Brigham Young University
BYU ScholarsArchive

Theses and Dissertations

2019-06-01

Toward Using Empirical Mode Decomposition to Identify Anomalies in Stream FlowData and Correlations with other Environmental Data

Saul Gallegos Ramirez
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Engineering Commons

BYU ScholarsArchive Citation
Ramirez, Saul Gallegos, "Toward Using Empirical Mode Decomposition to Identify Anomalies in Stream FlowData and Correlations with other Environmental Data" (2019). *Theses and Dissertations*. 7574.
https://scholarsarchive.byu.edu/etd/7574

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

المنارة للاستشارات
www.manaraa.com

Toward Using Empirical Mode Decomposition to Identify Anomalies in Stream Flow Data and Correlations with Other Environmental Data

Saul Gallegos Ramirez

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Gustavious P. Williams, Chair
Daniel P. Ames
Rollin H. Hotchkiss

Department of Civil and Environmental Engineering

Brigham Young University

Copyright © 2019 Saul Gallegos Ramirez

All Rights Reserved

المنارة للاستشارات

www.manaraa.com

ABSTRACT

Toward Using Empirical Mode Decomposition to Identify Anomalies in Streamflow Data and Correlations with Other Environmental Data

Saul Gallegos Ramirez
Department of Civil and Environmental Engineering, BYU
Master of Science

I applied empirical mode decomposition (EMD) and the Hilbert-Herbert transforms, as tools to analyze streamflow data. I used the EMD method to extract and analyze periodic processes and trends in several environmental datasets including daily stream flow, daily precipitation, and daily temperature on data from the watersheds of two rivers in the Upper Colorado River Basin, the Yampa and the Upper-Green rivers. I used these data to identify forcing functions governing streamflow. Forcing functions include environmental factors such as temperature and precipitation and anthropogenic factors such as dams or diversions. The Green and Yampa Rivers have similar headwaters, but the Yampa has minimal diversions or controls while Flaming George Dam on the Green river significantly affects flow. This provides two different flow regimes with similar large watersheds. In addition to flow data, I analyzed several time series data sets, including temperature and precipitation from Northeast Utah, North Western Colorado, and Southern Wyoming. These data are from the area that defines the Yampa River and Green River watersheds, which stretch from Flaming Gorge Dam to Ouray Colorado.

The EMD method is a relatively new technique that allows any time series data set, including non-linear and non-stationary datasets that are common in earth observation data, to be decomposed into a small quantity of composite finite data series, called intrinsic mode functions (IMFs). The EMD method can decompose any complicated data into several IMFs that represent independent signals in the original data. These IMFs may represent periodic forcing functions, such as environmental conditions or dam operations, or they may be artifacts of the decomposition method and not have an associated physical meaning. This study attempts to assign physical meaning to some IMFs resulting from the decomposition of the Green and Yampa flows where possible. To assign physical meaning to the IMFs, I analyzed frequencies of each IMF using the Hilbert-Hung transform, part of the Empirical Mode Decomposition method, and then compared frequencies of the IMFs with the known frequencies of physical processes. I performed these calculations on both flow, temperature, and precipitation.

I found significant correlation between IMF components of flow, precipitation, and temperature data with El Niño Southern Oscillation (ENSO) events. The EMD process also extracts the long-term trend in non-linear data sets that can provide insights into the effects of climate change on the flow system. Though in preliminary stages of research, these analysis methods may lead to further understanding the availability of water within the upper Yampa and Green River Watersheds.

Keywords: Empirical Mode Decomposition, Hilbert Haung Transforms, El Niño, La Niña, Yampa River, Yampa, Flaming Gorge, Green River

www.manaraa.com

ACKNOWLEDGEMENTS

They say that we stand on the shoulders of giants; I would not be where I am in life without the sacrifices of people that came before me. My family came with nothing to this country and sacrificed everything so that I could have the opportunity to obtain an education. In return for my family's sacrifice, I have put all my effort into obtaining an education. I don't in the slightest believe this work is a payment for their sacrifice, but this work is a symbol of my appreciation for my family for what they have done for me.

My second vote of appreciation goes out to Dr. Gus P. Williams, my advisor and mentor. This work required many long days and late nights. I would never have completed this work without his direction. Dr. Williams allowed me to experience discovery first hand by not giving me the answers. He let me bounce ideas off of him; then once I had direction, he let me wrestle with the problem. After the work was complete, Dr. Williams still gave me advice on how to achieve my goals. I can't imagine a professor that cares more about the success of his students.

Finally, I'd like to thank Brent Hargis from the Engineer Research and Development Center (ERDC) for sponsoring this research. More importantly, I would like to thank him for the great work he did during his time at Brigham Young University. His research was invaluable for establishing the direction this project would take.

المنارة للاستشارات

www.manaraa.com

TABLE OF CONTENTS

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

LIST OF TABLES

المنارة للاستشارات
www.manaraa.com

LIST OF FIGURES

www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

1 INTRODUCTION

1.1 Problem Statement

Some environmental processes create signals that can be observed in other processes that they influence. For example, analysis has shown that there is correlation of the El Niño Southern Oscillation (ENSO) process with precipitation and temperature datasets collected from gage data (Norden Eh Huang, 2014), that is the signal from the ENSO process can be observed in both precipitation and temperature signals. In the past to separate and identify such signals, scientists have analyzed data using wavelet analysis, and Fourier transforms (Norden E Huang et al., 1998). These methods, when applied to environmental data, have difficulties analyzing the non-linear, non-periodic signals that commonly occur in nature. Recently, researchers have extracted signals that correlate to the ENSO signal from precipitation data using a new analysis method called Empirical Mode Decomposition (EMD) (el-Askary, Sarkar, Chiu, Kafatos, & El-Ghazawi, 2004). This method works on signals that are non-linear and non-stationary (Norden E Huang et al., 1998). I applied Empirical Mode Decomposition (Norden E Huang et al., 1998) to extract and analyze trends and periodic processes in several environmental datasets including daily streamflow, daily precipitation, and daily temperature on two rivers in the Upper Colorado River Basin. I analyzed the ability of this technique to extract signals even after Flaming Gorge Dam, a very large dam on the Green River, significantly altered the processes.

www.manaraa.com

The Empirical Mode Decomposition method is a relatively new empirical signal processing technique that allows any time series data set, including non-linear and non-stationary datasets that are common in earth observation data, to be decomposed into a small quantity of composite finite data series called intrinsic mode functions (IMFs) (Norden E Huang et al., 1998). The EMD process also extracts a long-term trend in non-linear data sets that can provide insights into the effects of climate change or other long-term processes on the flow system. These trends and individual IMFs may or may not represent any physical processes.

The algorithm EMD uses is simple and does not assume stationary or linear signals. First, all local extrema are identified and a function to define the upper and lower envelopes is generated by interpolating between the local maxima and minima. Next, a temporary signal is created by interpolating an average of the upper and lower envelopes at every data interval (Kim & Oh, 2009). Next, the temporary signal is subtracted from the original dataset to obtain a partial IMF. The process is repeated until the number of extrema and zero-crossings differ by a maximum of one, the local average of the resulting equation must be zero. Once both requirements are satisfied, the partial IMF becomes a complete IMF, and the remaining data becomes a remainder. The IMF is then subtracted from the data set, and the processes is repeated. This process is known as sifting. I include a more comprehensive description in section 2.2.

1.2 Gaps in Research

Time series data can be used to analyze and understand the behavior of a processes over time using various signal processing techniques. Sometimes signal processing can reveal

المنارة للاستشارات
www.manaraa.com

valuable information by decomposing time series dataset into subsets. These subsets can identify independent processes or isolate and characterize the effects of a smaller component.

Historically, researchers have commonly used several tools for processing periodic signals including spectrograms, wavelet analysis, and the Fourier analysis. The resulting components from these analyses can sometimes represent simple real processes that can be combined to create the original signal. In other cases, these components, which are simpler than the original signal, do not represent any real processes, but are just artifacts of the mathematical decomposition. Many of these methods, such as wavelets or Fourier transform methods, require a number of conditions be met for application. Common conditions are linearity and stationarity. Each of these methods make assumptions that are not always valid when analyzing physical data (Norden E Huang et al., 1998).

Some of the challenges that arise when attempting to analyze physical data is that most real periodic environmental data are non-linear and non-stationary. Non-linear signals are defined as signals described by differential equations with time varying coefficients. Linear signals have constant coefficients. For example, the El Niño Southern Oscillation (ENSO) events have a periodic structure, but the structure period changes from 5 to 7 years on a seeming random basis. Non-stationary datasets have averages, variances, or covariances that change over time. Most environmental data sets have an underlying trend that makes them non-stationary.

EMD is a signal processing method that can process non-linear and non-stationary data and exactly recreates the original signal when the resulting component IMFs are summed. This is in contrast to approaches such as wavelets or Fourier analysis that require linear and stationary signals and only approximate the original signal when summing the resulting components. These methods, wavelets or Fourier analysis, can be only used on non-linear and non-stationary signals

المنارة للاستشارات
www.manaraa.com

by using a large number of components to represent the non-linear or non-stationary components over some finite signal length. This representation is only approximate, meaning the recreated signal, usually created by sum of the decomposed signal components, will not reproduce the original signal exactly.

By using EMD, environmental data sets with non-linear and non-stationary data can be processed. Most environmental data sets are non-stationary, that is they exhibit trends, and do not always meet other statistical assumptions required for non-empirical analysis. Because of the relaxed assumptions, EMD is a more mathematically sound approach to process environmental signals. However, a gap in research lies in the fact that, due to the relatively young age of EMD, signal processing with this method is not widely applied to streamflow or precipitation data.

1.3 Objectives

I evaluated whether I could use EMD methods to quantify and characterize signals from different environmental forcing functions and processes and determine if the signals were present in flow data from different locations that have similar watersheds but different amounts of anthropogenic control. Specifically, I will explore whether the EMD method can be used to analyze river flow data and identify various forcing functions such as dam operations, El Niño Southern Oscillation (ENSO), and also determine if the same periodic processes present in the flow data can be identified in local temperature or precipitation data. I evaluated whether processes like large reservoirs obscured or hid signals from nature processes, such as ENSO, or whether these signals were present in the flow data below the dam.

I selected the Green and Yampa Rivers as they have similar headwaters, but the Yampa has minimal diversions or controls while Flaming George Dam on the Green River significantly

www.manaraa.com

affects flow. This provides flow data from two different river regimes with similar large watersheds.

The objective of this study is to process several datasets, including flow from two rivers, one controlled and one mostly wild, along with temperature and precipitation records from the area comprising the watersheds of these two rivers. I will extract long-term trends and attempt to identify signals from natural processes. I will explore whether this method can be used to separate the effects of dams and other man-made regulatory structures from environmental processes. Once I have decomposed the original signals into independent, periodic IMFs, I will attempt to associate physical meaning, or physical processes, with selected IMFs. This may be difficult, since periodic processes within individual IMFs may or may not have physical meaning (Norden E Huang et al., 1998). However, if signals can be identified across the two rivers or different data sets, it is more likely that the IMF represents an actual process, either natural or anthropogenic

المنارة للاستشارات
www.manaraa.com

2 BACKGROUND

2.1 El Niño Southern Oscillation Data

The El Niño Southern Oscillation is an environmental phenomenon caused by bands of warm water that occur in the Pacific Ocean between the International Date Line and 120°W (Service, 2019). Every few years, El Niño affects the likelihood of floods in the western coast of South America, as well as the likelihood of drought in South East Asia. This climate pattern repeats on an approximate five- to seven-year period, and affects precipitation, and thus river flow, in western North and South America.

To understand what El Niño is, one must understand the normal climate pattern in the Southern Pacific Ocean. The winds in the Pacific Ocean travel from South America in the direction of Australia, this is a phenomenon known as the South East Trade winds, which occur between 30° latitude and the equator. The area is known for consistent winds of about 11 to 13 miles per hour. The South East Trade Winds are named after the ability to quickly move trading ships across the ocean and are caused by the Coriolis effect. As wind moves from east to west, the shear force moves the warm water near the ocean surface piling it on the western side of Australasia. Near South America this creates an upwelling, the rising of seawater in this upwelling region causes the cooler water to rise to the surface and replace the warm water being transported. This creates a temperature gradient where warm surface water is in Australasia, and cool surface water is near South America. The warm water increases air temperature, resulting in

المنارة للاستشارات
www.manaraa.com

rising air, cloud formations, and ultimately more rainfall. This pattern reinforces the wind circulations because warm moist air will rise in the west and cool, drier, air will descend near the Eastern part of the Pacific Ocean causing a steady state scenario. This is the normal weather pattern, with moisture being drawn away from the Americans, creating drier climate in the Americas (Trenberth, 1997) (Pizarro & Lall, 2002).

El Niño occurs when the trade winds are weakened. As trade winds are weakened, there is less transport of warm water, and less upwelling of cool water. Therefore, the centroid of the warm water moves eastward toward South America until the temperature gradient is eliminated, ultimately changing the precipitation and wind patterns in the Pacific Ocean. This results in increased precipitation in western North and South America.

La Niña occurs when trade winds are strengthened, moving or containing the warm water at the western part of the tropical Pacific; this process causes an increase of upwell in the Pacific Ocean. Air will rise over this warm area, but will dump precipitation near the center of the ocean. La Niña is the opposite of El Niño, in the fact that it will increase drought probability in South America, and flood probability in Australasia as shown in Figure 2.1-1 (NOAA, 2018).

The change of wind pattern can affect the likelihood of floods due to precipitation within the United States as shown in Figure 2.1-2 (NOAA, 2018).

The goal of my research was to determine if I could use EMD to analyze the environmental datasets and extract the signals that characterize the El Niño or La Niña events from within these periodic environmental processes, river flow, rainfall, and temperature in the presence of man-made control structures such as a large dam.

www.manaraa.com

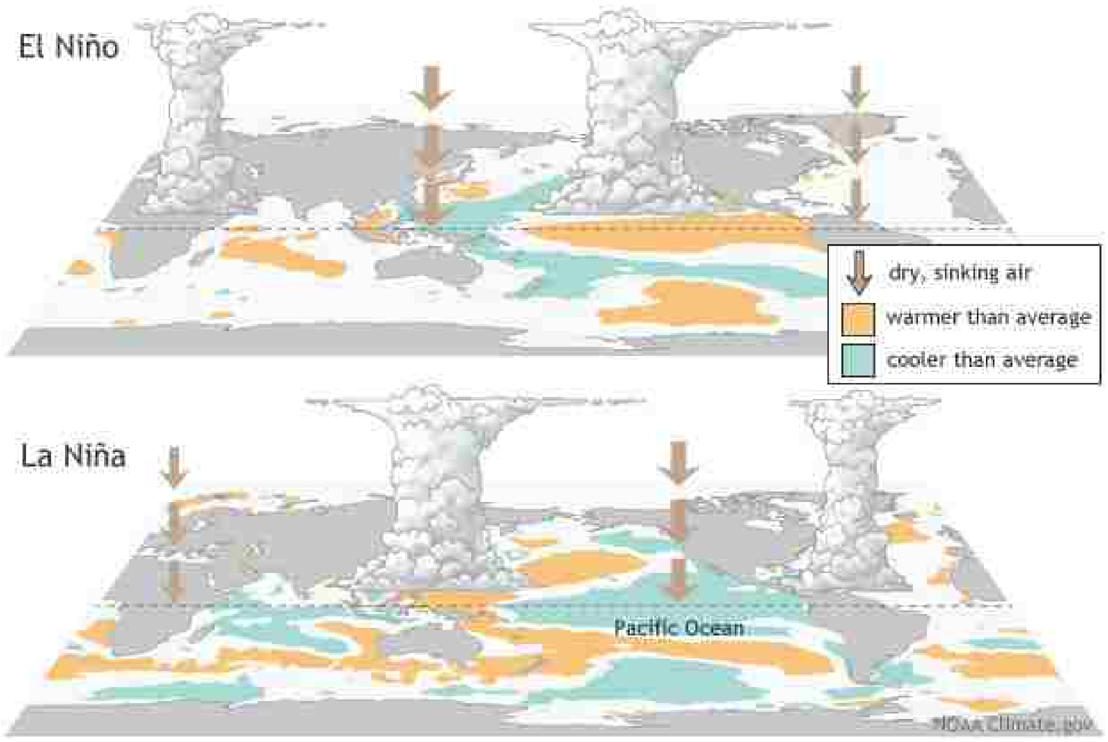


Figure 2.1-1 Precipitation Patterns Caused by El Niño and La Niña (NOAA, 2018)

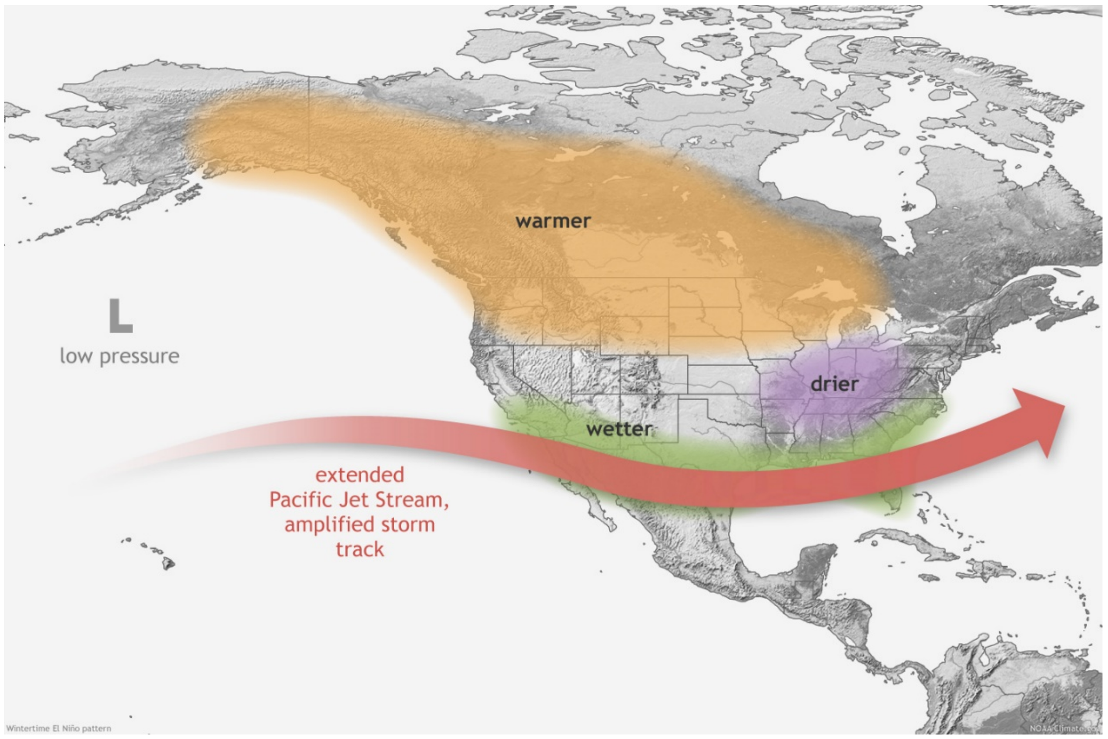

warmer
L
low pressure
drier
wetter
extended Pacific Jet Stream, amplified storm track
Wintertime El Niño pattern
NOAA Climate.gov

Figure 2.1-2 Precipitation Patterns Caused by El Niño in the United States (NOAA, 2018)

المنارة للاستشارات
www.manaraa.com

Quantitatively, El Niño and La Niña events are measured and standardized by the National Oceanic and Atmospheric Administration (NOAA) using the Oceanic Niño Index (ONI). The ONI is a running 3-month sea surface temperature (SST) mean in the Niño 3.4 region which is located at 5ºN-5ºS, 120º-170ºW. Values are based on a centered 30-year normal SST which is updated every 5 years. El Niño events are defined as 5 consecutive ONI periods with a value of +0.5º while La Nina events have a 5 consecutive ONI periods with a value of -0.5º. The SST measurements are obtained between 1 millimeter and 20 meters below the sea surface. The ONI measurements begin in December so that the first value will be centered in January (Service, 2019).

2.2 Proposed Method of Time Series Analysis

2.2.1 Empirical Mode Decomposition

The EMD method is a relatively new empirical technique that allows any time series data set, including non-linear and non-stationary datasets, that are common in earth observation data, to be decomposed into a small quantity of composite finite data series called IMFs (Norden E Huang et al., 1998). The EMD process also extracts a long-term trend in non-linear data sets that can provide insights into the effects of climate change or other long-term processes on the flow system. These trends may or may not represent any physical processes. The algorithm EMD uses is simple and does not assume stationary or linear signals.

In the EMD algorithm, first all local extrema are identified, both local maximums and minimums (Figure 2.2.1-1a). In the next step, the algorithm interpolates a curve through the local maxima or minima to create the upper and lower envelopes (Figure 2.2.1-1b). For this task I used the piecewise cubic hermite interpolating polynomial (PCHIP) algorithm to interpolate

www.manaraa.com

the local maxima and minima. Next, the average of these two envelopes is computed, by interpolating the upper and lower envelopes at every data interval (Figure 2.2.1-1c). This average is subtracted from the original dataset to obtain a partial IMF (Figure 2.2.1-1d). The process is repeated until the number of extrema and zero-crossings differ by one and the local average is zero (Figure 2.2.1-1e). Once both requirements are satisfied, the partial IMF becomes a complete IMF, and the remaining data becomes a residual and the process starts again (Figure 2.2.1-1f). This process is known as sifting.

After an IMF is created, it is subtracted from the original data and the process is repeated on the remaining data until the stopping rule is satisfied. The stopping rule is a convergence criterion that stops sifting when the current relative tolerance is less than the specified tolerance or a maximum number of IMFs are created. Once the stopping rule is satisfied, the remainder becomes the residual, which describes the dataset's long-term trend.

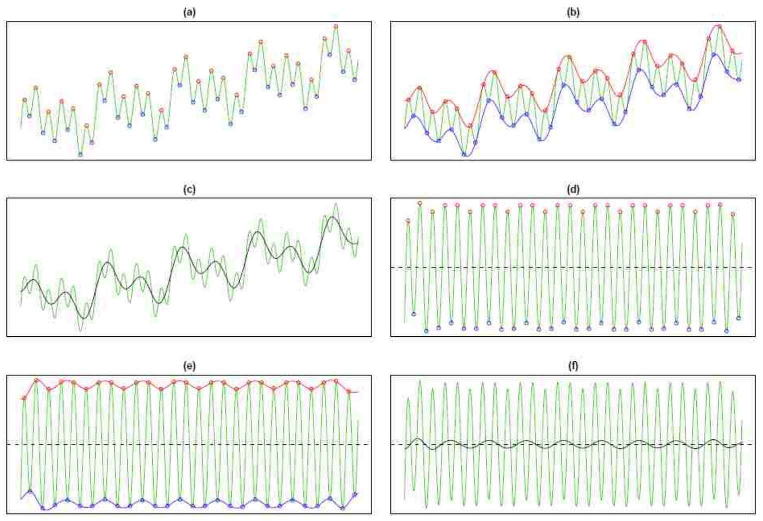


Figure 2.2.1-1: Sifting/Decomposition Process Visualized (Kim & Oh, 2009)

www.manaraa.com

I applied this technique to flow, precipitation, and temperature gage datasets within the area of interest to highlight trends, and attempted to assign a physical explanation to constant frequencies within oscillations. To ensure accurate results, I included only data sets with over 60 years of data in these calculations.

When EMD is applied to environmental data without a boundary condition, there are a myriad of outcomes based on the sifting tolerance. Due to human influence causing drastic changes in river flow, I used data from the stream gage at Flaming Gorge to select a value of 0.2 to use as the sifting tolerance. This tolerance of 0.2 is also suggested by Dr. Haung (Norden E Huang et al., 1998). Setting the sifting tolerance too low will overdecompose the signal making it likely to lose any physical meaning, and making the tolerance too high will superimpose signals that potentially are different (Norden E Huang et al., 1998). Haung suggests this value normally be between 0.2 and 0.3, therefore the value I selected is in the recommended bounds. I selected this tolerance value because the resulting residual most accurately depicted historical markers such as the filling of Flaming Gorge Reservoir in 1963, the move from peak hour flooding to constant flow for power generation in 1992, and the move to a more natural flow pattern that increased river flow in 2006. I did not set a limit on the number of IMFs created, letting the algorithm run to completion.

2.2.1 Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)

The EMD code requires an interpolation method to generate the data envelopes. I used PCHIP during IMF generation. Most EMD code uses cubic spline interpolation, but for environmental data, cubic spline methods can result in significant under and overshoot in the interpolated data. The results of PCHIP are a shape-preserving piecewise cubic interpolation

www.manaraa.com

that does not overshoot or undershoot as other data interpolators. A simple example of why this interpolator was chosen over others is shown in Figure 2.2.1-1 and Figure 2.2.1-2.

```
1 x = -3:3;
2 y = [-1 -1 -1 0 1 1 1];
3 xq1 = -3:.01:3;
4 p = pchip(x,y,xq1);
5 s = spline(x,y,xq1);
6 plot(x,y,'o',xq1,p,'-',xq1,s,'-.')
7 legend('Sample Points','pchip','spline','Location','SouthEast')
```

Figure 2.2.1-1 Sample Code Creating Data Comparing Overshoot and Undershoot of the Spline Method to the PCHIP Method Which Does Not Under or Overshoot the Data Envelope

Line 1 of the code in Figure 2.2.1-1, sets a value of x from -3 to 3 by a step of 1, while line 2 is a set of random y values. The variable xq1 is a query point for our interpolation, or the x-coordinates at which Matlab will apply the interpolation function. Figure 2.2.1-2 shows overshoot from the spline interpolation from -3 to -2 and from 1 to 2. The spline interpolation experiences undershoot from -2 to -1 and from 2 to 3. However, the PCHIP interpolation connects our data without any visible overshoot or undershoot. Overshoot and undershoot not only introduce error, but can obscure physical meaning in datasets such as flow where negative flow has no meaning.

المنارة للاستشارات
www.manaraa.com

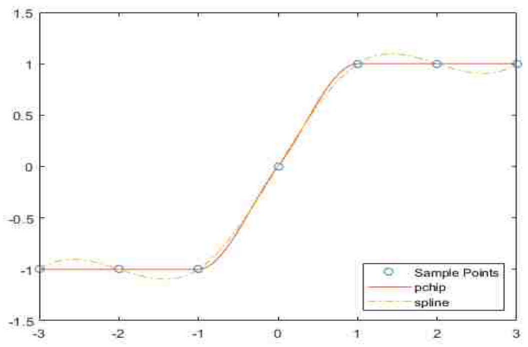


Figure 2.2.1-2 Plotting of Sample Data Comparing Overshoot and Undershoot of the Spline Method and How the PCHIP Method Does Not Over or Undershoot the Data Envelope

2.2.2 Hilbert-Haung Transform

The Hilbert-Haung Transform (HHT) is a computational method used to solve for the instantaneous frequency changes within one oscillation cycle (Norden Eh Huang, 2014). The HHT gives insight to the dominant signal frequency while allowing a non-constant phase. Paired with the EMD, the HHT analyzes data by processing the IMF subsets through the HHT algorithm. Like EMD, the HHT method can be used on non-linear data, which allows the method to be adaptive. However, the adaptive nature of the HHT makes it difficult to describe theoretically (Norden Eh Huang, 2014). The HHT can be used to generate frequency plots similar to spectrograms for IMF data.

2.3 Common Methods of Environmental Time Series Data Analysis

2.3.1 Fourier Transforms

The Fourier Transform is the most common transformation for time series data analysis. This transformation method assumes that any signal can be transformed into a summation of

المنارة للاستشارات
www.manaraa.com

sinusoidal functions (Bracewell & Bracewell, 1986). These sinusoidal functions have a stationary basic period with time independent frequency. This means that the Fourier Transform is only able to process stationary data. Perhaps its greatest limitation of the Fourier Transform is that approximations are used to model datasets, using sinusoidal functions. When the sinusoidal decompositions are used to reconstruct the original signal, the result contains approximation errors and does not exactly represent the original dataset.

2.3.2 Spectrogram Method

A spectrogram method of analysis uses the Fourier decomposition as its base. It implements a limited time window-width Fourier spectral analysis (Stankovic, 1994) (Norden E Huang et al., 1998). By continuously plotting frequencies along the time axis, you can compute and visualize a time-frequency distribution. Spectrograms are often used to visualize time series data, emphasizing amplitude, intensity, and overtones.

Spectrogram analysis relies on the same assumption of Fourier spectral analysis: that data are piecewise stationary. This implies that the window of analysis perfectly captures stationary time scales. However, this assumption is violated when analyzing non-stationary, non-linear signals.

2.3.3 Wavelet Transforms

The Wavelet Transform is another relatively recent signal processing method. Similar to a Spectrogram, the wavelet analysis uses an adjustable window that fits a repeating function to the data at different scales, similar to the way Fourier analysis fits sine functions at different scales (Labat, 2008). This approach is useful for analyzing gradual changes in frequency within data as, unlike Fourier analysis, wavelets can represent these frequencies changes. Most

المنارة للاستشارات
www.manaraa.com

applications of the Wavelet Transforms have been used in image compression and edge detection (Norden E Huang et al., 1998). Though there have been applications to river flow (Coulibaly & Burn, 2004; Melesse, Abtew, Dessalegne, & Wang, 2010; Smith, Turcotte, & Isacks, 1998).

Wavelet analysis has difficulty adapting over time for environmental data, as it is not designed for non-linear data. Because the wavelet analysis is based on a summation of functions, it inherits some of the shortcomings of the Fourier Transform Analysis.

2.4 Summary of Signal Analysis Methods

A summary of the different transform methods can be seen in Table 2.3-1 (Norden E Huang et al., 1998).

Table 2.3-1: A Summary of Different Signal Processing Methods and Descriptions

	Fourier	**Wavelet**	**Hilbert**
Basis	*a priori*	*a priori*	adaptive
Frequency	convolution: global uncertainty	convolution: regional uncertainty	differentiation: local, certainty
Presentation	energy-frequency	energy-time-frequency	energy-time-frequency
Nonlinear	No	No	Yes
Nonstationary	no	yes	yes
Feature Extraction	no	discrete: no; continuous: yes	yes
Theoretical base	theory complete	theory complete	empirical

www.manaraa.com

3 SOFTWARE RESOURCES USED

3.1 MATLAB

MATLAB® is a high level programing language and numerical analysis environment developed by Mathworks (MathWorks, 2019). MATLAB is commonly used by engineers and scientists to develop algorithms to solve mathematics problems. I chose MATLAB as the primary software for developing the EMD code due to its built-in Empirical Mode Decomposition functions released with the MATLAB 2018R package.

3.2 ArcGIS Pro

ArcGIS is a geographic information system (GIS) developed by ESRI (ESRI, 2019). Its primary application in this research was to compile geospatial information such as gage locations and gage types, and visually present it as a map.

3.3 Excel

Excel is a spreadsheet program developed by Microsoft as part of the Microsoft Office Suite (Microsoft, 2019). The program is arranged in rows and columns that be used to input data and manipulate those data mathematically. I used Excel primarily to organize data before inputting the datasets into MATLAB.

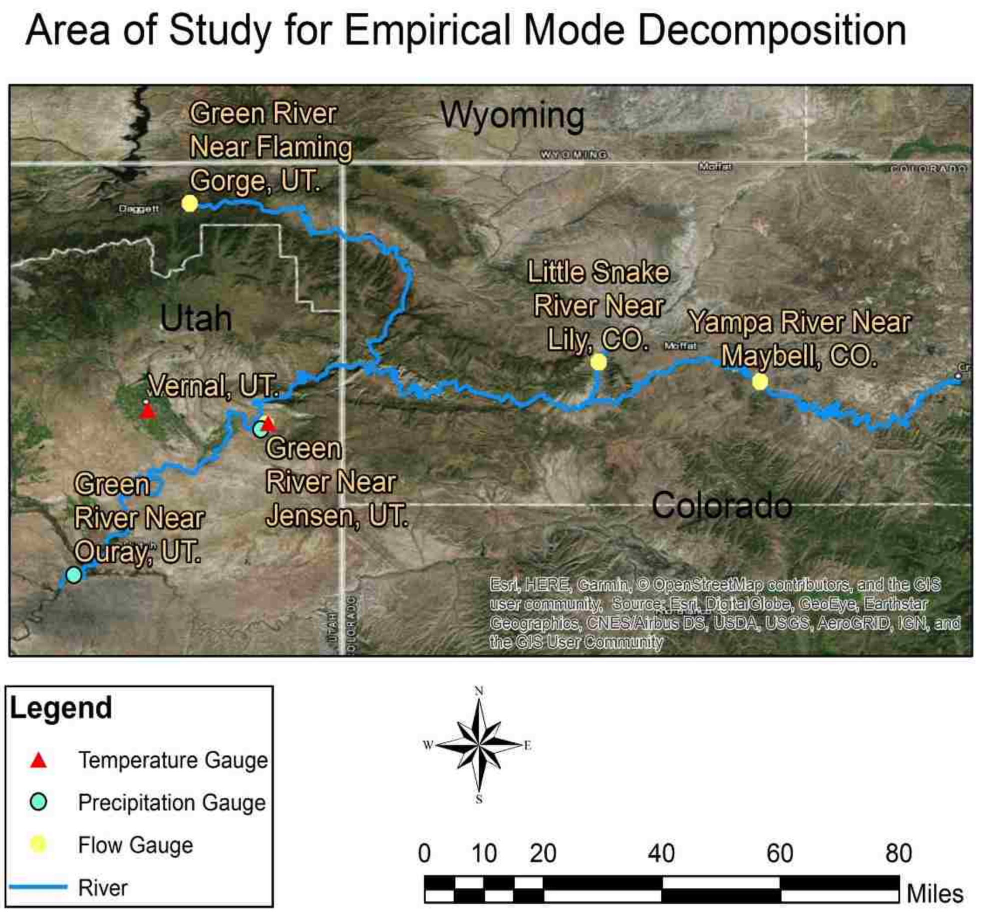
www.manaraa.com

4 IMPLEMENTATION

4.1 Study Area

For this project, I analyzed data from several stream gages along the Green and Yampa Rivers. Data were provided by the United States Geological Survey (USGS) (Survey, 2019). My original selection of gages was: Flaming Gorge, Green River near Jensen Utah, Green River near Ouray, Yampa near Deer Lodge Colorado, Little Snake near Lily Colorado, Yampa, near Maybel Colorado, Yampa near Craig Colorado, Yampa above Elk Head near Hayden Colorado and Yampa river near Elk Head (Figure 4.1-1). I selected these gages because they characterize flow in the Green and Yampa Rivers, and flows in the Green River after the confluence where the Yampa joins the Green. As noted, these rivers have similar headwaters, but the Yampa has minimal diversions or controls while Flaming George Dam on the Green River significantly affects flow. This provides two different flow regimes with similar large watersheds. The reason for selecting these gages is to evaluate if the rivers, before joining, exhibit different periodic processes, if these processes can be correlated with periodic processes extracted from the temperature or precipitation data, and if these processes can be extracted from the data after the confluence of the two rivers. I targeted the El Niño processes with a period of 5 to 7 years to determine if I could identify these processes in the data from these gages and if these signals could be tracked downstream.

المنارة للاستشارات
www.manaraa.com



Area of Study for Empirical Mode Decomposition
Green River Near Flaming Gorge, UT.
Wyoming
Utah
Little Snake River Near Lily, CO.
Yampa River Near Maybell, CO.
Vernal, UT.
Green River Near Jensen, UT.
Green River Near Ouray, UT.
Colorado
Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS user community, Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community
Legend
Temperature Gauge
Precipitation Gauge
Flow Gauge
River
N
S
E
W
0 10 20 40 60 80
Miles

Figure 4.1-1 Overview of Study Within the Yampa River and Green River Watersheds

I also analyzed precipitation data from gages within the watershed at Flaming Gorge, Jensen, Ouray, Maybel, and Craig (Figure4.1-1). I selected precipitation data from meteorological stations as close to the flow area of study as possible. Precipitation data were obtained through the Utah Climate Center, which provided the Vernal and Ouray datasets (Center, 2019).

Another interesting set of data that I analyzed was air temperature. Temperature was more difficult to obtain because there are not many temperature gages located in my study area. I selected temperature gages at Vernal, Utah; Grand Junction, Colorado; Jensen, Utah; and Flaming Gorge Dam, Utah (Center, 2019).

المنارة للاستشارات
www.manaraa.com

4.2 Data Collection

Stream discharge data were obtained from the USGS database (Survey, 2019). Precipitation and Temperature data were obtained from the Utah Climate Center of Utah State University (Center, 2019). For a dataset to be included in my analysis, they needed to meet a set of criteria. I developed these criteria based on previous research to ensure data quality data and to minimize missing data (Hargis, 2014).

Criteria for dataset (e.g., site) selection: Each site must have

- Approximately 60 years or more of data available
- Daily data
- Must be from a valid or accredited source

Discharge gage data are presented as daily averages, in cubic feet per second. Temperature data were obtained as daily max and minimum and calculated as daily average in degrees centigrade. Precipitation data were obtained as daily average values, but I calculated yearly precipitation data in millimeters per year to provide mathematical continuity during decomposition. A summary of the datasets is provided in Table 4.2-1.

Managing and choosing a precipitation data reduction method was challenging. Before I chose total yearly precipitation depth as my method for processing, I used total quarterly precipitation to satisfy the continuity requirement of EMD. When I plotted and analyzed the data using EMD, this method did not show the expected results due to rapidly changing crests and troughs in the signal. This was because the precipitation data were so sparse and precipitation in a given quarter could be very low and the data can appear to be noisy. Results from the quarterly

المنارة للاستشارات
www.manaraa.com

EMD datasets are shown in Appendix B. Based on this preliminary effort, I used total yearly precipitation as the reduction basis for this data set.

