Title
Hydrologic Response to Climate Change in California: Observational and Modeling Studies

Author
Daniels, Bruce Kenneth

Publication Date
2014

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

HYDROLOGIC RESPONSE TO CLIMATE CHANGE IN CALIFORNIA: Observational and Modeling Studies

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

EARTH SCIENCES

by

Bruce K. Daniels

December 2014

The Dissertation of Bruce K. Daniels
is approved:

Professor Andrew T. Fisher, Chair

Assistant Professor Noah Finnegan

Dr. Jeff Kiehl, Senior Scientist, NCAR

Tyrus Miller
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

FIGURES and TABLES

ABSTRACT

Hydrologic response to climate change in California:

observational and modeling studies

Bruce K. Daniels

The lack of adequate quantity and quality of water is a world-wide problem, which fosters concerns about the impacts of climate change. Obtaining forecasts of future water changes are important to allow early impact mitigation and adaption efforts.

This study forecasts precipitation changes, not through climate models, but by analysis of observations to derive trends of three metrics: event Intensity, event Duration, and lull Pause. From 50 long-term stations in California we obtained median trends of Intensity -0.45% per decade, Duration 0.50% per decade, and Pause 0.13% per decade.

One problem in the analysis of observations was proper techniques to handle gaps from missing data. Multiple Imputation (MI) was applied through fitting of Weibull probability distributions to the three metrics. This was tested by artificially injecting gaps into the mostly complete Sacramento record. MI partially restored deviations caused by that injection.

Permutation resampling techniques were applied with MI to derive significance p-values for each trend. Significance at 95% for Intensity was from 11 of the 50 stations, Duration from 16, and Pause from 19, of which 12 were 99% significant. Trends were combined by weighting them with the reciprocal of their p-values. Significance weighted California trends are Intensity -4.61% per decade, Duration 3.49% per decade, and Pause 3.58% per decade.

Two California basins with hydrologic models were studied: Feather River in the northern Sierra Nevada mountains and central coast Soquel-Aptos. Most hydrologic components between the two basins were shown to behave differently primarily because of climate differences.

Three metric trends were computed for each basin by combining trends from nearby observations. Each metric was changed without change to other metrics or the total precipitation and input into the models. Most hydrologic impacts were modest with magnitudes less than half the corresponding precipitation changes.

Feather River Basin's critical supply to Lake Oroville and the State Water Project were benefited from a streamflow increase by 0.5%. Soquel-Aptos Basin's value for water supply was harmed by groundwater recharge decrease by -2.5% and streamflow decrease by -1.1%. Neither of these impacts seem amenable to mitigation, thus adaptation is indicated.

ACKNOWLEDGEMENTS and DEDICATION

I would like to extend profound thanks to Professor Andy Fisher, the current chair of my dissertation committee and my co-adviser. His constant support, advice, and willingness to help have made this work what it is today. I also extend my profound thanks to Professor Lisa Sloan, the previous chair of my dissertation committee and my other co-adviser. She has always been my best advocate and booster.

Thanks are due to my other committee members, Prof. Noah Finnegan for his perseverance through this whole long process, and Dr. Jeff Kiehl for his amazing willingness to volunteer to jump in and help at the last minute.

I am grateful for funding from the David and Lucile Packard Foundation, UCSC, the National Science Foundation, the California Energy Commission, Benjamin and Ruth Hammett, the UCSC Services for Transfer And Re-entry Students (STARS), NOAA, and the National Water Research Institute.

I dedicate this work to my mother, almost-Doctor Lavenia Daniels. She inspired and fostered my interest in geology and my passion for science and learning.

INTRODUCTION

1. Motivation

The world climate is a complex system that encompasses numerous Earth surface systems, i.e. the atmosphere, hydrosphere, cryosphere, lithosphere, biosphere, etc. The preponderance of scientific evidence indicates that anthropogenic actions such as excess generation of greenhouse gases are causing serious climatic change. The typical manner to study climate change is through the construction and operation of Global Climate Models (GCMs) (IPCC, 2013).

Water is a critical resource to humans, commerce, wildlife, and the environment. This resource is already over-used and stressed in much of the world (Kondili et al., 2010). Changes of many of the atmospheric climate elements, such as temperature, radiation, wind speed and direction, humidity, and precipitation can impact water resources. A thorough understanding of the relationship between climate change and water resources could help improve management and improvement of water acquisition and delivery systems (Huntington, 2006).

Precipitation is the primary driver of the hydrologic cycle. Precipitation is typically characterized and measured in terms of the total annual precipitation. However, precipitation can be classified and quantified by other features such as the intensity of precipitation events, duration of events, and the lull time between events (Trenberth et al., 2003).

The terrestrial hydrologic cycle is that portion of the global hydrologic cycle that takes place on the land surface and the immediate sub-surface. It is an intricate system with many transport and storage components, such as evaporation, overland flow, infiltration, soil moisture, plant uptake, transpiration, percolation, stream base flow, and groundwater recharge, each of which has many uses and dependencies (Loaiciga et al., 1996).

It is important to understand how climate change will impact water resources. The typical manner to achieve such understanding is to hook-up the output of one or more GCMs (especially its precipitation output) to the input of hydrology models (Markstrom et al., 2011), (Synder et al., 2004), (Flint and Flint, 2012). The deficiency with this approach is that GCMs currently lack skill in predicting precipitation. GCMs often do not have substantial agreement on even the sign of change, i.e. wetter vs. drier (Figure 0-1).

One approach to climate understanding that has shown some significance is the analysis of changes of precipitation features. For example, the observed intensity of precipitation events is quite dramatic (Figure 0-2).

2. Thesis

The thesis of this study is that it is important to understand precipitation features in order to understand hydrology. This is motivated in part by the fact that such precipitation features are already observed to be changing. Even if annual precipitation is unchanging, precipitation features can still change. These precipitation feature changes by themselves can have significant impact on how water is partitioned and available for the hydrologic cycle components. The goal of this study is to compute the quantitive impacts on these hydrologic components from changes in precipitation features.

A subsidiary thesis is that an observation based approach to precipitation feature trend estimation is needed. Even if GCMs had precipitation skill, the coarse granularity of GCMs causes a fundamental "drizzle" problem where individual precipitation events get averaged over the entire grid cell and thereby obscured and lost (Perkins et al., 2007) (Mejia et al., 2014). A secondary goal of this study is to derive significant trends for the future changes of precipitation feature metrics from precipitation observations.

3. References

Flint, L.E., and Flint, A.L., 2012, USGS Scientific Investigations Report 2012–5132: Simulation of Climate Change in San Francisco Bay Basins, California: Case Studies in the Russian River Valley and Santa Cruz Mountains:.

Huntington, T.G., 2006, Evidence for intensification of the global water cycle: Review and synthesis: Journal of Hydrology, v. 319, p. 83–95, doi: 10.1016/j.jhydrol.2005.07.003.

IPCC, 2013, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental

Panel on Climate Change: Cambridge University Press, 1535 p.

Kondili, E., Kaldellis, J.K., and Papapostolou, C., 2010, A novel systemic approach to water resources optimisation in areas with limited water resources: Desalination, v. 250, p. 297–301, doi: 10.1016/j.desal.2009.09.046.

Loaiciga, H.A., Valdes, J.B., Vogel, R., Garvey, J., and Schwarz, H., 1996, Global warming and the hydrologic cycle: Journal of Hydrology, v. 174, p. 83–127, doi: 10.1016/0022-1694(95)02753-X.

Markstrom, S.L., Hay, L.E., Ward-Garrison, C.D., Risley, J.C., Battaglin, W.A., Bjerklie, D.M., Chase, K.J., Christiansen, D.E., Dudley, R.W., Hunt, R.J., Koczot, K.M., Mastin, M.C., Regan, R.S., Viger, R.J., et al., 2011, Integrated Watershed-Scale Response to Climate Change for Selected Basins Across the United States: U.S. Geological Survey Scientific Investigations Report 2011–5077.

Mejia, J.F., Niswonger, R.G., and Huntington, J., 2014, Uncertainty Transfer in Modeling Layers: From GCM to downscaling to hydrologic surface-groundwater modeling, *in* Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, CA, USA, International Environmental Modelling and Software Society (iEMSs).

Perkins, S.E., Pitman, A.J., Holbrook, N.J., and McAneney, J., 2007, Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions: Journal of Climate, v. 20, p. 4356–4376, doi: 10.1175/JCLI4253.1.

Synder, M.A., Sloan, L.C., and Bell, J.L., 2004, MODELED REGIONAL CLIMATE CHANGE IN THE HYDROLOGIC REGIONS OF CALIFORNIA: A CO2 SENSITIVITY STUDY: Journal of the American Water Resources Association(JAWRA),.

Trenberth, K.E., Dai, A., Rasmussen, R.M., and Parsons, D.B., 2003, The Changing Character of Precipitation: Bulletin of the American Meteorological Society, v. 84, p. 1205, doi: 10.1175/BAMS-84-9-1205.

U.S. GCRP, 2009, Global Climate Change Impacts in the United States: U.S. Global Change Research Program.

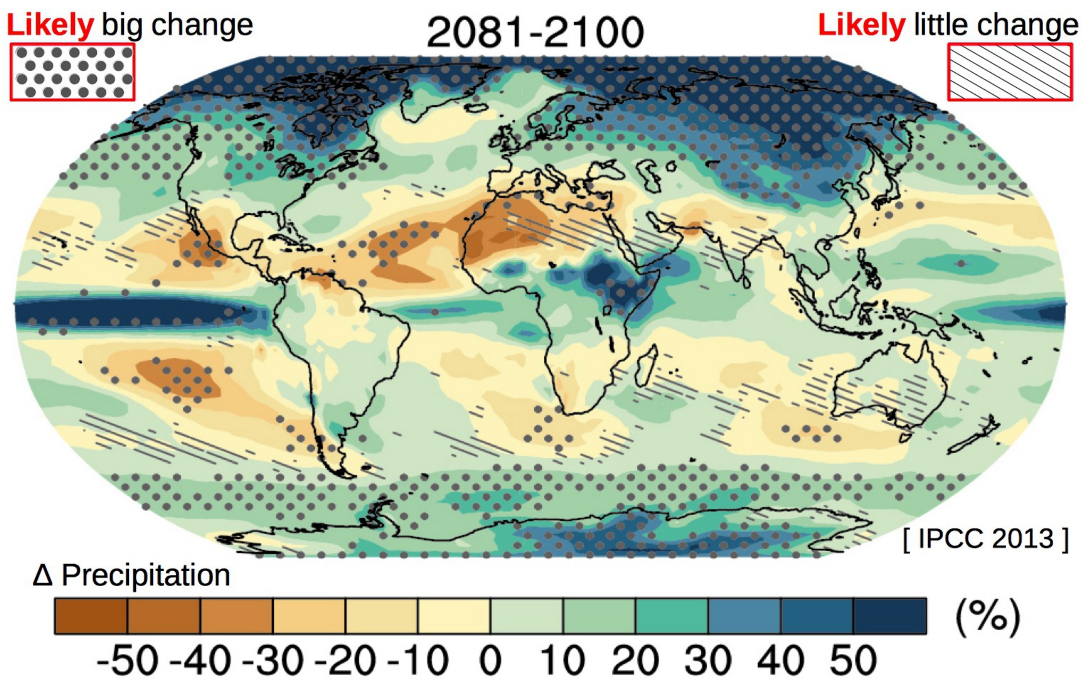


Figure 0-1 GCM precipitation agreements – Global map of CMIP5 multi-model mean results for the scenario RCP8.5 in 2081–2100 of average percent change in annual mean precipitation. Hatching indicates regions where at least 80% of the individual models agree on a change smaller than two standard deviations of natural internal variability (i.e., less than two standard deviations of variability in 20-year means). Stippling indicates regions where the multi-model mean is large compared to variability and where at least 90% of models agree on the sign of change. Regions that are not so marked have no such agreement (IPCC, 2013).

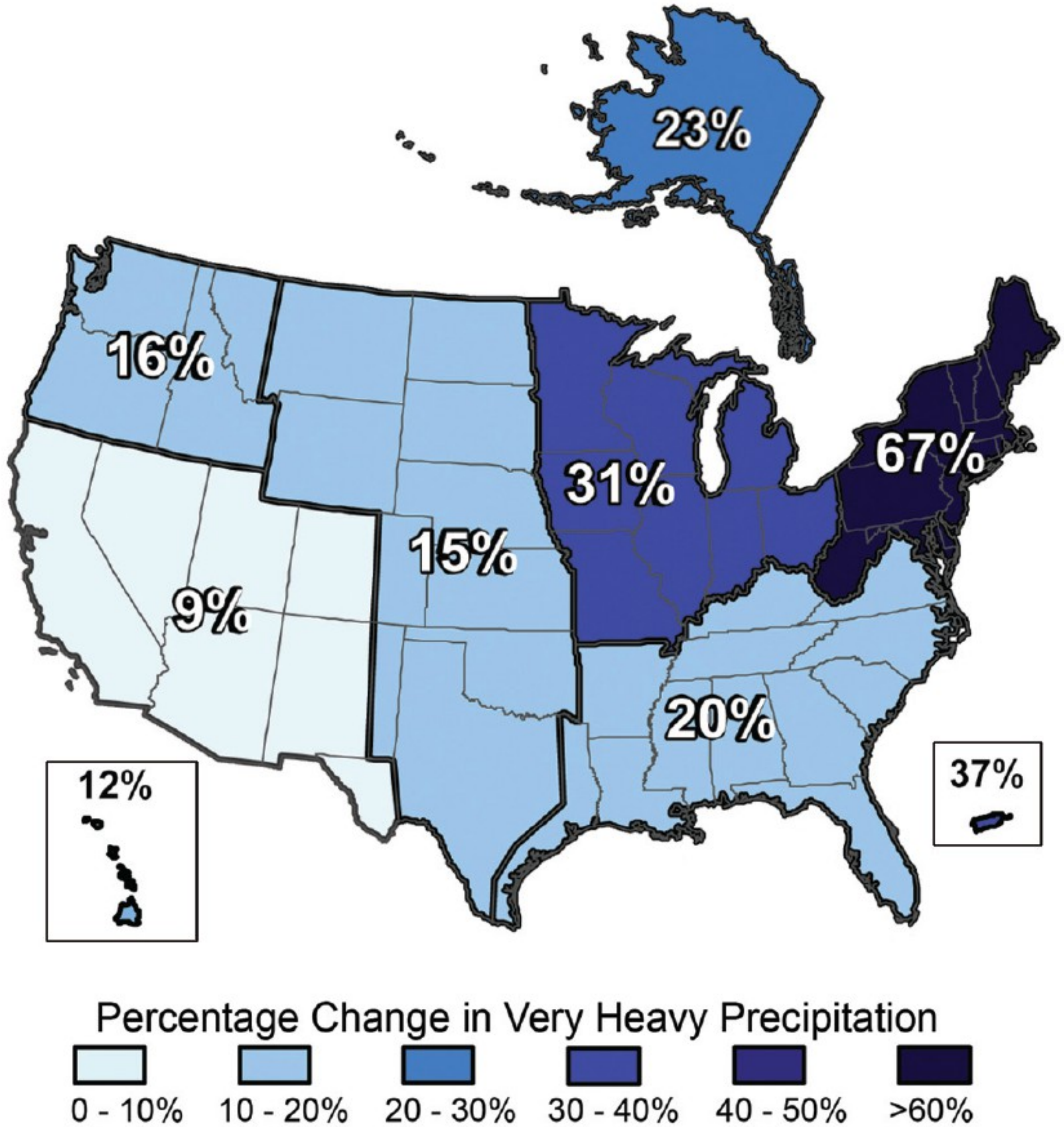


Figure 0-2 Very heavy precipitation regional changes - Increase in the amount of precipitation in very intense storms from 1958 to 2007, i.e. the top 1% of all daily events. This was described as:

"*The amount of rain falling in the heaviest downpours has increased approximately 20 percent on average in the past century, and this trend is very likely to continue, with the largest increases in the wettest places.*" (U.S. GCRP, 2009)

Chapter One

QUANTIFYING THE INFLUENCE OF MISSING DAILY PRECIPITATION DATA AND TECHNIQUES FOR ESTIMATING TRENDS IN INCOMPLETE RECORDS

Abstract

It is important to have observational trends for precipitation statistics, but this is imperiled by datasets having missing data. This work investigates the calculation of trends for precipitation event intensity, event duration, and lull pause using data sets that have gaps. From 50 long-term stations in California we obtained median trends of Intensity -0.45% per decade, Duration 0.50% per decade, and Pause 0.13% per decade.

Methods for missing data processing by Exclusion, Time Series Analysis (ARIMA), and Simple Imputation were found to be ineffective. The Multiple Imputation with Monte Carlo repetition method was shown to produce reasonable results when tested by artificially injecting missing values into the 99.8% complete Sacramento dataset. The initial intensity trend of 0.73% per decade got changed to 0.31% by missing injections and then MI moved it backwards to 0.47%. Duration started at 0.38% per decade, missing injections changed to 0.22%, and MI restored back to 0.40%. Pause started at 0.40% per decade, injections flipped it around to -0.40%, and MI brought it halfway back to 0.00% per decade. MI restored some of the impacts caused by that missing data injection and yielded a deviation from their original values by an average 0.76 of the standard deviation of the MI probability distribution process.

1.1 Introduction

Precipitation is important to understand because it is the basic driver of the hydrological cycle and is therefore fundamental for water resources. There are many water resource components of possible interest, including stream flows, flooding, soil moisture, evaporation, or groundwater recharge. But no matter which component is of concern, knowledge of precipitation is critical. In particular, to know the future of water resources, insight on precipitation changes is required.

Like many observational datasets, precipitation data records are usually incomplete. This is particularly to be expected for long-term observations back into the 19th century or the first half of the 20th century when observations were commonly recorded manually. Human error might cause observations to be forgotten or mislaid, and obviously erroneous values might be recorded and then later discarded. Even modern automatic recording devices can break or otherwise malfunction and thus miss observations.

Therefore, it is expected to find missing gaps in precipitation time series. These can be gaps of a few isolated days scattered here and there across the time range. A long sequence of missing data can happen because of human administrative or technical problems, such as an employee quits or is fired, funding is lost, or equipment breaks and takes a long time to be repaired. Such long down times can start and end at seemingly random times. However, some missing sequences are more systematic, such as gaps that start at the first day of a month and continue until the last day.

Daily precipitation data can, and often is, treated as an undifferentiated sequence of values. All of the typical statistical metrics such as mean, median, maximum, and standard deviation, can then be applied to that data sequence and used for inference. Other more specialized metrics such as extremes and trends can also be defined and applied to the data.

Daily precipitation data can be grouped into values for time periods longer than the daily periods, such as totals for precipitation in each month, or each quarter, or each year. Statistical metrics can then be applied to these period values.

An important consideration for the validity of such statistics is the completeness of the data record. The concern is that missing values may add bias or yield results that are not representative or change natural variability and spread.

There has been a long and continuing academic interest and efforts focused on the correct statistical handling of missing data (e.g. Baraldi and Enders, 2010; Graham, 2009; Pigott, 2001; Little and Rubin, 1987). That said, many of the problems, methods, and results are specific to the particular field of interest.

The goal of this chapter is to consider a spectrum of methods to be tested for their analysis of precipitation data when some of those data values are known to be missing (Schafer and Graham, 2002). The investigation will focus on the derivation of valid statistical trends of three specific precipitation metrics Intensity, Duration and Pause when working with data sets that are missing data.

1.2 Materials

The primary material necessary for this work is the source and properties of available observational precipitation datasets. Also important is computation for the statistical processing of this data and, in particular, the proper handling of missing data and the derivation of trends of precipitation metrics.

1.2.1 Precipitation Data

For this study we utilize climate data collected, analyzed, and made available by the National Climatic Data Center (NCDC), which is a part of the federal National Oceanic and Atmospheric Administration (NOAA) (Karl and Koss, 1984; Karl et al., 1986). The Global Historical Climatology Network (GHCN) inventory contains data from over 91,000 sites world-wide, with almost 50,000 located in the United States.

There is monthly data available, but its low frequency makes the data too coarse to be able to identify the precipitation features of concern in this study. Hourly data is available for some stations, but these high-frequency records are not provided at enough locations for long-term analysis. For example, there are only three stations in California that provide hourly data and extend back 75 years (to 1939).

For this study we use the GHCN daily climate records from the Global Historical

Climatology Network-Daily dataset, which is also known as GHCND (Klein Tank et al., 2002). There are 2,339 GHCND stations in California, of which 2018 provide precipitation records. Each GHCND station is identified with an 11 character identification code, e.g. the id for Santa Cruz, California is 'USC00047916'. Detailed information about GHCND stations and their data is found in Appendix 1-1.

Of the GHCND stations in California, there are 50 precipitation stations that have both long periods of at least 85 years and no more than 7% of their daily data missing. We focus on these 50 stations for this study (**Table 1.1**). Of the 50 sites chosen, the average quantity of missing data is 4.1%. Only four of those sites have less than 1% of their daily records missing. The lowest two of these stations, San Francisco and Sacramento, are particularly noteworthy and important in this work for their statistical inference properties.

Sacramento has 138 years of precipitation records, which is the second longest running station in California. This extent is more than twice as long as the full PDO cycles of 60 years. It also has a very low 0.2% of data missing, which gives it the capability to provide reliable, long-term statistical results.

San Francisco has a much shorter 94 year precipitation record. But its unique feature is that there is only one day missing, April 12th, 1947. Therefore, it can be expected to provide statistical results that are virtually identical with what would be produced if no data were missing.

1.2.2 Software

One of the major components in this study is a custom built software system called *ClimateData,* which is used for the access, manipulation, and statistical processing of precipitation and other climate data. This ClimateData software is described in Appendix A1.2.

1.3 Methods

1.3.1 Station Joining

The first technique to describe is a method to "join" stations. Often climate recording, which has been performed at one position for some time, ceases there. This recording later is performed at another place, which may be identified in the GHCND as a different Station with a unique Station Id. There are some stringent conditions that are followed to help ensure that this is essentially a single recording that has just been relocated. The new Station must be a recording that starts after the old one has ended and begins no more than two calendar years later, i.e. first could end in 1934 and second must begin no later than in 1936. The two Locations must be no further apart than 5 km and with an elevation difference of less than 150 m. If two or more such related Stations are found, then their records are automatically concatenated and treated as a single record whenever the first, initial Station's precipitation data is loaded.

1.3.2 Data Loading

There are a number of issues that must be resolved and properly handled in the loading and processing of GHCND data. The data is supplied as text files with each line of text tagged with and containing data for one metric. This work is just concerned with precipitation data. Other data types are not analyzed.

Each line is also tagged with and contains data for one month. Sometimes months get repeated or are skipped. In the first case, duplicate months are ignored. In the second case, skipped months are installed into the DataSet and daily values are marked as “missing” in the precipitation dataset, using the special IEEE floating number format "Not-a-Number" value (NaN).

Associated with each value in the file is a measurement-flag. The flag “T” indicates a trace amount of precipitation. We indicate this condition by storing a 0.05 mm value, i.e. half way between zero and the lowest representable value of 0.1 mm for precipitation.

1.3.3 Precipitation Features

In this work, the precipitation data is not treated just as an undifferentiated sequence of values or calendar period summaries. Instead we define and identify specific precipitation features within that data sequence.

Daily precipitation values are typically categorized and studied in particular value ranges. Traditional precipitation analysis often distinguishes the properties of wet days vs. non-wet days. Some researchers have defined a wet day to be any day with precipitation strictly >1 mm/day (Dai, 2001). Others consider a wet day to be precipitation ≥1 mm/day (Leander et al., 2013; Klein Tank et al., 2002); the latter definition is applied to the present study.

The first defined precipitation feature is a precipitation Event, the maximal sequence of adjoining and uninterrupted wet days. People might talk about such a precipitation Event colloquially as a rain storm or snow storm.

The second precipitation feature is a precipitation Lull, the maximal sequence of adjoining and uninterrupted non-wet days.

The third precipitation feature is a precipitation Gap. This is a maximal, uninterrupted sequence of days within the precipitation data sequence for which there is no data available, i.e. missing data.

None of these features can be preceded nor followed by a feature of the same class since their union would be a single feature of that same class but a longer sequence. Therefore a precipitation event must be preceded and followed by either Lulls or Gaps (or else be at the beginning or end of the entire sequence).

The end result of this feature identification process is a chronological sequence of features. Each such feature corresponds to one or more days in the chronological precipitation data sequence. Each day in that data sequence is contained within exactly one feature.

1.3.3.1 Uncertainty Associated with Gaps

If there is a Gap feature in the data record, then the precipitation quantity for that missing day (or multiple days) is by definition unknown. This missing data introduces statistical uncertainty about inferences for the dataset.

However, the existence of a Gap also removes statistical certainty about the two Features located on either side of the Gap. If these two adjoining Features were of the same type, then the Gap could have either joined them together into a much longer Feature of that same type, or else kept them separated. When the two adjoining features are of different types, then at least one of them would have been lengthened. So the uncertainty introduced by missing data infects adjoining Features and so has a much bigger impact on statistical certainty.

Missing data can be particularly harmful when it is associated with the Gaps between storms. Gaps last for ten days on average and at times they can persist for several months during the dry California summers. Because summers are often dry, a gap of a single day can influence interpretation of a long time period.

1.3.4 Metrics

For each precipitation Feature, there are various methods to specify and compute many statistical metrics of interest. For this work, we will utilize just three basic metrics, two metrics for precipitation Events and one metric for precipitation Lulls.

The first precipitation Event metric is Intensity, defined as the mean rate of precipitation for such an Event. This is just computed as the sum of the precipitation amounts for all the days of that Event divided by the length of that Event in days. This gives an Intensity metric in units of mm/day.

The second precipitation Event metric is Duration, the length of that Event in units of days.

The single precipitation Lull metric is Pause, defined to be the length of a Lull, measured in days.

1.3.4.1 Set Metrics

For sets of multiple Features of the same kind, there are various methods to specify and compute metrics of interest for such a set. For a set of precipitation Events, we define its Intensity of that set to be the mean of the set of Intensity values. Similarly, the Duration of a set of Events is the mean value of Event Durations. As an aside, these two examples of a mean value calculated from a set of mean values is termed a Grand Mean, which has interesting statistical properties discussed in Chapter 2.

For a set of Lulls, we define the Pause to be the median value of the set of Lull Pause values. We use the median rather than the mean because of the nature of precipitation patterns in California. In California, most days of a year are dry. Across the 50 stations chosen for this study, an average of 81% of the days are dry. So the number of dry days in a year is four times larger than the number of rain days. If the number of rain days were to change by **X**%, the fractional change in dry days would be **-X/4**%. Thus the number of dry days is relatively constant. Since the mean is this number of dry days divided by the count of Lulls, then the mean is primarily determined by just the count of Lulls in a year. In contrast, using the median of Lull Pauses provides some measure of the actual variability of the length of these Lulls instead of their count.

1.3.5 Robust Statistics

Precipitation and its various statistics do not have a normal distribution. The statistical methods and tools that depend on normality must be avoided since their built-in assumptions and results might well be invalid. Instead it is advisable to use non-parametric or distribution-free methods that avoid these normality assumptions.

In addition, methods commonly applied to normal distributions such as standard deviation or least-squares regression give extra added weight or emphasis to data outliers. Methods designed for use with non-normal data sets are better because they tend to de-emphasize the influence of outliers.

Instead of the typical use of standard deviation as a measure of scale, spread, or dispersion, this work employs the median absolute deviation (MAD). The MAD is defined as the median of the absolute value of the deviations from the median of the samples X, or alternately:

$$MAD = median_i(|X_i - \hat{X}|), \qquad where\ \hat{X} = median_j(X_j)$$

Instead of the typical linear fit estimated by least-squares regression, a preferable method used in this work is the least absolute deviations approach (Birkes and Dodge, 1993).

1.3.6 Missing Data Handling

It is clear that some process is needed to handle the missing data. We can instantly dismiss allowing missing data to just remain in the dataset because of fatal calculation problems.

1.3.6.1 Exclusion

A frequent process to handle missing data is ignore the gaps and calculate statistics on the available data. Exclusion is how a statistical computation system such as the R Project for Statistical Computing provides capabilities for the handling of missing data (R Core Team, 2013).

One statistical criteria typically used to check if ignoring missing data might be justified is to verify if the data is Missing Completely At Random (MCAR). This MCAR holds if there is no correlation between whether a particular data item is missing and any other data items either existing or missing. In other words, missing items are a random subset of the dataset. A weaker criteria called just Missing At Random (MAR) allows missing to be dependent on other variables, such as a station's temperature record. The opposite of these two criteria is called Missing Not At Random (MNAR) (Heitjan and Basu, 1996).

1.3.6.2 Time Series Analysis

A standard method for handling discrete chronological data is time series analysis that involves the creation and application of Auto-Regressive Integrated Moving Average (ARIMA) models. This technique is often used and applied for predicting future values in a time series and is most particularly applied to econometrics.

An ARIMA model forecasts a value in a time series as a linear combination of its past data values and past data deviations (or equivalently its past moving average values). The foundation of ARIMA, the Auto Regressive AR portion, consists of a weighted sum of previous values. The Moving Average or MA portion of ARIMA is a weighted sum of past forecast errors, which are differences between the past actual data and their forecasted values. There is a convertibility between AR and MA models such that any stationary AR(p) model of 'p' terms can be transformed into an MA(∞) model of infinite terms. Similarly an MA(q) model can be an AR(∞) model. The Integrated portion of ARIMA means that the AR and MA are applied to an initial differencing of the data. For all these ARIMA forms, the model's effectiveness is always dependent and predicated on having highly significant autocorrelations. In fact, analysis of data auto-correlations is typically performed first to determine which ARIMA forms and their parameters fit best (Bradley, 1985; Chatfield, 2003).

Note that although ARIMA model analysis is traditionally used for forecasting, it also has been recommended as a method to infill the missing values from a time series (Rosen and Porat, 1989; Walter et al., 2013). Applying a conventional ARIMA model to replace missing values would only use part of available information, i.e. past values. More effective missing replacement methods might use bi-directional ARIMA models to incorporate following values.

1.3.6.3 Simple Imputation

Simple imputation is a general term for the replacement of a missing data value with a single estimated possible value. Simple imputation can be further categorized as either endogenous or exogenous.

Endogenous imputation, also known as hot deck, is where the estimate is derived just from examining other records of the same dataset, e.g. other existing days of the same precipitation record. The time series ARIMA analysis method is a very general example of such endogenous imputation since it makes use of time shifted values from the same data.

In contrast, exogenous imputation, also known as cold deck, allows examination and use of other outside datasets for estimation. An example might be where precipitation records from other stations would be utilized. This could give the distinct advantage that information about weather conditions on the exact same day that is missing from some station might still be derived from other stations. If these other locations are relatively nearby, have similar elevations, and are not separated by high orographic features, then they might be particularly effective for estimation (Donders et al., 2006).

There are many imputation methods in use and some are quite simple, such as replacement with the mean or mode value, or repeating the last value. The more effective simple imputation methods consist of some algebraic expression involving a number of other relevant variables. This is most frequently just a linear weighted combination of such variables and is often derived through some kind of regression process (Luengo et al., 2012).

1.3.6.4 Multiple Imputation

Multiple imputation (MI) is a more sophisticated technique that can be used to understand and produce values for missing data. Unlike single imputation, which attempts to produce just a single estimate for a missing data item, multiple imputation tries to infer the probability distribution expected for the unknown value's population (Rubin, 1996). In particular, MI tries to identify the form of the distribution that fits the missing data, such as a Gaussian/Normal, exponential, Poisson, or Weibull distribution. It must also derive values for the distribution's parameters, e.g. statistical distribution center and spread. This MI method is performed by generating several imputed value estimates using different varying computational probabilistic approaches. From such a set of proposed replacements, a likely center and spread for the missing value can be derived (Schafer, 1999). Multiple imputation research has shown that 25 such values in an environment with even as high as 50% missing values is sufficient to be able to determine those two statistical quantities with a 99% confidence (Bodner, 2008).

In this work it is important to try to compute what the value of a trend would likely have been if none of the data were missing. We can never know exactly what the values were for any missing records. The best that can be done is to estimate the likely value of this missing data and its probabilistic variability.

Unfortunately it is not sufficient in this work to estimate just a single distribution for a station's missing precipitation. For a seasonal time series such as precipitation, and particularly for highly seasonal climates, the metric's probability distribution and in particular its parameter values in January would be expected to be rather different from that needed for July. In this work it is necessary to have different distribution parameter values for each month of the year and each would be used to estimate missing data in that month.

When the whole purpose of estimation is the evaluation of possible trends, then this admits to the possibility that values could either be increasing or decreasing with time. Therefore, it is also not reasonable to compute and use a single probability distribution for every January a long term record. To utilize just a single distribution for each particular month would only help to dampen or underestimate any trend by injecting the same mixture of values for missing data across the full time period. Therefore it is also necessary to have a distinct set of distribution parameters for each month within each year (Zhang, 2003).

Testing has been performed on how to compute the multiple imputations of statistic values. This test checked how many values from the previous and following months would be most effective for imputation. It was found that using the two months on either side of a target month could produce the best estimate and that adding more months caused the deviations to increase. Similarly, using values from the same month in the previous and following years was tested and showed that using up to eight years on either side produced the best estimates.

We can use all combinations of these possibilities to generate our multiple imputations. For each possibility of zero, one, or two months on both sides of the target month, we combine with the possibility of zero to eight years before and after. Thus producing 26 different multiple imputation values. From this set of 26 values it is straightforward to extract possible center and spread metrics. This MI technique is needed in order to take account of the possible range of errors introduced by missing data and its impact on a computed statistic.

1.3.6.5 Monte Carlo MI Injection

This work does not employ multiple imputation for the estimation of each actual missing daily precipitation value and then try to derive useful metrics from that imputed data. Doing so would take all errors in the estimation of all those daily precipitation values and compound and increase them in the process of metric calculation. Instead, multiple imputation will be used to directly estimate the value of the desired statistic.

Consider the computation of the long-term trend statistic of the Duration metric. First we examine that Duration metric for each month of the record. If all of the precipitation Event data used to compute the metric is Certain, then its metric value is computed and accepted. If the month has any Uncertain Feature data, then MI is used to compute and inject an estimated value that completes the metric. This is accomplished as a proportional process that combines the metric computed from the month's Certain Feature(s) and the MI computation of the Uncertain/missing Feature(s). For a 30 day month where 20 days are Certain and 10 are Uncertain, their direct computed and MI values are combined with weights 2/3 and 1/3 respectively.

