

## **Abstract**

Title of Thesis:       DEVELOPING A SYNTHETIC CONTINUOUS DAILY  
                              STREAMFLOW HYDROGRAPH TECHNIQUE FOR  
                              MARYLAND.

Pallavi Pathak, Master of Science, 2005

Thesis directed by:   Associate Professor Kaye L. Brubaker

Department of Civil and Environmental Engineering

One stumbling block in water resources engineering is the inadequacy of stream flow records. Records available from many streams are of insufficient length for use in design and prediction, and for some streams no streamflow data have been recorded. Risk-based methods may require large sets of possible stream flow scenarios. The goal of this study is to develop a method to extend gage records realistically, and allow synthesis of realistic daily discharge hydrographs. The original purpose of the study was to provide long duration (100-year) daily discharge hydrographs for a mathematical model that predicts bridge pier scour. A Markov model with time-varying parameters is developed and implemented in Matlab. The model captures the annual cycle and day-to-day persistence of daily streamflow, using 10 parameters estimated from gage records, without precipitation input or watershed modeling. The Matlab program reads text files of stream flow records, estimates the model parameters, and saves results in spreadsheet format. Seventy-eight stream gages in Maryland, with drainage areas from 2 to 27000  $mi^2$ , are analyzed. Several examples demonstrate the use of the model in synthesis mode. The model will be useful for any prediction or design tool requiring realistic synthetic discharge hydrographs.

**Developing a Synthetic Continuous Daily Streamflow Hydrograph Technique for  
Maryland**

**By**

**Pallavi Pathak**

**Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Science  
2005**

**Advisory Committee:**

**Associate Professor Kaye L. Brubaker, Chair  
Professor Richard H. McCuen  
Associate Professor Glenn E. Moglen**

**© Copyright by**

**Pallavi Pathak**

**2005**

## ACKNOWLEDGEMENTS

Throughout the life of this work, my professors, friends and family have given me priceless assistance. I would especially like to take this opportunity to thank the people whose contributions and support proved invaluable to the success of this thesis. I would like to thank my advisor, Dr. Kaye. L. Brubaker, for the opportunity, both financially and intellectually to accomplish my goals in my educational endeavors. Dr. Brubaker's patience and understanding gave me a tremendous support in completing this work.

My parents and sister have motivated and supported me through out the work. Without their assurance and guidance, it would be difficult to think of where I am now.

I would also like to thank my friends, both leisure and at work, for encouragement, and sharing valuable informations, which helped me to complete this work successfully.

Maryland State Highway has provided the funding for this project Grant Number SP107B4E. Their support is greatly acknowledged.

## TABLE OF CONTENTS

CHAPTER 1	INTRODUCTION .....	1
1.1	RESEARCH CONTEXT .....	1
1.2	RESEARCH OBJECTIVE .....	2
1.3	DESCRIPTION OF THE STUDY AREA .....	3
1.4	RESEARCH IMPLICATIONS .....	3
1.5	OVERVIEW .....	4
CHAPTER 2	HYDROGRAPHS: CONCEPTS AND EXISTING APPROACHES .....	5
2.1	NEED FOR SYNTHETIC CONTINUOUS HYDROGRAPH .....	5
2.2	DEFINITIONS.....	5
2.2.1	HYDROGRAPH AND HYETOGRAPH.....	5
2.2.2	EVENT HYDROGRAPH.....	6
2.2.3	CONTINUOUS HYDROGRAPH.....	7
2.2.4	UNIT HYDROGRAPH .....	7
2.3	UNIT HYDROGRAPH METHOD .....	8
2.4	WATERSHED MODELLING METHODS.....	9
2.5	MARKOV MODEL.....	12
2.6	SALAS'S METHOD .....	16
2.7	THIS STUDY'S CONTRIBUTION.....	17
CHAPTER 3	THEORY AND METHODS .....	18
3.1	OVERVIEW .....	18
3.2	ASSEMBLING STUDY DATA.....	18
3.3	DATA TRANSFORMATION.....	18
3.4	IDENTIFYING AND REMOVING SEASONALITY .....	18
3.5	ANALYZING TEMPORAL CORRELATION .....	19
3.6	CURVE FITTING SEASONAL STATISTICS .....	20
3.7	DECORRELATING THE DATA .....	21
3.8	ANALYZING THE RANDOM NOISE.....	22
3.9	DEVELOPING THE SYNTHETIC HYDROGRAPH MODEL .....	28
3.10	USE OF THE MODEL FOR PREDICTION AND DESIGN .....	29
3.11	MODEL SOFTWARE.....	29
3.12	EXAMPLE IMPLEMENTATION.....	30
CHAPTER 4	QUANTITATIVE RESULTS AND DEMONSTRATION OF MODEL ..	42
4.1	OVERVIEW .....	42
4.2	MODEL PARAMETERS.....	42
4.3	DISTRIBUTION OF PARAMETERS.....	42
4.4	PARAMETER INVESTIGATION .....	53
4.5	SUMMARY OF PARAMETERS .....	57
4.6	SYNTHETIC CONTINUOUS DAILY HYDROGRAPH.....	58
CHAPTER 5	CONCLUSION AND RECOMMENDATION.....	63
5.1	OVERVIEW .....	63
5.2	ACCOMPLISHMENT OF RESEARCH OBJECTIVES.....	63

5.2.1	COLLECT STREAM DATA.....	63
5.2.2	CREATE A MODEL .....	63
5.2.3	ANALYSE THE RANDOM COMPONENT OF THE MODEL .....	64
5.2.4	DEMONSTRATE THE MODEL'S CAPABILITY .....	64
5.2.5	SOFTWARE DEVELOPMENT .....	65
5.3	DISCUSSION .....	65
5.3.1	INPUT TO THE MODEL .....	65
5.3.2	OUTPUT FROM THE MODEL .....	66
5.4	FUTURE RESEARCH .....	67
	APPENDIX A .....	69
	APPENDIX B .....	78
	APPENDIX C .....	94
	REFERENCES .....	100

# CHAPTER 1 INTRODUCTION

## 1.1 RESEARCH CONTEXT

This thesis is a subproject of a larger study conducted by the Department of Civil & Environmental Engineering (Profs. Brubaker and Goodings) for the Maryland State Highway Administration, *Estimation Of Long-Term Scour At Maryland Bridges Using EFA/SRICOS* (Brubaker et al., 2004). The goal of that study was to determine whether a new set of hardware and software tools developed at Texas A&M University are appropriate for use in Maryland. The Erosion Function Apparatus (EFA) is a hardware tool that estimates the erodibility of soils (Briaud et al. 2001). Scour Rate in Cohesive Soils (SRICOS) is a software tool that uses EFA output and stream discharge to predict long-term scour (Briaud et al., 1999).

Scour is the erosion of the soil that supports a bridge pier. Structural failures due to scour can lead to great expense and inconvenience, as well as loss of life. Therefore, methods to predict scour have received increased interest from bridge design agencies such as the Maryland State Highway Administration. Widely used scour prediction methods were developed for non-cohesive, or sandy, soils. The EFA/SRICOS method of Briaud et al. (1999, 2001) is designed to predict scour in cohesive soils (clay).

The SRICOS model requires a long time series of daily discharge, perhaps as long as 100 years. Most bridge crossing sites in Maryland does not have long-term stream discharge records. In order to apply the SRICOS method to these sites, it was necessary to develop a method that would extend gage records realistically, and allow the synthesis of realistic daily discharge hydrographs for ungaged streams.

## 1.2 RESEARCH OBJECTIVE

The goal of the study was to develop a method for generating realistic long-term daily discharge hydrographs for locations where the gage record is inadequate or unavailable. The method is based on deterministic and random properties of the discharge alone, and does not require precipitation data as input. This thesis describes the analysis phase of the project and demonstrates its use in synthesis. A separate phase of the study, not included in this thesis, analyzes the relationship of model parameters to physical characteristics of the watersheds, based on information that can be obtained from GIS HYDRO 2000 (Moglen and Casey, 2000). Ultimately, equations will be developed that allow the parameters of the model, thus a realistic daily discharge hydrograph, to be estimated for any watershed in Maryland.

The research objectives of this phase were:

- 1) Obtain daily discharge records for a large number of Maryland streams and rivers with long records.
- 2) Model the discharge data for each site using a model that depicts the daily discharge hydrograph's seasonal cycle, interannual variability, and persistence. This objective consists of identifying parameters that control the deterministic part of the model for each gage under study.
- 3) Determine an appropriate statistical distribution for the random part of the model, and estimate its parameters for the watersheds under study.
- 4) Show that synthetic streamflow generated with the model possesses the same statistical characteristics (seasonal cycle, interannual variability, persistence, mean, standard deviation and skew) as the measured data.



### 1.3 DESCRIPTION OF THE STUDY AREA

The study area is located in the state of Maryland, U.S.A. There are 16 major watersheds in Maryland. The study analyzes gage data from 78 rivers, representing all sixteen major watersheds. Fig. 1.1 below gives an overview of the Maryland watersheds.

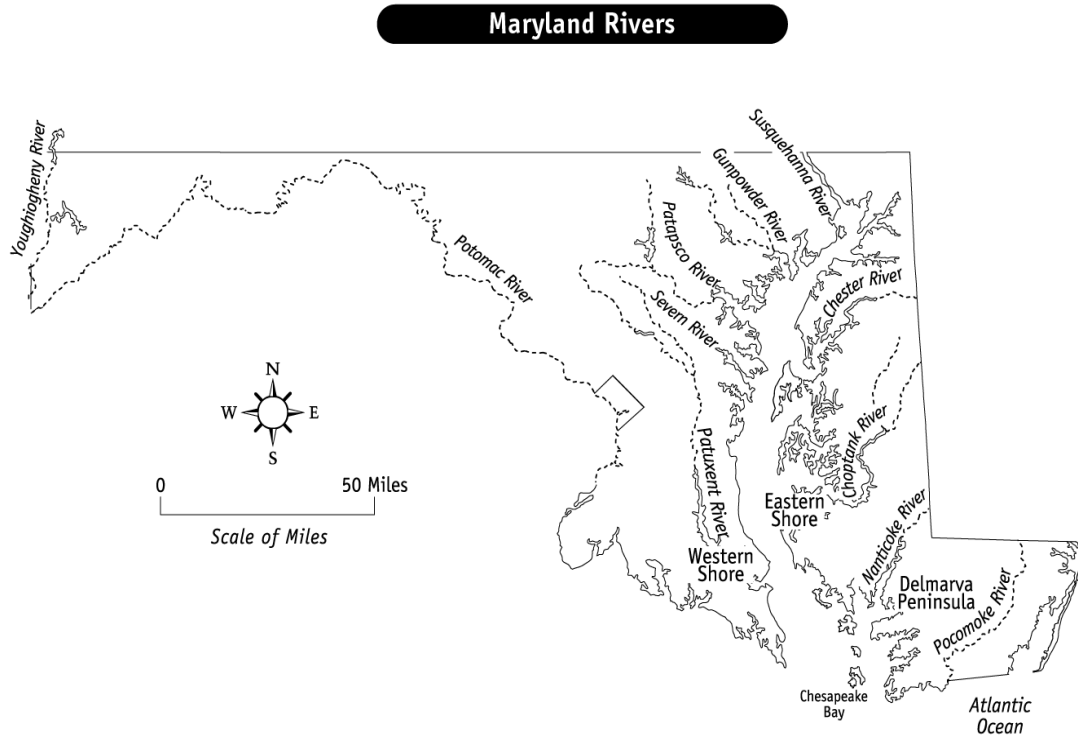


Figure 1.1. Maryland Rivers (Chapelle 2003, used with permission).

### 1.4 RESEARCH IMPLICATIONS

The synthetic hydrograph method would be useful not only for the EFA/SRICOS study, but for any investigation requiring realistic, long-term daily discharge hydrographs. The ability to synthesize many possible realizations of streamflow will be useful for probabilistic prediction methods.

## **1.5 OVERVIEW**

The thesis work is restricted to the generation of the synthetic continuous daily hydrograph for rivers with gaged data. This work provides the foundation for extending to the generation of the synthetic hydrographs for ungaged rivers.

The following chapters provide a detailed explanation of the steps taken in the project. Chapter 2 is an overview of issues in synthetic hydrograph generation, and relevant time series concepts. Chapter 3 explains the theory and procedures. The steps are demonstrated with examples: the rivers with highest and lowest discharge. In Chapter 4 the results for all analyzed rivers are presented and discussed. Chapter 5 summarizes the findings and conclusions and discusses future steps and applications.

## **CHAPTER 2 HYDROGRAPHS: CONCEPTS AND EXISTING APPROACHES**

### **2.1 NEED FOR SYNTHETIC CONTINUOUS HYDROGRAPH**

One stumbling block in water resources engineering is the inadequacy of stream flow records. Records available from many streams are of insufficient length and for some streams no streamflow data have been recorded (Fuller, 1978). Wurbs and James (2002) observe that the majority of hydrological engineering applications deal with ungaged watersheds. The solution for such applications is the synthetic hydrograph. The flows generated in the synthetic technique are representative of the likely flow; they are neither the prediction of future flows nor the reconstruction of historical flows in a stream. A unit hydrograph is one way of estimating a synthetic hydrograph. A unit hydrograph can be measured directly and indirectly. When the parameters such as watershed area, observed outflow hydrograph at the watershed outlet, base flow parameters and precipitation data for the input are available, then a unit hydrograph can be computed directly. But when there is an absence of any of these data, then the approach becomes indirect and it is called an indirect method or synthetic method.

### **2.2 DEFINITIONS**

#### **2.2.1 HYDROGRAPH AND HYETOGRAPH**

A hydrograph is a graph showing the variation in the discharge or flow of a river with respect to time. Hydrograph analysis deals with the runoff records at the stream gages. It is a record and graphical representation of discharge as a function of time at a specific location, for example, discharges at a point in a stream or discharges from a pumping well.

A hydrograph provides a way of seeing hourly, daily, seasonal or yearly changes in the flow or discharge of a waterway. Hydrographs can also be made for lakes, water wells, springs and other bodies of water. The stream discharge hydrograph is the area of concern in this project. Graphical representation of rainfall intensity versus time is called a hyetograph.

### 2.2.2 EVENT HYDROGRAPH

An event hydrograph describes the response to a single storm. The duration of the storm may be from few hours to a few days. A typical surface runoff hydrograph has a fast rising limb, and a slow receding limb that connects the base flow (Fig. 2.1). The total runoff hydrograph has two parts--runoff and baseflow. Baseflow is the runoff that occurs even when there is no rainfall. The direct runoff hydrograph is obtained after the deduction of baseflow from the total runoff hydrograph.

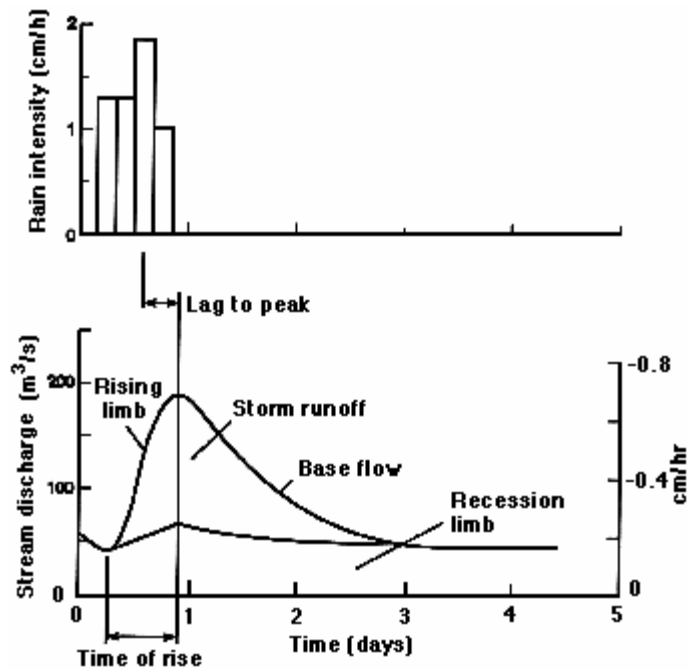


Figure 2.1 Event hydrograph

### 2.2.3 CONTINUOUS HYDROGRAPH

A continuous model simulates a longer period, predicting watershed response both during the storms and between the storms. Fig. 2.2 shows an example of a continuous daily hyetograph and hydrograph.

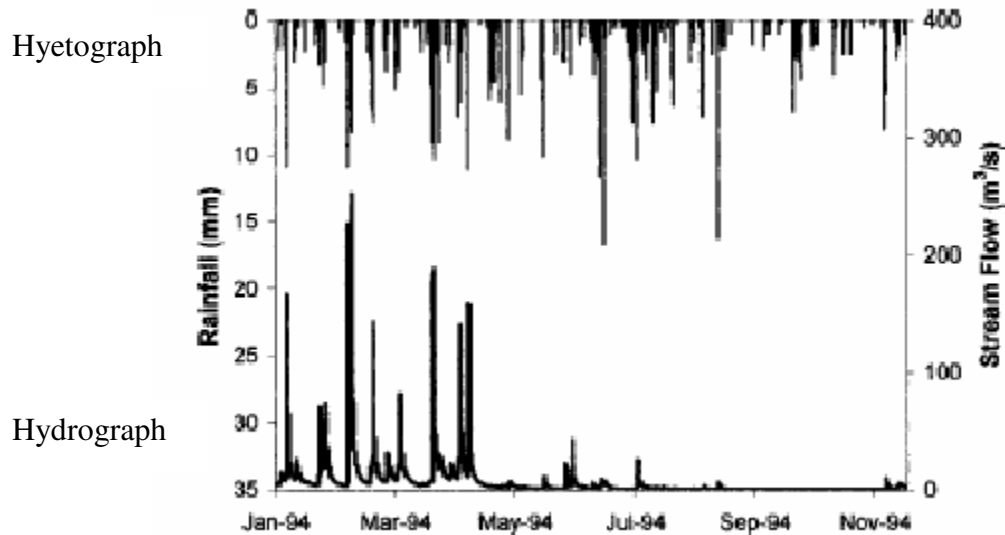


Figure 2.2 Continuous daily hydrograph ( Fleming et.al., 2004)

### 2.2.4 UNIT HYDROGRAPH

A unit hydrograph is the result from a unit excess rainfall occurring uniformly over the entire watershed with a uniform rate for a given period of time. The purpose of the unit hydrograph is to develop hydrographs for selected storms (Wurbs and James, 2002). The area under the unit hydrograph should be equivalent to 1 area-in. Rainfall data are convoluted using a unit hydrograph to calculate discharge at outlet. There are different types of unit hydrographs. Some most common types are T-hour unit hydrograph, instantaneous unit hydrograph, and dimensionless unit hydrograph. A T-hour unit hydrograph is based on the T-hour storm, which results in 1 area-in of runoff. The instantaneous unit hydrograph is a type of T-hour unit hydrograph with rainfall excess,

which is delivered over an infinitesimally small period. The third most common is a dimensionless unit hydrograph. The ordinates of unit hydrographs are usually given in terms of ratio of time to time to peak and, similarly, ratio of discharge to peak discharge. An instantaneous unit hydrograph needs to be converted to a T-hour unit hydrograph before use (McCuen, 1998).

### **2.3 UNIT HYDROGRAPH METHOD**

In the unit hydrograph approach, the rainfall hyetograph data are the input to the hydrologic model to develop a runoff hydrograph. In order to do this procedure, a transfer function is required. The unit hydrograph is the transfer function transforming the rainfall excess into the direct runoff hydrograph. According to McCuen (1998), there are four aspects for developing a unit hydrograph from observed data: the rainfall excess should be 1 inch; rainfall should have uniform spatial distribution over the watershed; the rainfall excess rate should be constant with time; and a duration of rainfall excess must be specified. During computation these aspects should be taken into consideration. The hydrograph volume should be equivalent to 1 area-in of direct runoff. Spatially non-uniform precipitation will not produce the proper temporal characteristics, so it is necessary for spatial distribution to be uniform. It is difficult to find storms of constant excess rate, so there could be some relaxation in this assumption (McCuen, 1998).

#### **2.3.1 CONVOLUTION**

A small unit hydrograph duration (D) is selected to reflect the variation of rainfall and rainfall excess and flows over time. The process of combining the rainfall excess of each D-hour increment of rainfall with D-hour unit hydrograph is called convolution (Wurbs and James, 2002). The convolution process scales the unit hydrograph in

proportion to runoff volume and lags incremental hydrograph to reflect the time sequencing of each D-hour increment of rainfall. In result of this process, the hydrograph developed reflects the runoff from the entire rainfall event.

### **2.3.2 UNIT HYDROGRAPH FROM GAGED FLOW DATA**

When a stream gage station is available, a rainfall event is selected which is considered to be representative of the storms that occur. Hydrographs from the rainfall events of same duration are averaged and one composite hydrograph is developed which represents the watershed. The first step in developing a unit hydrograph from gaged flow data is to remove the base flow, the second step is to determine the runoff volume by integrating the hydrograph and the last step is to develop the unit hydrograph by scaling the discharges in proportion in runoff volume. The rainfall duration for the unit hydrograph is same as observed for the hydrograph (Wurbs and James, 2002)

### **2.3.3 UNIT HYDROGRAPH FROM UNGAGED FLOW DATA**

There are various types of methods to develop synthetic unit hydrographs. Figure 2.3 summarizes the various methods in the form of a flow chart. Dimensionless unit hydrograph methods are used in developing the synthetic unit hydrographs. There are some indirect methods to develop unit hydrographs (Prakash, 1997).

## **2.4 WATERSHED MODELLING METHODS**

Models that develop hydrographs need meteorological data as input. Spatial distribution is another factor in deciding the model for the hydrograph. Data can be lumped or distributed. In a distributed model, geographic variations of processes and characteristics are included whereas; in a lumped model these variations are averaged or ignored. Based on these inputs, models can be differentiated.

Continuous Simulation models provide a long-term simulation involving multiple storm events, soil moisture monitoring, etc (Gupta, 2001). Some widely used lumped and distributed continuous models are listed below: -

- 1) National Weather Service River Forecasting System (NWSRFS) provided by National Weather Service (NWS) is a lumped parameter model (NWS, 2005).
- 2) Stream flow Synthesis and Reservoir Regulation (SSARR) provided by U.S. Army Corps of Engineers (USACE) is a lumped model (USACE, 2002).
- 3) Simulation for Water Resources in Rural Basins (SWRRB) provided by U.S. Department of Agriculture (USDA) is a distributed model for large basins (Arnold et.al., 1990).
- 4) Distributed Routing Rainfall-Runoff Model (DR3M) is a distributed model for small basins by U.S Geological Survey (USGS) (USGS, 2005b).
- 5) Precipitation –Runoff Modeling System (PRMS) is a U.S. Geological Survey model was released is a distributed model (USGS, 2005a).
- 6) Hydrological Simulation Program-FORTRAN (HSPF) is developed by U.S. Environment Protection Agency (EPA) allows integrated simulation of runoff and soil contamination. It can be run in lumped or semi-distributed model (USEPA, 2003).
- 7) Storm Water Management Model (SWMM) is developed by U.S. Environment Protection Agency (EPA) is a single event or continuous model that can be run in lumped or semi-distributed mode (James and James, 2000).

Statistical hydrology of natural stream flows is covered by HEC models: the Multi-Linear Regression Program (MLRP) and HEC-4 Monthly Stream flow Simulation models. MLRP



model is based on regression analysis to determine relationships among rainfall-runoff parameters and basin characteristics. This Model automatically deletes the least significant

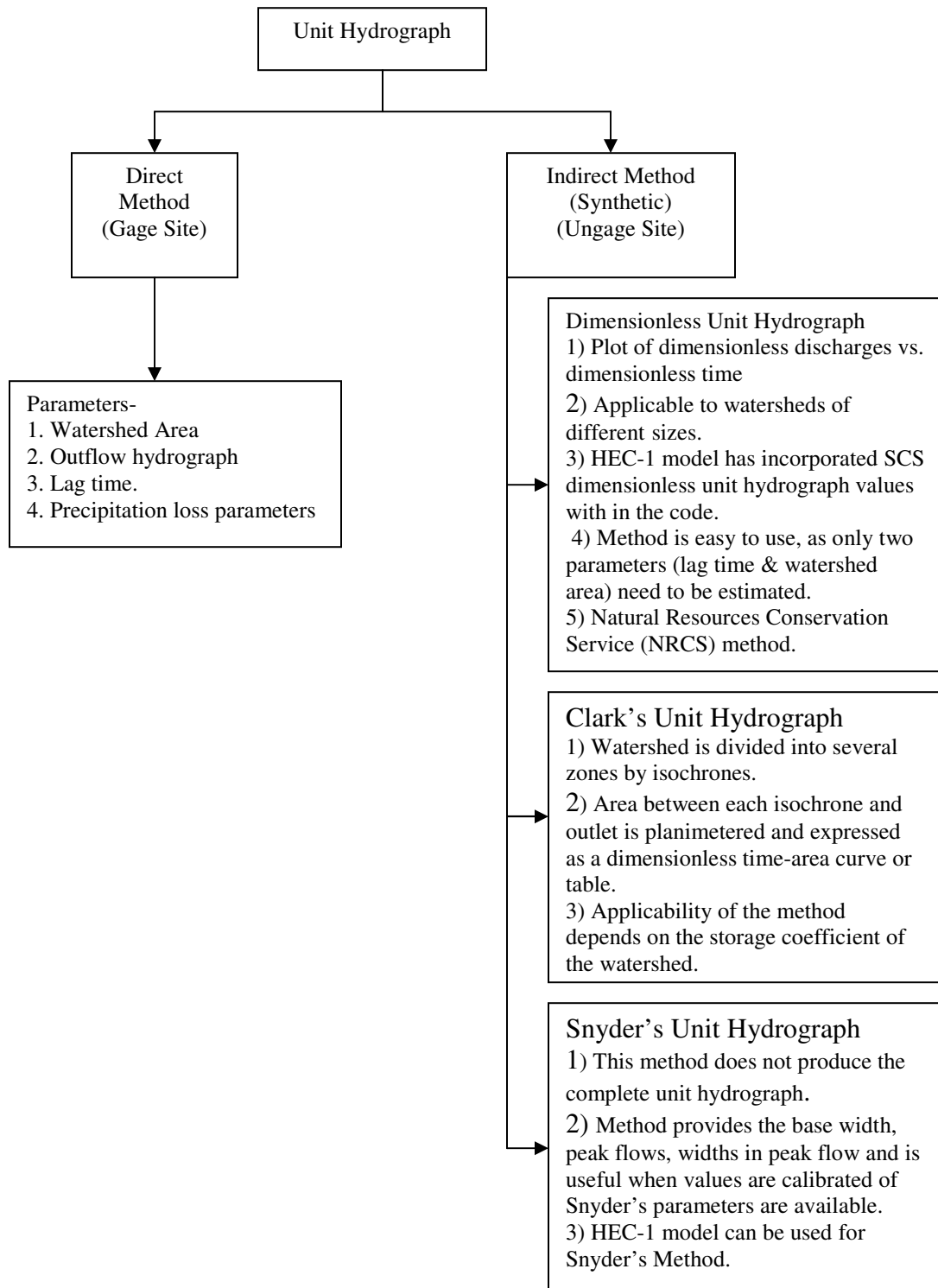


Figure 2.3 Unit Hydrograph Methods

variables after every iteration. The HEC-4 model on the other hand performs analysis of monthly stream flow at a number of interrelated stations to determine statistical parameters like- standard deviation, mean, and skew. Correlation of two stations fills the missing data from other stations. Using the historical data, the model generates a hypothetical sequence for a station with the same statistical characteristics.

All these models help to solve the problems for both gaged and ungaged sites. They all require meteorological data as input.

## **2.5 MARKOV MODEL**

The Markov process or Autoregressive model is a method to generate synthetic stream flow without rainfall or meteorological inputs. It is the basic method considered in this project. The Markov process considers that the value of a stream flow event at one time is correlated with the value of the event at an earlier period (Gupta, 2001). This is also called as the existence of the autocorrelation in the time series. A Markov process considers that there is an existence of correlation in the values of two successive events. The value of a variable in one time period is dependent on the value of the preceding time period plus a random component (Gupta, 2001). A Markovian process can be applied for years, seasons, months, days or other time increments. In this project a daily flow model with lag of one day is considered. This means that the magnitude of the current day's flow is significantly correlated with the previous day's flow. Multiple season models divide the yearly flow into season or months. A multilag model on the other hand needs a long memory. A multilag model considers that the influence on the current flow is correlated beyond the previous time period.

In Markov process, the data synthesis follows 4 steps-

Analysis: -

- 1) Determination of statistical parameters from the analysis of the historical record.

The four statistical parameters (mean, standard deviation, coefficient of skew, correlation coefficient) play an important role in the data synthesis of stream flows. Mean is the mean of the historical data. Standard deviation is the measure of the variability of the data. Coefficient of skew is the measure of the lack of the symmetry in the data. Correlation coefficient is a measure of the extent at which flow at particular day is affected by the previous day's flow.

- 2) Identifying the frequency distribution of data.

In stream flow, the data are often normal, lognormal or gamma distributed. The appropriate distribution may be identified using a probability plot (McCuen 1998).

Pearson proposed a distribution that fits many distributions including normal, beta and gamma. A form of Pearson function, which is similar to Gamma function, is known as Log Pearson-III distribution. It is a distribution in three parameters with a limited range in left direction, unbounded to the right with large skew. The characteristic of this distribution is the distinct positive skew.

Hamlin and Kottegoda (1971) found that assuming a normal distribution is generally a deviation from the reality as hydrologic data, especially river flow data, has positive skew. This is due to the fact that negative values are not possible and thus the left tail of the probability density curve is restricted when compared to the right tail, which allows for high flows. In their study they tested the data for all the distributions and found that gamma or Log Pearson-III gave the best fit for the Teme river data.

Synthesis: -

- 3) Generating random numbers of the same distribution and statistical characteristics.

Random numbers can be generated using the computer based pseudo random number generator or using the random number tables. The tables for uniformly distributed or normally distributed or gamma distributed are different.

Normal random numbers have zero mean and a standard deviation of one known as standard normal deviates. A transformed random variable with zero mean and variance of random variable of  $S_e^2$  can be given by,

$$n_i = S_e t_i \quad (2.1)$$

where,

$n_i$  = Transformed random variable

$S_e$  = Standard deviation of random numbers.

$t_i$  = Standard normal variate (with zero mean and unit variance)

- 4) Constituting the deterministic part considering the persistence and combining with the random part.

The Markovian process considers that there is an existence of the persistence in the flow pattern of streamflow. And it is very well reflected in the serial correlation. Thus in the p-order model the effects run through p terms, and the autoregressive model, AR (p), takes the form, (Gupta, 2001)

$$a_i = \phi_{p,1} a_{i-1} + \phi_{p,2} a_{i-2} + \phi_{p,3} a_{i-3} + \dots + \phi_{p,p} a_{i-p} + n_i \quad (2.2)$$

Where,

$a_i$  =  $i$ th variable of the stochastic sequence (dimensionless)

$n_i$  = Random number at  $i$ th time.

$\phi_{p,1}, \phi_{p,2}$  = Autoregressive parameters or weights.

The dimensionless stochastic component can be related to the flow at any time by

$$x_i = \bar{X} + Sa_i \quad (2.3)$$

where,

$\bar{X}$  = Mean of the historical sequence

$S$  = Standard deviation of the historical sequence.