Table 4.2-1: Summary of Data

Dataset	Gage Name	Beginning Date	Ending Date	Appendix
Stream Discharge	Flaming Gorge Flow	10/1/1946	3/15/2019	B.1
	Jensen Flow	10/1/1946	5/16/2018	B.2
	Little Snake Lily Flow	10/1/1921	5/2/2018	B.3
	Yampa Maybel Flow	5/1/1916	5/21/2018	B.4
Yearly Precipitation	Ouray Yearly Precipitation	6/1/1941	5/31/2018	B.7
	Jensen Yearly Precipitation	3/27/1925	5/31/2018	B.8
Temperature	Jensen Temperature	10/1/1939	6/15/2018	B.9
	Vernal Temperature	7/27/1900	4/30/2018	B.10

4.3 Missing Data Analysis

For the EMD algorithm to properly work, continuous data must be provided. However, there were several gaps from missing data in the temperature and precipitation datasets. Previous work, showed that missing values with gap periods up to a year have little impact on the overall EMD/IMF interpolation (Hargis, 2014). For missing daily values in the temperature dataset, I linearly interpolated between missing days. Missing daily values in the precipitation dataset were assumed to be zero, when entire years were missing, the yearly precipitation was linearly interpolated. Linear interpolation is the simplest and often least accurate method for replacing these data, however, previous research showed that these shorter gaps, less than 1-year, have

www.manaraa.com

limited impact when filled with linear data. Because of these minimal impacts, I selected linear interpolation for the ease of implementation. Interpolation introduces error into the data sets because the interpolated data do not represent reality, however, since the signals I am attempting to analyze have a period of 5 to 7 years, small gaps in data do not affect the analysis. Table 4.3-1 describes the amount of missing data within this study.

Table 4.3-1: Table of Missing Values

Dataset	Gage Name	Beginning Date	Ending Date	Missing Values	Completeness
Stream Discharge	Flaming Gorge Flow	10/1/1946	3/15/2019	0	100%
	Jensen Flow	10/1/1946	5/16/2018	0	100%
	Little Snake Lily Flow	10/1/1921	5/2/2018	0	100%
	Yampa Maybel Flow	5/1/1916	5/21/2018	0	100%
Yearly Precipitation	Ouray Yearly Precipitation	6/1/1941	5/31/2018	5329	81%
	Jensen Yearly Precipitation	3/27/1925	5/31/2018	3659	89%
Temperature	Jensen Temperature	10/1/1939	6/15/2018	557	98%
	Vernal Temperature	7/27/1900	4/30/2018	4035	91%

The IMF's show the gap, with each subsequent IMF also showing the gap from the original dataset. Essentially, then analyzing the resulting IMFs, the values in the gap regions are not valid, these effects extend a little into the actual data.

I compared IMFs resulting from decomposing the full data set (large gap at the beginning) and a shortened data set (does not include the gaps). Figure 4.3.1-3 shows the IMFs that result from decomposing a shortened Ouray dataset that does not include the first three years of data or the 9 years of missing data. Figure 4.3.1-3 shows, by comparing the long dataset and the

المنارة للاستشارات
www.manaraa.com

shortened dataset, several differences. First, the shape of the IMF near the missing data is affected as seen in IMF 3 of the shortened dataset. Next, the magnitudes vary slightly, which can be seen by comparing IMF 2 and 3 of the long and short datasets. Additionally, the rate of change and concavity of the two residuals differ. The concavity of the long dataset residual may suggest that an event occurred causing a peak in the data. The short dataset would not mislead researchers to make that assumption.

Overall, I suggest that datasets not include early data if large gaps exist since the mathematical error introduced into the IMFs may not be worth the few more years of additional data. If gaps bisect a dataset, a method of interpolation could be a viable solution to save physical meaning (Nelsen, Williams, Williams, & Berrett, 2018). While I used linear interpolation, since my gaps were relatively small compared to the periods I was interested in, more advanced data imputation methods could be used.

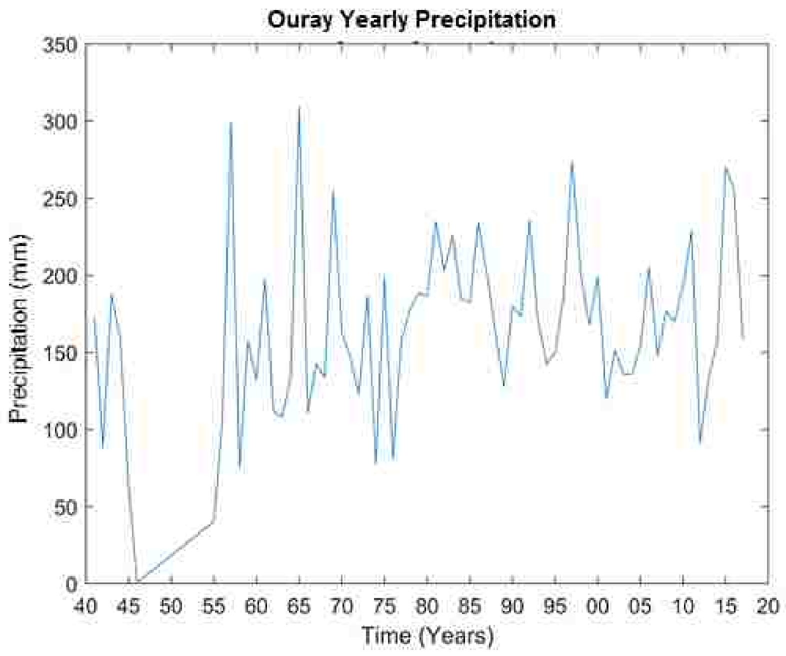


Figure 4.3.1-1: Complete Ouray Precipitation Dataset

www.manaraa.com

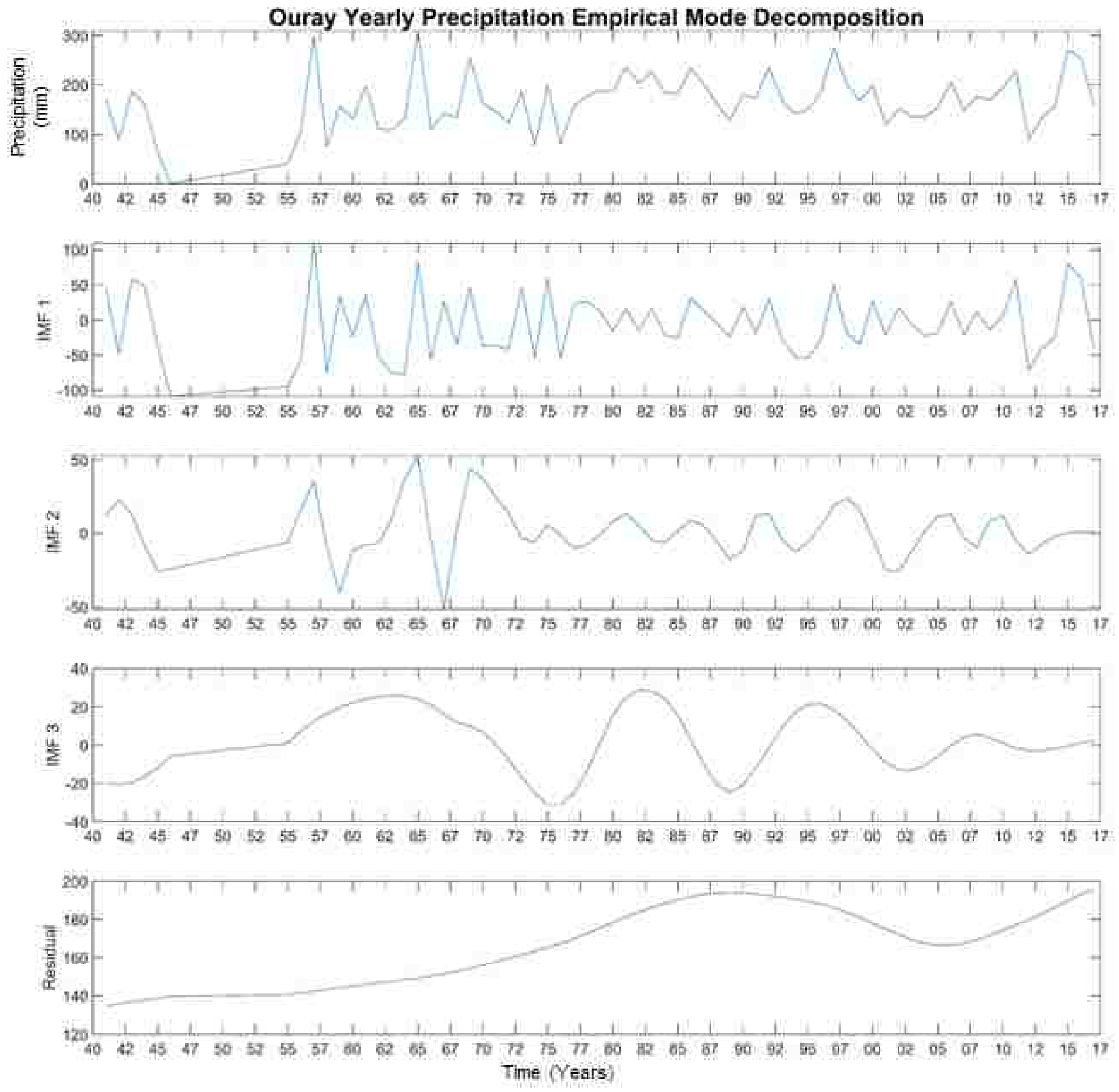


Figure 4.3.1-2: Complete Ouray Precipitation Dataset Decomposition Showing Missing Data Gaps Being Transferred from the Original Dataset to the IMFs and Then the Residual

www.manaraa.com

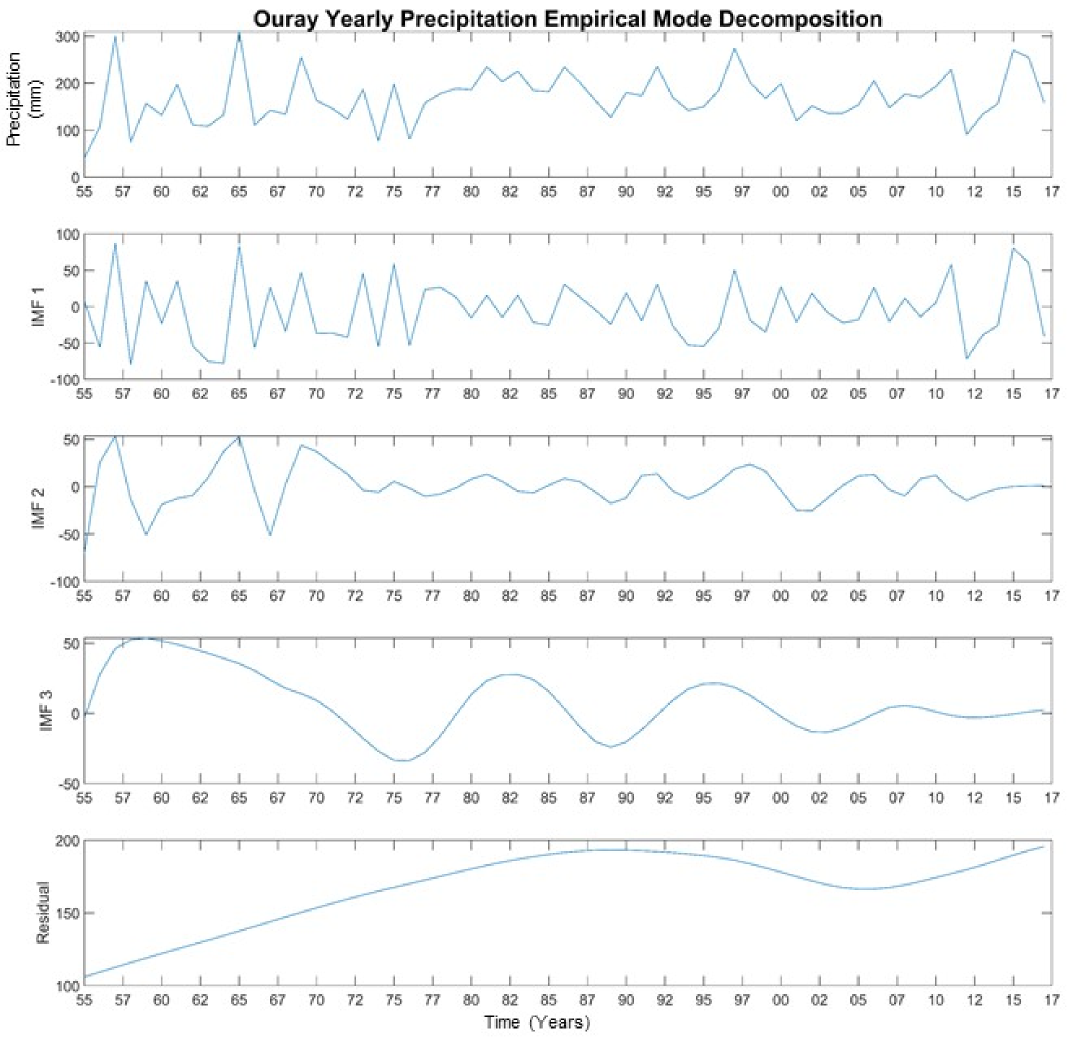


Figure 4.3.1-3: Shortened Ouray Precipitation Dataset Decomposition Ignoring the 9 Years of Missing Data and Three Years of Corresponding Data Before the Gap.

www.manaraa.com

5 MATLAB PROCESSING

5.1 MATLAB Code

I created custom MATLAB functions for data processing and analysis. I created MATLAB scripts to customize figures and generate data reports. This code is contained in the appendix. The MATLAB scripts and code are generally divided into three groups: Data Readers, Data Processors, and Plotters.

5.1.1 Data Readers

The Data Readers are the first part the EMD analysis process. The Data Reader scripts convert previously prepared Excel files (.csv) into MATLAB Data (.mat) files. MATLAB data files, significantly speedup IMF and HHT processing since MATLAB can load a ".mat" file much faster than reading a ".csv" file.

First, the Data Reader creates two string matrices containing names of the data used in the labeling and storage of data; a sample of the organization of each vector is shown in figure 5.1.1-1 and figure 5.1.1-2. The first string matrix is containing names of gages and file names. The second matrix contains gage names and file locations used for loading and reading data.

Next, the Data Reader references the string vectors to create a data structure which stores dates and associated measurements (Figure 5.1.1-3). In addition to the measurement data, this

www.manaraa.com

data structure will store the IMFs and Residual data after processing. The complete Data Reader code is located in the appendix.

Table 5.1.1-1: The First Matrix Contains a Column for Gage Name, Filename, and the 3-Letter Acronym for the Gage

String Vector 1		
Column1	**Column2**	**Column3**
Long Name	File name	Shortened Code Name

Table 5.1.1-2: Matrix 2 is Organized as a Column of Gage Name, Data Location, and the 3-Letter Acronym for the Gage to Ease Coding

String Vector 2		
Column1	**Column2**	**Column3**
Long Name	Exact Data Location	Shortened Code Name

Table 5.1.1-3: Data Reader Input File and Structure

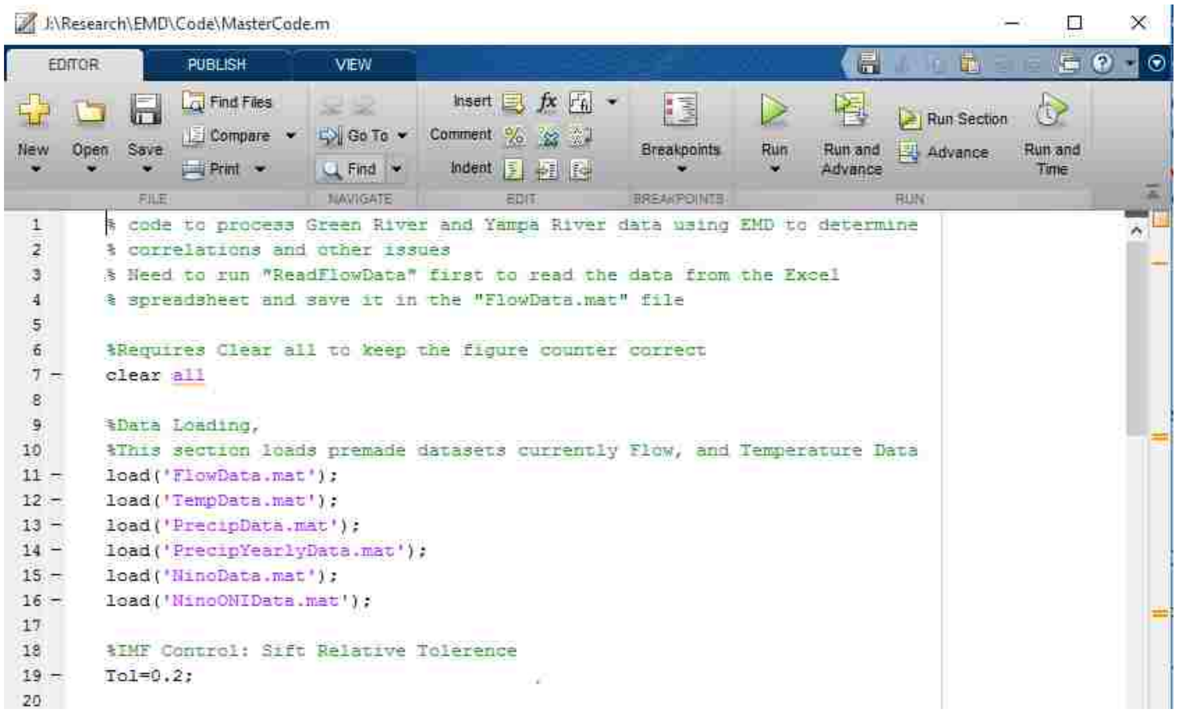
Data Structure			
Column1	**Column2**	**Column3**	**Column4**
Dates	Associated Data	IMF	Residual

5.1.2 Data Processors

The Data Processor, located in a MATLAB script called “MasterCode”, implements the EMD processing and generates the associated IMFs and HHTs. The code loads the previously

المنارة للاستشارات
www.manaraa.com

made ".mat" files, and creates a series of labels used for plotting (lines 11-16). Line 24 creates a matrix containing labels for plotting. The code then creates a series of plots. These plots are used as a check in order to see that loaded data matches the data from the Excel files. The code then decomposed each time series into IMFs and a Residual. These results are stored in the Data Structure and associated with the gage site and data dates. Next, the code analyzes the IMF and Residual signals with the HHT to generate time-frequency plots. Lines 44-70 create the IMFs, and HHTs for discharge data. I repeated this three-step process for each of the four data sets: Stream Discharge, Temperature, Quarterly Precipitation, and Yearly Precipitation.



```
1      % code to process Green River and Yampa River data using EMD to determine
2      % correlations and other issues
3      % Need to run "ReadFlowData" first to read the data from the Excel
4      % spreadsheet and save it in the "FlowData.mat" file
5
6      %Requires Clear all to keep the figure counter correct
7  -   clear all
8
9      %Data Loading,
10     %This section loads premade datasets currently Flow, and Temperature Data
11 -   load('FlowData.mat');
12 -   load('TempData.mat');
13 -   load('PrecipData.mat');
14 -   load('PrecipYearlyData.mat');
15 -   load('NinoData.mat');
16 -   load('NinoONIData.mat');
17
18     %IMF Control: Sift Relative Tolerence
19 -   Tol=0.2;
20
```

Figure 5.1.2-1 Mastercode, Generates the Associated IMFs and HHTs and Creates a Series of Labels Used in Plotting

5.1.3 Data Plotters

I wrote three Plotting Scripts to analyze data: CustomPloter, Residualcode, and DataNormalizer. I used each of these scripts to analyze data in different ways.

www.manaraa.com

The CustomPlotter script, allows the user to select a data set, the IMFs of interest and plot just that part of the data. Some of the features of this script include:

- Ability to Superimpose IMFs, as well as Residuals
- Extract examined data and create a .mat file from it
- Compare selected data to El Niño using but the ONI method and Magnitude Method
- Save the figure

Figure 5.1.3-1 Customplotter Input Settings

The Residualcode allows for the plotting of data residuals from respective datasets. It also has the ability to compare residuals across the four datasets. The DataNormalizer is a companion script to Residualcode that normalizes data between -1 and 1 in order to give better idea of the rate of change that is occurring in the residuals since flow residuals tend to be much greater than the other dataset's residuals

www.manaraa.com

6 CHARACTERIZING THE RESULTS

6.1 Discussion of Results

I applied EMD to all the listed datasets to qualitatively identify a relationship between IMFs and the El Niño ONI dataset. In these results, I present specific data sets which show extracted signals from the El Niño weather pattern in the Discharge, Temperature, and Precipitation datasets within the study area.

The standard presentation of EMD components is to present the original signal at the top of a figure, with subsequent IMFs below, and finally the long-term residual at the bottom in vertical format. In this paper, the original signal will be presented in the figure, followed by the complete decomposition in the next figure, and selected IMFs of interest in the final figure. This pattern will be repeated for different data types and locations.

6.2 Discussion of Stream Discharge Results

Figure 6.2-1 shows discharge data from Flaming Gorge Dam. From examining the dataset, several changes in the reservoir operations can be seen. The downstream stream gauge shows the highly varied natural flow of the Green River. Construction of the dam began in 1958 and the Green River's diversion tunnel was completed by November 19, 1959. The Flaming Gorge Dam began operation after its completion in 1964. The period after the dam was completed until the mid-1990 represent a period where the dam was operated primarily to provide peak power,

www.manaraa.com

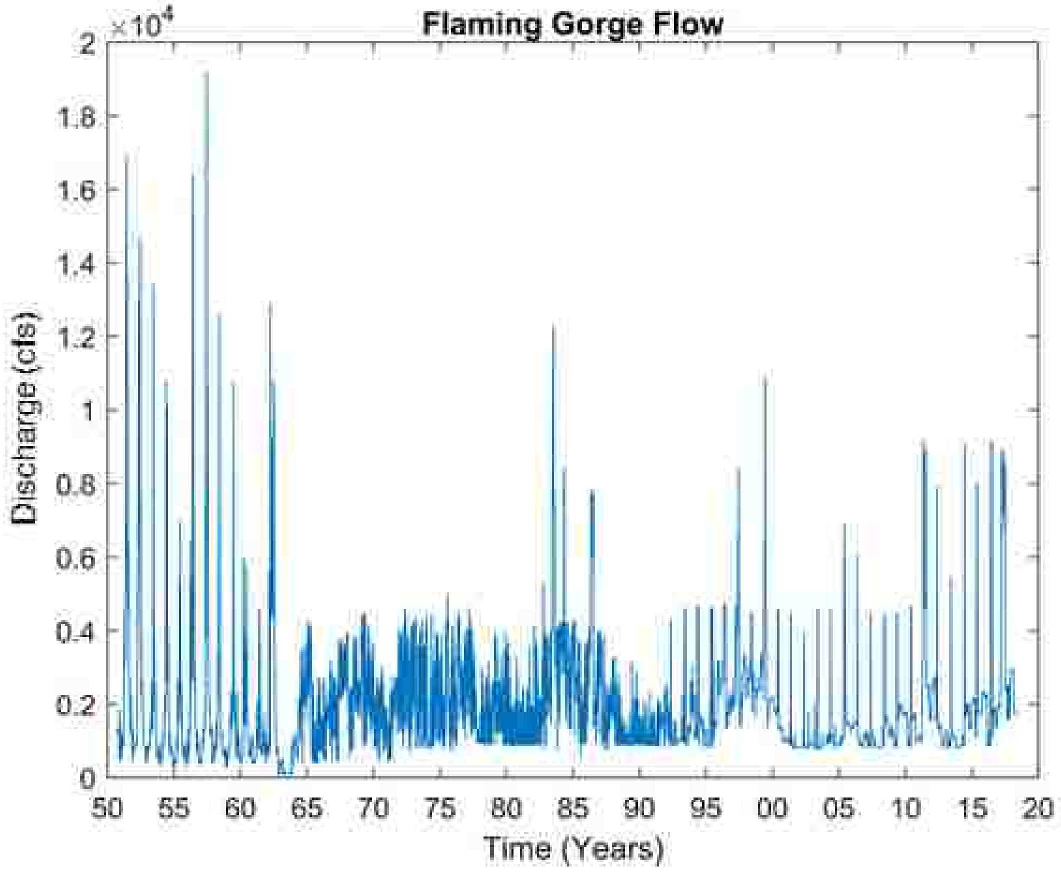


Figure 6.2-1: Flaming Gorge Stream Discharge Data

Figure 6.2-2 shows the EMD decomposition of the Flaming Gorge Dam discharge dataset that contains 11 IMFs and the residual. IMFs 7, 8 and 9 showed temporal patterns similar to the ONI. I summed these three IMFs and compared the resulting signal to the El Niño ONI dataset. The second to the bottom panel of Figure 6.2-3 shows the summed IMFs, while the bottom panel provides a visual comparison between the summed IMF's and the El Niño ONI dataset. Qualitatively the ONI dataset matches the summed IMF.

Visually there is good correlation between peaks in the ONI dataset and peaks in the composite IMF (Figure 6.2-3). The ONI often peaks before the IMF. This intuitively makes

www.manaraa.com

sense, if a wet year is caused by El Niño, more water will come into the reservoir, therefore, more water will be discharged from the dam the following year. This reasoning suggesting that the El Niño signal within the streamflow IMFs is present despite the structure, and operational influences. La Niña is not as easily seen in the composite IMF, though traces are present. A possible explanation for this is that La Niña is often a drier year and the reservoir is discharging minimum flow, or being filled, partially hiding the presence of La Niña signal in the release data and thus in the composite IMF.

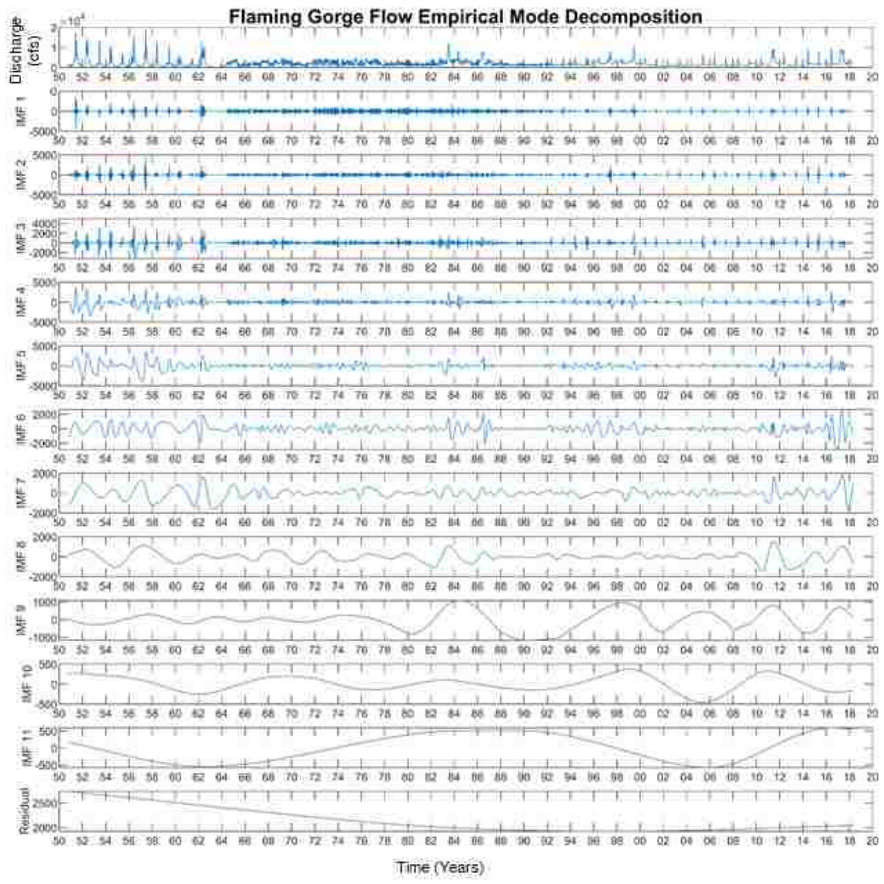


Figure 6.2-2: Flaming Gorge Stream Discharge Complete EMD

المنارة للاستشارات
www.manaraa.com

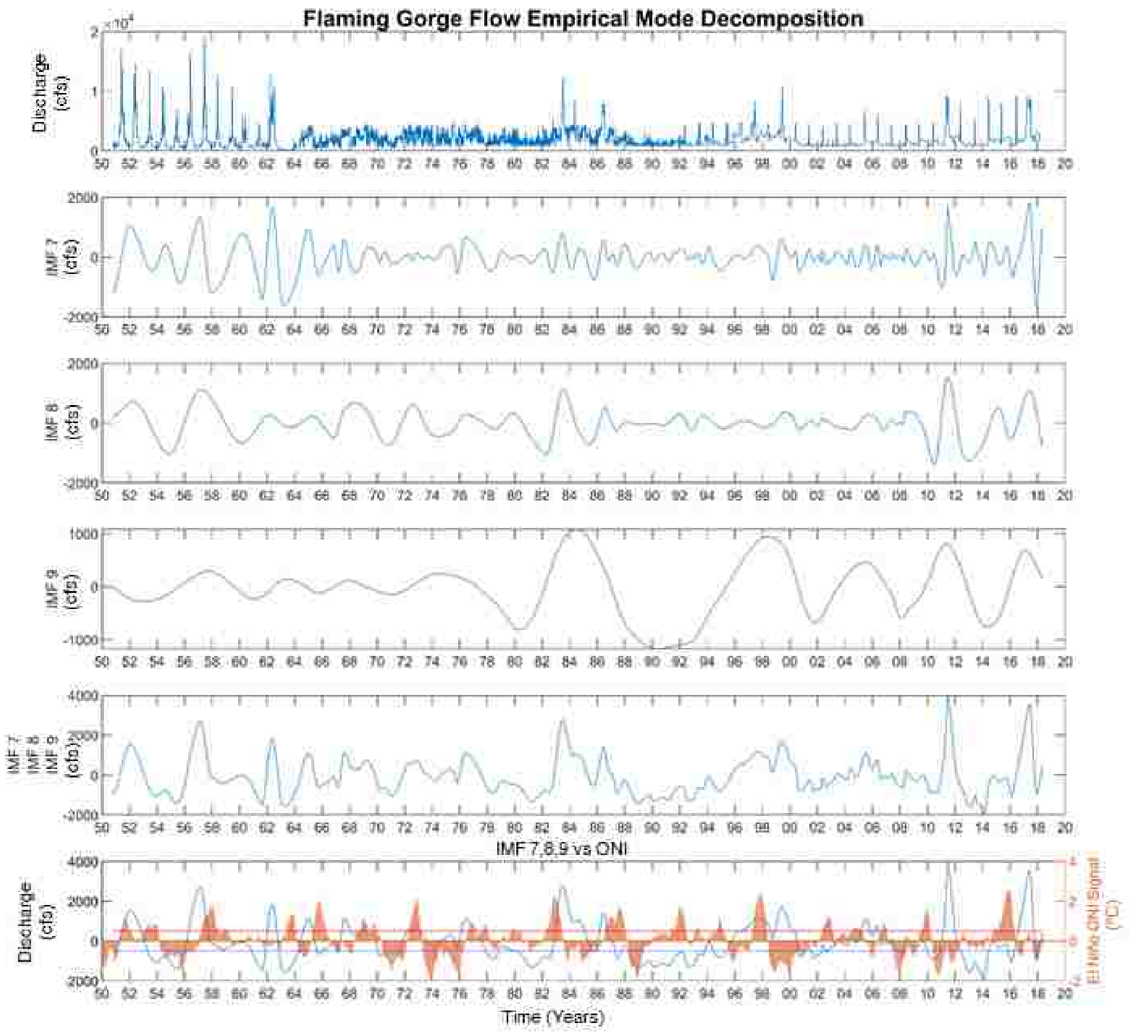


Figure 6.2-3: Flaming Gorge Stream Discharge IMF 7, IMF 8, IMF 9 Super Imposed and Compared to El Niño ONI (Bottom Panel)

The signal can be tracked downstream of Flaming Gorge Dam. One example of this is at the Jensen gage. This gage is important because it is approximately 40 miles south of the confluence of the Green River and the Yampa River and over 50 miles downstream of Flaming Gorge Dam. I created a composite IMF from IMF 7 and 8 (Figure 6.2-4). This composite IMF is similar to the ONI dataset suggesting that the El Niño signal that is passed through Flaming

www.manaraa.com

Gorge Dam is preserved, even after mixing with the Yampa River influx. It may also be slightly strengthened, as the Yampa River is not impacted from dam operations and the signal is less attenuated.

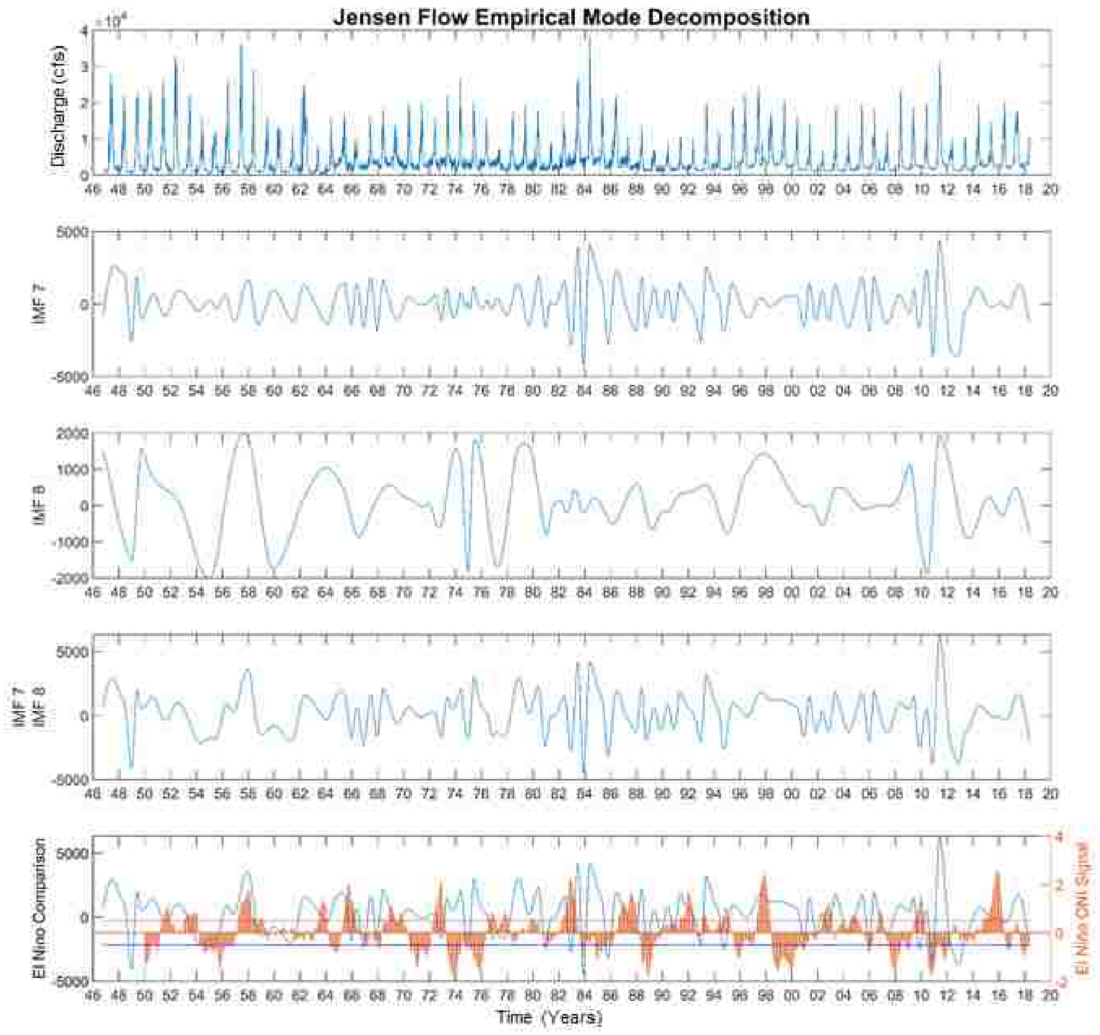


Figure 6.2-4: Jensen Streamflow Discharge IMF 7, IMF 8 Super Imposed and Compared to El Niño ONI (Bottom Panel)

www.manaraa.com

6.3 Discussion of Temperature Results

The temperature dataset at Vernal was the longest dataset analyzed (Figure 6.3-1). In preparation for the EMD process, I arithmetically averaged daily high and low temperatures. I then decomposed these data as shown in Figure 6.3-2. Figure 6.3-3 shows a correlation between temperature variations and the ONI index. This relationship is interesting since both datasets correspond to variations of temperature. The figure shows that an increase in the ONI corresponds with increased temperatures in the composite dataset of IMF 9 and 10.

Figure 6.3-4 is a comparison of the Vernal and Jensen temperature dataset residuals. It shows that the residuals represent an increase of approximately 1 degree centigrade over the last 35 years and evidence of a cooling during the mid-70's –80's.

