ABSTRACT

Title of Thesis	THE EFFECT OF HYDROLOGIC MODEL AND DATA COMPLEXITY ON WATER QUANTITY AND QUALITY MODEL PREDICTION ACCURACY
	Kristin Leigh Gilroy, Master of Science, 2008
Thesis Directed by:	Richard H. McCuen Department of Civil and Environmental Engineering

Hydrologic modeling is central to the solution of many flooding and water quality issues. As the complexity of these issues increases, model complexity increases. The purpose of this research was to determine the effects of model and data complexity on hydrologic model prediction accuracy. A complex hydrologic model was developed and then simplified based on structural complexity and the change in accuracy was assessed. Analyses of data complexity were also conducted. The results showed that complex models containing excessive low sensitivity parameters did not significantly improve prediction accuracy. However, a lack of complete representation of the physical processes of the hydrologic cycle did affect prediction accuracy. Data analyses revealed that misalignments between rainfall and runoff gauges may cause poor prediction of peaks and grab samples may adequately represent the mean value but not the distribution of a population. Guidelines were developed to improve future development and application of hydrologic models.

THE EFFECT OF HYDROOGIC MODEL AND DATA COMPLEXITY ON WATER QUANTITY AND QUALITY PREDICTION ACCURACY

by

Kristin Leigh Gilroy

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
2008

Advisory Committee:

Dr. Richard H. McCuen, Chair
Dr. Kay Brubaker
Dr. Alan Davis

TABLE OF CONTENTS

CHAPTER 1 .... **1**

INTRODUCTION **1**

1.1 INTRODUCTION 1
1.1.1 Problem: Components of Model Development and Application *2*
1.1.2 Problem: Complexity and Model Development *3*
1.1.3: Problem: Complexity of Data *4*
1.1.4: Problem: Complexity and Model Constraints *4*
1.1.5 Problem: Complexity and Calibration Criteria Selection *5*
1.1.6 Problem: Complexity and Model Calibration Fitting Methods *6*
1.1.7 Summary of Issues in Model Complexity *7*
1.2 GOALS AND OBJECTIVES 7
1.3 IMPLICATIONS 8

CHAPTER 2 .... **9**

LITERATURE REVIEW **9**

2.1 INTRODUCTION 9
2.2 MODEL STRUCTURE 9
2.3 FITTING METHODS 11
2.4 MODEL CONSTRAINTS 12
2.5 DATA COMPLEXITY 13
2.6 HYDROLOGIC PROCESSES 14

CHAPTER 3 .... **17**

METHODOLOGY **17**

3.1 MODEL DEVELOPMENT 17
3.1.1 Model Components *17*
3.1.2 Rainfall *18*
3.1.3 Interception *19*
3.1.4 Surface Runoff *19*
3.1.5 Infiltration *20*
3.1.6 Evaporation and Transpiration *21*
3.1.7 Outflow *25*
3.2 STEPS IN THE MASS BALANCE PROCESS 27
3.3 SUBJECTIVE OPTIMIZATION CALIBRATION 29
3.3.1 Goodness-of-Fit and Model Assessment *30*
3.3.2 Peak Discharge Prediction Accuracy *31*
3.3.3 Low Flow Prediction Accuracy *32*
3.3.4 Water Balance Prediction Accuracy *32*
3.3.5 Graphical Analysis *34*
3.3.5 Optimization Guidelines *36*
3.4 MODEL COMPLEXITY REDUCTION PROCESS 38
3.5 GRAB SAMPLE ANALYSIS PROCEDURE 39
3.5.1 Analysis Procedure *40*
3.5.2 Analysis Procedure of Grab Samples Using Systematic Elimination *43*

CHAPTER 4 .... **44**

MODEL CALIBRATION **44**

4.1 MODEL 1 44
4.1.1 Prediction Accuracy of Model 1 .... *45*
4.1.2 Calibration Process of Model 1 .... *46*
4.1.3 Sensitivity Analysis of Model 1 .... *48*
4.2 MODEL 2 50