The first order Autoregressive model will reduce to the form

$$a_i = \phi_{1,1}a_{i-1} + n_i \quad (2.4)$$

When the conditions of mean  $a_i=0$ , variance  $a_i=1$  and expectation  $E(n_i, a_{i-1})=0$  are included for the sequence  $a$ , the following relations are derived:

$$\phi_{1,1} = r_1 \quad (2.5)$$

Where  $r_1$  is the lag-one correlation of the stochastic component. The random and deterministic part of the Markov model can be represented as

$$x_i = \underbrace{\bar{X} + r_1(x_{i-1} - \bar{X})}_{\text{deterministic}} + \underbrace{S\sqrt{1-r_1^2}t_i}_{\text{random}} \quad (2.6)$$

Where,

$x_i$  = Streamflow at  $i$ th time.

$\bar{X}$  = Mean of recorded flows.

$r_1$  = Lag 1 serial or autocorrelation coefficient.

$S$  = Standard deviation of recorded flows.

$t_i$  = Random variate from an appropriate distribution with a mean of zero and variance of unity.

Rodriguez-Iturbe, (1969) considered the Markov process for the estimation of statistical parameters for annual river flows.

## 2.6 SALAS'S METHOD

Salas (1993) proposes a more complex Markov approach that accounts for seasonal variation in the statistics. In this project, the Salas approach is applied. The steps in the Salas approach are:

- 1) Identifying and removing seasonality in the mean and standard deviation.

Many rivers exhibit a periodic seasonal variation in their average value and standard deviation, when these statistics are considered across years. In other words,  $X$  and  $S$  in Eq.2.5 would not be assumed constant but would change in a deterministic periodic way depending on the day of the year. When many years of data are available, the interannual mean and standard deviation can be estimated, and the data standardized by subtracting the mean and dividing by the standard deviation

$$z_t = \frac{(y - \bar{y}_t)}{S_t} \quad (2.7)$$

$z_t$  is free of seasonality.

- 2) Removing Autocorrelation

The  $Z$  data created by Eq.2.7 have zero mean and unit variance, but are likely to exhibit autocorrelation, as measured by the lag-one correlation coefficient. It is possible that the lag-one correlation varies with time of the year. The  $Z$  data for day's  $t$  and  $t-1$  are related to according to

$$z_t = r_1 z_{t-1} + \varepsilon_t \quad (2.8)$$

Where,

$r_{1,t}$  = Lag One Correlation for day i

$\varepsilon_t$  = Residual series (random noise) free of autocorrelation.

The residual series is free of periodicities and is normally or non-normally distributed..

Uncorrelated normal random variable is also called noise, error term, and series of shocks

## **2.7 THIS STUDY'S CONTRIBUTION**

The author, advisor, and State Highway Administration colleagues were unaware of any previous applications of Markov/Salas models to synthesize continuous daily streamflow for Maryland streams. Previous studies (Rodriguez-Iturbe, 1969) had been done for the annual or monthly stream flows, or to estimate peak flows corresponding to selected exceedence probabilities. This project provides a new tool that will be useful for the EFA/SRICOS scour study and for other projects requiring long time series of realistic daily streamflow. The automated implementation in Matlab is also unique.

## CHAPTER 3 THEORY AND METHODS

### 3.1 OVERVIEW

In this chapter the methods and procedures are discussed. All procedures were performed using Matlab Scripts and MS Excel spreadsheets.

### 3.2 ASSEMBLING STUDY DATA

Daily streamflow ( $Q$ ) data were obtained from USGS (2005) for 78 stream gages in Maryland having records of data of more than 30 years. Data for each gage were arranged as a matrix with each year as a column and each day of the year (1 to 366, including Leap Year Day) as a row (Appendix B).

### 3.3 DATA TRANSFORMATION

Daily streamflow data generally follow skewed distributions; thus it is common to apply a logarithmic transformation. In this study, a log with base 10 transformation was applied to all data, resulting in 366-row matrices of  $\log Q$  data. Reported discharge values of zero would have introduced difficulties at this stage; however, no zero values were encountered in any of the discharge data analyzed. At this point, no assumption was made about the distribution of  $\log Q$ .

### 3.4 IDENTIFYING AND REMOVING SEASONALITY

Following Salas (1993), the mean and standard deviation of the transformed data were computed for each day of the year, across all years of record:

$$M[\log Q(d)] = \frac{\sum_{all\ y} \log Q(d, y)}{number\ of\ years} \quad (3.1)$$



$$StdDev[\log Q(d)] = \sqrt{E\{\log Q(d,y) - M[\log Q(d)]\}^2} \quad (3.2)$$

Where

$d$  and  $y$  are the indices for day and year, respectively, row and column of the data matrix.

$$\sum_{all\ y} \log Q(d, y) \quad = \text{Sum of log data for a given day, across all years}$$

$$M[\log Q(d)] \quad = \text{Mean of transformed data for a given day, across years}$$

$$StdDev[\log Q(d)] \quad = \text{Standard deviation of transformed data for a given day,}$$

across years

$E\{\}$  is the expectation, or average, operator

The seasonal effects are removed from the  $\log Q$  data by subtracting the interannual average and dividing by that day's standard deviation.

$$Z(y, d) = \frac{\log Q(y, d) - M[\log Q(d)]}{StdDev[\log Q(d)]} \quad (3.3)$$

$Z$  is the deseasonalized variate. Each row (day) of the  $Z$  matrix has zero mean and unit variance.

### 3.5 ANALYZING TEMPORAL CORRELATION

As a measure of persistence of flow, the tendency of daily discharge to be similar from day to day is quantified by a correlation term. The correlation of each day's  $Z$  value to the preceding day, across all years, is computed.

$$R_z(d) = \sqrt{E[Z(d)Z(d-1)]} = \frac{\sum_{all\ years} Z(d, y)Z(d-1, y)}{number\ of\ years} \quad (3.4)$$

The correlation coefficients computed in Eq. (3.3) are different from the typical autocorrelation coefficient of a time series. Eq. (3.4) represents the correlation between row  $d$  and row  $d-1$  of the data matrix, where rows correspond to days and columns to years. This approach allows for the possibility that persistence (or temporal correlation) could have a seasonal pattern.

### 3.6 CURVE FITTING SEASONAL STATISTICS

A cosine wave model is fit to the 366-day sequence for each of the three parameters, average annual log discharge  $M[\log Q (d)]$ , standard deviation of log Discharge  $StDev[\log Q (d)]$ , and one-day correlation of the deseasonalized variates  $R_z (d)$ ,

$$Y_{pred} = \bar{y} + y_a \cos \left[ \frac{2\pi}{366} (d - \tau) \right] \quad (3.5)$$

Where,

$Y_{pred}$  = The time varying parameter,  $M [\log Q (d)]$ ,  $StDev [\log Q (d)]$ , or  $R_z (d)$

$\bar{y}$  = Mean of the curve fit.

$y_a$  =Amplitude of the curve fit.

$\tau$  =Day of maximum.

$D$  = Day of the year.

A Matlab script (Appendix B) is used to perform the curve-fit. The mathematical routine may return a negative value of the phase shift,  $\tau$ . In such cases, the value is adjusted by 366 days to make it positive, corresponding to the day of the year on which that parameter takes its maximum value. Similarly, the amplitude of the curve fit can also take negative values. The sign is changed from negative to positive and day of the maximum is shifted by 186

days. (Future implementations will place constraints on the parameters to guarantee that  $t$  lies between 1 and 366, and that the amplitude is positive.)

### 3.7 DECORRELATING THE DATA

The deseasonalized  $Z$  data of each day are assumed to be made up of two terms: deterministic persistence from the previous day and random uncorrelated (“white”) noise,

$$Z(d, y) = \rho(d)Z(d-1, y) + \varepsilon(d, y) \quad (3.6)$$

where

$Z(d, y)$  = Deseasonalized variate for day  $d$  in year  $y$ .

$Z(d-1, y)$  = Deseasonalized for the previous day.

$\varepsilon(d)$  = Random noise with zero mean

$\rho(d)$  = lag-one autocorrelation

Models such as Eq. (3.6) are known as autoregressive (AR) or Markov models (Gupta, 2001) and are widely used in Hydrology, as discussed in Chapter 2. The notation here emphasizes that the lag-one autocorrelation may vary throughout the year.

The next step in the analysis is the removal of the deterministic contribution, i.e. decorrelation of the data, to isolate the random noise,

$$\varepsilon(d, y) = Z(d, y) - R_z(d)Z(d-1, y) \quad (3.7)$$

In Eq. (3.7), the sample lag-one correlation is used, leaving a sequence of uncorrelated random noise,  $\varepsilon(d)$ , with zero mean.

The noise term can be modeled as proportional to a random variable,  $K$ , with zero mean and standard deviation of one,

$$\varepsilon(d, y) = \sigma_\varepsilon(d)K(d, y) \quad (3.8)$$

The sample  $K$  values are determined by scaling the noise by its standard deviation

$$K(d,y) = \frac{\epsilon(d,y)}{s_\epsilon(d)} \quad (3.9)$$

By Eq. (3.9), the standard deviation of the noise is related to the lag-one correlation through the fact that both  $Z(d)$  and  $Z(d-1)$  have variance equal to one,

$$s_\epsilon^2(d) = 1 - R_z^2(d) \quad (3.10)$$

Therefore, the decorrelated, zero-mean, unit standard deviation noise variable  $K(d,y)$  can be estimated from the time series of  $Z$  as follows:

$$K(d,y) = \frac{Z(d,y) - R_z(d)Z(d-1,y)}{\sqrt{1 - R_z^2(d)}} \quad (3.11)$$

The sequence of  $K(d,y)$  will have zero mean and standard deviation of one across all years.

The random number  $K(d,y)$  is drawn from a distribution with zero mean and unit variance;

the specific distribution remains to be analyzed.

### 3.8 ANALYZING THE RANDOM NOISE

The final step of the analysis is to determine an appropriate statistical distribution for the random noise  $K$  in Eq. (3.11). Auto-regressive models such as Eq. (3.6) often use Gaussian noise; however, a normal distribution is not necessarily appropriate for this application. Calculation of the sample skew and exploratory plots of the  $K$  data for arbitrary days indicated positively skewed distributions.

Using MS Excel, a detailed analysis was conducted on the  $K$  data for two rivers, the USGS 01585100 White Marsh Run, White Marsh, MD and the USGS 01603000 North Branch Potomac, Cumberland, MD, for dates in December, February, June, and September. Probability plots were constructed for both the normal distribution and the Pearson III distribution.

To create a normal probability plot in Excel, the NORMINV function was used to calculate and plot the standard normal variate for selected probabilities. These values were used to construct the grid for plotting the data. The  $K$  data were ranked and their percentiles or plotting positions calculated using the Weibull formula,

$$P_i = \frac{i}{n+1} \quad (3.12)$$

where

$P_i$  =Exceedence probability for an event with rank  $I$

$i$  =Rank of the event

$n$  =Sample size.

The data were then plotted on the constructed probability graph. If the data are drawn from a normal distribution, they should lie along a straight line on this graph.

Similarly, the Pearson III probability plot was constructed using the GAMMAINV function in MS Excel, as a visual test of whether it is reasonable to assume that the  $K$  data are drawn from a Pearson III distribution with mean 0, standard deviation 1, and skew equal to the observed value. The Pearson III is a shifted gamma distribution with three parameters:  $\alpha$  (shape),  $\beta$  (scale), and  $shift$ ,

$$f(x; \alpha, \beta, shift) = \frac{1}{B^\alpha \Gamma(\alpha)} (x - shift)^{(\alpha-1)} e^{-\left(\frac{x - shift}{\beta}\right)} \quad (3.13)$$

The parameters of the distribution are estimated using the method of moments. For the Pearson III distribution, the parameters are related as follows (Abramowitz and Stegun, 1970),

$$\alpha = \left(\frac{2}{g}\right)^2 \quad (3.14)$$

$$\mu = shift + \alpha\beta \quad (3.15)$$

$$\sigma = \sqrt{\alpha\beta^2} \quad (3.16)$$

where  $g$  is the skew. The mean and standard deviation are set to 0 and 1, and the sample skew is calculated from the  $K$  data using the formula,

$$g = \frac{n \sum (x - \bar{x})^3}{(n-1)(n-2)s^3} \quad (3.17)$$

where,

$g$  = Skew

$n$  = Number of years

$\sum (x - \bar{x})^3$  = Third central moment of the data

$s$  = Sample standard deviation

The parameter  $\alpha$  is calculated using Eq. (3.14), and then the remaining parameters can be calculated by solving Eq. (3.15) and (3.16) with  $\sigma=1$  and  $\mu=0$ :

$$\beta = \sqrt{\frac{1}{\alpha}} \quad (3.18)$$

$$shift = -\alpha\beta \quad (3.19)$$

Using MS Excel's GAMMAINV function, selected percentiles are computed and used to create a plotting grid. The ranked sample data are plotted on the graph according to their plotting positions. If the data are drawn from a Pearson III distribution with mean 0, standard deviation 1, and skew  $g$ , they should lie on a straight line on this plot.

Examples of this procedure are shown in Fig. 3.1 and 3.2 for the North Branch of the Potomac River. Data were analyzed for the date December 1 for the years 1929-2001; there were 73 years of data available; each point on the graph represents (1 Dec) for a

given year. By visual inspection, the linear fit in Fig. 3.2 is very similar to Fig. 3.1. to quantify this comparison, the Probability Plot Correlation Coefficient Test, also known as the Vogel test (Stedinger et.al., 1993) is performed. The Vogel test is essentially the linear correlation between the data values and the population values corresponding to their estimated percentiles,

$$r = \frac{\sum (x_i - \bar{x})(w_i - \bar{w})}{\left[ \sum (x_i - \bar{x})^2 \sum (w_i - \bar{w})^2 \right]^{0.5}} \quad (3.20)$$

The values of the Vogel test are included in the graphs in Fig. 3.1 and 3.2, 0.9942 for the normal plot, and 0.9948 for the Pearson III. In this particular example, the values are nearly equal; however, this is not always the case, as discussed below.

Each of 366 days for all 78-gage records needed to be analyzed. Visual inspection of 54168 (2x366x74) probability plots would be tedious. The Vogel test provides a single value summarizing what is learned by visual inspection. Therefore, the Vogel test was performed on the  $K$  values for each data set, to test whether the Pearson III distribution is appropriate for describing this variable in all cases.

When the skew is small, the Vogel test statistics for the normal and Pearson III distribution are very similar, and the superiority of the Pearson III distribution is not compelling.

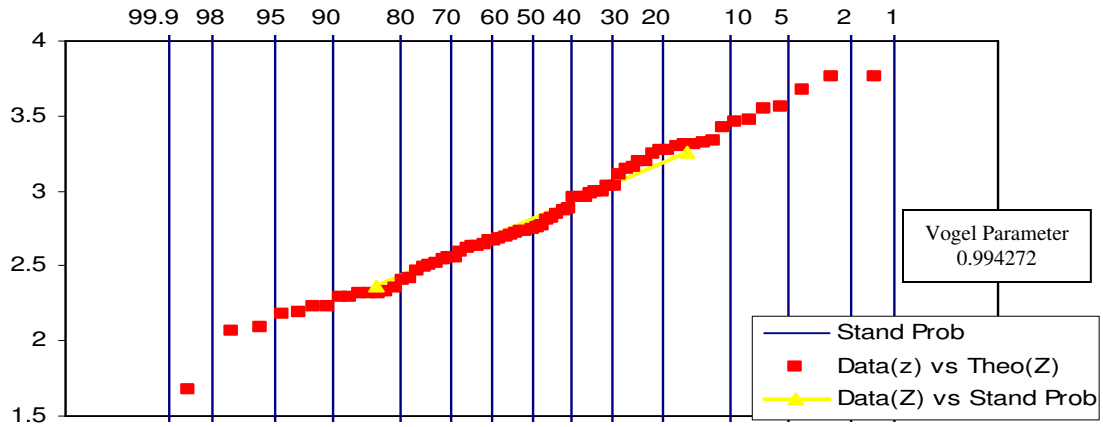


Figure 3.1 Normal Probability Plot for deseasonalized, decorrelated data (noise) for the North Branch Potomac River for 1 Dec. (73 years). Points show data for the watershed and the line is for the standard normal probability.

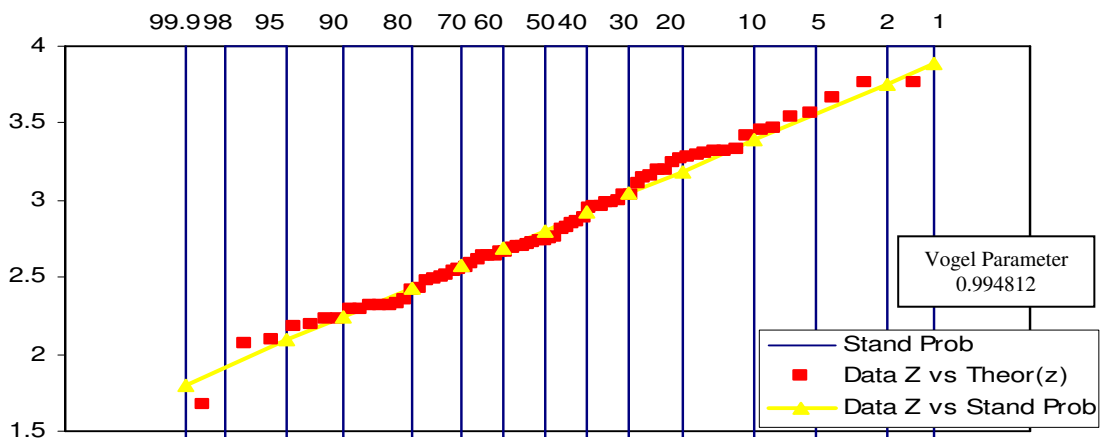


Figure 3.2 Pearson III Probability Plot for deseasonalized, decorrelated data (noise) for the North Branch Potomac River for 1 Dec. (73 years). Data for the watershed are shown by points and the line is for the Pearson III probability.

Examples, including the North Branch of Potomac case plotted in Fig. 3.1 and 3.2, are summarized in Table 3.1. It would be possible to use a normal distribution to model the noise when  $g$  is small, and a Pearson III otherwise. However, for simplicity in the algorithm, a Pearson III was applied in all cases. When the skew,  $g$ , approaches 0, the Pearson III and normal probability density functions become identical (Fig. 3.3).



In analyzing the decorrelated data, two computational issues arose: First, some of the  $K$  data sets showed negative sample skew,  $g < 0$ . Other hydrologic variables, such as annual peaks, can have negatively-skewed Log Pearson III distributions (McCuen 1998); however, by physical reasoning, the noise term in the daily flow model should be unbounded at high values, rather than low values, dictating a positive skew. Second, the computational tools, Matlab and Excel, were unable to invert the gamma distribution for very small values of  $g$  ( $g < 0.16$ ). For these two reasons, whenever a sample skew was less than 0.16, the value was set to 0.16 for the analysis. The implicit assumption is that very small or negative sample skews represent sampling error, because the population from which the  $K$  values are drawn must have a positive skew. A Pearson III pdf with mean 0, standard deviation 1, and skew 0.16 is very similar to the standard normal pdf (Fig. 3.3).

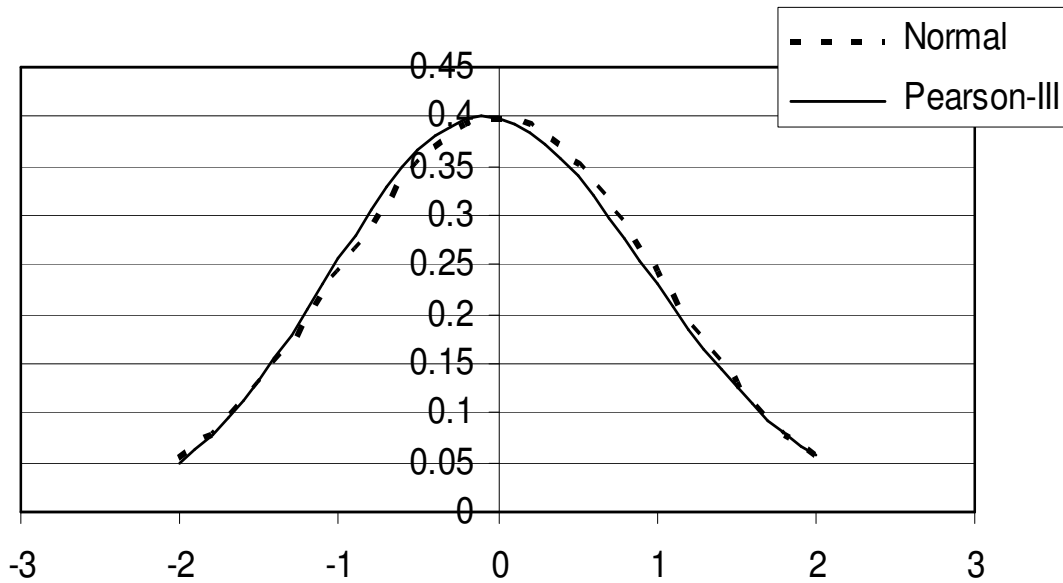


Figure 3.3 Plot to compare the Normal and Pearson-III distribution for skew=0.16.

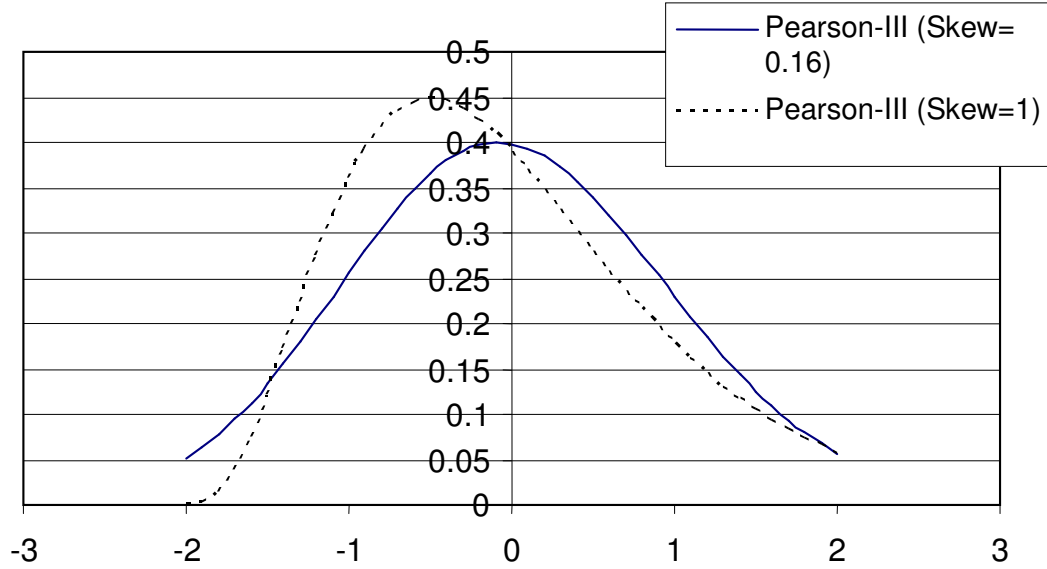


Figure 3.4 Plot to compare the Pearson-III distribution for skew of 0.16 and 1

Using Matlab, the sample skew is computed for all 366 days, for each of the 74 study watersheds. In addition, the Probability Plot Correlation Coefficient (Vogel test) is computed based on both the normal and the Pearson III distribution for each day. The results are stored as Excel spreadsheets. For each watershed, the 366 daily skew values are averaged to obtain a single value of skew to be used in the Markov model.

### 3.9 DEVELOPING THE SYNTHETIC HYDROGRAPH MODEL

The analysis described in sections 3.3 – 3.7 produces nine curve-fit parameters for the interannual average  $\log Q$ , interannual standard deviation of  $\log Q$ , and one-day correlation of the deseasonalized data. For each watershed, the 366 values of interannual skew (of the decorrelated data) are averaged. These ten values are the parameters of the Markov model. The complete model is given by the following equations:

$$\log Q(d, y) = M[\log Q(d)] + Z(d, y) * StDev[\log Q(d)] \quad (3.21)$$

$$M[\log Q(d)] = \bar{y}_M + y_{a,M} \cos\left[\frac{2\pi}{366}(d - \tau_M)\right] \quad (3.22)$$

$$StDev[\log Q(d)] = \bar{y}_s + y_{a,s} \cos \left[ \frac{2\pi}{366} (d - \tau_s) \right] \quad (3.23)$$

$$Z(d,y) = \rho_z(d)Z(d-1,y) + \sqrt{1 - \rho_z^2(d)}K(d,y) \quad (3.24)$$

$$\rho_z(d) = \bar{y}_R + y_{a,R} \cos \left[ \frac{2\pi}{366} (d - \tau_R) \right] \quad (3.25)$$

where  $K(d,y)$  is a Pearson III random variate with mean 0, standard deviation 1, and skew  $g$ .

### 3.10 USE OF THE MODEL FOR PREDICTION AND DESIGN

Given the ten parameters for a particular watershed, the model may be used to generate a realistic synthetic daily discharge hydrograph for any number of years. A random seed,  $Z(0)$ , is generated as a starting point;  $Z(0)$  is used in generating  $Z(1,1)$ , the first day's deseasonalized discharge value. For year  $y$ , day  $d$ , Eq. (3.25) is used to compute  $\rho_z(d)$ , and a random variate  $K(d,y)$  is drawn from the Pearson III distribution with 0 mean, unit variance and the prescribed skew, using the GAMMAINV function. The deseasonalized discharge,  $Z(d,y)$ , is computed using Eq.(3.24);  $M[\log Q(d)]$  and  $StDev[\log Q(d)]$  are evaluated using Eqs. (3.22) and (3.23), and  $\log Q(d,y)$  is computed with Eq. (3.21). Finally, the day's discharge is obtained by taking 10 to the power  $\log Q(d,y)$ .

### 3.11 MODEL SOFTWARE

An automated routine downloads data from the USGS site and organizes it into folders and files. Data from the USGS site is in text format and then it is stored in MS Excel in the matrix of 366 rows and number of years as columns (Appendix C). The Matlab script reads the text file and arranges it into a matrix of 366 rows and number of years as

columns, and saves results as Excel readable spreadsheets. Missing data are stored as NA() in Excel, and NaN (Not a Number) in Matlab.

The Matlab script performs all the steps discussed in this Chapter: log transformation; calculate the mean, standard deviation, and one-day correlation; deseasonalize the data; curve fit the seasonally varying parameters; decorrelate the deseasonalized data, calculate the skew, and the Probability Plot Correlation Coefficients.

It was first necessary to ensure that the program is taking leap years into consideration properly. Leap years occurs every four years; therefore the row corresponding to Feb. 29 (Day 60) has only  $\frac{1}{4}$  the data count of the other rows, reducing the number of values used for computing  $M$  [ $\log Q$ ] and  $StdDev$  [ $\log Q$ ] for Day 60, and  $R_z$  for both Days 60 and 61. The output report (Appendix B) shows that the correct numbers of data are being taken into account in the calculations.

### **3.12 EXAMPLE IMPLEMENTATION**

Each step described in sections 3.3 through 3.8 is now shown in detail for one particular site, the North Branch of the Potomac River at Cumberland (USGS Site 01603000).

The script computes the average and standard deviation of the log-transformed data across rows (by day across all years), producing vectors of  $M$  [ $\log Q$ ] and  $StdDev$  [ $\log Q$ ], as shown in Fig. 3.5 and 3.6. For the North Branch Potomac, the mean flow shows a seasonal cycle, with a maximum in March (Day 82) due to maximum precipitation and snow melt, and a minimum in August (Day 255) due to summer evapotranspiration. For better understanding of concept of interannual mean and standard deviation (“across rows”), histogram plots of two rows are plotted: the day of maximum mean and the day of

minimum mean (Fig. 3.7, Fig. 3.8). The interannual standard deviation of streamflow takes a minimum in spring and a maximum in fall.

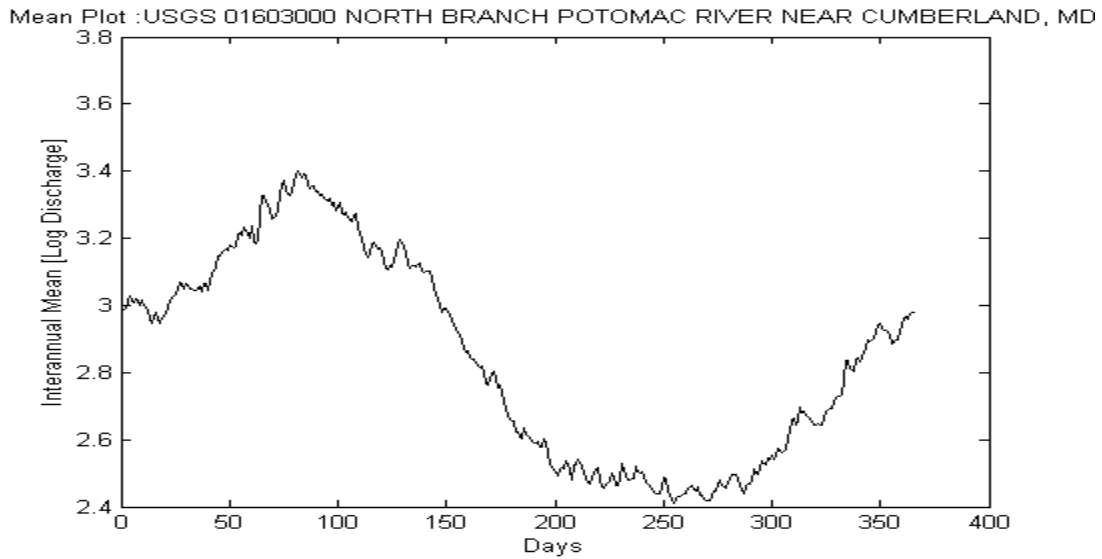


Figure 3.5 Interannual Mean of log-transformed daily stream flow, North Branch Potomac.

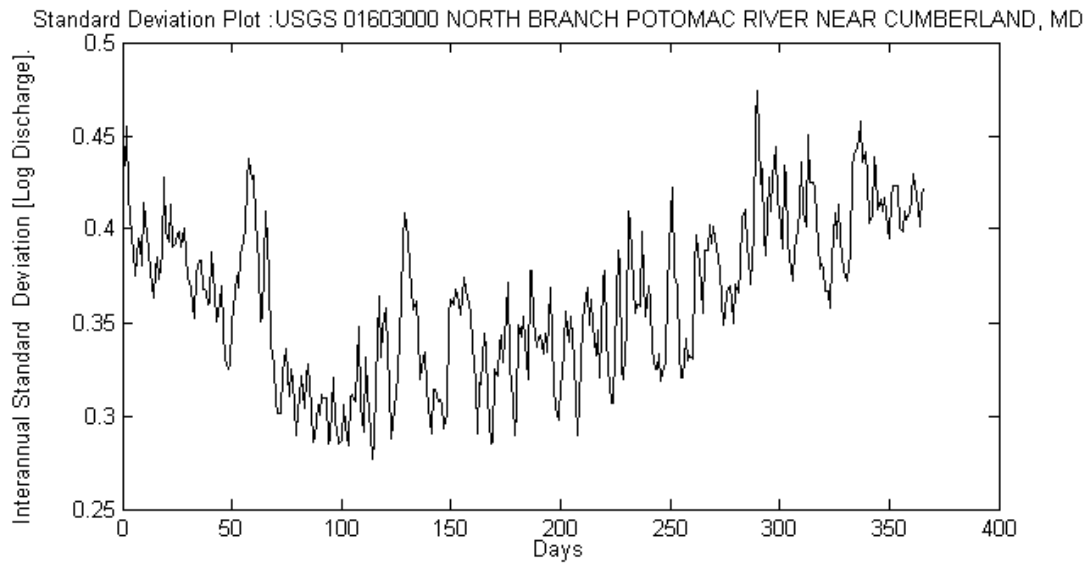


Figure 3.6 Interannual Standard Deviation of log-transformed daily stream flow, North Branch Potomac.