Note that in computing the metric for a particular month some Features may cross that month's start/end and be only partially contained in that month. In that case, the month in question gets partial credit for the fractional part of the Feature that is contained in that month. So a precipitation Event of length 5 days of which only 2 days is located in the month in question, then only 2/5 of that Event's metric will be utilized by that month and 3/5 by the adjoining month. In calculating that month's mean Duration, to the numerator will be added 2/5 of the event's 5 day Duration (i.e. 2 days) and to the denominator is added 2/5 of the count of its single Feature (i.e.

0.25). By doing calculations this way, each month gets the proper credit for exactly what happens in that month and none of the metric values are omitted or double counted.

This method can be used to calculate likely values of the metric for a month by using these randomly chosen monthly MI values. This can be repeated to compute the metric for each month of the entire record. Then to compute a metric statistic such as the historical trend of Event Durations, it is necessary to remove the seasonality of these monthly metric values. So twelve monthly metric values are combined to yield a yearly value. This can either be done for a calendar year (January-December) or even what the California Department of Water Resources (DWR) terms a water-year (October-September). This study uses the water year, which is better by far because of the seasonality of precipitation. Using these yearly values, a trend can be calculated that has replacement of any missing or Uncertain data with reasonably likely MI filled data values instead.

Finally, this MI injection method and statistic calculation can be performed repeatedly in a Monte Carlo process. Upon each iteration, a different set of values are randomly chosen from the metric MI probability distribution. For this work, this is performed ten thousand times, then it will produce a large population of possible trend statistic values. Each value is derived from both the concrete values of Certain data and also the filling of some likely values derived from that metric's MI probability distribution. Each of these ten thousand statistic values will be different and each is equally likely. From this population, the most likely value for that trend statistic is its median. From the spread of that population of trend statistic values about its median, we get the likely expected range for that trend value. What has been produced is an MI probability distribution for the statistic itself.

This MI Monte Carlo process yields a trend value, which does not simply exclude and ignore any missing or Uncertain data. It also produces a measure of the variability of this result (Schunk, 2008).

1.4 Results

1.4.1 Exclusion

The simplest method to handle and account for missing data records is to exclude them from the data to be analyzed.

Many locations are missing data early in the period of record, including Santa Cruz, California (**Figure 1-1**). For others, such as California State University, Chico, most of their missing values are later during the period of record (**Figure 1-2**). Missing data are often not distributed randomly through out the days of the year. For example, in some California locations such as San Jose the missing data tends to be clustered in summer months (**Figure 1-3**), perhaps because summers are typically dry and precipitation monitoring might be suspended.. For some other locations such as Hetch Hetchy in California's Yosemite National Park, data gaps are concentrated in the winter months (**Figure 1-4**). Such mountainous areas might be in locations that are difficult to reach and so winter storms could interfere with their daily access.

The length of clusters of missing precipitation data might be expected to follow a binomial distribution, which declines rapidly for longer gaps. For the long-term records and relatively modest missing percentages of the 50 California stations, they are expected to experience isolated missing singles over 90% of the time, doubles less than 10%, triples a few tenths of a percent, and essentially no clusters larger. But for these stations, cluster distributions tend to be less concentrated, highly spread out, and more idiosyncratic. For example, Santa Cruz, California station 'USC00047916' has a mean cluster size of 10 days and it shows a decidedly strong spike at both 30 and 31 days in duration (one month) and at other month multiples (**Figure 1-5**). In fact, there are 23 months in the Santa Cruz record that are missing all data. These characteristics are indicative of non-random gap distribution.

The statistical impact of the simple exclusion of missing data can be even more directly demonstrated by taking the long and largely complete Sacramento dataset and marking as missing those days that are missing at another station, e.g. Santa Cruz, which is missing data from 7.0% of the period of record. Inserting gaps in the Sacramento data set, changes the trend for the annual Intensity of precipitation events from 1.73% per decade to 3.08% trend per decade (**Figure 1-6**). Applying the same strategy to introduce gaps into the Sacramento data changes the apparent trend of Event Duration from -0.28% per decade value to -0.05% (**Figure 1-7**). Clearly it is not acceptable to simply ignore significant data gaps when calculating trends in these precipitation metrics.

1.4.2 Time Series Analysis

The applicability of time series analysis methods for estimation is dependent on good correlation with past values. So the analysis for ARIMA starts with autocorrelation for verification and then for parameter selection and fitting.

For the San Francisco station, the most complete precipitation record in the state, the autocorrelation of daily precipitation data with a lag of 1 day is 30.2% and a 2-day lag value is 17.4% (**Figure 1-8**). The autocorrelation of daily precipitation continues to decrease with longer lags, and by 7 days it has dropped down to a single digit percentage. The Santa Cruz station, which has the least complete record in this study (7.0% missing), shows almost the same results (**Figure 1-9**), suggesting that autocorrelation analysis is not heavily impacted by the presence of gaps. The theoretical autocorrelation significance levels are calculated as:

$$\pm z(1-\alpha/2) \div \sqrt[2]{N}$$

where N is the sample size, z is the cumulative distribution function of the standard normal distribution and α is the significance level. These two autocorrelation results are much higher than this autocorrelation significance level, which for 95% significance and 85 years of daily data is about ±1%. However, the 1 day lag plots for San Francisco (**Figure 1-10**) and for Santa Cruz (**Figure 1-11**) demonstrate that day-to-day variability is considerable for these stations, suggesting that filling gaps based on autocorrelation using a one day lag for a single station is unlikely to be beneficial.

Thus simple ARIMA solutions for estimating San Francisco and for Santa Cruz precipitation have error rates of at least 50%. Given the even lower autocorrelation values for lags greater than 1 day, the inclusion additional regressive or moving average terms, would as expected add complexity but did not improve results at all.

Note that if this ARIMA model is used to just estimate a single isolated missing data value, then such a level of error might be better than the alternative of repeating as estimate the mean value. Applying this approach to multiple data gaps of varying lengths is going to introduce additional errors as gaps become longer and more common in the record. In applying such a process to estimate values for one of the 30-day missing sequences that are observed, most of the estimated values will have essentially zero confidence.

1.4.3 Simple Imputation

Simple imputation is the replacing of missing data with an estimate. Exogenous imputation, in particular, is where the missing values from one dataset would be estimated using data from another. This is effective only when there is high correlation between the two records.

But precipitation in California is not only highly variable in time, but also variable in space. For example, consider the use of station "USC00049473" in Watsonville, California to impute missing precipitation values in Santa Cruz. Many of the physical location relationships are well satisfied. Both are located about 1-2 km inland along the north shore of Monterey Bay in central California. The distance between them is a modest 20.8 km. The Watsonville station is located at an elevation of 29 m and Santa Cruz is at 40 m. In the travel along State Highway 1 between them, there is just a single notable hill of height 140 m.

So one might expect that these two precipitation data sets should be similar and well correlated. But that does not appear to be true. The cross-plot of these two stations (**Figure 1-12**) shows no apparent relationship. In fact, the Pearson Correlation Coefficient between the Watsonville and Santa Cruz records is -0.0063, suggesting a vanishingly small correlation.

More sophisticated imputation techniques were applied to fill data gaps in the Santa Cruz record. This involved the use of multiple nearby stations. Some 89 stations were found within 50 km of Santa Cruz and with no more than a 250 m elevation change. All these stations were given weighting factors determined by how well their precipitation data were correlated to the Santa Cruz records. The estimates from all stations were combined by using several formulations of their relative weights. Even this more complex imputation technique was not able to accomplish significant improvement in imputation error reduction.

1.4.4 Multiple Imputation

The multiple imputation technique is used in this work to categorize, generate, and fill statistically plausible data values into time series data in place of whatever missing data was originally there. These imputed values are drawn from a probability distribution that is specified to match that of those missing data locations.

The first task was to select a suitable probability distribution form or forms to model the data. Since there are three Metrics of interest chosen for investigation, there could well be three different distribution forms. The Weibull distribution was chosen for consideration because it incorporates the Exponential distribution, which these three metrics seem to follow and is highly adaptable by use of its shape parameter (**Figure 1-13**). The Weibull distribution has frequently been applied in reliability and in lifetime and failure data analysis. In addition, the Weibull has already been found to be appropriate in modeling precipitation data (Duan et al., 1998; Wilks, 1989) and it is used widely in hydrology (Singh, 1987).

The Weibull distribution has a formula with two parameters that must be specified.

$$f(x;k,s)=k/s*(x/s)^{k-1}*\exp(-(x/s)^{k})$$

The shape parameter 'k' selects a single distribution curve out of a whole family of widely different shapes. The scale parameter 's', like most scale parameters of other distributions, causes a stretching out of the distribution curve and a related decrease of its height.

These two Weibull parameters can be determined by first computing a metric's empirical probability density function pdf(x). Then for the corresponding cumulative density function cdf(x), a Weibull Plot of

$$\ln\left(-\ln\left(1-cdf\left(x\right)\right)\right) \text{ versus } \ln\left(x\right)$$

is created. Verification that a Weibull probability distribution is appropriate comes from the observation that the data plots as a straight line. The slope of the linear regression line for that plot determines the Weibull shape parameter 'k'. Finally, the zero intercept can be used to compute the scale parameter 's' as

$$s=\exp\left(-intercept/k\right)$$

The Weibull Plot created for the probability density function for the two event metrics demonstrates good fit and the acceptability of Weibull distributions as a model. The Sacramento Event Intensity plot fit has a Least Absolute Deviation Regression Coefficient of Determination (a robust analog of the familiar R^2 value from least squares regression) value of 95.5% (**Figure 1-14**). The Sacramento Event Intensity plot yields a shape parameter value of 0.988. The Sacramento Event Duration shape parameter value is 0.742 with a fit coefficient of 96.5% (**Figure 1-15**). The match of these fitted Weibull curves to the actual data pdf are shown for Event Intensity (**Figure 1-17**) and for Event Duration (**Figure 1-18**).

The Sacramento Lull Pause shape is 0.535 with fit of 69.2% (**Figure 1-16**). The Weibull fit appears most linear up to values of ln(x) = 3.5 or so (about 33 days). After that time the data seems to follow a different distribution. For the time interval from about 30 days to 150 days, the data falls below the Weibull distribution. At longer durations the data returns to the Weibull distribution which is apparent in the plot of the fitted Lull Pause Weibull curve to its actual data pdf (**Figure 1-19**).

This Weibull Lull fit deviation suggests that it might be prudent to not go beyond a single month for imputation of Lull length (Pause). Conveniently that is exactly how these Weibull distribution fits are used in MI. They are applied to only impute the missing metric values for a single month or more frequently just a fraction of a month. So this lack of close fit for the longer time lengths is true but irrelevant for this study.

The Weibull shape is often found to be highly specific to the physical process that is involved in the underlying physical mechanism and its metrics (Papadopoulou et al., 2006; Smith and Naylor, 1987). It is thus expected that each of the three precipitation features of this study might well have different Weibull shapes. In fact, not just the median values, but also the spread of the 50 station Weibull shape values for each of the three metrics are quite distinct. The Inter-Quantile Range (IQR) for the Intensity shape extends from 0.89 to 1.04, the Duration shape IQR extends from 0.74 to 0.78, and Pause extends from 0.52 to 0.59.

The second Weibull parameter for scale can easily be determined from knowledge of the Mean or the Standard Deviation value for a feature metric. It would be inappropriate however to compute and use a single mean statistic for the entire time series. As expected, the values of these features can and do change throughout an annual cycle. Depending on where in that cycle a particular missing value(s) is located, a mean value suitable for that location should be used. In this work, a month by month categorization of that annual cycle seemed sufficient. In addition, if there is in fact a trend for a feature metric, then a mean value for an early year should be different than its value for a later year. So a mean value is collected to compute a Weibull scale parameter for each month of each year in the dataset, e.g. 1,200 values for 100 years of data.

Multiple Imputation (MI) is a statistical technique that is described as a convenient and flexible paradigm for analyzing data with missing values (Schafer, 1999). This is the technique to be applied to compute these monthly mean values. For each month, multiple estimation algorithms are performed, each of which generates an estimate of the metric for that time period. Multiple imputation research has shown that 25 such values in an environment with as much as 50% missing values is sufficient to be able to determine those quantities with a 99% confidence efficiency (Little and Rubin, 1987) The real advantage of this MI technique is that it produces not just a single representative central value, but also a measure of the variability for that substituted value.

Testing has been performed on the effectiveness of examining values from a previous and following month for imputation. It was found that using the two months on either side of a target month could produce the lowest deviation estimate for that target's value. Adding more months caused the estimate deviation to increase. Similarly, using values of the same month from a previous and following year was tested using some dozen different stations. This showed that the best numbers of years to use was a range of from 7 to 14 years on either side to produce the lowest deviation estimates. So this work has selected a median value of 10 years on either side of a target month to employ for MI estimation.

We use all combinations of these possibilities to generate our multiple imputations. For each possibility of zero, one, or two months on either side of the target month, we combine with the possibility of zero to ten years before or after. Thus this produces 33 different multiple imputation values. From this set of 33 values, we calculate a possible center (mean) and a spread (absolute deviation) metrics. This technique takes account of the possible range of errors introduced by missing data and its impact on a statistic.

The final result of this MI work is a table of Weibull distribution scale parameters for each month of each year of the record. Using this table it is then possible to construct a relevant probability distribution for a particular month and year and then to generate a random value from this probability function for that period.

1.4.5 Confirmation

We performed a test to determine the effectiveness of the Monte Carlo – Multiple Imputation filling of data gaps. This test was completed by taking a virtually complete data set, and inserting data gaps based on the pattern of gaps observed in another data set.

The complete Sacramento dataset was used to compute the trends for the three metrics: Intensity, Duration and Pause. Next, the Sacramento dataset was repeatedly gapped to have the same missing days as each of the other 48 stations with missing samples. Each of these 48 sets of gapped Sacramento data were used to compute the same three trend values from which the median was calculated. The deviation between the original data trend values and the gapped median values was considerable. The Intensity trend went from a 0.73% per decade to 0.31% per decade. Duration went from a 0.38% per decade to 0.22% per decade. The Pause trend went from a 0.39% per decade to -0.40% per decade, approximately the same magnitude but opposite sign.

The Multiple Imputation Monte Carlo filling process for missing data was applied to each of the 48 Sacramento gapped data sets, trends were calculated, and the median values calculated again. In all three cases this filling process helped restore at least some of the negative impact on the calculated trend values, which was imposed by the insertion of data gaps. The best case was the Duration trend, which was nearly restored to that calculated from the full dataset. Duration started originally at 0.38% per decade, gapping took it to 0.22% per decade, and then with MI filling it came back to 0.40% per decade. The Intensity trend started at 0.73% per decade, gapped to 0.31% per decade, and then with MI filling brought this value back to 0.47% per decade. The most extreme impact was to the Pause trend. The complete data set indicated a Pause trend of 0.39% per decade, gapping increased this to to -0.40% per decade, and MI filling resulted in a trend of 0.00% per decade.

1.5 Discussion

1.5.1 Time Series Analysis

The negative results observed from the autocorrelation 1-day lag plots are a measure of the highly variable temporal behavior of precipitation time series. Precipitation events are rarely organized to have a gradual increase and decrease in intensity through time. Gradual changes are more common in the economic forecasting that typically utilizes the ARIMA methods. Instead, precipitation can, and sometimes does, go from nothing one day to high intensity the next day and then return the third day to zero again.

However, some discussion is appropriate to understand why there is no autocorrelation visible in the lag plots, whereas the autocorrelation versus lag analysis shows a nontrivial 30% correlation for the 1-day lag, values in the teens for lags of 2-6 days, and all other values well over the 1% significance level. The key to understanding how both of these relations can exist within the same data set is in the long periods during which there is no rain. Throughout California during the period of record, 81% of the data days have no rain. The autocorrelation is based almost entirely on the likelihood of a dry day following another dry day, whereas wet days are virtually uncorrelated day to day. Another way to describe the latter relation is that most precipitation events either end after a day, or (if they persist) have a significantly different intensity day by day.

1.5.2 Simple Imputation

The negative precipitation correlation result such as is observed from the cross plot of Watsonville versus Santa Cruz daily precipitation is a measure of the highly variable and uncorrelated spatial behavior of California precipitation. Given the close proximity of these two stations, this lack of spatial correlation seems surprising. Up to 50% of California precipitation is delivered via long and very thin bands of atmospheric moisture that has been termed Atmospheric Rivers (AR) (Dettinger et al., 2011). Perhaps 90% of the water vapor moved toward the polar regions across the midlatitudes occurs in such narrow storms (Dettinger, 2011).

The existence and strong importance of such spatially, restricted precipitation delivery mechanisms in California could help to explain local variability in observed precipitation patterns. One station may lie within an AR storm track, whereas an adjacent station lie just outside the track. It would be interesting to use the historical and spatial data set to assess more extensively the nature of spatial correlation between nearby stations.

1.5.3 Multiple Imputation

The MI Monte Carlo filling process provided the least benefit when data gaps impacts were the most extensive. In the analysis using Sacramento station data, with gaps added using the pattern from the Santa Cruz data, the metric of event Duration was impacted the least, with a 0.16% per decade change and all of that was restored. Intensity was next with a 0.42% per decade change, of which 0.26% (about 2/3) was restored. Pause had the largest change of 0.79% per decade and only 0.39% (less than half) was restored.

Multiple Imputation works by determining a probability distribution for the missing data. The values drawn from this distribution will typically be its most frequent values and these are what fills in for the missing data. If the actual values had been from these frequent distribution values, then MI could fill-in with similar frequent values and this might be effective at restoring the statistics. But if, by random happenstance or the unintended consequence of decisions of what data to miss, the actual values had been extreme or infrequent, then MI would still replace with the frequent distribution values which might not be a good statistical match. The largest change from gapping for Pause might perhaps indicate that its actual data was rather extreme and infrequent.

The MI filling method does reduce the impact of missing data. This can be shown by looking at the range of the impacts on the trends resulting from data gaps. The original data Intensity trend was 0.73% per decade and the gapping impact had an inter-quartile range (IQR) from 0.05% to 0.71%, which is a spread of 0.65%. In comparison, the MI filling results ranged from 0.26% to 0.57%, a spread of 0.31%

less than half. For Duration the IQR spread was reduced slightly from 0.23% to 0.21%. The Pause results IQR spread dropped from 0.86% to 0.46%.

This MI filling method is itself a statistical process and these three final trend results deviate from their original true values by an average of 0.76 of the standard deviation of the probability distribution of that filling process. So the results are well with the expected capabilities of the method.

1.6 Conclusions

The three initial data handling methods of Exclusion, Time Series Analysis, and Single (exogenous) Imputation provide little benefit in filling data gaps for calculation of trends in precipitation metrics. In contrast, filling in the missing data with estimates generated by Multiple Imputation with a Monte Carlo process was more effective. This method does not provide a means to replace missing data, but it can restore at least some of the degradation on trend statistics caused by data gaps. As an added benefit, the method provides probability bounds on the reliability of its estimates.

References

Baraldi, A.N., and Enders, C.K., 2010, An introduction to modern missing data analyses: Journal of School Psychology, v. 48, p. 5–37, doi: 10.1016/j.jsp.2009.10.001.

Birkes, D., and Dodge, Y., 1993, Alternative Methods of Regression: New York, John Wiley & Sons, 240 p.

Bodner, T.E., 2008, What Improves with Increased Missing Data Imputations? Structural Equation Modeling: A Multidisciplinary Journal, v. 15, p. 651–675.

Bradley, C., 1985, The Absolute Correlation Coefficient: The Mathematical Gazette, v. 69, p. 12, doi: 10.2307/3616441.

Chatfield, C., 2003, The Analysis of Time Series: An Introduction, Sixth Edition: Boca Raton, FL, Chapman and Hall/CRC, 352 p.

Dai, A., 2001, Global Precipitation and Thunderstorm Frequencies. Part I: Seasonal and Interannual Variations: Journal of Climate, v. 14, p. 1092–1111.

Dettinger, M., 2011, Climate Change, Atmospheric Rivers, and Floods in California – A Multimodel Analysis of Storm Frequency and Magnitude Changes1: JAWRA Journal of the American Water Resources Association, v. 47, p. 514–523, doi: 10.1111/j.1752-1688.2011.00546.x.

Dettinger, M.D., Ralph, F.M., Das, T., Neiman, P.J., and Cayan, D.R., 2011, Atmospheric Rivers, Floods and the Water Resources of California: Water, v. 3, p. 445–478, doi: 10.3390/w3020445.

Donders, A.R.T., van der Heijden, G.J.M.G., Stijnen, T., and Moons, K.G.M., 2006, Review: a gentle introduction to imputation of missing values: Journal of Clinical Epidemiology, v. 59, p. 1087–1091, doi: 10.1016/j.jclinepi.2006.01.014.

Duan, J., Selker, J., and Grant, G.E., 1998, Evaluation of Probability Density Functions in Precipitation Models for the Pacific Northwest1: JAWRA Journal of the American Water Resources Association, v. 34, p. 617–627, doi: 10.1111/j.1752-1688.1998.tb00959.x.

Graham, J.W., 2009, Missing Data Analysis: Making It Work in the Real World: Annual Review of Psychology, v. 60, p. 549–576, doi: 10.1146/annurev.psych.58.110405.085530.

Heitjan, D.F., and Basu, S., 1996, Distinguishing “Missing at Random” and “Missing Completely at Random”: The American Statistician, v. 50, p. 207–213, doi: 10.2307/2684656.

Karl, T.R., and Koss, W.J., 1984, Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983.: National Climatic Data Center Historical Climatology Series 4-3, 38 p.

Karl, T.R., Williams, C.N., Young, P.J., and Wendland, W.M., 1986, A Model to Estimate the Time of Observation Bias Associated with Monthly Mean Maximum, Minimum and Mean Temperatures for the United States: Journal of Climate and Applied Meteorology, v. 25, p. 145–160, doi: 10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., et al., 2002, Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment: International Journal of Climatology, v. 22, p. 1441–1453, doi: 10.1002/joc.773.

Leander, R., Buishand, T.A., and Tank, A.M.G.K., 2013, An Alternative Index for the Contribution of Precipitation on Very Wet Days to the Total Precipitation: Journal of Climate, v. 27, p. 1365–1378, doi: 10.1175/JCLI-D-13-00144.1.

Little, R.J.A., and Rubin, D.B., 1987, Statistical Analysis with Missing Data: New York, J. Wiley & Sons.

Luengo, J., García, S., and Herrera, F., 2012, On the Choice of the Best Imputation Methods for Missing Values Considering Three Groups of Classification Methods: Knowl. Inf. Syst., v. 32, p. 77–108, doi: 10.1007/s10115-011-0424-2.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and Francis, R.C., 1997, A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production: Bulletin of the American Meteorological Society, v. 78, p. 1069–1079, doi: 10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Minobe, S., 1997, A 50–70 year climatic oscillation over the North Pacific and North America: Geophysical Research Letters, v. 24, p. 683–686, doi: 10.1029/97GL00504.

Papadopoulou, V., Kosmidis, K., Vlachou, M., and Macheras, P., 2006, On the use of the Weibull function for the discernment of drug release mechanisms: International Journal of Pharmaceutics, v. 309, p. 44–50, doi: 10.1016/j.ijpharm.2005.10.044.

Pigott, T.D., 2001, A Review of Methods for Missing Data: Educational Research and Evaluation, v. 7, p. 353–383, doi: 10.1076/edre.7.4.353.8937.

R Core Team, 2013, R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing, 3551 p.

Rosen, Y., and Porat, B., 1989, Optimal ARMA parameter estimation based on the sample covariances for data with missing observations: IEEE Transactions on Information Theory, v. 35, p. 342–349, doi: 10.1109/18.32128.

Rubin, D.B., 1996, Multiple Imputation After 18+ Years: Journal of the American Statistical Association, v. 91, p. 473, doi: 10.2307/2291635.

Schafer, J., 1999, Multiple imputation: a primer: Statistical Methods in Medical Research, v. 8, p. 3–15.

Schafer, J.L., and Graham, J.W., 2002, Missing data: Our view of the state of the art: Psychological Methods, v. 7, p. 147–177.

Schunk, D., 2008, A Markov chain Monte Carlo algorithm for multiple imputation in large surveys: AStA Advances in Statistical Analysis, v. 92, p. 101–114, doi: 10.1007/s10182-008-0053-6.

Singh, V.P., 1987, On application of the Weibull distribution in hydrology: Water Resources Management, v. 1, p. 33–43, doi: 10.1007/BF00421796.

Smith, R.L., and Naylor, J.C., 1987, A Comparison of Maximum Likelihood and Bayesian Estimators for the Three- Parameter Weibull Distribution: Applied Statistics, v. 36, p. 358, doi: 10.2307/2347795.

Walter, Y., Kihoro, J.M., Athiany, K.H.O., and W, K.H., 2013, Imputation of incomplete non-stationary seasonal time series data: Mathematical Theory and Modeling, v. 3, p. 142–154.

Wilks, D.S., 1989, Rainfall Intensity, the Weibull Distribution, and Estimation of Daily Surface Runoff: Journal of Applied Meteorology, v. 28, p. 52–58, doi: 10.1175/1520-0450(1989)028<0052:RITWDA>2.0.CO;2.

Zhang, P., 2003, Multiple Imputation: Theory and Method: International Statistical Review, v. 71, p. 581–592, doi: 10.1111/j.1751-5823.2003.tb00213.x.

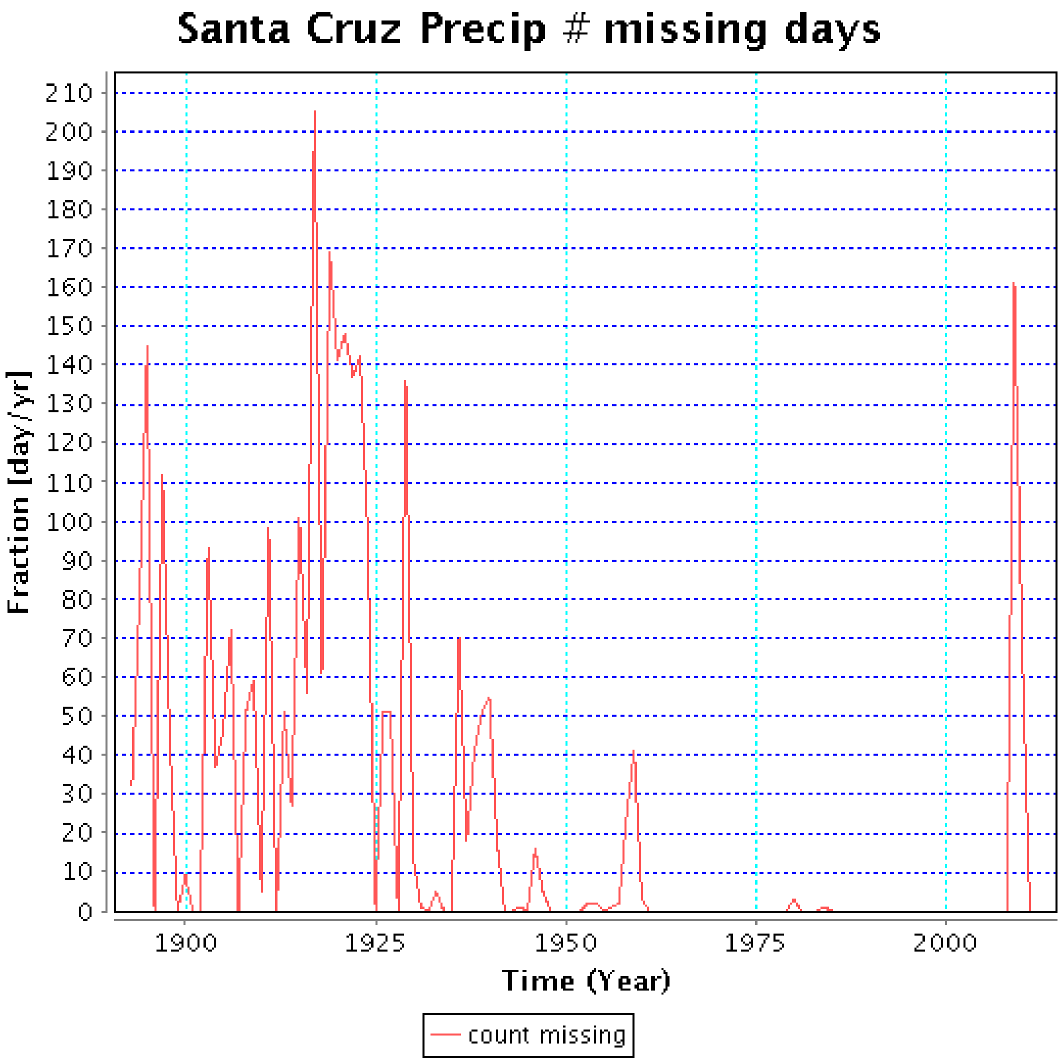


Figure 1-1 Early occurrence of missing days. Count of missing days within each year for station 'USC00047916' in Santa Cruz, California. This diagram shows that almost all its missing days occur early in the period of record. This is represents a decidedly Missing Not At Random (MNAR) situation.

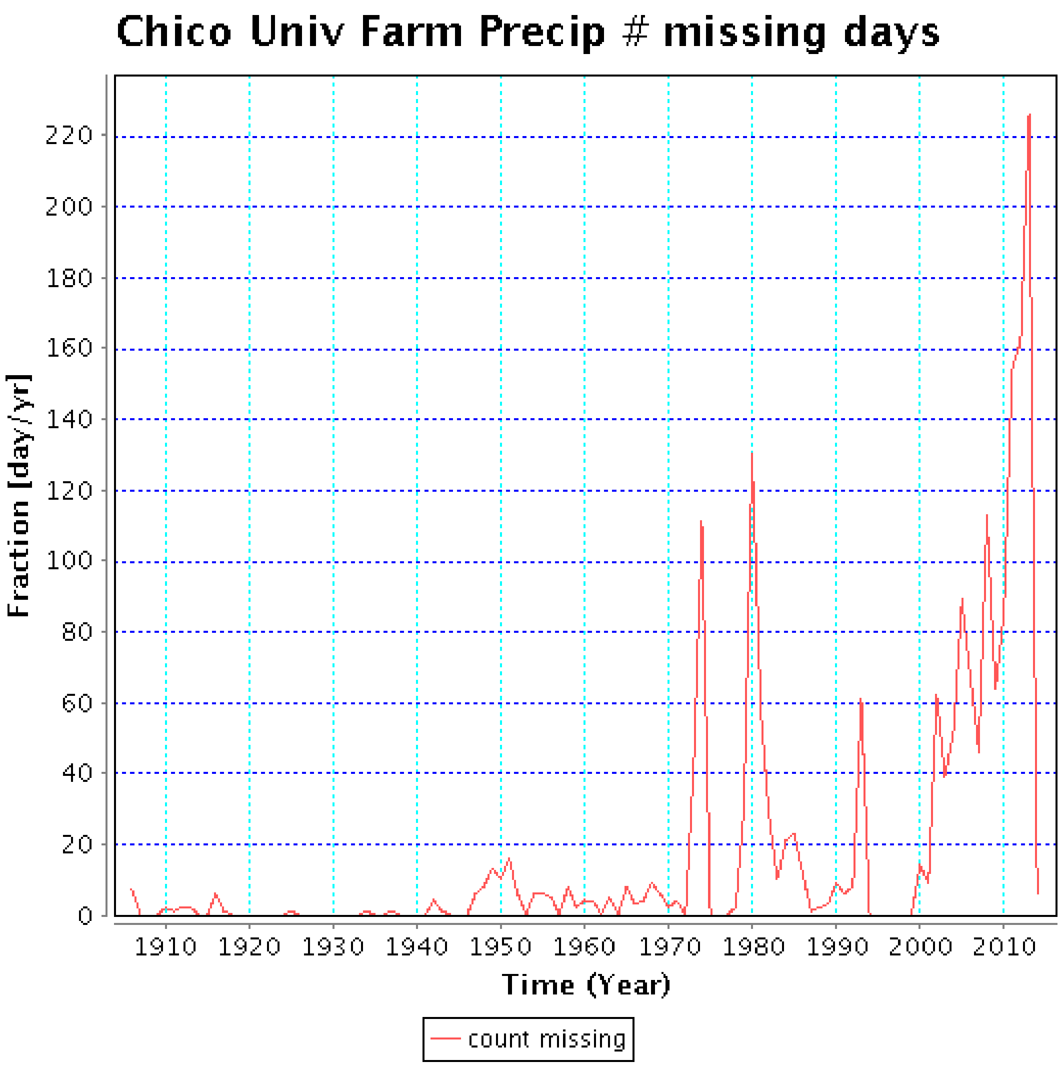


Figure 1-2 Late occurrence of missing days. Count of missing days within each year for station 'USC00041715' in Chico, California. This diagram shows that most of its missing days occur late in the station's period of record. This represents a decidedly Missing Not At Random (MNAR) situation.