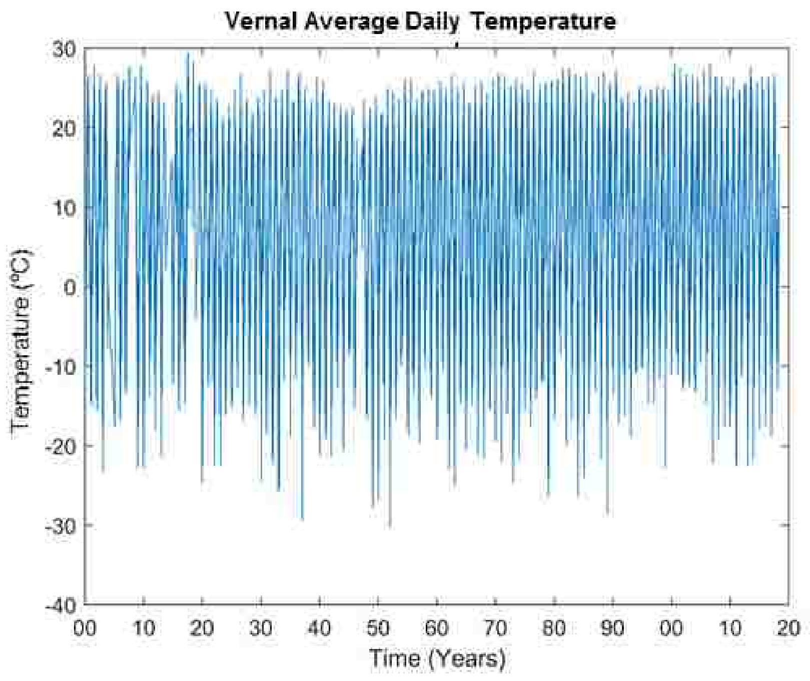


Figure 6.3-1: Vernal Temperature Dataset

www.manaraa.com

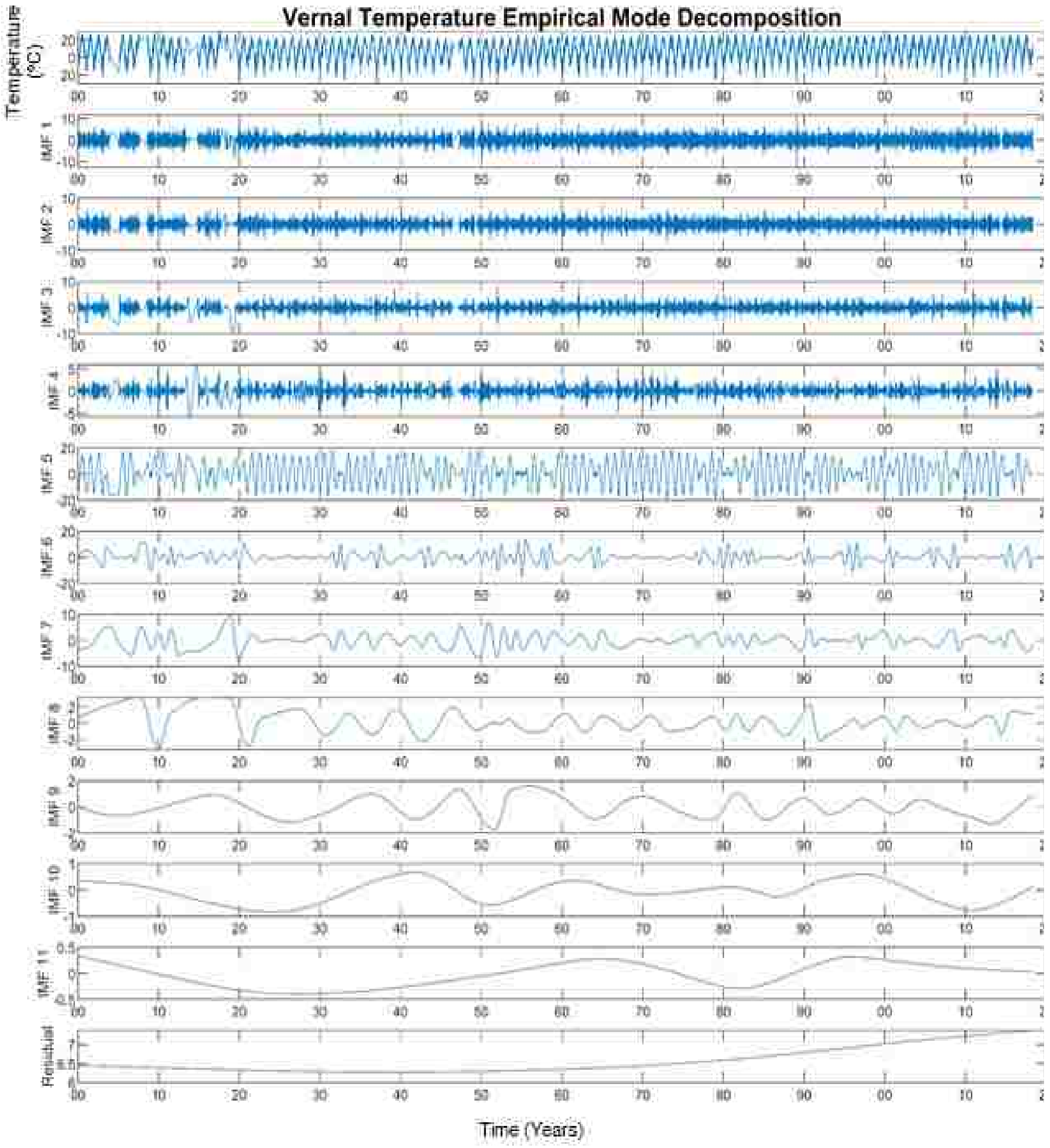


Figure 6.3-2: Decomposed Vernal Temperature Dataset

المنارة للاستشارات
www.manaraa.com

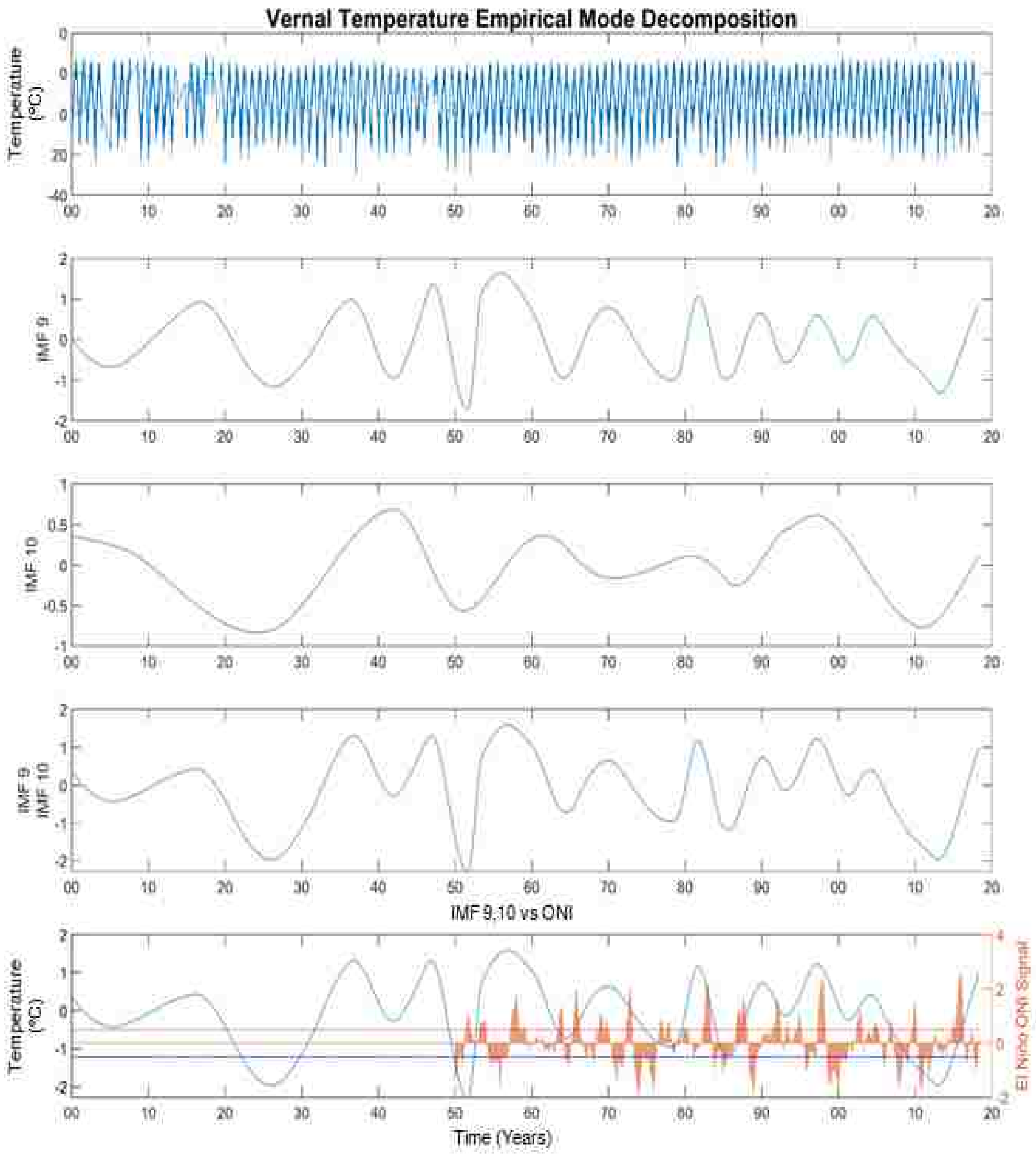


Figure 6.3-3: Vernal Temperature Decomposition Compared to the ONI Dataset (Bottom Panel).

www.manaraa.com

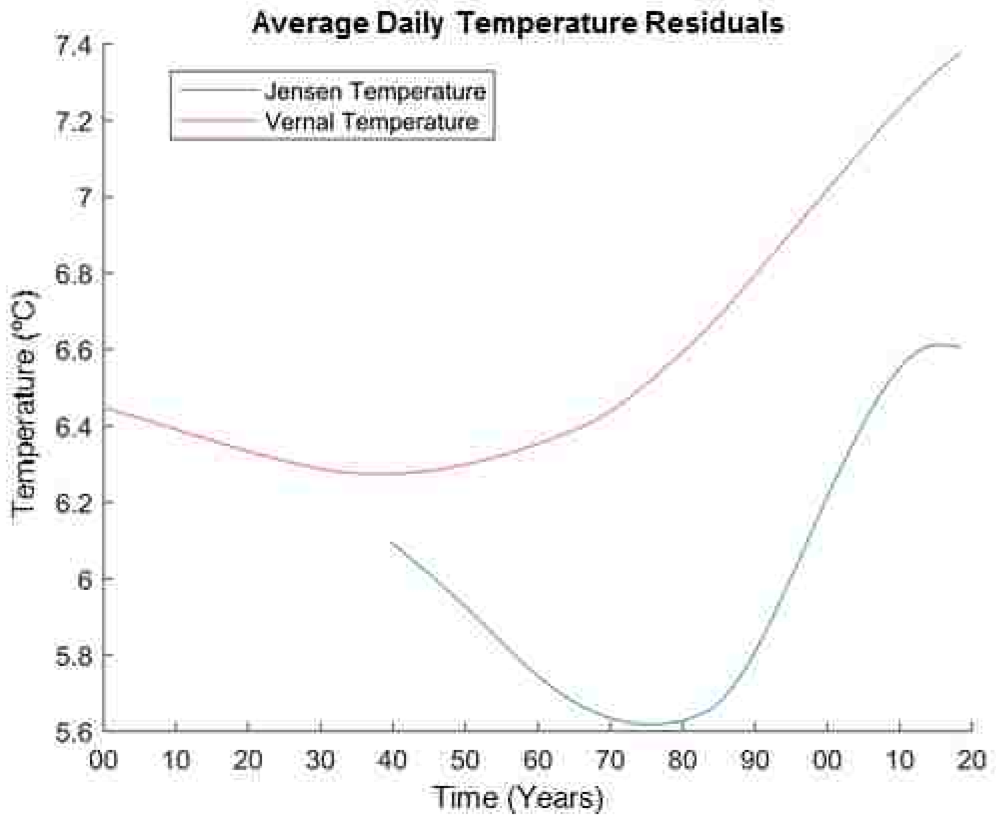


Figure 6.3-4: Residual of Jensen and Vernal Datasets

6.4 Discussion of Precipitation Results

Figure 6.4-1 shows the Jensen yearly precipitation dataset. Figure 6.4-2 shows the decomposed dataset, with Figure 6.4-3 showing the IMF/ONI comparison and the residuals (Figure 6.4-4).

The IMF/ONI comparison (Figure 6.4-3) shows IMF 2 compared to the ONI dataset. The data comparison suggests that ONI peaks resulting from El Niño correlate to increases in precipitation near the Jensen area. The data also suggests that La Niña does not necessarily forecast a drier year, however, El Niño can be an indicator of a wetter year, especially if it is a very strong event.

www.manaraa.com

The residuals shown below (Figure 6.4-4) suggest that yearly trend in total precipitation depth has increased over the 93-year span, though the area. However, annual averages may not exhibit this trend as several of the period IMFs are in the minimum part of their cycle in recent years, overriding the long-term trend.

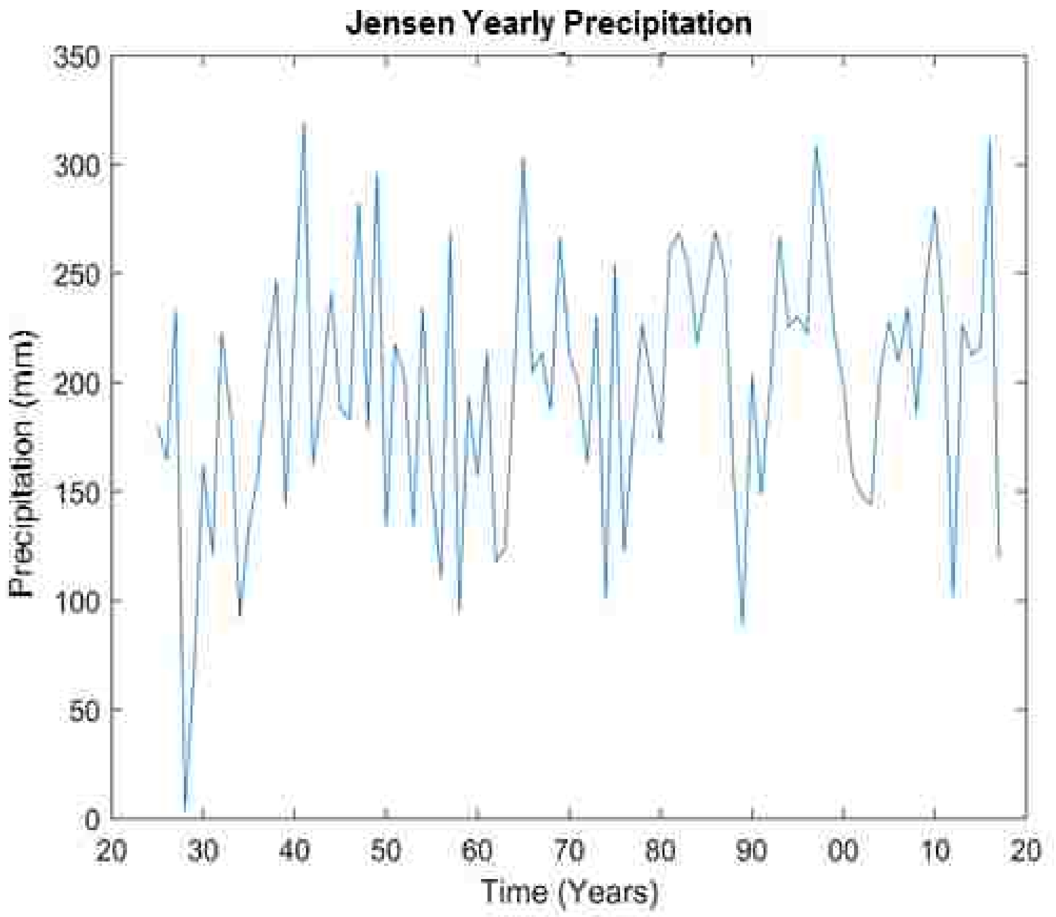


Figure 6.4-1: Jensen Yearly Precipitation Dataset. Every Point Is the Total Yearly Precipitation Depth

www.manaraa.com

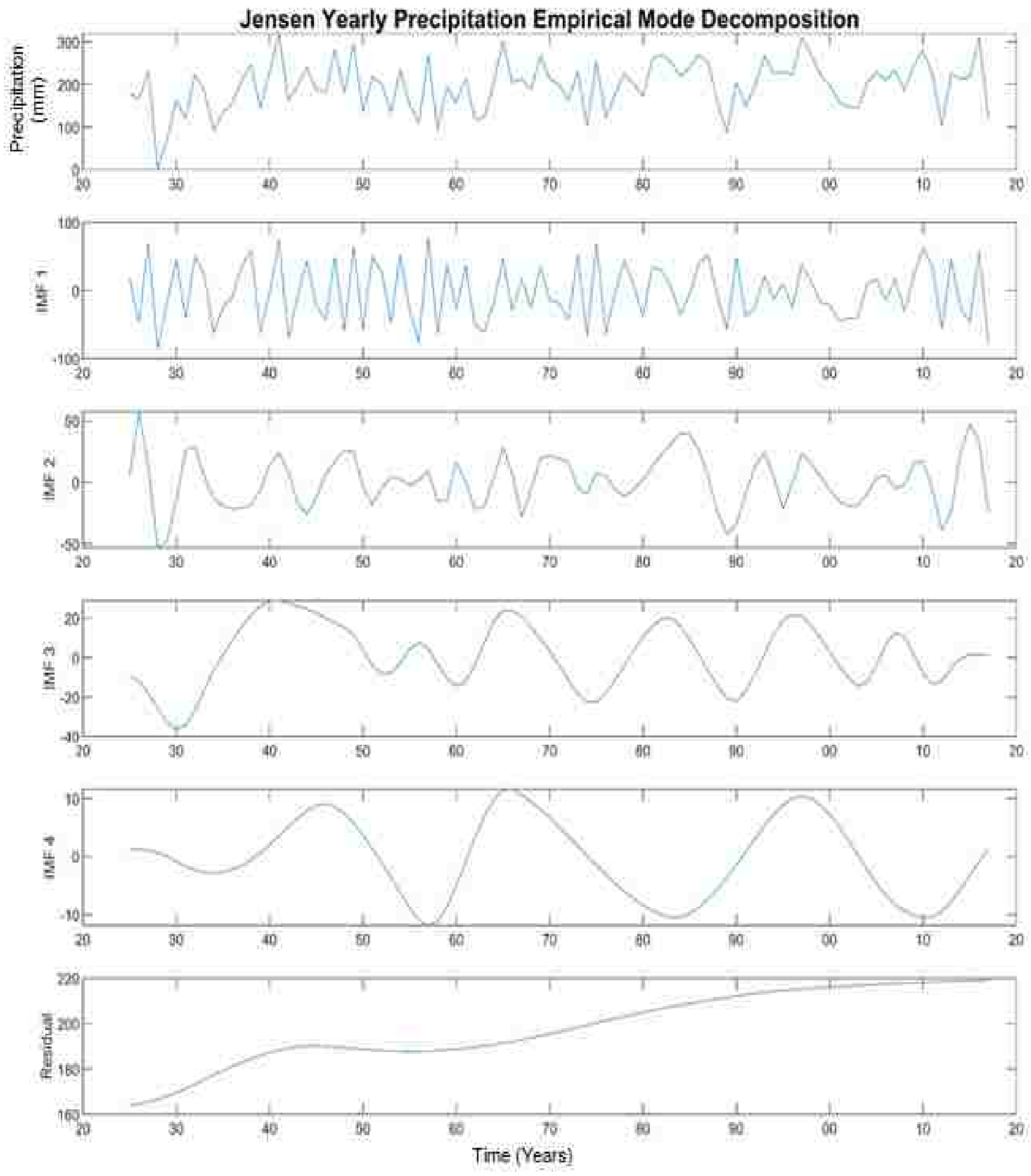


Figure 6.4-2: Jensen Precipitation Decomposition

www.manaraa.com

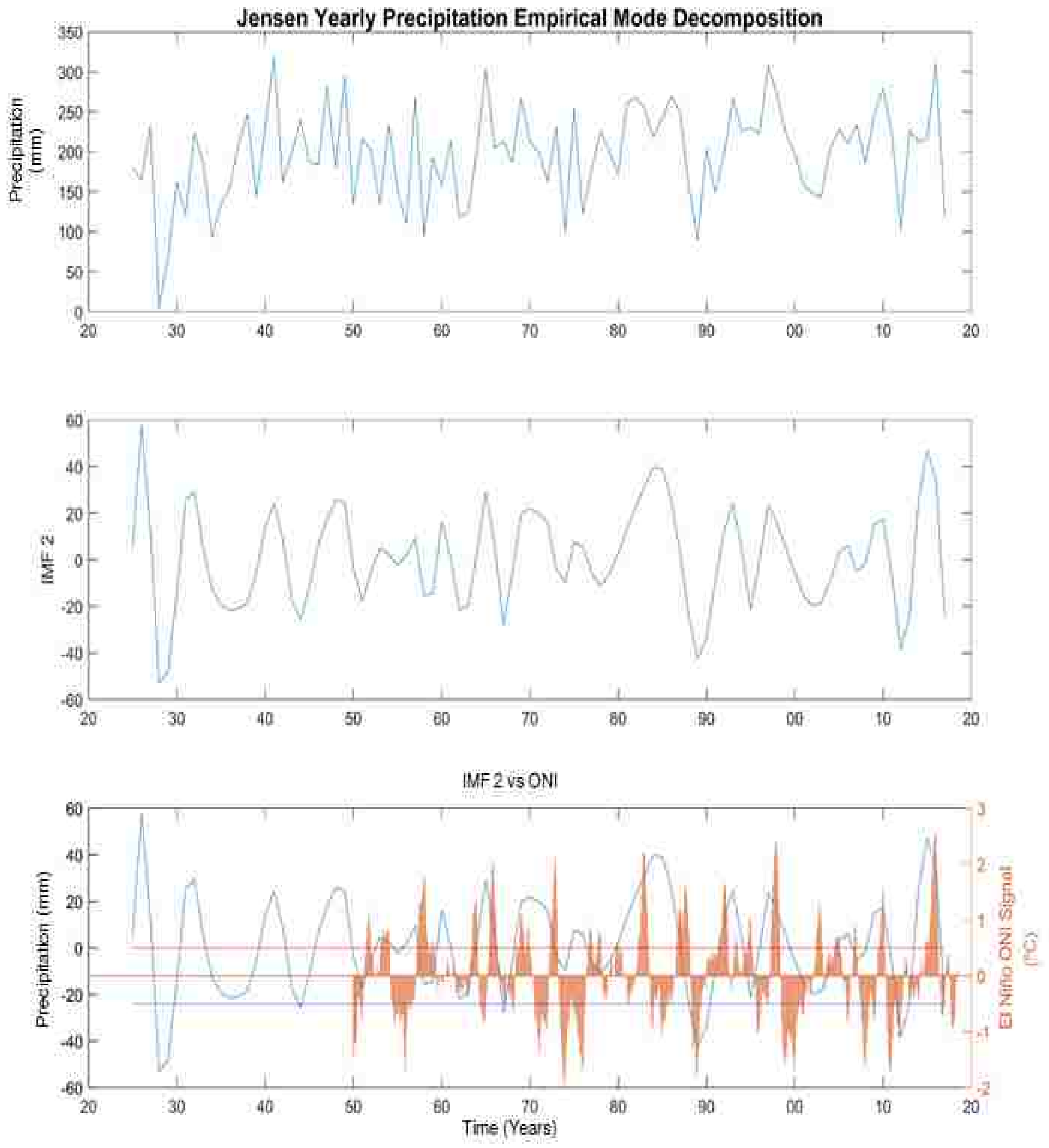


Figure 6.4-3: Jensen IMF 2 Compared to the ONI Dataset (Bottom Panel)

www.manaraa.com

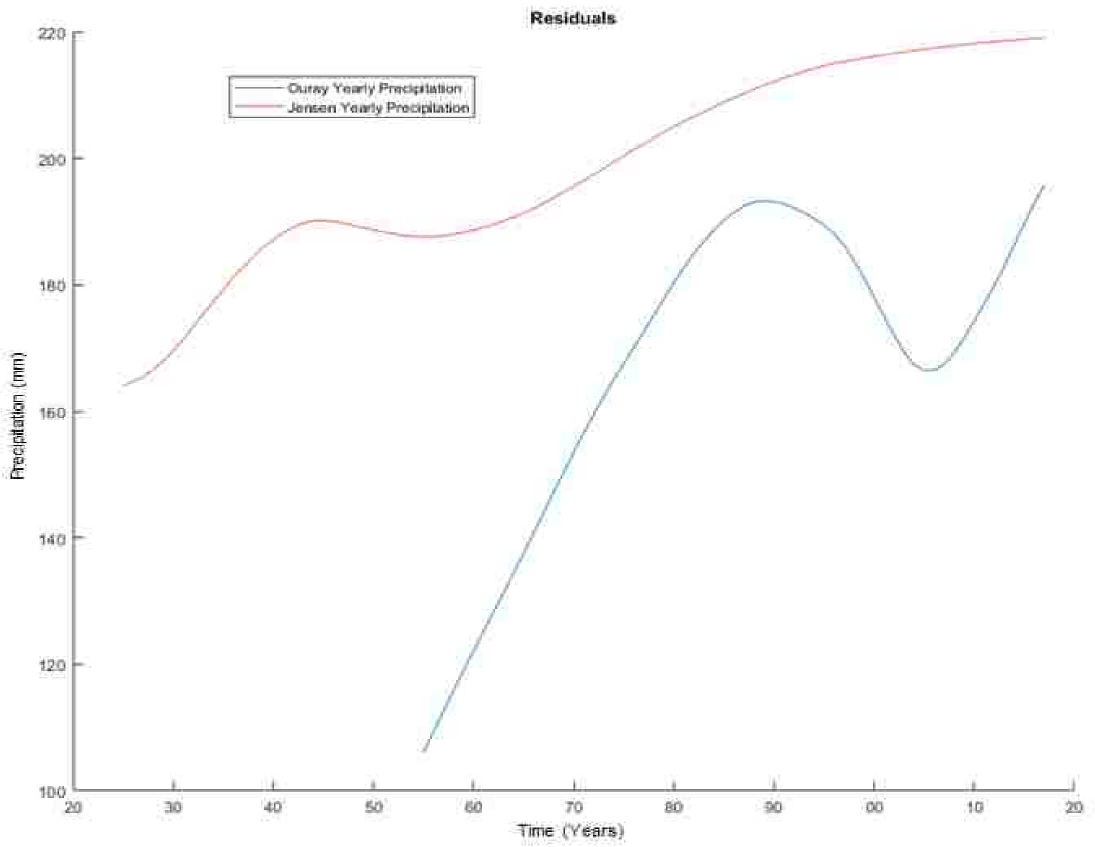


Figure 6.4-4: Ourey and Jensen Precipitation Residuals

www.manaraa.com

7 CONCLUSIONS

By applying the EMD method to an area with a control group of stream gages (little anthropogenic influence) and an experimental group of stream gages (heavily regulated) gave much insight into how dam operations affect signals from natural phenomenon. By using EMD, I was able to extract the El Niño signal from streamflow discharge data and compare it to the ONI dataset. Additionally, I tracked the El Niño signal through the dam and downstream. This signal was preserved despite the mixing of rivers, the detention structure (dam), and dam discharge operations.

Using EMD analysis I showed a a strong positive correlation between yearly precipitation, and daily average temperature IMFs to El Niño demonstrating how this global process (El Niño) can affect local temperature patterns. The residual or long-term trend in the temperature data represents an increase of 1 degree centigrade over the last 35 years and evidence of a cooling during the mid-70's –80's. The long-term trend in precipitation suggests the region that precipitation is increasing. This is not seen in the composite data set (measured data) as a number of IMFs are in the minimum part of their cycles, masking the trend. It does suggest that precipitation may increase in the future if the long-term trend continues and the IMFs move towards the maximums in their cycles.

المنارة للاستشارات
www.manaraa.com

8 FUTURE WORK

8.1 Recommendations

One of the limitations for this study is that there are not many gages in the area with the criteria of quality data as defined in section 4.2.

- Must have ~60 years or more of data available
- Must have daily data
- Must be a valid or accredited source

Trends were not easily seen in data sets smaller than 50 years long due to edge effects (boundary error). Though it is necessary to have continuous data, the guidelines above are based on previous research and not hard rules. It's necessary to research the effects of boundary errors on IMFs and residuals. Additionally, it is will be of interest to further research how missing data affects IMFs and residuals. By understanding edge effects and missing data, the criteria above can be scientifically founded.

As more data becomes available, ONI and gage, it may be worthwhile to apply this method to the same area of study and record the differences.

Since the issue is not having sufficient long data sets, using Palaeohydrology to infer historical hydrological data and applying the EMD method may reveal interesting trends. Climate models could also be used to generate past climate data in order to estimate ONI values.

المنارة للاستشارات

www.manaraa.com

Another recommendation is the use of machine learning to recognize patterns within the IMFs and fill missing data gaps. This is more feasible than attempting to fill the original data gaps since IMFS are more likely to contain recognizable patterns.

المنارة للاستشارات
www.manaraa.com

REFERENCES

Bracewell, R. N., & Bracewell, R. N. (1986). *The Fourier transform and its applications* (Vol. 31999): McGraw-Hill New York.

Center, U. S. C. (2019). *Climate Data* [Precipitation, Temperature]. Retrieved from: https://climate.usu.edu/mapGUI/mapGUI.php

Coulibaly, P., & Burn, D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research, 40*(3). doi:10.1029/2003wr002667

el-Askary, H., Sarkar, S., Chiu, L., Kafatos, M., & El-Ghazawi, T. (2004). Rain gauge derived precipitation variability over Virginia and its relation with the El Nino southern oscillation. *Advances in space research, 33*(3), 338-342.

ESRI. (2019). Documentation. Retrieved from http://desktop.arcgis.com/en/documentation/

Hargis, B. H. (2014). Analysis of Long-Term Utah Temperature Trends Using Hilbert-Haung Transforms.

Huang, N. E. (2014). *Hilbert-Huang transform and its applications* (Vol. 16): World Scientific.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., . . . Liu, H. H. (1998). *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.* Paper presented at the Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences.

Kim, D., & Oh, H.-S. (2009). EMD: a package for empirical mode decomposition and Hilbert spectrum. *The R Journal, 1*(1), 40-46.

المنارة للاستشارات
www.manaraa.com

Labat, D. (2008). Wavelet analysis of the annual discharge records of the world's largest rivers. *Advances in Water Resources, 31*(1), 109-117. doi:https://doi.org/10.1016/j.advwatres.2007.07.004

MathWorks. (2019, 2019). R2018b. Retrieved from https://www.mathworks.com/help/documentation-center.html

Melesse, A., Abtew, W., Dessalegne, T., & Wang, X. (2010). Low and high flow analyses and wavelet application for characterization of the Blue Nile River system. *Hydrological Processes, 24*(3), 241-252. doi:10.1002/hyp.7312

Microsoft. (2019). Excel 2019. *What's new in Excel 2019 for Windows.* Retrieved from https://support.office.com/en-us/article/what-s-new-in-excel-2019-for-windows-5a201203-1155-4055-82a5-82bf0994631f

Nelsen, B., Williams, D., Williams, G., & Berrett, C. (2018). An Empirical Mode-Spatial Model for Environmental Data Imputation. *Hydrology, 5*(4), 63.

NOAA. (2018). El Niño & La Niña (El Niño-Southern Oscillation). Retrieved from https://www.climate.gov/enso

Pizarro, G., & Lall, U. (2002). El Niño-induced flooding in the U.S. West: What can we expect? *Eos, Transactions American Geophysical Union, 83*(32), 349-352. doi:doi:10.1029/2002EO000255

Service, N. W. (2019). *Cold & Warm Episodes by Season*. El Nino ONI. Retrieved from: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

Smith, L. C., Turcotte, D. L., & Isacks, B. L. (1998). Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes, 12*(2), 233-249. doi:10.1002/(sici)1099-1085(199802)12:2<233::aid-hyp573>3.0.co;2-3

Stankovic, L. (1994). A method for time-frequency analysis. *IEEE Transactions on signal processing, 42*(1), 225-229.

Survey, U. S. G. (2019). *USGS Water Data for the Nation* [Flow Data]. Retrieved from: https://waterdata.usgs.gov/nwis?