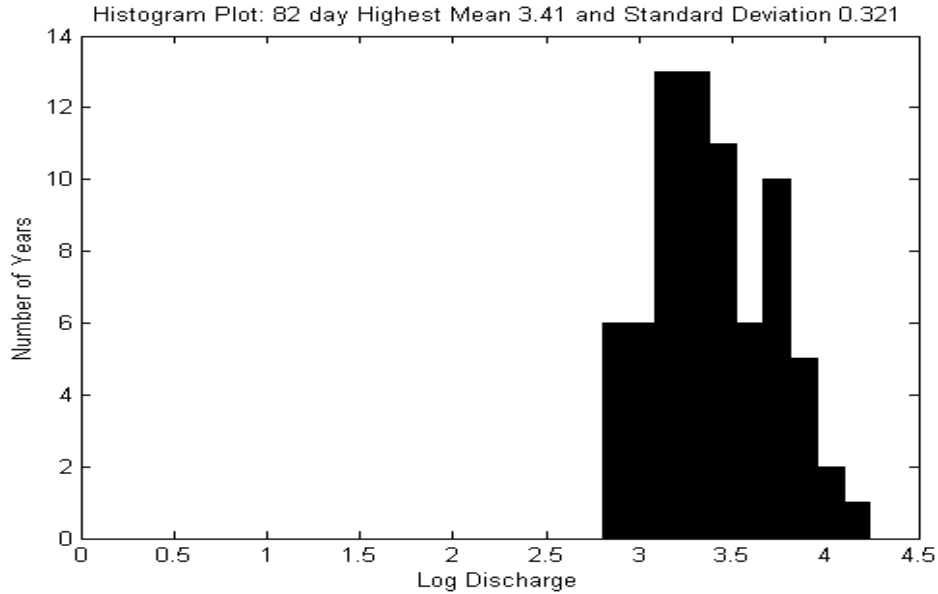


Figure 3.7 Histogram plot of data with maximum Mean (82 Day)

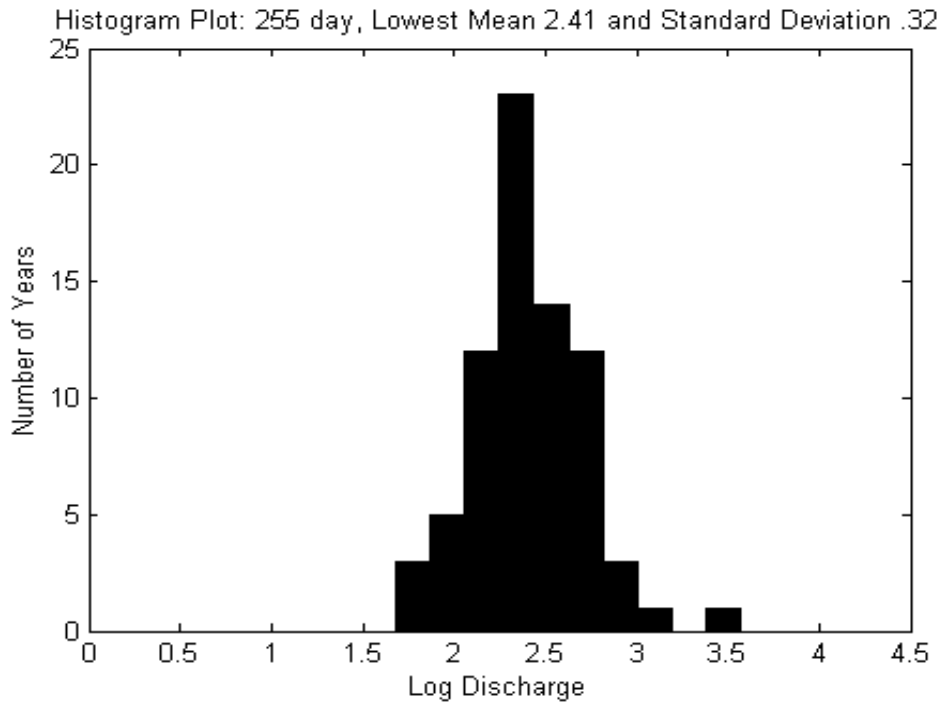


Figure 3.8 Histogram plot of data with lowest Mean (255 day)

The script computes the one-day correlation of the log-transformed streamflow, by taking the correlation between rows of the data matrix, producing a vector of 366 values of  $R_z(d)$ . For the North Branch Potomac (Fig. 3.9), the lag-one correlation is quite high, about

0.98, and consistent throughout the year. In Fig. 3.10 and 3.11, the days with maximum and minimum mean are plotted with the previous day to show the correlation between the previous day and today. Fig. 3.11 shows a good correlation with the previous day. The unusually low value (0.5) on day 59 is due to the smaller sample size (73/4) on Leap Year Day (Fig. 3.12).



Figure 3.9 Deseasonalized Data (Z) for North Branch Potomac

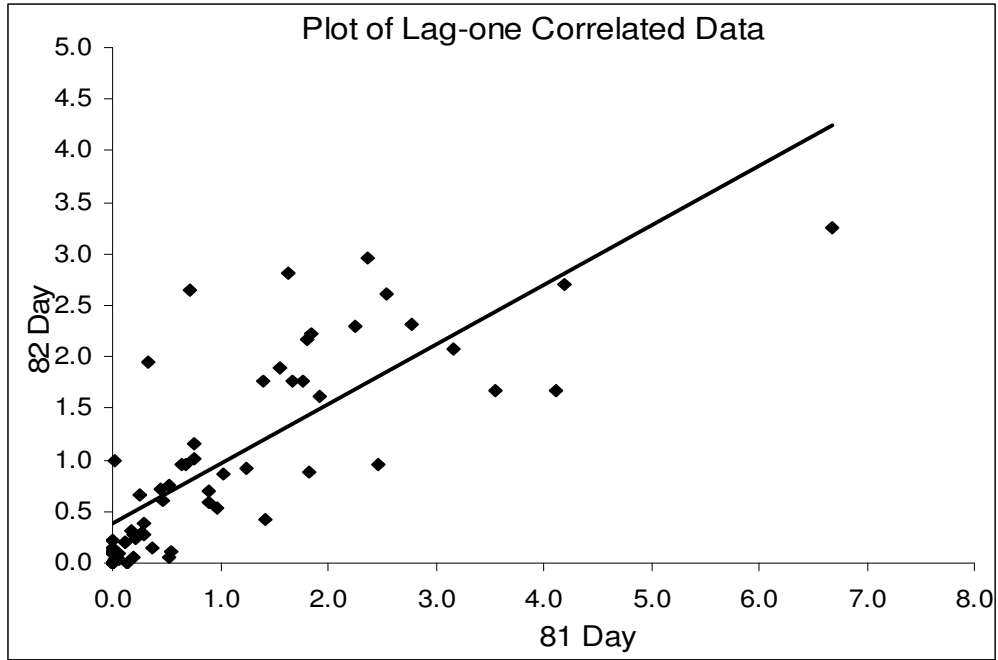


Figure 3.10 Scatter plot showing correlation between deseasonalized data on Days 81 and 82.

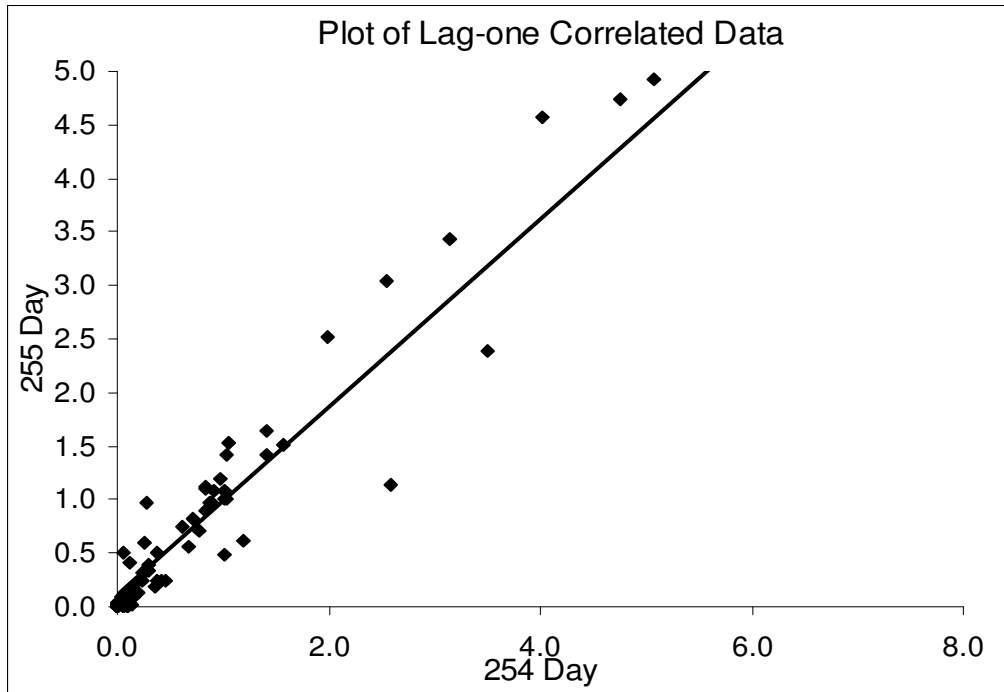


Figure 3.11 Scatter plot showing correlation between deseasonalized data on Days 254 and 255.

The script uses a nonlinear optimization to perform the cosine curve-fit of the interannual mean, standard deviation and correlation data. The resulting curve fits for North Branch



Potomac are shown in Fig. 3.12 and 3.13. Each curve fit model has three parameters: mean, amplitude, and day of maximum. The parameters for the North Branch Potomac are listed in Table 3-1.

Table 3. 1 Sample table of the skew analysis output, Days 1-23, North Branch Potomac.

Days	Skew	Vogel-Normal	Vogel-Pearson
1	3.20E+00	7.00E-01	7.77E-01
2	3.09E+00	7.05E-01	7.79E-01
3	2.63E+00	6.83E-01	7.37E-01
4	1.93E+00	6.96E-01	7.27E-01
5	2.82E+00	6.93E-01	7.56E-01
6	3.20E+00	6.93E-01	7.71E-01
7	3.17E+00	6.99E-01	7.76E-01
8	2.67E+00	6.99E-01	7.56E-01
9	2.17E+00	7.04E-01	7.43E-01
10	2.82E+00	7.01E-01	7.64E-01
11	1.82E+00	6.95E-01	7.23E-01
12	3.45E+00	6.92E-01	7.81E-01
13	1.60E-01	6.28E-01	6.28E-01
14	1.50E+00	7.05E-01	7.25E-01
15	1.75E+00	6.89E-01	7.15E-01
16	2.30E+00	6.99E-01	7.43E-01
17	6.36E-01	6.42E-01	6.45E-01
18	1.48E+00	7.05E-01	7.24E-01
19	5.67E+00	7.02E-01	9.07E-01
20	2.38E+00	6.92E-01	7.38E-01
21	1.73E+00	6.98E-01	7.24E-01
22	1.99E+00	7.04E-01	7.38E-01
23	2.12E+00	6.99E-01	7.37E-01

The deseasonalized data ( $Z$ ) (Eq. 3.3) for North Branch Potomac are plotted in Fig. 3.9. In this figure, 73 years of data are superimposed. Persistence in low values is notable in several years between Days 200 and 366. The decorrelated data ( $K$ , following Eq. 3.11) are graphed in Fig. 3.14. Again, the same 73 years of data as in Fig. 3.9 are superimposed. Persistent runs of low values are no longer visible in the graphs. For all days, the data show a

central tendency or clustering slightly below zero, and high values are more likely than low values,

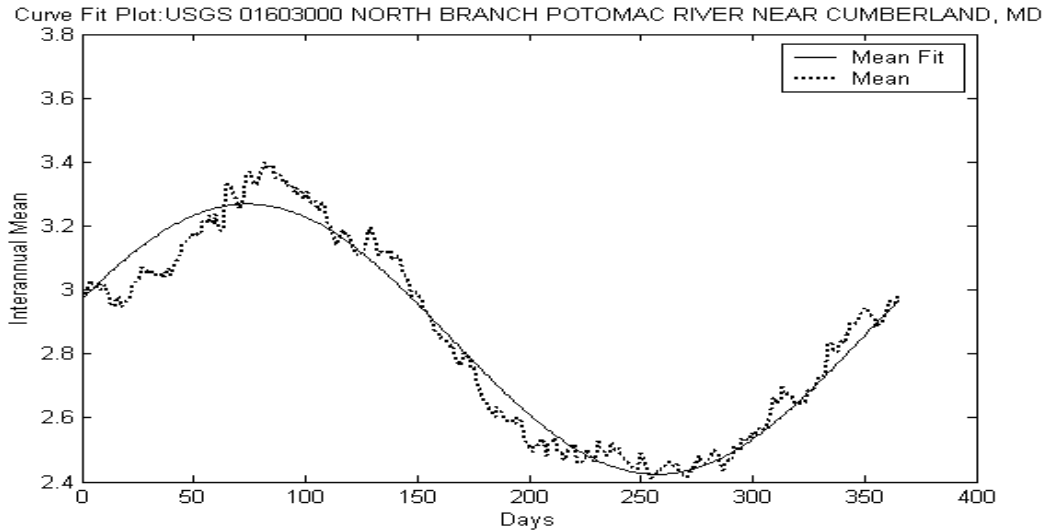


Figure 3.12 Curve Fit Plot of interannual mean, North Branch Potomac at Cumberland.

indicating positive skew. Figs. 3.14 and 3.15 are included to show the distribution of K values on Days 82 and 255, corresponding to the maximum and minimum interannual mean. These plots indicate that data are clustered around 0 for both the days.

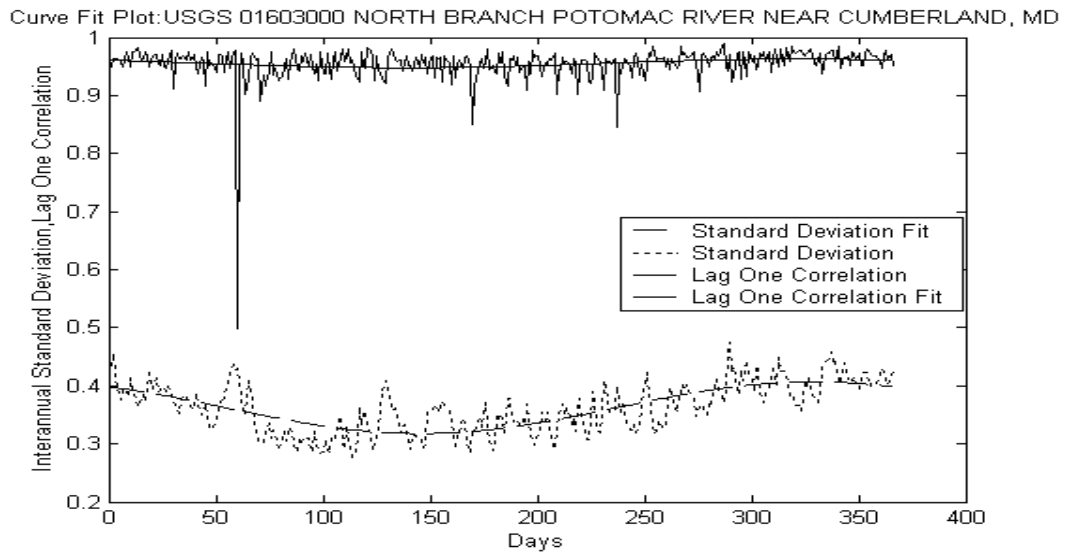


Figure 3.13 Curve Fit Plots of interannual standard deviation and lag-one correlation, North Branch Potomac at Cumberland.

Decorrelated Data Plot:USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD

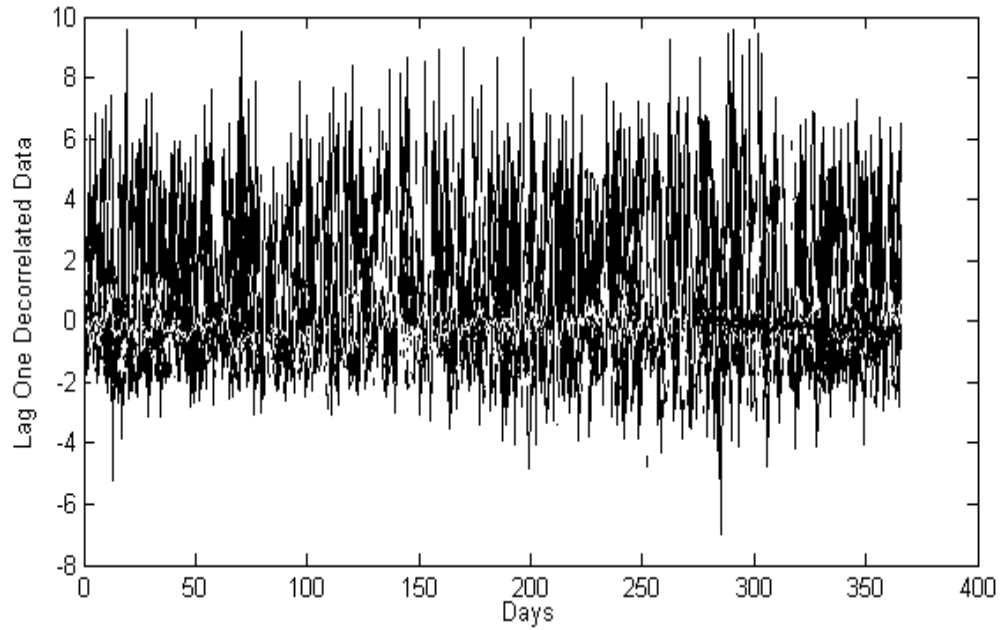


Figure 3.14 Decorrelated Data ( $K$ ) for North Branch Potomac

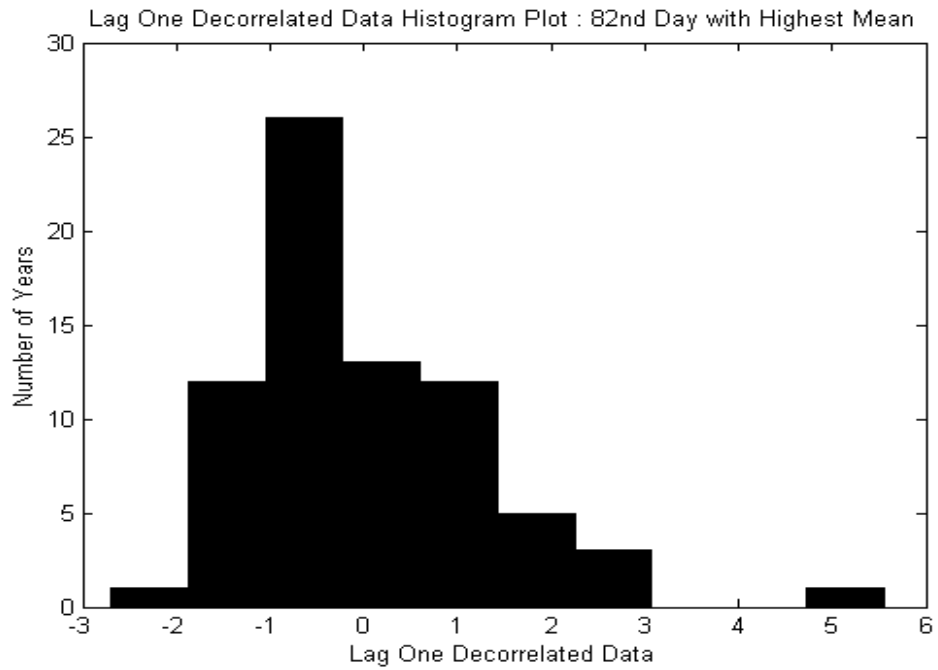


Figure 3.15 Histogram plot of lag-one decorrelated data of Day 82 (day of maximum mean)

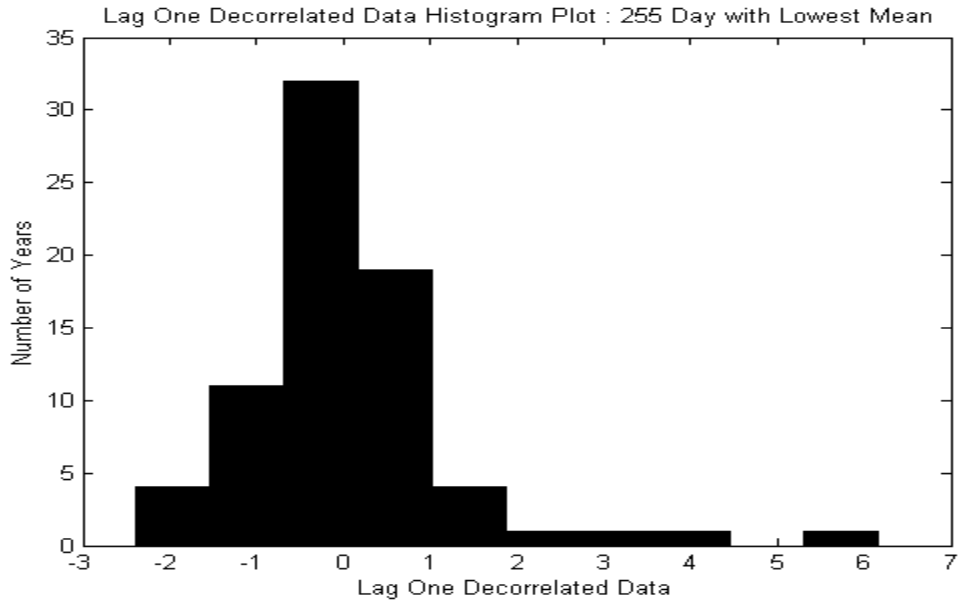


Figure 3.16 Histogram plot of lag-one decorrelated data of Day 255 (day of minimum mean)

The next step is the calculation of the skew and Vogel parameters for each day of the year. As mentioned above, the routine converts any negative skew to 0.16, the value determined to be the minimum value acceptable for the Matlab and Excel gamma distribution inversion routines. Whenever the routine identifies a data set with negative skew, a histogram is plotted to allow inspection of the data. In all such cases, the data tend to cluster around the mean, and the negative skews are very small in magnitude, lending support to (or at least insufficient to contradict) the belief that the underlying population is positively skewed. An example is shown (Fig. 3.17) for Day 252, North Branch Potomac; for this particular day, the sample over 73 years did not include any rare high values; the sample histogram is nearly symmetrical, with a sample skew of (-0.621). The analysis routine uses  $g=0.16$  in the Vogel test for the Pearson III distribution.

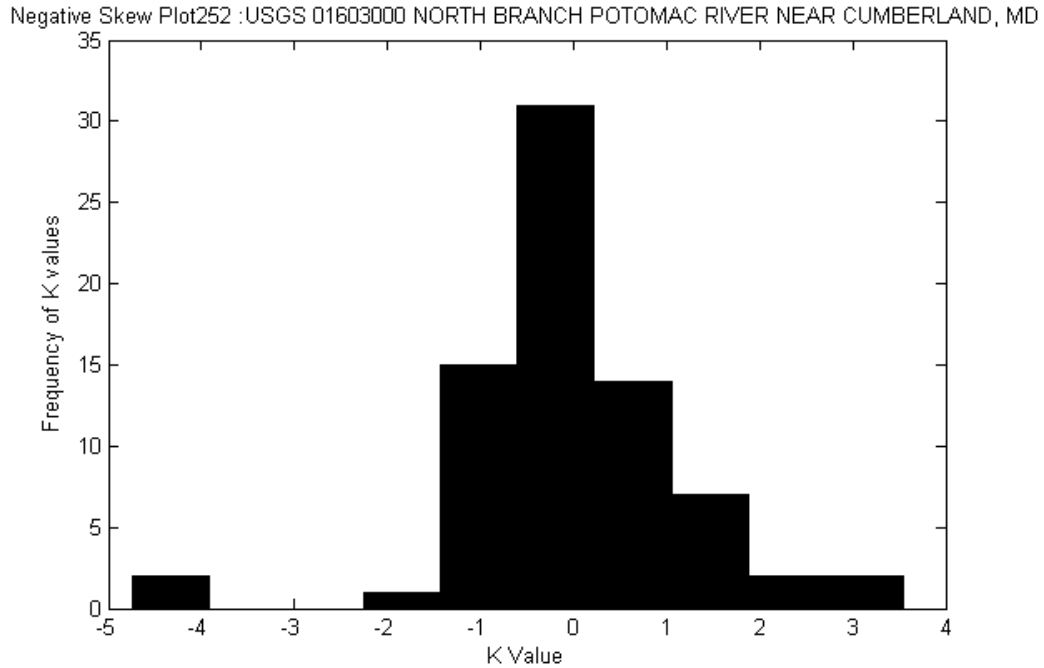


Figure 3.17 Histogram plots of sample data with negative skew, Day 252 for North Branch Potomac (total of 73 years).

The Matlab routine computes the Vogel test statistic (Probability Plot Correlation Coefficient) for both the standard normal and the Pearson III (using the sample skew or 0.16, whichever is greater). These values are provided as an output table.

The final product of the Matlab analysis routine is a table providing the interannual statistics for the decorrelated, deseasonalized, log-transformed data on a daily basis: Mean (should be 0), Standard Deviation (should be 1), skew, and Vogel statistics for the normal and Pearson III distribution. An example is shown in Table 3-1, which is the output for Days 1-23 for North Branch Potomac River, Cumberland, MD. The sample skew (column 2) ranges from 0.16 to 5.17 and the Vogel statistic for the Pearson III distribution (column 4) is in all cases greater than for the normal distribution (column 3), except day 13 when the skew is smallest. All 366 daily values of skew were averaged to obtain  $g = 2.29$  for the North Branch Potomac at Cumberland, providing the tenth parameter included in Table 3-1.

A complete graphical check on the pdf goodness-of-fit can be done in MS Excel for dates of interest. Several cases are shown below. On some days, the skew is nearly zero, and the Vogel parameters for normal and Pearson III are equal (Fig. 3.18). For days with large skew, there is a corresponding difference between the Vogel statistics, with the Pearson III giving a higher value (better fit) than the Normal (Fig. 3.20). In most cases, the difference between the Vogel parameters is not large, and the Pearson III value is only slightly better than the normal (Fig. 3.19). For the North Branch Potomac, the Vogel parameters ranged between 0.6 and 0.8; a value of 1.0 would indicate a perfect fit. As shown in Fig. 3.18, 3.19, and 3.20 Pearson III does not consistently match the sample data perfectly, in some cases only slightly better than the normal distribution. However, as a general model to be applied for all cases, the Pearson III is more flexible than the normal and appears to be a good choice for this application.

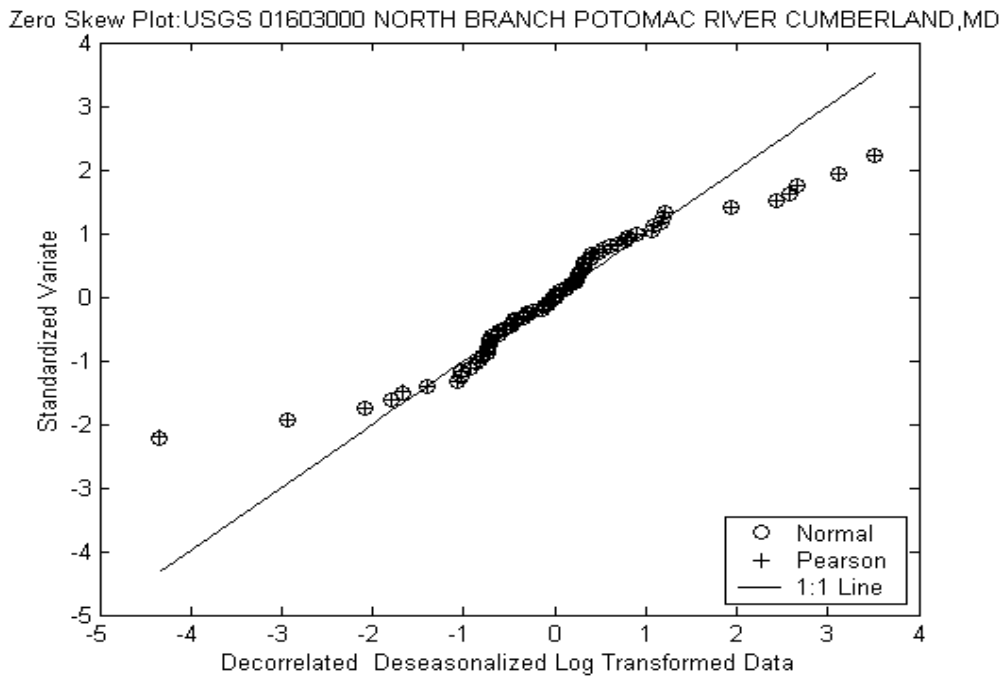


Figure 3.18 Probability Plot for Day 182 North Branch Potomac (Skew =0.02). Normal and Pearson III describes the data equally well.

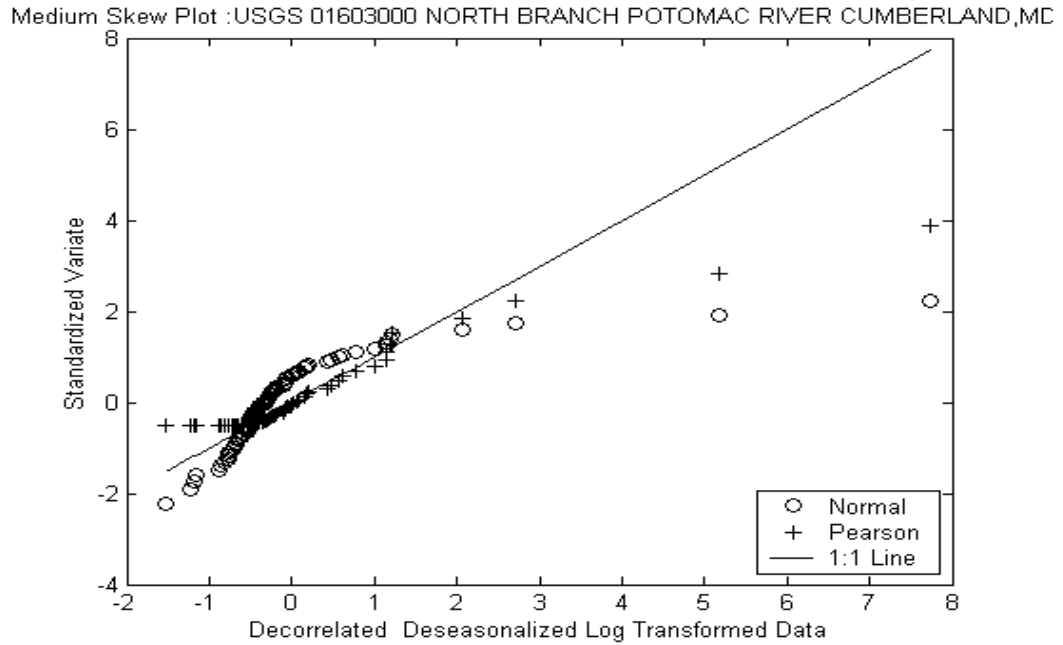


Figure 3.19 Probability Plot for Day 45 North Branch Potomac Medium Skew Plot (Skew=3.56)  
 Pearson III fits the sample slightly better than normal.