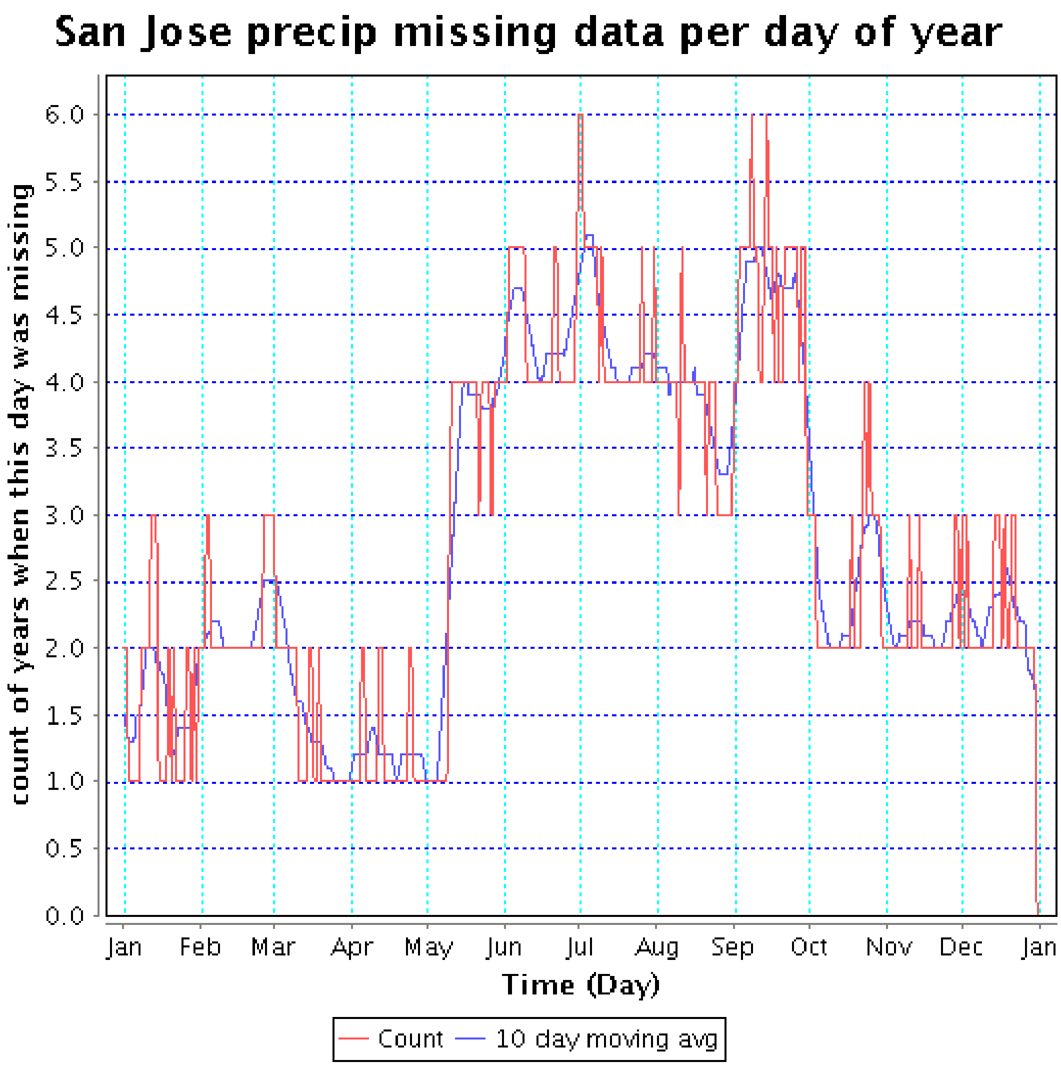


Figure 1-3 Summer occurrence of missing days. Count of missing days within each day of the year for station 'USC00047821' in San Jose, California. This diagram shows that many of the data gaps occur in the summer season. This represents a Missing Not At Random (MNAR) situation.

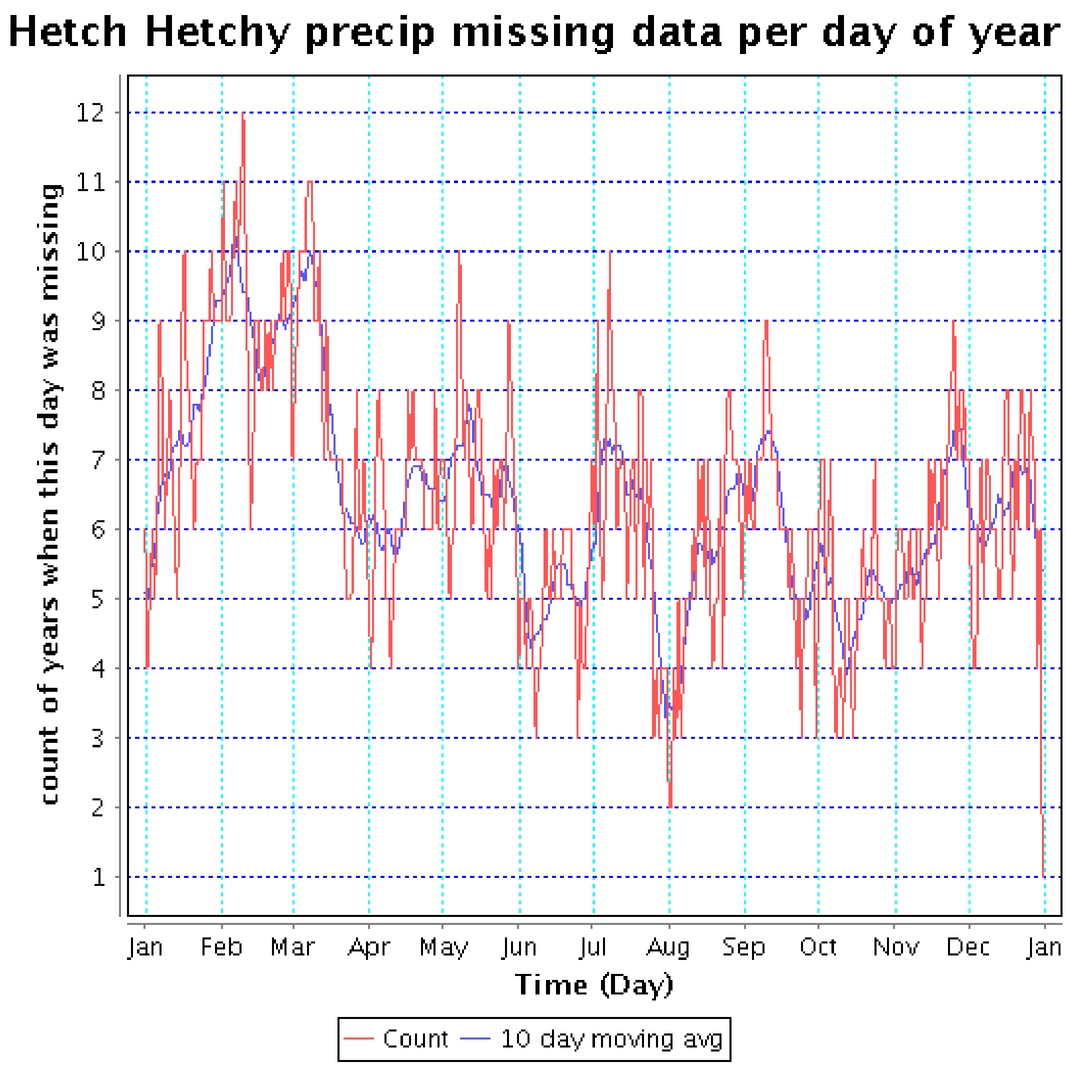


Figure 1-4 Winter occurrence of missing days. Count of missing days within each day of the year for station 'USC00043939' in Hetch Hetchy, California. This diagram shows that more of the data gaps occur in the winter season. This lends some support to a Missing Not At Random (MNAR) situation.

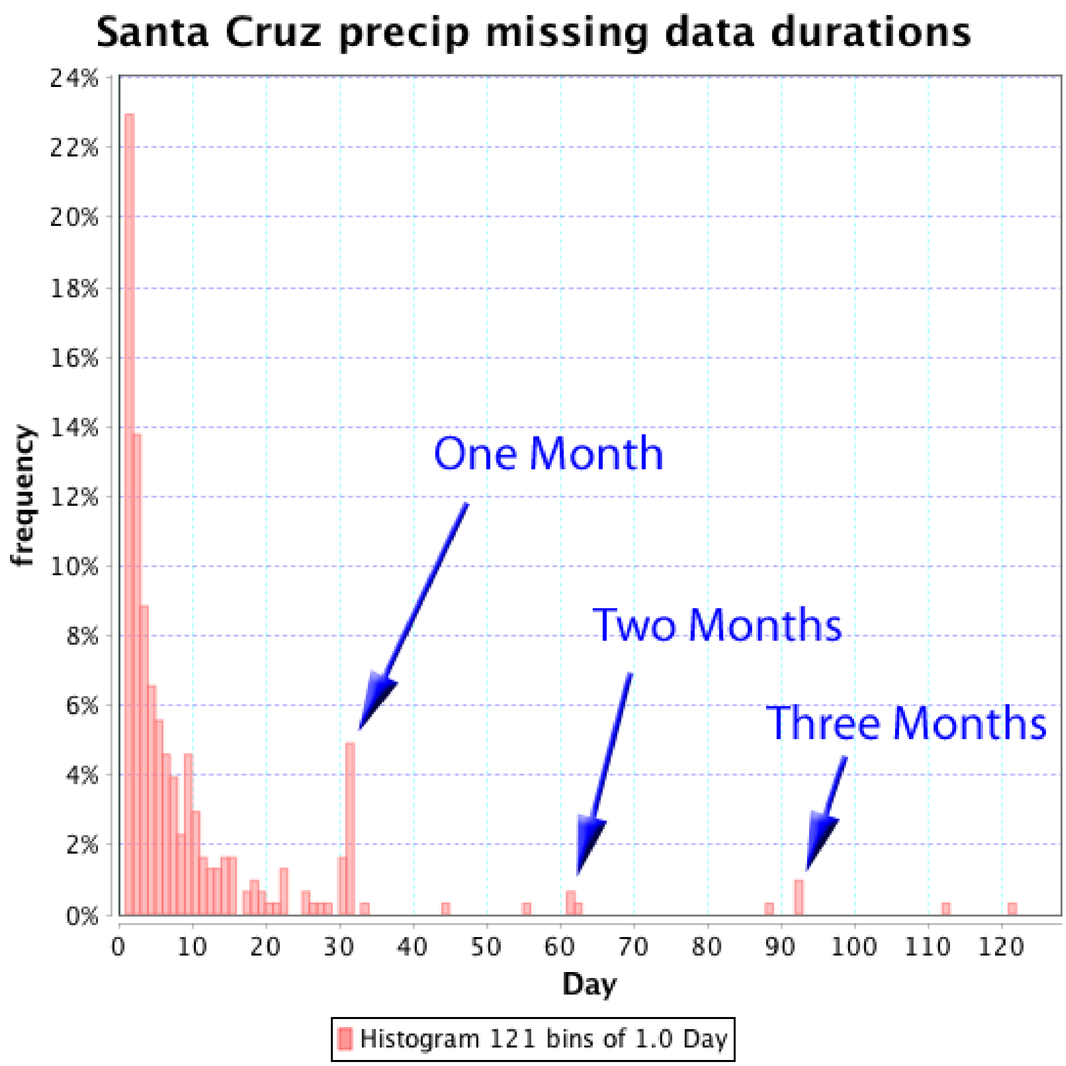


Figure 1-5 Length of clusters of the missing days for station 'USC00047916' in Santa Cruz, California. Of the 43,494 days in its record there are 3,055 missing (7.0%). If these missing days were randomly placed, then a little over 93% would be expected to occur as isolated singletons, somewhat over 6% as twins, perhaps a couple tenths percent as triplets, and essentially nothing larger. In contrast, the mean length of clusters in this Santa Cruz record is 10 days. Note the distinctive peaks occurring at lengths of one, two, and three months. All this is very strong support of Missing Not At Random (MNAR).

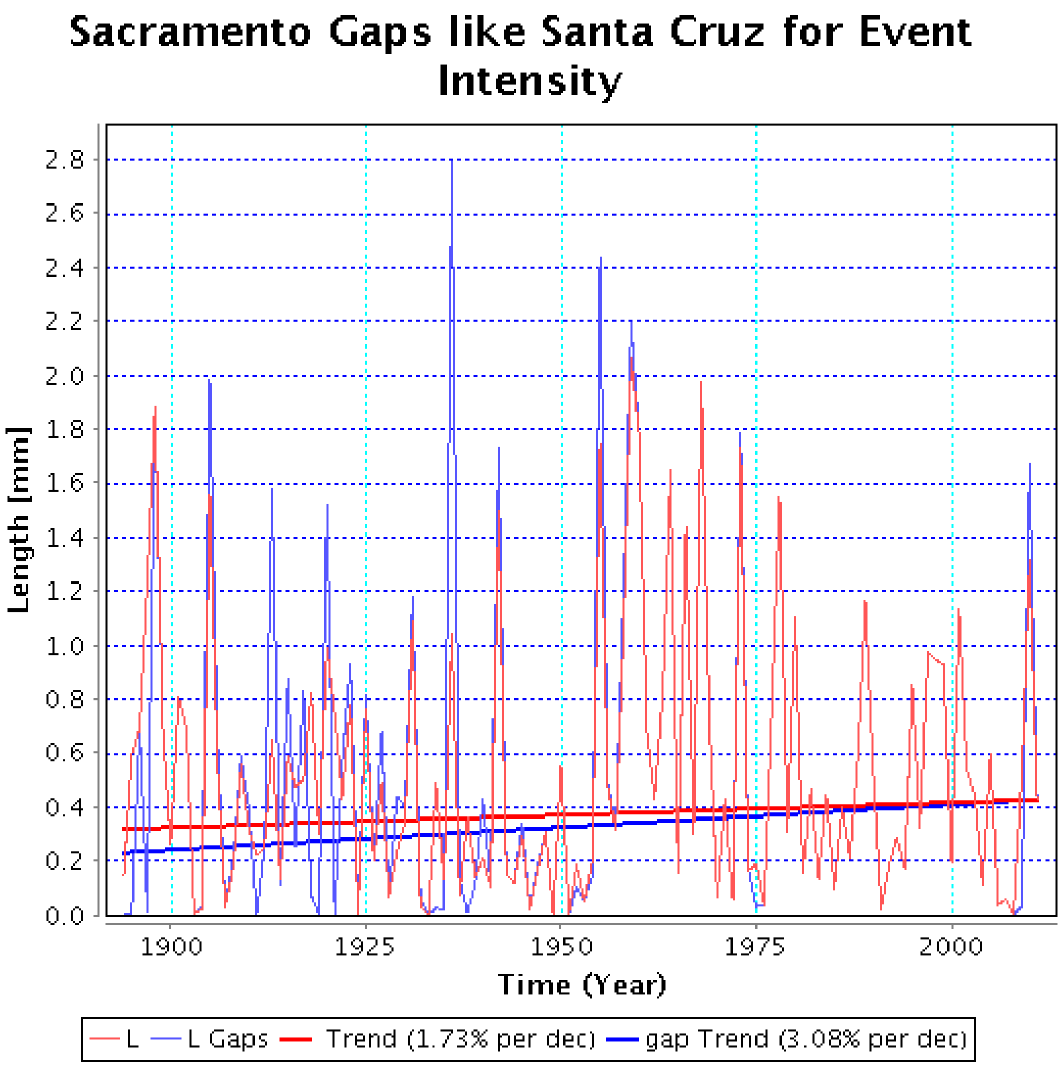


Figure 1-6 Event Intensity trends before and after gapping. The plot of annual values of Event Intensity for station "USW00023271" in Sacramento, California together with its long-term trend, both shown in red. The blue color shows the same information about Sacramento but after the 3,055 days missing in the record for station "USC00047916" in Santa Cruz are then marked as missing within the Sacramento record. So Sacramento has gone from missing 72 days (0.2%) to missing 3,095 days (7.1%). This demonstrates the serious impact on the trend value from the original 1.7% per decade to 3.08% per decade after the missing additions.

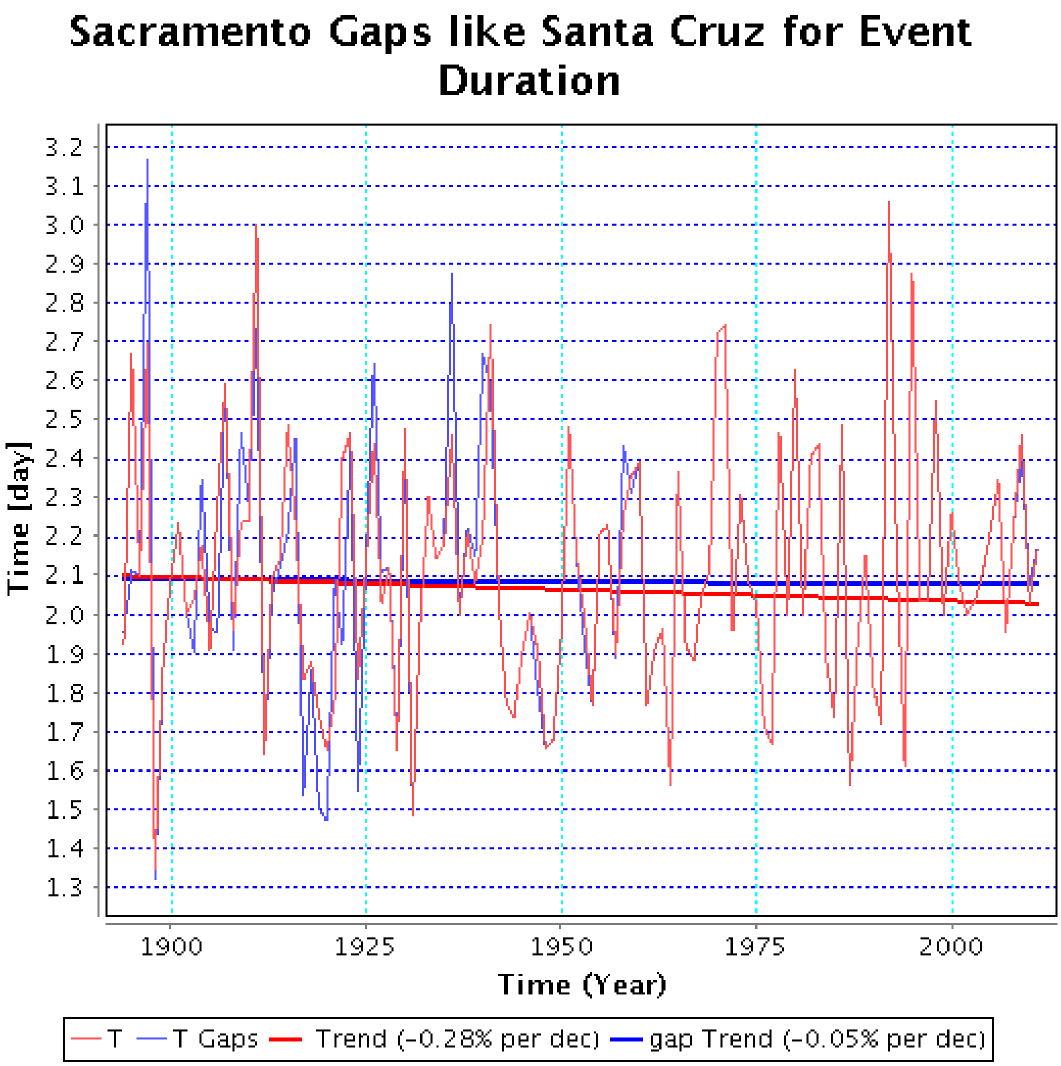


Figure 1-7 Event Duration trends before and after gapping. The plot of annual values of Event Duration for station "USW00023271" in Sacramento, California together with its long-term trend, both shown in red. The blue color shows the same information about Sacramento but after the 3,055 days missing in the record for station "USC00047916" in Santa Cruz are then marked as missing within the Sacramento record. So Sacramento has gone from missing 72 days (0.2%) to then missing 3,095 days (7.1%). This demonstrates the serious impact on the trend value from the original -0.28% per decade to -0.05% per decade after the missing additions.

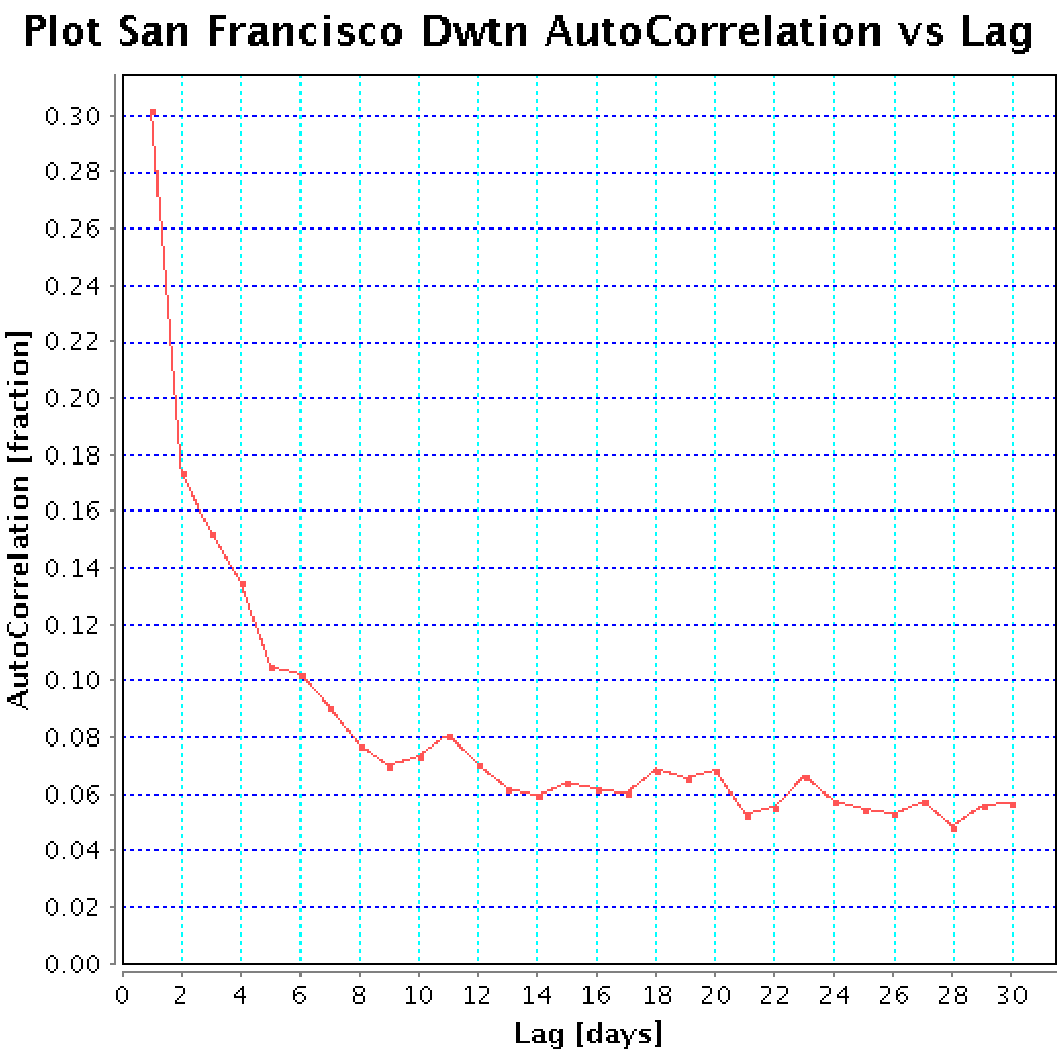


Figure 1-8 Autocorrelation of the daily precipitation record for station "USW00023272" in San Francisco, California for lags from 1 day to 30.

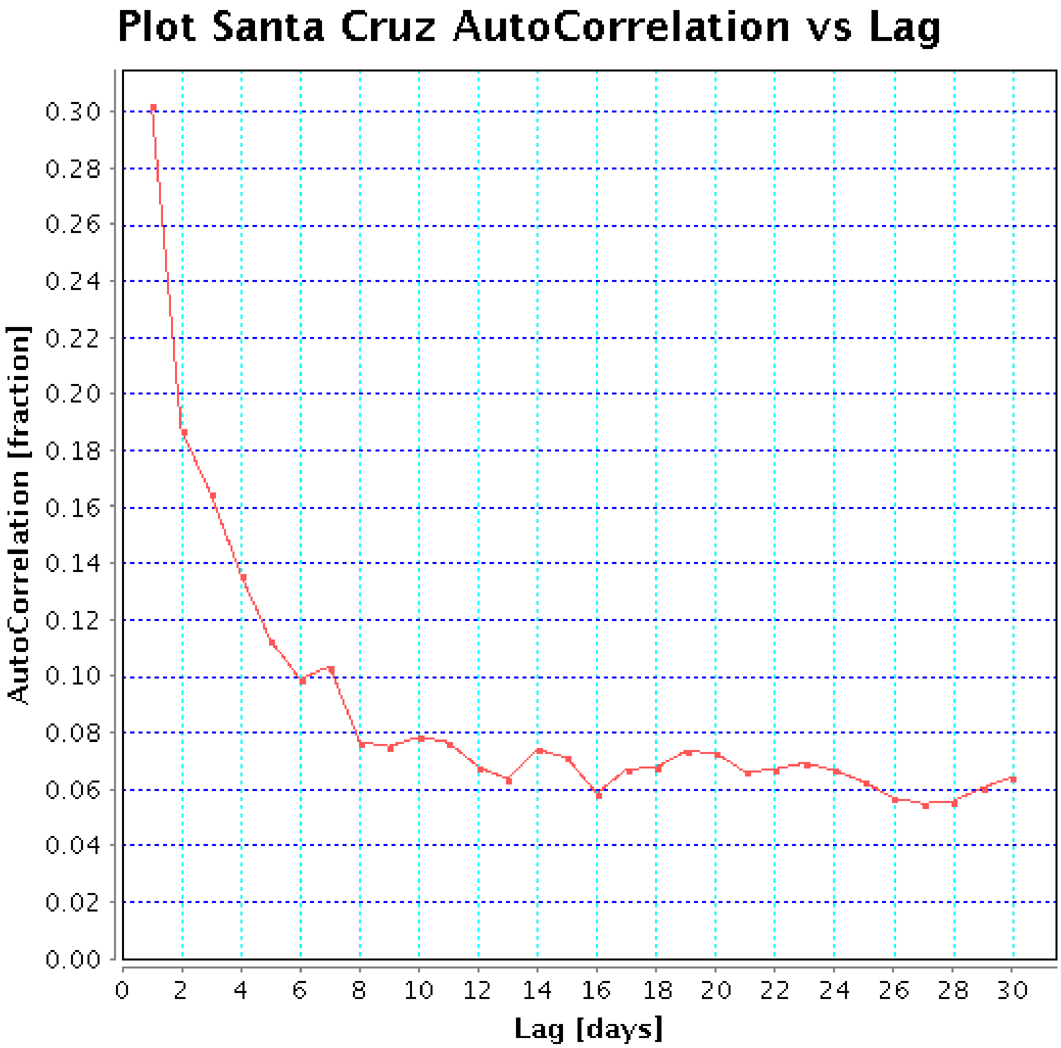


Figure 1-9 Autocorrelation of the daily precipitation record for station "USC00047916" in Santa Cruz, California for lags from 1 day to 30.

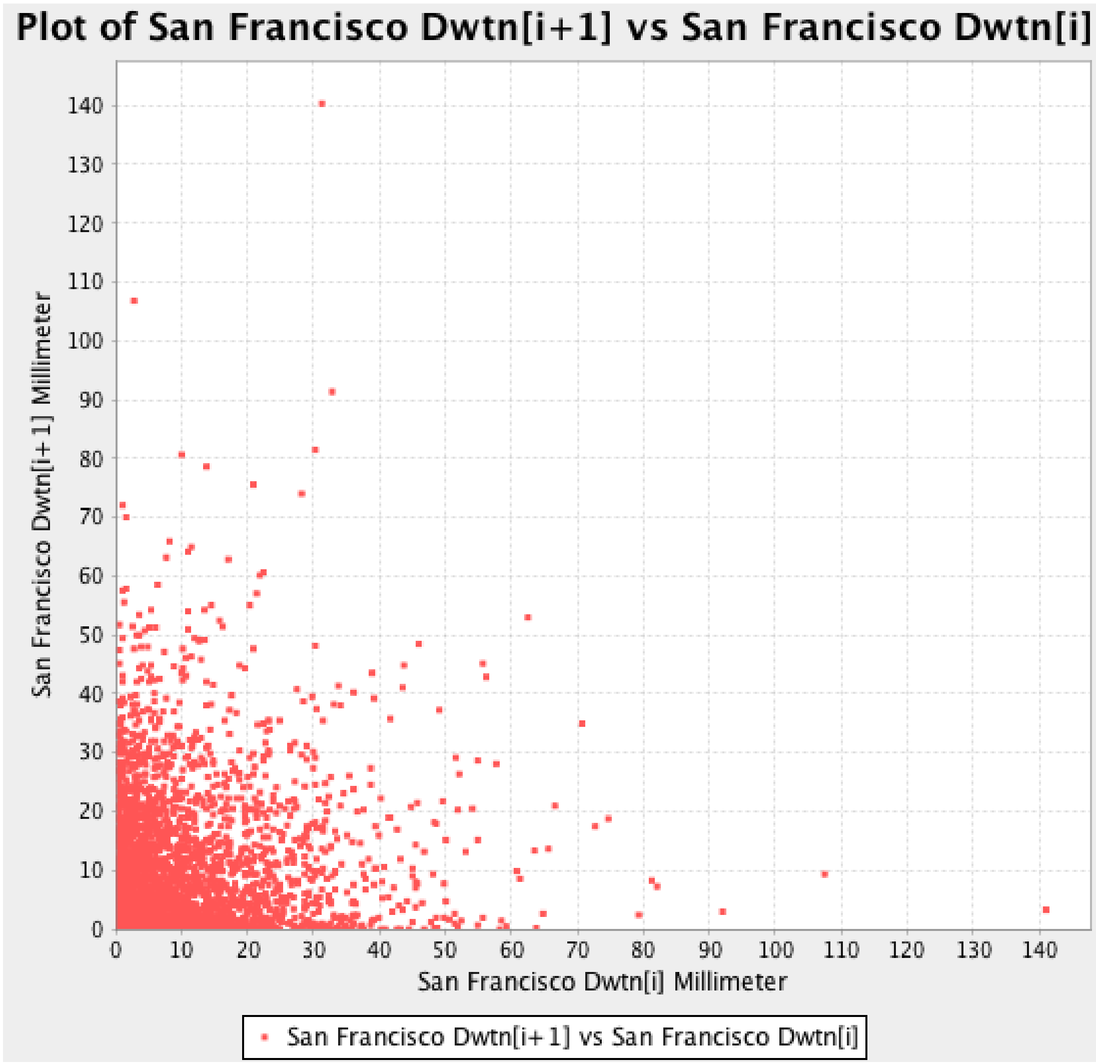


Figure 1-10 Autocorrelation 1 day lag cross-plot of the daily precipitation record for station "USW00023272" in San Francisco, California. Each point shows the relationship of the value for one day of the record to that for the next day. Note the lack of any trend tendency.

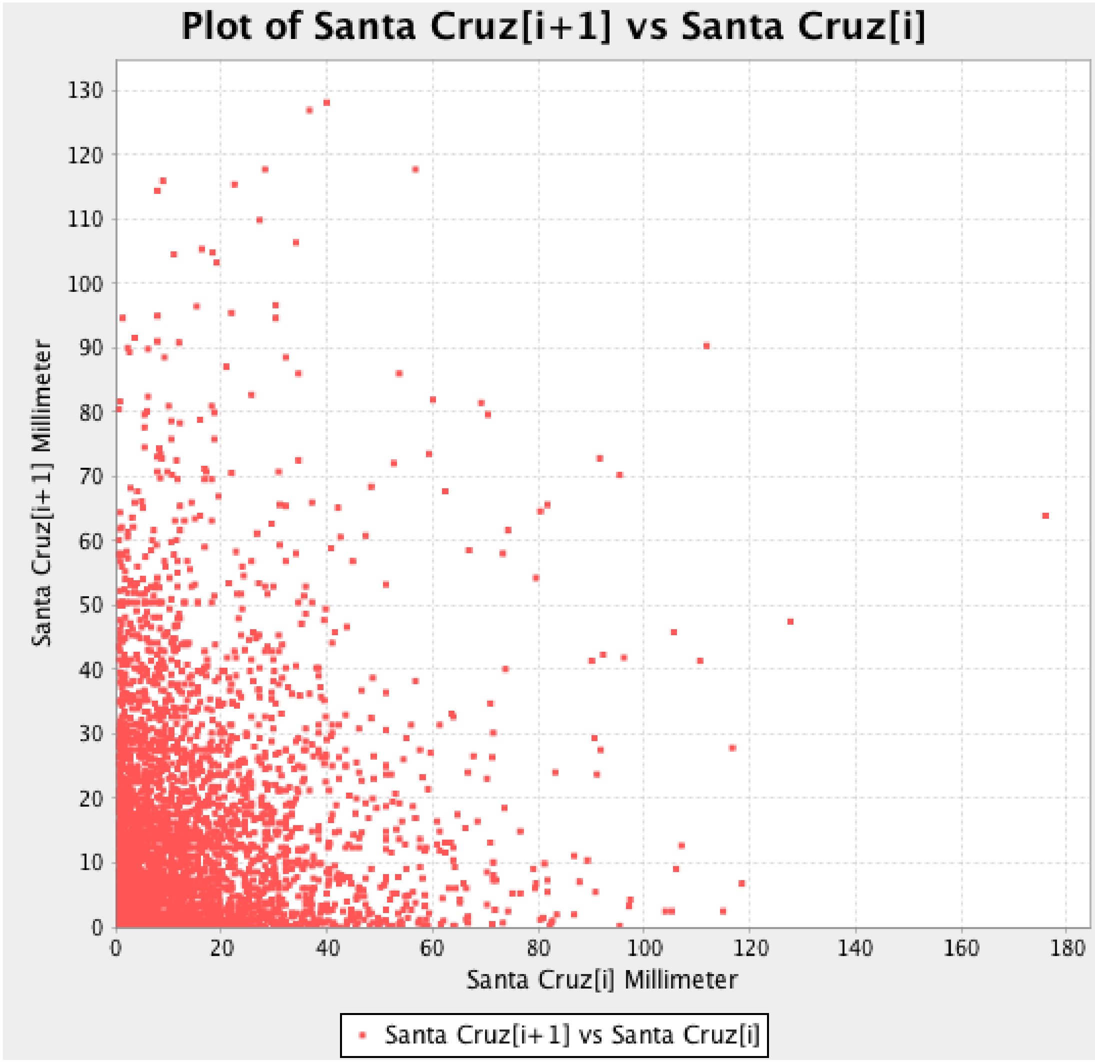


Figure 1-11 Autocorrelation 1 day lag plot of the daily precipitation record for station "USC00047916" in Santa Cruz, California. Each point shows the relationship of the value for one day of the record to that for the next day. Note the lack of any trend tendency.