المنارة للاستشارات

www.manaraa.com

Trenberth, K. E. (1997). The Definition of El Niño. *Bulletin of the American Meteorological Society, 78*(12), 2771-2778. doi:10.1175/1520-0477(1997)078<2771:tdoeno>2.0.co;2

المنارة للاستشارات
www.manaraa.com

APPENDIX A. MATLAB CODE

A.1 Custom Plotter

```
%Variable key for selecting "n" or "m".
%n is a variable used to select a specific dataset out of the data
%structure. "n" has been used to distinguish flow data from "m" which is
%temperature data

%Flow Data Key
%{
Max of 10 sets
n=1 "Flaming Gorge";
n=2 "Jensen";
n=3 "Little Snake Lily";
n=4 "Yampa Maybel";
Commented Out
n=5 "Yampa Deerlodge";
n=6 "Yampa Craig";
n=7 "Green River Ouray"
%}
```

المنارة للاستشارات
www.manaraa.com

```

%Temperature Data Key
%{
10 offset
n=11 "Jensen";
n=12 "Vernal";
Commented Out
n=13 "Grand Junction";
n=14 "Flaming Gorge";
n=15 "Maybel";
n=16 "Craig";
%}

%Precipitation Quarterly Data Key
%{
20 offset
n=21 "Ourey";
n=22 "Jensen";
Commented Out
n=23 "Maybel";
n=24 "Flaming Gorge";
n=25 "Craig";
%}
```

www.manaraa.com

```

%Precipitation Yearly Data Key
%{
30 offset
n=31 "Ourey";
n=32 "Jensen";
Commented Out
n=33 "Maybel";
n=34 "Flaming Gorge";
n=35 "Craig";
%}

%Custom Plot creator
%Must run Mastercode first
%The user must input n,V, and SuperPosition=1/0 Res=1/0
%When working with flow data select n value associated with the data of
%interest, enter the value of V as a vector with the imf datasets of
%interest and if user desires to impose the datasets on eachother and plot
%make the value of SuperPosition 1, otherwise SuperPosition should be 0
%For best practice, only use one of the codes at a time.


%Data Plotting Settings
```

www.manaraa.com

```

V=[9,10];      %Which IMF datasets do you want to look at,
n=12;          %Which area are you interested in- Key above
SuperPosition=1;    %Superimpose IMFs of interest Y(1)/N(0)
Res=0;         %Residual visible Y(1)/ N(0)
Dataextract=0;     %Saves Superimposed Vector
NinoRef=1;       %Displays barchart showing known El Nino Years
NinoMethod=0;      %Custom Plotter using ONI (0) or Nino intensity (1)
ResidualAdd=0;     %Add Residual to superimposed Y(1)/ N(0)
SaveFigure=1;      %Saves the Current Figure


%bar(Ninodata.date,Ninodata.flow)
%When Extracting Data, import by hand and rename workspace name
%figure
%plot(ExtractDataFlow(:,1),ExtractDataFlow(:,2),ExtractDataTemp(:,1),..
%ExtractDataTemp(:,2))
%datetick('x',11)
if NinoMethod==0;
%____________________________Using ONI Data____________________________
if n<=10
  Comparex=NinoONIdata.date;
  Comparey=NinoONIdata.flow;
```

www.manaraa.com

```
  Cplotimf(V,Gagedata(n).date,Gagedata(n).flow,Gagedata(n).imf,...
    Gagedata(n).resid,...
  Gage1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>10 && n<=20
  n=n-10;
  Comparex=NinoONIdata.date;
  Comparey=NinoONIdata.flow;
  Cplotimf(V,Tempdata(n).date,Tempdata(n).flow,Tempdata(n).imf,...
    Tempdata(n).resid,...
  Temp1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>20 && n<=30
n=n-20;
  Comparex=NinoONIdata.date;
  Comparey=NinoONIdata.flow;
  Cplotimf(V,Precipdata(n).date,Precipdata(n).flow,Precipdata(n).imf,...
    Precipdata(n).resid,...
  Precip1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>30 && n<=40
n=n-30;
  Comparex=NinoONIdata.date;
```

المنارة للاستشارات
www.manaraa.com

```
  Comparey=NinoONIdata.flow;
  Cplotimf(V,PrecipYearlydata(n).date,PrecipYearlydata(n).flow,...
   PrecipYearlydata(n).imf,PrecipYearlydata(n).resid,...
  PrecipY1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)

end
else
  NinoMethod==1;
%_______________________________Nino
Magnitude_______________________________

if n<=10
  Comparex=Ninodata.date;
  Comparey=Ninodata.flow;
  Cplotimf(V,Gagedata(n).date,Gagedata(n).flow,Gagedata(n).imf,...
Gagedata(n).resid,...
  Gage1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>10 && n<=20
  n=n-10;
  Comparex=Ninodata.date;
  Comparey=Ninodata.flow;
```

www.manaraa.com

```
  Cplotimf(V,Tempdata(n).date,Tempdata(n).flow,Tempdata(n).imf,...
  Tempdata(n).resid,...
  Temp1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>20 && n<=30
n=n-20;
  Comparex=Ninodata.date;
  Comparey=Ninodata.flow;
  Cplotimf(V,Precipdata(n).date,Precipdata(n).flow,Precipdata(n).imf,...
  Precipdata(n).resid,...
  Precip1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)
elseif n>30 && n<=40
n=n-30;
  Comparex=Ninodata.date;
  Comparey=Ninodata.flow;
  Cplotimf(V,PrecipYearlydata(n).date,PrecipYearlydata(n).flow,...
  PrecipYearlydata(n).imf,PrecipYearlydata(n).resid,...
  PrecipY1(n,1),Lab,SuperPosition,Res,NinoRef,Comparex,Comparey,...
  Dataextract,ResidualAdd,SaveFigure)

end

```

المنارة للاستشارات
www.manaraa.com

```
end



%_________________________FUNCTION____________________________________

function Cplotimf(A,Date,Flow,IMF,Residual,Name,Label,SuperPosition,...
  Res,Nino,Compx,Compy,extract,ResidualAdd,SaveFigure)
figure1=figure('Position',[0,0,1024,1200]);
imfrows=length(A);
r=Res;
Super=SuperPosition;
SP=0;


%This section Plots the original time series signal
subplot(imfrows+1+Super+r+Nino,1,1);
plot(Date,Flow);
dateFormat=11;
datetick('x',dateFormat)
title(strcat('\fontsize{14}',Name+Label(1,3)));
ylabel('Signal');

```

المنارة للاستشارات
www.manaraa.com

```

%This section sets the size and plots the IMFs
for i=1:imfrows
 j=A(1,i);
  subplot(imfrows+1+Super+r+Nino,1,i+1);
  plot(Date,IMF(:,j));
  ylabel(strcat('IMF '+ string(j)));
  dateFormat=11;
  datetick('x',dateFormat);
end

%Option for displaying Residual
if r==1
  subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r);
  plot(Date,Residual);
  ylabel('Residual');
  dateFormat=11;
  datetick('x',dateFormat);
end

%Option for adding IMFs
if SuperPosition == 1
  %Adding the Residual to the IMF
```

www.manaraa.com

```
if ResidualAdd==0
  for i=1:imfrows
  j=A(1,i);
  SP=SP+IMF(:,j);
  end
%Saving Superimposed Datasets
if extract==1
  ExtractData=[Date,SP];
  save('J:\Research\EMD\Extracts\'+Name+' IMF '+mat2str(A)+'.mat',...
    'ExtractData')
end

subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r+Super);
plot(Date,SP);
subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r+Super);
plot(Date,SP);
subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r+Super);
plot(Date,SP);
ylabel(strcat('IMF '+ string(A)));
dateFormat=11;
datetick('x',dateFormat);

elseif ResidualAdd==1
```

www.manaraa.com

```
    for i=1:imfrows
    j=A(1,i);
    SP=SP+IMF(:,j);
    end
    SP=SP+Residual;

    if extract==1
    ExtractData=[Date,SP];
    save('J:\Research\EMD\Extracts\'+Name+' IMF '+mat2str(A)+'.mat',...
      'ExtractData')
    end

  subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r+Super);
  plot(Date,SP);

  ylabel(strcat('IMF '+ string(A)+' Residual'));
  dateFormat=11;
  datetick('x',dateFormat);
  end

else

end
```

www.manaraa.com

```
if Nino==1
  if SP==0
    SP=IMF(:,A);
  end
  subplot(imfrows+1+Super+r+Nino,1,imfrows+1+r+Super+Nino);
  plot(Date,SP);
  ylabel('El Nino Comparison');
  datetick('x',dateFormat);
  %yyaxis left
  hold on
  yyaxis right
  bar(Compx,Compy,'FaceAlpha',.80);
  %plot(Compx,Compy);
  %El Nino Reference
  %El Nino event is efined by ONI of +/- 0.5
    TimeRef=[min(Date);max(Date)];
    ELNINO=[0.5;0.5];
    LANINA=[-0.5;-0.5];
    plot(TimeRef,ELNINO,'-.','Color','red','MarkerSize',0.1);
    plot(TimeRef,LANINA,'-.','Color','blue','MarkerSize',0.1);
  ylabel('El Niño ONI Signal')
  dateFormat=11;
```

www.manaraa.com

```
  datetick('x',dateFormat);
  hold off
end

if SaveFigure==1
  Folder='J:\Research\Writing\Figures\Custom Plots\';
  Tar=strcat(Folder,Name,' IMF ',mat2str(A));
  print(gcf,Tar,'-dpng','-r1000')
end

end



```

A.2 Data Interpolator

```
clear all

%load('FlowData.mat');
load('TempData.mat');
%load('PrecipData.mat');
```

www.manaraa.com

```
%load('NinoData.mat');

St=Tempdata;
t=1;
n=1;

%for i=1:length(Gage1)
for i=5:6

%for i=1:1

  ax=St(i).date(1);
  b=length(St(i).date);
  cx=St(i).date(b);
  Intx=ax:cx;
  Inty=interp1(St(i).date,St(i).flow,ax:t:cx,'linear');

  Data1=[St(i).date,St(i).flow];
  Data2=[Intx.',Inty.'];
  [DataFinal(i).original,DataFinal(i).Int]=ReadGage(Data1,Data2);

  %figure(i)=figure
  plot(St(i).date,St(i).flow,"o")
```

المنارة للاستشارات
www.manaraa.com

```
  title("Original"+( i));
hold on

plot(Intx,Inty,"*")
  title("Interpolate"+( i));
hold off

%Precipdata(i).date=x
%Precipdata(i).flow=y
 %}
end


%Flow and Date importation
function [days, flow]=ReadGage(Data1,Data2)
days=Data1; %Importing Date data
flow=Data2; %Importing Flow Data
end
```

A.3 Data Normalizing Code

```
%This example normalize is used to compare trends between datasets
```

المنارة للاستشارات
www.manaraa.com

```
%In this case Yampa Deerlodge flow IMF [12] and Grand Junction IMF [11]
SaveFigure=0;

figure
x1=ExtractDataFlow(:,1)
y1=ExtractDataFlow(:,2)
y1Max=max(ExtractDataFlow(:,2))
y1Min=min(ExtractDataFlow(:,2))
y1Range=y1Max-y1Min
y1Norm=(y1-y1Min)/y1Range

x2=ExtractDataTemp(:,1)
y2=ExtractDataTemp(:,2)
y2Max=max(ExtractDataTemp(:,2))
y2Min=min(ExtractDataTemp(:,2))
y2Range=y2Max-y2Min
y2Norm=(y2-y2Min)/y2Range


plot(x1,y1Norm,x2,y2Norm)
datetick('x',11)

if SaveFigure==1
```

للاستشارات المنارة

www.manaraa.com

```
  Folder='J:\Research\Writing\Figures\Custom Plots\';
  Tar=strcat(Folder,Name,' IMF ',mat2str(A));
  print(gcf,Tar,'-dpng','-r1000')
end
```

A.4 Master Code

```
% code to process Green River and Yampa River data using EMD to determine
% correlations and other issues
% Need to run "ReadFlowData" first to read the data from the Excel
% spreadsheet and save it in the "FlowData.mat" file

%Requires Clear all to keep the figure counter correct
clear all

%Data Loading,
%This section loads premade datasets currently Flow, and Temperature Data
load('FlowData.mat');
load('TempData.mat');
load('PrecipData.mat');
load('PrecipYearlyData.mat');
load('NinoData.mat');
```

المنارة للاستشارات
www.manaraa.com

```
load('NinoONIData.mat');

%IMF Control: Sift Relative Tolerence
Tol=0.2;

% Array for Figure Lables
%This matrix will contain the labels used for any figures

Lab=["Date","Flow (cfs)"," Empirical Mode Decomposition",...
" Hilbert Spectrum","Temp Figure"];

% Comptue IMFs using EMD
% Best Tolerances are 0.02 and .001, unstable afterwards (Crash)

% ____________________________Flow Gage Data_________________________________
%Ploting Flow Data
%This Plots the initial flow time series data
%Changing the x-axis to a year format

%{
for i=1:length(Gage1)
%for i=1:3
```

www.manaraa.com

```
 PlotFlow(Gagedata(i).date,Gagedata(i).flow,Gage1(i,1),i)
 dateFormat = 'yy';
end
%}

%EMD for flow data
for i=1:length(Gage1)
%for i=1:3
 [imf, resid, info]=emd(Gagedata(i).flow,'Interpolation','pchip',...
  'SiftRelativeTolerance',Tol,'MaxNumIMF',500,'SiftmaxIterations',...
  900);%,'Display',0);
 Gagedata(i).imf=imf;
 Gagedata(i).resid=resid;
end


%Plotting Flow Data IMF
%Calls the User defined 'PlotIMF' function and applies it to Flow Data
for i=1:length(Gage1)
%for i=1:3
 PlotIMF(Gagedata(i).date,Gagedata(i).flow,Gagedata(i).imf,...
  Gagedata(i).resid,...
 Gage1(i,1),Lab);
```

www.manaraa.com

```
end


%Ploting HHT
for i=1:length(Gage1)
%for i=1:3
  h = findobj('type','figure');
  n = length(h);
  PlotHHT(Gagedata(i).imf,Gage1(i,1),Lab)
end

%}


%________________________________Temperature___________________________________
%Ploting Flow Data
%This Plots the initial flow time series data
%Changing the x-axis to a year format
%{
for i=1:size(Temp1,1)
%for i=1:3
  PlotFlow(Tempdata(i).date,Tempdata(i).flow,Temp1(i,1),i);
  dateFormat = 'yy';
```

www.manaraa.com

```
end
%}
%EMD for Temperature Data


for i=1:size(Temp1,1)
 [imf, resid, info]=emd(Tempdata(i).flow,'Interpolation','pchip',...
   'SiftRelativeTolerance',Tol,'MaxNumIMF',500,'SiftmaxIterations',...
   900);%,'Display',0);
 Tempdata(i).imf=imf;
 Tempdata(i).resid=resid;
end


%Plotting Temperature based IMF
%for i=1:length(Temp1)
for i=1:2
 PlotIMF(Tempdata(i).date,Tempdata(i).flow,Tempdata(i).imf,...
   Tempdata(i).resid,...
 Temp1(i,1),Lab);
end


```

www.manaraa.com

```
%Ploting HHT
for i=1:size(Temp1,1)
%for i=1:3
  h = findobj('type','figure');
  n = length(h);
  PlotHHT(Tempdata(i).imf,Temp1(i,1),Lab)
end

%}



%___________________Precipitation_Quarterly________________________________
%Ploting Precipitation Data
%This Plots the initial flow time series data
%Changing the x-axis to a year format
%{
for i=1:size(Precip1,1)
%for i=1:3
  PlotFlow(Precipdata(i).date,Precipdata(i).flow,Precip1(i,1),i);
  dateFormat = 'yy';
end
%}
```

www.manaraa.com

```

%EMD for Quarterly Precipitation Data

%for i=1:length(Precip1)
for i=1:size(Precip1,1)
  [imf, resid, info]=emd(Precipdata(i).flow,'Interpolation','pchip',...
    'SiftRelativeTolerance',Tol,'MaxNumIMF',500,'SiftmaxIterations',...
  900);%'Display',0);
  Precipdata(i).imf=imf;
  Precipdata(i).resid=resid;
end


%Plotting Quarterly Precip based IMF
%for i=1:length(Precip1)
for i=1:size(Precip1,1)
 PlotIMF(Precipdata(i).date,Precipdata(i).flow,Precipdata(i).imf,...
   Precipdata(i).resid,...
 Precip1(i,1),Lab);
end


%Ploting HHT
```

المنارة للاستشارات
www.manaraa.com

```
for i=1:size(Precip1,1)
%for i=1:3
  h = findobj('type','figure');
  n = length(h);
  PlotHHT(Precipdata(i).imf,Precip1(i,1),Lab)
end

%}


%____________________Precipitation_Yearly____________________________________
%Ploting Precipitation Data
%This Plots the initial flow time series data
%Changing the x-axis to a year format
%{
for i=1:size(PrecipY1,1)
%for i=1:3
  PlotFlow(PrecipYearlydata(i).date,PrecipYearlydata(i).flow,...
PrecipY1(i,1),i);
  dateFormat = 'yy';
end
%}

```

www.manaraa.com

```
%EMD for Yearly Precip Data
for i=1:size(PrecipY1,1)
  [imf, resid, info]=emd(PrecipYearlydata(i).flow,'Interpolation',...
    'pchip','SiftRelativeTolerance',Tol,'MaxNumIMF',500,...
    'SiftmaxIterations',900);%'Display',0);
  PrecipYearlydata(i).imf=imf;
  PrecipYearlydata(i).resid=resid;
end


%Plotting Yearly Precip Data based IMF
for i=1:size(PrecipY1,1)
 PlotIMF(PrecipYearlydata(i).date,PrecipYearlydata(i).flow,...
   PrecipYearlydata(i).imf,PrecipYearlydata(i).resid,...
 PrecipY1(i,1),Lab);
end


%Ploting Yearly Precip Data HHT
for i=1:size(PrecipY1,1)
  h = findobj('type','figure');
  n = length(h);
  PlotHHT(PrecipYearlydata(i).imf,PrecipY1(i,1),Lab)
```

www.manaraa.com

```
end

%}





%______________________________Functions___________________________________

%Ploting Flow vs Date
function PlotFlow(days,flow,Name,number)
figure('Name',strcat(Name,'Flow Data'),'NumberTitle','off');
plot(days, flow);
%This can be changed to 'dd/mm/yyy' in order to plot data of specific dates
dateFormat = 'yy';
datetick('x',dateFormat);
  xlabel('Time (Years)');
  ylabel('Signal');
  title(strcat(Name));
  %%Comment to stop file Creation
  Tar=strcat('J:\Research\Writing\Figures\',Name);
%print(gcf,Tar,'-dpng','-r1000')
```

www.manaraa.com

```
end

%Plotting all IMFs
function PlotIMF(Date,Flow,IMF,Residual,Name,Label)
%Dimensioning Matrix
%x1=1:length(Flow);
imfrows=size([IMF],2);
figure1=figure('Position',[0,0,1024,1200],'Name',Name);


%Plots the signal (Flow as time series)
subplot(imfrows+2,1,1);
plot(Date,Flow);
dateFormat=11;
datetick('x',dateFormat)
title(strcat('\fontsize{14}',Name+Label(1,3)));
ylabel('Signal');

for i=1:imfrows
  subplot(imfrows+2,1,i+1);
  plot(Date,IMF(:,i));
  ylabel(strcat('IMF '+ string(i)));
  dateFormat=11;
```

www.manaraa.com

```
  datetick('x',dateFormat)
end
subplot(imfrows+2,1,imfrows+2);
plot(Date,Residual);
ylabel('Residual');
dateFormat=11;
datetick('x',dateFormat)
  %%Comment to stop file Creation
  Tar=strcat('J:\Research\Writing\Figures\',Name,Label(1,3));
%print(gcf,Tar,'-dpng','-r1000')
end

%Plotting HHTs
function PlotHHT(IMF,Name,Label) %ActiveFigureNumber,
figure('name',strcat(Name+Label(1,4)))
hht(IMF);
title(strcat(Name+Label(1,4)));
  %%Comment to stop file Creation
  Tar=strcat('J:\Research\Writing\Figures\',Name,Label(1,4));
%print(gcf,Tar, '-dpng','-r1000')
end
```

المنارة للاستشارات

www.manaraa.com

A.5 Read Flow Data

```
% ReadFlowData
% Reads flow data from Excel Spreadsheet and stores it in a .mat file

%Scraping data from Spreadsheets
%Data has been historically imported from .csv format spreadsheets with
%dates in 'excel serial number' format in column A and
%flow values in column B

clear all

% Gage is matrix of meta data, column 1 is gage name, Column 2 is File
% name, Column 3 is the 3 letter acronym for the gage
Gage(1,1)="Flaming Gorge Flow";    Gage(1,2)="FG1.csv"; Gage(1,3)="FG1";
Gage(2,1)="Jensen Flow";        Gage(2,2)="GJ2.csv"; Gage(2,3)="GJ2";
Gage(3,1)="Little Snake Lily Flow";  Gage(3,2)="YL4.csv"; Gage(3,3)="LS1";
Gage(4,1)="Yampa Maybel Flow";     Gage(4,2)="YM5.csv"; Gage(4,3)="YM5";
%Gage(5,1)="Yampa Maybel Flow E";    Gage(5,2)="YME5.csv"; Gage(5,3)="YME5";
%Gage(5,1)="Yampa Deerlodge Flow";   Gage(3,2)="YD3.csv"; Gage(3,3)="YD3";
%Gage(6,1)="Yampa Craig Flow";     Gage(6,2)="YC6.csv"; Gage(6,3)="YC6";
%Gage(7,1)="Green River Ouray Flow";  Gage(7,2)="GO9.csv"; Gage(7,3)="GR9";

% Gage1 is matrix of meta data, column 1 is gage name, Column 2 is File
```

للاستشارات المنارة
www.manaraa.com

```
% name, Column 3 is the 3 letter acronym for the gage
Gage1=["Flaming Gorge Flow",   'J:\Research\EMD\Flow_Data\FG1.csv', "FG1"; ...
   "Jensen Flow",       'J:\Research\EMD\Flow_Data\GJ2.csv', "GJ2"; ...
   "Little Snake Lily Flow", 'J:\Research\EMD\Flow_Data\YL4.csv', "LS1"; ...
   "Yampa Maybel Flow",    'J:\Research\EMD\Flow_Data\YM5.csv', "YM5";];
   %"Yampa Maybel Flow E",   'J:\Research\EMD\Flow_Data\YME5.csv', "YME5";];
   %"Yampa Deerlodge Flow",  'J:\Research\EMD\Flow_Data\YD3.csv', "YD3"; ...
   %"Yampa Craig Flow",    'J:\Research\EMD\Flow_Data\YC6.csv', "YC6";];
   %"Green River Ouray Flow", 'J:\Research\EMD\Flow_Data\GO9.csv', "GR9"];



for i=1:length(Gage1)
%for i=1:3
 [Gagedata(i).date, Gagedata(i).flow]=ReadGage(Gage1(i,2));
end



save('FlowData.mat');


%Functions below
```

www.manaraa.com

```
%---------------------------------------------------------------------

%Flow and Date importation
function [days, flow]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
flow=xlsread(filename, 'B:B'); %Importing Flow Data
end
```

A.6 Reader Nino Data

```
clear all

Nino(1,1)="El Nino";
Nino(1,2)="J:\Research\EMD\El_Nino_Events\Nino.csv";
Gage(1,3)="Nino";

Nino1=["Nino", 'J:\Research\EMD\El_Nino_Events\Nino.csv', "Nino";];

[Ninodata(1).date, Ninodata(1).flow]=ReadGage(Nino1(1,2));

save('NinoData.mat');

```

المنارة للاستشارات
www.manaraa.com

```
%Functions below
%-------------------------------------------------------------------

%Flow and Date importation
function [days, flow]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
flow=xlsread(filename, 'B:B'); %Importing Flow Data
end
```

A.7 Read ONI Data

```
clear all

NinoONI(1,1)="El Nino ONI";
NinoONI(1,2)="J:\Research\EMD\El_Nino_Events\NOAA.csv";  Gage(1,4)="Nino ONI";

NinoONI=["Nino ONI", 'J:\Research\EMD\El_Nino_Events\NOAA.csv', "Nino ONI";];

[NinoONIdata(1).date, NinoONIdata(1).flow]=ReadGage(NinoONI(1,2));

```

www.manaraa.com

```
save('NinoONIData.mat');



%Functions below
%-------------------------------------------------------------------

%Flow and Date importation
function [days, flow]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
flow=xlsread(filename, 'E:E'); %Importing Flow Data
end
```

A.8 Read Precipitation Quarterly

```
% ReadFlowData
% Reads flow data from Excel Spreadsheet and stores it in a .mat file

%Scraping data from Spreadsheets
%Data has been historically imported from .csv format spreadsheets with
%dates in 'excel serial number' format in column A and
%flow values in column B
```

www.manaraa.com

```

%Clear all used to clear cache of matlab to not save unneeded excess
clear all

% Gage is matrix of meta data, column 1 is gage name, Column 2 is File
% location, Column 3 is the 3 letter acronym for the gage


Precip(1,1)="Ouray Precipitation"; Precip(1,2)='J:\Research\EMD\Precipitation\OGP.csv'; Precip(1,3)="OGP";
Precip(2,1)="Jensen Precipitation"; Precip(2,2)='J:\Research\EMD\Precipitation\JGP.csv'; Precip(2,3)="JGP";
%Precip(3,1)="Maybel Precipitation"; Precip(3,2)='J:\Research\EMD\Precipitation\MGP.csv'; Precip(3,3)="MGP";
%Precip(4,1)="Flaming Gorge Precipitation"; Precip(4,2)='J:\Research\EMD\Precipitation\FGP.csv'; Precip(4,3)="FGP";
%Precip(5,1)="Craig Precipitation"; Precip(5,2)='J:\Research\EMD\Precipitation\CGP.csv'; Precip(5,3)="CGP";

% Gage1 is matrix of meta data, column 1 is gage name, Column 2 is File
% name, Column 3 is the 3 letter acronym for the gage
%This Matrix will used to Scrape the data from the spreadsheets

```

www.manaraa.com

```
Precip1=["Ourey Precipitation",      'J:\Research\EMD\Precipitation\OGP.csv', "OGP";...
    "Jensen Precipitation",     'J:\Research\EMD\Precipitation\JGP.csv', "JGP";];
%{
    "Maybel Precipitation",     'J:\Research\EMD\Precipitation\MGP.csv', "MGP";...
    "Flaming Gorge Precipitation",  'J:\Research\EMD\Precipitation\FGP.csv', "FGP";...
    "Craig Precipitation",      'J:\Research\EMD\Precipitation\CGP.csv', "CGP";];
 %}

%The length function will only work if the matrix has more rows than
%columns

%for i=1:length(Precip1)
for i=1:2
 [Precipdata(i).date, Precipdata(i).flow]=ReadGage(Precip1(i,2));
end



save('PrecipData.mat');



```

www.manaraa.com

```
%Functions below
%--------------------------------------------------------------------


%Flow and Date importation
function [days, temp]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
temp=xlsread(filename, 'B:B'); %Importing Flow Data
end

```

A.9 Read Precipitation Data Yearly

```
% ReadFlowData
% Reads flow data from Excel Spreadsheet and stores it in a .mat file

%Scraping data from Spreadsheets
%Data has been historically imported from .csv format spreadsheets with
%dates in 'excel serial number' format in column A and
%flow values in column B

%Clear all used to clear cache of matlab to not save unneeded excess
```

المنارة للاستشارات
www.manaraa.com

```
clear all


% Gage is matrix of meta data, column 1 is gage name, Column 2 is File
% location, Column 3 is the 3 letter acronym for the gage



PrecipY(1,1)="Ouray Yearly Precipitation";
PrecipY(1,2)='J:\Research\EMD\PrecipitationYearly\OGYP.csv'; PrecipY(1,3)="OGYP";
PrecipY(2,1)="Jensen Yearly Precipitation";
PrecipY(2,2)='J:\Research\EMD\PrecipitationYearly\JGYP.csv'; PrecipY(2,3)="JGYP";
%Precip(3,1)="Maybel Precipitation";
Precip(3,2)='J:\Research\EMD\Precipitation\MGP.csv'; Precip(3,3)="MGP";
%Precip(4,1)="Flaming Gorge Precipitation";
Precip(4,2)='J:\Research\EMD\Precipitation\FGP.csv'; Precip(4,3)="FGP";
%Precip(5,1)="Craig Precipitation";
Precip(5,2)='J:\Research\EMD\Precipitation\CGP.csv'; Precip(5,3)="CGP";


% Gage1 is matrix of meta data, column 1 is gage name, Column 2 is File
% name, Column 3 is the 3 letter acronym for the gage
%This Matrix will used to Scrape the data from the spreadsheets



PrecipY1=["Ouray Yearly Precipitation",
```

www.manaraa.com

```
'J:\Research\EMD\PrecipitationYearly\OGYP.csv', "OGYP";...
    "Jensen Yearly Precipitation",    'J:\Research\EMD\PrecipitationYearly\JGYP.csv', "JGYP";];
%{
    "Maybel Precipitation",    'J:\Research\EMD\Precipitation\MGP.csv', "MGP";...
    "Flaming Gorge Precipitation",  'J:\Research\EMD\Precipitation\FGP.csv', "FGP";...
    "Craig Precipitation",     'J:\Research\EMD\Precipitation\CGP.csv', "CGP";];
 %}

%The length function will only work if the matrix has more rows than
%columns

%for i=1:length(Precip1)
for i=1:2
 [PrecipYearlydata(i).date, PrecipYearlydata(i).flow]=ReadGage(PrecipY1(i,2));
end



save('PrecipYearlyData.mat');



```

المنارة للاستشارات
www.manaraa.com

```
%Functions below
%----------------------------------------------------------------------


%Flow and Date importation
function [days, temp]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
temp=xlsread(filename, 'B:B'); %Importing Flow Data
end

```

A.10 Reader Temperature Data

```
% ReadFlowData
% Reads flow data from Excel Spreadsheet and stores it in a .mat file

%Scraping data from Spreadsheets
%Data has been historically imported from .csv format spreadsheets with
%dates in 'excel serial number' format in column A and
%flow values in column B

%Clear all used to clear cache of matlab to not save unneeded excess
```

المنارة للاستشارات
www.manaraa.com

```
clear all

% Gage is matrix of meta data, column 1 is gage name, Column 2 is File
% location, Column 3 is the 3 letter acronym for the gage
%This is just a means of organization, not necessary


Temp(1,1)="Jensen Temperature";
Temp(1,2)='J:\Research\EMD\Temperature\JGT.csv'; Temp(1,3)="JGT";
Temp(2,1)="Vernal Temperature";
Temp(2,2)='J:\Research\EMD\Temperature\VGT.csv'; Temp(2,3)="VGT";
%Temp(3,1)="Grand Junction Temperature";
Temp(3,2)='J:\Research\EMD\Temperature\GJT.csv'; Temp(3,3)="GJT";
%Temp(4,1)="Flaming Gorge Temperature";
Temp(4,2)='J:\Research\EMD\Temperature\FGT.csv'; Temp(4,3)="FGT";
%Temp(5,1)="Maybel Temperature";
Temp(5,2)='J:\Research\EMD\Temperature\MGT.csv'; Temp(5,3)="MGT";
%Temp(6,1)="Craig Temperature";
Temp(6,2)='J:\Research\EMD\Temperature\CGT.csv'; Temp(6,3)="MGT";

% Gage1 is matrix of meta data, column 1 is gage name, Column 2 is File
% name, Column 3 is the 3 letter acronym for the gage
%This Matrix will used to Scrape the data from the spreadsheets
```

www.manaraa.com

```
Temp1=["Jensen Temperature",     'J:\Research\EMD\Temperature\JGT.csv', "JGT";...
    "Vernal Temperature",     'J:\Research\EMD\Temperature\VGT.csv', "VUT";];
  %{
    "Grand Junction Temperature", 'J:\Research\EMD\Temperature\GJT.csv', "GJT";...
    "Flaming Gorge Temperature",  'J:\Research\EMD\Temperature\FGT.csv', "FGT";...
    "Maybel Temperature",     'J:\Research\EMD\Temperature\MGT.csv', "MGT";...
    "Craig Temperature",      'J:\Research\EMD\Temperature\CGT.csv', "CGT";];
  %}

%The length function will only work if the matrix has more rows than
%columns

%for i=1:length(Temp1)
for i=1:2
  [Tempdata(i).date, Tempdata(i).flow]=ReadGage(Temp1(i,2));
end



save('TempData.mat');

```

www.manaraa.com

```


%Functions below
%--------------------------------------------------------------------


%Flow and Date importation
function [days, temp]=ReadGage(filename)
days=xlsread(filename, 'A:A'); %Importing Date data
temp=xlsread(filename, 'B:B'); %Importing Flow Data
end

```

A.11 Residual Code

```
%plot(Gagedata(1).date,Gagedata(1).resid)

Normalized=0;

Flow=1;
Temperature=1;
PrecipitationQuarter=1;
PrecipitationYearly=1;
```

المنارة للاستشارات
www.manaraa.com

```
SaveFigure=0;
  if SaveFigure==1
    Name='';
  end

Color=1;

figure(700);
  xlabel('Time (days)');
  ylabel('Signal');
  title('Residuals');

hold on
if Normalized==1
if Flow==1
for i=1:length(Gage1)

  minF=min(Gagedata(i).resid);
  maxF=max(Gagedata(i).resid);
  rangeF=maxF-minF;
  resF=Gagedata(i).resid;
  %Next line normalizes data between 0 and 1

```

www.manaraa.com

```
  %Individual Line Plot Options
  if Color==0
  plot(Gagedata(i).date,((resF-minF)/rangeF),'DisplayName',Gage1(i,1))
  elseif Color==1
  %Color Option
  plot(Gagedata(i).date,((resF-minF)/rangeF),'b','DisplayName',Gage1(i,1))
  end
end
end


if Temperature==1
for y=1:length(Temp1)
  %hold on
  minT=min(Tempdata(y).resid);
  maxT=max(Tempdata(y).resid);
  rangeT=maxT-minT;
  resT=Tempdata(y).resid;
  %Individual Line Plot Options
  if Color==0
  plot(Tempdata(y).date,((resT-minT)/rangeT),'DisplayName',Temp1(y,1))
  %Color Option
  elseif Color==1
```

المنارة للاستشارات
www.manaraa.com

```
  plot(Tempdata(y).date,((resT-minT)/rangeT),'r','DisplayName',Temp1(y,1))
  %plot(Tempdata(y).date,Tempdata(y).resid,'DisplayName',Temp1(y,1))
  end
end
end

if PrecipitationQuarter==1
%for j=1:length(Precip1)
for j=1:2
  %hold on
  minP=min(Precipdata(j).resid);
  maxP=max(Precipdata(j).resid);
  rangeP=maxP-minP;
  resP=Precipdata(j).resid;
  %Individual Line Plot Options
  if Color==0
  plot(Precipdata(j).date,((resP-minP)/rangeP),'DisplayName',Precip1(j,1))
  %Color Option
  elseif Color==1
  plot(Precipdata(j).date,((resP-minP)/rangeP),'c','DisplayName',Precip1(j,1))
  end
end
end
```

www.manaraa.com

```

if PrecipitationYearly==1
%for k=1:length(PrecipY1)
for k=1:2
  %hold on
  minPy=min(PrecipYearlydata(k).resid);
  maxPy=max(PrecipYearlydata(k).resid);
  rangePy=maxPy-minPy;
  resPy=PrecipYearlydata(k).resid;
  %Individual Line Plot Options
  if Color==0
  plot(PrecipYearlydata(k).date,((resPy-minPy)/rangePy),'DisplayName',PrecipY1(k,1))
  %Color Option
  elseif Color==1
  plot(PrecipYearlydata(k).date,((resPy-minPy)/rangePy),'o','DisplayName',PrecipY1(k,1))
  end
end
end

else
if Flow==1
for i=1:length(Gage1)

```

المنارة للاستشارات
www.manaraa.com

```
  minF=min(Gagedata(i).resid);
  maxF=max(Gagedata(i).resid);
  rangeF=maxF-minF;
  resF=Gagedata(i).resid;
  %Next line normalizes data between 0 and 1

  %Individual Line Plot Options
  if Color==0
  plot(Gagedata(i).date,Gagedata(i).resid,'DisplayName',Gage1(i,1))
  elseif Color==1
  %Color Option
  plot(Gagedata(i).date,Gagedata(i).resid,'b','DisplayName',Gage1(i,1))
  end
end
end


if Temperature==1
%for y=1:length(Temp1)
for y=1:2
  %hold on
  minT=min(Tempdata(y).resid);
  maxT=max(Tempdata(y).resid);
```

المنارة للاستشارات
www.manaraa.com

```
  rangeT=maxT-minT;
  resT=Tempdata(y).resid;
  %Individual Line Plot Options
  if Color==0
  plot(Tempdata(y).date,Tempdata(y).resid,'DisplayName',Temp1(y,1))
  %Color Option
  elseif Color==1
  plot(Tempdata(y).date,Tempdata(y).resid,'r','DisplayName',Temp1(y,1))
  %plot(Tempdata(y).date,Tempdata(y).resid,'DisplayName',Temp1(y,1))
  end
end
end

if PrecipitationQuarter==1
%for j=1:length(Precip1)
for j=1:2
  %hold on
  minP=min(Precipdata(j).resid);
  maxP=max(Precipdata(j).resid);
  rangeP=maxP-minP;
  resP=Precipdata(j).resid;
  %Individual Line Plot Options
  if Color==0
```

المنارة للاستشارات
www.manaraa.com

```
  plot(Precipdata(j).date,Precipdata(j).resid,'DisplayName',Precip1(j,1))
  %Color Option
  elseif Color==1
  plot(Precipdata(j).date,Precipdata(j).resid,'c','DisplayName',Precip1(j,1))
  end
end
end

if PrecipitationYearly==1
%for k=1:length(PrecipY1)
for k=1:2
  %hold on
  minPy=min(PrecipYearlydata(k).resid);
  maxPy=max(PrecipYearlydata(k).resid);
  rangePy=maxPy-minPy;
  resPy=PrecipYearlydata(k).resid;
  %Individual Line Plot Options
  if Color==0
  plot(PrecipYearlydata(k).date,PrecipYearlydata(k).resid,'DisplayName',PrecipY1(k,1))
  %Color Option
  elseif Color==1
plot(PrecipYearlydata(k).date,PrecipYearlydata(k).resid,'o','DisplayName',PrecipY1(k,1))
```