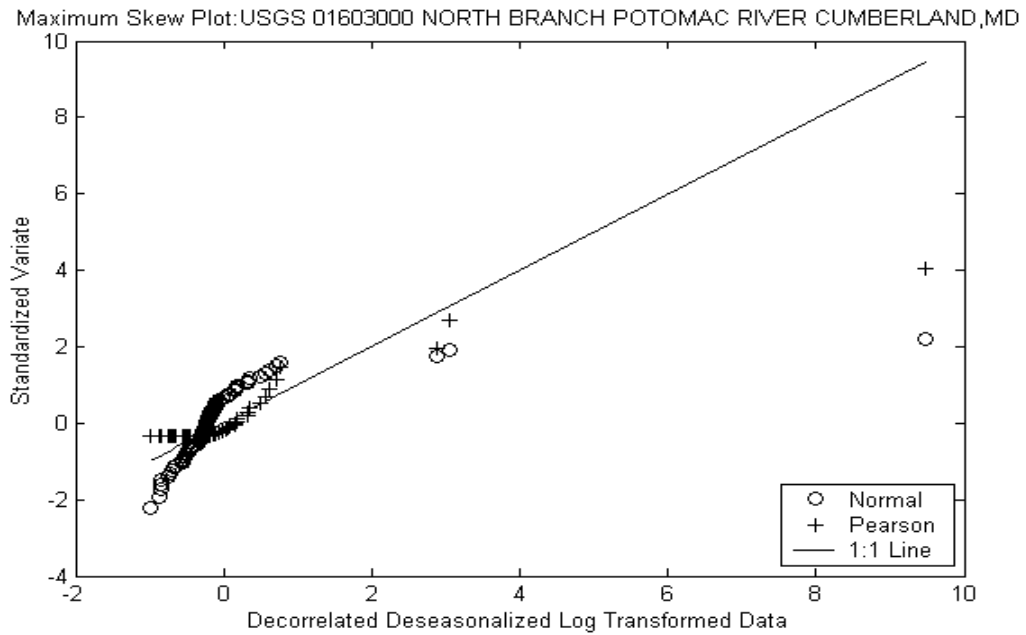


Figure 3.20 Probability Plot for Day 288, North Branch Potomac Maximum Skew Plot (Skew=5)  
 Pearson III fits the sample better than normal.

The final output table for all study watersheds and summary figures are included and discussed in Chapter 4.

## **CHAPTER 4 QUANTITATIVE RESULTS AND DEMONSTRATION OF MODEL**

### **4.1 OVERVIEW**

After applying the procedures of Chapter 3, ten parameters were generated for each of the 78 watersheds in the study. The ten-parameter model can be used to generate synthetic continuous daily discharge hydrographs that have the same statistical properties as the data record. This chapter describes the complete results, explores the variation in model parameters, and demonstrates the use of the model in hydrograph synthesis.

### **4.2 MODEL PARAMETERS**

For each analyzed watershed, the program returned three parameters for each of the cosine models (interannual average, interannual standard deviation, lag-one correlation), as well as the average of 366 values of skew for the decorrelated data. The complete results of the analysis (ten parameters for each watershed) are tabulated in Appendix A. The appendix also includes the drainage area for each of the study watersheds, because it was expected that some of the model parameters would be mathematically related to drainage area.

### **4.3 DISTRIBUTION OF PARAMETERS**

In order to have better understanding of the parameter distribution among watersheds, the parameters are plotted in different ways. Fig. 4.1 is a frequency histogram of the mean of the interannual average cosine model. This parameter has a range from 0.2 to 4.3. Most of the watersheds have values between 1.5 and 2. Very few watersheds had values more than 2.5. In Fig. 4.2, the means of the cosine model are plotted against the drainage area on a log



scale. Each point represents one of the study watersheds; the open circle indicates North Branch Potomac at Cumberland, the gage for which detailed analysis was provided in Chapter 3. The average log discharge is proportional to the log of the drainage area.

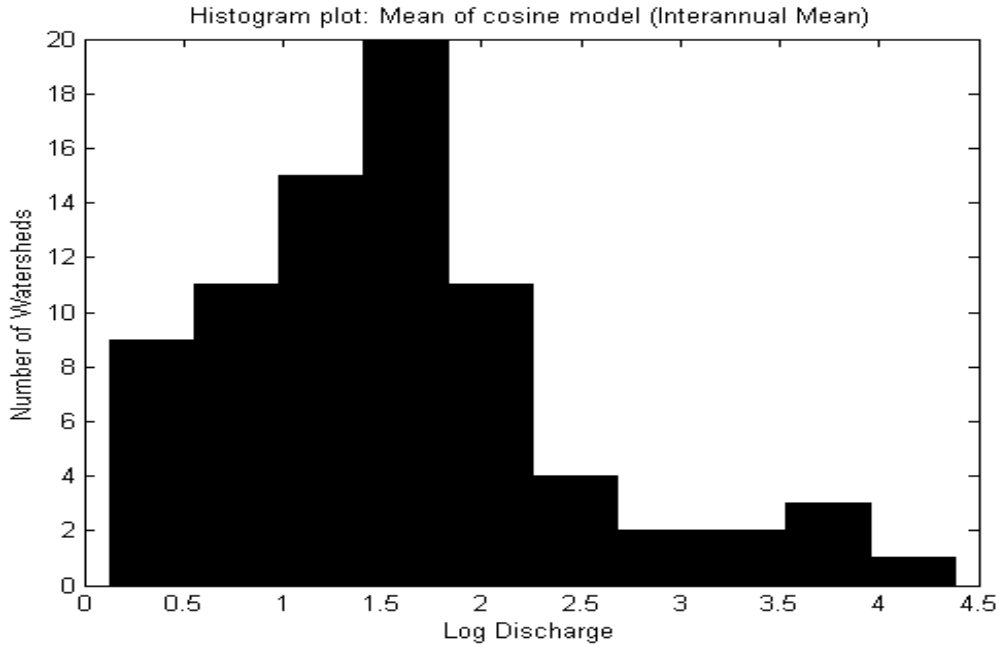


Figure 4.1 Frequency distribution of mean of the cosine curve fit for Interannual Average

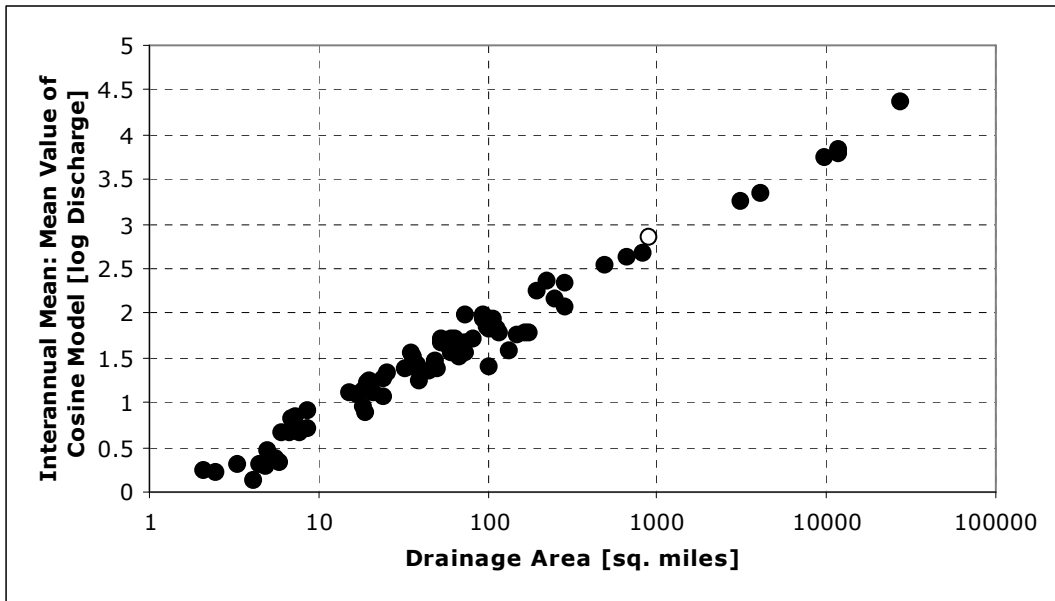


Figure 4.2 Mean of Cosine model for Interannual Average versus drainage area

The amplitude of the cosine model for interannual mean has the largest number of watersheds in the range of 0.2-0.25 (Fig. 4.3). The amplitude does not show a clear relationship to drainage area (Fig. 4.4). Each point in Fig. 4.4 represents a watershed. The y axis is the amplitude of the interannual mean ( $y_{a,M}$  in Eq. 3.22); this parameter varies from 0 to 0.8.. The larger basins show less variation in this parameter than the smaller basins, but this may be due to the fact there are fewer large basins in the study set.

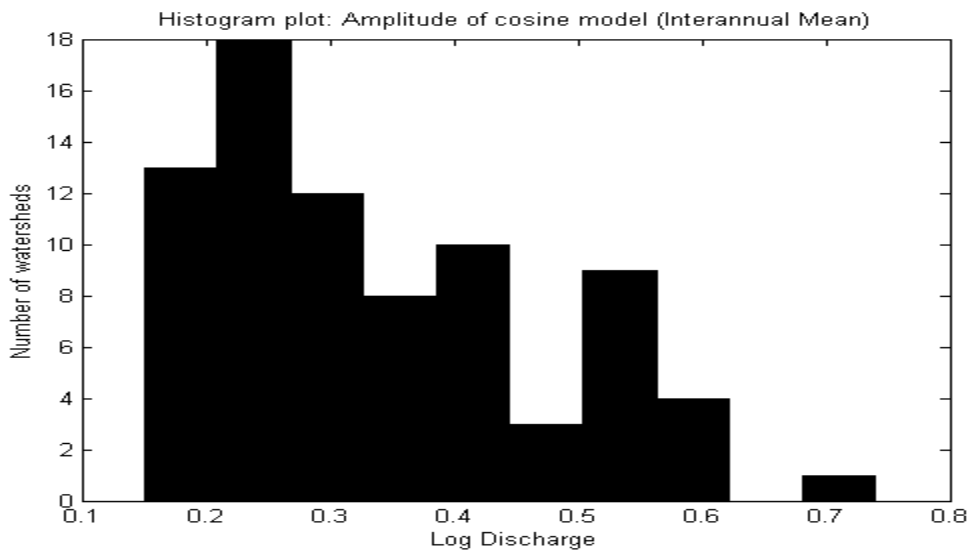


Figure 4.3 Frequency distribution of amplitude of the curve fit (Interannual Average)

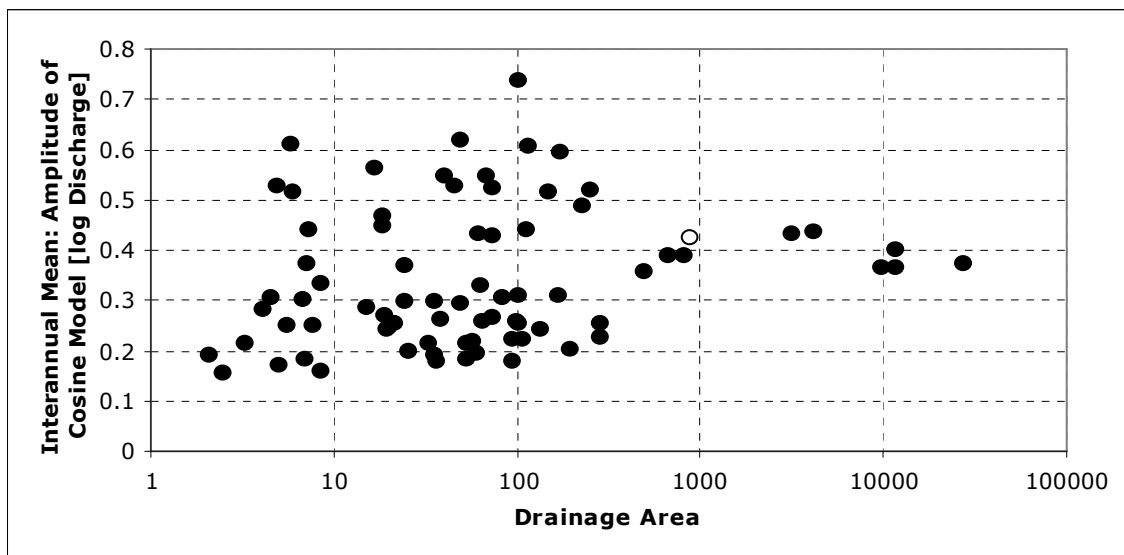


Figure 4.4 Interannual average: Amplitude of cosine model (log discharge) versus drainage area. Open circle is North Branch Potomac at Cumberland.

The day of the maximum for interannual average is generally during the winter/early spring i.e. between Days 70 and 80 with some outliers at Day 100 (Fig. 4.5). The day of maximum does not appear related to drainage area (Fig. 4.6). The y axis in Fig. 4.6 is the day of maximum for the interannual average ( $\tau_M$  in Eq. 3.22); this parameter varies from Day 50 to Day 100.

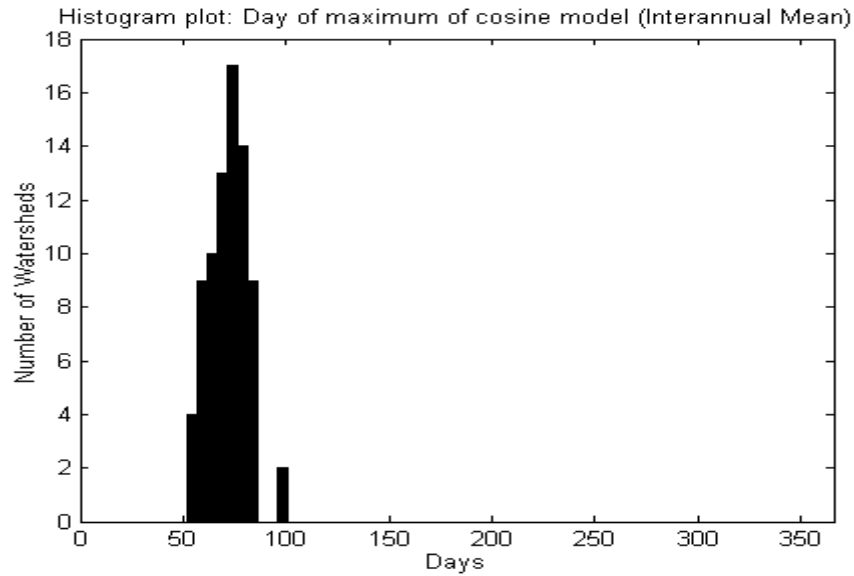


Figure 4.5 Frequency distribution of day of the maximum (Interannual Average)

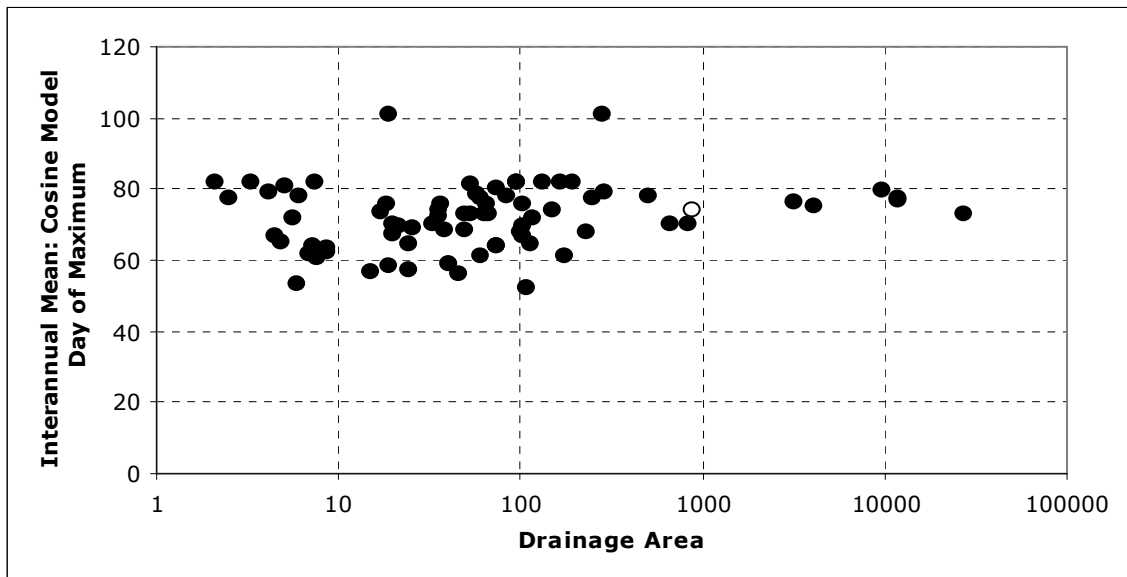


Figure 4.6 Interannual Mean: Day of maximum of cosine model versus drainage area

The mean of the cosine model fit for interannual standard deviation ranges from 0.25 to 0.85 (Fig. 4.7). The majority of the watersheds have values between 0.35 and 0.4. This parameter is also more variable for smaller watersheds, and is not directly proportional to drainage area (Fig. 4.8). In Fig. 4.8, the y-axis is the mean of interannual standard deviation ( $\bar{y}_s$  in Eq. 3.23); this parameter varies from 0.2 to 0.9.

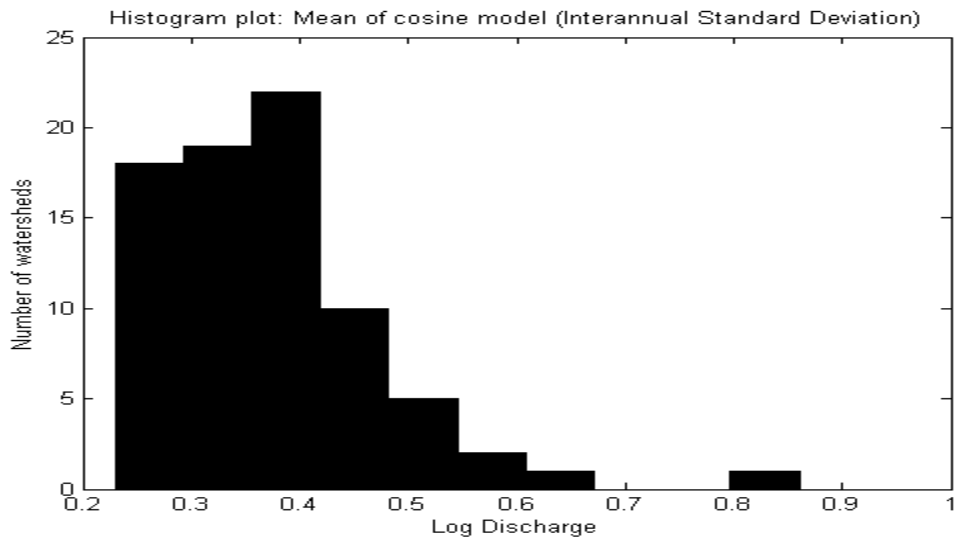


Figure 4.7 Frequency distribution of mean of the curve fit (Interannual Standard Deviation)

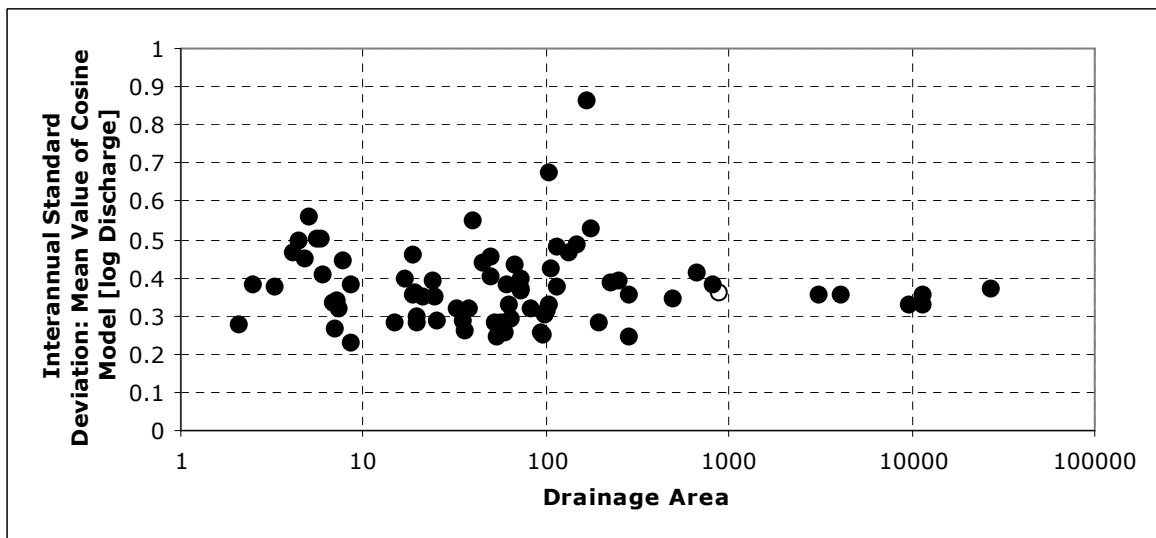


Figure 4.8 Interannual Standard Deviation: Mean of the cosine model (log discharge) versus drainage area. The open circle is for North Branch Potomac at Cumberland.

The amplitude of the cosine model for interannual standard deviation varies from 0 to 0.27. More than 20 watersheds have values in the range 0.025 to 0.05. The higher values of this parameter are associated with drainage areas of order 10 to 100 mi<sup>2</sup>. (Fig. 4.10). In Fig. 4.10, the y axis is the amplitude of interannual standard deviation ( $y_{a,S}$  in Eq. 3.23); this parameter varies from 0 to 0.3.

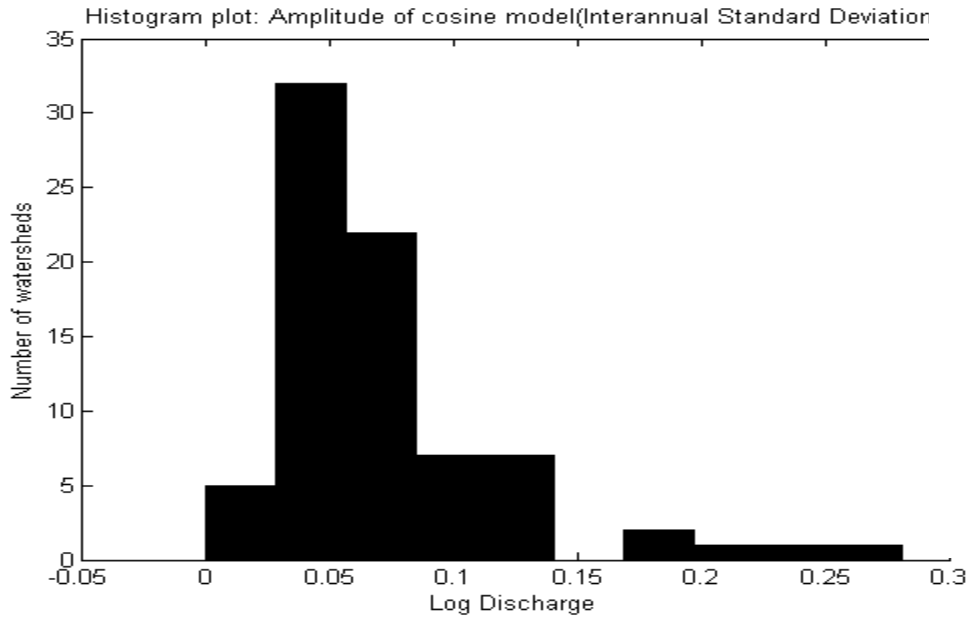


Figure 4.9 Frequency distribution of amplitude of the curve fit (Interannual Standard deviation)

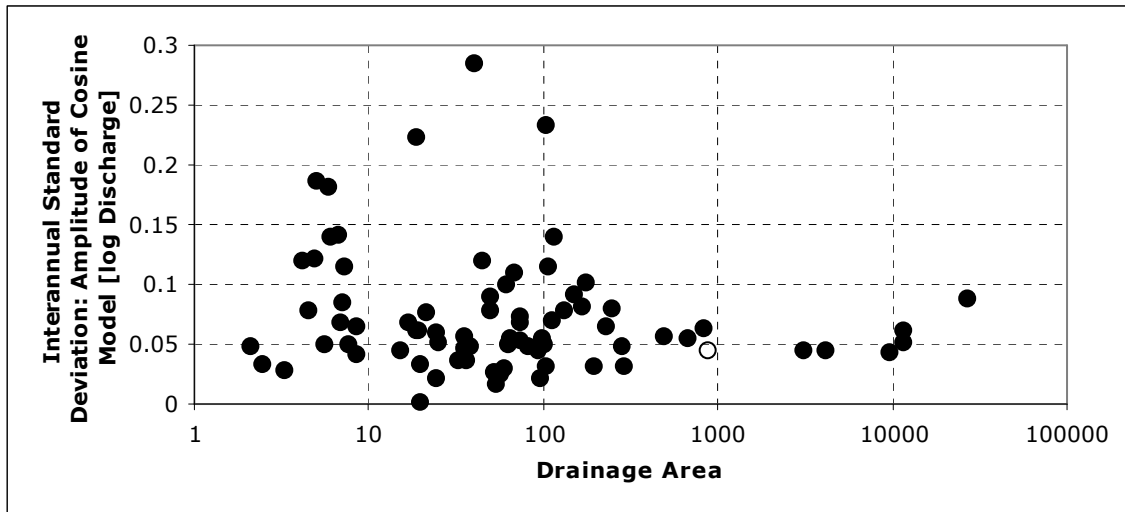


Figure 4.10 Interannual Standard Deviation: Amplitude of cosine model (log discharge) versus drainage area. Open circle is for North Branch Potomac at Cumberland.

The day of the maximum for interannual standard deviation has a range of 0-350 days (Fig. 4.11). More than 20 watersheds have the day of the maximum between Days 250 and 300 (mid-August through Sept.). Three outliers have maximum standard deviation earlier in the year (Day 50 to 100). By visual inspection, there is a slight linear relationship between this parameter and drainage area (Fig. 4.12).

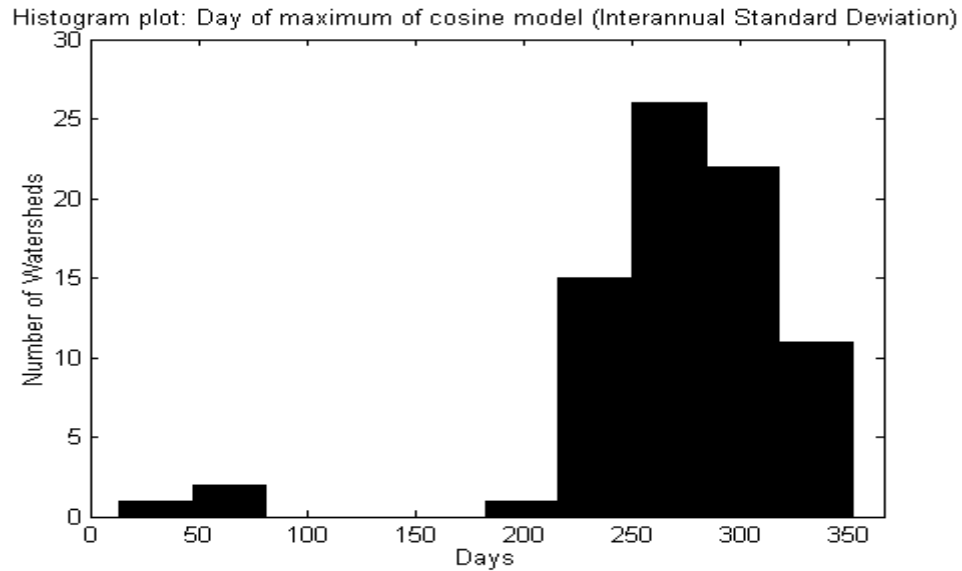


Figure 4.11 Frequency distribution of day of the maximum (Interannual Standard deviation)

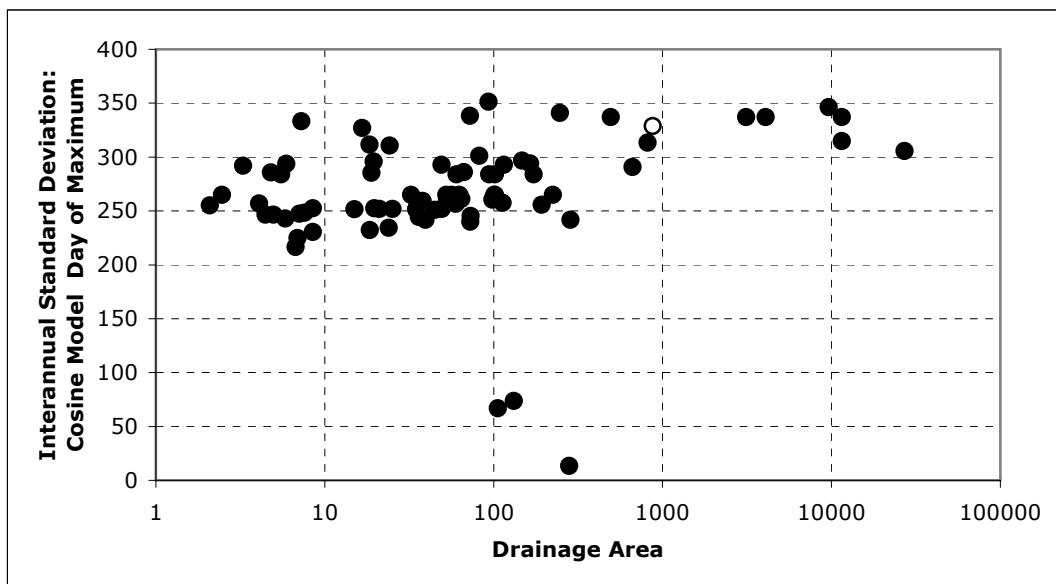


Figure 4.12 Interannual Standard Deviation: Day of Maximum of cosine model versus drainage area. The open circle is for North Branch Potomac at Cumberland.

In Fig. 4.12, the y-axis is the day of maximum interannual standard deviation ( $\tau S$  in Eq. 3.23); this parameter varies from Day 0 to Day 350. The 3 outliers are Patuxent River at Laurel, Savage River near Bloomington, and Antietam Creek near Sharpsburg; dams or major diversions affect these gage records, as discussed in Section 4.4.

The mean of the cosine model for lag-one correlation ( $\bar{y}_R$  in Eq. 3.25) ranges from 0.7 to 0.975. More than 18 watersheds have values in the range of 0.95-0.975 (Fig. 4.13). The high values of lag-one correlation indicate strong persistence in daily streamflow. Lower values occur with smaller drainage area (Fig. 4.14).