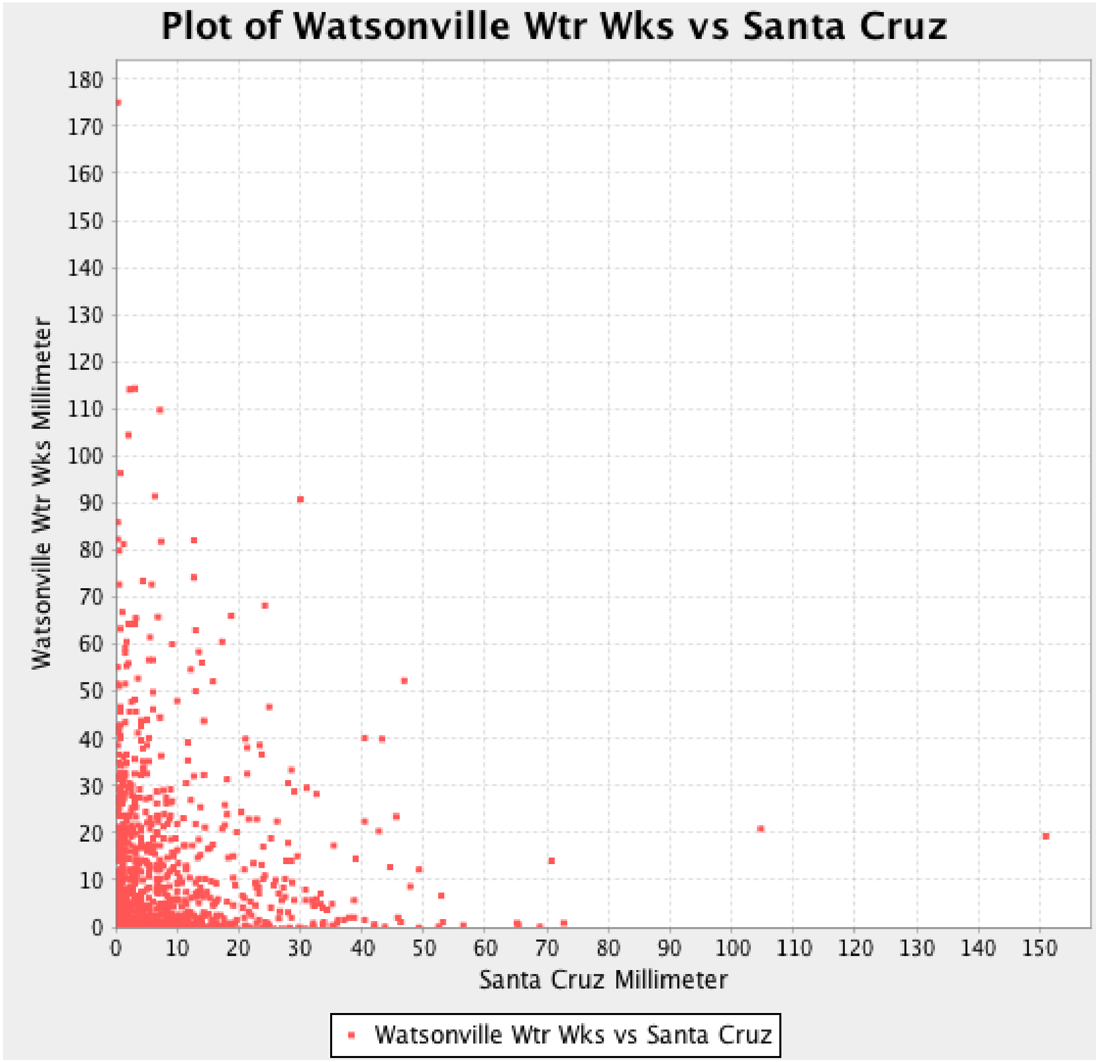


Figure 1-12 Cross plot of station "USC00049473" in Watsonville, California versus station "USC00047916" in Santa Cruz, California. Each point shows the relationship of the value for one day of the Watsonville record to that for the same day for Santa Cruz. Note the lack of any trend tendency.

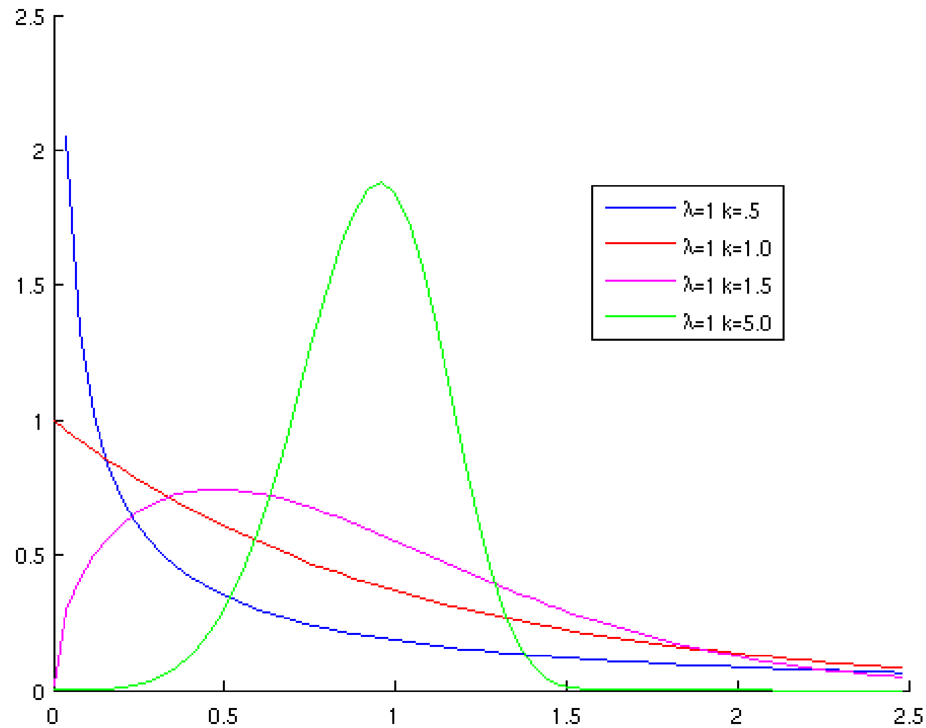


Figure 1-13 Weibull family curves showing the variability that comes from different shape parameters 'k'. The shape parameter k=1.0 produces an Exponential probability distribution. [Philip Leitch, Wikipedia]

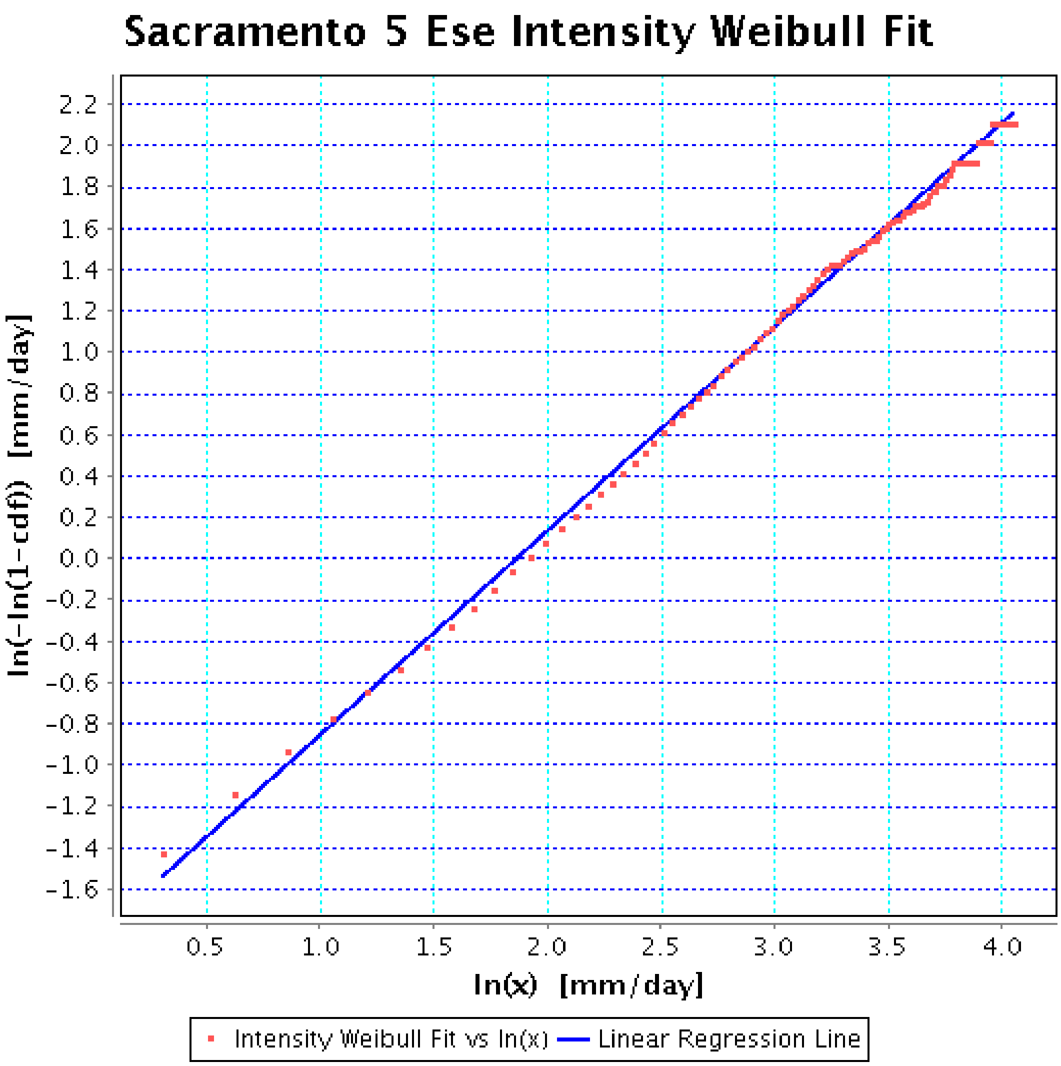


Figure 1-14 Weibull Plot of a Weibull fit for the Event Intensity for Sacramento precipitation. This fit is a cross-plot of the ln(-ln(1-cdf(x))) versus ln(x) where x is the Intensity values. The calculated Least Absolute Deviation Coefficient of Determination for this regression fit is 95.5%. The fitted line's slope gives a Weibull shape parameter 'k' of 0.988 and the intercept yields a scale parameter 's' of 6.75.

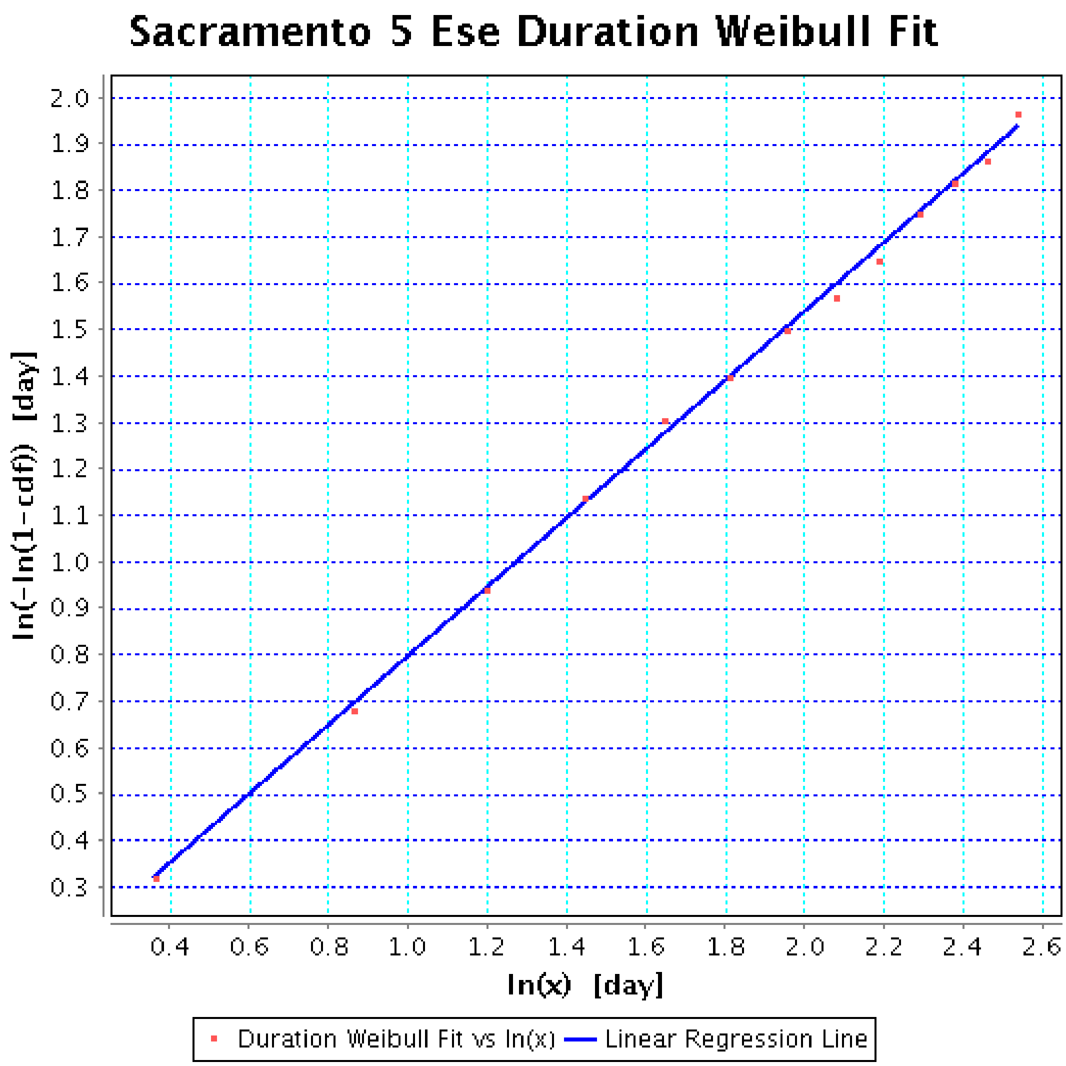


Figure 1-15 Weibull Plot of a Weibull fit for the Event Duration for Sacramento precipitation. This fit is a cross-plot of the ln(-ln(1-cdf(x))) versus ln(x) where x is the Duration values. The calculated Least Absolute Deviation Coefficient of Determination for this regression fit is 96.5%. The fitted line's slope gives a Weibull shape parameter 'k' of 0.742 and the intercept yields a scale parameter 's' of 1.137.

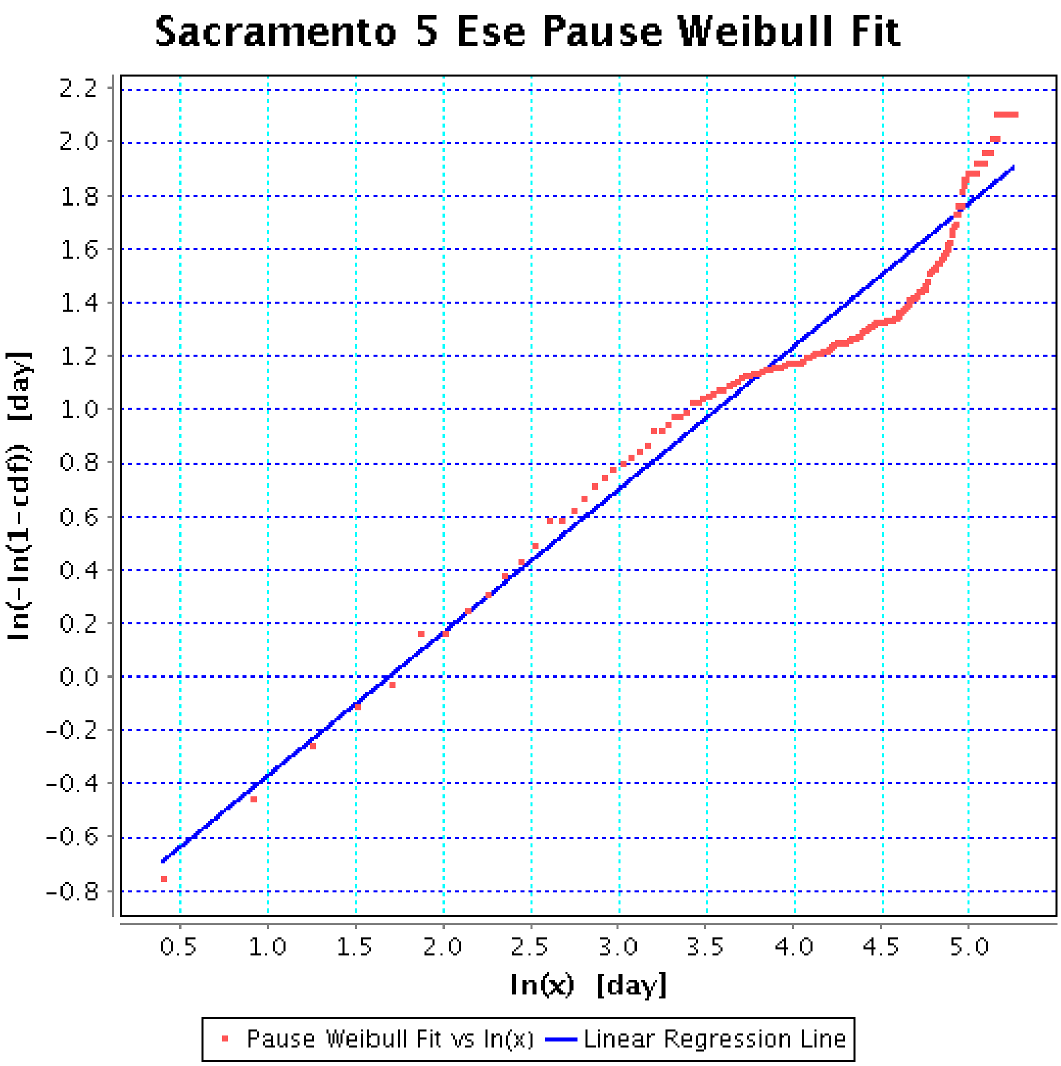


Figure 1-16 Weibull Plot of a Weibull fit for the Lull Pause for Sacramento precipitation. This fit is a cross-plot of the ln(-ln(1-cdf(x))) versus ln(x) where x is the Pause values. The calculated Least Absolute Deviation Coefficient of Determination for this regression fit is 69.2%. The fitted line's slope gives a Weibull shape parameter 'k' of 0.535 and the intercept yields a scale parameter 's' of 5.746.

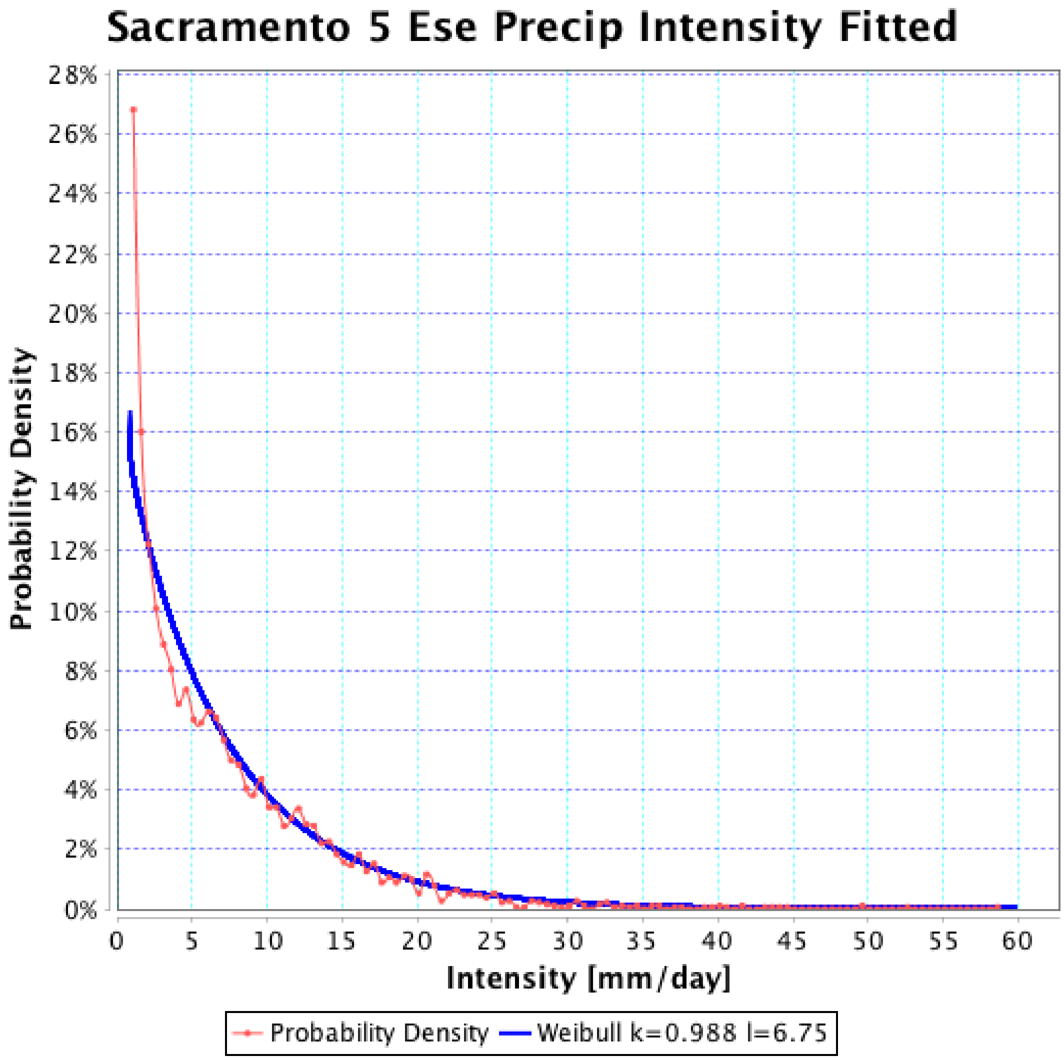


Figure 1-17 The empirical probability density for the Event Intensity of the Sacramento station precipitation shown with its Weibull distribution fit. This Weibull distribution fit has a shape parameter 'k' of 0.988 and a scale parameter 's' of 6.75.

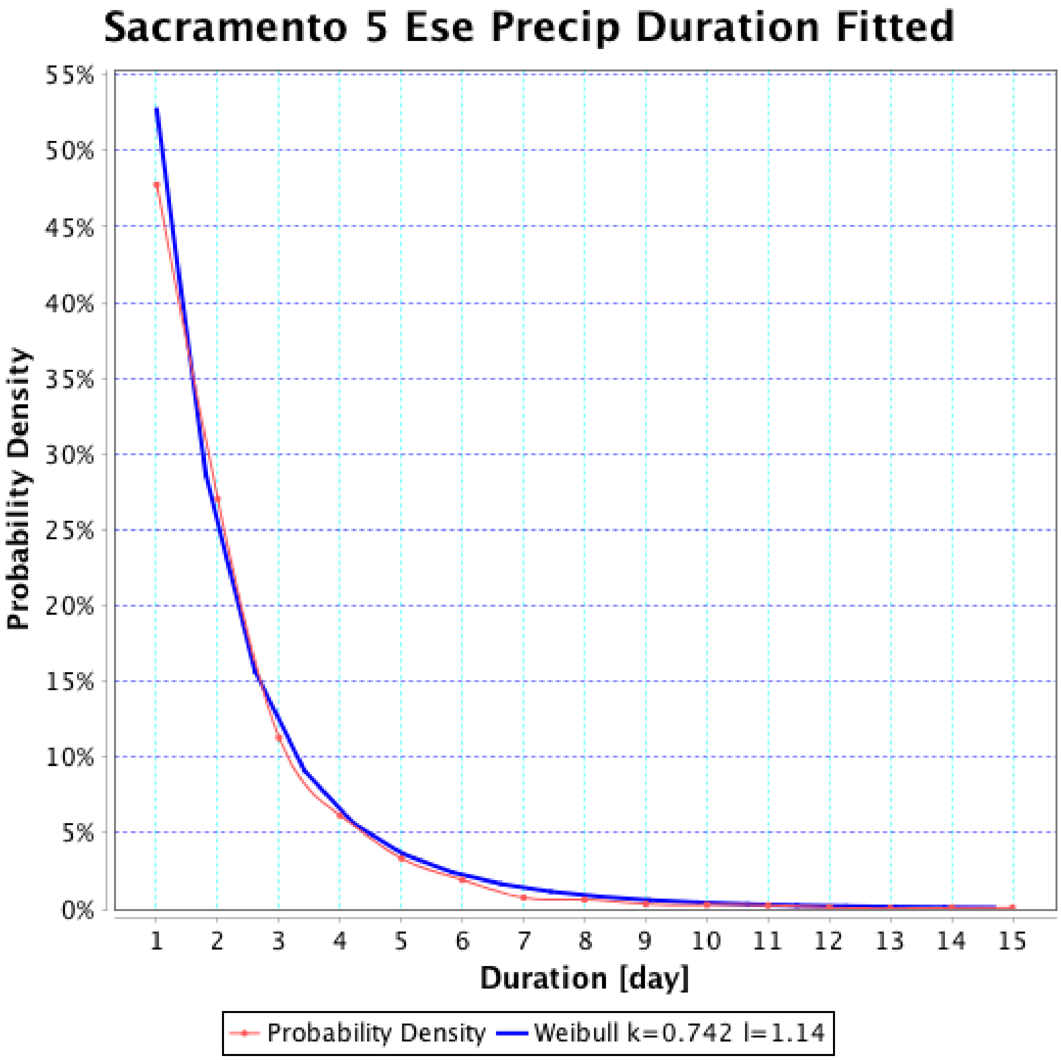


Figure 1-18 The empirical probability density for the Event Duration of the Sacramento station precipitation shown with its Weibull distribution fit. This Weibull distribution fit has a shape parameter 'k' of 0.742 and a scale parameter 's' of 1.14.

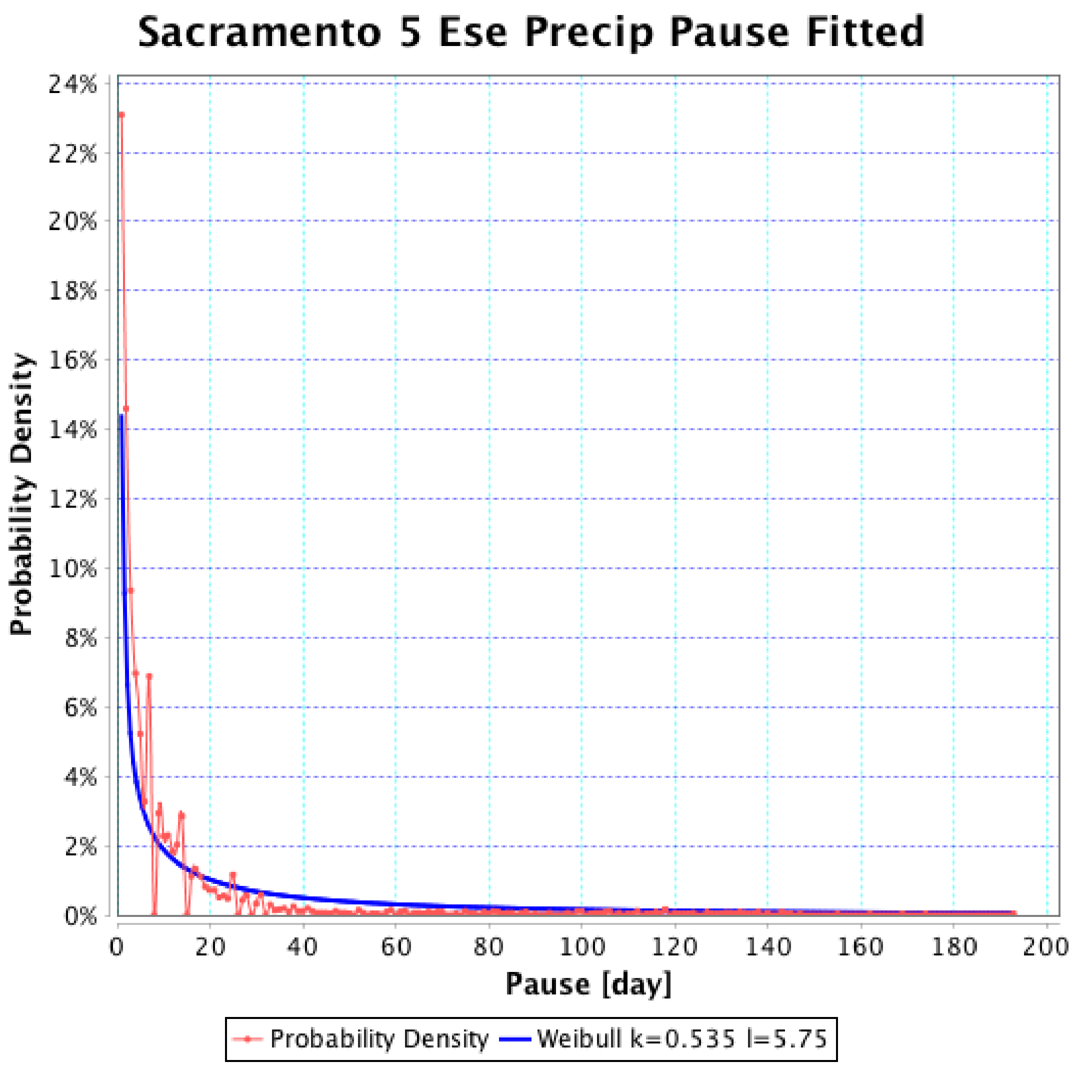


Figure 1-19 The empirical probability density for the Lull Pause of the Sacramento station precipitation shown with its Weibull distribution fit. This Weibull distribution fit has a shape parameter 'k' of 0.535 and a scale parameter 's' of 5.75.

Table 1-1 The 50 GHCND long-term, low-missing stations in California

Site Name	GHCND Id	Years	Missing	County
North Coast Drainage				
Graton	USC00043578	89	0.3%	Sonoma
Scotia	USC00048045	89	0.6%	Humboldt
Ukiah	USC00049122	121	3.5%	Mendocino
Orleans	USC00046508	112	4.1%	Humboldt
Santa Rosa	USC00047965	112	6.7%	Sonoma
Sacramento Drainage				
Sacramento 5 Ese	USW00023271	138	0.2%	Sacramento
Stony Gorge Rsvr	USC00048587	89	1.4%	Glenn
Nevada City	USC00046136	122	2.7%	Nevada
De Sabla	USC00042402	109	2.9%	Butte
Colgate Ph	USC00041916	109	3.2%	Yuba
Nicolaus	USC00046193	104	3.7%	Sutter
East Park Rsvr	USC00042640	94	4.0%	Colusa
Volta Ph	USC00049390	89	4.1%	Shasta
Canyon Dam	USC00041497	101	4.4%	Plumas
Chico Univ Farm	USC00041715	109	4.8%	Butte
Red Bluff Muni Ap	USW00024216	123	5.3%	Tehama
Orland	USC00046506	112	5.7%	Glenn
Willows 6 W	USC00049699	109	5.7%	Glenn
Hat Creek	USC00043824	94	5.9%	Shasta
Chester	USC00041700	105	6.1%	Plumas
Grass Valley	USC00043571	122	6.6%	Nevada
Northeast Interior Basins				
Tahoe City	USC00048758	112	6.0%	Placer
Doyle	USC00042504	92	6.1%	Lassen
Central Coast Drainage				
San Francisco Dwtn	USW00023272	94	0.0%	San Francisco
Livermore	USC00044997	112	1.2%	Alameda
San Jose	USC00047821	115	2.4%	Santa Clara
Priest Valley	USC00047150	109	4.8%	Monterey
Big Sur Stn	USC00040790	100	6.2%	Monterey

Site Name	GHCND Id	Years	Missing	County
San Joaquin Drainage				
Madera	USC00045233	87	2.7%	Madera
Elliott	USC00042760	89	2.8%	San Joaquin
Ash Mtn	USC00040343	88	3.5%	Tulare
Camp Pardee	USC00041428	89	3.6%	Calaveras
Calaveras Big Trees	USC00041277	86	3.8%	Calaveras
Tiger Creek Ph	USC00048928	93	4.2%	Amador
Lindsay	USC00044957	102	4.4%	Tulare
Porterville	USC00047077	103	5.3%	Tulare
Hetch Hetchy	USC00043939	105	6.3%	Tuolumne
Auberry 2 Nw	USC00040379	100	6.7%	Fresno
South Coast Drainage				
Chula Vista	USC00041758	97	1.4%	San Diego
Laguna Beach	USC00044647	87	2.3%	Orange
Newport Beach Harbor	USC00046175	94	3.5%	Orange
Redlands	USC00047306	117	4.2%	San Bernardino
San Gabriel Canyon Ph	USC00047776	98	4.2%	Los Angeles
San Bernardino F S 22(	USC00047723	112	4.2%	San Bernardino
Ojai	USC00046399	110	6.1%	Ventura
Tustin Irvine Rch	USC00049087	113	6.9%	Orange
Southeast Desert Basin				
Haiwee	USC00043710	92	3.3%	Inyo
Greenland Rch	USC00043603	104	4.2%	Inyo
Trona	USC00049035	95	4.4%	San Bernardino

The 50 GHCN-Daily stations in California that contain at least 85 years of precipitation data record and with no more than 7% of its daily records missing. These are grouped by their climatic regions. These 50 stations are those used for the investigations in this study.

Chapter Two

THE SIGNIFICANCE OF TRENDS IN PRECIPITATION METRICS IN THE OBSERVATIONAL RECORD OF CALIFORNIA FROM 1877 TO 2014

Abstract

This work applies statistical techniques to quantify confidence in interpreted trends in long-term precipitation event Intensity, event Duration, and lull Pause for 50 stations in California. The permutation technique is used to generate a trend's empirical distribution, which provides exact inference of its significance p-values. This Monte Carlo process is repeated 50,000 times to ensure p-value accuracy to 0.09%.

This permutation technique has been enhanced to allow statistical consideration of the effect of missing data for calculation of these p-values for trends in precipitation metrics. Within the Monte Carlo process, a step is added to repeatedly fill-in missing data with Multiple Imputation (MI) estimates using Weibull probability density distributions of likely values. For the Intensity metric, 11 of the 50 stations had trends 95% significant. The Duration trends had 16 stations 95% significant. Pause trends had 19 stations 95% significant of which 12 were 99% significant.

Sets of trend results are summarized through use of p-value reciprocal weighing. The state-wide summary for the Intensity trend is -4.61% per decade with a p-value of 0.16% and shows a west-east increase. The Duration trend of 3.49% per decade has a p-value of 0.03% and shows both a strong west-east and a north-south increase. The Pause trend of 3.58% per decade has a p-value of 0.09% but no directional tendencies. The study shows that for California as a whole storms are getting more gentle, lasting longer, and perhaps getting separated longer in time.