4.2.1 Calibration Process of Model 2 .... *51*
4.2.2 Prediction Accuracy of Model 2 .... *54*
4.3 MODEL 3........ 54
4.3.1 Calibration Process of Model 3 .... *55*
4.4 MODEL 4........ 59
4.4.1 Calibration Process of Model 4 .... *59*
4.5 MODEL 5........ 62
4.5.1 Calibration Process of Model 5 .... *62*
4.6 MODEL 6........ 64
4.6.1 Calibration Process of Model 6 .... *65*
4.7 MODEL 7........ 70
4.7.1 Calibration Process of Model 7 .... *71*
4.7.1.2 Calibration Run 7B *73*
4.7.2 SENSITIVITY ANALYSIS OF MODEL 7 75
4.8 MODEL 8........ 77
4.8.1 Calibration Process of Model 8 .... *78*
4.8.1.2 Calibration Run 8B *79*
4.9 FURTHER REDUCTIONS IN COMPLEXITY 80
4.10 SENSITIVITY TRACK 81
4.10.1 Model 9 .... *81*
4.10.2 Model 10 .... *82*
4.10.3 Model 11 .... *84*
4.11 SENSITIVITY VS. HYDROLOGIC PROCESSES........ 86
4.11.1 Model 12 .... *86*
4.12 SENSITIVITY PROCESSES........ 88
4.12.1 Model 13 .... *88*
4.12.3 Model 14 .... *91*
4.13 HYDROLOGIC PROCESSES 93
4.13.2 Model 16 .... *95*

CHAPTER 5 97

ANALYSES AND RESULTS........97

5.1 RESULTS OF MODEL STRUCTURE COMPLEXITY SIMPLIFICATIONS........ 97
5.2 EFFECT OF RAINFALL MISALIGNMENT 100
5.3 OPTIMIZATION FOR DESIGN CRITERIA 107
5.3.1 Calibration for Peak Discharge Prediction Accuracy *108*
5.3.2 Calibration for Base Flows Prediction Accuracy *111*
5.4 THE IMPORTANCE OF PARAMETRIC SENSITIVITY........ 115
5.4.1 High Parameter Sensitivity *116*
5.4.2 Low Parameter Sensitivity *117*
5.4.3 Discussion of Results *118*
5.5 HYDROLOGIC RATIONALITY OF PARAMETERS........ 119
5.5.1 Interception Zone *120*
5.5.2 Surface Zone *121*
5.5.3 Root Zone *122*
5.5.4 Vadose Zone *124*
5.6 CONFIRMATION 125
5.6.1 Runoff Bias *126*
5.6.2 Water Balance *126*
5.6.3 Relative Standard Error and Correlation Coefficient *127*
5.6.4 Relative Bias *128*
5.6.5 Confirmation Analysis *128*
5.7 ASSESSMENT OF DATA COMPLEXITY IN WATER QUALITY DATA 132
5.7.1 Results for Random Elimination of Samples *134*
5.7.4 Analysis of Data Complexity of Grab Samples *138*

CHAPTER 6 ...... **143**

GUIDELINES FOR EFFICIENT MODELING **143**

6.1 GUIDELINES ON DATA ANALYSIS 143
6.2 GUIDELINES ON CALIBRATION 147
6.3 GUIDELINES FOR ASSESSING GOODNESS OF FIT 151
6.4 GUIDELINES ON MODEL DEVELOPMENT 154

CHAPTER 7 ...... **159**

CONCLUSIONS AND RECOMMENDATIONS **159**

7.1 CONCLUSIONS 159
7.2 RECOMMENDATIONS 162

APPENDICES **164**

APPENDIX A: GRAPHICAL ANALYSES OF THE HYDROLOGIC MODEL 164
APPENDIX B: SENSITIVITY ANALYSES 168
APPENDIX C: MODEL SIMPLIFICATIONS 171
APPENDIX D: GRAB SAMPLE ANALYSES TABLES 172