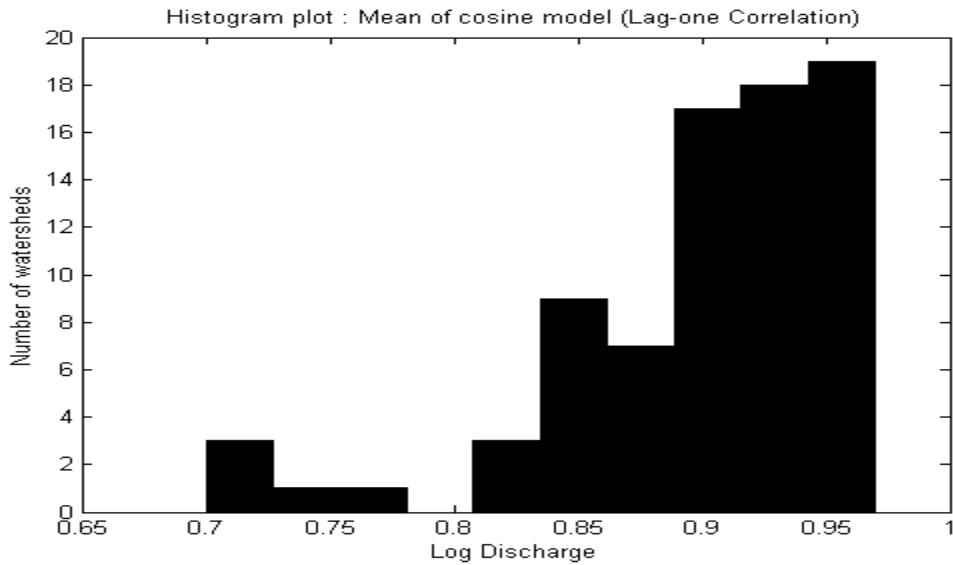


Figure 4.13 Frequency distribution of mean of the curve fit (Lag One Correlation)

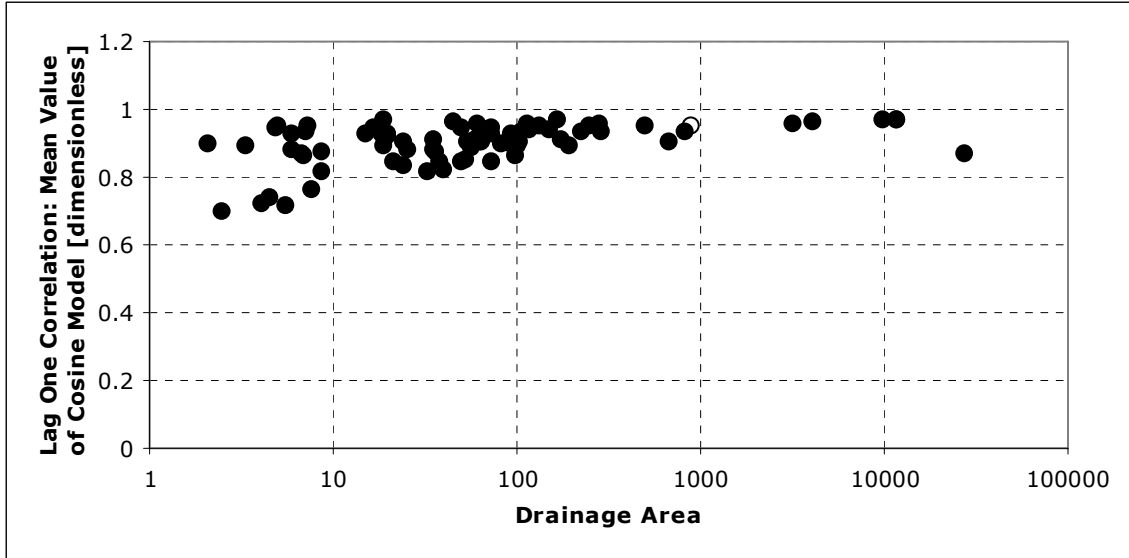


Figure 4.14 Lag One Correlation: Mean of the cosine model (Log Discharge) versus drainage area. The open circle for North Branch Potomac at Cumberland.

The amplitude of the curve fit for lag-one correlation varies from 0 to 0.04. Most of the watersheds have values between 0.00 and 0.02 (Fig. 4.15). The small amplitude (a few percent of the mean value) indicates that lag-one correlation is nearly constant throughout the year, and perhaps making this a constant parameter rather than using a cosine model could simplify the model. The largest amplitude of lag-one correlation is associated with the largest drainage area (Fig. 4.16).

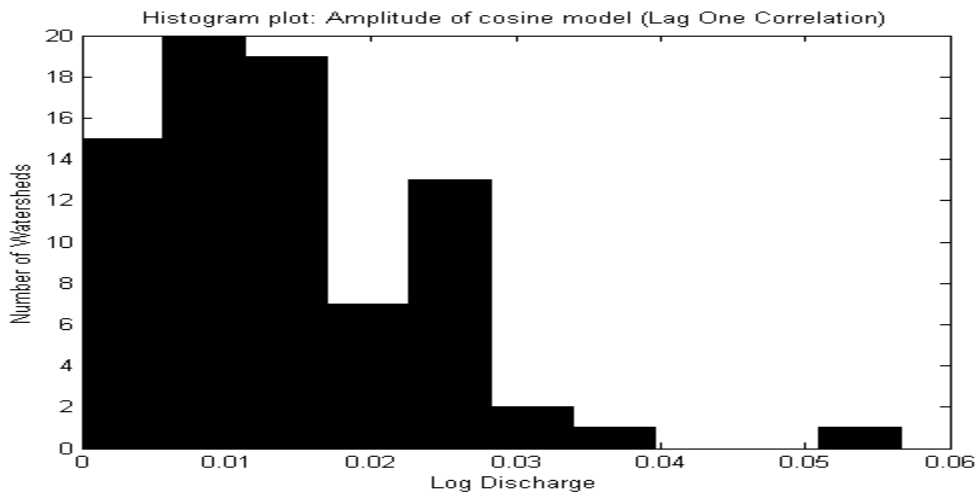


Figure 4.15 Frequency distribution of amplitude of the curve fit (lag-one correlation)



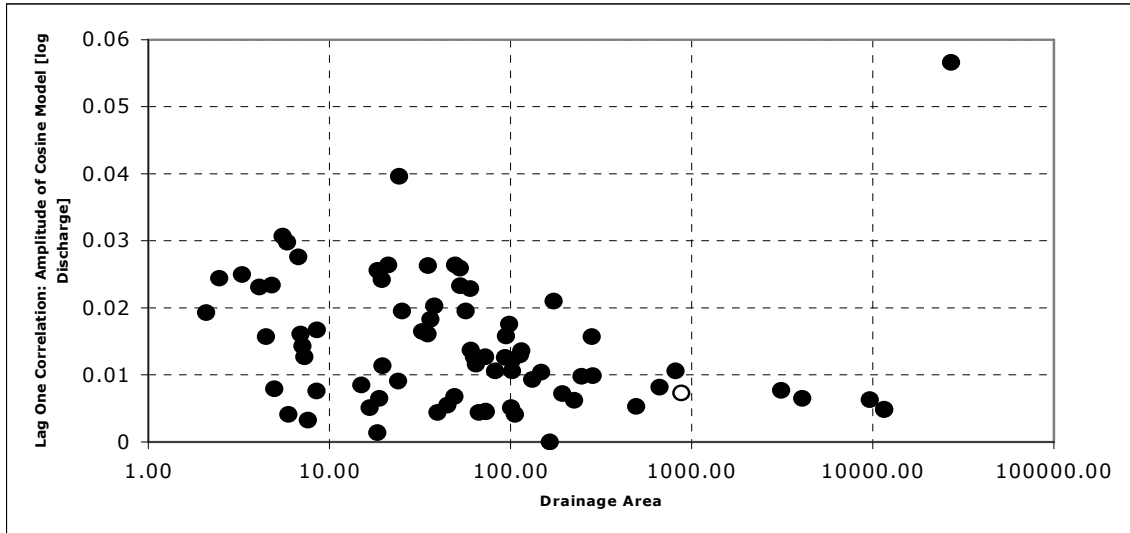


Figure 4.16 Lag-one correlation: Amplitude of cosine model (log discharge) versus drainage area. The open circle is for North Branch Potomac at Cumberland.

The day of the maximum for lag-one correlation ( $\tau_R$  in Eq. 3.25) is in the range of 175 to 200 for than 15 watersheds. This parameter values from 0 to 350 (Fig. 4.17). The day of maximum for lag-one correlation does not vary systematically with drainage area (Fig. 4.18), although the larger basins appear to cluster around day 300 (except for the largest basin). There is a noticeable gap around day 100. The day of maximum may not be very meaningful for lag-one correlation, because it is nearly constant throughout the year (small amplitude).

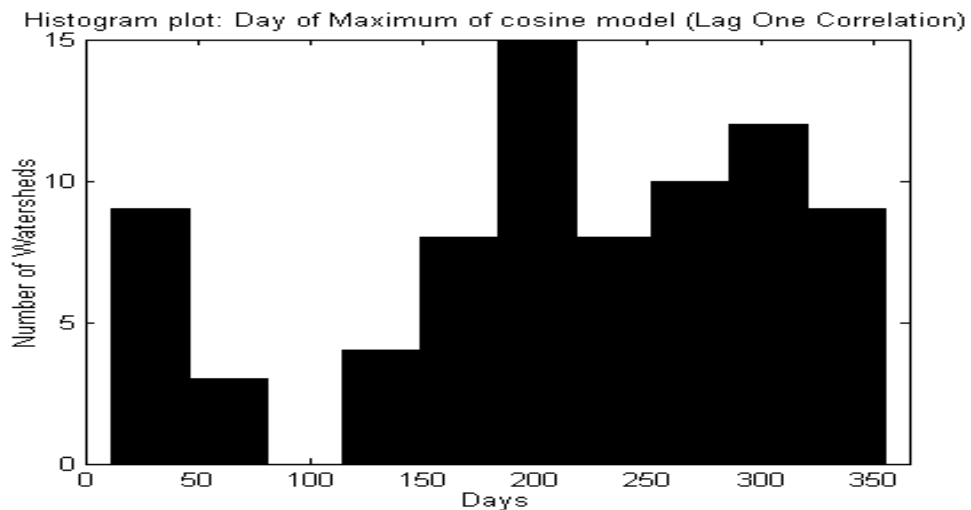


Figure 4.17 Frequency distribution of day of the maximum (lag-one correlation)

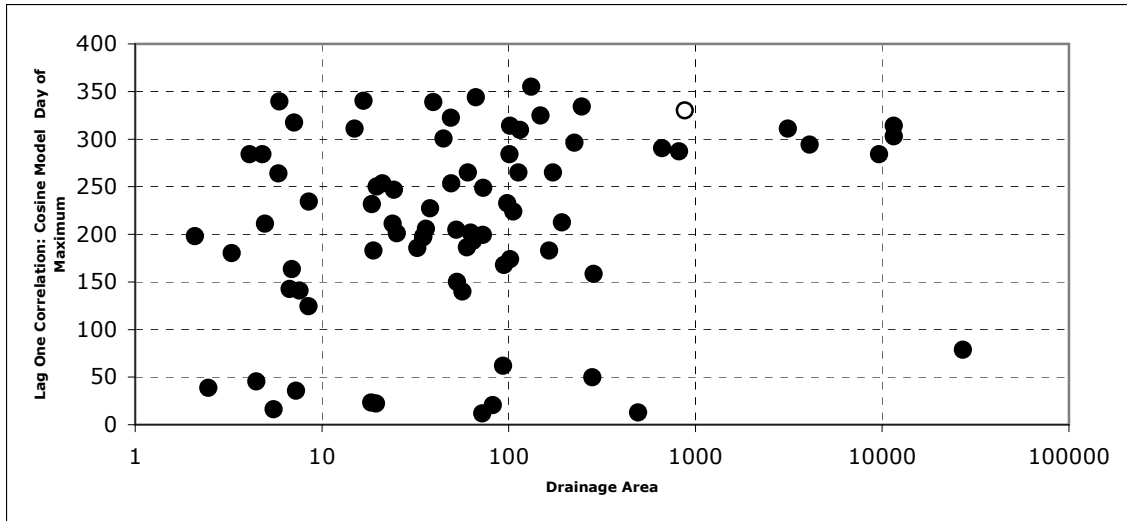


Figure 4.18 Lag-one correlation: Day of maximum of cosine model (log discharge) versus drainage area. The open circle is for North Branch Potomac at Cumberland.

For each watershed, 366 values of skew were computed for the decorrelated daily data. An average value of the 366 is reported for each watershed. Mean skew for watersheds is mainly in the range of 1.575 to 2.25, with the overall range of 0.5 to 2.75 (Fig. 4.19). There is a slight linear relationship between skew and drainage area (Fig. 4.20). Again, the value for the largest watershed appears to be an outlier.

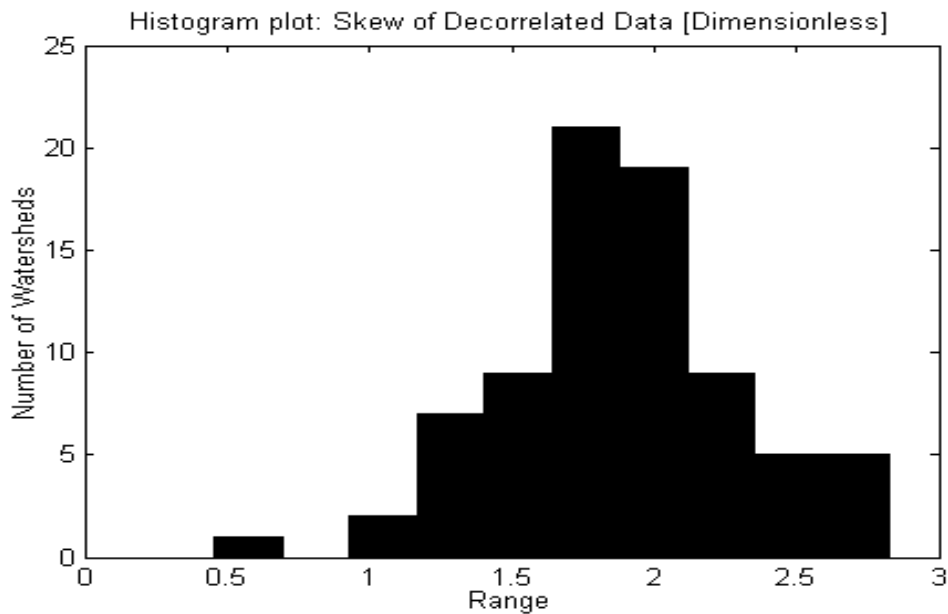


Figure 4.19 Frequency distribution of skew of decorrelated data (Average 366 values for each watershed).

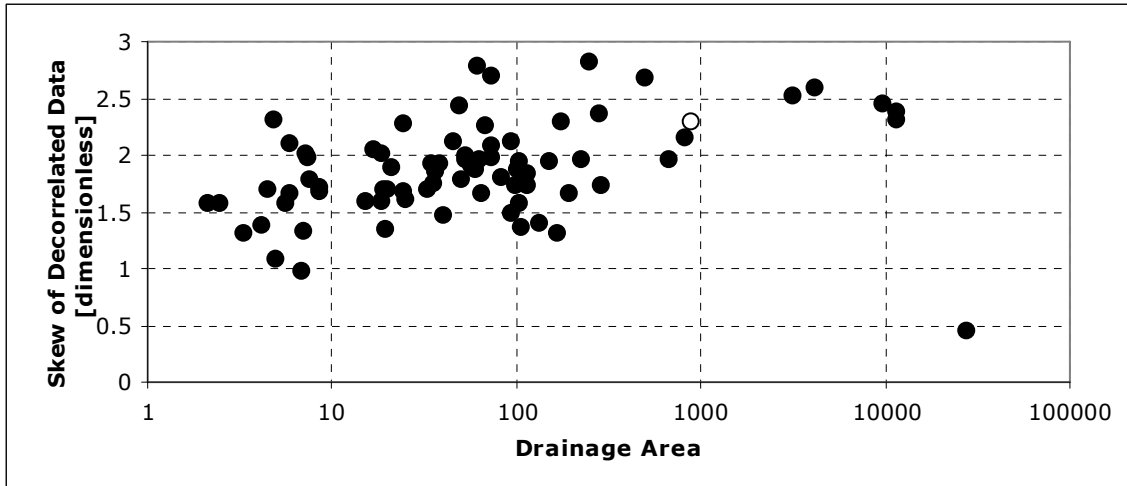


Figure 4.20 Skew of decorrelated data versus drainage area. The open circle is for North Branch Potomac at Cumberland.

#### 4.4 PARAMETER INVESTIGATION

Although it was not the goal of this project to analyze the physical meaning of the model parameters, it is interesting to investigate some of the unusual results. This section examines the effect of artificial storage on the model parameters.

When the study gages were selected, they were not screened for dams or diversions. The Little Patuxent at Guilford (Fig. 4.21), Little Patuxent at Savage (Fig. 4.22), and Patuxent near Laurel (Fig. 4.23) are from the same basin (Patuxent). The Little Patuxent River at Guilford has a drainage area of  $38 \text{ mi}^2$  and peak interannual average of 56.2 cfs and Little Patuxent River downstream at Savage, with drainage area of  $98.4 \text{ mi}^2$  has peak interannual average of 158.5 cfs (approx.). But the Patuxent River has a drainage area of  $132 \text{ mi}^2$  and still has peak interannual mean of only 100 cfs, because of the control structure and diversion on Patuxent River. The reservoir just upstream of this gaging station provides water supply to Metropolitan Washington, DC, and the withdrawals are not accounted for in the streamflow record (USGS 2005). The control structure, storage, and withdrawal of water supply have a strong effect on variability in streamflow. The two Little Patuxent gages have

maximum variability in late summer and early fall as shown by the peak interannual standard deviation at that time (Figs. 4.24 and 4.25). The standard deviation for Patuxent at Laurel is at a minimum in late summer/early fall, because the function of the water supply reservoir is to assure water supply throughout the year. The water supply reservoir explains one of the outliers in Figs. 4.11 and 4.12. The other two are the Savage River below Savage Dam near Bloomington, MD (drainage area 106 mi<sup>2</sup>) and Antietam Creek near Sharpsburg (281 mi<sup>2</sup>). Unlike the other two, the Antietam Creek gage is not located below a dam (USGS 2005c); however, it receives municipal discharge from the city of Hagerstown. The city's water supply is diverted from the Potomac; municipal water use suppresses the natural cycle of variability in Antietam Creek:-

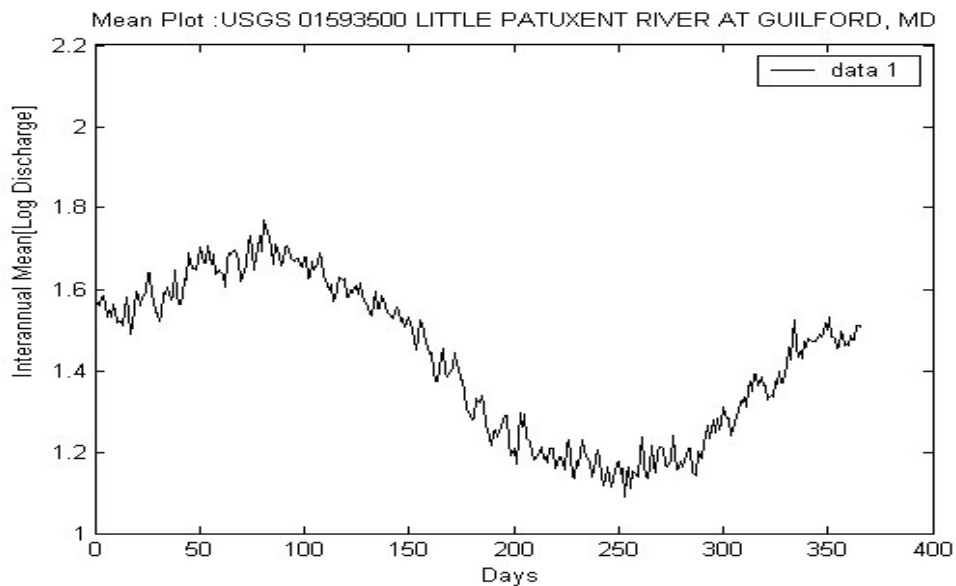


Figure 4.21 Interannual Average plot of Little Patuxent River at Guilford

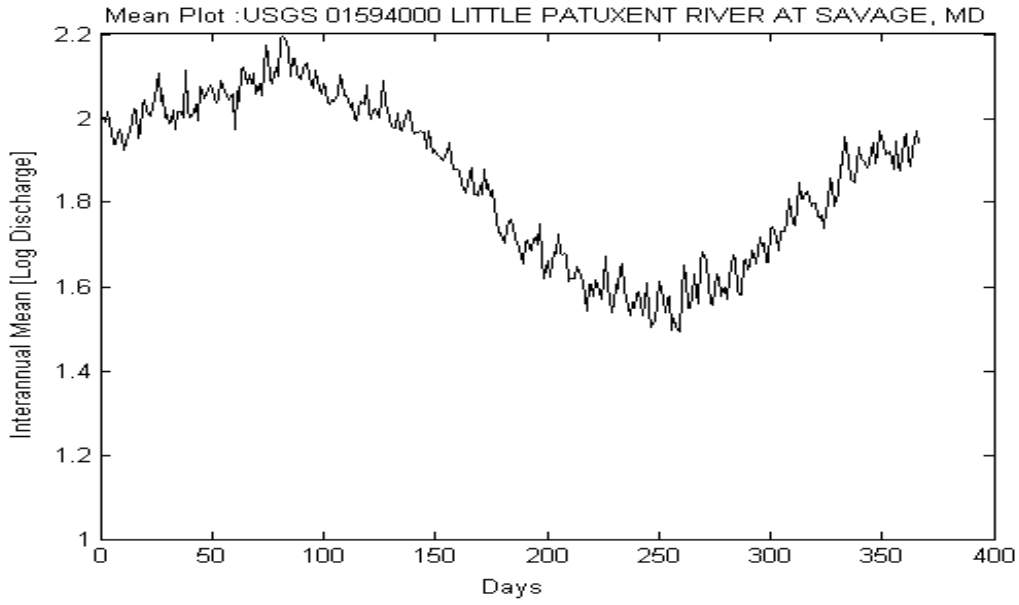


Figure 4.22 Interannual Average plot of Little Patuxent River at Savage

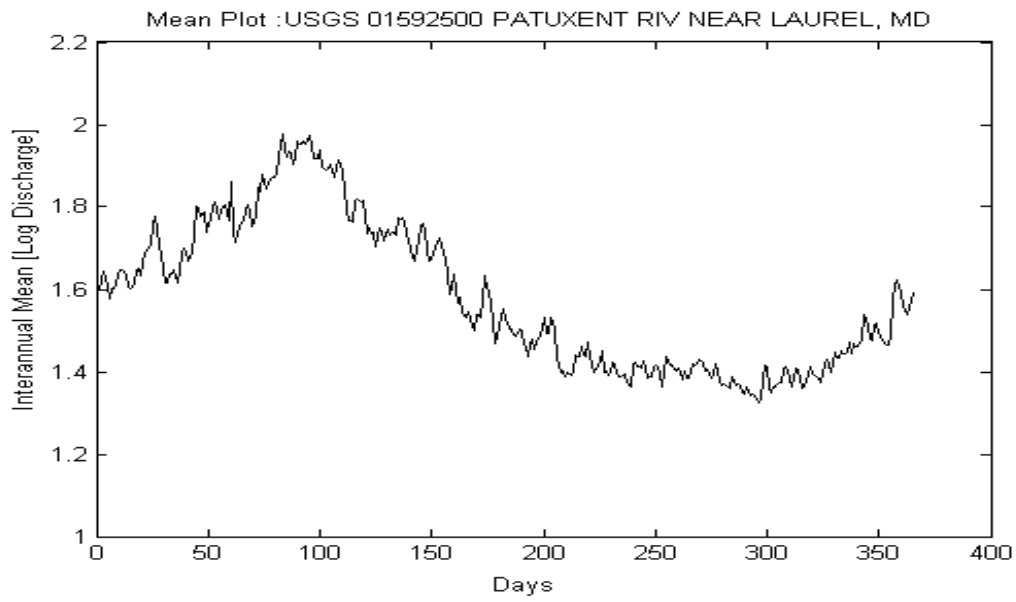


Figure 4.23 Interannual Average plot of Patuxent River near Laurel

Standard Deviation Plot :USGS 01593500 LITTLE PATUXENT RIVER AT GUILFORD, MD

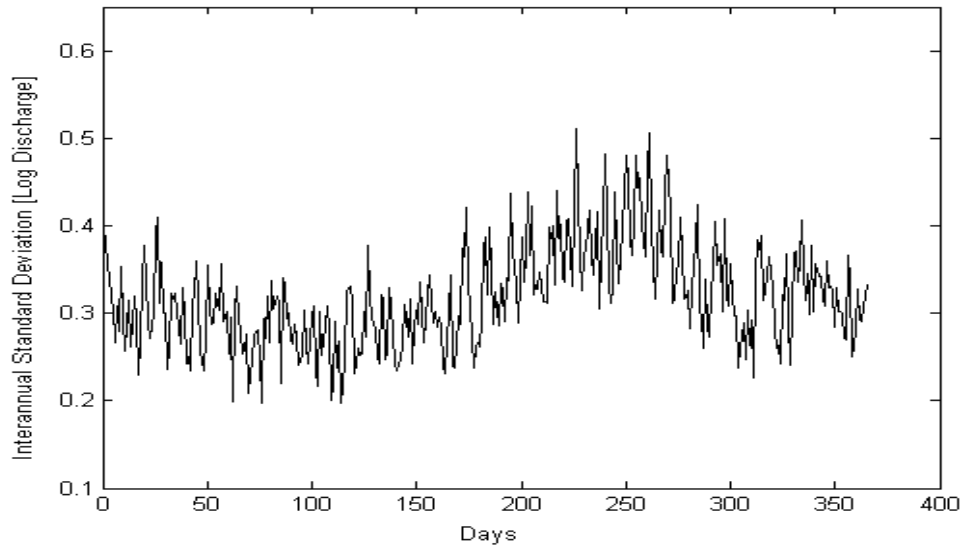


Figure 4.24 Interannual Standard Deviation plot of Little Patuxent River at Guilford

Standard Deviation Plot :USGS 01594000 LITTLE PATUXENT RIVER AT SAVAGE, MD

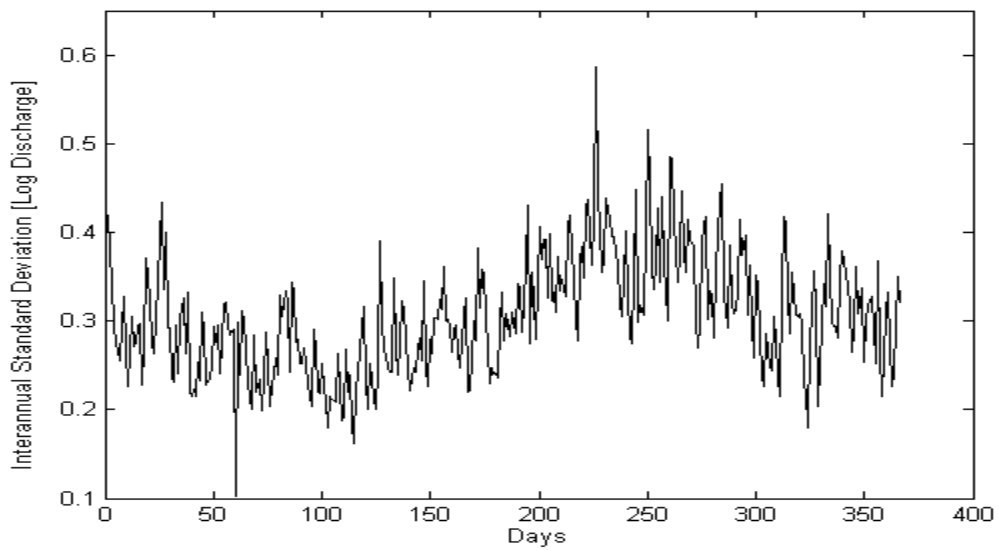


Figure 4.25 Interannual Standard Deviation plot of Little Patuxent River at Savage

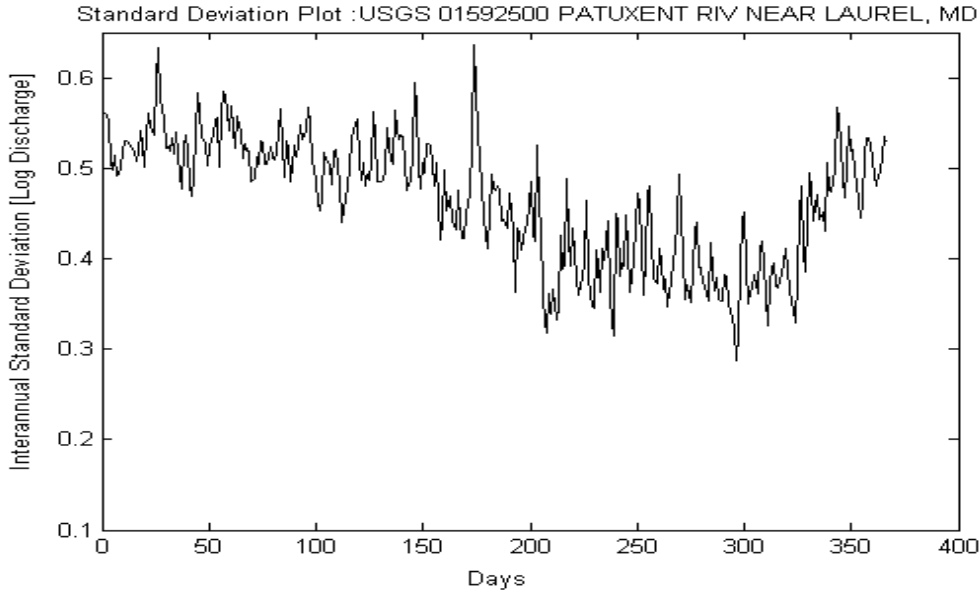


Figure 4.26 Interannual Standard Deviation plot for Patuxent River near Laurel

#### 4.5 SUMMARY OF PARAMETERS

The range of the parameters is shown in Table 4-2. The maximum drainage area is 27000  $mi^2$  for USGS Gage 0178310 Susquehanna River at Conowingo, MD and minimum of 2.09  $mi^2$  USGS Gage 0158000 Slade Run near Glyndon, MD. Table 4-2 gives an idea of how the drainage area is or is not associated with the ten model parameters.

Table 4-1 Range of Parameters

Range	Interannual Mean			Interannual Standard Deviation			Lag-one Correlation			Decorrelated Data Skew
	Mean	Amplitude	Day of Maximum days	Mean	Amplitude	Day of Maximum days	Mean	Amplitude	Day of Maximum days	
Max	4.3814	0.7379	101.0014	0.8629	0.2842	351.3432	0.9723	0.0566	354.8459	2.8248
Min	0.1314	0.1505	52.1983	0.2309	0.0013	13.3479	0.7005	0.0000	11.7241	0.9873

Table 4-2 Parameters of maximum and minimum drainage area

Name of the Location	Drainage Area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag-one Correlation			Decorrelated Data Skew
		Mean	Amplitude	Day of the maximum days	Mean	Amplitude	Day of the maximum days	Mean	Amplitude	Day of the maximum days	
USGS 01578310 SUSQUEHANNA RIVER AT CONOWINGO, MD	27,100	4.38	0.37	73	0.37	0.08	306.	0.86	0.06	79	0.4558
USGS 01583000 SLADE RUN NEAR GLYNDON, MD	2.09	0.25	0.19	82	0.27	0.05	255	0.90	0.02	198	1.5859

#### 4.6 SYNTHETIC CONTINUOUS DAILY HYDROGRAPH

The use of the model is demonstrated by generating 80 years of synthetic data for three watersheds of different sizes using the parameters identified in the analysis phase: North Branch Potomac at Cumberland (877 mi<sup>2</sup>), White Marsh Run at White Marsh (7.61 mi<sup>2</sup>), and Beaver Dam Branch at Matthews (5.85 mi<sup>2</sup>). The synthesis technique is described in the previous chapter.