2.1 Introduction

The goal of this chapter is to resolve trends of precipitation metrics where some of the data is missing, including an assessment of confidence. Even when data gaps are non-missing at random, it has been shown that a likely value can still be calculated for a trend statistic. This calculation also produces bounds for the range of likely values. But the question here is: can it be determined if such a likely value is statistically significant and not just the result of random occurrence?

Trend values are sometimes desired for practical resource management. These values become important when resource decisions are based on these results. These decisions may entail considerable investments in time, effort, and money. So it is critical to verify if such trends are significant.

Determining significance for trends is particularly challenging for this research because it uses observation records of precipitation in California. Precipitation in California is highly variable in space and time. One study (Dettinger et al., 2011) assessed the values of the coefficient of variation (ratio of the standard deviation over the mean) for inter-annual precipitation across the contiguous U.S. The coefficient of variability in the eastern half of the country generally ranges from 0.1 to 0.2 and most of the western half from 0.2 to 0.3. The majority of values in California are from 0.3 to 0.5 and a few might be as high 0.7 (**Figure 2-1**). As a specific California example, the coefficient of variation for the Sacramento station is 0.34 (**Figure 2-2**).

This inter-annual variability of precipitation in California is often larger than the inter-annual change implied by precipitation trends. This makes it challenging to detect significant trends against this background of strong "noise" in the climate system. However if statistical methods can produce significant inferences in California, then they could be more applicable elsewhere.

2.2 Data and Metrics

The precipitation data described in the first chapter is used again in this chapter. The same 50 Global Historical Climatology Network – Daily (GHCND) stations in California are used; all have at least 85 years of daily records and no more than 7% missing data. Statistics of three precipitation metrics are derived from the records of those stations. The Intensity metric is the mean precipitation for an Event (a sequence of "wet" days with precipitation ≥1 mm/d). The Duration metric is the length in days of such an Event. The Pause metric is the length in days of a Lull (a sequence of "dry" days with precipitation <1 mm/d). The statistic of interest for these three metrics is their long-term trend, as defined by the slope of a straight line of the metric versus time.

2.2.1 Climate Divisions

Climate Divisions have been established by the National Climatic Data Center (NCDC) part of the federal National Oceanic and Atmospheric Administration (NOAA). They were defined for the contiguous United States to allow a collection of long-term temporally and spatially complete data to generate regional historical climate analyses (Guttman and Quayle, 1996). There are 344 climate divisions within the contiguous United States. The Global Historical Climatology Network (GHCN) stations in the United States are grouped by the state in which they are located. Within each state, the stations are further grouped by NCDC into from 1 to 10 individual state Climate Divisions. The stations in California are organized into these seven NCDC Climate Divisions:

1) North Coast Drainage
2) Sacramento Drainage
3) Northeast Interior Basins
4) Central Coast Drainage
5) San Joaquin Drainage
6) South Coast Drainage
7) Southeast Desert Basin

as shown on the NCDC map (**Figure 1-3**).

Information about which GHCN stations are within a NCDC Climate Division can be found by cross-correlating the GHCN station identifier with the station data listing of the stations of the Cooperative Observer Network (COOP). This station list is available from the NCDC Historical Observing Metadata Repository (HOMR) file found at:

<http://www.ncdc.noaa.gov/homr/file/coop-stations.txt>

This COOP listing also provides some additional station information such as the name of county in the state where it is located and its time zone.

This work uses these NCDC Climate Division assignments for the 50 California stations selected for this study. This allows investigation of whether and how the precipitation trend statistics of these stations might vary across the climate regions of the state.

2.2.2 Software

The ClimateData software system created by this author and described in the first chapter is used again in this chapter. That software base has been augmented for this chapter with an additional 2-3,000 lines of Java code also created by this author. The new code adds the capability to perform the statistical significance computations and produce the plots described in the Methods section.

2.3 Methods

The approach applied in this chapter is to determine linear trends in designated precipitation metrics, and calculate the p-values to assess the significance of the calculated trends. This section also describes the methods to summarize sets of p-values in order to compare them across climate regions.

2.3.1 Trend Significance

A powerful method for the analysis of statistical significance is called permutation testing, sometimes called exact or randomization tests (Ernst, 2004; Feinstein, 1993; Good, 2005). The permutation test method has the advantage that it is completely non-parametric, so it does not depend on the data being normally distributed. Since non-normality is clearly the case for the precipitation metrics chosen for this research, these permutation tests are suitable (Welch, 1990).

Permutation testing also offers the great benefit that it is applicable for any test statistic and not just useable for the usual statistics mean, standard deviation, least squares parameter fitting, etc. It works even if the data's probability distribution is unknown. Permutation tests accomplish this advantage by generating the metric's empirical distribution by Monte Carlo sampling the population of data permutations (Smyth and Phipson, 2010; Ojala and Garriga, 2010).

For determination of the significance of trends, this permutation method first computes a reference trend value **t** from the original observed, non-permuted sequence of annual metric values (Berry et al., 2011). This trend computation is carried out just as described in the previous chapter. Missing data for each month are filled in by random sampling from a specific monthly Weibull probability density distribution. After this filling procedure, the trend is calculated using the resulting filled annual values. These two operations are repeated 10,000 times as a Monte Carlo process from which the desired trend value **t** is taken as the median of all these estimates.

For the second permutation step, the annual data values are randomly permuted or shuffled and a new permuted trend value $\mathbf{pt_1}$ is computed. Within a second Monte Carlo process, such permutation is repeated again and again to generate permuted trend values $\mathbf{pt_2}$, $\mathbf{pt_3}$, … $\mathbf{pt_m}$. This research uses a permutation loop repetition count **m** of 50,000 times. The generated trend sequence $\mathbf{pt_1}$, $\mathbf{pt_2}$, $\mathbf{pt_3}$, … $\mathbf{pt_m}$ thereby constitutes a probability density distribution for the trend metric of interest (Gandy, 2009).

Finally, the significance of the actual trend $\mathbf{t_0}$ is derived by performing a count **B** of all the permuted trend values $\mathbf{pt_i}$ at least as extreme as the reference trend value **t**. This count **B** is a measure of how far the reference trend value **t** is found out on the tail of the permutation generated probability distribution. The p-value for significance is given by count **B** divided by the loop number **m** (Berger, 2000).

The repetition count **m** also determines the error bounds for the computed significance p-value. The permutation Monte Carlo process is probabilistic and has a standard deviation value of

$$\sqrt[2]{\frac{p(1-p)}{m}}$$

where **p** is the p-value of the significance being computed (Johnston et al., 2007; Gandy, 2009). Some extra certainty for these error bounds suggests allowing for 2 of these standard deviations. For example, for 50,000 repetition count for **m** and testing for 99% significance (a p-value of 0.01), the error bounds or precision of these significance values is 0.0009 or ±0.09%.

There is an additional innovation that accounts for the effects of missing data. Each time around the Monte Carlo permutation loop, Multiple Imputation is performed to again fill in missing data for each monthly metric value. This is the same MI method that is carried out in the fill-in phase of the **t** trend estimation method. In other words, within each of the 50,000 permutation repetitions, the monthly missing data is recomputed and filled with different metric values. These new monthly metric values come from random draws from the specific Weibull probability distribution for that month.

2.3.2 Significance Weighting

Measures of statistical significance are useful when dealing with sets of data values. The simple approach is to treat each value as equally important. This means producing a summary measure of central tendency such as using computed statistical mean, median, or mode.

However, if there are significance measures for the various values in the set, then some of these values are less likely to be random. Accommodation can insure that certain values are more likely to be correct. A method to perform this trust accommodation is to give different weights to each value, with stronger weights to those that are more significant (Zhang and Nguyen, 2005; Ghazanfar and Prugel-Bennett, 2010).

There are many formulas that accomplish such significance weighting of data values. However, it has been shown that for the advantages of weighting, “power is remarkably robust to misspecification of these weights” (Roeder and Wasserman, 2009). This work uses the reciprocal of the p-value, which is perhaps the simplest weighting formula. So a value that is more significant will have a p-value closer to zero yielding more weight for the corresponding data value.

In addition, it is possible to give different weights to the sets of p-values themselves. This gives an indication of the p-value or significance of such a weighted set of data values (Alves and Yu, 2011). Just as done with the data value trend weighting, each p-value is weighted by its own reciprocal.

2.4 Results

2.4.1 Trend Significance

For each of the 50 selected GHCND stations, metric trend values and the corresponding significance p-value were calculated. The 2-sigma error bounds (95% containment) was calculated for each computed p-value. These calculations were repeated for each of the three precipitation metrics Intensity, Duration, and Pause.

For the significance of the precipitation Event Intensity metric (Table 2-1), there were 5 stations at the traditional cut-off level of 95% significance with p-values <= 5%. An additional 6 stations were at the higher 99% very significant level with p-values <= 1%. This puts 11 of the 50 station set (22%) supplying results at least 95% significant. Also there were 4 stations that just missed being termed significant by having their p-values >5% but < 6%.

For the precipitation Event Duration metric (Table 2-2), there were 10 stations at the level of 95% significance with p-values <= 5%. An additional 6 stations were at the 99% very significant level with p-values <= 1%. So there are 16 of the 50 stations (32%) supplying results at least 95% significant.

For the precipitation Pause metric (Table 2-3), there were 7 stations at the level of 95% significance. A total of 12 stations were at the 99% very significant level. Of the 50 stations, 19 (38%) were at least 95% significant.

As a particular example, the Santa Cruz station has an event Intensity trend of -1.98% per decade. Its significance p-value of 0.53% is determined from its position on the left hand tail of the permutation probability density curve, where only 266 out of the 50,000 values have values less than or equal (Figure 2-4). The computed error bounds for that p-value are ±0.03%. The Santa Cruz station event Duration trend of 0.60% lies toward the right side of its permutation curve where 12,567 out of the 50,000 values are greater than or equal. This gives a p-value of 25.13% with bounds of ±0.19% (Figure 2-5). The Santa Cruz Pause trend 1.69% has 336 permutation values greater than or equal, giving a p-value of 0.67% with bounds of ±0.04% (Figure 2-6).

2.4.2 Significance Weighting & Climate Divisions

For each of the three metrics, the set of 50 California stations were grouped into the 7 defined climate divisions and then the trend results and their significance p-values were summarized for each division. These climate divisions were further grouped and then summarized to assess if any consistent trend patterns occur across the state in the west-to-east direction or along the north-to-south direction (Figure 2-3).

These summaries were performed in two ways. First, the station results were combined using the standard median measure of central tendency. Second, the previously defined method of weighting values by the reciprocal of their p-value was applied. Both the trends and particularly the significance p-values produced by the weighting method are noticeably stronger than their median values.

2.5 Discussion

2.5.1 Trend Significance

The probability density distributions for Monte Carlo trend permutations appear quasi-symmetric. This follows from the fact that for every permutation order that yields some slope **X,** there exists the reversed order permutation, which must have the opposite slope **-X**. Therefore, if these plots were totally complete, i.e. included all possible permutations, then they would be precisely symmetric. Since a hundred year precipitation record will have factorial (100) or about 10^{158} possible permutations, such plots can never be complete. However, with a large number of permutations, such as the 50,000 performed here, there is a probabilistic tendency for the count around some trend value above zero to be balanced by a similar count for the corresponding below zero trend value.

There is another probabilistic factor that could explain the lack of precise symmetry: the filling in of missing data. Random values are drawn from the Weibull probability distribution chosen for the metric of interest and at the year-month time of that missing data. Therefore, the values being permuted also have their values changed somewhat upon each iteration.

Both of these random processes mean that the results for trend values and their p-values are not absolutely fixed or certain. Every time that these calculations are performed, the results will be slightly different. However, estimates of the spread of these variable results can be produced. The two-sigma spread of the calculated p-values is shown for the trend significances shown in (Tables 2-1, 2-2, and 2-3) in the Error column. For example, the Santa Cruz Intensity metric trend has a p-value of 0.53% with an error bound of ±0.03%. The Santa Cruz Duration metric has a p-value of 37.98% with an error of ±0.22% and its Pause p-value is 0.67% with error ±0.04%.

The variability of these statistical results is determined by the iteration count of the relevant Monte Carlo process. The iteration counts have been selected so that the spread of these results are relatively modest and, in particular, are much smaller than the result values themselves. If tighter bounds of variability are desired, then the iteration counts can be increased. At the current count of 50,000 iterations for p-value Monte Carlo, the computation time is quite acceptable. On the author's small Apple Mac Mini mid-2012 computer, the calculation for a metric of its trends and their permutation significance p-values for all 50 stations takes 5 ½ minutes, which is less than 7 seconds each.

2.5.2 Significance Weighting & Climate Divisions

It is worthwhile to consider the choice of weighting for the p-value of sets of results. The use of the reciprocal formula for weighting p-values provides an exponential weighting of trend constituents. So as the trend p-value approaches 0% its weight grows rapidly. Of course, there are many other possible weight formulas. For example, rather than the reciprocal of the p-value, the reciprocals of the p-value's square root or its cube root do not grow in value as quickly (Figure 2-7).

2.5.2.1 Intensity

For the Climate Divisions, precipitation event Intensity trends of the median measure of central tendency provide no statistically significant results. The p-values range from 13% to 31%. The trends themselves are mixed with two positive values and the other five negative and most magnitudes under 1% per decade. There is no state apparent directional tendency.

The weighted Intensity trends offer some significance with two p-values less than 5% and two less than 1%. The trends are all larger but still mixed. However, these weighted trends show a distinct west to east increase, but with decreasing magnitude trends of -5.0%, -2.4%, and -1.7% per decade. It might be expected that Intensity would naturally strengthen from orographic lifting effects as storms head inland. This lifting might be expected to still strengthen about as much the storms that are weaker arriving at the coast.

2.5.2.2 Duration

For the precipitation event Duration metric trends, the median measure provides just one result of any significance with p-value 1.5%. The Division Duration trends are almost all positive (one is at -0.01%). So it seems reasonable to conclude that there is a real state trend in Duration, i.e. storms are lasting longer, with a state-wide trend of 0.5% and p-value of 18.6%. There is no state directional tendency for these medians.

The weighted Duration trends again offer good significance with three p-values less than 5% and two less than 1%. All weighted Division trends are positive except for one anomaly. For Climate Division 6 South Coast Drainage, the mean trend is a positive 0.14% per decade while the weighted trend is -3.34%, i.e. a reversal and also different than other weighted Duration trends. The two stations with the best p-values are Ojai with trend of -3.3% and p-value of 0.00% and Chula Vista with -2.6% and p-value 0.42%. So the advantage of weighting is shown where the negative trends of those two highly significant stations overcome the weak positive trends of the other non-significant stations. However, the weighted state-wide Duration trend is a strong positive value of 3.5% per decade with a p-value of 0.03%.

The weighted trends show a very strong west to east increase of -3.3%, 2.5%, and 6.9% per decade. All these also have strong significance p-values of 0.04%, 2.43%, and 0.00%, respectively. Just as orographic lifting intensifies storms, it would be expected that a concentrated and shorter storm front at the coast could get dispersed and lengthened as it moves inland.

There is also an increasing Duration trend from north to south of 1.2%, 2.7%, and 3.5% and their significance also increases in this direction from a p-value 5.7% north to 0.01% south. So this tendency could just be an artifact of that increasing significance. The one climate change effect suspected in that southmost area is a poleward expansion of the tropical belt (Seidel et al., 2008; Lu et al., 2007). This movement could engender a north to south tendency. An encroaching descending branch of the Hadley cell might extinguish smaller precipitation events and thus increase the Duration.

2.5.2.3 Pause

For the precipitation lull Pause metric trends, the median measure provides just one result of any significance with p-value 2.5%. The Division Duration trends are mixed with four positive and three negative. The median state-wide trend in Duration is a weak value of 0.1% per decade and a p-value of 12.3%. There is no state directional tendency for these median trends.

The weighted Pause trends are quite significant with three p-values less than 5% and three less than 1%. The individual Division weighted trends are just as mixed as they were in the median trend case. However, the state-wide weighted Pause trend is a 3.6% per decade value with a weighted p-value of 0.1%. There is no directional tendency for these weighted trends.

2.5.3 Attributions

The majority of the infrequent mentions in literature of precipitation event Duration and event Intensity definition are when they are lumped together as Intensity-Duration-Frequency (IDF), used in determining the return period of storms. Duration and Intensity are rarely measured or analyzed separately and there is only study we have seen that purports to measure their trends (Palecki et al., 2005). This study used a definition of storm as any 15-minute interval with 2.54 mm of precipitation, which is completely different and considerably more limiting and extreme than our definition of any day with 1 mm of precipitation. So there seems no effective way to compare our results to any other known studies.

We can speculate on reasons for the consistent increase in event Duration and the decrease of Intensity. Both changes would seem to follow if storm events were moving slower. Duration would clearly be longer since a slow moving storm would take longer to pass through a location. The import of new moisture needed to feed the storm precipitation could also be reduced by a slower storm velocity.

There have been recent work on a concept called Arctic amplification where the Arctic is observed to warm faster than the mid-latitudes thereby reducing the temperature gradient (Francis and Vavrus, 2012; Cohen et al., 2014). It is claimed that this gradient change has caused a 14% decrease in upper-level winds leading to a slowing of west to east movement of large-scale Rossby waves. Also implicated is an increase of atmospheric blocking events. All these effects could be connected with such a slower movement of storms.

Someone might argue that our measured decrease in precipitation event intensity over the last 100+ years is contrary to many published studies of the anticipated influence of climate change on the hydrologic cycle, which suggest that greater intensity is to be expected.

Most such theoretical and observational studies have based such estimates on country, continental, or world-wide scales (Alexander et al., 2006; Barnett et al., 2006; IPCC, 2013; Min et al., 2011; Toreti et al., 2013; U.S. GCRP, 2014; Karl and Knight, 1998; Groisman et al., 1999, 2004, 2005). The differences between the Southwest region and the rest of the United States is illustrated by large storm intensity measures (**Figure 2-8).** The differences between California and the rest of continental United States is also seen in **Figure 2-1**. The statistics of this study are only for California precipitation.

Both larger scale studies and those that strictly study California (Pierce et al., 2013; Russo et al., 2013) typically investigate the statistics of just the very largest of storms. Study of the full range of storm events shows that the largest storms can increase, while the small and even moderate events could decrease **Figure 2-8**. An increase in the very largest storms implies little about the average trends of all storms, such as performed in this study.

Almost all these other studies are quantifying complete storms. These are measured as the total precipitation of a storm. In contrast, this study is perhaps unique in that it bifurcates that total storm metric in order to research separately the length of the storm event (Duration metric) and its daily average (Intensity metric). There can be confusion because the commonly used total storm metric is often called intensity too. The Intensity metric of this study is measuring something different than the other use of the word.

2.6 Conclusions

The permutation Monte Carlo method was used to identify significant trends in precipitation metrics. It demonstrates the possibility of extracting significant p-value results in the California region, despite its highly variable climate. With inclusion of Multiple Imputation to repeatedly estimate and fill-in for missing data values, it can also properly account for effects of modest amounts of missing data.

The provision of data significance p-values allows all data to not have to be treated the same. Instead, greater weight can be given to those results with demonstrated higher significance. This significance weighing can provide a summary value for a set of samples giving more credence to those contributors that are more significant. This can provide a significance p-value to the summary value itself.

References

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., et al., 2006, Global observed changes in daily climate extremes of temperature and precipitation: Journal of Geophysical Research: Atmospheres, v. 111, p. n/a–n/a, doi: 10.1029/2005JD006290.

Alves, G., and Yu, Y.-K., 2011, Combining Independent, Weighted P-Values: Achieving Computational Stability by a Systematic Expansion with Controllable Accuracy: PLoS ONE, v. 6, p. e22647, doi: 10.1371/journal.pone.0022647.

Barnett, D.N., Brown, S.J., Murphy, J.M., Sexton, D.M.H., and Webb, M.J., 2006, Quantifying uncertainty in changes in extreme event frequency in response to doubled CO2 using a large ensemble of GCM simulations: Climate Dynamics, v. 26, p. 489–511, doi: 10.1007/s00382-005-0097-1.

Berger, V.W., 2000, Pros and cons of permutation tests in clinical trials: Statistics in Medicine, v. 19, p. 1319–1328, doi: 10.1002/(SICI)1097-0258(20000530)19:10<1319::AID-SIM490>3.3.CO;2-S.

Berry, K.J., Johnston, J.E., and Mielke, P.W., 2011, Analysis of Trend: A Permutation Alternative to the F Test: Perceptual and Motor Skills, v. 112, p. 247–257, doi: 10.2466/03.PMS.112.1.247-257.

Cohen, J., Screen, J.A., Furtado, J.C., Barlow, M., Whittleston, D., Coumou, D., Francis, J., Dethloff, K., Entekhabi, D., Overland, J., and Jones, J., 2014, Recent Arctic amplification and extreme mid-latitude weather: Nature Geoscience, v. 7, p. 627–637, doi: 10.1038/ngeo2234.

Dettinger, M.D., Ralph, F.M., Das, T., Neiman, P.J., and Cayan, D.R., 2011, Atmospheric Rivers, Floods and the Water Resources of California: Water, v. 3, p. 445–478, doi: 10.3390/w3020445.

Ernst, M.D., 2004, Permutation Methods: A Basis for Exact Inference: Statistical Science, v. 19, p. 676–685.

Feinstein, A.R., 1993, Permutation tests and “statistical significance”: M. D. Computing: computers in medical practice, v. 10, p. 28–41.

Francis, J.A., and Vavrus, S.J., 2012, Evidence linking Arctic amplification to extreme weather in mid-latitudes: Geophysical Research Letters, v. 39, p. L06801, doi: 10.1029/2012GL051000.

Gandy, A., 2009, Sequential Implementation of Monte Carlo Tests With Uniformly Bounded Resampling Risk: Journal of the American Statistical Association, v. 104, p. 1504–1511, doi: 10.1198/jasa.2009.tm08368.

Ghazanfar, M., and Prugel-Bennett, A., 2010, Novel Significance Weighting Schemes for Collaborative Filtering: Generating Improved Recommendations in Sparse Environments.

Good, P., 2005, Permutation, Parametric and Bootstrap Tests of Hypotheses: Springer.

Groisman, P.Y., Karl, T.R., Easterling, D.R., and Knight, R.W., 1999, Changes in the Probability of Heavy Precipitation: Important Indicators of Climatic Change: Climatic Change, v. 42.

Groisman, P.Y., Knight, R.W., Easterling, D.R., Karl, T.R., Hegerl, G.C., and Razuvaev, V.N., 2005, Trends in Intense Precipitation in the Climate Record:.

Groisman, P.Y., Knight, R.W., Karl, T.R., Easterling, D.R., Sun, B., and Lawrimore, J.H., 2004, Contemporary Changes of the Hydrological Cycle over the Contiguous United States: Trends Derived from In Situ Observations: Journal of Hydrometeorology, v. 5, p. 64–85.

Guttman, N.B., and Quayle, R.G., 1996, A Historical Perspective of U.S. Climate Divisions: Bulletin of the American Meteorological Society, v. 77, p. 293–303, doi: 10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2.

IPCC, 2013, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change: Cambridge University Press, 1535 p.

Johnston, J.E., Berry, K.J., and Mielke, P.W., 2007, Permutation tests: precision in estimating probability values: Perceptual and Motor Skills, v. 105, p. 915–920, doi: 10.2466/pms.105.3.915-920.

Karl, T.R., and Knight, R.W., 1998, Secular trends of precipitation amount, frequency, and intensity in the United States.: Bulletin of the American Meteorological Society, v. 79, p. 231.

Lu, J., Vecchi, G.A., and Reichler, T., 2007, Expansion of the Hadley cell under global warming: Geophysical Research Letters, v. 34, p. L06805, doi: 10.1029/2006GL028443.

Min, S.-K., Zhang, X., Zwiers, F.W., and Hegerl, G.C., 2011, Human contribution to more-intense precipitation extremes: Nature, v. 470, p. 378–381, doi: 10.1038/nature09763.

Ojala, M., and Garriga, G.C., 2010, Permutation Tests for Studying Classifier Performance: J. Mach. Learn. Res., v. 11, p. 1833–1863.

Palecki, M.A., Angel, J.R., and Hollinger, S.E., 2005, Storm Precipitation in the United States. Part I: Meteorological Characteristics: Journal of Applied Meteorology, v. 44, p. 933–946, doi: 10.1175/JAM2243.1.

Pierce, D.W., Cayan, D.R., Das, T., Maurer, E.P., Miller, N.L., Bao, Y., Kanamitsu, M., Yoshimura, K., Snyder, M.A., Sloan, L.C., Franco, G., and Tyree, M., 2013, The Key Role of Heavy Precipitation Events in Climate Model Disagreements of Future Annual Precipitation Changes in California: Journal of Climate, v. 26, p. 5879–5896, doi: 10.1175/JCLI-D-12-00766.1.

Roeder, K., and Wasserman, L., 2009, Genome-Wide Significance Levels and Weighted Hypothesis Testing: Statistical science : a review journal of the Institute of Mathematical Statistics, v. 24, p. 398–413, doi: 10.1214/09-STS289.

Russo, T.A., Fisher, A.T., and Winslow, D.M., 2013, Regional and local increases in storm intensity in the San Francisco Bay Area, USA, between 1890 and 2010: Journal of Geophysical Research: Atmospheres, p. 1–10, doi: 10.1002/jgrd.50225.

Seidel, D.J., Fu, Q., Randel, W.J., and Reichler, T.J., 2008, Widening of the tropical belt in a changing climate: Nature Geoscience, v. 1, p. 21–24, doi: 10.1038/ngeo.2007.38.

Smyth, G.K., and Phipson, B., 2010, Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn : Statistical Applications in Genetics and Molecular Biology:.

Toreti, A., Naveau, P., Zampieri, M., Schindler, A., Scoccimarro, E., Xoplaki, E., Dijkstra, H.A., Gualdi, S., and Luterbacher, J., 2013, Projections of global changes in precipitation extremes from Coupled Model Intercomparison Project Phase 5 models: Geophysical Research Letters, v. 40, p. 4887–4892, doi: 10.1002/grl.50940.

U.S. GCRP, 2014, Climate Change Impacts in the United States: The Third National Climate Assessment: U.S. Global Change Research Program 3.

U.S. GCRP, 2009, Global Climate Change Impacts in the United States: U.S. Global Change Research Program.

Welch, W.J., 1990, Construction of Permutation Tests: Journal of the American Statistical Association, v. 85, p. 693–698, doi: 10.2307/2290004.

Zhang, J., and Nguyen, T.N., 2005, A New Term Significance Weighting Approach: Journal of Intelligent Information Systems, v. 24, p. 61–85, doi: 10.1007/s10844-005-0267-y.

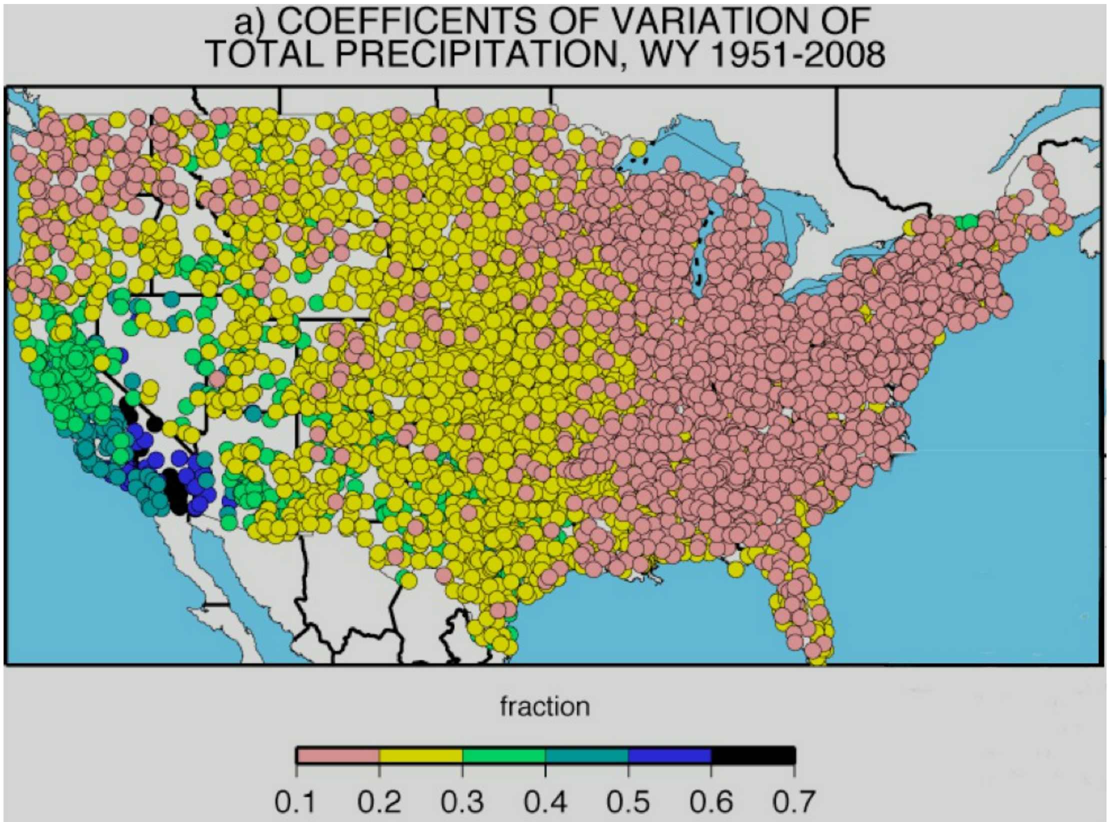


Figure 2-1 Illustration showing how large and exceptional is the inter-annual coefficient of variation of precipitation in California compared to the rest of the contiguous United States. The coefficient of variation is defined as the standard deviation of inter-annual water-year precipitation divided by its mean. This diagram is from (Dettinger et al., 2011).

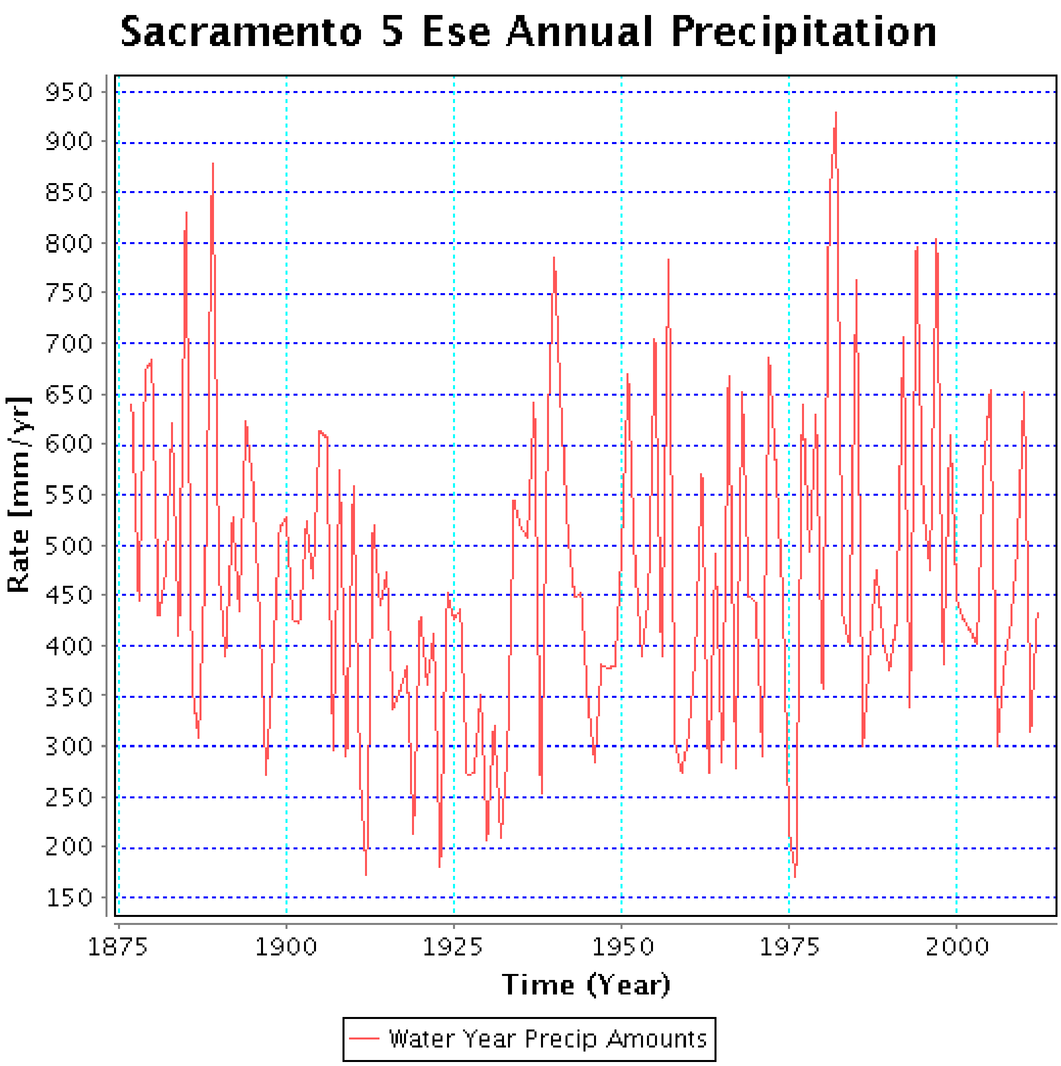


Figure 2-2 Chart of the annual water year precipitation in Sacramento, CA. This location's inter-annual precipitation standard deviation of 159 mm divided by its annual mean of 464 mm yields a coefficient of variation value of 34.3%. This large value illustrates how California inter-annual precipitation is significantly more variable than elsewhere in the contiguous United States.

Figure 2-3 Map of California with its County boundaries. This map is overlain with red lines and numbers, which define the boundaries of the 7 GHCN defined Climatic Divisions. These are #1 North Coast Drainage, #2 Sacramento Drainage, #3 Northeast Interior Basins, #4 Central Coast Drainage, #5 San Joaquin Drainage, #6 South Coast Drainage, #7 Southeast Desert Basin.

Divisions #1, #4, and #6 in this work are termed the Western coastal region. Divisions #2 and #5 are the Central Valley. Divisions #3 and #7 form the Eastern boundary with Nevada and Arizona. Divisions #1-3 are termed North, #4 and #5 are Middle, and #6 and #7 are South.