المنارة للاستشارات
www.manaraa.com

```
  end
end
end
if SaveFigure==1
  Folder='J:\Research\Writing\Figures\Custom Plots\';
  Tar=strcat(Folder,Name);
  print(gcf,Tar,'-dpng','-r1000')
end
end

hold off
legend;
dateFormat=11;
datetick('x',dateFormat);
```

www.manaraa.com

APPENDIX B. PROCESSED DATASETS BY LOCATION AND DATA

B.1 Flaming Gorge Flow Gage

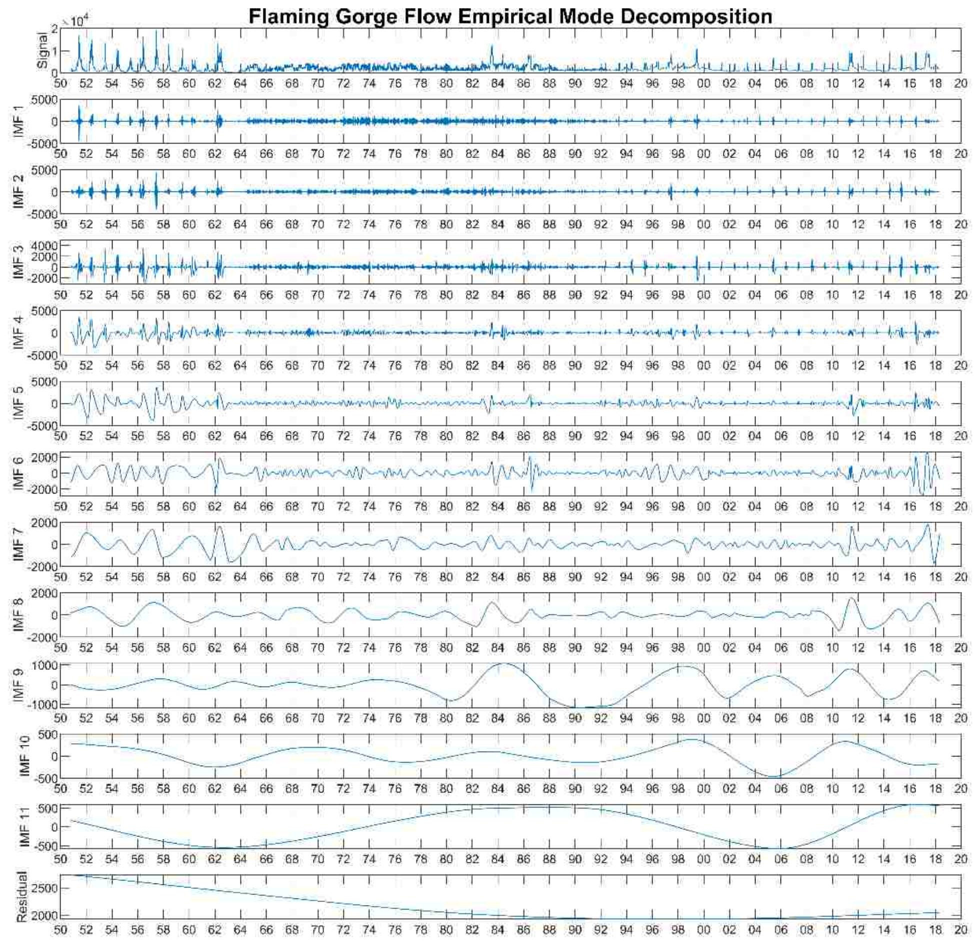


Figure B.1-1:Flaming Gorge Flow Empirical Mode Decomposition

www.manaraa.com

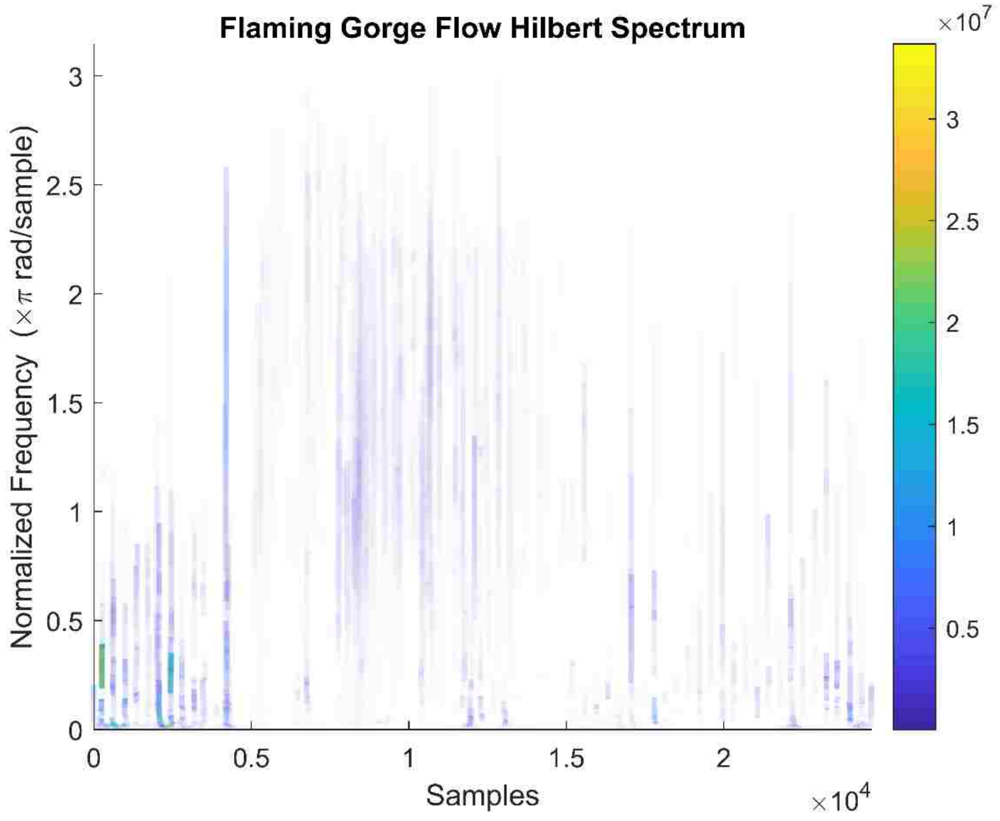


Figure B.1-2:Flaming Gorge Flow HHT

www.manaraa.com

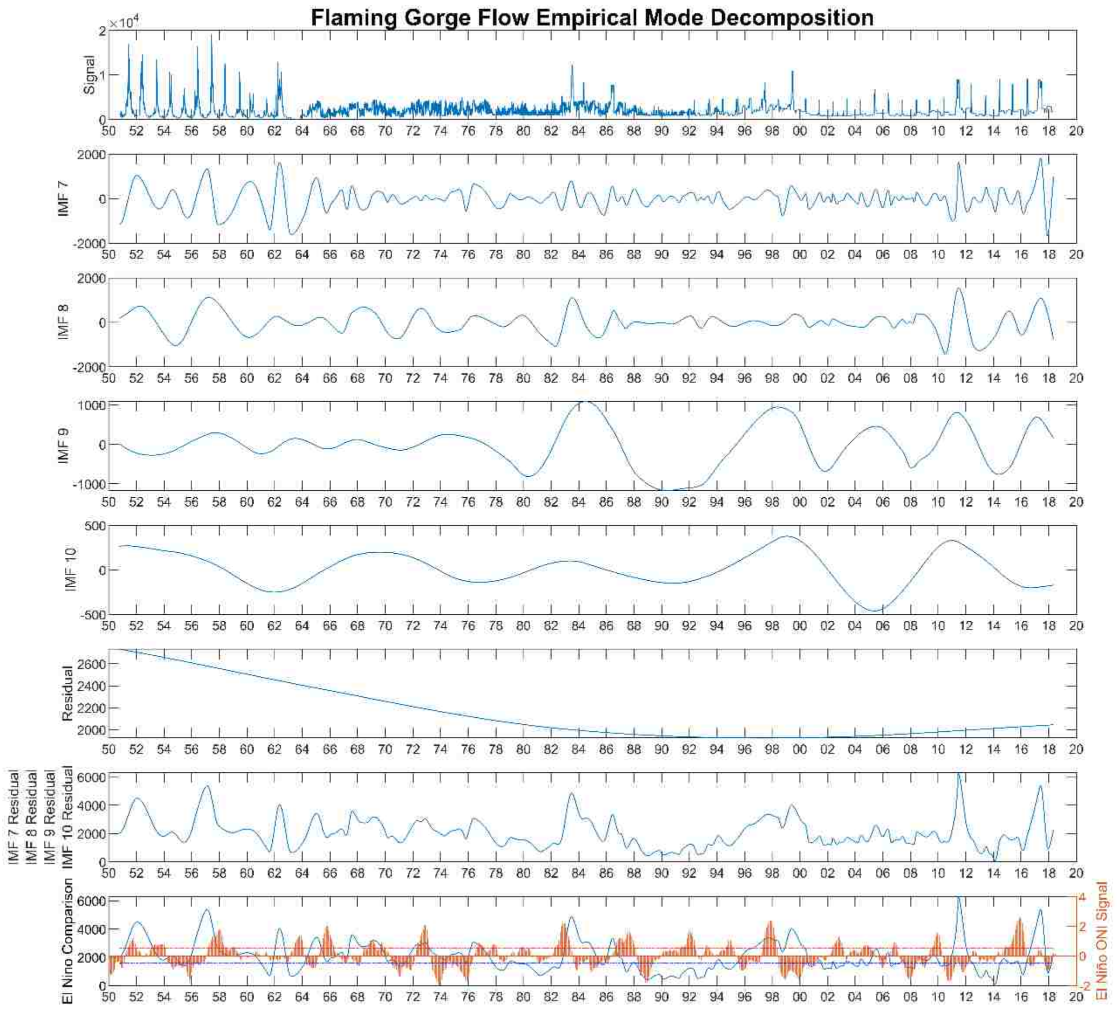


Figure B.1-3:Flaming Gorge Flow Empirical Mode Decomposition IMF 7,8,9 and 10

www.manaraa.com

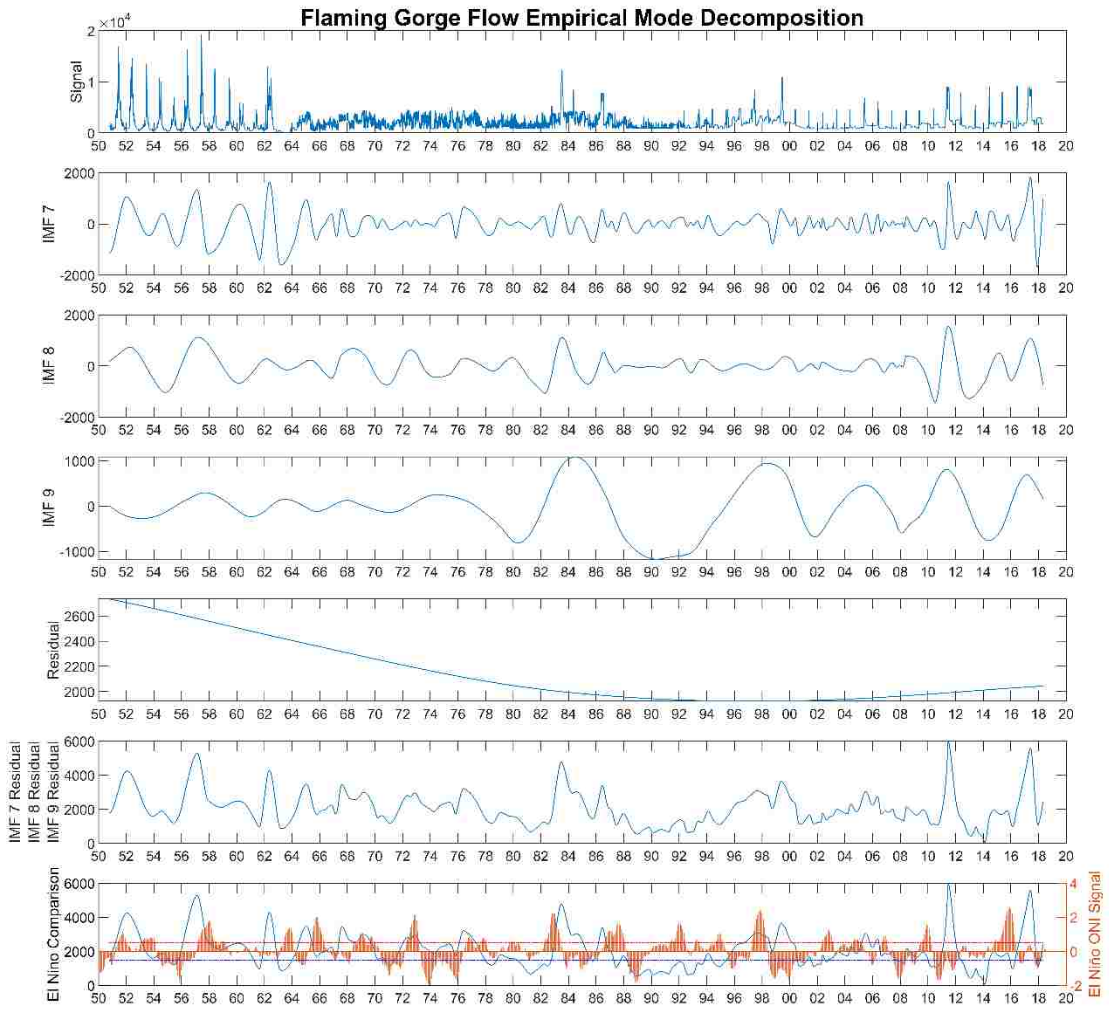


Figure B.1-4:Flaming Gorge Flow Empirical Mode Decomposition Residual Superimposed on IMF 7 8 and 9

