville, FL: University Press of Florida.

Rasmusson, Eugene M. and Thomas H. Carpenter (1982). Variations in Tropical Sea Surface Temperature and Surface Wind Fields Associated with the Southern Oscillation/El Nino. Monthly Weather Review. 110, 354-384.

Rogers, Jeffery C (1984). The Association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere. Monthly Weather Review. 112, 1999-2015.

Ropelewski, Chester F. and Michael S. Halpert (1987). Global and Regional Scale Precipitation Patterns Associated with the El Nino/Southern Oscillation. Monthly Weather Review. 115, 1606-1626.

_____. (1996). Quantifying Southern Oscillation-Precipitation Relationships. Journal of Climate. 9, 1043-1059.

Schmidt Nancy, E. Lipp, J. Rose and M. Luther (2001). ENSO Influences on Seasonal Rainfall and River Discharge in Florida. Journal of Climate. 14, 615-628.

Sharma, A. (2000). Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1: A strategy for system predictor identification. Journal of Hydrology. 239, 232-239.

Simpson, James H. and Debra C. Colodner (1999). Arizona precipitation response to the Southern Oscillation: A potential water management tool. Water Resources Research. 35(12), 3761-3769.

Sun Hongbing, and David J. Furbish (1997). Annual precipitation and river discharges in Florida in response to El Nino- and La Nina-sea surface temperature anomalies. Journal of Hydrology. 199, 74-87.

Treat, Sally F. (Ed.) (1998). Proceedings of the Charlotte Harbor Public Conference and Technical Symposium. Southwest Florida Water Management District. Technical Report No 98-02. Brooksville, FL.

Unwin, David. (1975). An Introduction to Trend Surface Analysis. Concepts and Techniques in Modern Geography. Norwich: Geo Abstracts Ltd.

Viessman, Warren Jr., Gary L. Lewis and John W. Knapp (1989). Introduction to Hydrology. (3rd ed.). (pp. 517-528, 675-754). New York: Harper Collins Publishers.

Wanner, Heinz, S. Bronnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmutz, D. Stephenson, and E. Zoplaki (2001). North Atlantic Oscillation-Concepts and Studies. Surveys in Geophysics. 22, 321-382.

Zorn, Matthew R. and Peter Waylen (1997). Seasonal Response of Mean Monthly Streamflow to El Nino/Southern Oscillation in North Central Florida. Professional Geographer. 49(1), 51-62.

BIOGRAPHICAL SKETCH

David Coley (born Indianapolis, Indiana; 1962), lived in Sarasota, Florida, from 1964 to 1998 before attending the University of Florida in Gainesville. Upon receiving his General Equivalency Diploma in 1981, Mr. Coley worked in the restaurant service industry for 15 years while sporadically attending community college. He received a Bachelor of Arts in May of 2001 and a Master of Science in August of 2003, both in the field of geography from the University of Florida. He has always had a strong interest in the earth sciences and hydrology, with additional academic and professional interests in statistics, Geographic Information Systems, cartography and remote sensing.