Two of the 80 synthetic years for the North Branch Potomac at Cumberland are shown in Fig. 4.27 (a) and (b). The continuous hydrograph shows realistic peaks and recessions. Fig. 4.27 (c) shows that the synthetic model successfully captures the statistics of the observed data: Interannual mean, and interannual standard deviation. The lag-one correlation (not shown) matched equally well.



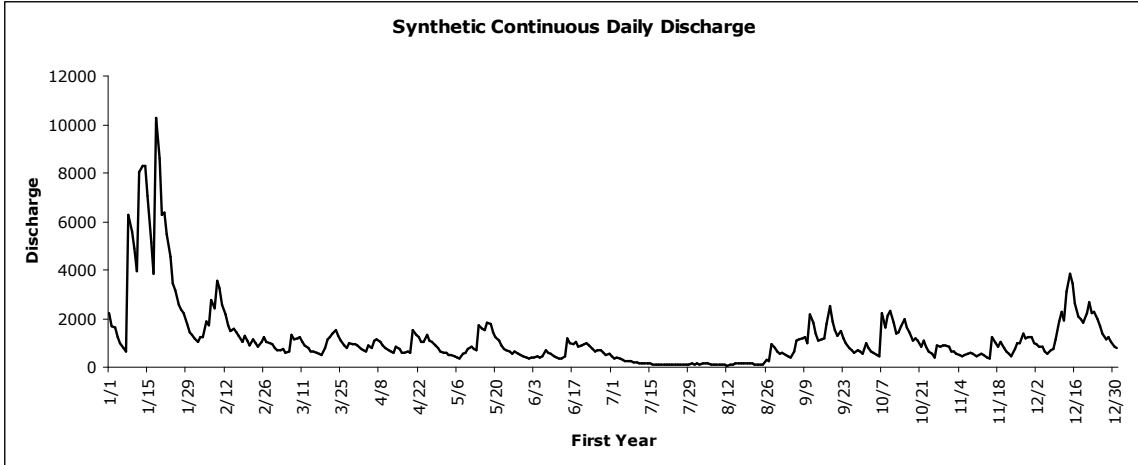


Figure 4.27(a) Continuous Daily Hydrograph for North Branch Potomac, Cumberland (First Year)

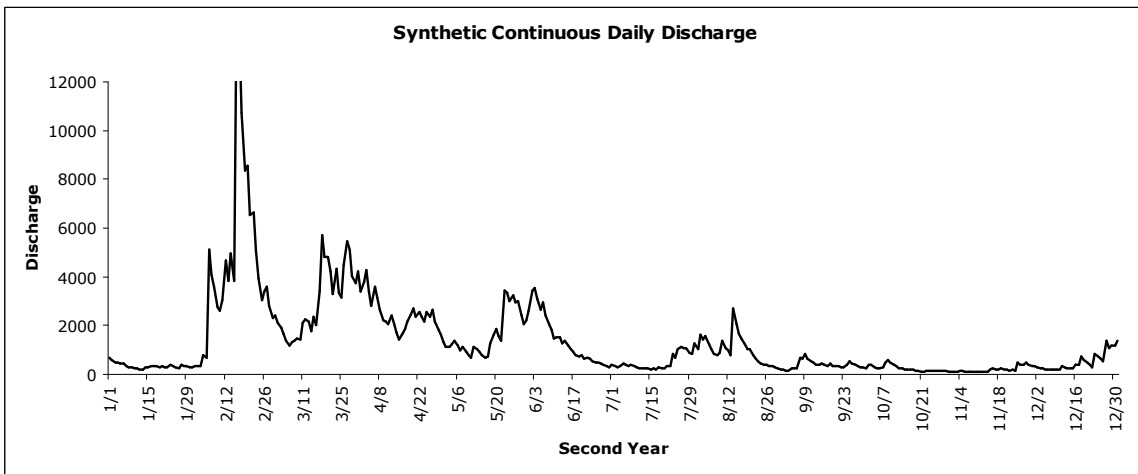


Figure 4.27(b) Continuous Daily Hydrograph for North Branch Potomac, Cumberland (Second Year)

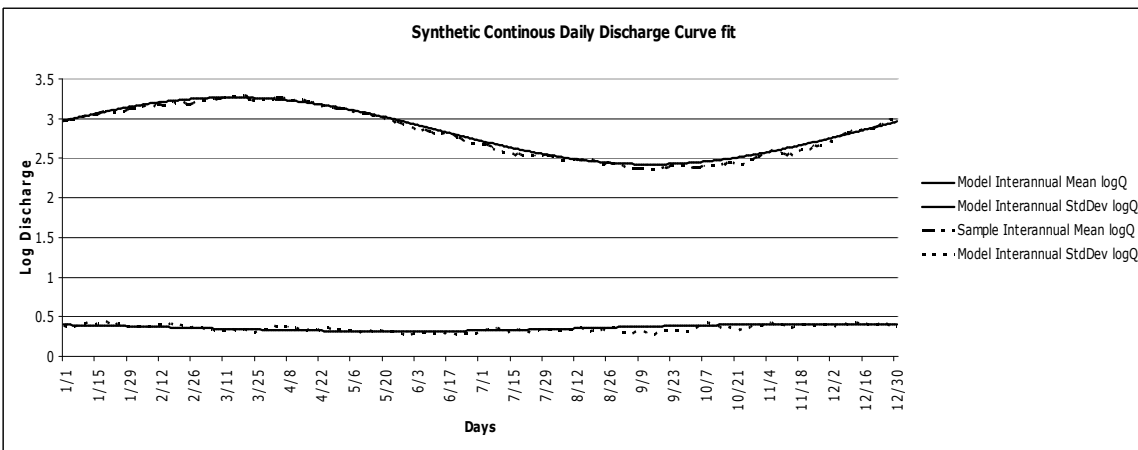


Figure 4.27(c) Continuous Daily Hydrograph from North Branch Potomac, Cumberland. The smooth curves are as derived from real data in the analysis phase. The somewhat irregular curves are the average and standard deviation of the 80 years of synthesized data.

White Marsh Run watershed is very small in comparison to North Branch Potomac, yet larger than Beaver Dam Creek, so this is the second example included to demonstrate the use of the model. Two sample years are shown in Figs. 4.28(a) and (b), and the interannual cycles of the synthetic data are compared to the curve fit observed values in Fig. 4.28(c). Each watershed's 80-year synthetic hydrograph is started from a different random number seed, and they are not based on precipitation inputs. Therefore, the synthetic hydrographs for the different watersheds are not meant to represent the same periods of time.

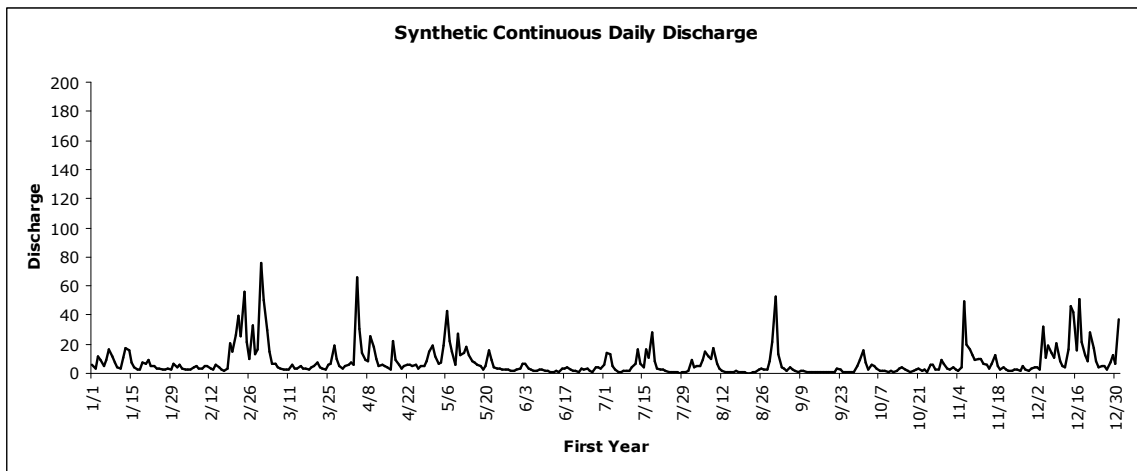


Figure 4.28(a) Synthetic Continuous Daily Hydrograph for White Marsh Run (First Year)

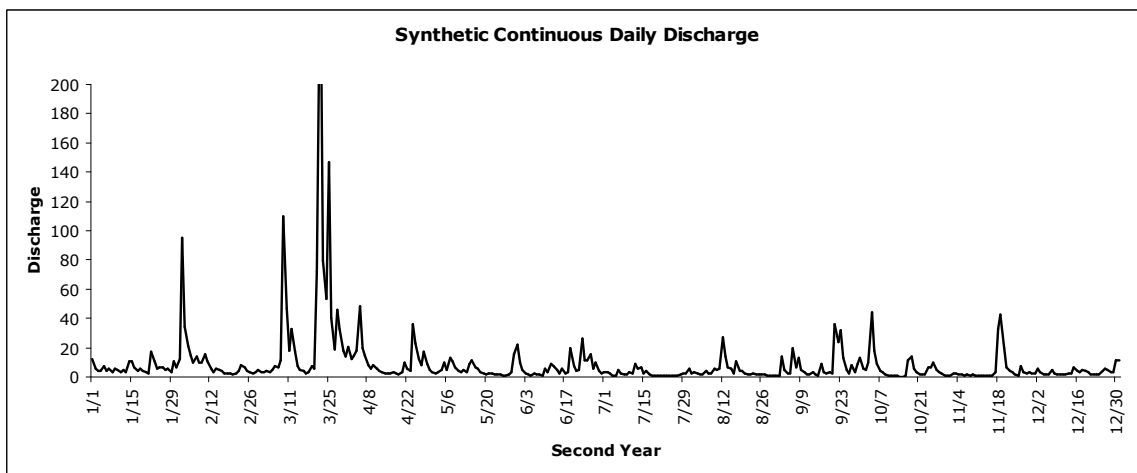


Figure 4.28(b) Synthetic Continuous Daily Hydrograph for White Marsh Run (Second Year)

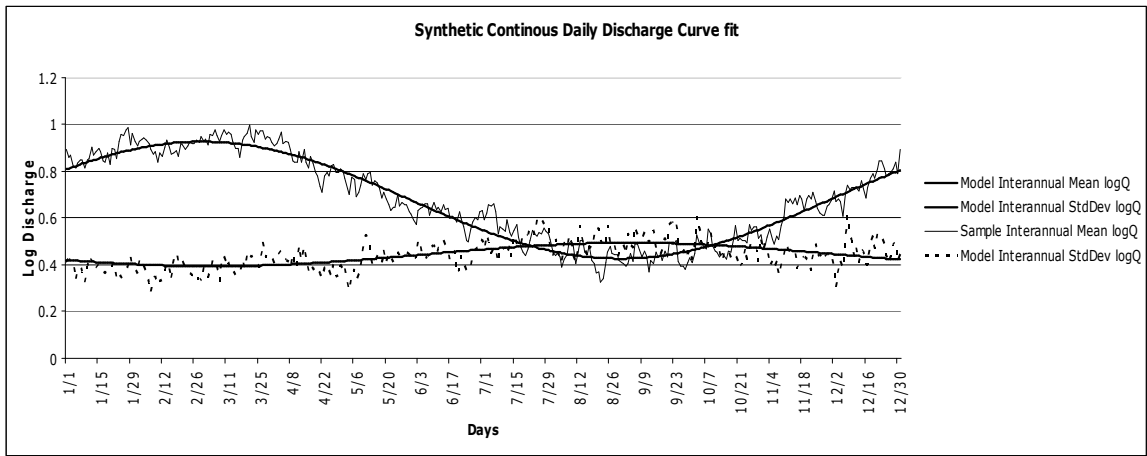


Figure 4.28(c) Continuous Daily Hydrograph from White Marsh Run. Comparison of interannual statistics from synthetic streamflow (80 years) to cosine models derived from real data.

The third example, Beaverdam Branch, is an even smaller watershed. Its discharge is less than White Marsh, but still realistic in its patterns of rise and fall [Fig. 4.29(a) and (b)] and successful in capturing the statistics of the observed data [Fig. 4.29(c)].

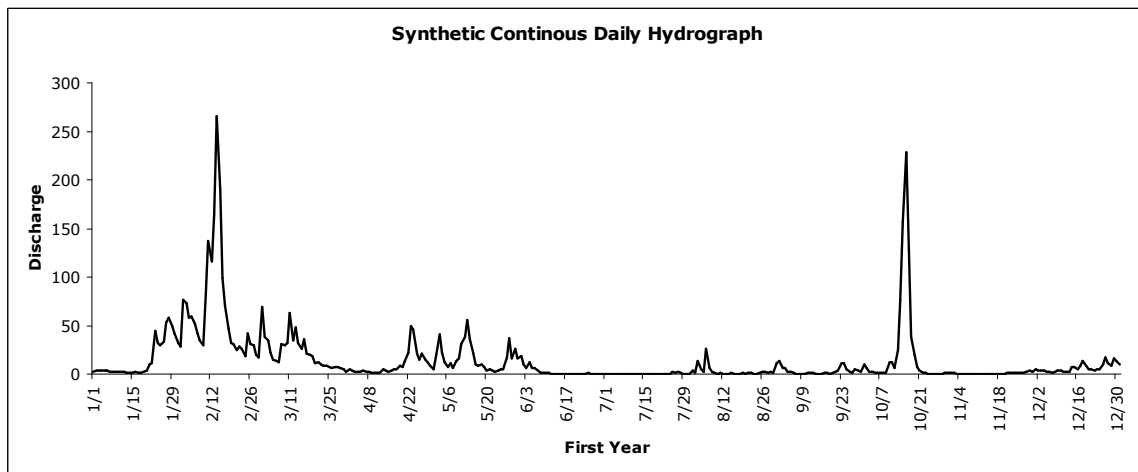


Figure 4.29(a) Synthetic Continuous Daily Hydrograph for Beaverdam Branch at Matthews (First Year)

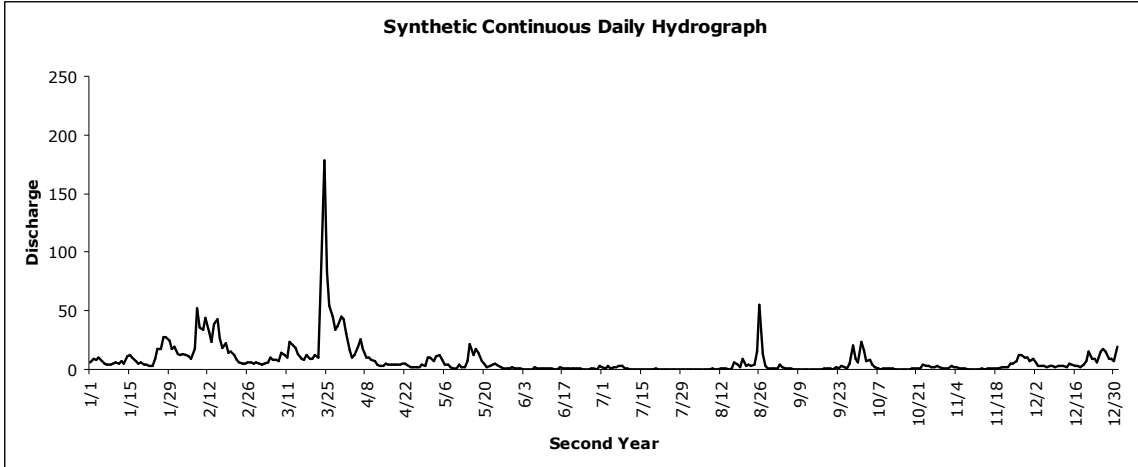


Figure 4.29(b) Synthetic Continuous Daily Hydrograph for Beaverdam Branch at Matthews (Second Year)

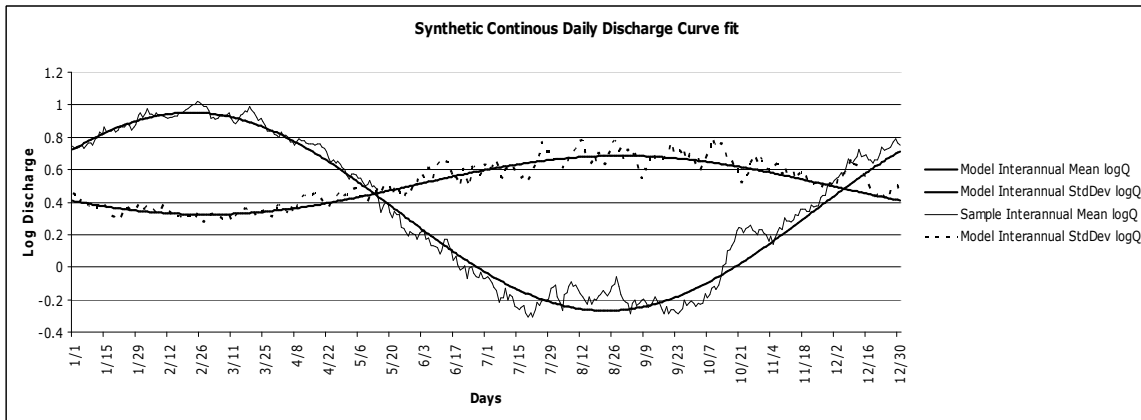


Figure 4.29(c) Continuous Daily Hydrograph from Beaverdam Branch at Matthews. Comparison of synthetic data (average of 80 years) with cosine model derived from real data.

## CHAPTER 5 CONCLUSION AND RECOMMENDATION

### 5.1 OVERVIEW

The goal of this project was to develop a continuous daily hydrograph model that captures the statistics of the gage record and can be used to generate long time series of realistic daily discharge. This chapter summarizes the research accomplishments, discusses lessons learnt, and suggests topics for future research.

### 5.2 ACCOMPLISHMENT OF RESEARCH OBJECTIVES

Chapter 1 outlined four specific research objectives. The research results are summarized here in terms of those objectives.

#### 5.2.1 COLLECT STREAM DATA

Discharge data were obtained from the USGS for 78 streams in Maryland. Each data set contained at least 30 years of record, and the streams represented the variety of watersheds in Maryland.

#### 5.2.2 CREATE A MODEL

This study developed a Markov model that uses cosine curves to represent deterministic variation in the mean, standard deviation, and lag-one correlation of log discharge, and a random noise term to represent the effects of variable precipitation (without explicitly modeling precipitation or requiring precipitation inputs). The model equations are:

$$\log Q(d, y) = M[\log Q(d)] + Z(d, y) * StDev[\log Q(d)] \quad (5.1)$$

$$M[\log Q(d)] = \bar{y}_M + y_{a,M} \cos \left[ \frac{2\pi}{366} (d - \tau_M) \right] \quad (5.2)$$

$$StDev[\log Q(d)] = \bar{y}_S + y_{a,S} \cos \left[ \frac{2\pi}{366} (d - \tau_S) \right] \quad (5.3)$$

$$Z(d,y) = \rho_Z(d)Z(d-1,y) + \sqrt{1 - \rho_Z^2(d)}K(d,y) \quad (5.4)$$

$$\rho_Z(d) = \bar{y}_R + y_{a,R} \cos \left[ \frac{2\pi}{366} (d - \tau_R) \right] \quad (5.5)$$

where  $K(d,y)$  in Eq. (5.4) is a random variate with mean 0, standard deviation 1, and skew  $g$ .

The model requires 10 parameters: three parameters that determine the cosine model describing the annual variation of the mean ( $y_M, y_{a,M}$  and  $\tau_M$ ); three that determine the cosine model for annual variation of the standard deviation ( $y_S, y_{a,S}$  and  $\tau_S$ ); three that determine the cosine model for annual variation of the lag-one correlation ( $y_M, y_{a,M}$  and  $\tau_M$ ), and the skew ( $g$ ) of the random variate.

The parameters of the model have been estimated for 78 streams in Maryland. Chapter 4 presents and discusses similarities and differences in the model parameters among sites.

### 5.2.3 ANALYSE THE RANDOM COMPONENT OF THE MODEL

The Pearson III distribution was found to be suitable to represent the random part of the model. The parameters of the Pearson III distribution (shape, scale, and shift) are dictated by the requirement for zero mean, unit standard deviation, and skew  $g$ .

### 5.2.4 DEMONSTRATE THE MODEL'S CAPABILITY

Examples with several study gages show that the 10-parameter Markov model can generate realistic synthetic hydrographs. The skewed random noise model generates

hydrograph rises of suitable frequency and steepness, and the lag-one persistence establishes appropriate event recessions.

### **5.2.5 SOFTWARE DEVELOPMENT**

The project has developed an automated tool that can minimize the amount of time in processing the large data and can store it in a form that can be accessed for future applications. The software program is written in MATLAB. The program used in this project is not restricted by the amount of data (provided that at least 30 years of record are available), and can deal with missing data within the record. It can be applied to areas other than Maryland. The output is stored in spreadsheet format.

## **5.3 DISCUSSION**

### **5.3.1 INPUT TO THE MODEL**

Sample size is one of the most important factors in a data analysis project. A large sample size is generally more reliable. This is also observed during the project. Only gages with a record of at least 30 years were considered in this project. Within the 30-year record, however, there were sometimes missing data. Missing data resulted in small sample sizes for some days at some sites. Some of these small samples were associated with negative sample skew in the analysis of random noise. Because the model assumed positive population skew in the random noise, these particular results were treated as sampling error and adjusted to conform to model assumptions.

One lesson learned from this project is the consideration of the Leap Year. The analysis used a 366-day year, with Day 60 (Feb. 29) treated as missing on non-Leap Years.

In order to deal with missing data “NaN” (Not a number) was used in MATLAB, and “NA()” in Excel. Mathematical operation with NaN or NA() was one of the challenges.

The other important factor is the data distribution. The Pearson and normal distributions both provide a good fit to the random noise, when examined graphically for selected watersheds. The Vogel test was performed to make a rational selection. Numerical results for all days in all watersheds demonstrated that the Pearson distribution nearly always fit the data as well as or at least slightly better than the normal. The Pearson distribution was applied for all the watersheds. When the skew is small, the Pearson distribution approaches the normal distribution, making the Pearson a more generally applicable model.

### **5.3.2 OUTPUT FROM THE MODEL**

Each watershed has different characteristics, as evident from the graphs developed for mean, standard deviation etc. in Chapter 4. Some distinctive examples are included in Chapter 3 and 4. It is understood that there are various physical factors that control the annual cycle and variability of the discharge, and the present study has not sought to understand those factors.

Seasonal variation is very well reflected from the interannual mean and standard deviation graphs. Peak average flows are observed from in late winter to early spring, when both infiltration and evapotranspiration are low. Average flow reaches a minimum in late summer, when watershed storage is depleted by evapotranspiration. Interannual variability is highest in late summer to early Fall, because that is the season experiencing both low and high extremes of streamflow (extended droughts in some years, and flooding due to tropical storms in others).



It was interesting to see the effects of control structures and diversions on the model parameters. The construction of dams controls the peak by storing the water. So even if the peak upstream is high, the downstream peak is not necessarily large. Regulated watersheds should not be used in analyzing the relationship between model parameters and physical characteristics of the watersheds. Therefore, care must be taken in the next stages of model development to check for artificial flow controls.

#### **5.4 FUTURE RESEARCH**

The project output is the generation of continuous daily hydrograph to extend the record at gaged sites, which is the first step of synthesis. This project is being extended to generate synthetic hydrographs for ungaged sites; a separate research effort is underway to identify the relationship of model parameters to physical characteristics of the watersheds. A set of rules (such as regression equations) could be used to estimate the model parameters, and thus generate realistic synthetic streamflow, for ungaged basins. (Brubaker et al. (2004) present a preliminary attempt at such regression equations, based on analysis by L. Guy at the University of Maryland.)

Human intervention is one of the major causes in the alteration of the discharge statistics. It can be in various forms, such as dams, canals, urbanization, deforestation, etc. Direct runoff and infiltration are affected mostly when there is any change in the area. Urbanization and deforestation cause more runoff and higher peak flows. When there is heavily vegetated land contributing to a stream, more infiltration reduces the peak in the stream, and flow is contributed more slowly later. Therefore, the land use in a watershed is expected to be an important factor in predicting the parameters of this model for ungaged watersheds.

Topographical relief is another deciding factor of the peak in stream. Steep-sloped areas have high stream flow speed. As the slope lessens the velocity also reduces. High-relief areas contribute fast to the downstream peak as compared to low relief areas. Therefore, topography (particularly slope) is expected to be an important factor in estimating the model parameters for ungaged basins.

This project did not try to identify or remove long-term trends in the data, such as the effects of land use change or climate change. This should be a focus of future work, because the consideration of change is important in the analysis and synthesis stages. Changing conditions in the gaged watersheds will need to be considered when doing regression between the parameters and physical characteristics (such as forested versus impervious area). If reliable relationships could be established, the model might be used to investigate the potential hydrologic effects of alternative development scenarios.

The software program written can be improved by adding some more features to create a stand-alone program, i.e. the user simply provides the file path and the program will run on its own without further inputs. Currently, the code reads files that are listed by the user. Features such as asking the user for the file path and then automatically creating an output file structure, or making a platform-independent executable file, would make this program more effective, efficient, and useful.

## APPENDIX A

APPENDIX  A  Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01578310 SUSQUEHANNA RIVER AT CONOWINGO, MD	27,100	4.38	0.37	73	0.37	0.09	306	0.87	0.06	79	0.46
USGS 01578500 OCTORARO CREEK NEAR RISING SUN, MD	193	2.26	0.2	82	0.28	0.03	256	0.9	0.01	213	1.67
USGS 01580000 DEER CREEK AT ROCKS, MD	94.4	1.98	0.18	82	0.25	0.02	284	0.9	0.02	168	2.11
USGS 01493000 UNICORN BRANCH NEAR MILLINGTON, MD	19.68	1.25	0.24	70	0.29	0	252	0.93	0.01	250	1.71
USGS 01493500 MORGAN CREEK NEAR KENNEDYVILLE, MD	N/A	0.84	0.15	57	0.29	0.02	209	0.85	0.02	201	1.74
USGS 01495000 BIG ELK CREEK AT ELK MILLS, MD	52.6	1.68	0.22	73	0.28	0.03	265	0.85	0.03	205	2.01
USGS 01496000 NORTHEAST CREEK AT LESLIE, MD	24.3	1.28	0.3	64	0.35	0.02	311	0.84	0.04	247	1.69
USGS 01589300 GWYNNNS FALLS AT VILLA NOVA, MD	32.5	1.38	0.21	70	0.32	0.04	265	0.82	0.02	186	1.7
USGS 01581500 BYNUM RUN AT BEL AIR, MD	8.52	0.7	0.33	64	0.38	0.07	252	0.82	0.02	234	1.72

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01581700 WINTERS RUN NEAR BENSON, MD	35	1.57	0.19	72	0.29	0.05	252	0.88	0.03	198	1.76
USGS 01582000 LITTLE FALLS AT BLUE MOUNT, MD	52.9	1.72	0.18	82	0.25	0.02	259	0.9	0.02	150	1.96
USGS 01583000 SLADE RUN NEAR GLYNDON, MD	2.09	0.25	0.19	82	0.28	0.05	255	0.9	0.02	198	1.59
USGS 01583500 WESTERN RUN AT WESTERN RUN, MD	59.8	1.72	0.19	77	0.26	0.03	256	0.91	0.02	186	1.89
USGS 01584500 LITTLE GUNPOWDER FALLS AT LAUREL BROOK, MD	36.1	1.51	0.18	76	0.26	0.04	244	0.88	0.02	206	1.85
USGS 01585300 STEMMERS RUN AT ROSSVILLE, MD	4.46	0.32	0.31	66	0.5	0.08	247	0.74	0.02	45	1.7
USGS 01585500 CRANBERRY BRANCH NEAR WESTMINSTER, MD	3.29	0.31	0.21	82	0.38	0.03	292	0.89	0.03	180	1.31
USGS 01586000 NORTH BRANCH PATAPSCO RIVER AT CEDARHURST, MD	56.6	1.66	0.22	78	0.28	0.03	265	0.89	0.02	140	1.91
USGS 01587000 NORTH BRANCH PATAPSCO RIV NEAR MARRIOTTVILLE,	165	1.78	0.31	82	0.86	0.08	294	0.97	0	183	1.32

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01587500 SOUTH BRANCH PATAPSCO RIVER AT HENRYTON, MD	64.4	1.71	0.26	76	0.29	0.05	261	0.9	0.01	193	1.67
USGS 01589000 PATAPSCO RIVER AT HOLLOFIELD, MD	285	2.07	0.25	79	0.35	0.03	242	0.93	0.01	158	1.74
USGS 01589100 EAST BRANCH HERBERT RUN AT ARBUTUS, MD	2.47	0.22	0.15	78	0.38	0.03	265	0.7	0.02	39	1.57
USGS 01589330 DEAD RUN AT FRANKLINTOWN, MD	5.52	0.39	0.25	72	0.5	0.05	284	0.72	0.03	16	1.57
USGS 01589440 JONES FALLS AT SORRENTO, MD	25.2	1.34	0.2	69	0.29	0.05	252	0.88	0.02	201	1.62
USGS 01589500 SAWMILL CREEK AT GLEN BURNIE, MD	4.97	0.46	0.17	81	0.56	0.19	246	0.95	0.01	211	1.08
USGS 01585100 WHITEMARSH RUN AT WHITE MARSH, MD	7.61	0.68	0.25	60	0.44	0.05	248	0.77	0	141	1.8
USGS 01590000 NORTH RIVER NEAR ANNAPOLIS, MD	8.5	0.91	0.16	62	0.23	0.04	230	0.88	0.01	124	1.68
USGS 01590500 BACON RIDGE BRANCH AT CHESTERFIELD, MD	6.92	0.82	0.18	61	0.27	0.07	225	0.86	0.02	163	1.34

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01491000 CHOPTANK RIVER NEAR GREENSBORO, MD	113	1.84	0.44	65	0.38	0.07	258	0.96	0.01	265	1.75
USGS 01492000 BEAVERDAM BRANCH AT MATTHEWS, MD	5.85	0.34	0.61	53	0.51	0.18	243	0.88	0.03	264	1.67
USGS 01591000 PATUXENT RIVER NEAR UNITY, MD	34.8	1.4	0.3	74	0.31	0.06	251	0.91	0.02	197	1.93
USGS 01592500 PATUXENT RIV NEAR LAUREL, MD	132	1.58	0.24	82	0.46	0.08	74	0.95	0.01	355	1.4
USGS 01593500 LITTLE PATUXENT RIVER AT GUILFORD, MD	38	1.43	0.26	69	0.32	0.05	259	0.85	0.02	227	1.94
USGS 01594000 LITTLE PATUXENT RIVER AT SAVAGE, MD	98.4	1.85	0.26	68	0.31	0.06	261	0.86	0.02	233	1.73
USGS 01594800 ST LEONARD CREEK NEAR ST LEONARD, MD	6.73	0.67	0.3	62	0.34	0.14	216	0.87	0.03	143	0.99
USGS 01486500 BEAVERDAM CREEK NEAR SALISBURY, MD	19.5	1.23	0.24	67	0.3	0.03	296	0.92	0.02	22	1.35
USGS 01490000 CHICAMACOMICO RIVER NEAR SALEM, MD	15	1.12	0.28	57	0.28	0.04	252	0.93	0.01	311	1.59