www.manaraa.com

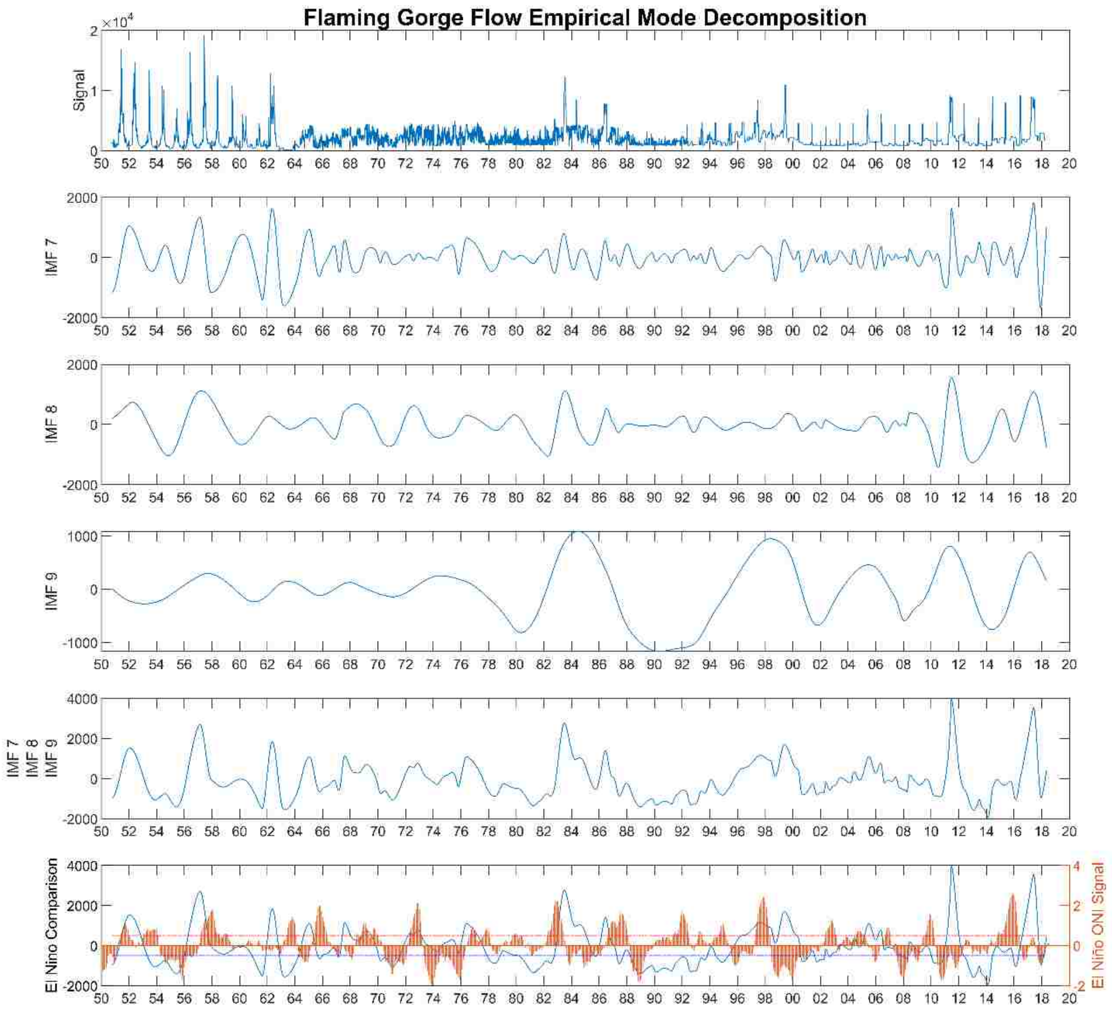


Figure B.1-5:Flaming Gorge Flow Empirical Mode Decomposition IMF 7 8 and 9

www.manaraa.com

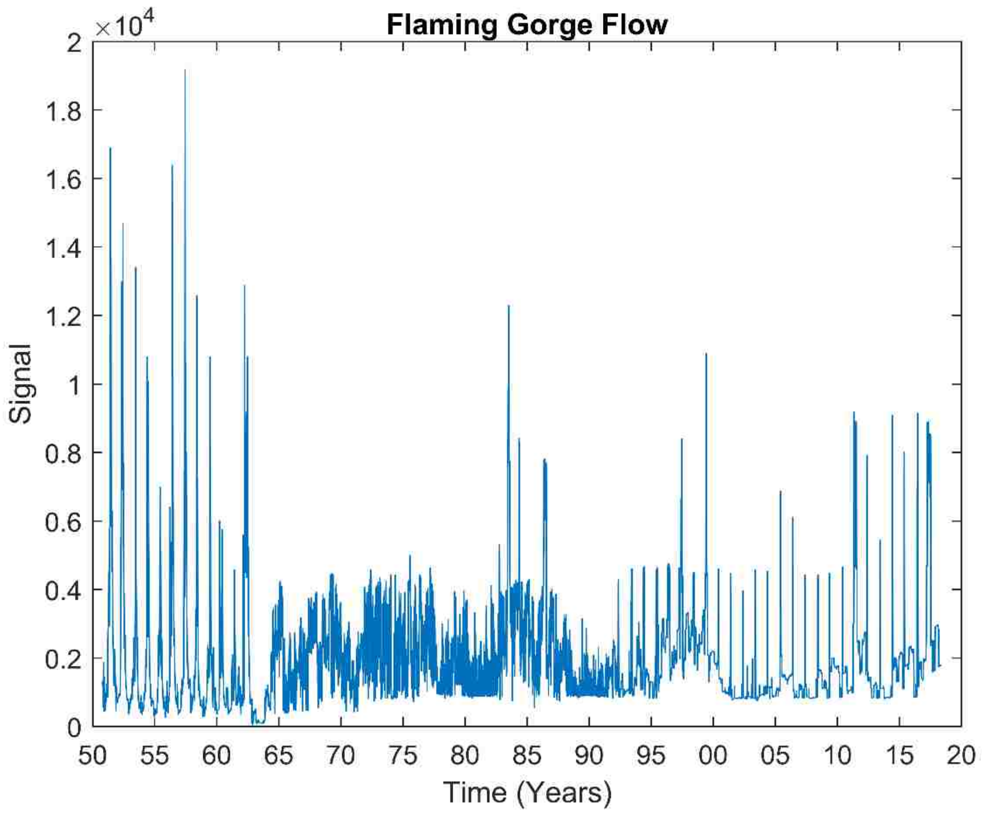


Figure B.1-6:Flaming Gorge Flow

المنارة للاستشارات
www.manaraa.com

B.2 Jensen Flow Gage

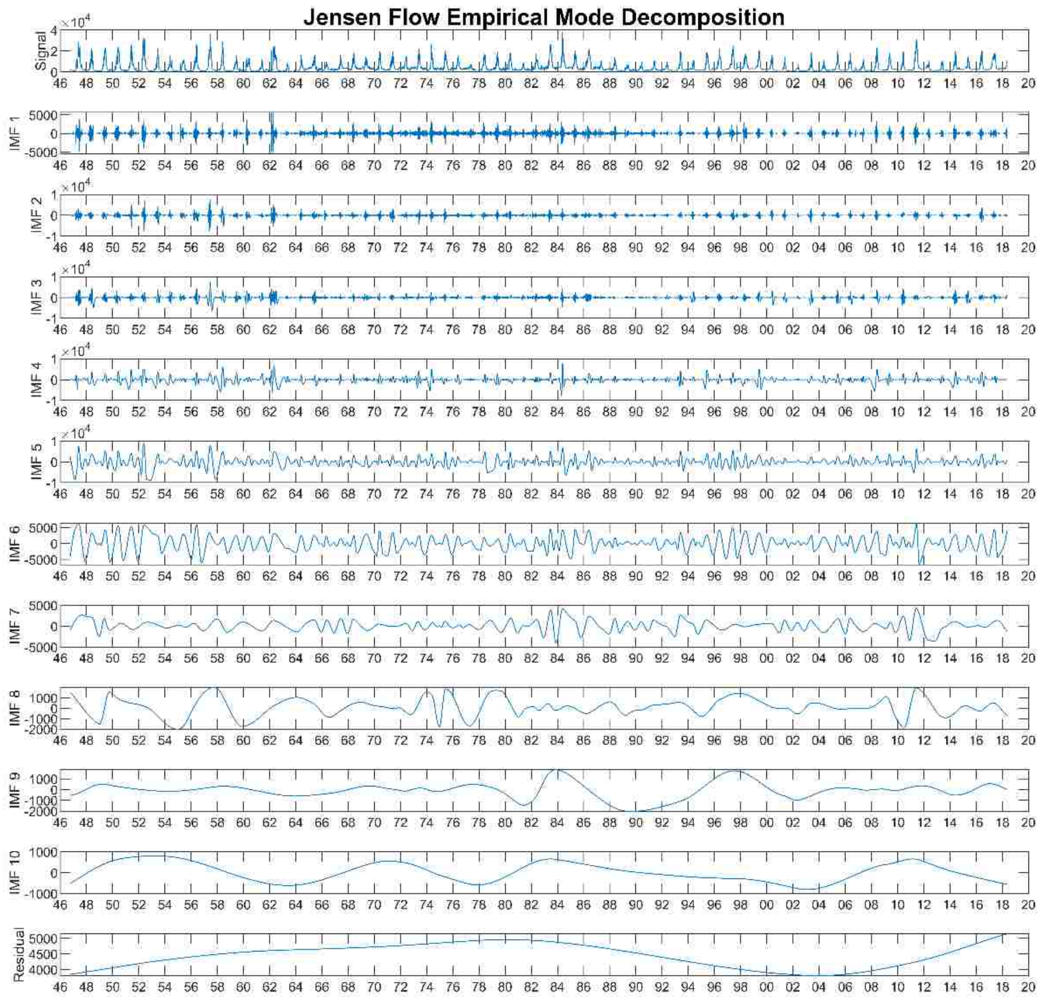


Figure B.2-1:Jensen Flow Empirical Mode Decomposition

www.manaraa.com

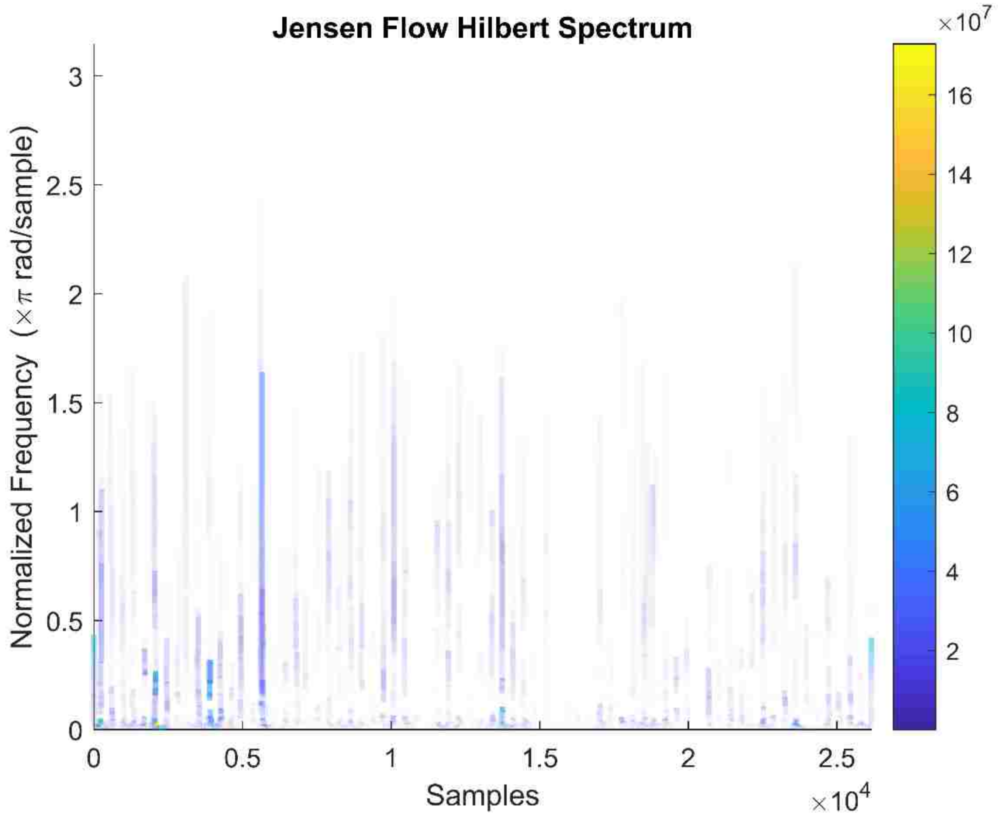


Figure B.2-2:Jensen Flow Empirical Mode Decomposition HHT

المنارة للاستشارات
www.manaraa.com

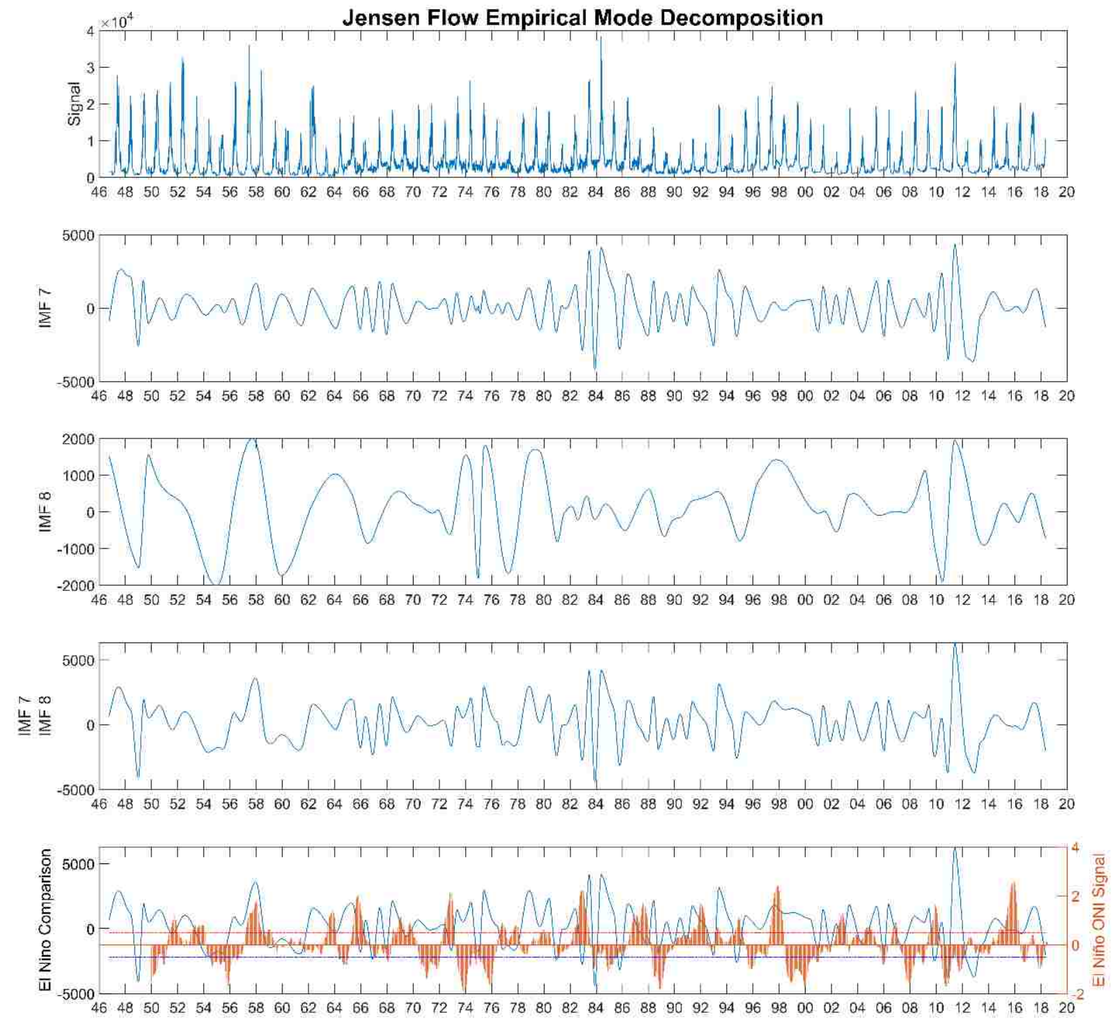


Figure B.2-3:Jensen Empirical Mode Decomposition IMF 7 and 8

المنارة للاستشارات
www.manaraa.com

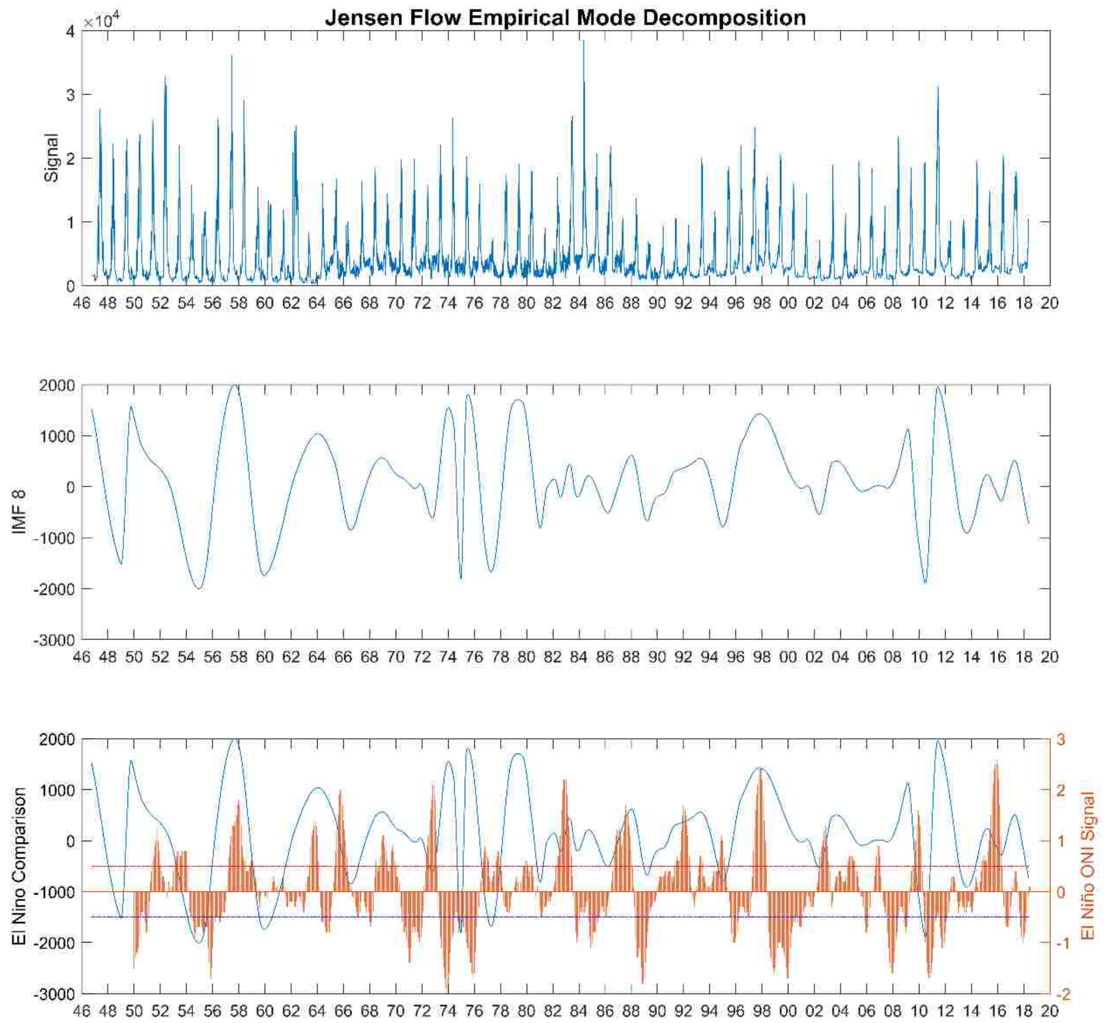


Figure B.2-4:Jensen Empirical Mode Decomposition IMF 8

www.manaraa.com

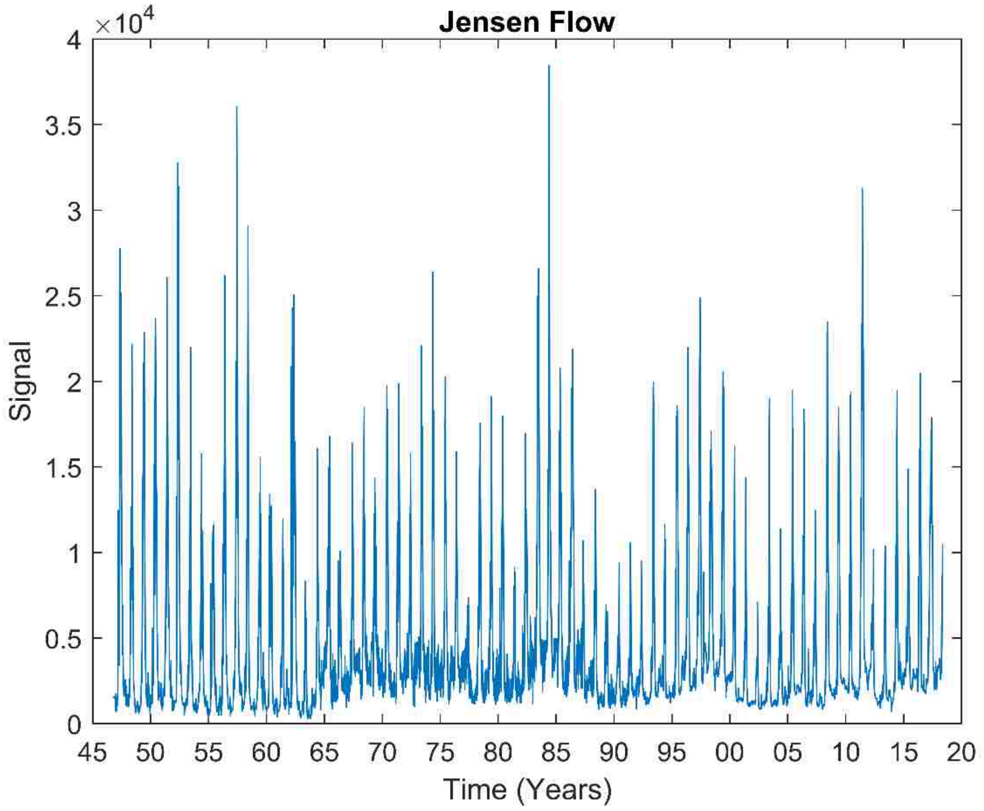


Figure B.2-5:Jensen Flow Data

www.manaraa.com

B.3 Little Snake- Lily Flow

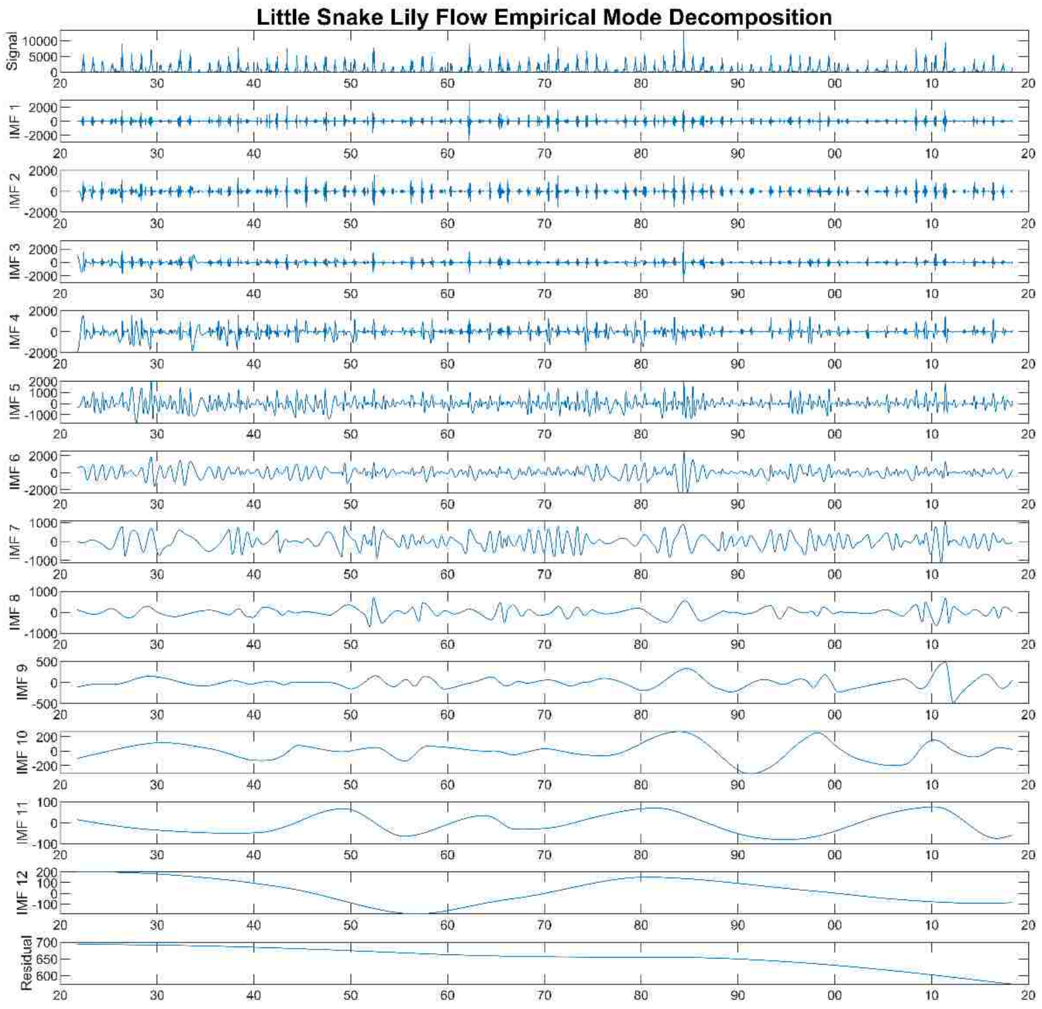


Figure B.3-1:Little Snake Lily Empirical Mode Decomposition

www.manaraa.com

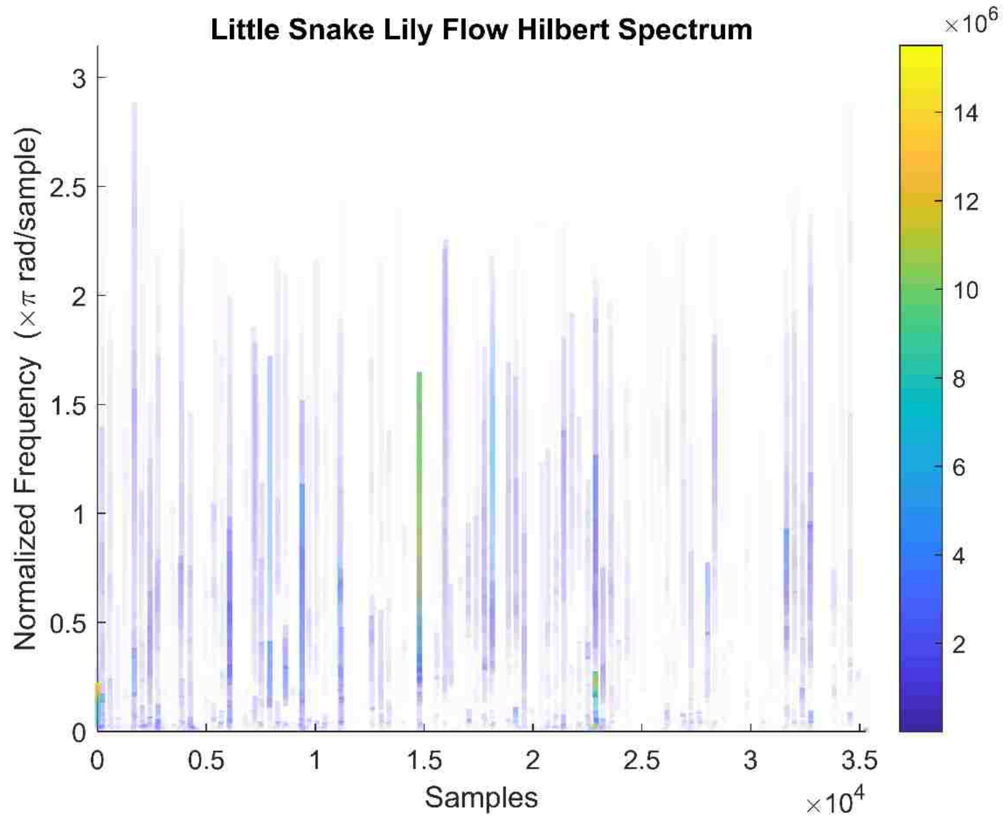


Figure B.3-2: Little Snake Lily Empirical Mode Decomposition HHT

المنارة للاستشارات
www.manaraa.com

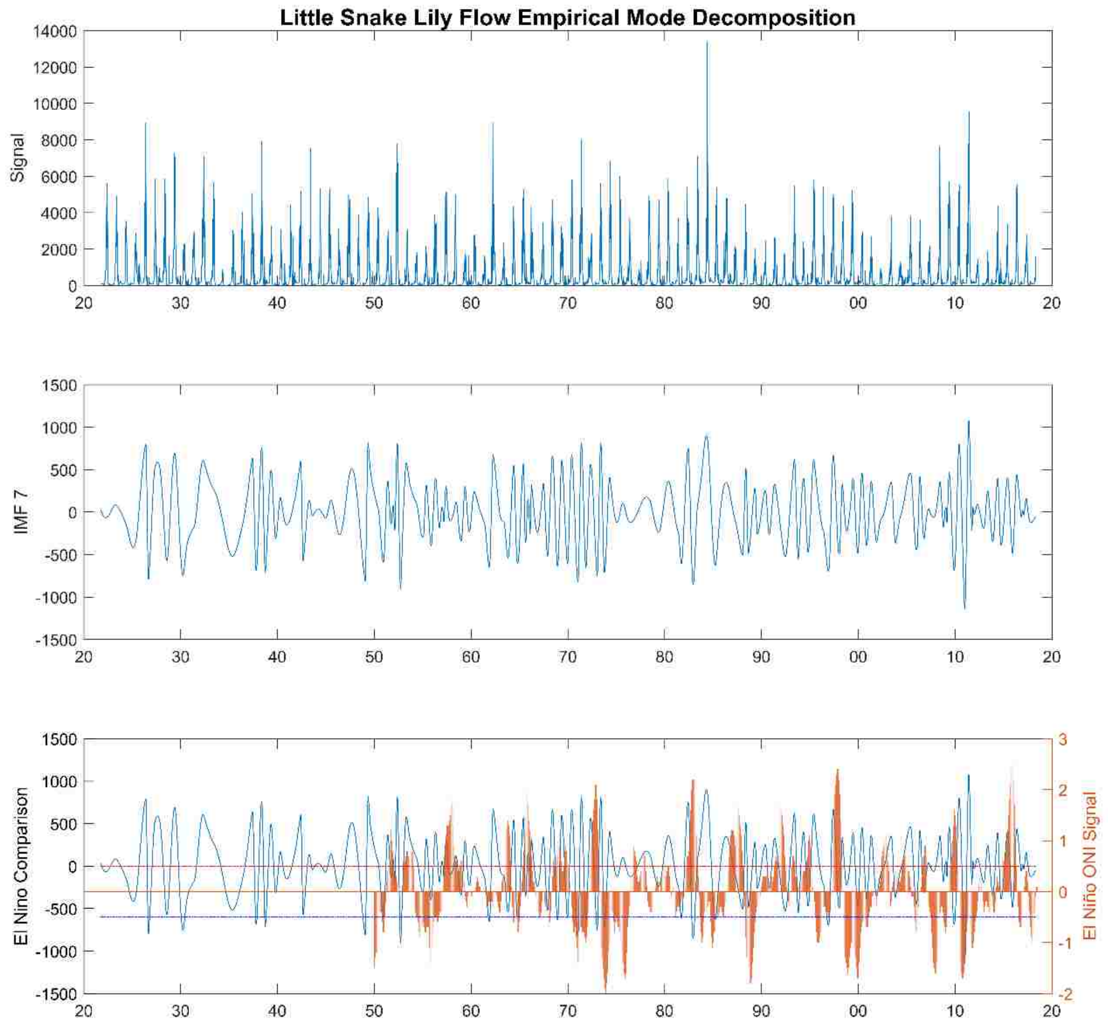


Figure B.3-3: Little Snake Lily Empirical Mode Decomposition IMF 7

www.manaraa.com

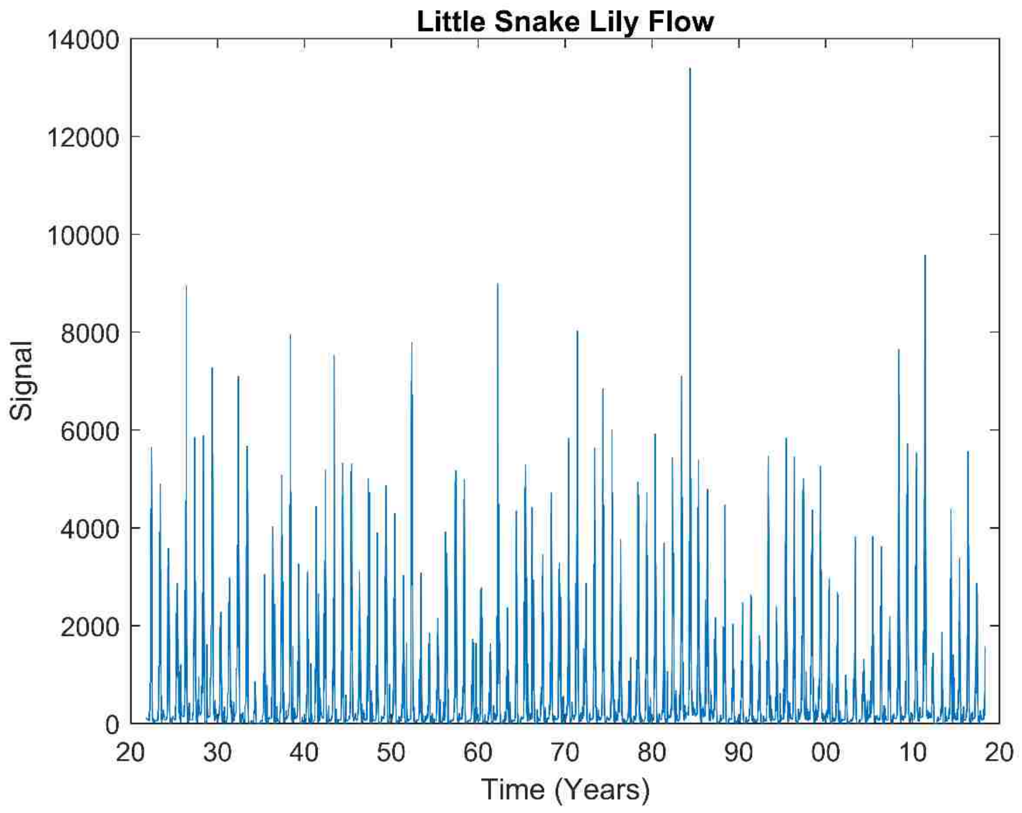


Figure B.3-4: Little Snake Lily Flow

www.manaraa.com

B.4 Yampa Maybel Flow

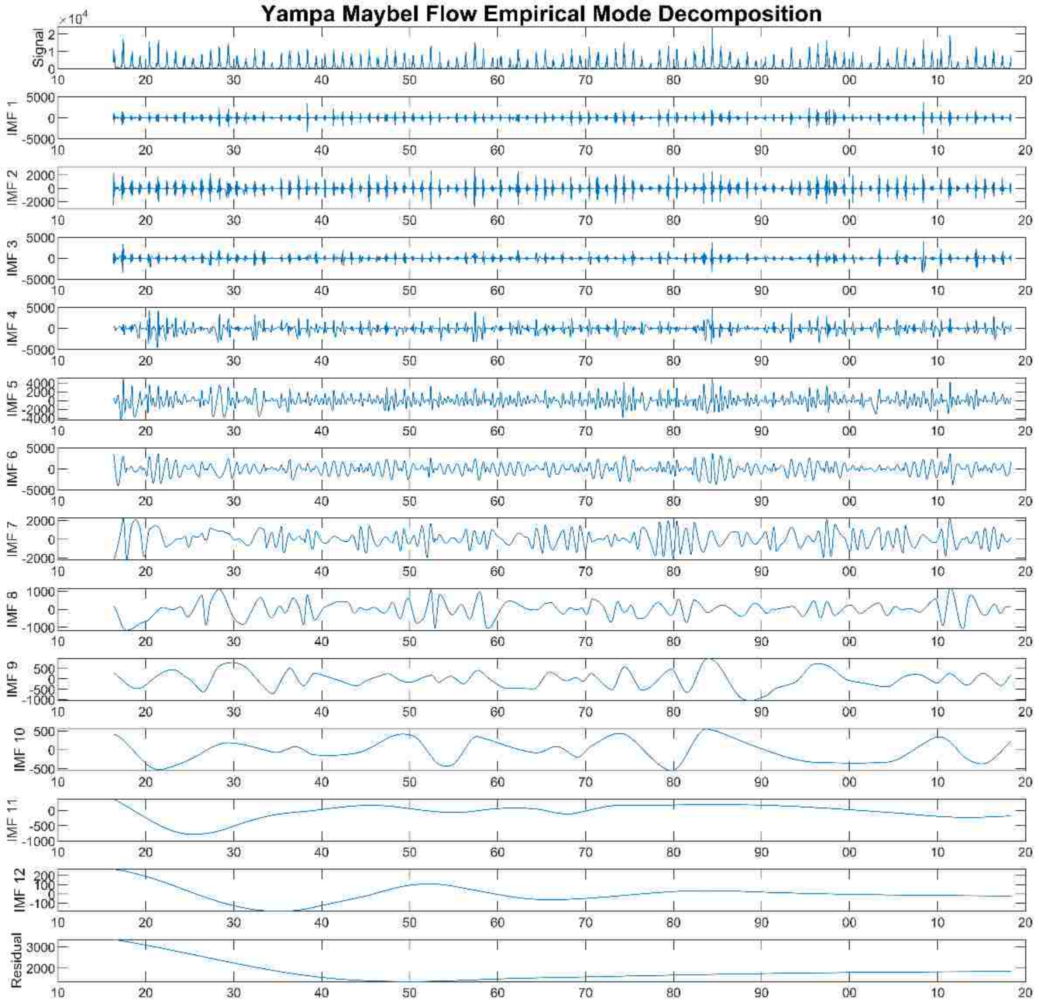


Figure B.4-1: Yampa Maybel Empirical Mode Decomposition

www.manaraa.com

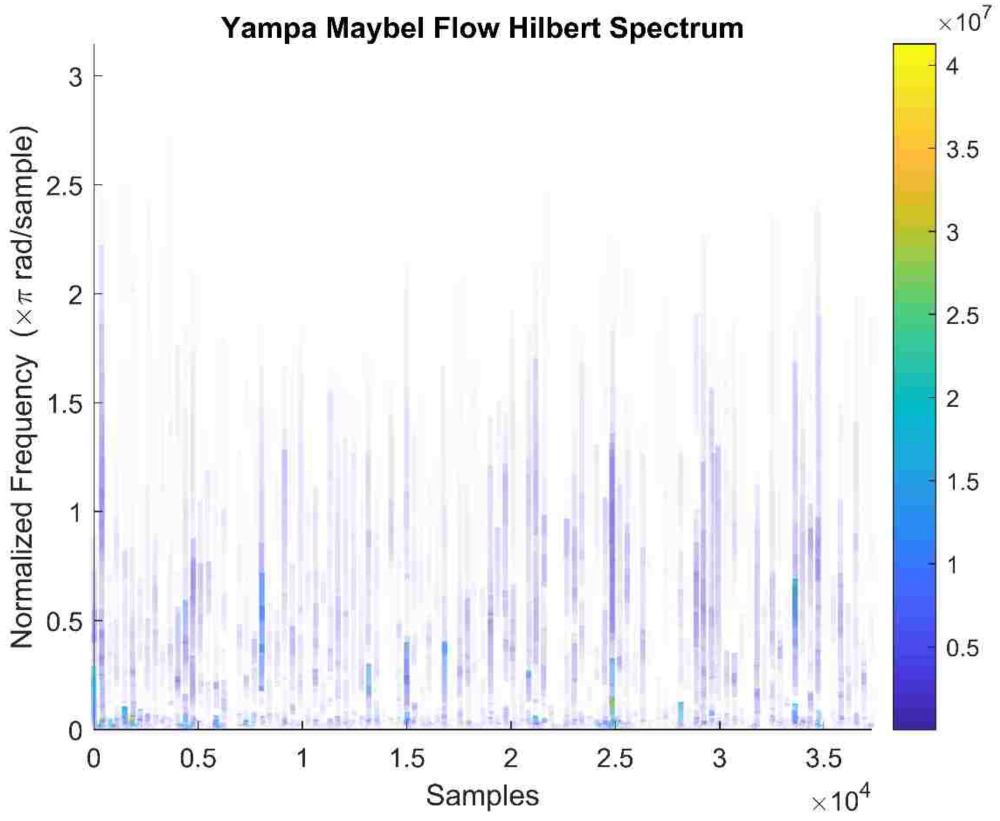


Figure B.4-2: Yampa Maybel Empirical Mode Decomposition HHT

www.manaraa.com

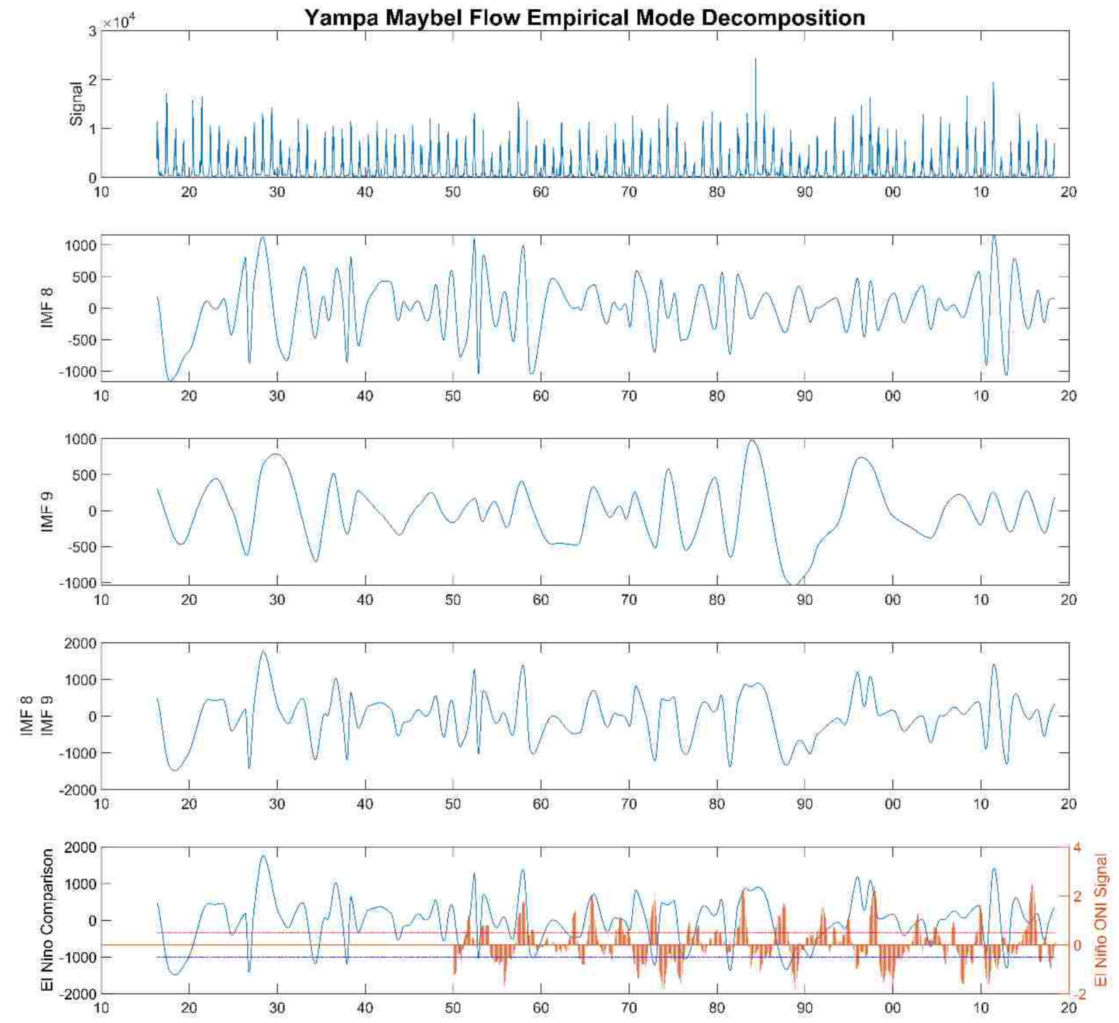


Figure B.4-3: Yampa Maybel Empirical Mode Decomposition IMF 8 and 9

www.manaraa.com

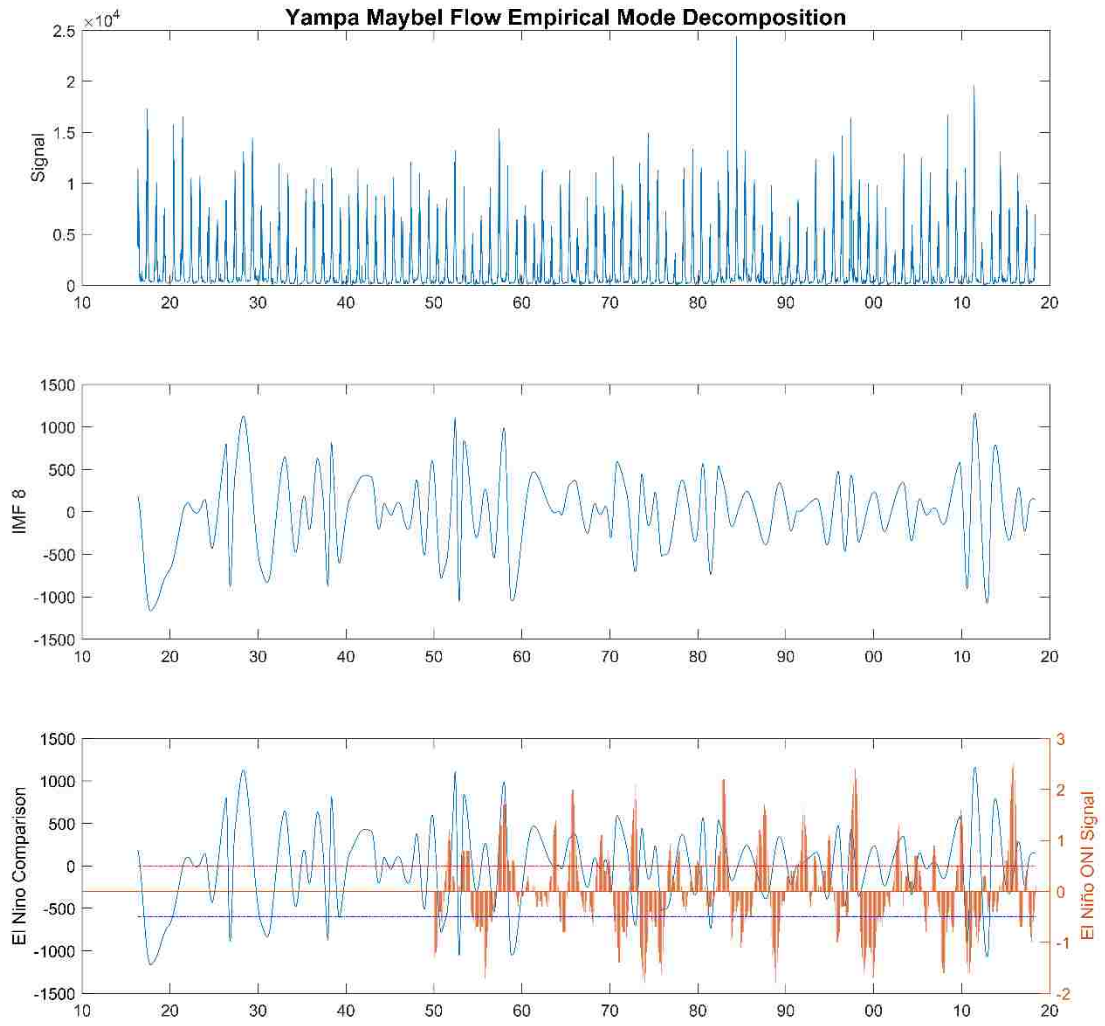


Figure B.4-4: Yampa Maybel Empirical Mode Decomposition IMF 8

www.manaraa.com

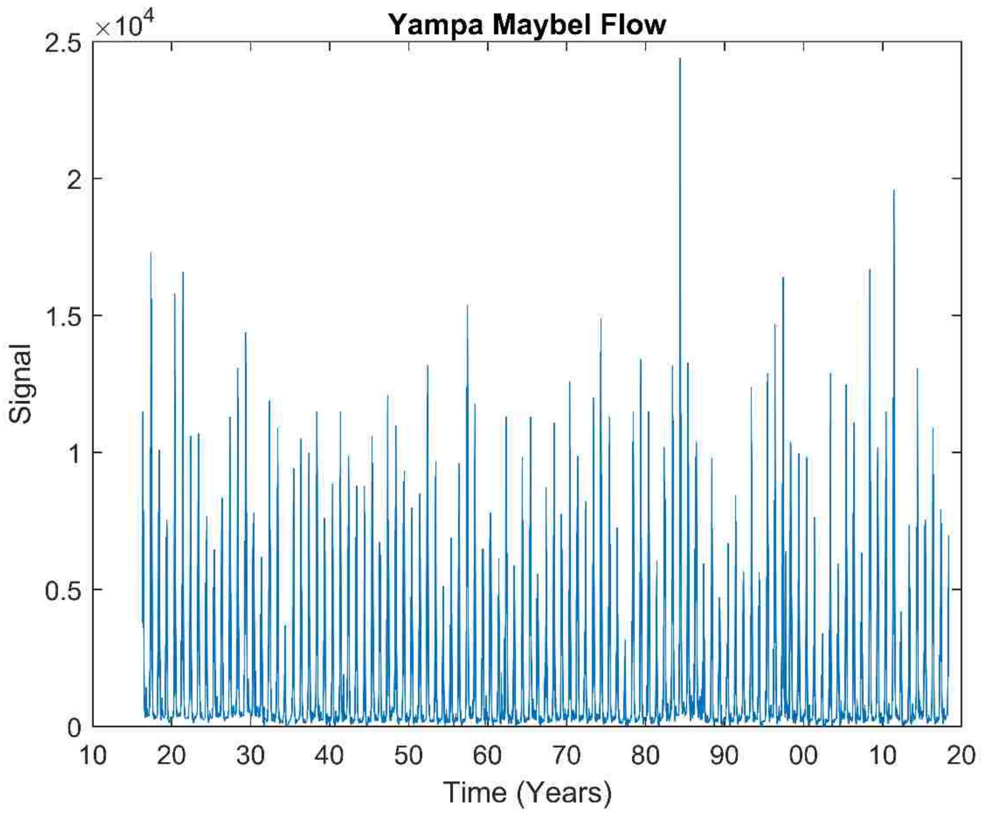


Figure B.4-5: Yampa Maybel Flow

www.manaraa.com

B.5 Ourey Quarterly Precipitation

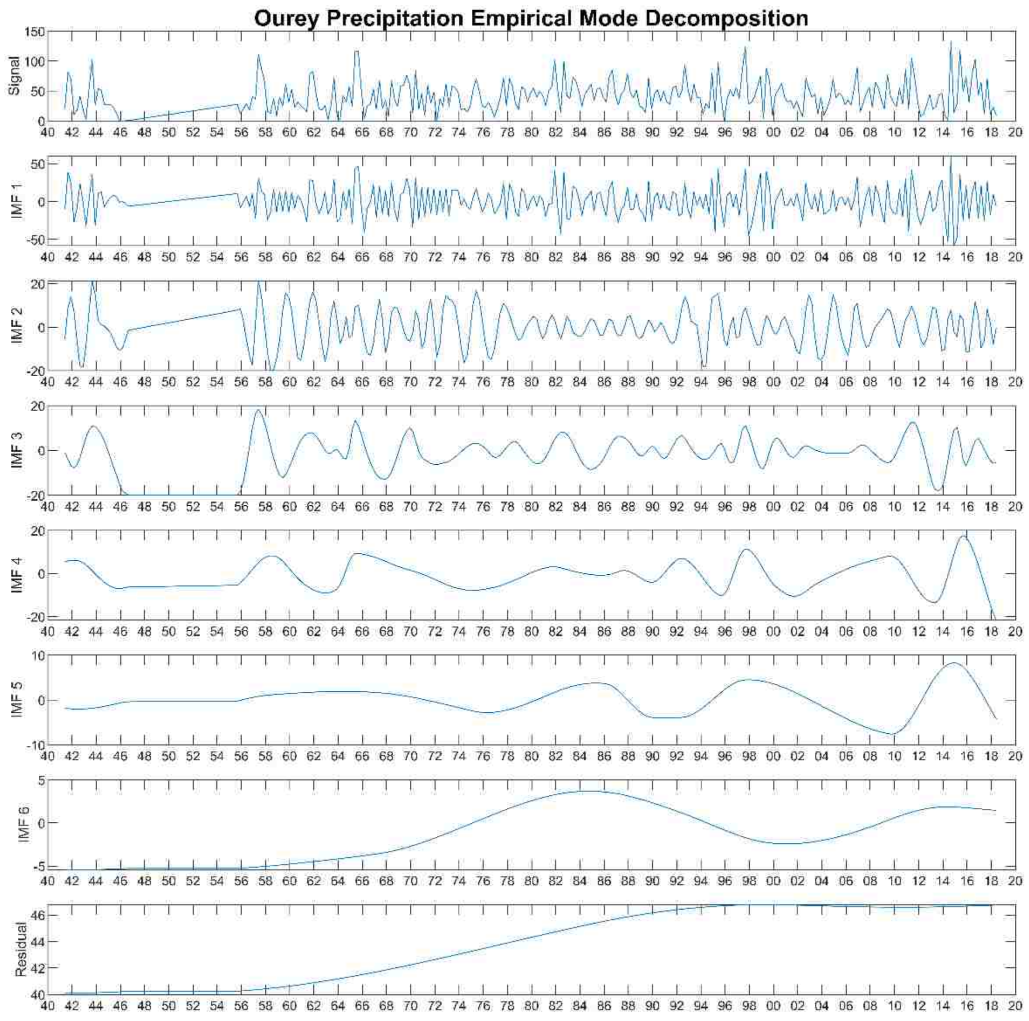


Figure B.5-1: Ourey Quarterly Precipitation Empirical Mode Decomposition