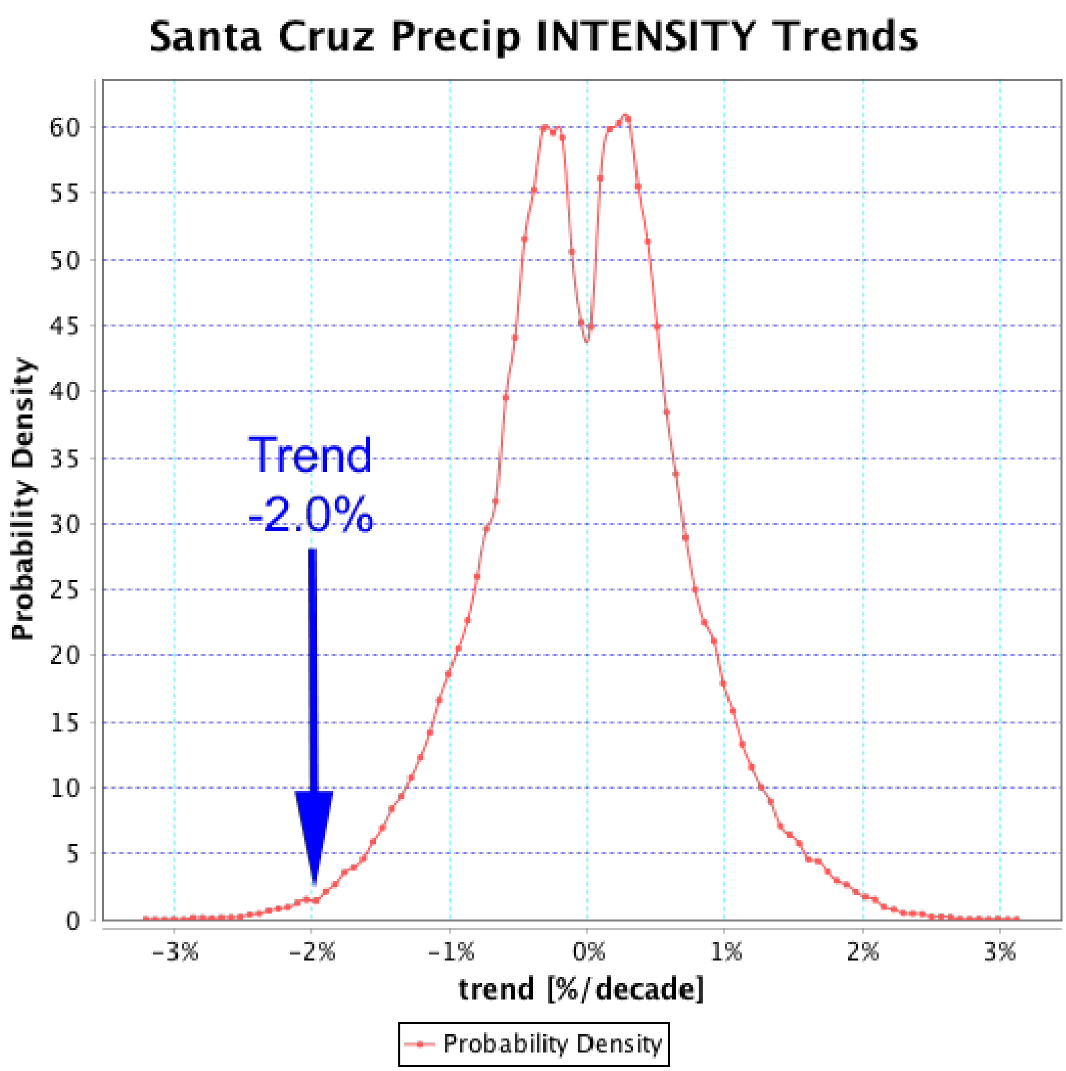


Figure 2-4 Probability density plot for Santa Cruz Intensity trend permutations. Shown as a blue arrow is where the original, non-permuted trend of about -2.0% per decade lies on this distribution's tail. At this location there are 266 of the 50,000 values that are equal or less than this trend value. This position yields a p-value of 266/50,000 or 0.53%, which represents 99.47% significance.

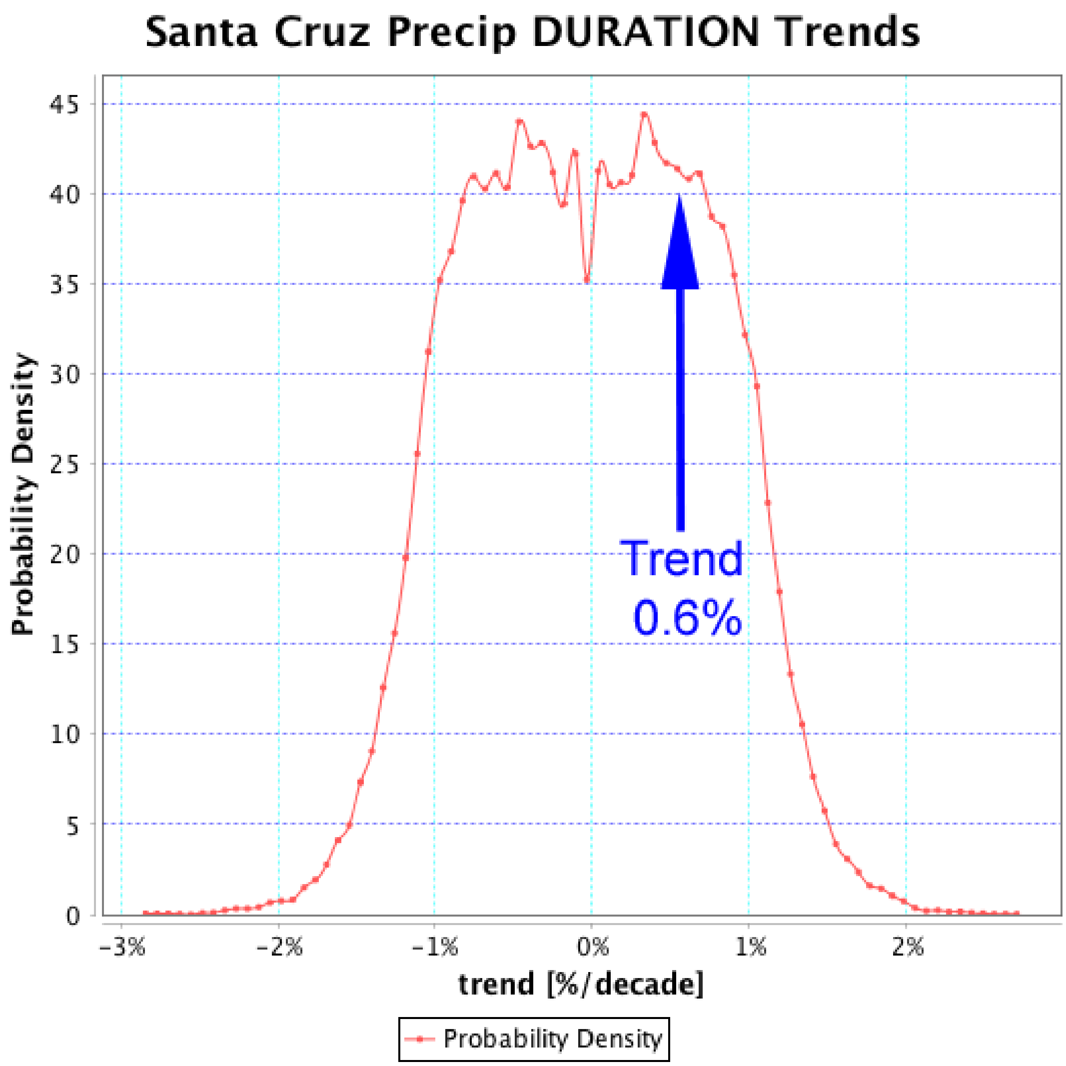


Figure 2-5 Probability density plot for Santa Cruz Duration trend permutations. Shown as a blue arrow is where the original, non-permuted trend of about 0.6% per decade lies on this distribution. At this location there are 18,992 of the 50,000 values that are equal or greater than this trend value. This position yields a p-value of 18,992/50,000 or 25.13%, which represents 74.87% significance.

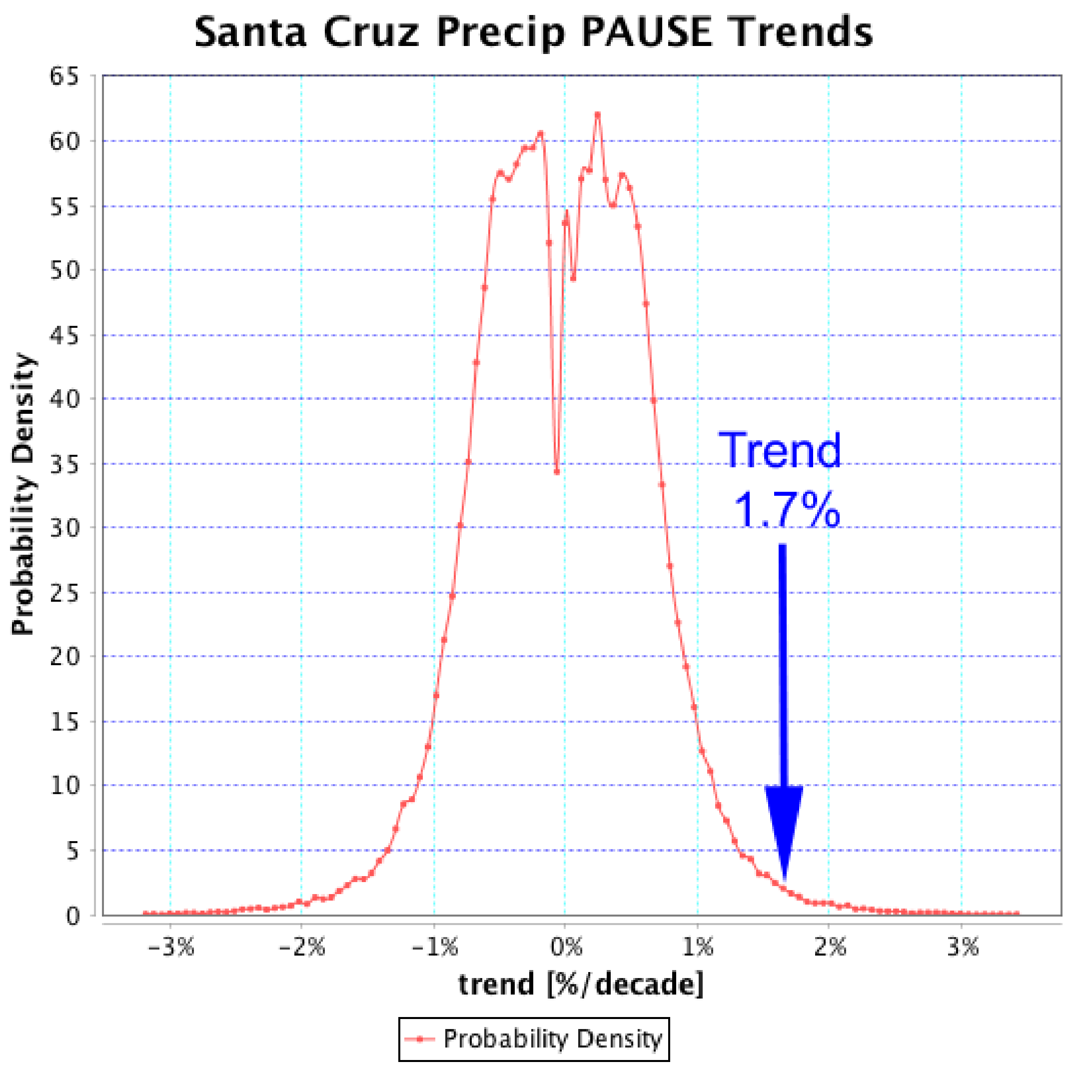


Figure 2-6 Probability density plot for Santa Cruz Pause trend permutations. Shown as a blue arrow is where the original, non-permuted trend of about 1.7% per decade lies on this distribution's tail. At this location there are 336 of the 50,000 values that are equal or greater than this trend value. This position yields a p-value of 336/50,000 or 0.67%, which represents 99.33% significance.

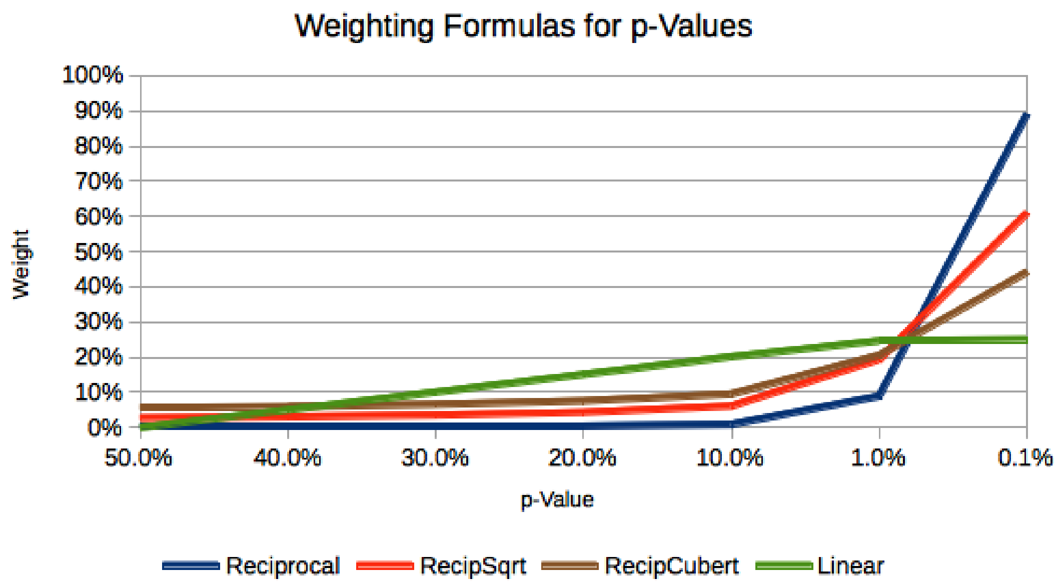


Figure 2-7 Formula options for giving relative weights to p-values. In addition to the direct Reciprocal of the p-value, shown is 'RecipSqrt' the reciprocal of its square root or 'RecipCubert' of its cube root. Also shown is the simple 'Linear' weighting by the formula

$$0.5 - \mathbf{p}$$

where **p** is the p-value. In all cases, the weight increases as the p-value goes from 50% or no significance towards 0% or very, very significant.

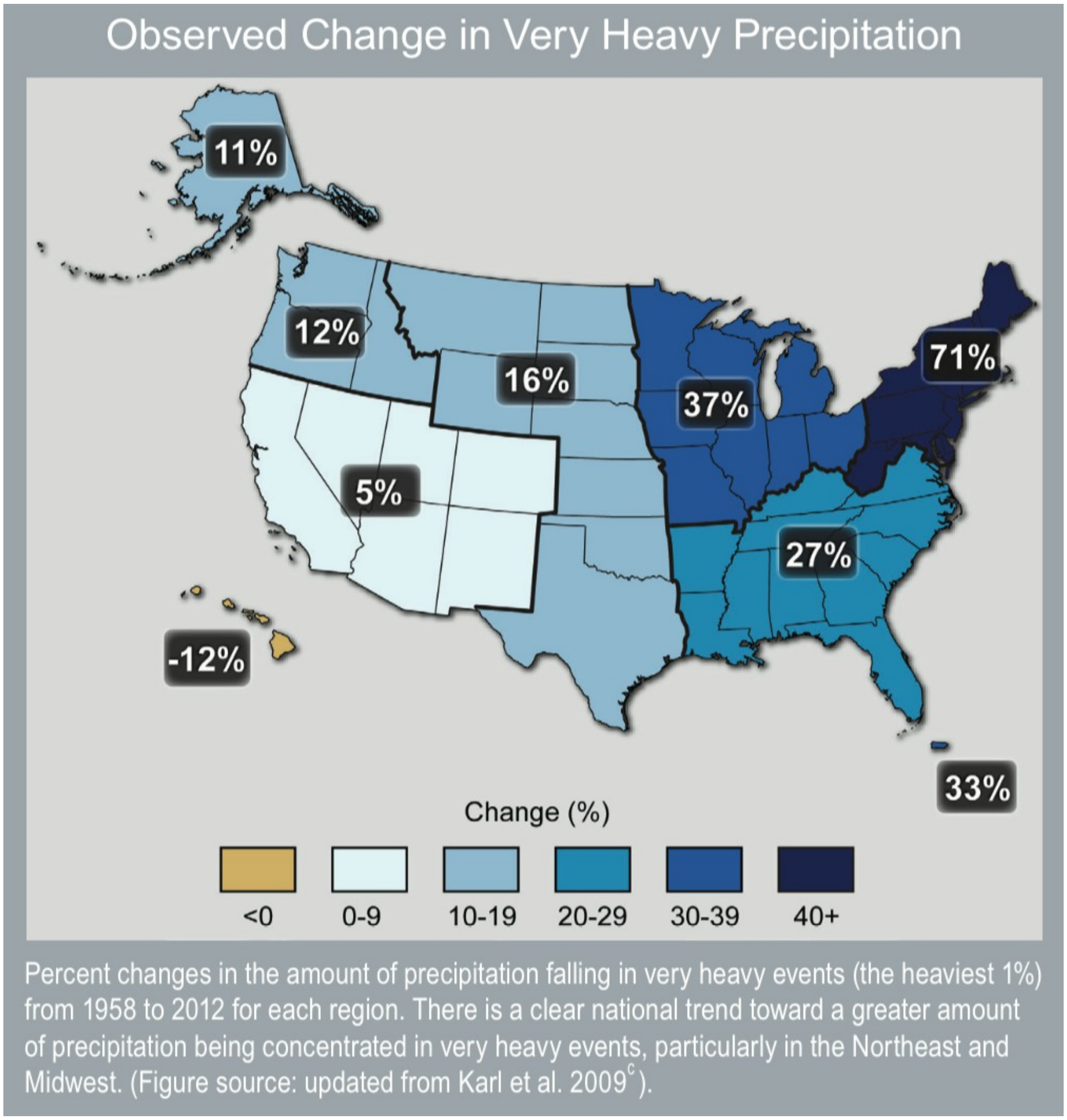
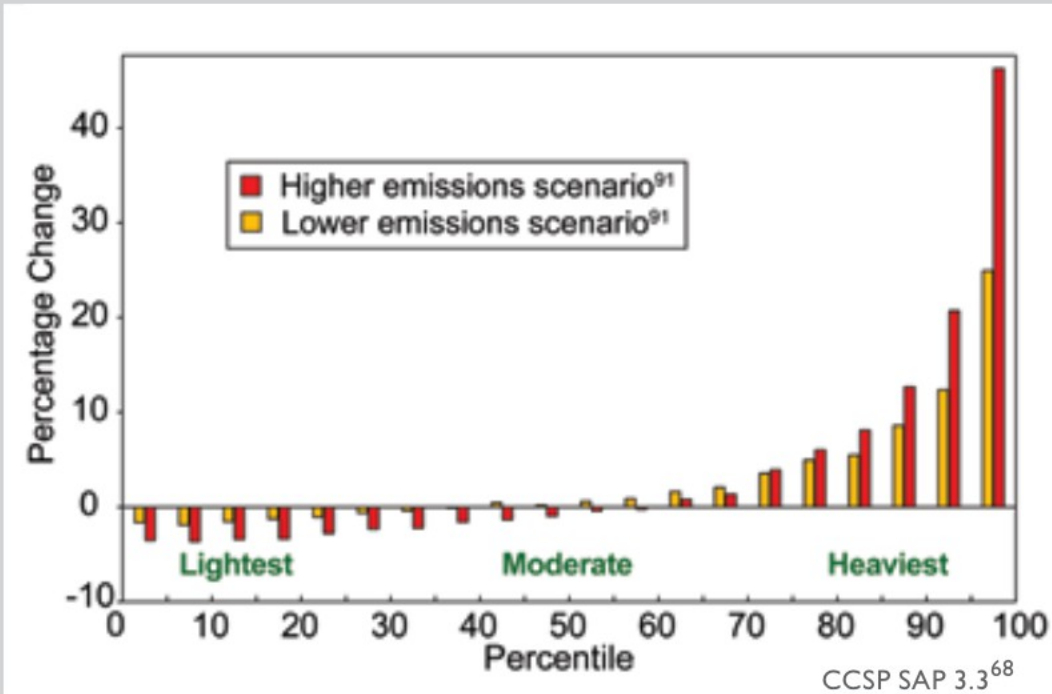


Figure 2-8 This is a map that shows the considerable disparity between the Southwest region and the rest of the United States of the trends for the heaviest 1% of storms (U.S. GCRP, 2014).

Projected Changes in Light, Moderate, and Heavy Precipitation (by 2090s)



The figure shows projected changes from the 1990s average to the 2090s average in the amount of precipitation falling in light, moderate, and heavy events in North America. Projected changes are displayed in 5 percent increments from the lightest drizzles to the heaviest downpours. As shown here, the lightest precipitation is projected to decrease, while the heaviest will increase, continuing the observed trend. The higher emission scenario[91] yields larger changes. Projections are based on the models used in the IPCC 2007 Fourth Assessment Report.

Figure 2-9 This is a graph of the distribution across relative storm size of their projected future change (U.S. GCRP, 2009).

Table 2-1 Intensity trend significance

Station	ClimDiv	Missing (%)	Trend (%/decade)	p-value	Error
Graton	1	0.3	-0.11%	47.94%	0.22%
Scotia	1	0.6	-1.58%	**2.11%**	0.06%
Ukiah	1	3.5	0.42%	29.67%	0.20%
Orleans	1	4.1	0.56%	21.81%	0.18%
Santa Rosa	1	6.7	-0.92%	15.60%	0.16%
Sacramento 5 Ese	2	0.2	0.54%	25.59%	0.20%
Stony Gorge Rsvr	2	1.4	-0.62%	29.94%	0.20%
Nevada City	2	2.7	0.53%	24.84%	0.19%
De Sabla	2	2.9	-1.14%	8.95%	0.13%
Colgate Ph	2	3.2	0.98%	7.44%	0.12%
Nicolaus	2	3.7	0.59%	26.63%	0.20%
East Park Rsvr	2	4.0	2.88%	5.32%	0.10%
Volta Ph	2	4.1	2.39%	**1.38%**	0.05%
Canyon Dam	2	4.4	0.36%	39.59%	0.22%
Chico Univ Farm	2	4.8	-3.17%	**0.03%**	0.01%
Red Bluff Muni Ap	2	5.3	-0.63%	22.57%	0.19%
Orland	2	5.7	-0.19%	42.11%	0.22%
Willows 6 W	2	5.7	0.32%	31.39%	0.21%
Hat Creek	2	5.9	0.63%	28.06%	0.20%
Chester	2	6.1	-1.18%	13.12%	0.15%
Grass Valley	2	6.6	-1.18%	5.69%	0.10%
Tahoe City	3	6.0	-0.45%	30.07%	0.21%
Doyle	3	6.1	1.16%	19.44%	0.18%
San Francisco Dwtn	4	0.0	-0.69%	23.00%	0.19%
Livermore	4	1.2	-0.19%	37.44%	0.22%
San Jose	4	2.4	0.30%	38.69%	0.22%
Priest Valley	4	4.8	-1.87%	**1.91%**	0.06%
Big Sur Stn	4	6.2	-1.11%	15.90%	0.16%
Santa Cruz	4	7.0	-1.98%	**0.53%**	0.03%

Station	ClimDiv	Missing	Trend	p-value	Error
Madera	5	2.7	2.52%	5.47%	0.10%
Elliott	5	2.8	3.33%	**0.27%**	0.02%
Ash Mtn	5	3.5	-0.48%	32.79%	0.21%
Camp Pardee	5	3.6	0.91%	18.23%	0.17%
Calaveras Big Trees	5	3.8	-2.35%	5.83%	0.10%
Tiger Creek Ph	5	4.2	-0.61%	22.61%	0.19%
Lindsay	5	4.4	-0.83%	23.59%	0.19%
Porterville	5	5.3	-2.52%	**0.99%**	0.04%
Hetch Hetchy	5	6.3	0.75%	24.23%	0.19%
Auberry 2 Nw	5	6.7	-2.47%	**2.17%**	0.07%
Chula Vista	6	1.4	-0.03%	48.72%	0.22%
Laguna Beach	6	2.3	-0.15%	45.05%	0.22%
Newport Beach Harbor	6	3.5	-2.10%	**4.00%**	0.09%
Redlands	6	4.2	-2.90%	**0.16%**	0.02%
San Gabriel Canyon	6	4.2	-0.74%	28.43%	0.20%
San Bernardino F S	6	4.2	-1.55%	6.80%	0.11%
Ojai	6	6.1	-5.08%	**0.00%**	0.00%
Tustin Irvine Rch	6	6.9	0.86%	18.65%	0.17%
Haiwee	7	3.3	-3.12%	5.36%	0.10%
Greenland Rch	7	4.2	-0.12%	47.64%	0.22%
Trona	7	4.4	-0.45%	39.26%	0.22%

The long-term observational trends of precipitation event Intensity. The 5 p-values shown in bold are significant at the 95% level (p-value less than 5%). The 6 shown in red are at the further 99% significance level (p-value less than 1%). See **Figure 2-3** for definitions of ClimDiv, the GHCN Climatic Divisions.

Table 2-2 Duration trend significance

Station	ClimDiv	Missing (%)	Trend (%/decade)	p-value	Error
Graton	1	0.3	-0.32%	**2.65%**	0.07%
Scotia	1	0.6	1.98%	23.23%	0.19%
Ukiah	1	3.5	0.49%	24.21%	0.19%
Orleans	1	4.1	0.71%	**0.79%**	0.04%
Santa Rosa	1	6.7	2.12%	35.11%	0.21%
Sacramento 5 Ese	2	0.2	0.25%	**1.82%**	0.06%
Stony Gorge Rsvr	2	1.4	2.19%	9.81%	0.13%
Nevada City	2	2.7	0.70%	17.15%	0.17%
De Sabla	2	2.9	0.75%	40.35%	0.22%
Colgate Ph	2	3.2	-0.20%	30.62%	0.21%
Nicolaus	2	3.7	0.55%	24.74%	0.19%
East Park Rsvr	2	4	-0.85%	**2.66%**	0.07%
Volta Ph	2	4.1	-2.02%	39.63%	0.22%
Canyon Dam	2	4.4	-0.20%	**4.29%**	0.09%
Chico Univ Farm	2	4.8	1.26%	42.38%	0.22%
Red Bluff Muni Ap	2	5.3	-0.14%	**3.23%**	0.08%
Orland	2	5.7	1.18%	5.20%	0.10%
Willows 6 W	2	5.7	1.46%	25.81%	0.20%
Hat Creek	2	5.9	-0.54%	38.23%	0.22%
Chester	2	6.1	-0.25%	35.10%	0.21%
Grass Valley	2	6.6	0.30%	8.63%	0.13%
Tahoe City	3	6	0.67%	22.66%	0.19%
Doyle	3	6.1	-0.70%	15.38%	0.16%
San Francisco Dwtn	4	0	0.62%	39.09%	0.22%
Livermore	4	1.2	-0.19%	43.74%	0.22%
San Jose	4	2.4	0.07%	**3.32%**	0.08%
Priest Valley	4	4.8	1.43%	**0.45%**	0.03%
Big Sur Stn	4	6.2	2.43%	25.13%	0.19%
Santa Cruz	4	7	0.60%	37.98%	0.22%

Station	ClimDiv	Missing	Trend	p-value	Error
Madera	5	2.7	-0.49%	20.12%	0.18%
Elliott	5	2.8	-1.06%	22.44%	0.19%
Ash Mtn	5	3.5	0.51%	13.82%	0.15%
Camp Pardee	5	3.6	1.04%	**0.13%**	0.02%
Calaveras Big Trees	5	3.8	3.27%	46.28%	0.22%
Tiger Creek Ph	5	4.2	0.08%	**3.37%**	0.08%
Lindsay	5	4.4	-1.13%	**1.88%**	0.06%
Porterville	5	5.3	1.66%	34.76%	0.21%
Hetch Hetchy	5	6.3	0.28%	16.28%	0.17%
Auberry 2 Nw	5	6.7	0.67%	**0.42%**	0.03%
Chula Vista	6	1.4	-2.58%	34.36%	0.21%
Laguna Beach	6	2.3	0.49%	9.11%	0.13%
Newport Beach Harbor	6	3.5	-0.96%	25.68%	0.20%
Redlands	6	4.2	0.64%	22.81%	0.19%
San Gabriel Canyon	6	4.2	0.96%	39.21%	0.22%
San Bernardino F S	6	4.2	-0.22%	**1.70%**	0.06%
Ojai	6	6.1	1.47%	**0.00%**	0.00%
Tustin Irvine Rch	6	6.9	-3.35%	**0.00%**	0.00%
Haiwee	7	3.3	5.48%	**0.00%**	0.00%
Greenland Rch	7	4.2	8.40%	**4.42%**	0.09%
Trona	7	4.4	-3.18%	103.18%	0.09%

The long-term observational trends of precipitation event Duration. The 10 p-values shown in bold are significant at the 95% level (p-value less than 5%). The 7 in shown red are at the further 99% significance level (p-value less than 1%). See **Figure 2-3** for definitions of ClimDiv, the GHCN Climatic Divisions.

Table 2-3 Pause trend significance

Station	ClimDiv	Missing (%)	Trend (%/decade)	p-value	Error
Graton	1	0.3	2.54%	**0.88%**	0.04%
Scotia	1	0.6	-3.76%	**2.20%**	0.07%
Ukiah	1	3.5	-0.10%	44.29%	0.22%
Orleans	1	4.1	0.18%	43.42%	0.22%
Santa Rosa	1	6.7	0.35%	36.07%	0.21%
Sacramento 5 Ese	2	0.2	0.31%	25.61%	0.20%
Stony Gorge Rsvr	2	1.4	-1.23%	18.51%	0.17%
Nevada City	2	2.7	1.64%	**0.33%**	0.03%
De Sabla	2	2.9	-0.11%	45.91%	0.22%
Colgate Ph	2	3.2	0.09%	44.80%	0.22%
Nicolaus	2	3.7	-1.83%	**1.29%**	0.05%
East Park Rsvr	2	4	2.29%	**1.50%**	0.05%
Volta Ph	2	4.1	-1.87%	**2.99%**	0.08%
Canyon Dam	2	4.4	2.21%	**1.09%**	0.05%
Chico Univ Farm	2	4.8	-2.19%	**0.15%**	0.02%
Red Bluff Muni Ap	2	5.3	-0.09%	41.52%	0.22%
Orland	2	5.7	0.30%	28.36%	0.20%
Willows 6 W	2	5.7	0.67%	30.33%	0.21%
Hat Creek	2	5.9	1.45%	8.33%	0.12%
Chester	2	6.1	0.83%	22.35%	0.19%
Grass Valley	2	6.6	0.70%	15.80%	0.16%
Tahoe City	3	6	-0.71%	21.83%	0.18%
Doyle	3	6.1	-0.94%	29.58%	0.20%
San Francisco Dwtn	4	0	-1.37%	8.79%	0.13%
Livermore	4	1.2	0.00%	50.00%	0.22%
San Jose	4	2.4	0.82%	6.45%	0.11%
Priest Valley	4	4.8	-0.45%	18.62%	0.17%
Big Sur Stn	4	6.2	-0.19%	42.96%	0.22%
Santa Cruz	4	7	1.69%	**0.67%**	0.04%

Station	ClimDiv	Missing	Trend	p-value	Error
Madera	5	2.7	-0.65%	26.78%	0.20%
Elliott	5	2.8	-0.14%	45.00%	0.22%
Ash Mtn	5	3.5	-0.56%	28.62%	0.20%
Camp Pardee	5	3.6	-1.18%	7.63%	0.12%
Calaveras Big Trees	5	3.8	-2.61%	**0.37%**	0.03%
Tiger Creek Ph	5	4.2	0.76%	23.00%	0.19%
Lindsay	5	4.4	3.91%	**0.00%**	0.00%
Porterville	5	5.3	0.70%	16.13%	0.16%
Hetch Hetchy	5	6.3	1.37%	5.60%	0.10%
Auberry 2 Nw	5	6.7	0.30%	35.86%	0.21%
Chula Vista	6	1.4	1.45%	**3.41%**	0.08%
Laguna Beach	6	2.3	-2.70%	**0.97%**	0.04%
Newport Beach Harbor	6	3.5	2.04%	**3.97%**	0.09%
Redlands	6	4.2	-0.07%	46.31%	0.22%
San Gabriel Canyon	6	4.2	-0.11%	42.91%	0.22%
San Bernardino F S	6	4.2	2.42%	**0.35%**	0.03%
Ojai	6	6.1	2.00%	**0.40%**	0.03%
Tustin Irvine Rch	6	6.9	3.31%	**0.03%**	0.01%
Haiwee	7	3.3	-2.47%	**0.96%**	0.04%
Greenland Rch	7	4.2	-2.51%	**0.12%**	0.02%
Trona	7	4.4	2.03%	6.46%	0.11%

The long-term observational trends of precipitation lull Pause. The 7 p-values shown in bold are significant at the 95% level (p-value less than 5%). The 12 shown in red are at the further 99% significance level (p-value less than 1%). See **Figure 2-3** for definitions of ClimDiv, the GHCN Climatic Divisions.

Table 2-4 Intensity trend and p-value summaries

Climate Division	--- Median --- Trend (% per decade)	p-value	--- Weighted --- Trend (% per decade)	p-value
1 North Coast Drainage	-0.11%	23.42%	-1.20%	7.83%
2 Sacramento Drainage	0.34%	19.54%	-2.97%	**0.43%**
3 Northeast Interior Basins	0.36%	24.75%	0.53%	23.61%
4 Central Coast Drainage	-0.90%	19.58%	-1.87%	**2.34%**
5 San Joaquin Drainage	-0.55%	13.62%	1.50%	**1.76%**
6 South Coast Drainage	-1.15%	18.98%	-5.02%	**0.03%**
7 Southeast Desert Basin	-0.45%	30.75%	-2.56%	12.88%
California	-0.45%	20.62%	-4.61%	**0.16%**
West Coast	-0.74%	18.65%	-4.98%	**0.07%**
Central Valley	0.07%	20.40%	-2.38%	**0.60%**
East Border	-0.45%	30.07%	-1.73%	15.74%
Northern	0.32%	22.57%	-2.93%	**0.60%**
Middle	-0.65%	17.07%	0.45%	**1.94%**
Southern	-0.74%	18.65%	-5.02%	**0.04%**

The summaries of precipitation event Intensity trends for GHCND Climate Divisions and state directions. The 3 weighted p-values shown in bold are significant at the 95% level (p-value less than 5%). The 7 shown in red are at the further 99% significance level (p-value less than 1%).