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01489000 FAULKNER BRANCH AT FEDERALSBURG, MD	7.1	0.73	0.37	64	0.34	0.08	247	0.93	0.01	317	2.02
USGS 01485000 POCOMOKE RIVER NEAR WILLARDS, MD	60.5	1.57	0.43	61	0.38	0.1	284	0.96	0.01	265	2.78
USGS 01485500 NASSAWANGO CREEK NEAR SNOW HILL, MD	44.9	1.36	0.53	56	0.44	0.12	251	0.97	0.01	301	2.12
USGS 01486000 MANOKIN BRANCH NEAR PRINCESS ANNE, MD	4.8	0.3	0.53	65	0.45	0.12	286	0.94	0.02	284	2.31
USGS 01595000 NORTH BRANCH POTOMAC RIVER AT STEYER, MD	73.1	1.98	0.43	64	0.37	0.07	245	0.93	0	249	1.98
USGS 01595500 NORTH BRANCH POTOMAC RIVER AT KITZMILLER, MD	225	2.36	0.49	68	0.39	0.07	265	0.93	0.01	296	1.96
USGS 01596500 SAVAGE RIVER NEAR BARTON, MD	49.1	1.46	0.62	68	0.45	0.09	293	0.95	0.01	322	2.43
USGS 01597000 CRABTREE CREEK NEAR SWANTON, MD	16.7	1.1	0.56	73	0.4	0.07	327	0.94	0.01	340	2.05
USGS 01597500 SAVAGE RIV BL SAVAGE RIV DAM NEAR BLOOMINGTON,	106	1.95	0.22	52	0.42	0.11	67	0.94	0	224	1.36

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01598000 SAVAGE RIVER AT BLOOMINGTON, MD	115	1.8	0.61	72	0.48	0.14	293	0.94	0.01	309	1.84
USGS 01599000 GEORGES CREEK AT FRANKLIN, MD	72.4	1.57	0.52	80	0.4	0.07	338	0.95	0.01	12	2.7
USGS 01601500 WILLS CREEK NEAR CUMBERLAND, MD	247	2.18	0.52	77	0.39	0.08	341	0.95	0.01	334	2.82
USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD	877	2.85	0.42	74	0.36	0.04	329	0.96	0.01	330	2.29
USGS 01609000 TOWN CREEK NEAR OLDTOWN, MD	148	1.77	0.51	74	0.49	0.09	297	0.94	0.01	325	1.95
USGS 01610000 POTOMAC RIVER AT PAW PAW, WV	3129	3.26	0.43	76	0.35	0.05	337	0.96	0.01	311	2.52
USGS 01610155 SIDELING HILL CREEK NEAR BELLEGROVE, MD	102	1.4	0.74	67	0.67	0.23	265	0.93	0.01	314	1.57
USGS 01613000 POTOMAC RIVER AT HANCOCK, MD	4090	3.34	0.43	75	0.36	0.04	337	0.96	0.01	294	2.6
USGS 01614500 CONOCOHEAGUE CREEK AT FAIRVIEW, MD	494	2.54	0.36	78	0.34	0.06	337	0.95	0.01	13	2.69



Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01617800 MARSH RUN AT GRIMES, MD	18.9	0.9	0.27	101	0.36	0.06	285	0.97	0.01	183	1.69
USGS 01619000 ANTIETAM CREEK NEAR WAYNESBORO, PA	93.5	1.95	0.22	82	0.26	0.05	351	0.93	0.01	62	1.49
USGS 01619500 ANTIETAM CREEK NEAR SHARPSBURG, MD	281	2.34	0.23	101	0.25	0.05	13	0.96	0.02	50	2.37
USGS 01637500 CATOCTIN CREEK NEAR MIDDLETOWN, MD	66.9	1.53	0.54	73	0.43	0.11	286	0.94	0	344	2.27
USGS 01638500 POTOMAC RIVER AT POINT OF ROCKS, MD	9651	3.75	0.37	80	0.33	0.04	346	0.97	0.01	284	2.45
USGS 01645000 SENECA CREEK AT DAWSONVILLE, MD	101	1.85	0.25	76	0.31	0.05	265	0.89	0.01	284	1.88
USGS 01646500 POTOMAC RIVER NEAR WASH, DC LITTLE FALLS PUMP STA	11560	3.8	0.4	77	0.36	0.06	315	0.97	0	303	2.31
USGS 01646502 POTOMAC RIVER (ADJUSTED) NEAR WASH, DC	11560	3.85	0.36	77	0.33	0.05	337	0.97	0	314	2.39
USGS 01646550 LITTLE FALLS BRANCH NEAR BETHESDA, MD	4.1	0.13	0.28	79	0.47	0.12	257	0.72	0.02	284	1.38

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01639000 MONOCACY RIVER AT BRIDGEPORT, MD	173	1.79	0.59	61	0.53	0.1	284	0.91	0.02	265	2.29
USGS 01639500 BIG PIPE CREEK AT BRUCEVILLE, MD	102	1.84	0.31	69	0.33	0.03	284	0.9	0.01	174	1.95
USGS 01640500 OWENS CREEK AT LANTZ, MD	5.93	0.66	0.51	78	0.41	0.14	294	0.93	0	339	2.1
USGS 01641000 HUNTING CREEK AT JIMTOWN, MD	18.4	1.14	0.45	76	0.36	0.06	312	0.93	0	23	2.01
USGS 01641500 FISHING CREEK NEAR LEWISTOWN, MD	7.29	0.85	0.44	82	0.32	0.12	333	0.95	0.01	36	1.98
USGS 01642000 MONOCACY RIVER NEAR FREDERICK, MD	665	2.64	0.39	70	0.41	0.05	291	0.9	0.01	290	1.97
USGS 01642500 LINGANORE CREEK NEAR FREDERICK, MD	82.3	1.73	0.31	78	0.32	0.05	301	0.9	0.01	21	1.81
USGS 01643000 MONOCACY RIVER AT JUG BRIDGE NEAR FREDERICK, MD	817	2.68	0.39	70	0.38	0.06	313	0.93	0.01	287	2.16
USGS 01643500 BENNETT CREEK AT PARK MILLS, MD	62.8	1.62	0.33	73	0.33	0.05	265	0.91	0.01	202	1.97

Name of the Site	Drainage area mi <sup>2</sup>	Interannual Mean			Interannual Standard Deviation			Lag One Correlation			Decorrelated Data Average Skew
		Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	Mean	Amplitude	Day of the Maximum (days)	
USGS 01649500 NORTH EAST BRANCH ANACOSTIA RIVER AT RIVERDALE, MD	72.8	1.68	0.27	64	0.37	0.05	240	0.85	0	199	2.09
USGS 01650500 NW BRANCH ANACOSTIA RIVER NEAR COLESVILLE, MD	21.1	1.12	0.25	69	0.35	0.08	252	0.85	0.03	254	1.89
USGS 01651000 NW BRANCH ANACOSTIA RIVER NEAR HYATTSVILLE, MD	49.4	1.39	0.29	73	0.4	0.08	252	0.85	0.03	254	1.78
USGS 01653600 PISCATAWAY CREEK AT PISCATAWAY, MD	39.5	1.24	0.55	59	0.55	0.28	242	0.82	0	339	1.47
USGS 01661050 ST CLEMENT CREEK NEAR CLEMENTS, MD	18.5	0.96	0.47	58	0.46	0.22	232	0.89	0.03	232	1.59
USGS 01661500 ST MARYS RIVER AT GREAT MILLS, MD	24	1.08	0.37	57	0.39	0.06	234	0.9	0.01	211	2.28

## APPENDIX B

### Code

Code is in the same order in which it is to be run.

#### 1) Reading text file (Code.m)

The code reads data from user specified file and stores it into a 366xN matrix, where N is the number of years. This matrix is then used through out the program.

```
[site,sitenumber,year,day,disch,expon]=textread('C:\Documents and Settings\pallavi\Desktop\Project\02060003
Gunpowder-Patapsco\USGS01583000\discharge.txt','%s%s%d%s%f%s');
len = length(site);
len
yrno=[];
num_yrs = 0;
match =0;
for i = 1:len,
    match =0;
    if (num_yrs ~= 0)
        for j = 1:num_yrs,
            if (yrno(j) == year(i))
                match = 1;
            end
        end
        if match == 0
            num_yrs = num_yrs + 1;
            yrno(num_yrs) = year(i);
        end
    else
        num_yrs = num_yrs +1;
        yrno(num_yrs) = year(i);
    end
end
%yrno
%extracted year numbers into array yrno, num_yrs contains number of distinct years in the data
month = [0,31,60,91,121,152,182,213,244,274,305,335];
[site,sitenumber,ddate,discharge,expon]=textread('C:\Documents and
Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\discharge.txt','%s%s%s%f%s');
DM = zeros(366,num_yrs);
for i = 1: len,
    dt = char(ddate(i)),
    [y,m,d] = stread(dt,'%d%d%d','delimiter','-'),
    pos = 0,
    j =1,
    found = 0,
    while (j<=num_yrs & found~=1),
        if (yrno(j) ~= y)
            j = j+1;
        else
```

```

        pos = j,
        found = 1,
        DM(month(m)+d,pos) = discharge(i);
    end
end

end

2) Calculating Mean (Meantab.m)

```

This program computes mean

```

[p,q] = size(DM);
rowsum = 0;
countnum=0;
h= yrno;
s= 'Mean';
st='USGS 01583000 SLADE RUN NEAR GLYNDON, MD';
for m = 1:p,
    for n =1:q,
        if(DM(m,n)==0),
            DM(m,n)=NaN;
        end
    end
end
mp= [];
mp=log10(DM);      % transform data to log
rowmean = [];
for i=1:p,        % p is row index
    for j= 1:q,   % q is column index
        if(isnan(mp(i,j)) == 1),
            disp("");
        else
            rowsum=rowsum + mp(i,j);
            countnum = countnum+1;
        end
    end
    rowmean(i) = rowsum/countnum;
    countnum=0;
    rowsum=0;
end

```

3) Calculating Standard Deviation(Stdev.m)

This function computes standard deviation and stores the data into file name as Mean & Standard deviation.xls as well as save the plots of mean and standard deviation.

```

[p,q] = size(DM);
s='Standard Deviation';
for m = 1:p,
    for n =1:q,
        if(DM(m,n)==0),
            DM(m,n)=NaN;
        end
    end
end

```

```

        end
    end
end
stdv = [];
ssqr=0;
ssum=0;
snum=0;
stdev=[];
stdv= log10(DM); % log transform data
for i = 1:p, % p is row index
    for j = 1:q, % q is column index
        if (isnan(stdv(i,j)) ==1), % traps nonnumeric values
            disp("");
        else
            ssqr= ssqr + (stdv(i,j)*stdv(i,j));
            ssum= ssum+stdv(i,j);
            snum=snum+1;
        end
    end
    stdev(i)=sqrt((ssqr-(ssum*ssum)/snum)/(snum-1)); %unbiased estimate of Std. Dev.
    ssqr=0;
    ssum=0;
    snum=0;
end
fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Mean&Standev.xls','wt');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%s',st);
fprintf(fid,'\n');
fprintf(fid,'%s','Days');
fprintf(fid,'\t');
fprintf(fid,'%s','Mean');
fprintf(fid,'\t');
fprintf(fid,'%s',s);
fprintf(fid,'\t');
fprintf(fid,'%d \t',yrmo(1,:));
fprintf(fid,'\n');
for i = 1:p,
    fprintf(fid,'%d \t',i);
    fprintf(fid, '%d \t', rowmean(i));
    fprintf(fid, '%d\t', stdev(i));
    fprintf(fid, '%d \t', log10(DM(i,:)));
    fprintf(fid,'\n');
end
fclose(fid);
figure(2);
plot(stdev);
title('Standard Deviation Plot :USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('Standard Deviation. ');
saveas(figure(2),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Standard Deviation','fig');

```

#### 4) Lag one correlation Calculation (zdatamatrixtab.m)

Lag one correlation is computed in this function. Output is then stored as zproduct.xls. Data are plotted and saved in the specified location.

```
[p,q] = size(DM);
for m = 1:p,
    for n = 1:q,
        if(DM(m,n)==0),
            DM(m,n)=NaN;
        end
    end
end
zdatamtrx = [];
zdatamtrx = log10(DM); % transform to log
for i=1:p, % p is row index, q is col index
    for j= 1:q,
        if(isnan(zdatamtrx(i,j)) == 1),
            disp("");
        else
            zdatamtrx(i,j)= (zdatamtrx(i,j)-rowmean(i))/stdev(i);
        end
    end
end

end
rowsum2=0;
countnum2=0;
stdevzdata=ones(366,1);
meanzdata=zeros(366,1);

zproduct =[];
for m =2 :59
    for n = 1 :q,
        zproduct(m,n)=zdatamtrx(m,n)*zdatamtrx(m-1,n);
    end
end
for m =60
    for n = 1 :q,
        if (isnan(zdatamtrx(60,n)) ==1),
            disp("");
        else
            zproduct(m,n)=zdatamtrx(m,n)*zdatamtrx(m-1,n);
        end
    end
end
for m =61
    for n = 1 :q,
        if (isnan(zdatamtrx(60,n)) ==1),
            zproduct(m,n)=zdatamtrx(m,n)*zdatamtrx(59,n);
        else
            zproduct(m,n)=zdatamtrx(m,n)*zdatamtrx(60,n);
        end
    end
end
```

```

        end
    end
end
for m =62:p
    for n = 1 :q,
        zproduct(m,n)=zdatamtrx(m,n)*zdatamtrx(m-1,n);
    end
end

m =1 ;
for n =2:q,
    zproduct(m,n)= zdatamtrx(m,n)*zdatamtrx(366,n-1);
end

zdm=[];
zdm =zproduct;
rowsum1=0;
countnum1=0;

for i=1:p,
    for j= 1:q,
        if(isnan(zdm(i,j)) == 1),
            disp("");
        else
            rowsum1=rowsum1 + zdm(i,j);
            countnum1 = countnum1+1;
        end
    end
    Rmean(i) = sqrt(rowsum1/countnum1);
    countnum1=0;
    rowsum1=0;
end
Rmean=Rmean';

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\zproduct.xls','wt');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%s',st);
fprintf(fid,'\n');
fprintf(fid,'%s','Days');
fprintf(fid,'\t');
fprintf(fid,'%s','ZdataMean');
fprintf(fid,'\t');
fprintf(fid,'%d \t',yrmo(1,:));
fprintf(fid,'\n');
for i = 1:p,
    fprintf(fid,'%d \t',i);
    fprintf(fid, '%d \t', (Rmean(i)));
    fprintf(fid, '%d \t', (zproduct(i,:)));
    fprintf(fid,'\t');

    fprintf(fid,'\n');
end

```



```

fclose(fid);
figure(3);
plot(zdatamtrx);
title('Zdatamatrix Plot:USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('ZData');
saveas(figure(3),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapasco\USGS01583000\Zdatamatrix','fig');

```

## 5) Curve fitting the data (Curvefit.m)

After Lag one correlation, curve fit is done. Data are stored in curve fit.xls and plot is stored in the desired location.

```

xdata=[1:366];
r= 'Mean';
s= 'Standard Deviation';
re='Lag One Correlation';
res = 'Residual';
ybar='Ybar';
ya='Ya';
phe='Phi';
ydata= rowmean;
F = inline('x(1) + x(2)*cos((2*pi/366)*(xdata-x(3)))','x','xdata');
[x ,resnorm]= lsqcurvefit(F,[0,0,0], xdata,ydata)

if (x(1,3)) <0, % Correction for day of maximum from negative to positive.
    disp('x(1,3) is less than zero');
    n = fix((-x(1,3))/(366));
    x(1,3)=(x(1,3)+(366*(n+1)));
end
if(x(1,2))<0,% Correction for amplitude from negative to positive.
    x(1,2)=-1*(x(1,2));
    if(x(1,3)<186)
        x(1,3)=x(1,3)+186;
    else
        x(1,3)=x(1,3)-186;
    end
end

ydata1= stdev;

F = inline('x1(1) + x1(2)*cos((2*pi/366)*(xdata-x1(3)))','x1','xdata');
[x1 ,resnorm1]= lsqcurvefit(F,[0,0,0], xdata,ydata1)

if (x1(1,3)) <0,% Correction for day of maximum from negative to positive.
    disp('x1(1,3) is less than zero');
    n1 = fix((-x1(1,3))/(366));
    x1(1,3)=(x1(1,3)+(366*(n1+1)));
    x1(1,3)
end
if(x1(1,2))<0,% Correction for amplitude from negative to positive.

```

```

x1(1,2)=-1*(x1(1,2);
if(x1(1,3)<186)
    x1(1,3)=x1(1,3)+186;
else
    x1(1,3)=x1(1,3)-186;
end
end

ydata2= Rmean';

F = inline('x2(1) + x2(2)*cos((2*pi/366)*(xdata-x2(3)))','x2','xdata');
[x2 ,resnorm2]= lsqcurvefit(F,[0,0,0], xdata,ydata2)

if (x2(1,3)) <0,% Correction for day of maximum from negative to positive.
    disp('x2(1,3) is less than zero');
    n2 = fix((-x2(1,3))/(366));
    x2(1,3)=(x2(1,3)+(366*(n2+1)))
end
if(x2(1,2))<0,% Correction for amplitude from negative to positive.
    x2(1,2)=-1*(x2(1,2);
    if(x2(1,3)<186)
        x2(1,3)=x2(1,3)+186;
    else
        x2(1,3)=x2(1,3)-186;
    end
end

meanfit = x(1)+ x(2)*cos(((2*pi/366)*(xdata- x(3))));
stdevfit= x1(1)+ x1(2)*cos(((2*pi/366)*(xdata- x1(3))));
Rmeanfit= x2(1)+ x2(2)*cos(((2*pi/366)*(xdata- x2(3))));

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\curvefit.xls','wt');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%s',st);
fprintf(fid,'\n');
fprintf(fid,'\t');
fprintf(fid,'%s',ybar);
fprintf(fid,'\t');
fprintf(fid,'%s',ya);
fprintf(fid,'\t');
fprintf(fid,'%s',phe);
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%s',res);
fprintf(fid,'\n');
fprintf(fid,'%s',r);
fprintf(fid,'\t');
fprintf(fid,'%d',x(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x(2));
fprintf(fid,'\t');

```

```

fprintf(fid,'%d',x(3));
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%d',resnorm);
fprintf(fid,'\n');
fprintf(fid,'%s',s);
fprintf(fid,'\t');
fprintf(fid,'%d',x1(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x1(2));
fprintf(fid,'\t');
fprintf(fid,'%d',x1(3));
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%d',resnorm1);
fprintf(fid,'\n');
fprintf(fid,'%s',re);
fprintf(fid,'\t');
fprintf(fid,'%d',x2(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x2(2));
fprintf(fid,'\t');
fprintf(fid,'%d',x2(3));
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%d',resnorm2);
fclose(fid);

figure(4);
plot(meanfit,'r');
hold on;
plot(rowmean,'c');
hold on;
plot(stdevfit,'b');
hold on;
plot(stdev,'m');
hold on;
plot(Rmean,'g');
hold on;
plot(Rmeanfit,'k');
hold off;
legend('Mean Fit','Mean','Standard Deviation Fit','Standard Deviation','Lag One Correlation','Lag One
Correlation Fit');
title('Curve Fit Plot:USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('Mean,Standard Deviation,Regression');
saveas(figure(4),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Curve Fit','fig');

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\final_output.xls','w');
fprintf(fid,'\t');

```

```

fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%s','x1');
fprintf(fid,'\t');
fprintf(fid,'%s','x2');
fprintf(fid,'\t');
fprintf(fid,'%s','x3');
fprintf(fid,'\t');
fprintf(fid,'%s','x4');
fprintf(fid,'\t');
fprintf(fid,'%s','x5');
fprintf(fid,'\t');
fprintf(fid,'%s','x6');
fprintf(fid,'\t');
fprintf(fid,'%s','x7');
fprintf(fid,'\t');
fprintf(fid,'%s','x8');
fprintf(fid,'\t');
fprintf(fid,'%s','x9');
fprintf(fid,'\t');
fprintf(fid,'%s',ms);
fprintf(fid,'\n');
fprintf(fid,'%s',st);
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'%d',x(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x(2));
fprintf(fid,'\t');
fprintf(fid,'%d',x(3));
fprintf(fid,'\t');
fprintf(fid,'%d',x1(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x1(2));
fprintf(fid,'\t');
fprintf(fid,'%d',x1(3));
fprintf(fid,'\t');
fprintf(fid,'%d',x2(1));
fprintf(fid,'\t');
fprintf(fid,'%d',x2(2));
fprintf(fid,'\t');
fprintf(fid,'%d',x2(3));
fprintf(fid,'\t');
fprintf(fid,'\n');

fclose(fid);

```

## 6) Decorrelation of data (Décor.m)

Decorrelation is runs after curvefit.m function. Data are stored with a name as décor.xls and graph is saved as decorrelated.m

```

options = optimset('largescale','on');

[p,q] = size(DM);
for m = 1:p,
    for n = 1:q,
        if(DM(m,n)==0),
            DM(m,n)=NaN;
        end
    end
end

ztoday=[];
zyest=[];
Kcorr=[];
[xx,yy] = size(Rmean);

for m = 2 :p,          % m is row index, n is column)
    for n = 1 :q,
        Kcorr(m,n)=(zdatamtrx(m,n)-zdatamtrx(m-1,n)*Rmean(m))/(sqrt(1-Rmean(m)*Rmean(m)));
    end
end
m = 1 ;
for n = 2:q,
    Kcorr(m,n)= (zdatamtrx(m,n)-zdatamtrx(366,n-1)*Rmean(m))/(sqrt(1-Rmean(m)*Rmean(m)));
end
Krowsum = 0;
Kcountnum=0;
Krowmean = [];
for i=1:p,          % p is row index
    for j= 1:q,      % q is column index
        if(isnan(Kcorr(i,j)) == 1),
            disp("");
        else
            Krowsum=Krowsum + Kcorr(i,j);
            Kcountnum = Kcountnum+1;
        end
    end
    Krowmean(i) = Krowsum/Kcountnum;
    Kcountnum=0;
    Krowsum=0;
end
Kssqr=0;
Kssum=0;
Ksnum=0;
Kstdev=[];
for i = 1:p,      % p is row index
    for j = 1:q, % q is column index
        if (isnan(Kcorr(i,j)) ==1), % traps nonnumeric values
            disp("");
        else
            Kssqr= Kssqr + (Kcorr(i,j)*Kcorr(i,j));
            Kssum= Kssum+Kcorr(i,j);
            Ksnum= Ksnum+1;
        end
    end
end

```

```

        end
    end
    Kstdev(i)=sqrt((Kssqr-(Kssum*Kssum)/Ksnum)/(Ksnum-1)); %unbiased estimate of Std. Dev.
    Kssqr=0;
    Kssum=0;
    Ksnum=0;
end

```

```

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\De_correl.xls','wt');
fprintf(fid,'\t');
fprintf(fid,'%s ',st);
fprintf(fid,'\n');
fprintf(fid,'%d \t',yrmo(1,:));
fprintf(fid,'\t');
fprintf(fid,'%s \t','Krowmean');
fprintf(fid,'%s \t','Kstdev');
fprintf(fid,'\n');
for i = 1:p,
fprintf(fid, '%d \t ', (Kcorr(i,:)));
fprintf(fid, '\t');
fprintf(fid,'%d',Krowmean(i));
fprintf(fid, '\t');
fprintf(fid,'%d',Kstdev(i));
fprintf(fid,'\n');
end
fclose(fid);
figure(5);
plot(Kcorr);
title('Decorrelated Data Plot:USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('Decorrelated');
saveas(figure(5),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Decorrelated','fig');

```

## 7) Skew computation (Pearson.m)

Skew is calculated in the Pearson .m function and negative skew row's data are plotted as

histogram.

```

options = optimset('largescale','on');
ms='Mean Skew';
[p,q] = size(DM);
for m = 1:p,
    for n =1:q,
        if(DM(m,n)==0),
            DM(m,n)=NaN;
        end
    end
end
end

Kcountnum=0;
sttsum=0;

```

```

stt=[];
skewg=[];
Kalpha=[];
Kbeta=[];
Kshift=[];
    % transform data to natural log
for i=1:p,      % p is row index
    for j= 1:q, % q is column index
        if(isnan(Kcorr(i,j)) == 1),
            disp("");
        else
            Kcountnum = Kcountnum+1;
            stt(i,j)= (Kcorr(i,j)-Krowmean(i))^3;
            sttsum=stt(i,j)+sttsum;
        end
    end
skewg(i)=(Kcountnum*sttsum)/((Kcountnum-1)*(Kcountnum-2)*(Kstdev(i)^3));
if skewg(i)<0,
    figure(i+2);

    i
    hist(Kcorr(i,:));
    xlabel(i);
    title(strcat('Negative Skew Plot ',num2str(i),' :USGS 01583000 SLADE RUN NEAR GLYNDON, MD'));
    xlabel('Days');
    ylabel('Decorrelated Day with negative skew');
    saveas(figure(i+2),strcat('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\NegativeSkew',num2str(i)), 'fig');
    skewg(i)=0.16;
end
Kcountnum=0;
sttsum=0;
end
for m=1:p,
    Kalpha(m)=(2/skewg(m))*(2/skewg(m)); %alpha
    Kbeta(m)=sqrt(1/Kalpha(m)); %beta
    Kshift(m)=-1*Kalpha(m)*Kbeta(m); %shift
end

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\final_output.xls','a');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');

fprintf(fid,'\t');

fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');
fprintf(fid,'\t');

```

```
fprintf(fid,'%d',mean(skewg));
fprintf(fid,'\n');
fclose(fid);
```

## 8) Vogel Parameters Calculation (Vogel.m)

In Vogel.m normal and Pearson distribution is checked. Their values are stored for comparison in vogel.xls file.

```
[p,q] = size(Kcorr);
meanxnorm=[];
meanxpearson=[];
```

```
for a=1:p,
    sortKrow=sort(Kcorr(a,:));%sorting of row
    sortK=[];
```

```
    count=q;
    sumxnorm=0;
    sumxpearson=0;
    ssqrsum=0;
```

```
    for j= 1:q, %counting Nans
        if(isnan(sortKrow(j)) == 1),
            count=count-1;
        end
    end
```

```
    for Kr=1:count,
        probrow(Kr)=Kr/(count+1);
```

```
        xnorm(Kr)=norminv((1-probrow(Kr)),0,1);
        xpearson(Kr)=gaminv((1-probrow(Kr)),Kalpha(a),Kbeta(a))+Kshift(a);
        sortK(count-Kr+1)=sortKrow(Kr); %since sort arranges in ascending order thats why this step is
```

done to match the data with the rank.

```
        sortdata(a,Kr)=(sortK(Kr)-Krowmean(a));
        ssqrsort(a,Kr)=sortdata(a,Kr)^2;
        ssqrsum=ssqrsum+ssqrsort(a,Kr);
```

```
        sumxnorm=sumxnorm+(xnorm(Kr));
        sumxpearson=sumxpearson+xpearson(Kr);
```

```
    end
    qrsum(a)=ssqrsum;
    meanxnorm(a)=sumxnorm/count;
    meanxpearson(a)=sumxpearson/count;
```

```
npcount=0;
nmsqrsum=0;
pmsqrsum=0;
```



```

nprodsum=0;
pprodsum=0;

for Kr=1:count,
    nmoment(a,Kr)=(xnorm(Kr)-meanxnorm(a)); %vogel formula-(w-wbar)
    nmsqr(a,Kr)=nmoment(a,Kr)^2;          %(w-wbar)^2
    nmsqrsursum=nmsqr(a,Kr)+nmsqrsursum;    %Sum of (moment)

    pmoment(a,Kr)=(xpearson(Kr)-meanxpearson(a));
    pmsqr(a,Kr)=pmoment(a,Kr)^2;
    pmsqrsursum=pmsqr(a,Kr)+pmsqrsursum;

    nprod(a,Kr)=nmoment(a,Kr)*sortdata(a,Kr);
    nprodsum=nprodsum+nprod(a,Kr);

    pprod(a,Kr)=pmoment(a,Kr)*sortdata(a,Kr);
    pprodsum=pprodsum+pprod(a,Kr);
end

vogeln(a)=nprodsum/(sqrt(nmsqrsursum*qrsum(a)));
vogelp(a)=pprodsum/(sqrt(pmsqrsursum*qrsum(a)));

% figure(a+500);
% plot(sortK(1:Kr),xnorm(1:Kr),'o');
% hold on;
% plot(sortK(1:Kr),xpearson(1:Kr),'ro');
%hold on;
%plot(sortK(1:Kr),sortK(1:Kr),'k');
%title(strcat('Skew Plot ',num2str(a),' :plot'));
%xlabel('Decorrelated Data');
%ylabel('Standardized Normal Variate');

%hold off

a % to see whether program is running fine or not and also to hold patience for 366 days computation.
end

fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Vogel.xls','wt');
fprintf(fid,'\t \t \t');
fprintf(fid,'%s',st);
fprintf(fid,'\n');
fprintf(fid,'%s \t','Days');
fprintf(fid,'%s \t','Skew');
fprintf(fid,'%s \t','Mean');
fprintf(fid,'%s \t','StdDev');
fprintf(fid,'%s \t','Vogel-Normal');
fprintf(fid,'%s \t','Vogel-Pearson');
fprintf(fid,'%d \t',yrno(1,:));
fprintf(fid,'\n');
for i = 1:p,
    fprintf(fid,'%d \t',i);          %days
    fprintf(fid, '%d \t ',skewg(i)); %skew
    fprintf(fid,'%d',Krowmean(i)); %rowmean=0
    fprintf(fid,'\t');

```

```

fprintf(fid,'%d \t',Kstdev(i));           %standard deviation=1
fprintf(fid,'%d \t',vogeln(i));
fprintf(fid,'%d \t',vogelp(i));
fprintf(fid,'%d \t',Kcorr(i,:));
fprintf(fid,'\n');
end
fclose(fid);
a
disp('finish')
figure(2);
plot(skewg);
title('Skew Plot: USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('Skew(K)');
saveas(figure(2),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Skew','fig');

```

## 9) Generation of Synthetic Hydrograph (Markov.m)

This is the last function of the program. This function generates the continuous daily hydrograph and save both the output and figure as Discharge.

```

p = 366; % # of days in year
q = 80; % # of years for simulation (doesn't have to match data)

mskew = mean(skewg);
Kalpha = (2/mskew)*(2/mskew); % same expressions you had before
Kbeta =sqrt(1/Kalpha);
Kshift =-1*Kalpha*Kbeta;

w = 0;

% starting random number drawn from Person distribution

Zlast = gaminv((1-rand),Kalpha,Kbeta) + Kshift;

for n=1:q % n is year index
    for m = 1:p % m is Day of Year index
        w = w+1; % w is the running time-series index
        xpear = gaminv((1-rand),Kalpha,Kbeta)+Kshift; % new random noise
        Zfinal = Zlast*Rmeanfit(1,m)+sqrt(1-(Rmeanfit(1,m))^2)*xpear;
        logQ(w) = meanfit(1,m) + stdevfit(1,m)*Zfinal;
        QF(w) = 10^logQ(w);
        Q(m,n) = QF(w);
        Zlast = Zfinal;
    end
end
fid = fopen('C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Discharge.xls','wt');
fprintf(fid,'\t \t \t');
fprintf(fid,'%s',st);
fprintf(fid,'\n');