www.manaraa.com

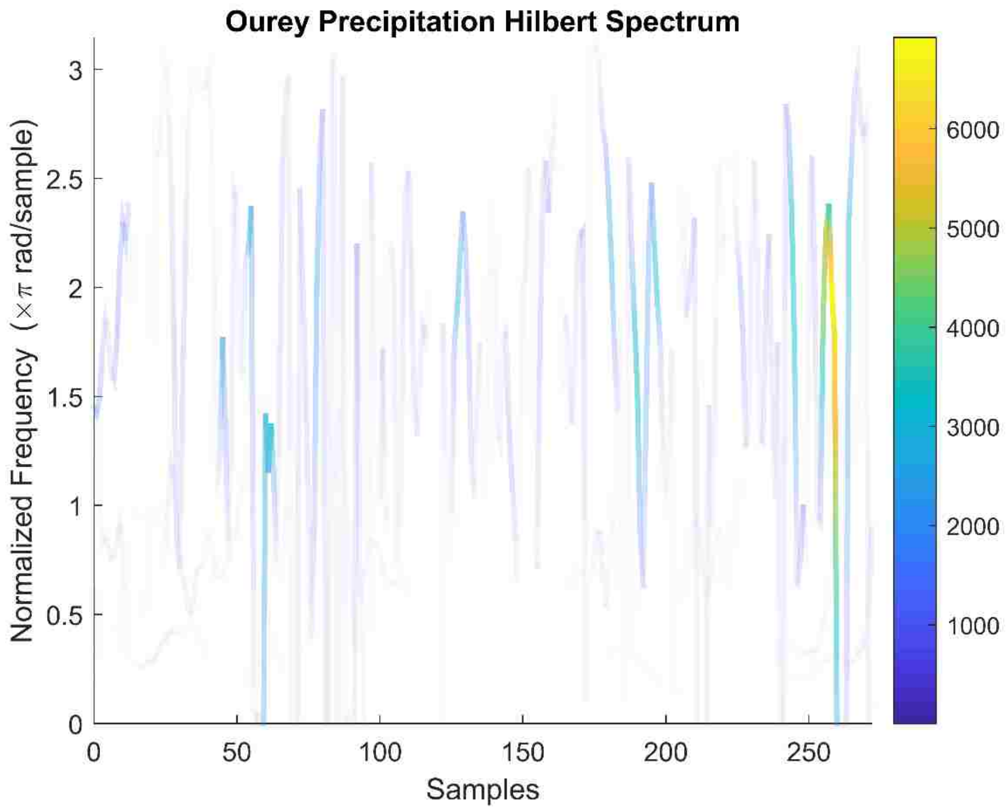


Figure B.5-2: Ourey Quarterly Precipitation Empirical Mode Decomposition HHT

www.manaraa.com

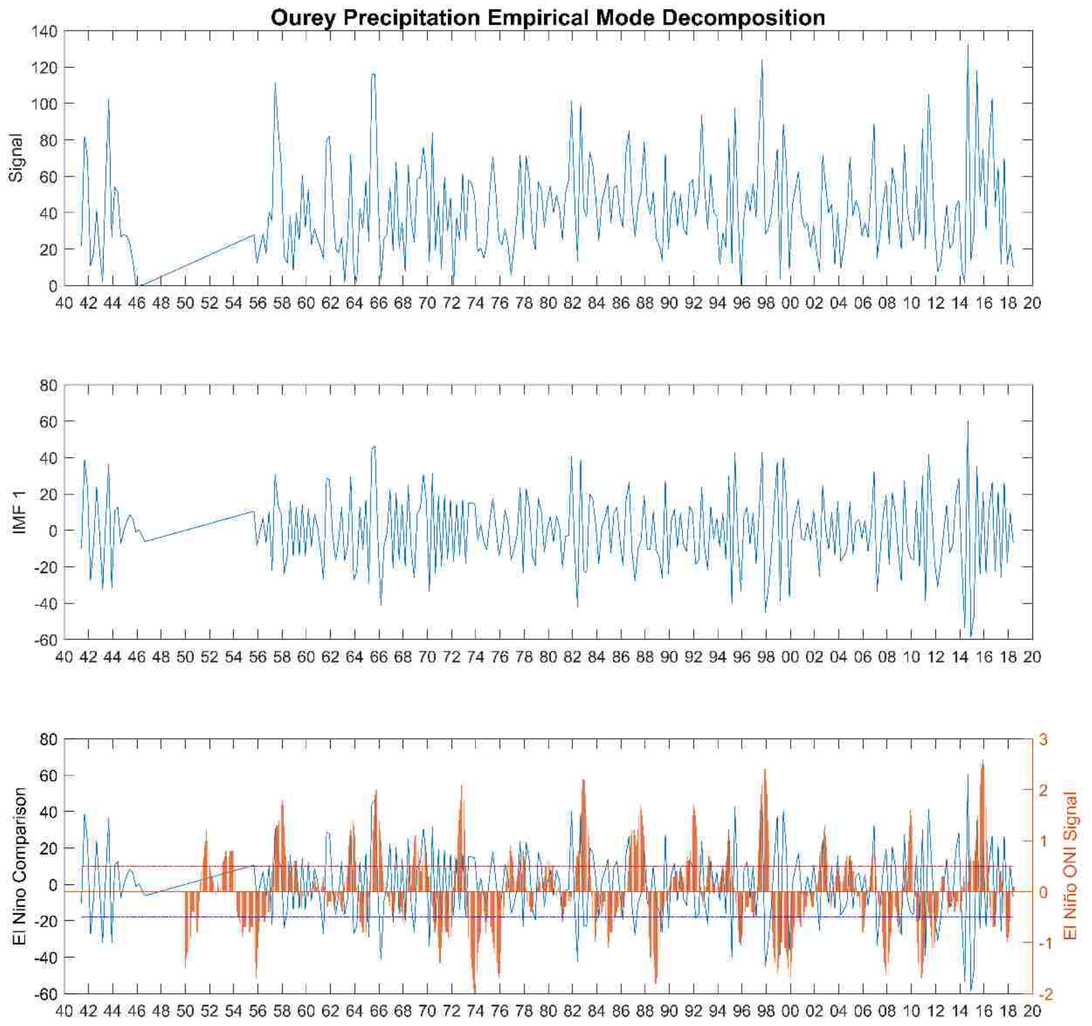


Figure B.5-3: Ourey Quarterly Precipitation Empirical Mode Decomposition IMF 1

www.manaraa.com

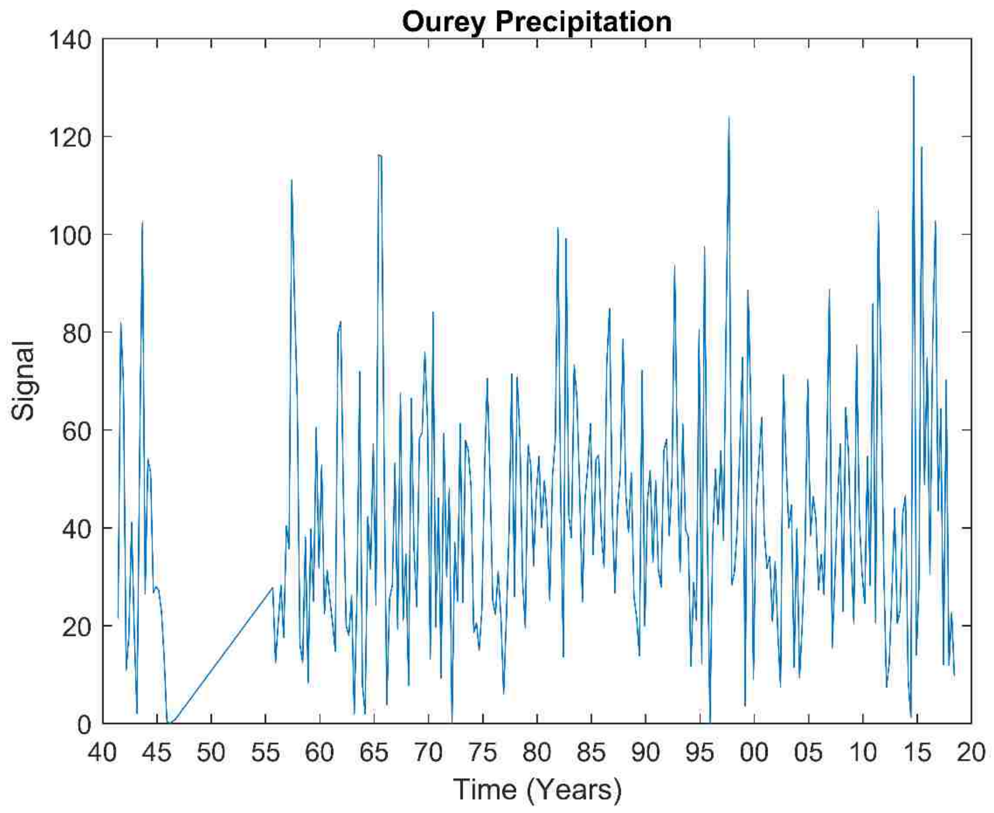


Figure B.5-4: Ourey Quarterly Precipitation

www.manaraa.com

B.6 Jensen Quarterly Precipitation

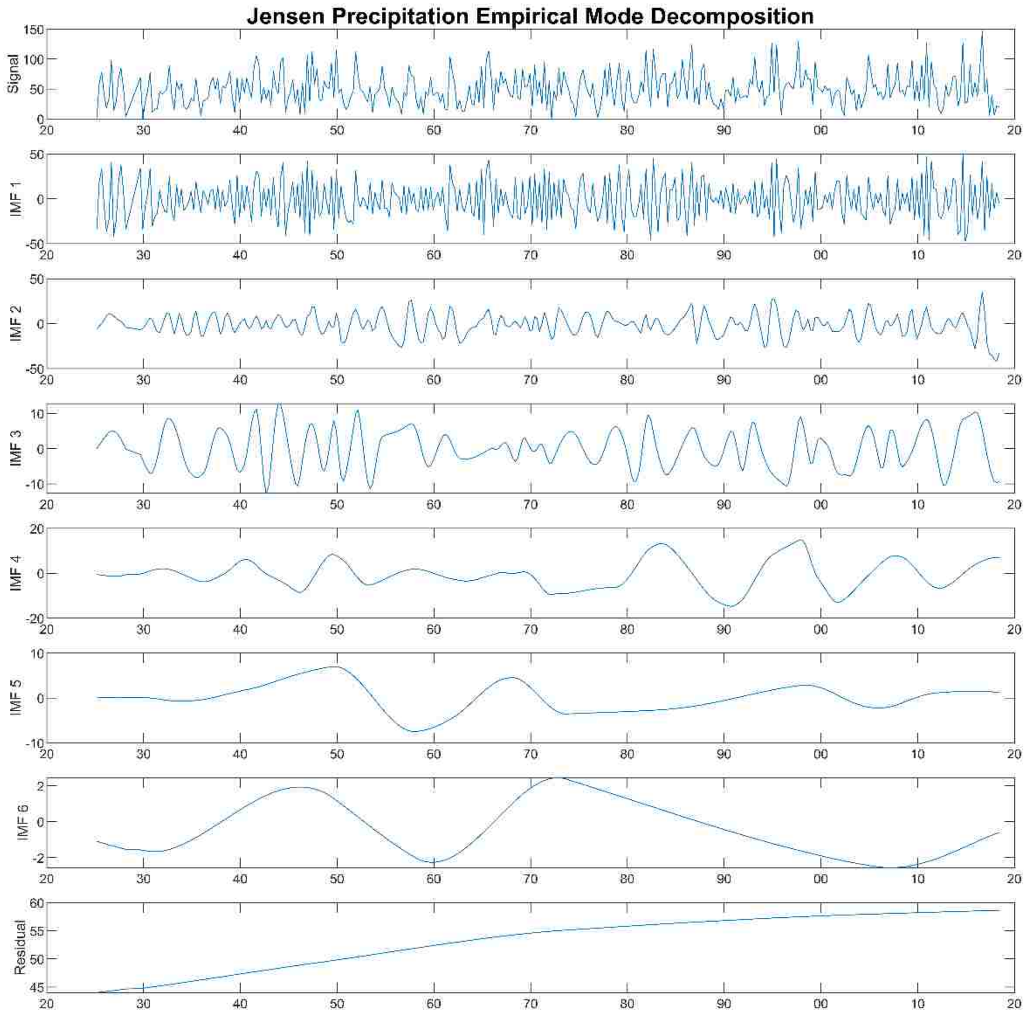


Figure B.6-1: Jensen Quarterly Precipitation Empirical Mode Decomposition

www.manaraa.com

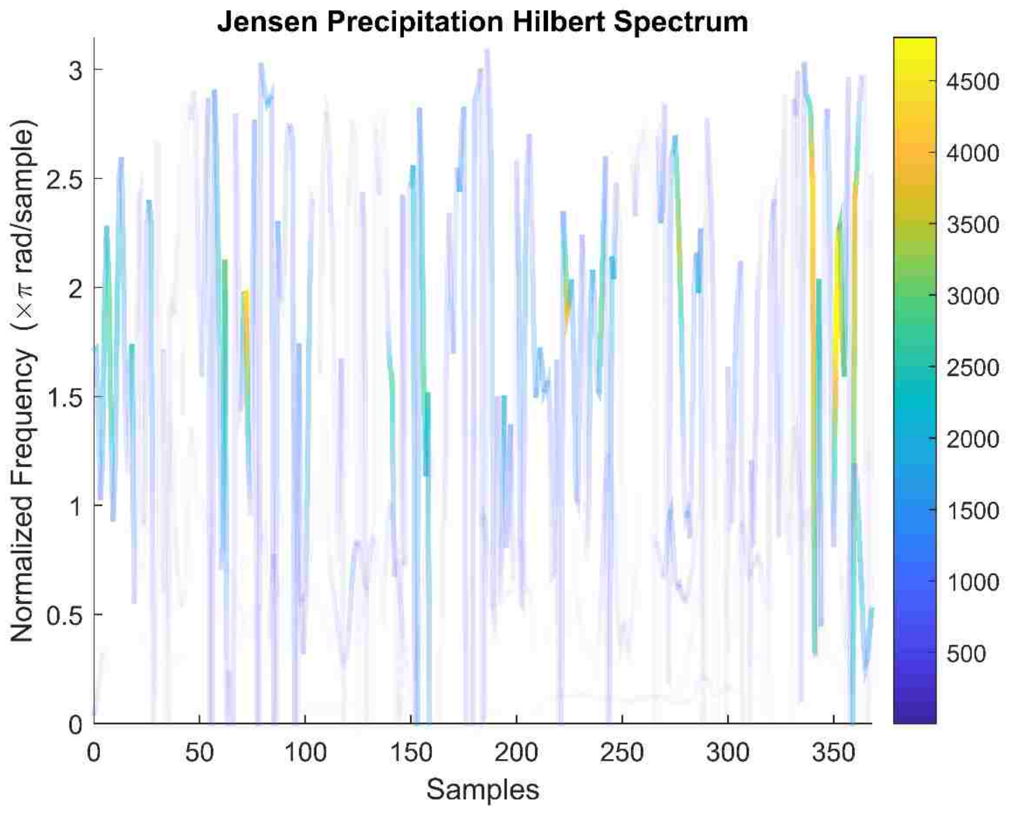


Figure B.6-2: Jensen Quarterly Precipitation Empirical Mode Decomposition HHT

www.manaraa.com

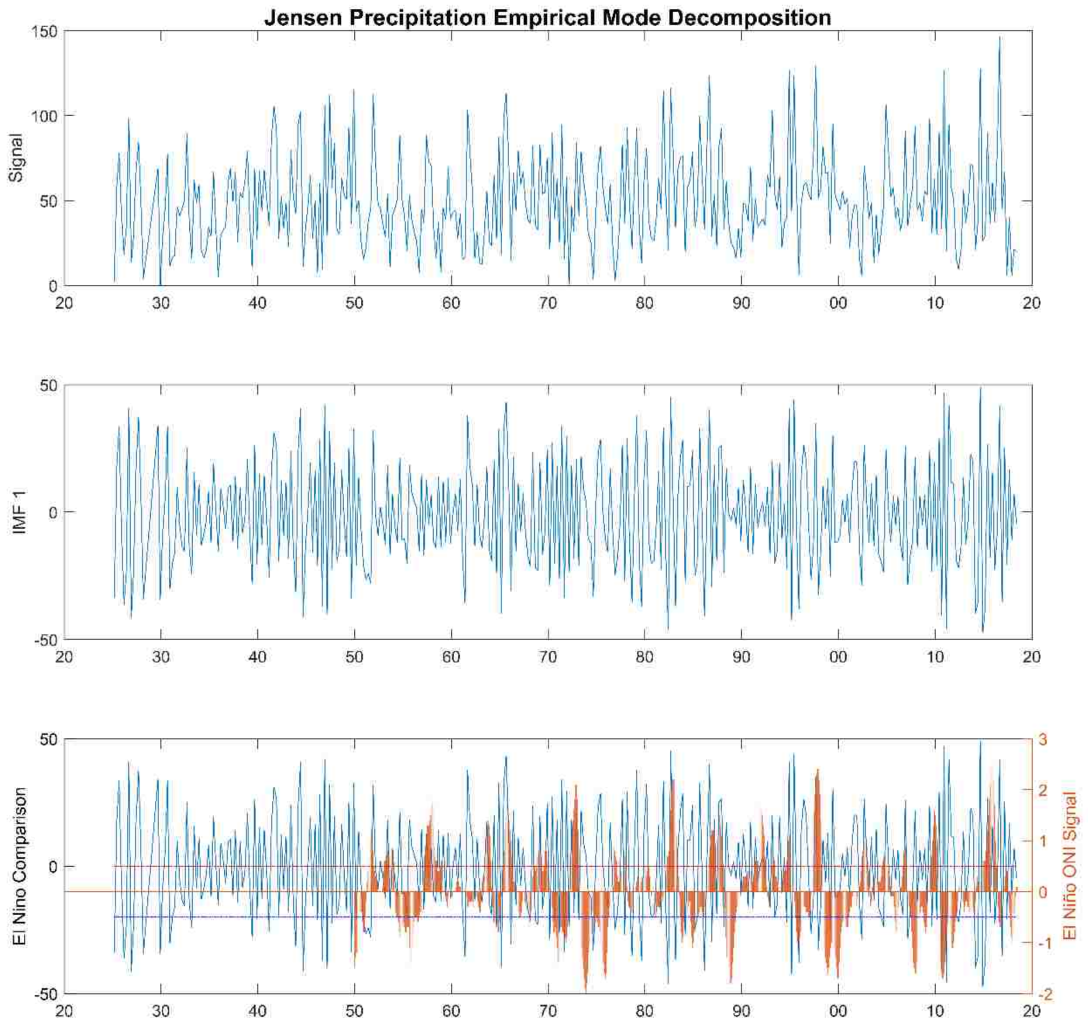


Figure B.6-3: Jensen Quarterly Precipitation Empirical Mode Decomposition IMF 1

www.manaraa.com

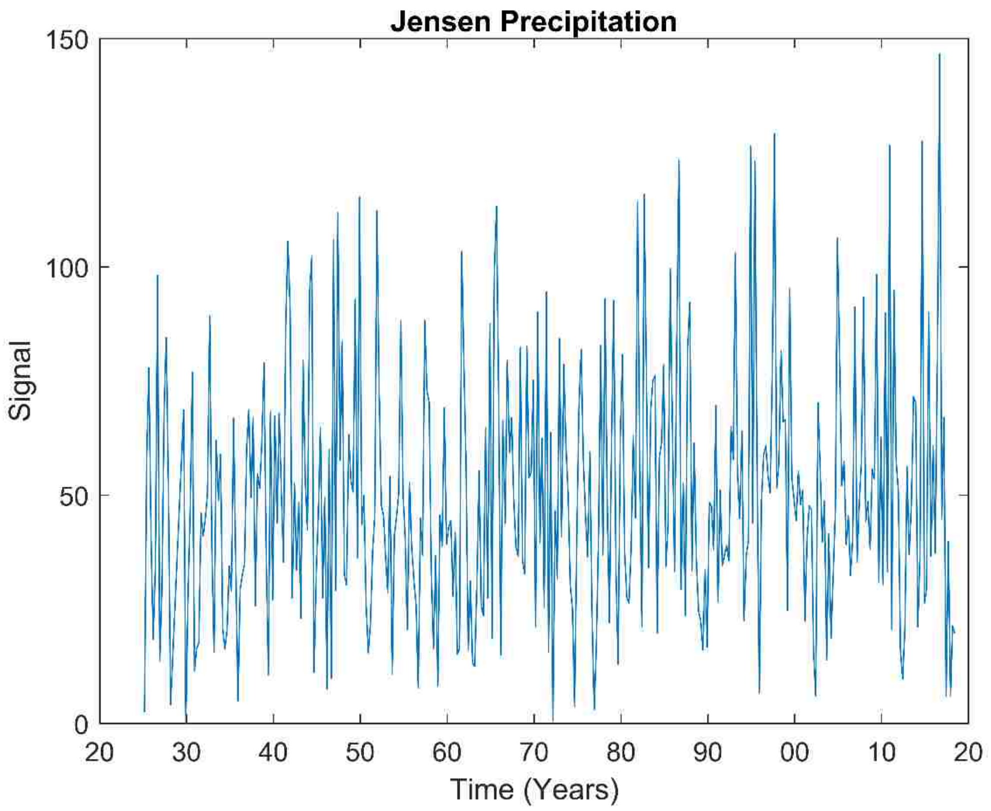


Figure B.6-4: Jensen Quarterly Precipitation

www.manaraa.com

B.7 Ouray Yearly Precipitation

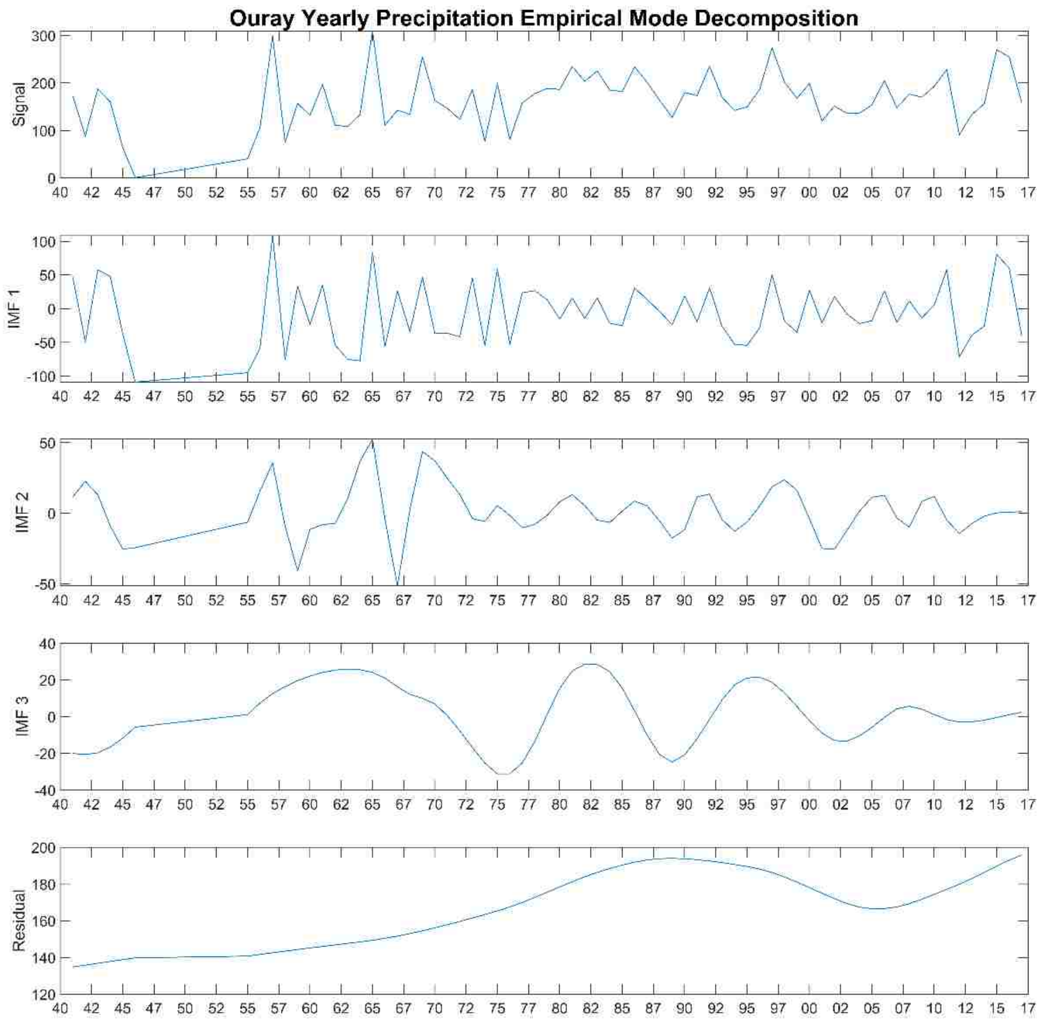


Figure B.7-1: Ouray Yearly Precipitation Empirical Mode Decomposition

www.manaraa.com

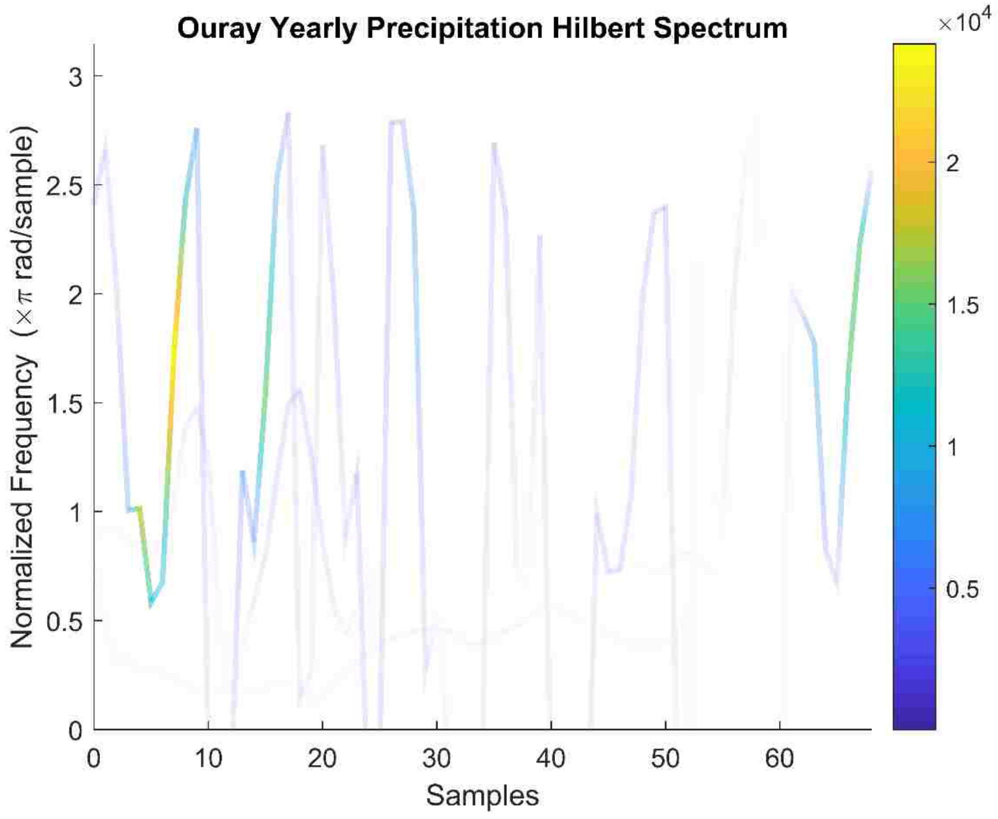


Figure B.7-2: Ouray Yearly Precipitation Empirical Mode Decomposition HHT

www.manaraa.com

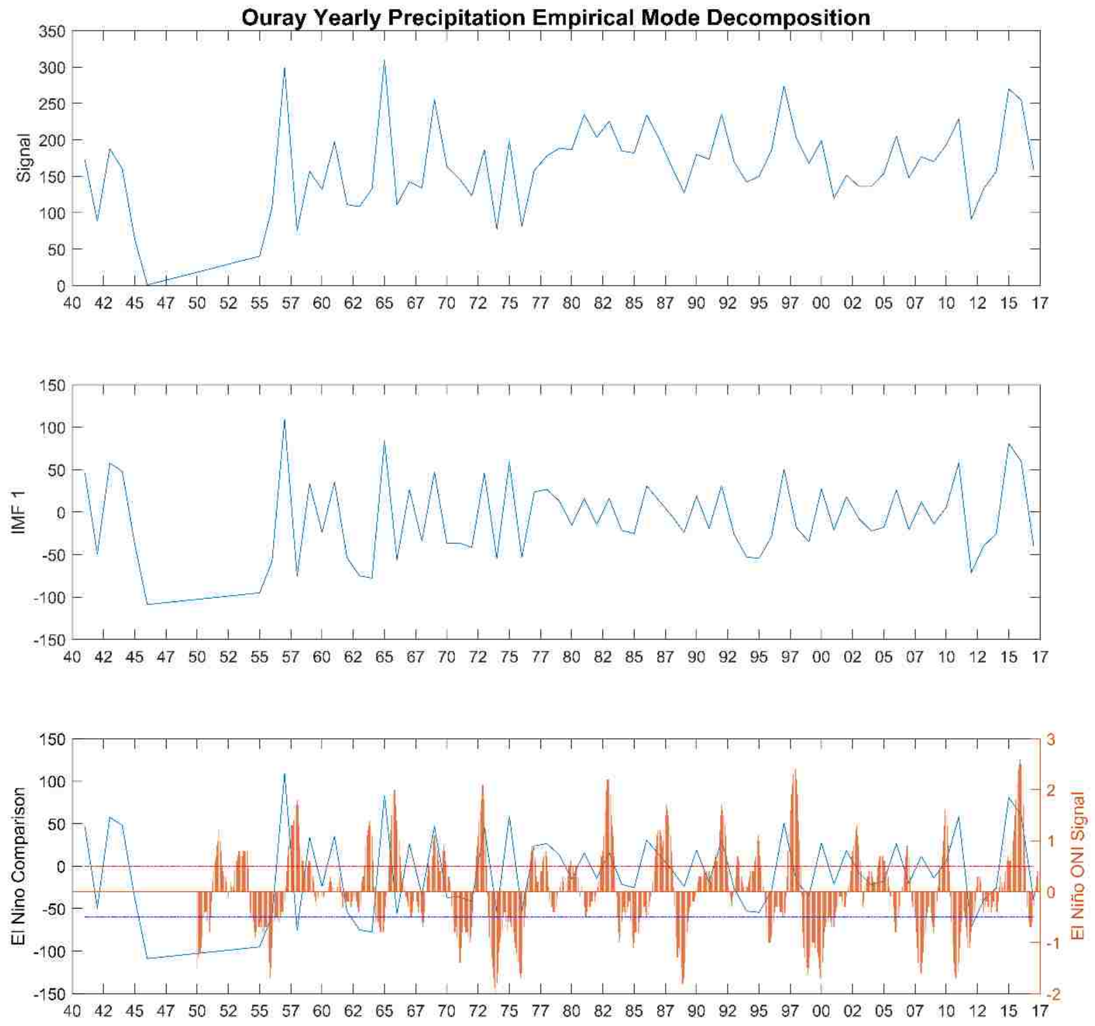


Figure B.7-3: Ouray Yearly Precipitation Empirical Mode Decomposition IMF 1

www.manaraa.com

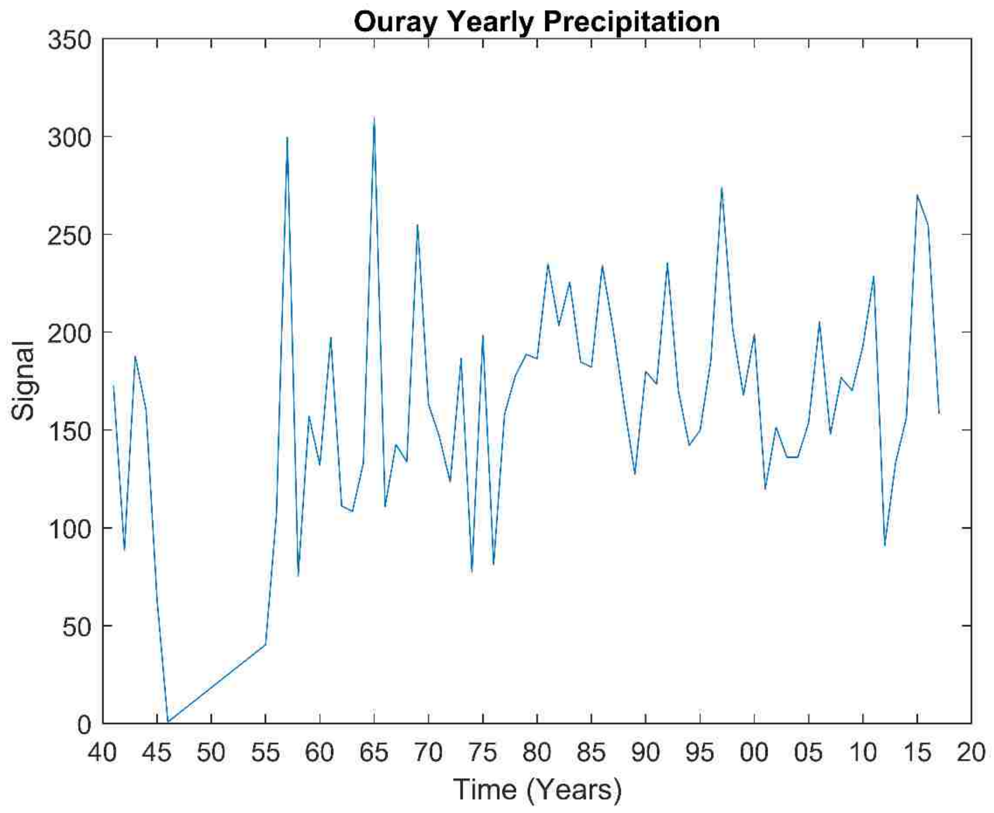


Figure B.7-4: Ouray Yearly Precipitation

المنارة للاستشارات

www.manaraa.com

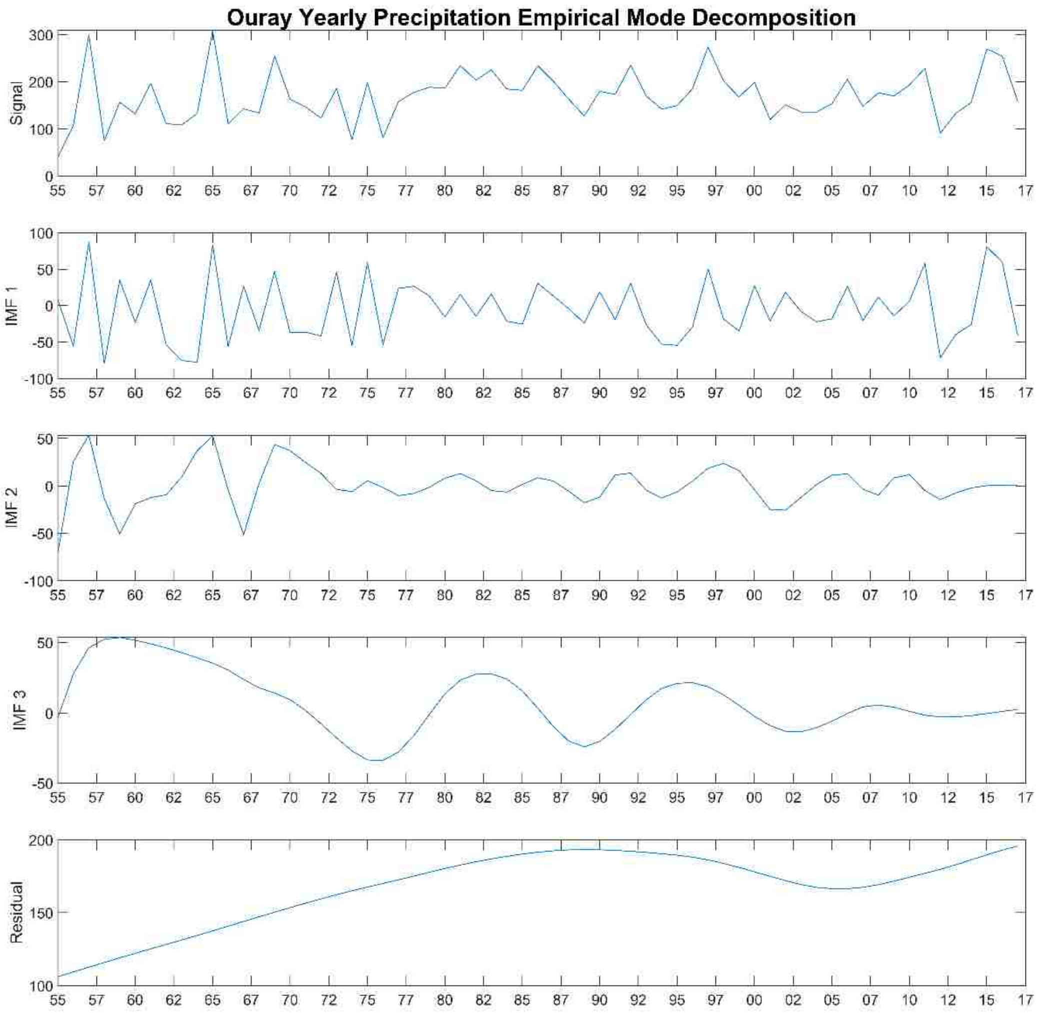


Figure B.7-5: Ouray Yearly Precipitation Empirical Mode Decomposition Shortened to Account for Missing Data

www.manaraa.com

B.8 Jensen Yearly Precipitation

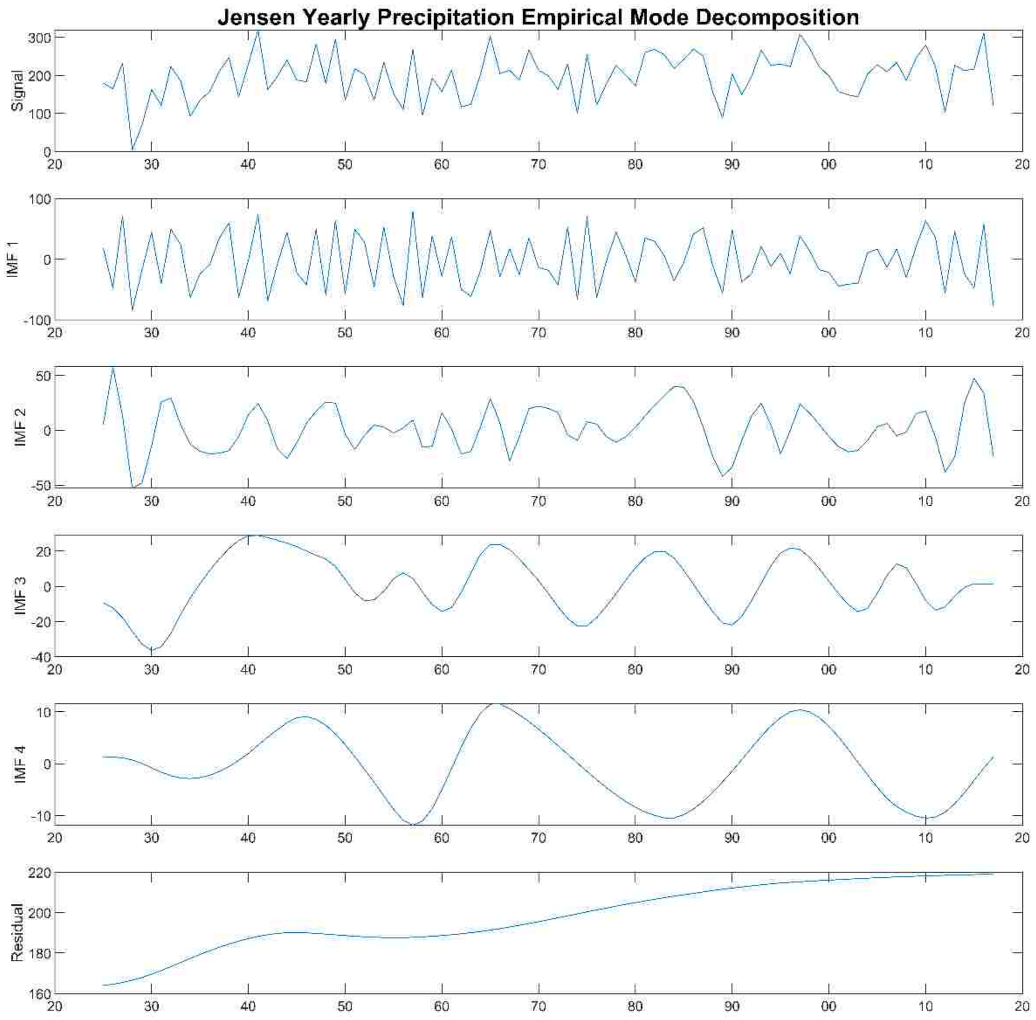


Figure B.8-1: Jensen Yearly Precipitation Empirical Mode Decomposition

www.manaraa.com

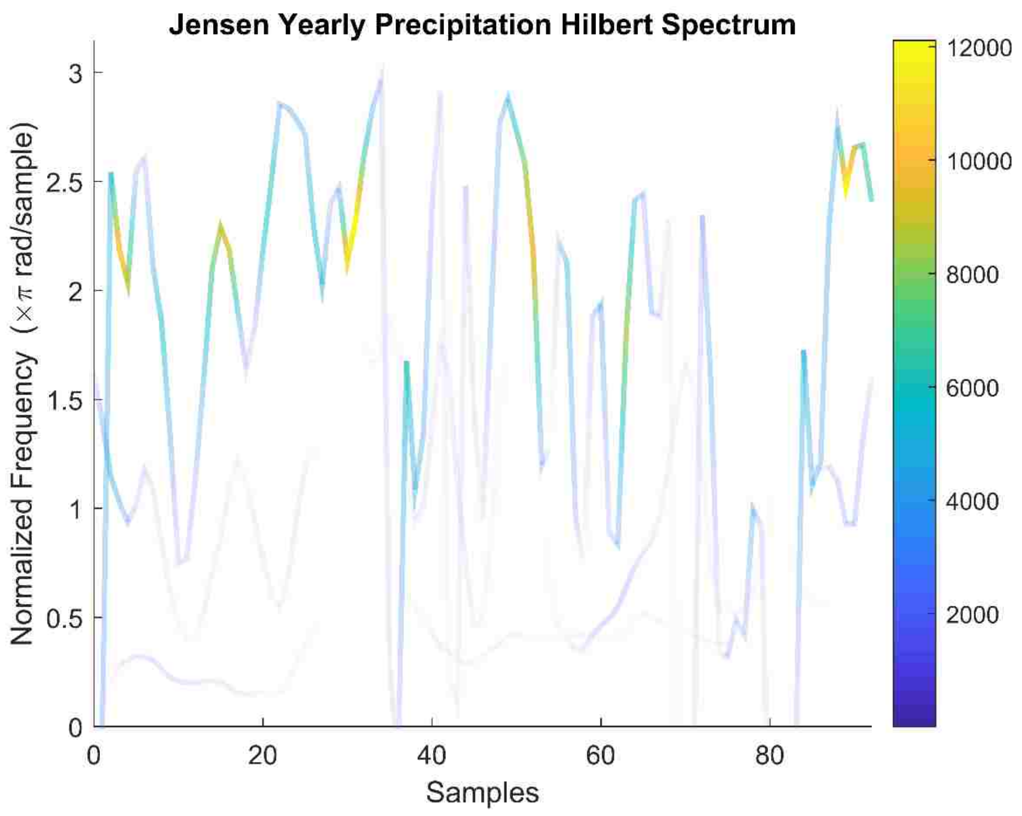


Figure B.8-2: Jensen Yearly Precipitation Empirical Mode Decomposition HHT

www.manaraa.com

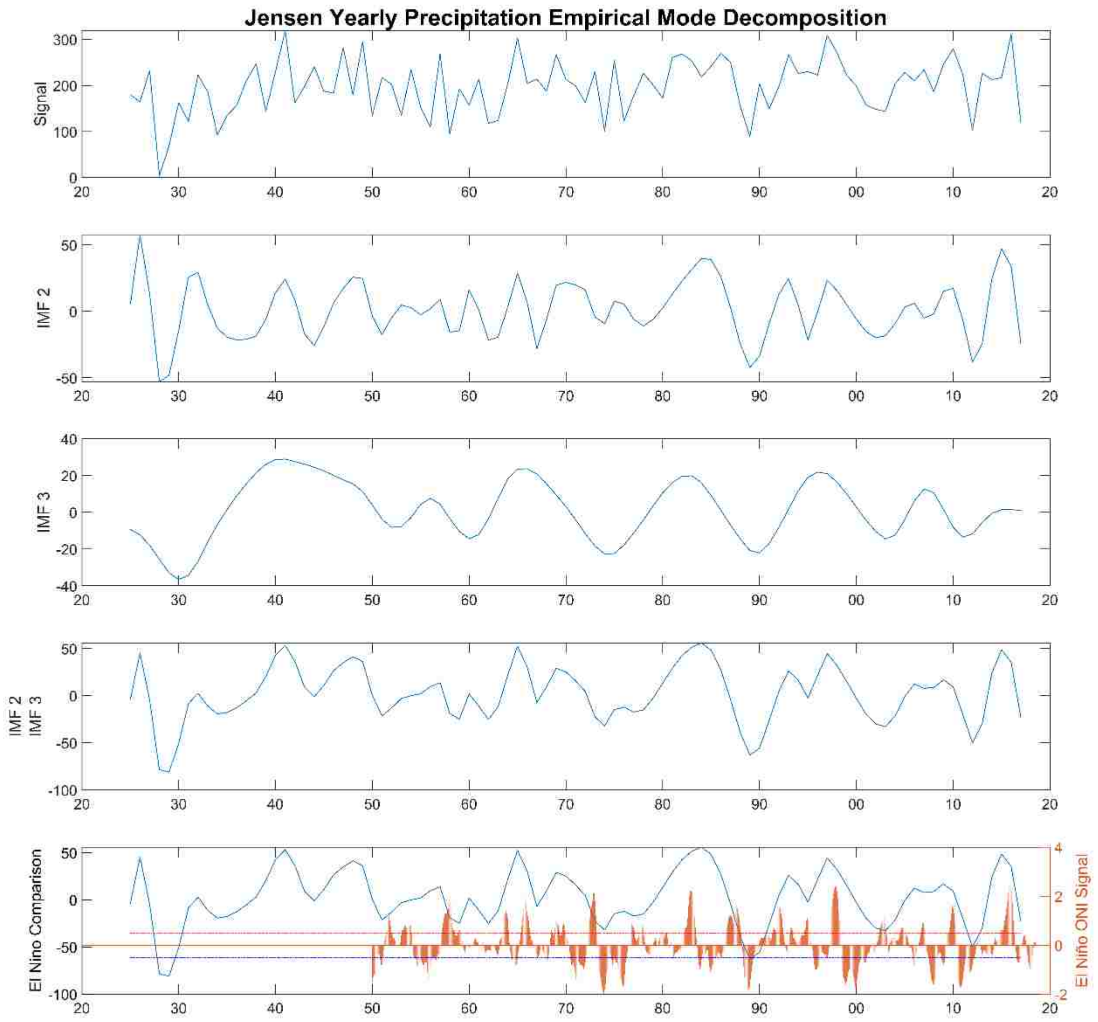


Figure B.8-3: Jensen Yearly Precipitation Empirical Mode Decomposition IMF 2 and 3