Table 2-5 Duration trend and p-value summaries

Climate Division	--- Median --- Trend (% per decade)	p-value	--- Weighted --- Trend (% per decade)	p-value
1 North Coast Drainage	0.71%	16.05%	1.97%	**2.84%**
2 Sacramento Drainage	0.28%	22.26%	0.70%	7.66%
3 Northeast Interior Basins	-0.01%	15.64%	0.29%	12.50%
4 Central Coast Drainage	0.61%	21.19%	2.20%	**2.26%**
5 San Joaquin Drainage	0.40%	19.71%	2.91%	**1.16%**
6 South Coast Drainage	0.14%	16.66%	-3.34%	**0.02%**
7 Southeast Desert Basin	5.48%	**1.48%**	6.94%	**0.00%**
California	0.50%	18.63%	3.49%	**0.03%**
West Coast	0.60%	22.81%	-3.29%	**0.04%**
Central Valley	0.29%	21.28%	2.48%	**2.43%**
East Border	0.67%	**4.42%**	6.94%	**0.00%**
Northern	0.49%	23.23%	1.24%	5.73%
Middle	0.55%	18.20%	2.74%	**1.42%**
Southern	0.49%	**4.42%**	3.50%	**0.01%**

The summaries of precipitation event Duration trends for GHCND Climate Divisions and state directions. The 3 median and 5 weighted p-values shown in bold are significant at the 95% level (p-value less than 5%). The 6 weighted p-values shown in red are at the further 99% significance level (p-value less than 1%).

Table 2-6 Pause trend and p-value summaries

Climate Division	--- Median --- Trend (% per decade)	p-value	--- Weighted --- Trend (% per decade)	p-value
1 North Coast Drainage	0.18%	25.37%	0.72%	**3.00%**
2 Sacramento Drainage	0.31%	18.05%	-0.62%	**1.25%**
3 Northeast Interior Basins	-0.82%	25.71%	-0.81%	25.12%
4 Central Coast Drainage	-0.10%	21.25%	1.33%	**3.24%**
5 San Joaquin Drainage	0.08%	18.90%	3.87%	**0.02%**
6 South Coast Drainage	1.72%	12.29%	3.02%	**0.18%**
7 Southeast Desert Basin	-2.47%	**2.51%**	-2.43%	**0.31%**
California	0.13%	12.30%	3.58%	**0.09%**
West Coast	0.18%	6.45%	2.88%	**0.39%**
Central Valley	0.30%	17.32%	3.76%	**0.05%**
East Border	-0.94%	6.46%	-2.42%	**0.52%**
North	0.18%	21.83%	-0.47%	**1.58%**
Middle	-0.07%	17.37%	3.86%	**0.03%**
South	1.45%	**0.97%**	2.07%	**0.20%**

The summaries of precipitation lull Pause trends for GHCND Climate Divisions and state directions. The 1 median and 4 weighted p-values shown in bold are significant at the 95% level (p-value less than 5%). The 1 median and 9 weighted p-values shown in red are at the further 99% significance level (p-value less than 1%).

Chapter Three

HYDROLOGIC RESPONSE TO CHANGES IN PRECIPITATION METRICS: TWO BASINS IN CALIFORNIA

Abstract

This study computes the relative impacts to hydrology flow components from the changes of three precipitation metrics: event Intensity, even Duration, and lull Pause. The analysis is performed using a Precipitation-Runoff Modeling System (PRMS) hydrology model for two California locations: the Feather River Basin in the northern Sierra Nevada mountains and the central coast Soquel-Aptos Basin. The three metric trends are computed for each basin by weighting and combining the trends from nearby precipitation station observations. Each metric is changed in the precipitation record for a basin without change to the other metrics or the total precipitation.

Most hydrologic components are shown to behave differently between the two basins primarily because of climate differences. Most hydrologic impacts are modest with magnitudes less than half the corresponding precipitation metric changes. The one exception is the Duration metric change in the Feather River Basin, which was by far the smallest metric change. This might suggest that small precipitation metric changes could have relatively larger hydrologic impacts.

The importance of the Feather River Basin to supply streamflow into Lake Oroville and the State Water Project is benefited from all three precipitation changes by an increase of mean of 0.5%. The value of the Soquel-Aptos Basin for water supply from groundwater recharge is harmed from all three precipitation changes by a decrease of mean of -2.5%. Soquel-Aptos streamflow is similarly harmed by a

decrease of mean of -1.1%. Neither of these Soquel-Aptos impacts seem amenable to mitigation, thus adaptation is indicated.

3.1 Introduction

The goal of this chapter is to compute the quantitive impacts on hydrologic conditions and processes from changes in precipitation metrics. This computation is performed by varying the precipitation metrics that are input to a hydrological model. It is a standard practice in the calibration of parameters of a hydrologic model to compute the sensitivity of model outputs to changes in these parameters. This study is concerned with the sensitivity of model outputs to changes of precipitation inputs.

There are practical values of such a hydrologic impact study. There may be actions which can be taken to reduce or eliminate impacts on water resources. Early notification of impacts can provide the time it might take to investigate, design, permit, finance, and implement facilities or policies to mitigate some or all of the impacts. A hydrologic study can also provide the early warning for actions needed to adapt to impacts, such as acquisition of additional water supplies, flood controls, or conservation.

This study performs a series of tests to determine the differential impact of precipitation metric changes on the hydrological components. Each basin test repeatedly invokes a hydrological model. First, the model is invoked with its base climate data, which is the same historical precipitation record with which the model was calibrated. Second, the model is invoked with that climate input but with one precipitation metric varied. The computed water flows of the model are analyzed to compute the change relative to the base results of each hydrological component. That second model invocation and analysis is repeated for each of the three precipitation metrics: Intensity, Duration, and Pause. The end result is a set of the relative sensitivities of the various hydrology components to change of each precipitation metric.

Tests are run for two watershed and groundwater basin locations to assess how the results vary between these two basins. Numerous physiological and biological characteristics could influence the nature of hydrologic response to changes in precipitation metrics, including topography, vegetation, drainage, or soils. The two basins are located in different climate areas, which provides additional information about the nature of hydrologic response.

3.2 Study Areas

This study investigates climate impacts on the hydrology at two locations in California (**Figure 3-1**). Each of them is located in a different NOAA climate division identified for California and each has a hydrologic model used in this study.

The Feather River Basin is located in the Global Historical Climatology Network Daily (GHCND) California Climate Division #2, the Sacramento Drainage Division. The Feather River basin is described in considerable detail in the report about its hydrologic model (Koczot et al., 2005). But in brief, the basin is located in north-east California on the western slope of the Sierra Nevada foothills. It has an area of about 45,000 square kilometers (3,600 square miles). It ranges from an elevation of (843 ft) at its outlet at Oroville Dam to (9,525 ft) near Mt. Lassen. The basin receives an average of about 114 cm (45 inches) of precipitation ranging from 33 cm (13 inches) in a rain-shadow area of the Sierra Nevada to 317 cm (125 inches) near Mt. Lassen. Almost 60% of the basin lies below its current average snowline. With this placement, temperature can easily change between above and below freezing. Some 68% of precipitation falls as snow with some melting on arrival and others persisting for months.

The Soquel-Aptos Basin is located in the GHCND California Climate Division #4, the Central Coast Drainage. The Soquel-Aptos Basin is described in the report about its hydrologic model (Hydrometrics, 2011). The study area was also detailed in an extensive earlier report (Johnson et al., 2004). This basin has an area of 820 square kilometers (65 square miles) in central California along the north shore of Monterey Bay. It ranges from sea-level up to 963 m (3,160 feet). The basin receives average precipitation of 30 inches (76 cm) near the coast up to 46 inches (117 cm) at higher elevations. There is occasional snow fall at the highest elevations, but no regular snowpack accumulation.

3.3 Materials

3.4 Methods

The primary methods of this chapter are the procedures for changing a precipitation dataset so as to adjust one of three metrics: Intensity, Duration, or Pause. The requirement is to change a metric to accomplish some computed metric trend percentage increase or decrease. The goal is to accomplish this change of a metric while leaving the other metrics and the total precipitation amount unchanged.

3.4.1 Hydrology Model Software

This study uses the U.S. Geological Survey (USGS) software Precipitation-Runoff Modeling System (PRMS). It is a modular, deterministic, distributed-parameter, physical process based system for modeling hydrology (Wenming et al., 2008; Leavesley et al., 1983). It was developed to evaluate the impacts of various combinations of precipitation, climate, and land use on streamflow, sediment yields, water-balance relationships, flow regimes, flood peaks and volumes, soil-water relationships, ground-water recharge, and general basin hydrology. PRMS models most hydrology components of interest from the first contact of precipitation (rain or snow) at the surface down to the water recharged into groundwater (Figure 3-2 and Figure 3-3).

PRMS represents four hydrologic zones (Figure 3-2): Surface, Soil, Ground, and

SurfaceWater. Each zone has its own input sources and output destinations. Often an output of one zone serves as an input into another.

PRMS does not provide detailed groundwater representation or modeling. The Ground zone contains just the one input, Recharge. For output, any excess that is not needed to discharge to streams as base flow can be discarded into a non-specific GroundWater Sink basin output.

For this modeling work, we can track and analyze these following 15 distinct PRMS flows (grouped by their zone):

Surface: Precipitation, Exfiltration (Dunnian), Plant Evaporation, Snow Sublimation, Surface Impervious Evaporation, Surface Runoff, Infiltration, Far Runoff

Soil: Infiltration, Soil ET, Exfiltration (Dunnian), Interflow, Recharge, Far Interflow

Ground: Recharge, Groundwater Sink, Baseflow

Surface Water: Surface Runoff, Interflow, Baseflow, Lake Precipitation, Lake Evaporation, Stream Outflow

Neither of the two models employed in this study make use of the flows Exfiltration (Dunnian), Lake Precip, Lake Evap, Far Runoff, or Far Interflow components. Since these are always zero, they are ignored and not shown.

We are also not including in our analyses the seven water stores (shown as blue ovals on Figure 3-2), but have chosen instead to focus on flows within and between zones. Changes in storage could be assessed by calculating a budget based on aggregate input and output flows.

PRMS has been used to create dozens of hydrologic models. A number of these PRMS models have been used to perform climate change impact studies, almost always by connecting outputs of GCMs to the PRMS model inputs (Bae et al., 2008; Chang and Jung, 2010; Hay et al., 2006; Im et al., 2010; Legesse et al., 2004; Walker et al., 2011). In particular, several such conventional climate investigations with PRMS have involved the Feather River Basin (Hay et al., 2010; Huang et al., 2012; Koczot et al., 2012; Walker et al., 2011). Another study, which utilized a different hydrology Basin Characterization Model (BCM), used two GCMs and included the Soquel-Aptos Basin (Flint and Flint, 2012).

Software was previously created by the author to supply inputs and extract outputs of hydrologic models that use the USGS GSFLOW (coupled Ground water and Surface-water FLOW) software system. This software could automatically incrementally increase or decrease supplied climate input data. The software could run the model with that modified input and then extract the resulting output. By comparing output from models with different input parameters, the relative impacts were computed. This software has been adjusted for use with the PRMS models of this hydrology testing study. It has also been enhanced to independently vary each of the three precipitation metrics: Intensity, Duration, and Pause.

3.4.2 Changing Intensity

The precipitation event Intensity metric is the mean value of the intensity of all the precipitation events. The intensity of a single event is the mean of its daily precipitation amounts, i.e. the sum of those daily precipitation amounts divided by its count of days. As such Intensity is the mean of the set of the means of intensity for all those events. This mean of means is referred to as a grand mean (Everitt, 2010). If the divisors of each of the set of means are different, as is the case with these precipitation events, then the Intensity grand mean can be changed without changing Duration or Pause metrics or total precipitation.

The Duration metric is computed as the total number of event days divided by the number of events. Our goal is to change Intensity, without changing either the total number of event days or the number of events, since that would change the Duration.

It is also not allowed to change the total of daily precipitation, since that would almost certainly produce different hydrologic outputs.

It is allowable to exchange days between events or to take some of the precipitation from one event day and transfer it to another event on a different day. To see that this exchange works, five events of random lengths and values were selected (**Table 3-1**). Repeatedly and randomly exchanging their daily values and searching for the rearrangement with the largest grand mean changes the value from 2.0791 to 2.7297 (**Table 3-2**). This simple exchange example has produced a 31% increase in its Intensity value.

It is currently unknown how to classify those exchanges that increase Intensity from those that decrease Intensity. All that is possible now is to perform an exchange, recompute the Intensity metric, and observe if its value has increased or decreased.

The Intensity change methodology operates by choosing 20 pairs of precipitation events and 20 pairs of days (one from each event). Some random portion of the precipitation from one day is redistributed to the other day. The Intensity is recomputed and if these redistributions moves the metric towards the desired goal, then they are retained, otherwise these 20 redistributions are discarded. This process is repeated until the desired change is obtained

Performing this daily value redistribution process has the great benefit that it does not change the lulls between the events or their time length values. It also does not change the event lengths or the count of events. So both the Pause and the Duration metrics are unchanged by changing the Intensity using this method.

3.4.3 Changing Pause

The precipitation Pause metric is defined as the median value of the time length of all the lulls between precipitation events. Let us say that the Lull metric is 'D' days in length. To decrease the Pause metric, we need to increase the number of lulls that are less than the Pause median D value and decrease the number of those greater than that value. To increase the Pause metric, we decrease the number of lulls less than D and increase the number of those greater than D.

A goal for the change of the Pause metric is to change its value without changing Duration or Intensity. The simplest way to change the Pause metric without affecting the other two metrics is to employ an event shifting technique moving a precipitation event either back or forward in time. When shifting an event backwards in time, the lull immediately preceding the event gets shorter in time length (length A decreases) and the lull following the event gets longer (length B increases). When shifting forwards, the effects are reversed.

In either case it is possible to increase a lull length that is less than the Pause median D into one greater than the median or to change a lull greater than median D into one less than D.

If both A and B are greater than the median value D (called **GG**), then a sufficient shift of the event will change one of the greater than the median D values (a **G**) into a value less than median D. This produces a situation with one lull greater than D and one less than D (called: **GG → GL)**.

In general, we actually need to distinguish between two **GL** cases, one in which the sum of the two lull lengths (A + B) is greater than 2*D+1 (called **GL>2D+1**), and one in which the sum is less than 2*D-1 (called **GL**): **GL<2D-1**. When **GL>2D+1**, a shift can accomplish a transition back to the **GG** case, i.e. **GL>2D+1 → GG**. When **GL<2D-1**, a shift can transition to a situation where both lulls have lengths less than

D (called **LL**): **GL<2D-1 → LL**.

We also need to distinguish between the **LL** case when the sum A plus B is greater than D+1 (called **LL>D+1**), then there can be a transition back to the **GL<2D-1** case or **LL>D+1 → GL<2D-1**. If instead A plus B is less than D+1 (called LL<D+1), then no worthwhile transition is possible.

In all the cases, whenever there is a change in the timing of an event, this causes two changes in the Pause metric, with the gain for **G** or **L** being a loss for the other one. This whole event shifting and transition scheme is illustrated (**Table 3-3**).

To decrease the Pause metric, which is calculated as the median of all lull lengths, we need to increase the number of lulls that are less than the initial Pause value and decrease those greater than that value. This requires performing shifts of precipitation events that meet those cases that start as GG or GL<2D-1 above. To increase the Pause metric, we would perform event shifts that start with the cases GL>2D+1 or LL>D+1.

There is a script of this event shift algorithm at work (**Table 3-4**). The script repeatedly performs a specified series of shifts increase or decrease the Pause metric by a desired decrease percentage. The script checks progress and continues adding shifts as needed to achieve the desired change in the Pause metric, slowing down the magnitude of each shift as it approaches the goal.

One new concept needed for this shift algorithm is the concept of an “interpolated” median. Both lull and precipitation event length metrics have a relatively small integral number of days. So any computed median will almost always be an integer such as 3, 4, 5, etc. With several thousand lulls occurring in a long precipitation record, these small integers will have hundreds of repetitions. A change of the median from a 4 to a 5 is a 25% increase. Therefore it would seem to be difficult to accomplish or measure a 10% change.

An expression has been devised to measure how close or far away a dataset is from a targeted change in the median value. This is motivated by the fact that if the middle of the dataset is exactly between the last 4 value and the first 5, then its normal median is 4.5. Similarly if the middle is between the last 3 and the first 4 then its value is 3.5. So a simple interpolation uses the location of that dataset middle within the long sequence of all the 4 values. If the dataset middle is right at the middle of the 4 values, then the interpolated median is a 4.0 quantity. As the dataset middle gets closer to the boundary with the 3 values, then the median is interpolated to be closer to a 3.5 quantity. On the other edge of the 4 values closer to the 5's, the interpolated median value changes closer to a 4.5 quantity. With this interpolation mechanism, we can precisely measure and accomplish small changes in the median within such heavily repeated integer dataset records.

The obvious advantage of this event shifting approach is that no changes need to made to precipitation events other than some slight shifting of their start times in the record. There are still just as many precipitation events as there were before. Each event still has exactly the same length and the same daily precipitation amounts. So this shifting has no influence on the Duration or Intensity metrics.

3.4.4 Changing Duration

The precipitation event metric Duration is the mean value of the time lengths of all the precipitation events. As such it is the sum of the number of days in all those events divided by the count of events.

Changing the number of total event days is problematical. For example, if a new event day is added to the climate record, this would increase the total precipitation. Changing the total precipitation while adjusting the Duration makes it difficult to tell if hydrologic changes occur because of the change in Duration or because of an increase in total precipitation.

Instead of giving some new precipitation amount to a newly added event day, we could move precipitation from one day to another day, either in the same event or in another event. However, this spreading of precipitation across more days would reduce the Intensity metric.

The only way to change the Duration metric without changing other metrics is to change the count of events. This is done by splitting an existing event into two new daughter events. We subtract events by merging two existing events into a single survivor event. Adding events decreases Duration, and subtracting events increases Duration.

3.4.4.1 Preserving Other Metrics

The effects of these Duration metric event mergers and splits will tend to change the Pause metric, but this effect can be neutralized. After the Duration change, the Pause change process can be invoked to remove the inadvertent change. As already noted, changes to the Pause metric will not affect the Duration metric.

Performing Duration change will also change the Intensity metric, but this effect can be removed as well, without affecting the Duration metric.

3.4.5 Hydrology Testing

For each hydrological model there are at least two nearby long-term climate stations with fairly complete records. For each of these stations, we have calculated the three climate metric trends and their statistical significance P-values. We combine the three model-local sets of station trend results for each metric by weighting them by the reciprocal of their significance p-value. This weighting yields a single trend value for each metric that is used for testing that hydrological model.

Each model is tested by using the observation precipitation dataset with which the model was designed and calibrated. One assumption here is that each model was carefully and professionally prepared and that it is a reasonably good representation of the basin hydrology.

First, the hydrology model is executed with its supplied precipitation dataset. This execution yields as output detailed data on resulting hydrology component flows. These flows are expressed in terms of centimeters of water over the entire model domain averaged over all the years of the climate data extent, i.e. cm per year. For example, for the Feather River Basin the precipitation was 100.41 cm/yr, the recharge was 21.98 cm/yr, and the stream outflow was 41.88 cm/yr. These two datasets, one for each model, we term the Base data for that model.

Second, each model is executed three additional times, using each of the three metrics to modify the precipitation dataset, This analysis thus produces six output hydrology impact datasets, one for each model and desired change in each metric.

Finally, each impact dataset is compared to its relevant model Base dataset. We take the difference for each hydrology flow component between its impact dataset value and the Base value and then divide by the Base value. This yields a proportional impact measure as a non-dimensional percentage for each component and for that model-metric combo.

3.5 Results

3.5.1 Basin Metric Trends

For each basin model there are at least two local long-term climate stations with fairly complete records. From each of these stations we computed an estimate of climate trends for Intensity, Duration, and Pause and their statistical significance p-values. The base trend statistics for each station are expressed on a percentage per decade basis are shown for the Feather River Basin and for the Soquel-Aptos Basin in (**Tables 3-5 and 3-6)**, respectively. The two values lightly highlighted are statistically significant at the 95% level. The seven values that are darkly colored are highly significant at the 99% level.

For each set of individual station metric trend results, the trends are combined by weighting them by the reciprocal of their p-values. This yields a composite trend value for each metric to apply to each hydrologic model. The individual p-values are themselves weighed as well to yield a significance measure for the combined trend values. The combined metric trend results and their composite p-values are also shown (**Table 3-5**) and (**Table 3-6**).

It is encouraging that six of nine combined trends have at least one contributor that is highly significant and a seventh value has a significant contributor. There are only two of combined metrics that have little statistical significance.

3.5.2 Impacts Extent

We wish to understand the future hydrologic impacts. The precipitation metric trends are all based on past observations. As shown in **Table 3-5** and **Table 3-6**, the combined trends are all quite significant, except for the one Duration trend from the Feather River Basin. Significance does give support that the results are not random and, hopefully, are not a result of internal variability. Therefore, we assume that it is justifiable to extend those metric trends into the future.

We need to select a time period of interest. Extrapolation of linear trends is risky because some climate changes are non-linear. In particular, water evaporation, which is the supplier of atmospheric water vapor and thus a driver of precipitation, is approximately exponential with respect to increases in temperature (Alduchov and Eskridge, 1996). Also it would be productive to select a time extent long enough to be a reasonable planning horizon for many institutional investment and operation endeavors.

An extent of 30 years has been selected for future climate trend projections and hydrology impact testing. Assuming that the climate trend values computed are to be applied for a 30 year period, the Feather River Basin metrics are expected to change as: Intensity = -8.21%, Duration = 0.22%, and Pause = 6.28%. The extrapolated Soquel-Aptos metrics are expected to change by: Intensity = -8.81%, Duration = 6.55%, and Pause = 5.00%.

3.5.3 Hydrologic Impacts

The three 30 year metric trends were applied separately to the Feather River Basin precipitation dataset. The relative impacts to the hydrologic components were computed (**Table 3-7**). Similar hydrology impacts were computed for the Soquel-Aptos basin (**Table 3-8**).

3.6 Discussion

It is encouraging that four of the six metric trend composites used for the hydrologic modeling have at least one precipitation station contributor that is highly statistically significant (**Tables 3-5** and **3-6**). An additional composite trend, the Feather River Pause metric, has a statistically significant contributor. Only one of the composites, the Feather River Duration metric, has no statistical significance support. The end result is strong in that three of the six composite metric trends for modeling are highly statistically significant, two are significant, and one is not significant.

3.6.1 Basin Comparison

Before discussing hydrologic impacts, it is important to understand the quite different hydrology of the two basins. Many of these differences can be seen from their relative flow rate components in **Table 3-9**.

Some of the differences follow directly from their different climate conditions. The Feather River Basin is located in the Sierra Nevada mountains and snow is a large fraction of its precipitation and a dominant hydrology factor.

Snow sublimation is about 1% of its precipitation, whereas Soquel-Aptos is at low coastal elevations and has essentially none. With most precipitation arriving as snow and melting slowly over time, Feather has less than 0.1% runoff, instead Soquel has almost 14%. As a consequence, Feather a has high infiltration of about 93% and Soquel infiltration is lower at 68%. Along with this higher Feather infiltration flow rate, Feather has higher rates of interflow, recharge, and baseflow.

Another climate effect comes from the Soquel precipitation arriving as rain instead of snow. The warmer rain, which wets the Soquel plant leaves and stems, evaporates about 17% of Soquel precipitation. With the Feather precipitation arriving as snowfall and colder rain, the corresponding plant surface evaporation is a smaller 6%.

A final difference is caused by a land use disparity. Feather is almost entirely rural and its impervious surface evaporation is at zero. Soquel is partly urban and has an impervious evaporation of about 1.7%.

Only two of the hydrologic flow components are somewhat similar between Feather and Soquel: soil evapotranspiration and stream outflow.

With so many clear differences between the two models, mainly from climate effects, we should not be surprised to see them respond in clearly different ways to our precipitation metric changes.

3.6.2 Feather River Hydrologic Impacts

The Feather River Basin hydrology impacts have some interest (**Table 3-7**). It is perhaps gratifying to see some hydrology impacts are similar in magnitude to the climate changes.

The Feather surface runoff decrease was almost double the Intensity decrease. With runoff less than 0.1% of precipitation, its process is weak and possibly quite sensitive to change. This is reinforced by the moderately strong effect that Duration also has on runoff. In any case, runoff is so small that it is of theoretical interest, but probably of no practical consequence in this basin.

All three climate metrics had fairly strong negative effects on Feather sublimation. Weaker and longer snow storm changes could cause fresh, colder snow to continue to be added to the top of the snowpack, so Intensity and Duration changes might well reduce sublimation. It is unclear why longer pause could also reduce sublimation.

Although the Duration metric change was the smallest (+0.22%), it had significant relative influences on most hydrology flow component impacts. Five components changed with a magnitude larger than the Duration change and three more had at least half its magnitude. Some of the hydrology component sensitivities to change are probably non-linear. This Duration result might suggest that hydrology response to small climate changes could be relatively bigger than the response to large climate changes.

Perhaps the most important economic value of the Feather River Basin is to fill Lake Oroville, the largest dam and second largest reservoir in California and the primary watershed for the State Water Project. The impact on streamflow from the Duration change is negative -0.07%, which is quite minimal. In contrast, the impact from Duration is 0.26% and from Pause is 1.28%. Since both the numeric and economic streamflow impacts are largely positive, probably no mitigation or adaptation efforts are worth consideration.

3.6.3 Soquel-Aptos Hydrologic Impacts

All the Soquel-Aptos Basin climate metric changes are large and significant (**Table 3-6**). However, the hydrologic impacts to Intensity are only modest and the impacts to Duration and Pause are weak.

The strongest Intensity responses are from recharge (-6.70%) and baseflow (-5.69%) with magnitudes at least half of the -8.81% Intensity trend. Intensity impacts to another five components, plant evaporation (3.65%), impervious evaporation (2.97%), surface runoff (-2.20%), interflow (-2.88%), and streamflow (-2.98%), are percentages in a non-trivial 2-4% range.

Both Duration and Pause impacts are anemic. There is only one impact, the impervious evaporation to Duration change, which is barely over 1%. It is puzzling why this would happen. It is almost as if this temperate hydrology is insensitive to any timing changes.

The only effect that seems to make any big difference to impacts is actual water amount changes. But note that the Intensity metric change does not change the total precipitation received. All that this change does is reduce mean Intensity by redistributing precipitation amounts. Such an Intensity reduction seems to be accomplished by moving rainfall amounts away from the short events into the longer events. This would make small storms weaker and big storms stronger. Stronger big storms could produce more soil saturation, which would be consistent with the two largest Intensity effects of reduced recharge (-6.70%) and baseflow (-5.69%). Soil saturation would be expected to decrease infiltration, which did happen but very weakly (-0.88%). However, soil saturation would also be expected to increase runoff and interflow, however both decreased: runoff (-2.20%) and interflow (-2.88%). The only impacts that increased were plant evaporation (3.65%) and impervious evaporation (2.97%), which do not follow logically. We have no cogent theory that would explain all the confusing impacts from this Intensity change.

The values for water in the Soquel-Aptos Basin are for domestic water supply, irrigation, and the environment. The news from this study are all bad.

If water supply comes from groundwater, this is bad news. All three recharge impacts are negative and especially Intensity at -5.69%. The only positive components for Intensity are the three evaporation impacts. These can not be reduced to mitigate the deficit. Effecting some sort of rain water detention solution such as berms or spreading basins could increase infiltration. But these solutions come at the cost of even more reduction to streamflow, which would harm riparian ecosystems, threatened fish, and its use for water supply. So the only solution seems to be to adapt to these shortages.

If water supply comes from streams, this is bad news. All three streamflow impacts are negative and especially Intensity at -2.98%. Once again the positive evaporation impacts can not be reduced. Routing more surface water into runoff would harm infiltration and groundwater recharge. So one must try to adapt to these shortages.

For irrigation and the environment, soil moisture is of importance. The news for soil moisture is not good. For all three metric changes, the gain (if any) from the major soil input (infiltration) is smaller than the loss from the major soil output (evapotranspiration). In particular for the Intensity change, the infiltration input decreases by -0.88% and the evapotranspiration output increases by 0.77%. Yet again, these Intensity impacts can not be easily mitigated, so one must try to adapt (somehow).

3.7 Conclusions

This study was designed to quantify the impacts of precipitation trends in two hydrologic basins in California. This was done by using computed trends in precipitation metrics, extrapolated over a 30 year period, to drive hydrologic models. The models for the two basins show large differences in behavior attributable to the differences in climate.

Our results suggest that hydrologic impacts are typically modest compared to changes of the three precipitation metrics. The magnitudes of most hydrologic impacts are less than half the magnitudes of the corresponding precipitation metric changes. The only exception was the Feather River Basin impacts to the Duration metric change. This change was far smaller than that for any other metrics, suggesting that perhaps small changes might have relatively larger hydrologic impacts.

Impacts are often subjective. Much depends on what it is that someone values, which determines the hydrologic components that matter. If the Feather River streamflow into the Lake Oroville and then feeds the State Water Project is all important, then impacts are largely positive and thus beneficial by a mean of 0.5%. For critical Soquel-Aptos water supply, the impacts of all three precipitation changes harm both groundwater recharge by -2.5% and surface streamflow by -1.1%. There seems to be no feasible ways to mitigate these impacts and thus adaptation to this climate change is indicated.

References

Alduchov, O.A., and Eskridge, R.E., 1996, Improved Magnus Form Approximation of Saturation Vapor Pressure: Journal of Applied Meteorology, v. 35, p. 601–609, doi: 10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Bae, D., Jung, I., and Chang, H., 2008, Potential changes in Korean water resources estimated by high-resolution climate simulation: CLIMATE RESEARCH, v. 35, p. 213–226, doi: 10.3354/cr00704.

Chang, H., and Jung, I., 2010, Spatial and temporal changes in runoff caused by climate change in a complex large river basin in Oregon: JOURNAL OF HYDROLOGY, v. 388, p. 186–207, doi: 10.1016/j.jhydrol.2010.04.040.

Everitt, B.S., 2010, Cambridge Dictionary of Statistics: Cambridge University Press, 478 p.

Flint, L.E., and Flint, A.L., 2012, USGS Scientific Investigations Report 2012–5132: Simulation of Climate Change in San Francisco Bay Basins, California: Case Studies in the Russian River Valley and Santa Cruz Mountains:.

Hay, L.E., Clark, M.P., Pagowski, M., Leavesley, G.H., and Gutowski, W.J., 2006, One-Way Coupling of an Atmospheric and a Hydrologic Model in Colorado: Journal of Hydrometeorology, v. 7, p. 569–589.

Hay, L.E., Markstrom, S.L., and Ward-Garrison, C., 2010, Watershed-Scale Response to Climate Change through the Twenty-First Century for Selected Basins across the United States: Earth Interactions, v. 15, p. 1–37, doi: 10.1175/2010EI370.1.

Huang, G., Kadir, T., and Chung, F., 2012, Hydrological response to climate warming: The Upper Feather River Watershed: Journal of Hydrology, v. 426-427, p. 138–150.

Hydrometrics, 2011, Estimation of Deep Groundwater Recharge Using a Precipitation-Runoff Watershed Model Soquel-Aptos, California:.

Im, E., Jung, I., Chang, H., Bae, D., and Kwon, W., 2010, Hydroclimatological response to dynamically downscaled climate change simulations for Korean basins: CLIMATIC CHANGE, v. 100, p. 485–508, doi: 10.1007/s10584-009-9691-2.

Johnson, N.M., Williams, D.E., Yates, E.B., and Thrupp, G., 2004, Groundwater assessment of alternative conjunctive use scenarios, draft technical memorandum 2: hydrogeologic conceptual model:.

Koczot, K.M., Jeton, A.E., McGurk, B., and Dettinger, M.D., 2005, Precipitation-Runoff Processes in the Feather River Basin, Northeastern California, and Streamflow Predictability, Water Years 1971-97: Scientific Investigations Report 2004-5202, 82 p.

Koczot, K.M., Markstrom, S.L., and Hay, L.E., 2012, Watershed Scale Response to Climate Change—Feather River Basin, California: U.S. GEOLOGICAL SURVEY Fact Sheet 2011–3125, 6 p.

Leavesley, G.H., Lichty, R.W., Troutman, B.M., and Saindon, L.G., 1983, Precipitation-Runoff Modeling System: User's Manual: U.S. Geological Survey Water-Resources Investigations 83-4238, 207 p.

Legesse, D., Vallet-Coulomb, C., and Gasse, F., 2004, Analysis of the hydrological response of a tropical terminal lake, Lake Abiyata (Main Ethiopian Rift Valley) to changes in climate and human activities: HYDROLOGICAL PROCESSES, v. 18, p. 487–504, doi: 10.1002/hyp.1334.

Walker, J.F., Hay, L.E., Markstrom, S.L., and Dettinger, M.D., 2011, Characterizing Climate-Change Impacts on the 1.5-yr Flood Flow in Selected Basins across the United States: A Probabilistic Approach: Earth Interactions, v. 15, p. 1–16, doi: 10.1175/2010EI379.1.

Wenming, Z., Zengvhuan, D., and Chengtao, Z., 2008, Distributed hydrologic modelling based on MMS/PRMS and its applications: Water Resources and Power, v. 26, p. 9–13.

Figure 3-1 Two study sites in California: the Soquel-Aptos Basin in Santa Cruz County and the Feather River Basin in Plumas, Butte, Lassen, Shasta, and Sierra Counties.