```

```

fprintf(fid,'%s \t','Days');
fprintf(fid,'%d \t',yrmo(1,:));
fprintf(fid,'\n');
for i = 1:p,
fprintf(fid,'%d \t',i);           %days
fprintf(fid,'%d \t ',skewg(i));   %skew
fprintf(fid,'%d',Krowmean(i));    %rowmean=0
fprintf(fid,'\t');
fprintf(fid,'%d \t',Kstdev(i));    %standard deviation=1
fprintf(fid,'%d \t',vogeln(i));
fprintf(fid,'%d \t',vogelp(i));
fprintf(fid,'%d \t',((Q(i))));
fprintf(fid,'\t');
fprintf(fid,'\n');
end

fclose(fid);
figure(6);
plot(QF);
title('Continuous Daily Hydrograph :USGS 01583000 SLADE RUN NEAR GLYNDON, MD');
xlabel('Days');
ylabel('Discharge');
saveas(figure(6),'C:\Documents and Settings\pallavi\Desktop\Project\02060003 Gunpowder-
Patapsco\USGS01583000\Discharge','fig');

```

## APPENDIX C

Sample from data of 366 days x 73 Years is taken as an example.

a) Mean & Standard Deviation Table

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD											
Days	Mean	Standard Deviation	1929	1930	1931	1932	1933	1934	1935	1936	1937
1	2.99E+00	4.34E-01	NaN	3.17	2.11	2.71	3.11	3.08	3.16	2.72	3.47
2	2.99E+00	4.55E-01	NaN	3.20	1.98	2.93	3.08	3.52	3.18	2.73	3.43
3	3.01E+00	4.18E-01	NaN	3.27	2.01	3.03	3.00	3.37	3.10	2.98	3.51
4	3.03E+00	3.96E-01	NaN	3.28	2.05	2.99	3.00	3.23	3.06	3.46	3.57
5	3.01E+00	3.89E-01	NaN	3.20	2.17	2.90	3.00	3.43	2.90	3.30	3.43
6	3.01E+00	3.75E-01	NaN	3.14	2.61	3.11	3.00	3.65	2.90	3.22	3.32
7	3.02E+00	3.95E-01	NaN	3.10	2.73	3.65	2.95	4.04	2.96	3.20	3.38
8	3.02E+00	3.93E-01	NaN	3.09	2.63	3.44	2.90	3.99	2.99	3.36	3.62
9	3.00E+00	3.80E-01	NaN	3.05	2.47	3.31	2.88	3.69	3.50	3.51	3.68
10	3.02E+00	4.14E-01	NaN	3.03	2.50	3.23	2.90	3.51	3.65	3.84	3.90
11	3.00E+00	3.97E-01	NaN	2.98	2.43	3.11	2.88	3.38	3.52	3.59	4.07
12	2.99E+00	3.88E-01	NaN	2.95	2.34	3.02	2.89	3.27	3.38	3.46	3.78
13	2.96E+00	3.78E-01	NaN	2.97	2.30	2.96	2.96	3.20	3.27	3.40	3.59
14	2.95E+00	3.63E-01	NaN	2.97	2.19	2.95	2.90	3.17	3.26	3.47	3.46
15	2.97E+00	3.79E-01	NaN	2.95	2.06	2.88	2.89	3.10	3.20	3.42	3.39
16	2.98E+00	3.85E-01	NaN	2.92	1.95	2.86	2.83	3.03	3.15	3.44	3.42
17	2.96E+00	3.73E-01	NaN	2.86	2.03	2.85	2.82	2.95	3.18	3.39	3.34
18	2.95E+00	3.87E-01	NaN	2.85	2.00	2.81	2.82	2.85	3.47	3.33	3.68
19	2.96E+00	4.28E-01	NaN	2.66	2.11	2.78	2.82	2.86	3.33	3.47	3.66
20	2.97E+00	3.99E-01	NaN	2.60	2.16	2.74	2.83	2.83	3.28	3.40	3.59
21	2.97E+00	3.94E-01	NaN	2.48	2.24	2.70	2.79	2.78	3.62	3.34	3.81
22	2.99E+00	4.13E-01	NaN	2.48	2.37	2.67	2.87	2.76	4.03	3.27	4.11
23	3.02E+00	3.90E-01	NaN	2.48	2.35	2.77	3.04	2.78	3.79	3.11	4.18
24	3.03E+00	3.93E-01	NaN	2.48	2.38	3.22	3.01	3.05	3.57	3.04	3.96
25	3.03E+00	3.98E-01	NaN	2.48	2.29	3.28	2.99	2.94	3.42	3.11	3.99
26	3.05E+00	3.99E-01	NaN	2.48	2.43	3.16	3.52	2.93	3.40	3.08	3.89
27	3.07E+00	3.91E-01	NaN	2.48	3.31	3.12	3.56	2.99	3.30	3.00	3.69
28	3.07E+00	4.00E-01	NaN	2.48	3.34	3.10	3.48	2.94	3.19	2.85	3.54
29	3.05E+00	3.87E-01	NaN	2.48	3.14	3.05	3.36	3.01	3.18	2.88	3.45
30	3.06E+00	3.74E-01	NaN	2.48	2.95	3.40	3.25	2.82	3.12	2.90	3.38
31	3.06E+00	3.69E-01	NaN	2.48	2.83	3.51	3.19	2.78	3.03	2.93	3.34
32	3.05E+00	3.59E-01	NaN	2.49	2.78	3.32	3.17	2.90	2.97	2.85	3.36
33	3.05E+00	3.52E-01	NaN	2.54	2.67	3.22	3.25	2.93	3.00	2.81	3.27
34	3.04E+00	3.76E-01	NaN	2.60	2.56	3.19	3.38	2.83	3.16	2.84	3.17
35	3.05E+00	3.83E-01	NaN	2.70	2.49	3.62	3.30	2.79	3.15	2.83	3.13
36	3.06E+00	3.84E-01	NaN	3.00	2.51	4.05	3.22	2.78	3.04	2.83	3.11
37	3.04E+00	3.68E-01	NaN	3.00	2.36	3.70	3.11	2.74	3.00	2.85	3.05
38	3.06E+00	3.67E-01	NaN	3.00	2.40	3.51	3.11	2.64	2.87	2.87	3.05
39	3.07E+00	3.59E-01	NaN	3.00	2.41	3.42	3.27	2.63	2.88	2.79	3.56

b) Lag One Correlation

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD										
Days	ZdataMean	1929	1930	1931	1932	1933	1934	1935	1936	1937
1	9.49E-01	0.00	0.23	3.68	0.45	0.11	-0.04	0.18	0.44	1.18
2	9.63E-01	NaN	0.19	4.51	0.09	0.06	0.25	0.16	0.36	1.07
3	9.64E-01	NaN	0.28	5.31	-0.01	-0.01	0.99	0.08	0.04	1.13
4	9.64E-01	NaN	0.39	5.91	0.00	0.00	0.43	0.01	-0.08	1.63
5	9.61E-01	NaN	0.30	5.36	0.03	0.00	0.54	-0.02	0.79	1.46
6	9.59E-01	NaN	0.16	2.27	-0.08	0.00	1.82	0.08	0.41	0.87
7	9.48E-01	NaN	0.07	0.77	0.44	0.00	4.41	0.04	0.26	0.75
8	9.75E-01	NaN	0.04	0.72	1.71	0.05	6.41	0.01	0.40	1.40
9	9.66E-01	NaN	0.02	1.36	0.88	0.09	4.46	-0.11	1.16	2.73
10	9.48E-01	NaN	0.00	1.74	0.42	0.09	2.14	2.01	2.66	3.82
11	9.75E-01	NaN	0.00	1.79	0.15	0.08	1.12	1.99	2.95	5.76
12	9.74E-01	NaN	0.01	2.38	0.02	0.08	0.69	1.31	1.78	5.45
13	9.78E-01	NaN	0.00	2.92	0.00	0.00	0.47	0.82	1.40	3.39
14	9.83E-01	NaN	0.00	3.67	0.00	0.00	0.40	0.69	1.66	2.34
15	9.58E-01	NaN	0.00	4.98	0.00	0.03	0.21	0.53	1.70	1.56
16	9.63E-01	NaN	0.01	6.37	0.07	0.08	0.05	0.27	1.41	1.27
17	9.75E-01	NaN	0.04	6.65	0.10	0.14	0.00	0.25	1.35	1.15
18	9.68E-01	NaN	0.06	6.07	0.11	0.13	0.01	0.78	1.11	1.90
19	9.48E-01	NaN	0.17	4.87	0.15	0.12	0.06	1.16	1.15	3.08
20	9.59E-01	NaN	0.64	4.05	0.25	0.12	0.08	0.67	1.27	2.56
21	9.69E-01	NaN	1.16	3.79	0.39	0.16	0.17	1.28	0.99	3.30
22	9.70E-01	NaN	1.58	2.83	0.53	0.14	0.28	4.10	0.61	5.70
23	9.54E-01	NaN	1.73	2.58	0.48	-0.02	0.35	4.95	0.17	8.08
24	9.51E-01	NaN	1.94	2.84	-0.30	0.00	-0.03	2.75	0.01	7.09
25	9.70E-01	NaN	1.97	3.12	0.30	0.00	-0.01	1.34	0.01	5.70
26	9.51E-01	NaN	2.00	2.89	0.17	-0.12	0.07	0.85	0.02	5.08
27	9.48E-01	NaN	2.17	-0.93	0.03	1.49	0.06	0.52	-0.01	3.35
28	9.59E-01	NaN	2.23	0.41	0.01	1.29	0.07	0.18	0.10	1.89
29	9.78E-01	NaN	2.17	0.16	0.00	0.82	0.03	0.10	0.25	1.23
30	9.12E-01	NaN	2.32	-0.07	0.01	0.39	0.07	0.05	0.19	0.87
31	9.61E-01	NaN	2.49	0.19	1.09	0.17	0.50	-0.01	0.15	0.64
32	9.72E-01	NaN	2.45	0.48	0.92	0.12	0.32	0.02	0.20	0.65
33	9.64E-01	NaN	2.22	0.82	0.36	0.19	0.14	0.03	0.39	0.55
34	9.64E-01	NaN	1.69	1.40	0.18	0.49	0.19	-0.04	0.38	0.21
35	9.68E-01	NaN	1.05	1.87	0.56	0.58	0.38	0.08	0.31	0.07
36	9.60E-01	NaN	0.13	2.08	3.87	0.28	0.48	-0.01	0.34	0.03
37	9.76E-01	NaN	0.02	2.66	4.64	0.07	0.59	0.01	0.30	0.00
38	9.60E-01	NaN	0.02	3.34	2.15	0.02	0.96	0.06	0.27	0.00
39	9.63E-01	NaN	0.03	3.30	1.18	0.07	1.43	0.27	0.41	-0.07
40	9.70E-01	NaN	0.02	1.61	0.96	0.49	1.81	0.15	0.62	2.81
41	9.51E-01	NaN	0.02	-0.87	0.58	0.23	2.51	-0.21	0.64	3.44
42	9.37E-01	NaN	0.07	0.24	0.27	0.04	3.11	0.54	0.82	2.05

c) Curve Fitting

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD				
	Ybar	Ya	Phi	Residual
Mean	2.85	0.42	74.08	1.33
Standard Deviation	0.36	0.04	328.76	0.30
Regression	0.96	0.01	330.02	0.19

d) Decorrelated Data

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD											
Days	Krowmean	Kstdev	1929	1930	1931	1932	1933	1934	1935	1936	1937
1	0.00	1.19	0.00	-0.33	-0.98	0.12	-0.24	1.25	-0.10	0.12	0.39
2	0.00	1.27	NaN	0.20	-0.98	1.73	-0.33	3.56	0.07	0.12	-0.45
3	0.00	1.26	NaN	0.66	-0.95	0.68	-0.81	-1.04	-0.68	1.77	0.99
4	0.00	1.26	NaN	0.17	-0.61	-0.52	-0.17	-1.17	-0.50	4.34	0.88
5	0.00	1.27	NaN	-0.49	0.73	-0.71	0.13	2.11	-1.28	-1.08	-0.93
6	0.00	1.28	NaN	-0.39	3.66	1.93	0.05	2.40	-0.01	-0.52	-0.70
7	0.00	1.30	NaN	-0.36	0.82	4.19	-0.46	3.08	0.38	-0.20	0.42
8	0.00	1.20	NaN	-0.14	-1.18	-2.12	-0.57	-0.21	0.28	1.86	2.90
9	0.00	1.26	NaN	-0.16	-1.73	-0.87	-0.17	-2.26	5.34	1.91	1.14
10	0.00	1.30	NaN	-0.28	0.24	-0.79	0.12	-1.65	0.89	2.30	1.45
11	0.00	1.20	NaN	-0.36	-0.97	-1.01	-0.18	-0.94	-0.86	-2.09	2.67
12	0.00	1.21	NaN	-0.30	-1.20	-0.86	0.13	-0.87	-1.18	-1.06	-2.61
13	0.00	1.17	NaN	0.67	-0.61	-0.32	1.30	-0.31	-0.79	-0.08	-1.50
14	0.00	1.10	NaN	0.20	-1.95	0.01	-0.83	-0.11	0.26	1.54	-1.28
15	0.00	1.28	NaN	-0.44	-1.35	-0.85	-0.25	-0.88	-0.67	-0.63	-0.80
16	0.00	1.27	NaN	-0.40	-1.37	-0.36	-0.68	-0.70	-0.63	0.16	0.23
17	0.00	1.20	NaN	-0.46	0.50	0.03	-0.03	-0.72	0.69	-0.08	-0.44
18	0.00	1.25	NaN	0.06	-0.11	-0.23	0.10	-0.91	3.19	-0.50	3.59
19	0.00	1.30	NaN	-1.49	0.99	-0.31	-0.06	0.04	-1.38	0.79	-0.46
20	0.00	1.28	NaN	-0.86	-0.38	-0.54	-0.05	-0.39	-0.13	-0.17	-0.04
21	0.00	1.24	NaN	-1.52	0.37	-0.57	-0.58	-0.71	3.56	-0.51	2.42
22	0.00	1.23	NaN	-0.09	1.26	-0.41	0.70	-0.34	3.76	-0.94	2.68
23	0.00	1.29	NaN	-0.64	-0.89	0.37	1.14	-0.23	-1.37	-1.29	1.40
24	0.00	1.29	NaN	-0.30	-0.12	3.49	-0.33	2.02	-1.60	-0.67	-1.56
25	0.00	1.24	NaN	-0.15	-1.08	0.66	-0.25	-1.10	-1.56	0.70	0.44
26	0.00	1.30	NaN	-0.33	0.79	-1.06	4.17	-0.29	-0.13	-0.37	-0.57
27	0.00	1.30	NaN	-0.50	6.44	-0.40	0.35	0.26	-0.76	-0.80	-1.29
28	0.00	1.28	NaN	-0.05	0.38	-0.10	-0.56	-0.42	-0.90	-1.33	-1.16
29	0.00	1.17	NaN	-0.19	-2.12	-0.39	-1.04	1.00	0.12	0.43	-0.64
30	0.00	1.32	NaN	-0.54	-1.25	2.16	-0.58	-1.35	-0.38	-0.05	-0.23
31	0.00	1.27	NaN	-0.28	-1.22	1.25	-0.45	-0.52	-0.81	0.20	-0.22
32	0.00	1.22	NaN	-0.03	-0.61	-1.76	0.04	1.42	-0.56	-0.92	0.59
33	0.00	1.26	NaN	0.21	-1.34	-0.96	0.86	0.26	0.30	-0.54	-0.78
34	0.00	1.26	NaN	0.77	-0.93	-0.30	1.32	-0.96	1.63	0.42	-1.06
35	0.00	1.24	NaN	0.99	-0.79	4.49	-0.83	-0.44	-0.16	-0.15	-0.43
36	0.00	1.27	NaN	2.57	-0.15	4.21	-0.73	-0.30	-1.06	-0.16	-0.20
37	0.00	1.19	NaN	0.12	-2.10	-3.51	-1.09	-0.54	-0.32	0.29	-0.52
38	0.00	1.28	NaN	-0.23	-0.09	-1.80	-0.15	-1.34	-1.53	-0.11	-0.28
39	0.00	1.27	NaN	-0.07	-0.34	-0.66	1.65	-0.40	0.04	-1.00	5.22
40	0.00	1.23	NaN	0.23	3.69	0.12	1.27	-1.16	0.79	-0.21	3.03
41	0.00	1.30	NaN	-0.18	5.90	-1.12	-1.80	-0.99	3.23	-0.10	-0.93
42	0.00	1.31	NaN	-0.61	-1.95	-0.27	-0.22	-0.63	0.24	-0.82	-0.96

e) Vogel Parameters

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD													
Days	Skew	Mean	StdDev	Vogel-Normal	Vogel-Pearson	1929	1930	1931	1932	1933	1934	1935	1936
1	3.20	0.00	1.19	0.70	0.78	0.00	-0.33	-0.98	0.12	-0.24	1.25	-0.10	0.12
2	3.09	0.00	1.27	0.70	0.78	NaN	0.20	-0.98	1.73	-0.33	3.56	0.07	0.12
3	2.63	0.00	1.26	0.68	0.74	NaN	0.66	-0.95	0.68	-0.81	-1.04	-0.68	1.77
4	1.93	0.00	1.26	0.70	0.73	NaN	0.17	-0.61	-0.52	-0.17	-1.17	-0.50	4.34
5	2.82	0.00	1.27	0.69	0.76	NaN	-0.49	0.73	-0.71	0.13	2.11	-1.28	-1.08
6	3.20	0.00	1.28	0.69	0.77	NaN	-0.39	3.66	1.93	0.05	2.40	-0.01	-0.52
7	3.17	0.00	1.30	0.70	0.78	NaN	-0.36	0.82	4.19	-0.46	3.08	0.38	-0.20
8	2.67	0.00	1.20	0.70	0.76	NaN	-0.14	-1.18	-2.12	-0.57	-0.21	0.28	1.86
9	2.17	0.00	1.26	0.70	0.74	NaN	-0.16	-1.73	-0.87	-0.17	-2.26	5.34	1.91
10	2.82	0.00	1.30	0.70	0.76	NaN	-0.28	0.24	-0.79	0.12	-1.65	0.89	2.30
11	1.82	0.00	1.20	0.70	0.72	NaN	-0.36	-0.97	-1.01	-0.18	-0.94	-0.86	-2.09
12	3.45	0.00	1.21	0.69	0.78	NaN	-0.30	-1.20	-0.86	0.13	-0.87	-1.18	-1.06
13	0.16	0.00	1.17	0.63	0.63	NaN	0.67	-0.61	-0.32	1.30	-0.31	-0.79	-0.08
14	1.50	0.00	1.10	0.71	0.72	NaN	0.20	-1.95	0.01	-0.83	-0.11	0.26	1.54
15	1.75	0.00	1.28	0.69	0.72	NaN	-0.44	-1.35	-0.85	-0.25	-0.88	-0.67	-0.63
16	2.30	0.00	1.27	0.70	0.74	NaN	-0.40	-1.37	-0.36	-0.68	-0.70	-0.63	0.16
17	0.64	0.00	1.20	0.64	0.64	NaN	-0.46	0.50	0.03	-0.03	-0.72	0.69	-0.08
18	1.48	0.00	1.25	0.70	0.72	NaN	0.06	-0.11	-0.23	0.10	-0.91	3.19	-0.50
19	5.67	0.00	1.30	0.70	0.91	NaN	-1.49	0.99	-0.31	-0.06	0.04	-1.38	0.79
20	2.38	0.00	1.28	0.69	0.74	NaN	-0.86	-0.38	-0.54	-0.05	-0.39	-0.13	-0.17
21	1.73	0.00	1.24	0.70	0.72	NaN	-1.52	0.37	-0.57	-0.58	-0.71	3.56	-0.51
22	1.99	0.00	1.23	0.70	0.74	NaN	-0.09	1.26	-0.41	0.70	-0.34	3.76	-0.94
23	2.12	0.00	1.29	0.70	0.74	NaN	-0.64	-0.89	0.37	1.14	-0.23	-1.37	-1.29
24	2.18	0.00	1.29	0.70	0.74	NaN	-0.30	-0.12	3.49	-0.33	2.02	-1.60	-0.67
25	2.04	0.00	1.24	0.70	0.74	NaN	-0.15	-1.08	0.66	-0.25	-1.10	-1.56	0.70
26	2.48	0.00	1.30	0.70	0.75	NaN	-0.33	0.79	-1.06	4.17	-0.29	-0.13	-0.37
27	2.62	0.00	1.30	0.68	0.73	NaN	-0.50	6.44	-0.40	0.35	0.26	-0.76	-0.80
28	3.34	0.00	1.28	0.70	0.79	NaN	-0.05	0.38	-0.10	-0.56	-0.42	-0.90	-1.33
29	0.80	0.00	1.17	0.70	0.70	NaN	-0.19	-2.12	-0.39	-1.04	1.00	0.12	0.43
30	3.17	0.00	1.32	0.69	0.77	NaN	-0.54	-1.25	2.16	-0.58	-1.35	-0.38	-0.05
31	1.78	0.00	1.27	0.70	0.73	NaN	-0.28	-1.22	1.25	-0.45	-0.52	-0.81	0.20
32	1.98	0.00	1.22	0.70	0.74	NaN	-0.03	-0.61	-1.76	0.04	1.42	-0.56	-0.92
33	2.36	0.00	1.26	0.70	0.75	NaN	0.21	-1.34	-0.96	0.86	0.26	0.30	-0.54
34	1.56	0.00	1.26	0.70	0.72	NaN	0.77	-0.93	-0.30	1.32	-0.96	1.63	0.42
35	1.81	0.00	1.24	0.70	0.73	NaN	0.99	-0.79	4.49	-0.83	-0.44	-0.16	-0.15
36	1.36	0.00	1.27	0.67	0.69	NaN	2.57	-0.15	4.21	-0.73	-0.30	-1.06	-0.16
37	0.65	0.00	1.19	0.69	0.70	NaN	0.12	-2.10	-3.51	-1.09	-0.54	-0.32	0.29
38	2.23	0.00	1.28	0.70	0.74	NaN	-0.23	-0.09	-1.80	-0.15	-1.34	-1.53	-0.11
39	2.82	0.00	1.27	0.69	0.76	NaN	-0.07	-0.34	-0.66	1.65	-0.40	0.04	-1.00
40	1.83	0.00	1.23	0.70	0.72	NaN	0.23	3.69	0.12	1.27	-1.16	0.79	-0.21
41	2.54	0.00	1.30	0.68	0.74	NaN	-0.18	5.90	-1.12	-1.80	-0.99	3.23	-0.10
42	2.54	0.00	1.31	0.69	0.74	NaN	-0.61	-1.95	-0.27	-0.22	-0.63	0.24	-0.82



f) Markov output.

USGS 01603000 NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND, MD								
Days	1	2	3	4	5	6	7	8
1	498.0	616.2	7158.4	1274.0	270.5	1295.0	1265.7	498.0
2	455.8	523.3	5604.5	1051.0	233.9	1077.7	1482.0	455.8
3	440.8	512.0	4296.1	1332.4	209.6	2354.3	1193.8	440.8
4	376.0	429.6	3439.4	1602.5	230.5	1856.3	1241.8	376.0
5	332.4	563.7	3251.2	2447.0	209.2	1504.5	1168.1	332.4
6	331.1	478.5	3189.6	1974.8	214.5	1205.3	970.5	331.1
7	770.5	418.6	2830.1	2497.1	205.1	1321.7	1026.0	770.5
8	653.1	416.3	2459.5	1990.4	296.9	1229.2	884.6	653.1
9	637.9	383.7	2014.2	1793.5	325.0	1263.1	1352.7	637.9
10	578.9	327.3	1658.6	1432.7	302.6	2043.5	1141.2	578.9
11	500.9	303.0	1529.0	1891.2	263.2	2115.0	1076.6	500.9
12	425.0	354.0	1268.0	3155.2	228.1	1791.0	881.6	425.0
13	378.1	376.2	1190.8	3003.1	528.8	1434.9	2134.0	378.1
14	478.3	468.1	1384.1	2808.4	745.4	2545.9	3435.3	478.3
15	485.5	454.5	1235.7	2652.6	743.8	2060.0	2802.0	485.5
16	412.6	461.7	1075.5	2483.6	1070.8	2283.2	2380.0	412.6
17	354.3	455.2	939.0	2396.1	1139.6	1804.9	1893.0	354.3
18	363.9	421.7	776.0	1913.3	992.4	1936.1	1760.7	363.9
19	315.9	676.8	669.8	1634.9	814.9	2137.5	1602.7	315.9
20	542.9	722.8	587.1	1317.3	813.9	1919.9	1532.4	542.9
21	569.0	658.2	502.4	1617.8	693.3	1707.3	1475.7	569.0
22	493.4	817.8	425.0	1362.5	926.1	1407.8	1448.0	493.4
23	1351.9	738.2	413.7	1354.3	781.3	1490.2	1195.5	1351.9
24	1730.9	836.9	383.7	1285.3	896.2	1627.7	1002.6	1730.9
25	1987.6	764.7	340.7	2324.7	1417.4	1794.4	832.8	1987.6
26	2281.1	726.1	344.7	2052.3	1450.4	1609.3	708.2	2281.1
27	1984.4	705.2	336.3	2008.8	1300.9	1826.5	602.8	1984.4
28	2103.1	662.1	433.7	1661.1	1676.9	1571.4	527.8	2103.1
29	2408.9	558.7	594.1	1384.9	1375.6	1630.6	529.4	2408.9
30	5459.8	592.1	702.3	1371.0	1413.6	1526.1	530.1	5459.8
31	4374.0	601.8	997.0	1505.9	1320.1	1408.5	626.0	4374.0
32	3970.6	652.9	1290.3	1526.9	1273.3	1218.9	1568.3	3970.6
33	3104.0	657.5	1229.4	1239.6	1393.3	1442.9	1552.4	3104.0
34	2831.4	724.1	1042.5	1170.1	1134.4	1551.3	1287.8	2831.4
35	2660.7	734.0	1129.9	1180.1	934.2	1266.6	1376.3	2660.7
36	2152.3	702.2	1291.3	1009.1	778.2	1424.4	1469.5	2152.3
37	1942.5	622.0	1066.9	1019.2	761.6	1863.9	1510.9	1942.5
38	2225.2	676.1	3064.9	852.8	804.2	1792.2	1417.8	2225.2
39	1899.9	947.5	2654.0	1040.5	2302.9	1564.3	1602.5	1899.9
40	1564.9	924.1	2356.9	1056.4	2756.7	1861.9	1302.2	1564.9
41	498.0	616.2	7158.4	1274.0	270.5	1295.0	1265.7	498.0
42	455.8	523.3	5604.5	1051.0	233.9	1077.7	1482.0	455.8

## REFERENCES

- Abramowitz, M. and Stegun, I.A., 1970. *Handbook of Mathematical Functions*, Dover Publications Inc., New York, pp 1046.
- Arnold, J.G., J.R. William, A.D. Nicks, and N.B. Sammons, 1990. *SWRRB: A Basin Scale Simulation Model for Soil and Water Resources Management*. College Station, TX, Texas A&M University Press, 142 pp.
- Briaud, J.L., F. Ting, H.C. Chen, R. Gudavalli, B.P. Kwak, S.W. Han, S. Perugu, G. Wei, P. Nurtjahyo, Y. Cao, and Y. Li, 1999. "SRICOS: Prediction of Scour Rate at Bridge Piers". Report 2937-1, Texas Department of Transportation.
- Briaud, J.L., F.C.K. Ting, H.C. Chen, S.W. Han, and K.W. Kwak, 2001 "Erosion Function Apparatus for Scour Rate Predictions," *Journal of Geotechnical and Geoenvironmental Engineering* 127:105-113.
- Brubaker, K.L., Ghelardi, V., Goodings, D., Guy, L., Pathak, P., Estimation Of Long-Term Scour At Maryland Bridges Using EFA/SRICOS, August 2004. Maryland State Highway Administration Research Report Project No. SP107B4E.
- Chapelle, S., 2003. "Maryland Rivers," from [www.gibbs-smith.com/textbooks/downloads/statemaps/MD/rivers.gif](http://www.gibbs-smith.com/textbooks/downloads/statemaps/MD/rivers.gif) (accessed May 2005 and used with permission). Gibbs Smith, Publisher, Salt Lake City, Utah, [www.gibbs-smith.com](http://www.gibbs-smith.com) .
- Fleming, M. and V. Neary, "Continuous hydrologic modeling study with the hydrologic modeling system," *Journal of Hydrologic Engineering*, June 2004, 9:175-183.
- Fuller, G. A., "Generation of ungaged streamflow data," *Journal of Hydraulics Division*, March 1978, 104:377-383.

- Gupta, R. S., 2001. *Hydrology and Hydraulics System*, Second edition, Waveland Press, Inc, Illinois, 867 pp.
- Hamlin, M. J., and Kottegoda, N. T., “Extending the record on the Teme,” *Journal of Hydrology*, December 1971, 12:100-116.
- James, W. and R.C. James, 2000. *Water Systems Models-A Guide to the Rain, Temperature and Runoff modules of the USEPA SWMM4*, Version 2.3, CHI Publications, Guelph, Ontario (Canada), 275 pp..
- McCuen, R.H., 1998. *Hydrologic Analysis and Design*, Second Edition, Prentice Hall, New Jersey, 814 pp.
- Moglen, G.E., and M.J. Casey, 2000 “GISHydro2000,” University of Maryland for Maryland State Highway Administration.
- National Weather Service (NWS), 2005. “National Weather Service River Forecast System (NWSRFS),” [http://hydrology.nws.noaa.gov/oh/hrl/hseb/ahps/info\\_nwsrfs](http://hydrology.nws.noaa.gov/oh/hrl/hseb/ahps/info_nwsrfs) (Accessed March, 2005).
- Prakash, A., 2004. *Water Resources Engineering*, ASCE Press, Reston, Virginia, 348 pp.
- Rodriguez-Iturbe, I., 1969. “Estimation of Statistical Parameters for Annual River flows,” *Water Resources Research* 5:1419-1421.
- Salas J.D., 1993, “Analysis and modeling of hydrologic time series” in *Handbook of Hydrology*, McGraw Hill, pp 19.1-19.62.
- Stedinger R.J., R.M. Vogel and E. Foufoula-Georgiou, 1993. “Frequency analysis of extreme events” in *Handbook of Hydrology*, McGraw-Hill, pp 18.1-18.27.

- U.S Army Corps of Engineers (USACE), 2002. "SSARR Program - Streamflow Synthesis and Reservoir Regulation, Corps of Engineers, North Pacific Water Management," <http://www.nwd-wc.usace.army.mil/report/ssarr.htm> (Accessed March, 2005).
- U.S. Environmental Protection Agency (USEPA), 2003. "EPA Exposure Assessment Models: HSPF," <http://www.epa.gov/ceampubl/swater/hspf/> (Accessed March, 2005).
- U.S. Geological Survey (USGS), 2005a. "Water Resources Application Software: Summary of PRMS," [http://water.usgs.gov/cgi-bin/man\\_wrdapp?prms](http://water.usgs.gov/cgi-bin/man_wrdapp?prms) (Accessed March, 2005).
- U.S. Geological Survey (USGS), 2005b. "Water Resources Applications Software: Summary of DR3M," [http://water.usgs.gov/cgi-bin/man\\_wrdapp?dr3m](http://water.usgs.gov/cgi-bin/man_wrdapp?dr3m) (Accessed March, 2005).
- U.S. Geological Survey (USGS), 2005c, "USGS Real-Time Water Data for USGS 01619500 Antietam Creek Near Sharpsburg, MD," [http://nwis.waterdata.usgs.gov/nwis/uv/?site\\_no=01619500&agency\\_cd=USGS](http://nwis.waterdata.usgs.gov/nwis/uv/?site_no=01619500&agency_cd=USGS) (Accessed March, 2005).
- Wurbs, R.A. and W.P. James , 2002. *Water Resources Engineering*, Prentice Hall, New Jersey, 828 pp.