www.manaraa.com

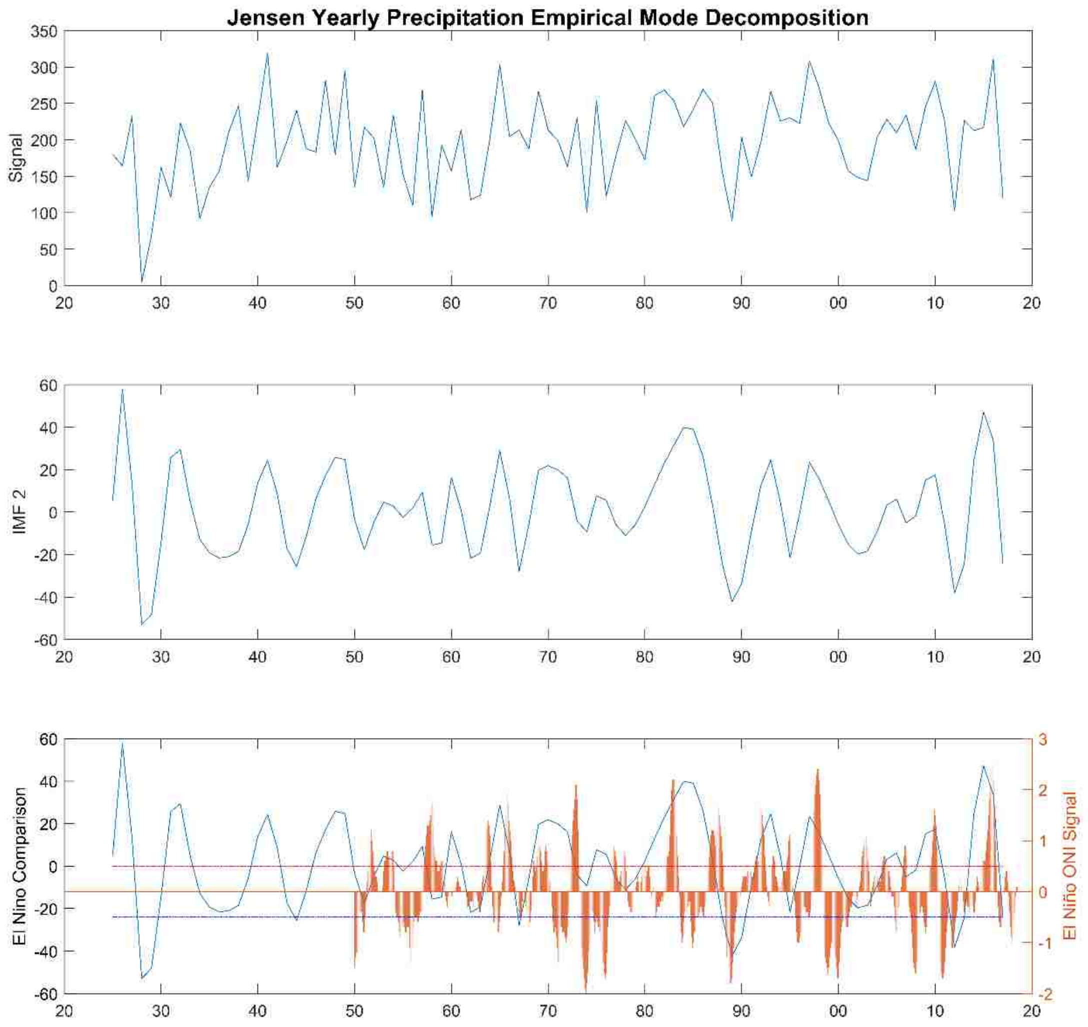


Figure B.8-4: Jensen Yearly Precipitation Empirical Mode Decomposition IMF 2

المنارة للاستشارات
www.manaraa.com

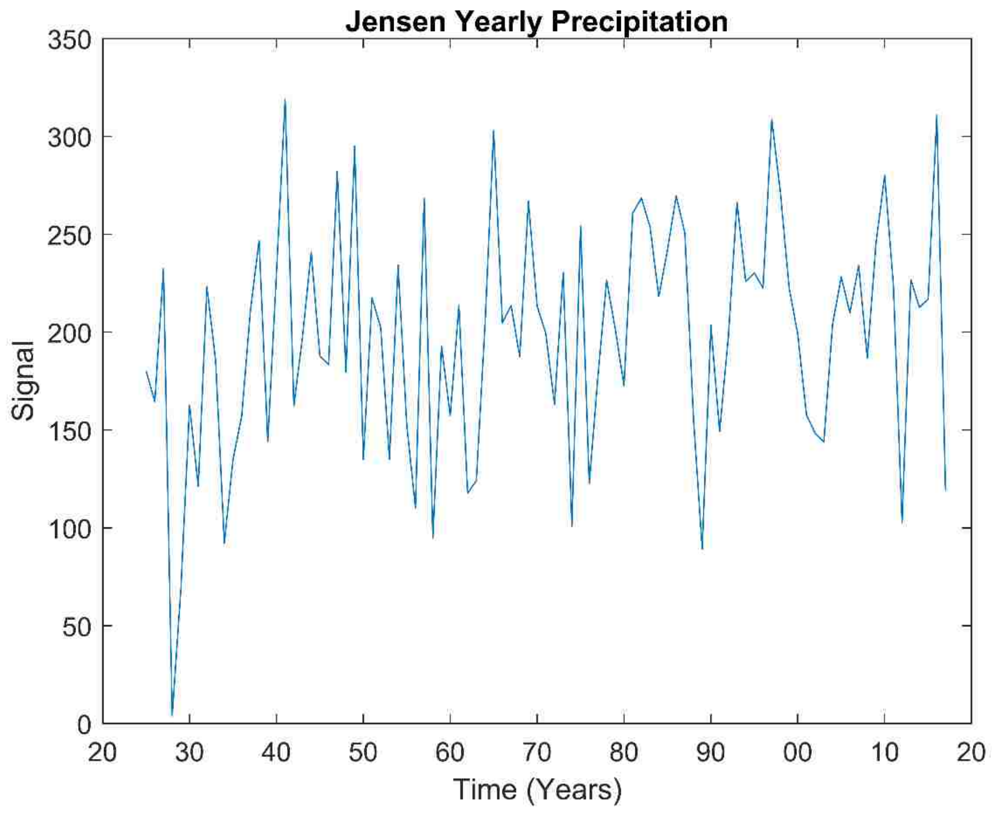


Figure B.8-5: Jensen Yearly Precipitation

www.manaraa.com

B.9 Jensen Temperature

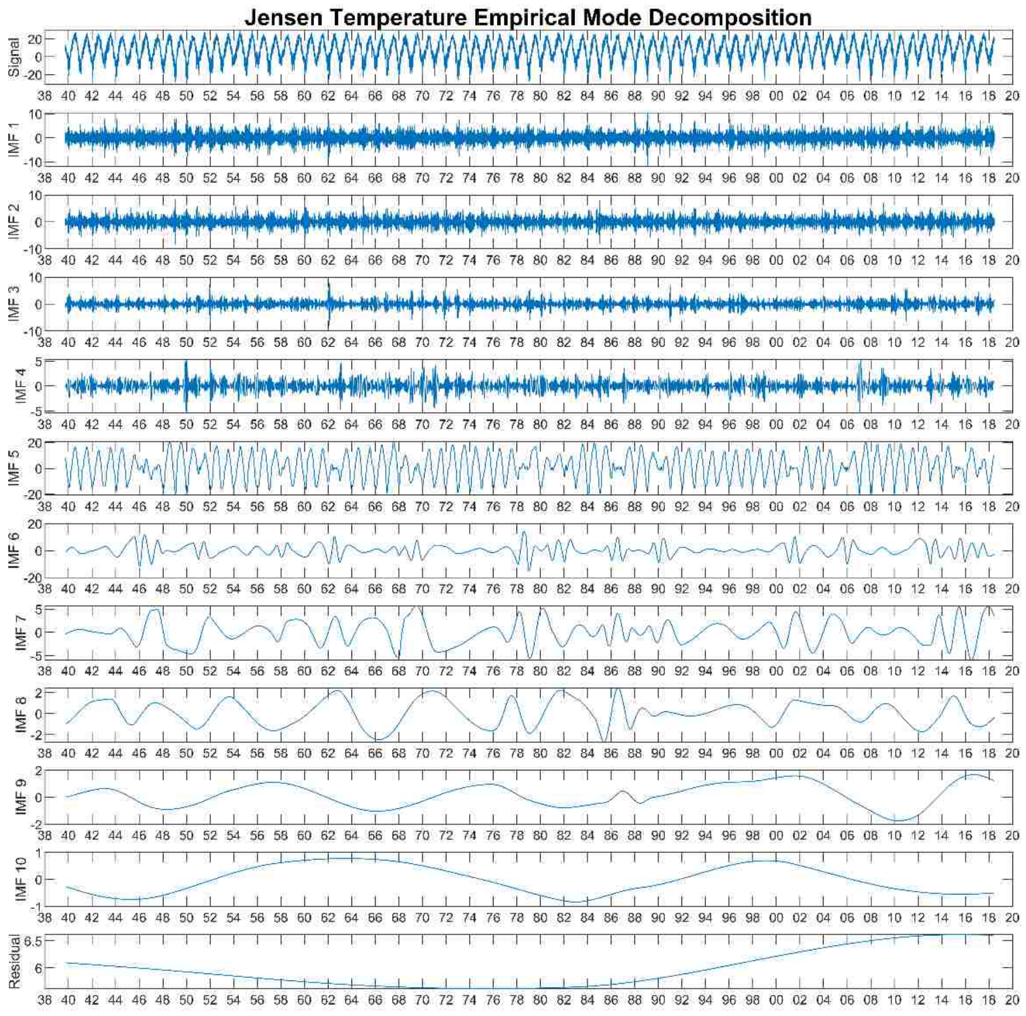


Figure B.9-1: Jensen Average Temperature Empirical Mode Decomposition

www.manaraa.com

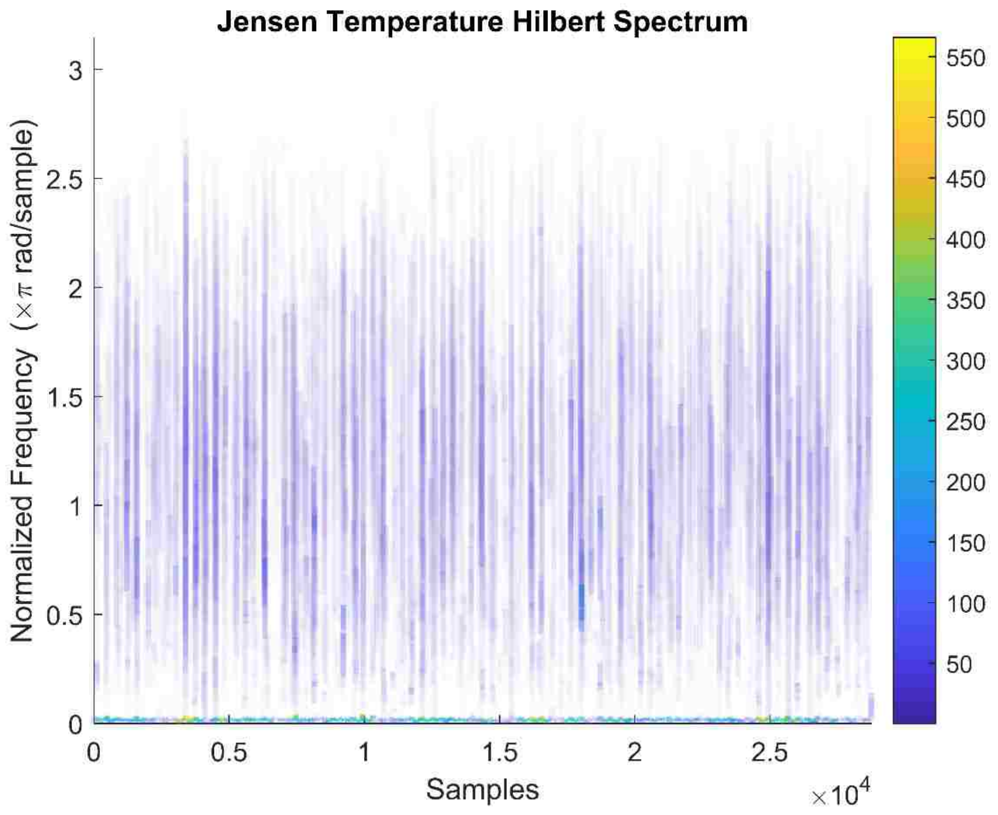


Figure B.9-2: Jensen Average Temperature Empirical Mode Decomposition HHT

المنارة للاستشارات

www.manaraa.com

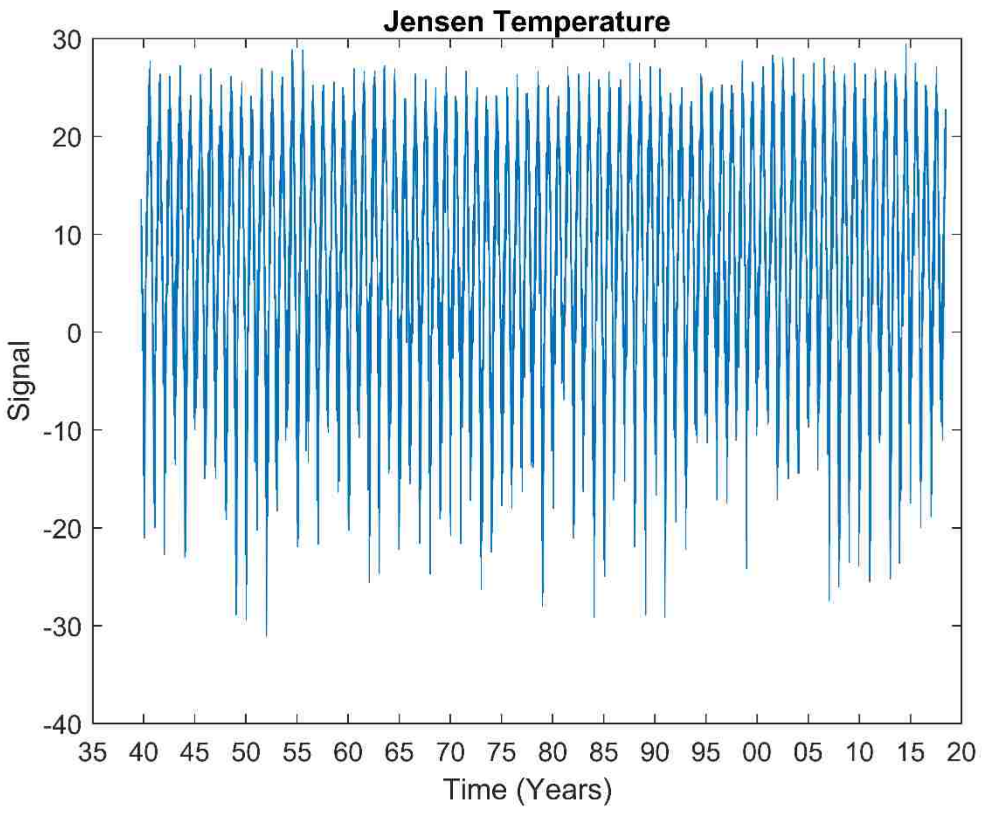


Figure B.9-3: Jensen Average Temperature

المنارة للاستشارات
www.manaraa.com

B.10 Vernal Temperature

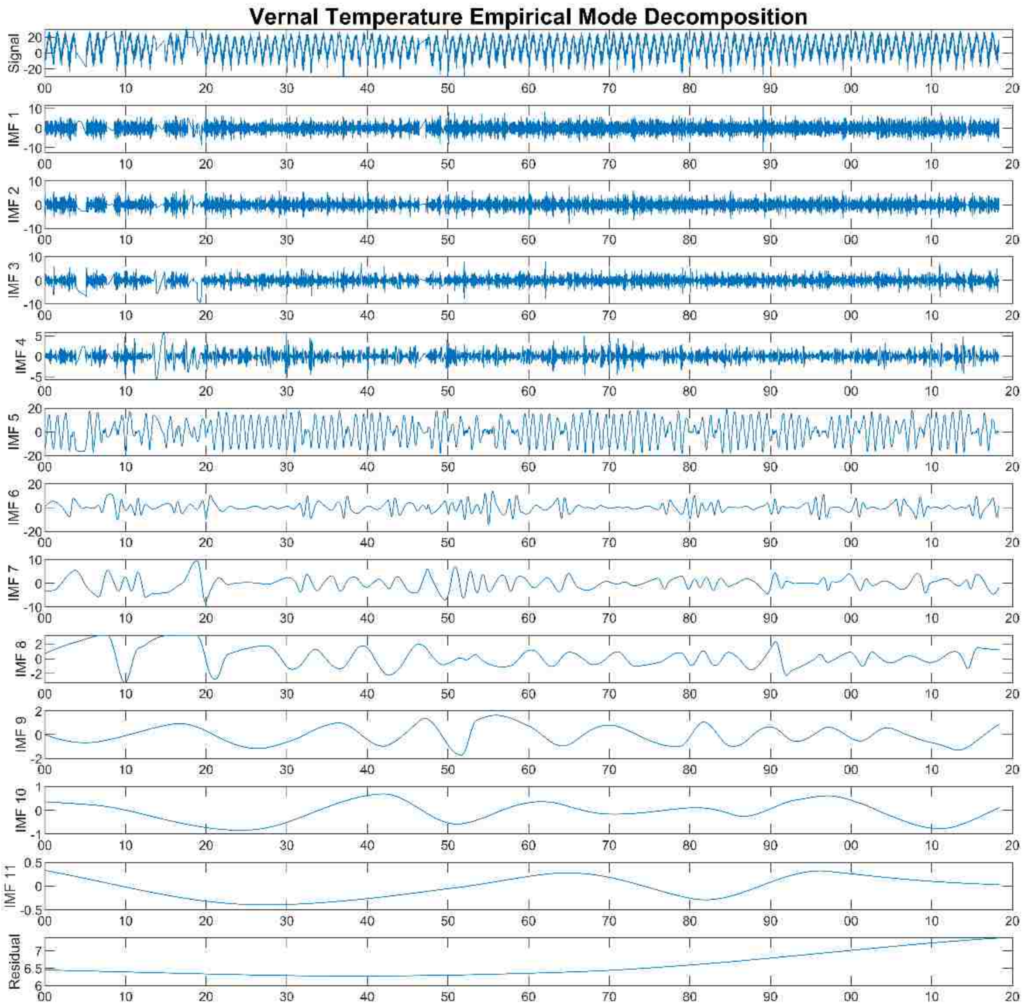


Figure B.10-1: Vernal Average Temperature Empirical Mode Decomposition

www.manaraa.com

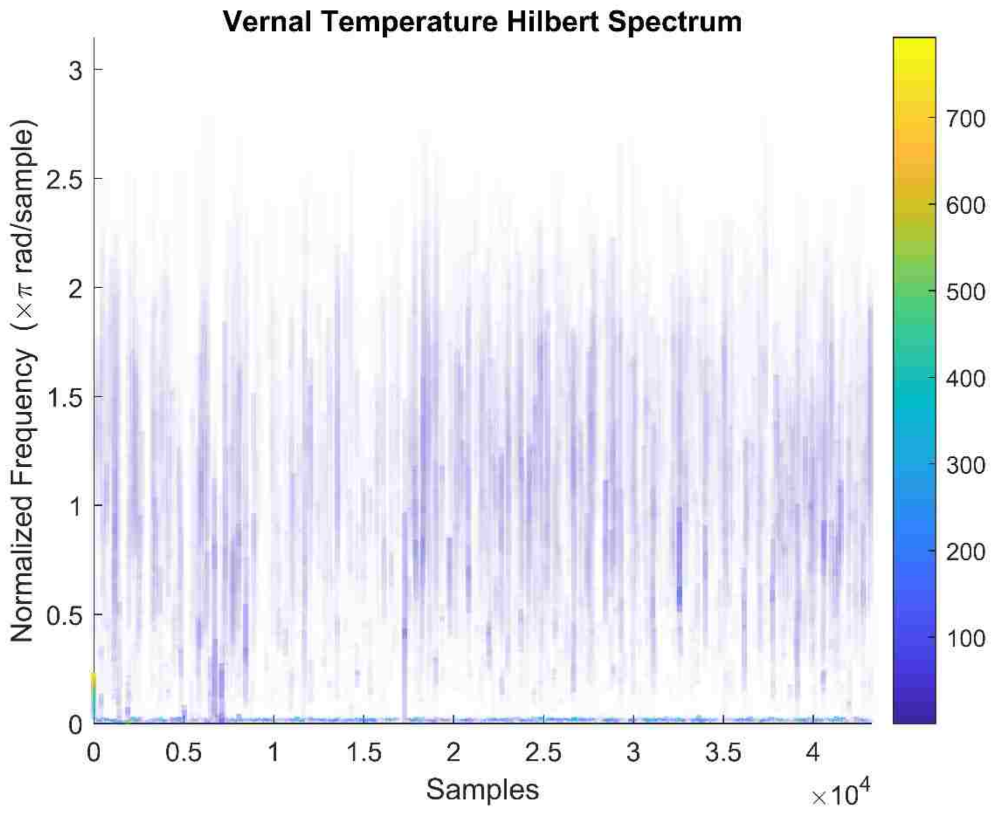


Figure B.10-2: Vernal Average Temperature Empirical Mode Decomposition HHT

www.manaraa.com

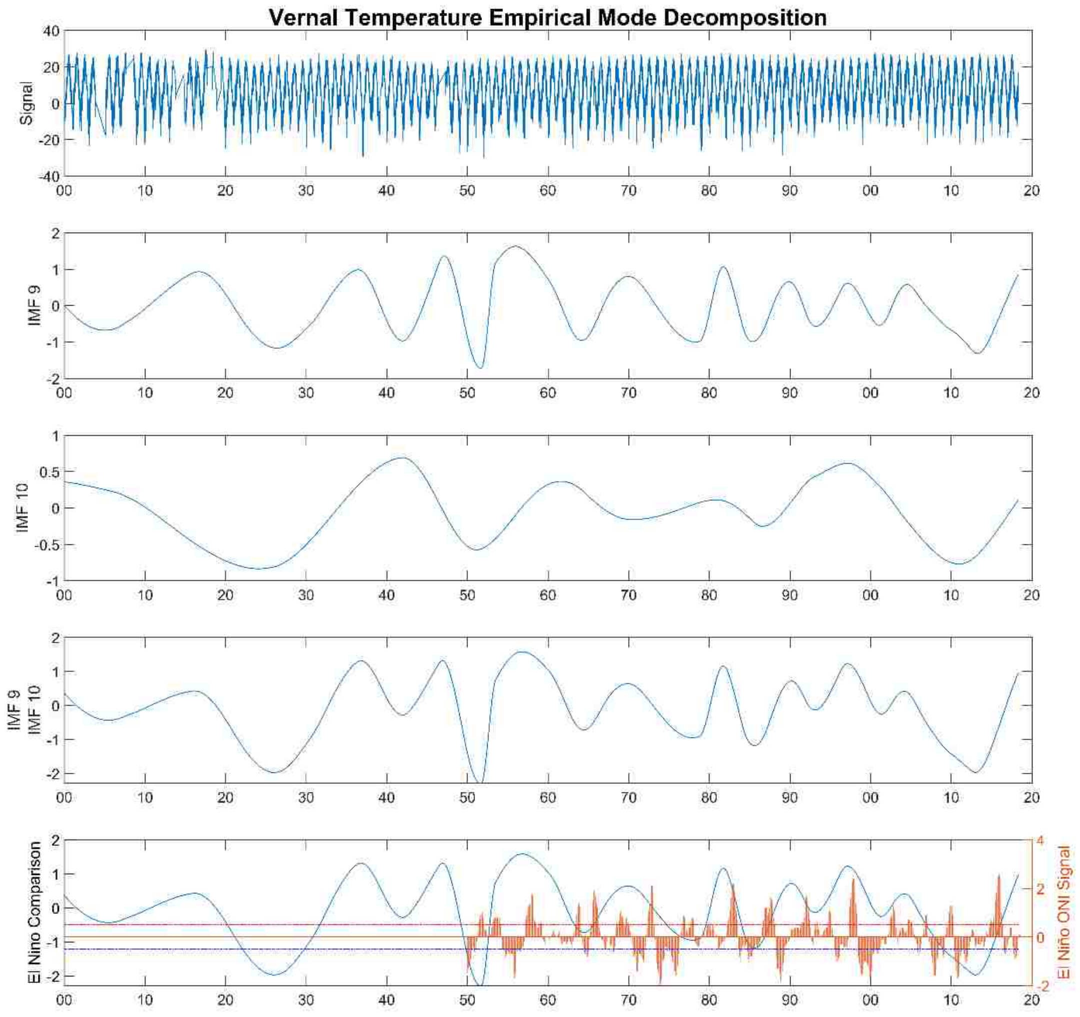


Figure B.10-3: Vernal Average Temperature Empirical Mode Decomposition IMF 9 and 10

المنارة للاستشارات
www.manaraa.com

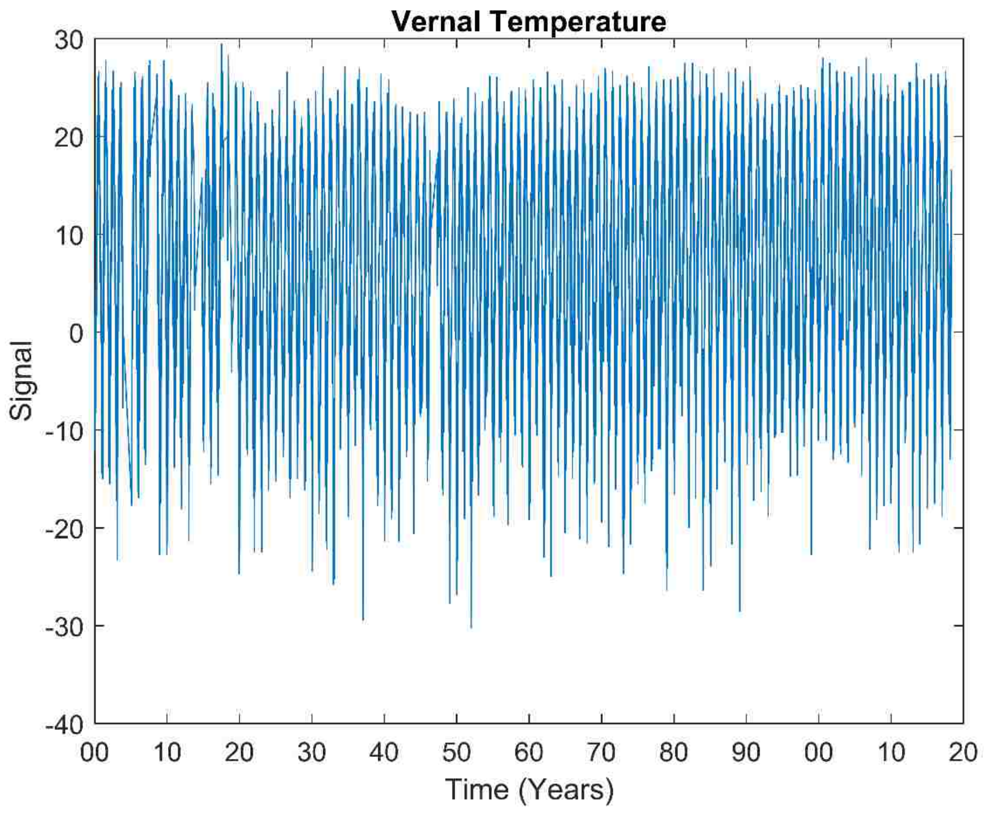


Figure B.10-4: Vernal Average Temperature

www.manaraa.com

B.11 Combination

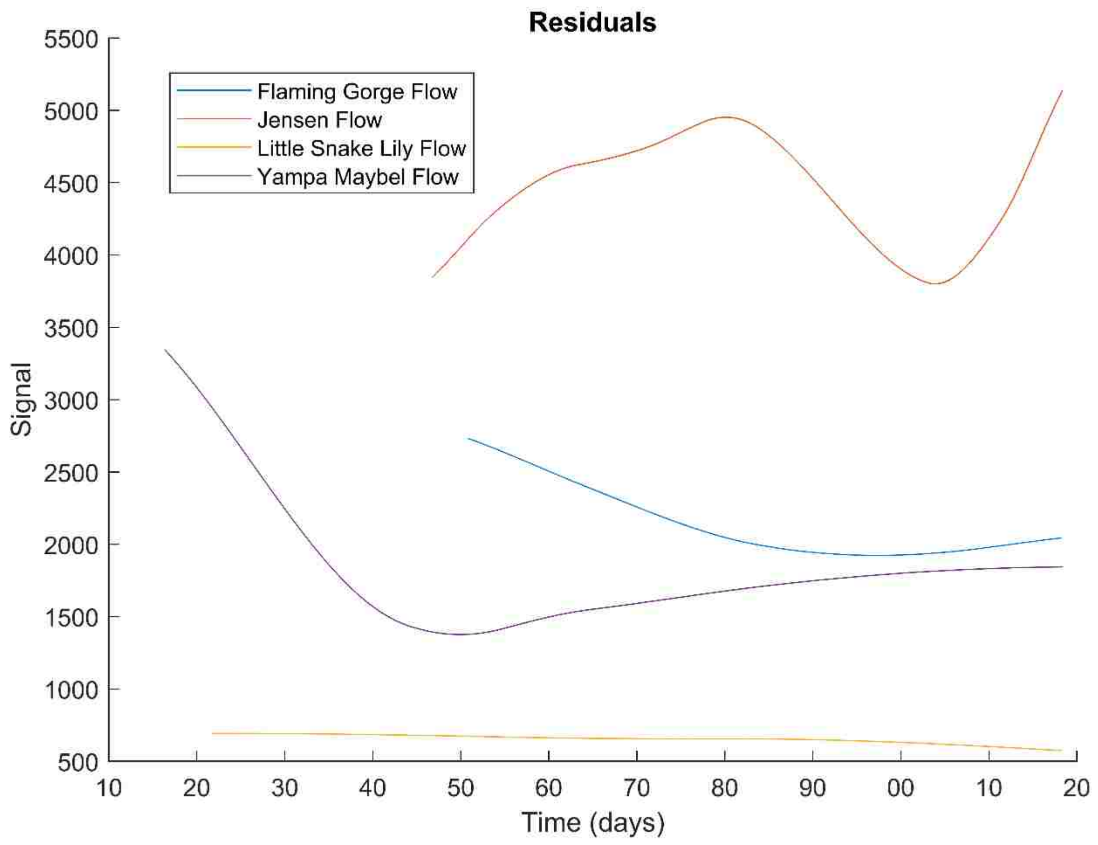


Figure B.11-1: Flow Residuals

www.manaraa.com

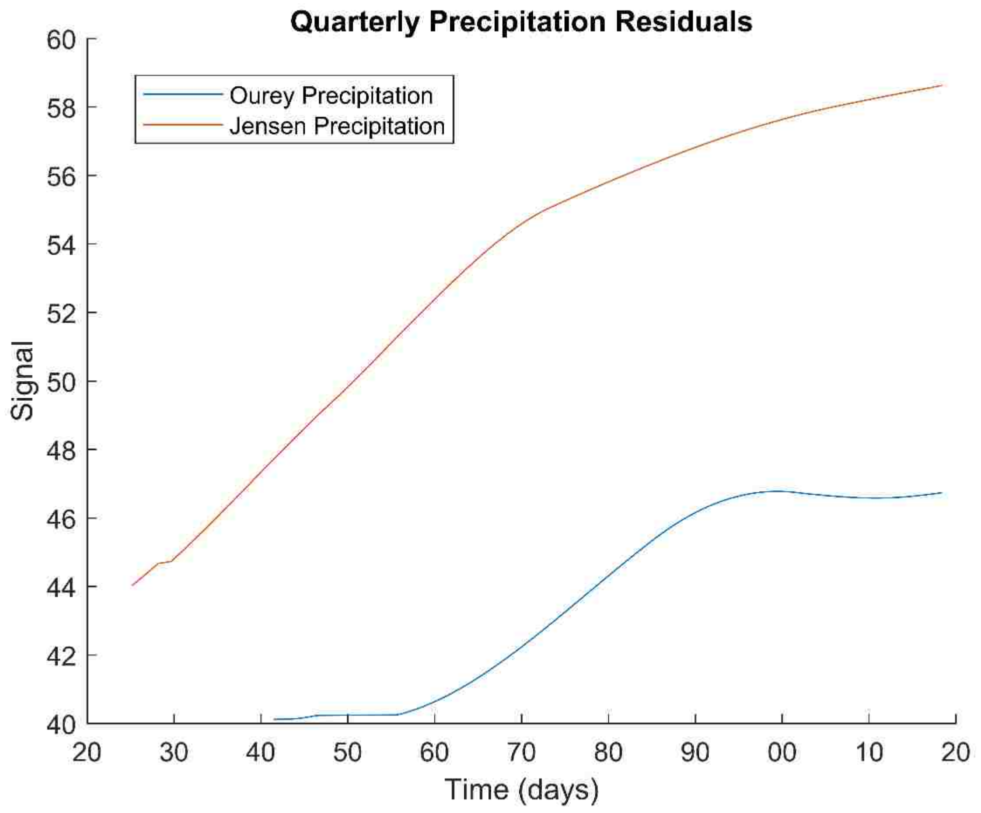


Figure B.11-2: Quarterly Precipitation Residuals

المنارة للاستشارات

www.manaraa.com

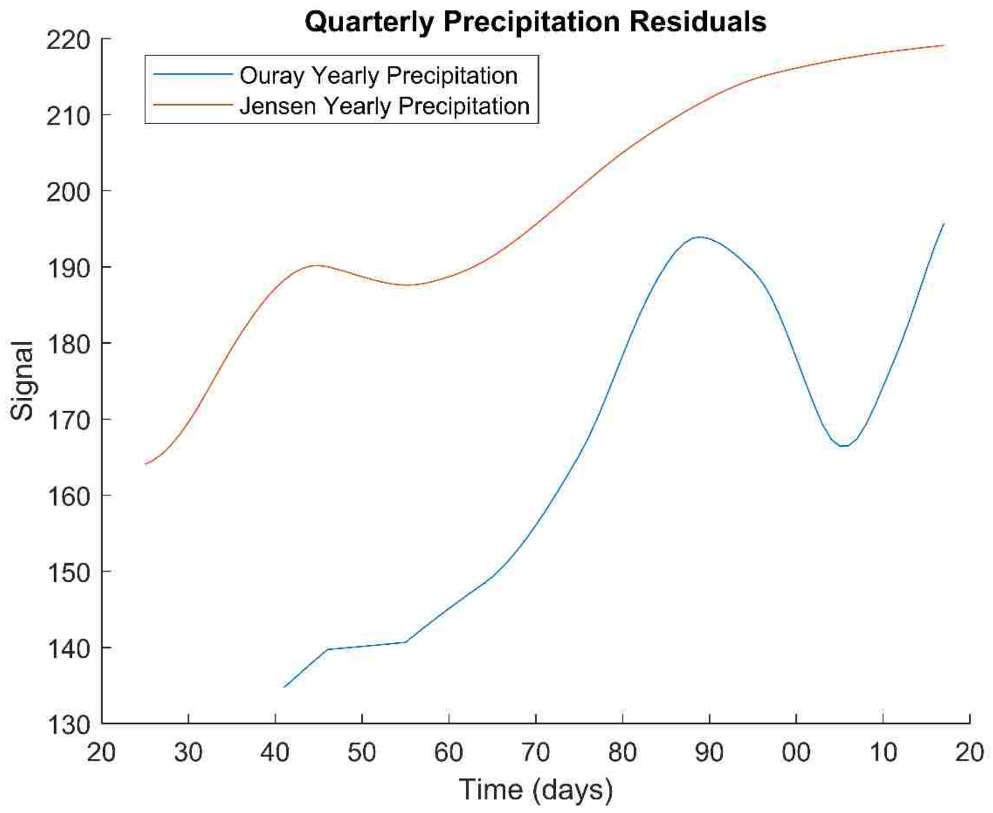


Figure B.11-3: Yearly Precipitation Residuals

www.manaraa.com

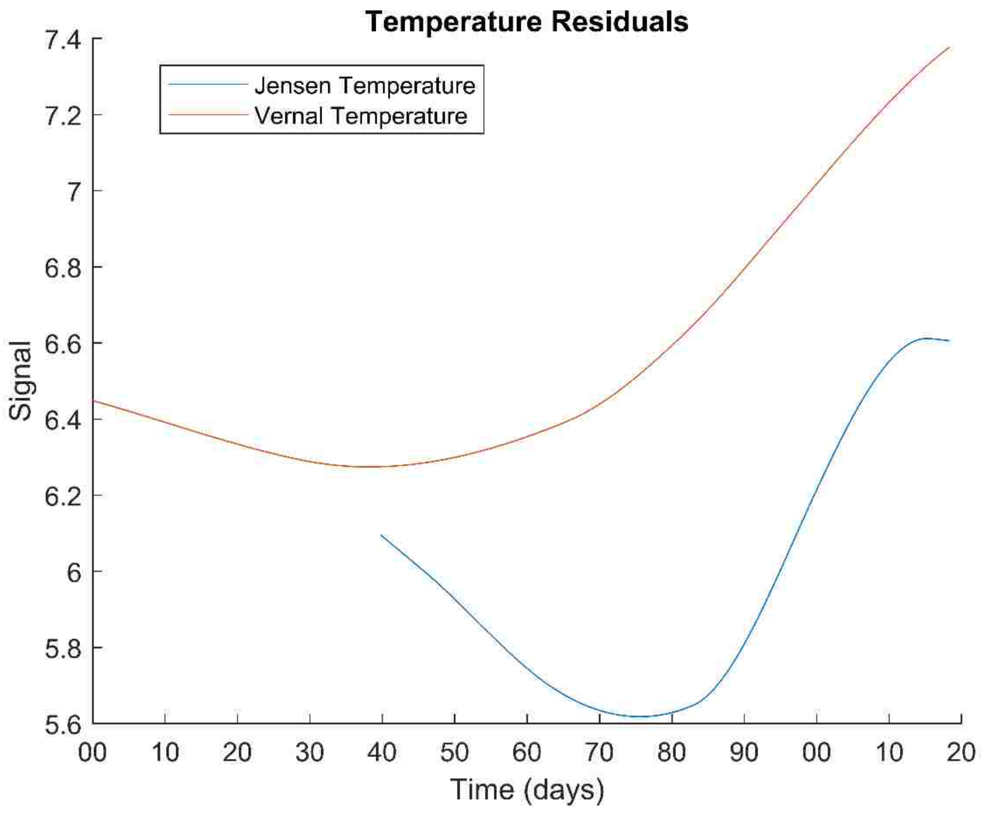


Figure B.11-4: Temperature Residuals

www.manaraa.com