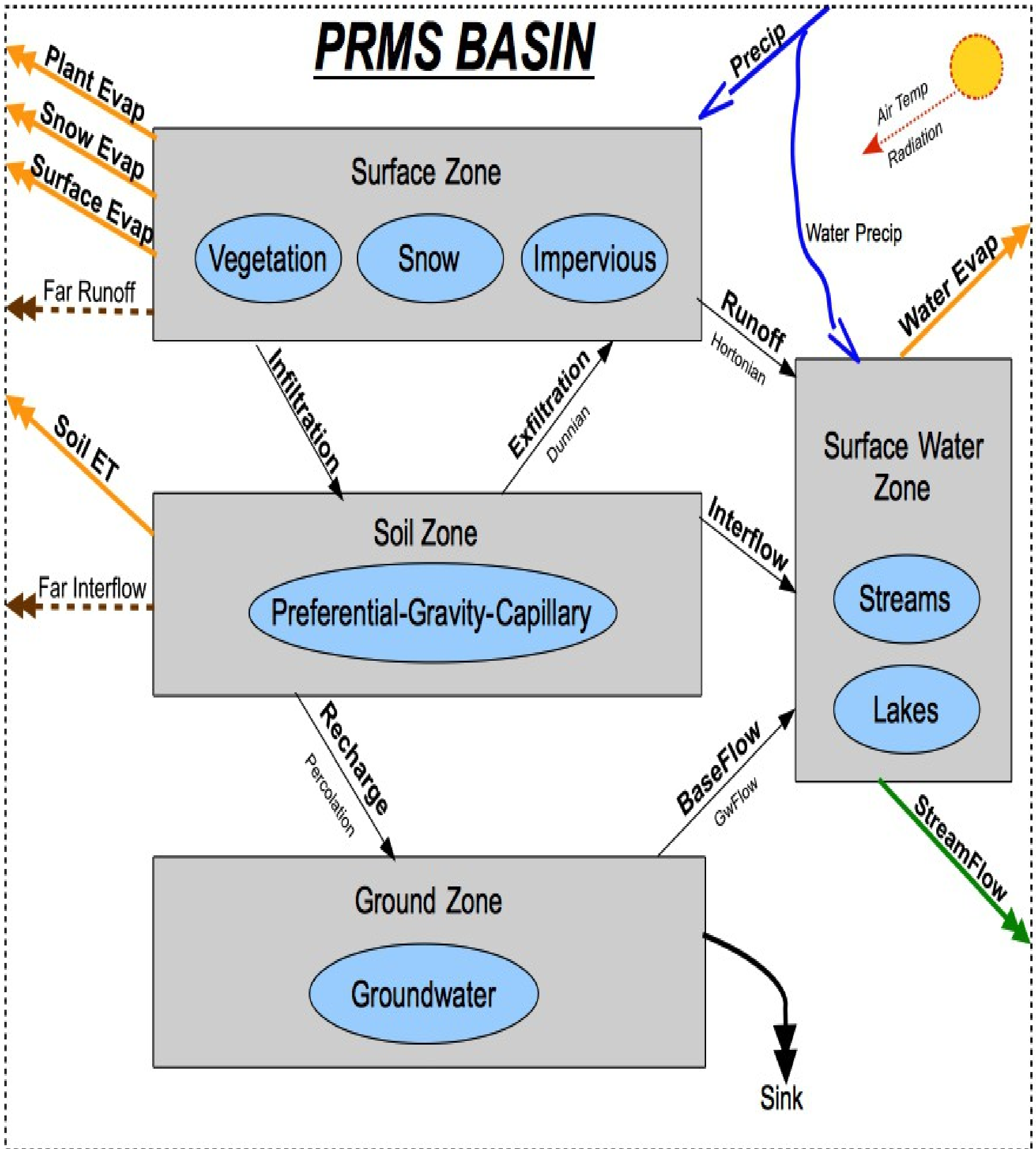


Figure 3-2 Schematic of the USGS Precipitation-Runoff Modeling System (PRMS) conceptual component architecture. The seven water store reservoirs are shown as blue ovals grouped into the four zones: Surface, Soil, Ground, and Surface Water. The one basin water input is precipitation. The nine water flow outputs are shown as thick lines with double arrow heads leading to the basin dotted boundary. The six internal water flows between zones are thin lines with single arrow heads.

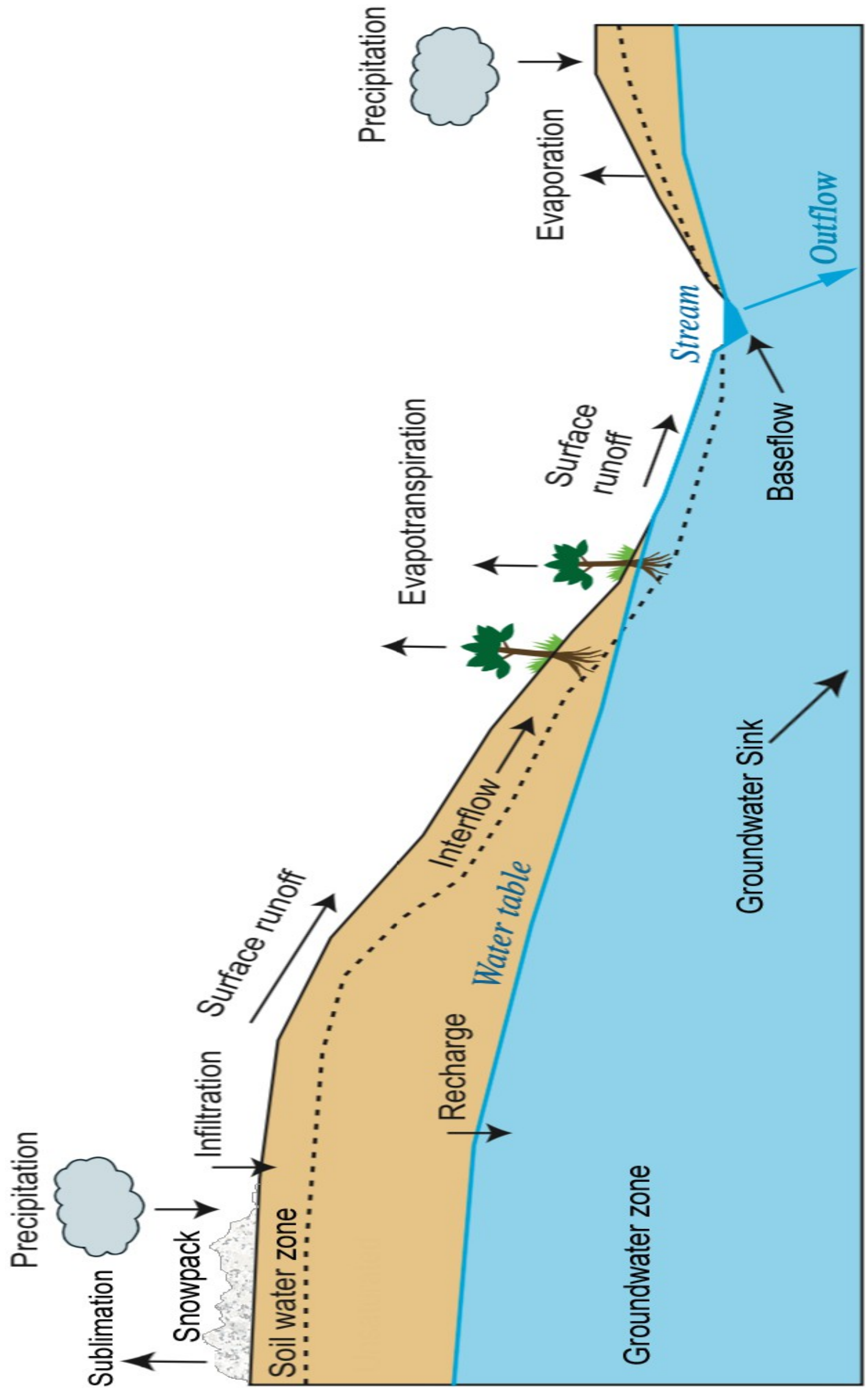


Figure 3-3 PRMS geographical component architecture view.

Table 3-1 Initial events and their Grand Mean.

Event #	Mean values	Element count	Element values
#1	0.99475	2	0.89135, 1.0982
#2	3.3977	3	4.0874, 3.9204, 2.1853
#3	1.2275	1	1.2275
#4	2.7390	3	4.8198, 2.3465, 1.0506
#5	2.0368	2	3.6771, 0.39642

This is an initial set of five events of random length and values. Their Grand Mean (the mean of this set of five mean values) is 2.0791.

Table 3-2 Final events and their maximized Grand Mean.

Event #	Mean values	Element count	Element values
#1	3.0118:	2	3.6771, 2.3465
#2	1.8481:	3	3.9204, 1.2275, 0.39642
#3	4.8198:	1	4.8198
#4	1.3757:	3	1.0506, 0.89135, 2.1853
#5	2.5928:	2	4.0874, 1.0982

This is the final set of five events each with the same count of elements as above. However, the element values have been repeatedly swapped between events in order to find this final arrangement that has the largest Grand Mean value of 2.7297.

Table 3-3 Event shift transitions for changing Pause metric

Case start	L changes	G changes	Net change (G–L)	Case end
GG	+1	-1	-2	GL>2D+1
GL>2D+1	-1	+1	+2	GG
GL<2D-1	+1	-1	-2	LL>D
LL>D+1	-1	+1	+2	GL<2D-1
LL<D+1	none	none	0	none

This the set of interesting cases for a pair of lulls that surround a precipitation event. Shown is transitions between cases when a precipitation event between two lulls is shifted. This is done in order to change a Pause metric.

A letter 'G' represents a lull with a length greater than the current Pause median value 'D'. A letter 'L' is a lull with length less than D. Transitions go from the start column to the corresponding end column with the associated change in the count of Ls, change in count of Gs and net change of the difference.

Table 3-4 Log of event shifts performed to change a Pause metric

Increase goal	Increase achieved	Shifts	Pause Metric
+10.000%	0.0%	0	4.4810
+10.000%	+3.554%	34	4.6402
+10.000%	+7.933%	42	4.8364
+10.000%	+9.393%	14	4.9019
+10.000%	+9.811%	4	4.9206
+10.000%	+10.019%	2	4.9299

Log of the process to increase Pause value by +10% by doing shifts of precipitation events between two lulls. This is for the Santa Cruz station precipitation dataset from January 1, 1893 to February 1, 2012. The Pause median starts with a value of exactly 4, but its interpolated value is 4.481. Note on each repetition it uses a different count of shifts to be done based on the progress achieved by the last shift count.

Table 3-5 Feather River Basin Climate Trends and Significance

Metric	Canyon Dam	Chester	Quincy	Combined
Intensity				
Trend/decade	0.36%	-1.18%	-2.75%	-2.74%
P-Value	39.59%	13.12%	**0.06%**	**0.17%**
Duration				
Trend/decade	-0.20%	-0.25%	0.45%	0.07%
P-Value	39.63%	38.23%	25.28%	32.98%
Pause				
Trend/decade	2.21%	0.83%	0.38%	2.09%
P-Value	**1.09%**	22.35%	34.05%	**3.04%**

There are three precipitation stations of note in the Feather River Basin area. Their trends for metrics Intensity, Duration, and Pause are combined by weighting with the reciprocal of their p-values. The two p-values shown in bold are significant at the 95% level (p-value less than 5%). The two shown in red are at the further 99% significance level (p-value less than 1%). Thus two of the three combined results are significant to at least the 95% level.

Table 3-6 Soquel-Aptos Basin Climate Trends and Significance

Metric	Watsonville	Santa Cruz	Combined
Intensity			
Trend/decade	-3.12%	-1.98%	-2.94%
P-Value	**0.10%**	**0.53%**	**0.17%**
Duration			
Trend/decade	2.19%	0.60%	2.18%
P-Value	**0.05%**	25.13%	**0.10%**
Pause			
Trend/decade	0.84%	1.69%	1.67%
P-Value	20.83%	**0.67%**	**1.30%**

There are two relevant precipitation stations in the Soquel-Aptos Basin area. Their trends for metrics Intensity, Duration, and Pause are combined by weighting with the reciprocal of their p-values. The one p-value shown in bold is significant at the 95% level (p-value less than 5%). The six shown in red are at the further 99% significance level (p-value less than 1%). All three of the combined results are significant to at least the 95% level.

Table 3-7 Feather River Basin Hydrology Impacts

Component	Base (cm/yr)	ΔIntensity	ΔDuration	ΔPause
Precipitation	100.41	0.23%	0.00%	0.02%
Plant Evap	6.13	-0.06%	-0.05%	0.40%
Sublimation	1.12	**-5.87%**	**-0.91%**	**-5.80%**
Imperv Evap	0.00	n/a	n/a	n/a
Surface Runoff	0.08	**-16.61%**	**0.15%**	-0.45%
Infiltration	93.08	0.33%	0.02%	0.07%
Soil ET	19.07	3.07%	**-0.14%**	-1.36%
Interflow	30.44	0.28%	**0.39%**	1.56%
Other ET	21.59	-1.58%	**-0.49%**	-2.46%
Recharge	21.98	-0.09%	**0.12%**	1.73%
Groundwater Sink	10.63	0.77%	**0.35%**	2.98%
Baseflow	11.35	-0.91%	-0.08%	0.55%
Stream Outflow	41.88	-0.07%	**0.26%**	1.28%

These are the impacts to the Feather River Basin hydrology components relative to their base values, i.e. the PRMS model outputs produced using unaltered observational climate inputs. These hydrology impacts are separately from the climate changes of the three metrics ΔIntensity = -8.21%, ΔDuration = +0.22%, and ΔPause = +6.28%. The five impacts with magnitude greater than half the trend's magnitude are shown in bold and the six greater than the trend's magnitude are shown in red.

Table 3-8 Soquel-Aptos Basin Hydrology Impacts

Component	**Base** (cm/yr)	**ΔIntensity**	**ΔDuration**	**ΔPause**
Precipitation	82.01	-0.24%	0.03%	0.01%
Plant Evap	13.69	3.65%	-0.66%	-0.47%
Sublimation	0.00	n/a	n/a	n/a
Imperv Evap	1.36	2.97%	1.05%	0.66%
Surface Runoff	11.33	-2.20%	-0.03%	0.01%
Infiltration	55.63	-0.88%	0.18%	0.12%
Soil ET	38.81	0.77%	0.39%	0.25%
Interflow	8.74	-2.88%	-0.13%	-0.07%
Recharge	8.02	**-6.70%**	-0.45%	-0.34%
Groundwater Sink	0.00	n/a	n/a	n/a
Baseflow	3.57	**-5.69%**	-0.32%	-0.34%
Stream Outflow	23.63	-2.98%	-0.11%	-0.07%

These are the impacts to the Soquel-Aptos Basin hydrology components relative to their base values, i.e. the PRMS model outputs produced using unaltered observational climate inputs. These hydrology impacts are separately from the climate changes of the three metrics ΔIntensity = -8.81%, ΔDuration = +6.55%, and ΔPause = +5.00%. The two impacts with magnitude greater than half the trend's magnitude are shown in bold.

Table 3-9 Comparison of Hydrology of Feather River versus Soquel-Aptos Basins

Component	Feather River	Soquel-Aptos	Comparison
Precipitation			
Plant Evap	6.10%	16.69%	-46%
Sublimation	1.11%	0.00%	100%
Imperv Evap	0.00%	1.66%	-100%
Surface Runoff	0.08%	13.81%	-99%
Infiltration	92.70%	67.83%	15%
Soil ET	40.50%	47.32%	-8%
Interflow	30.32%	10.65%	48%
Recharge	21.89%	9.78%	38%
Groundwater Sink	10.58%	0.00%	100%
Baseflow	11.31%	4.35%	44%
Stream Outflow	41.71%	28.81%	18%

Comparison of the two basins, Feather River and Soquel-Aptos, by showing their flow components expressed as a percentage of their Precipitation input. The Comparison column is the difference of the two percentages as a fraction of their sum. A positive value shows that Feather River has relatively more of that component and a negative shows that Soquel-Aptos has more.

CONCLUSIONS

The original vision for this study was simply to derive results for climate change impacts on water resources. We believe that some portion of that vision has been achieved. There is a logical chain of derivation steps of those results through all the chapters. Chapter 1 derived trends of three metrics from precipitation observations. Chapter 2 computed significance p-values for those trends. Chapter 3 used these p-values to combine local trend metrics, which produced the input into two hydrological basin models. Comparing the model outputs before and after these trend changes yielded relative climate change impact results.

There were some interesting new, or at least unusual, twists to these hydroclimatology results. These results did not utilize the predominate method of obtaining climate change information from the execution of global climate models (GCMs). Because GCMs seem to have problems with precipitation significance and skill, they were rejected. Instead, this study produced climate change results by obtaining trends from observations and projecting these forward into the future.

Another twist is the low-level precipitation metrics chosen for trend estimation. Precipitation is often analyzed at a coarse time granularity, such as quarterly or yearly totals. The finest granularity for study is usually the total accumulation of a complete storm event. This study divided that quantity down further into the Duration of an event and its Intensity, the daily mean precipitation amount. Another metric used in this study, but rarely studied elsewhere, was the Pause length of the lull time between storms.

A third twist has been the ability to define these three metrics so that their trends can be derived separately and then applied independently. This has allowed the hydrologic impacts from a single metric change to be considered in isolation from the other two metric changes and any change in total precipitation. Sometimes such isolated impact analysis can lead to insight, or sometimes, as happened in cases here, it can lead to puzzlement.

However, perhaps the most important accomplishment of this study has turned out to be its methodology. Throughout this study unexpected challenges have cropped up. These had to be resolved through the piecing together of suitable methodologies.

The first challenge was whether and how to deal with the ever-present missing data in precipitation records. It became clear that missing data could not simply be ignored for trend computation since that produced statistical deviations. A number of methodologies were investigated as possible solutions, but without success. The multiple imputation methodology, as applied in this work, appears to work admirably by supplying, not only a likely value for the missing data, but also its probability distribution, which can inform about the uncertainty range for the missing value.

The second challenge was how to measure statistical significance without ignoring the effects of missing data. The permutation algorithm is well suited to calculation of significance p-values. The permutation algorithm was enhanced by a strategic multiple imputation step which takes values from a distribution and fills in for the missing data.

A great attribute of these methodology contributions is they are fairly general data processing methods. Therefore, they are probably useful in many diverse fields and for all kinds of problems. The broad utility of the methodology in this study is satisfying since the climate change impact results seem less clear-cut and perhaps less valuable than we originally envisioned. However, we can be pleased that in California, with its highly variable climate, we were still able to successfully derive and utilize so many trends with such impressive levels of significance.

Future Work

The first thing that cries out to be studied is the application of the three metric changes at the same time to the hydrology models. All three trends were detected in the observations occurring at the same time. It seems that applying them together is perhaps even more justified than applying them in isolation. That would give a clearer picture of the total climate impacts to hydrology. We would then be able to tell if the separate impacts reinforce each other or cancel out.

We could apply these analysis methodologies to other metrics, such as the number of precipitation events in a yearly period or some measure of precipitation seasonality. We could also study metrics for the one missing precipitation feature:

drizzle instances where precipitation is less than 1 mm/day but greater than zero.

Another thing we have long wanted to investigate is the role of precipitation metrics in droughts. We all know that a drought is an unusual decrease in precipitation. But does this decrease stem from having fewer precipitation events, or from less intense events, or shorter events? This metric analysis could well be applied in other situations. For example, how does El Niño or La Niña affect the statistics of these metrics? What about the Pacific Decadal Oscillation? It seems these low-level metrics and their inference might give a new way to analyze problems and processes.

APPENDICES

Chapter 1

A1.1 Global Historical Climatology Network-Daily precipitation data

For this study we use the GHCN daily climate records from the Global Historical Climatology Network-Daily dataset, which is also known as GHCN-Daily or GHCND (Klein Tank et al., 2002). For this work we use the version of GHCN Daily denoted as "3.12-upd-2014040705", which was prepared and published on 7 April 2014, and retrieved from <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily>.

The GHCND-derived text file 'ghcnd-inventory.txt' contains 527,770 lines of text that describe the climate elements each station collects. In addition to precipitation data, stations can provide climate data on local temperatures, wind speeds, cloud cover, humidity, etc. These various climate elements are specified via four character tags. In this study, we focus on the precipitation element, which is denoted by the element tag 'PRCP'. The full set of element tags are listed and described in the GHCN supplied text file 'readme.txt'.

In the inventory file for Santa Cruz, California there is the text line:

```
USC00047916  36.9906 -121.9911 PRCP 1893 2014
```

This inventory line starts with the 11 character station identification number, which for Santa Cruz is 'USC00047916'. Next the station's Latitude and Longitude are supplied. Then there is the indication that this line is the description for the precipitation data denoted by tag 'PRCP'. Finally is the info that the Santa Cruz data provided for this precipitation element begins within the year 1893 and extends into 2014.

The actual climate data for this, or any other, station is contained in a separate text file named the same as the station identifier appended with the file suffix of '.dly', which stands for 'daily'. Each text line of that station data file contains the values for just one kind of climate element and for one month of record. Here is the first text line of the text file named 'USC00047916.dly' containing Santa Cruz data:

```
USC00047916189301PRCP    0P 6    0P 6    0P 6
       0P 6    0P 6    0P 6    0P 6    0P 6    0P 6
       0P 6    0P 6    0P 6    0P 6    0P 6  318  6
       0P 6    0P 6    0P 6    0P 6    0P 6    0P 6
       0P 6    0P 6    0P 6    0P 6  572  6   89  6
      38  6  216  6  114  6    0P 6
```

Immediately after the 11 character station identifier is found the four digit year number '1893' and the two digit month number '01', which corresponds to January. The fact that this text line supplies data for the precipitation element is indicated next by the tag 'PRCP'. The daily data for that month then follows and consists of 31 repeated 8 character fields for each day of that month. If the month has less than 31 days, then the extra day fields at the end of the line are set to the special value that denotes it is missing and can just be ignored.

These 8 character fields for each day consists of first a 5 character integer number for the precipitation value as expressed in tenths of a millimeter (mm). The GHCND has a special notation for a trace precipitation, which is a day with a value intermediate between the zero amount of 0 mm/day and the smallest non-zero value representable of 0.1 mm/day. For this work we store such a trace amount with an intermediate quantity of 0.05 mm/day. Any missing data item is indicated by a special value of -9999. After the number field, there follows three single character fields, which contain tags with information about that value's measurement (10 possible values), its quality (14 values), and its source (28 values). The various meanings for these flag values will not be described here but are specified in the GHCN text file 'readme.txt'.

The precipitation records in the text file 'USC00047916.dly' for Santa Cruz, CA extend from the month of January 1893 into March 2014, 121 years and 2 months and five days. This means that there are 1,455 months of precipitation data listed in that full record, each of which is provided by a single text line in that file. In particular, for the initial month of January 1893 that was shown above, the 31 provided precipitation values in mm are:

```
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,
   0.0, 0.0, 0.0, 31.8, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,
   0.0, 0.0, 0.0, 0.0, 57.2, 8.9, 3.8, 21.6, 11.4, 0.0
```

We see this January 1893 month had a relatively dry start with no precipitation until the 15th, which got a single 31.8 mm rain day, and then later there was a nice five day storm starting on 26th with a total of accumulation of 102.9 mm. In addition, we did have data for all 31 days of that month with no observations missing. When all the other months of this precipitation data are extracted from this Santa Cruz data file, then we get the annual rainfall chart as shown in (**Figure A1-1**).

There are 2,339 GHCN stations with precipitation records in California, of which 2018 provide precipitation records. Each GHCND station is identified with an 11 character identification code, e.g. the id for Santa Cruz, California is 'USC00047916'. The first two letters of that id are the FIPS code of the country containing the station. These country code abbreviations are detailed in the GHCN text file named 'ghcnd-countries.txt', e.g. 'US' stands for United States. The third character of a station identifier is a network code (9 possible values), which identifies the station numbering system used. For example, the 'C' network in the Santa Cruz station identifier is described in file 'readme.txt' as

C = U.S. Cooperative Network identification number.

The remaining eight characters of the id contain the actual station Id specifier string.

Associated with each station identifier in the 91,267 text line GHCN supplied file named 'ghcnd-stations.txt' is a single text line with some basic station info. For the Santa Cruz station id of 'USC00047916' is the text line:

```
USC00047916 36.9906 -121.9911 39.6 CA SANTA CRUZ HCN
```

This station is described as located at Latitude 36.9906°N and Longitude W -121.9911°. Then is an elevation value of 39.6m and info that it is located in the state of 'CA', i.e. California. This line shows the station has the name 'SANTA CRUZ'. The station name is not necessarily unique in the GHCND. For example, the station 'BR000253000' along with four others located in Brazil have the same name 'SANTA CRUZ'. Finally indicated is that this station is part of the 'HCN', which is the NOAA U.S. Historical Climatology Network (USHCN), a set of stations created to quantify national and regional climate changes.

In order to get valid statistics, we need stations with long-term precipitation records. The longer the record, the better to avoid effects from climate cycles such as the Pacific Decadal Oscillation (PDO), which has events that can persist for 20-30 years (Mantua et al., 1997). Therefore, it is important to be able to span a complete PDO cycle, which can be 60 years in length (Minobe, 1997). There are 214 California stations with at least 85 years of precipitation data, 159 stations with 100 years, 49 stations have 120 years, and only 3 stations have as much as 130 years of record.

Some California locations have most of their precipitation data missing, which makes them unusable for long-term trend analysis. For example, Bucks Lake in Plumas county is missing 93% of its 1916-1970 daily records. In Kern county Tehachapi has 91% missing and Lost Hills has 94% missing. There are four sites in Southern California with more than 90% missing. On average, California stations with at least 50 years of data are missing 19.6% of their daily records.

Of the GHCND stations in California, there are 50 precipitation stations that have both long periods of at least 85 years and no more than 7% of their daily data missing.

Of these 50 sites, the average quantity of missing data is 4.1%. Only four of those sites have less than 1% of their daily records missing. These are 'San Francisco Dwtn' with id 'USW00023272', which has 0.0% missing, 'Sacramento 5 Ese' with id 'USW00023271' has 0.2% missing, 'Graton' with id 'USC00043578' has 0.3%, and 'Scotia' id 'USC00048045' has 0.6%.

A1.2 ClimateData Software

One of the major components in this study is a custom built software system called *ClimateData,* which is used for the access, manipulation, and processing of precipitation and other climate data. ClimateData was designed and implemented by the author for this project and consists of some 25,000 lines of Java code.

Data processing systems typically have some fundamental data unit(s) for its data representation and processing. For example, the basic unit of a relational database management system (RDBMS) is a record. In ClimateData the fundamental unit is a container called a *DataSet,* which holds a fixed-sized collection of double precision floating point numbers.

The construct *NumSeq* is defined to reference a sequence, or perhaps just a sub-sequence, of the numerical values in a DataSet. In fact, there can be multiple such NumSeq's that all reference portions in the same DataSet. Those references can be disjoint, overlapping, nested, identical, etc. So one NumSeq can reference an entire multi-thousand days of precipitation quantities while others reference the data for each year or for each individual storm event. A change made through any of these NumSeq's can be shared and visible in all of the others which contain that data item.

A NumSeq also provides a host of processing functionality. You can add, subtract, multiply, or divide a NumSeq by either a single number or another NumSeq of the same size. A rather complete set of statistical operations are also provided for computing mean, median, sum, max, min, variance, standard deviation, absolute deviations, quarters, moving averages, summations, etc. Other operations provide the capabilities to perform data searches, sub-setting, selecting, and sorting.

A NumSeq provides the capability to specify a particular data item by an integer index or a sequence of data items by pair of indexes as a range. A *DateSeq* is an extension of NumSeq, which also allows such data specification by dates, e.g. get all data from 21 June 1907 up to 21 September 1912. In addition, a DateSeq provides addditional date operations such as to slice the data into multiple month or water-year chunks. There is also support there for reading in and processing data values from text files using the very specialized and complex GHCN data format.

Further extensions of the DateSeq data type are specialized for handling precipitation data record extents (a *Precip* object) and also individual precipitation storm events or the lull period between events (a *PrecipState* object). There is even an extension called a *PrecipStateFragment,* which can be used to represent pieces of a

PrecipState broken up so that no data crosses any time period boundary such as months or otherwise would make statistics for such periods somewhat arbitrary.

A set of geographic objects are closely associated with climate data. The *PrecipSource* object contains sets of the sources of data used in this study. Each of these sources is a *Station,* which contains its specific unique GHCN identification(s) together with information about the inventory of the different kinds of data it has to offer, e.g. PRCP, TMAX, TMIN, etc. Each Station is also a *Location,* which defines its position as a longitude, latitude, elevation, country, province/state, city, postalcode, timezone, etc. together with the functionality to find any Location(s) with particular values of those attributes or those within certain distances. Given two Locations it can even determine the great circle distance and the direction from one to the other.

ClimateData offers some additional higher level statistics functionality. There is a choice of several covariance, correlation, and auto-correlation methods. A standard set of probability distributions are available together with methods to create probability density functions from datasets. The usual parametric and some non-parametric regression techniques are present and can be utilized. Graphing of these statistical results is then supported by connections to the bundled open-source *JFreeChart* package (Gilbert, 2014).

Finally there is a subsystem that gives much of the functionality similar to that offered by a basic relational database, i.e. selection, projection, grouping, and aggregation. Someone can select from a sequence of all precipitation storm events (a *PStateList*) to create another PStateList containing those with intensity values greater than 5 mm/day or those with a duration of 3 days. A projection can extract a value such as the total storm accumulation amount from each element of such a PStateList to create a NumSeq of those values. Grouping allows the execution of an operation such as to group all precipitation storm events by the month of the year, which produces a mapping from each of the months January through December onto a sequence (a PStateList) of the storms that happened in that month. Aggregation is a calculation operation that allows such composites to be easily summarized. For example, in the grouping for each such month, compute the standard deviation of all the intensities extracted from each storm event in that month's list. This produces a NumSeq of the 12 monthly standard deviation quantities, which could then be displayed as a bar graph.

Most ClimateData capabilities are fairly standard computer science techniques and statistics capabilities. However in addition, there are several very specialized and somewhat complex statistical processing and analysis capabilities in ClimateData that have been researched and created just for this study of missing data and statistical trend inference. These facilities are described in the Methods section of Chapter 1.

References

Gilbert, D., 2014, JFreeChart: http://www.jfree.org/jfreechart/, Object Refinery Limited.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., et al., 2002, Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment: International Journal of Climatology, v. 22, p. 1441–1453, doi: 10.1002/joc.773.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and Francis, R.C., 1997, A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production: Bulletin of the American Meteorological Society, v. 78, p. 1069–1079, doi: 10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Minobe, S., 1997, A 50–70 year climatic oscillation over the North Pacific and North America: Geophysical Research Letters, v. 24, p. 683–686, doi: 10.1029/97GL00504.

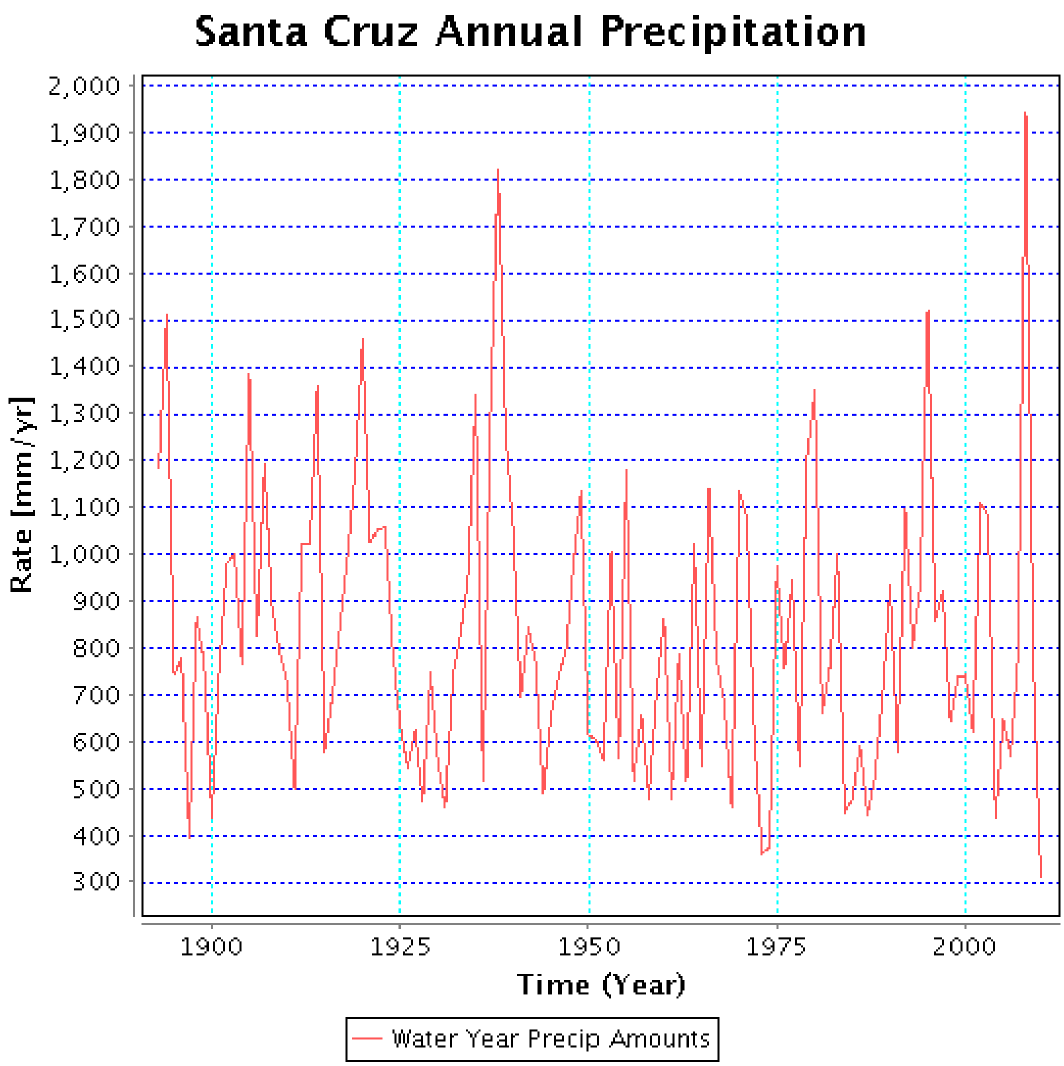


Figure A1-1 Annual precipitation from the NOAA GHCN daily dataset for station 'USC00047916' in Santa Cruz, California. Annual precipitation amounts are accumulated for water-years, which are defined as running from October 1st through September 31st of the next year. The annual precipitation mean value over the period of record is 827 mm/yr and the median is 768 mm/yr. The standard deviation is 306 mm/yr, the mean absolute deviation is 242 mm/yr, and its median absolute deviation is 210 mm/yr.