REFERENCES **178**

CHAPTER 1

INTRODUCTION

1.1 Introduction

Computer modeling enables engineers and scientists to analyze complex systems and changes to complex systems in order to achieve specific goals and objectives. Hydrologic modeling is an important tool for designing, planning, and regulating development. For example, with the increase in development and the decrease in pervious land, stormwater management for both water quantity and quality control is a growing issue for engineers and hydrologists. Water quantity modeling is necessary to predict runoff values, such as the peak discharge or total runoff volume, in order to assess the needs for storm water management. Accurate predictions of these hydrologic metrics are necessary to decrease flooding and, therefore, decrease the risk of property damage. In addition to development, the recent flooding that caused massive damage in the Midwest reveals the need for more accurate water quantity modeling to reduce the risk of loss of life as a result of poor water management planning. Accuracy in water quantity modeling is essential to reduce the risk of life and property loss due to flooding.

Water quality is another issue of concern in hydrologic modeling. Good water quality is becoming difficult to maintain with the increase in pollutants being released through municipal, industrial, and agricultural activities. Upon the generation of pollutants, physical transport into the water environment occurs. Chemicals from these sources can leach through the subsurface layer and contaminate groundwater. Storm runoff transports the pollutants that remain on the surface to nearby lakes and streams. The integration of these pollutants into water bodies can cause many water quality problems, such as low dissolved oxygen levels, high bacterial levels, eutrophication, and high toxic chemical levels. These changes can cause disease transmission, imbalances in the ecosystem, and aesthetic changes, such as color and odor of water (Thomann and Mueller 1987). Accurate water quantity modeling is essential in predicting the transport of pollutants through the water system and is necessary to develop water quality regulations and decrease water pollutants.

The number and complexity of problems in which the hydrologic processes are fundamental elements have increased in the last decade. Reduction of peak discharges and the control of eroded soil are no longer the only issues that concern society. Public concern has been raised over the levels of pharmaceuticals and even the concentration of sun blocking chemicals in streams. These more complex problems require more complex models, at least in terms of the physical, biological, and chemical processes that the model must include as components. More complex models require greater varieties of input data, more detailed fitting methods, a better understanding of criteria used to judge prediction accuracy, and more knowledge of the model user.

1.1.1 Problem: Components of Model Development and Application

While the need for accurate models is great, model development is a complicated process. The modeling process involves five components: (1) formulation of the model structure; (2) assembling the calibration data; (3) identifying model constraints; (4) selecting calibration criteria; and (5) selecting the calibration fitting method. Each component contributes to the process of efficiently developing a model that provides the desired prediction accuracy while remaining within the available resources. Inexperience with any of the five components may hinder the efficiency and accuracy of model development and calibration.

1.1.2 Problem: Complexity and Model Development

Formulating the model structure involves selecting the physical processes to represent the system and the equations that best describes these processes. Complex model structures often include more processes and nonlinear equations when deemed necessary. As a result, complex model structures tend to include more parameters than simpler models. The selection of processes and equations depends on the resources and physical knowledge available. Models are limited by the scientific knowledge available to represent these physical processes. Likewise, a lack of adequate resources may require a simpler model that contains fewer parameters and simplifies the physical processes.

The trend in hydrologic engineering is toward increased model complexity. While models such as the Rational method are still widely used for small project design, more complex models such as SWMM and HSPF are being required for many design projects. The trend is motivated by the fact that more complex models are conceptually more rational and allow for greater variation in design conditions. However, increased complexity adds a burden on the model user facing the task of calibrating the model. The user must have a better understanding of both the fitting process and the interpretation and balancing of an array of goodness-of-fit criteria. A lack of experience and knowledge of complex model fitting can result in a model that lacks optimal goodness of fit and has parameter values that do not reflect the physical processes being modeled. If the trend for requiring more complex models outstrips the growth of knowledge of the modeling process, then inaccurate designs are likely.

1.1.3: Problem: Complexity of Data

Data selection is an important component because poor quality data will most likely result in poor prediction capabilities of the model. Data that are unrepresentative of the watershed being modeled may result in calibrated parameters that do not truly represent the physical processes for the watershed. Incomplete or short data records also make it difficult to accurately calibrate the model. Selection of calibration data must take into account these issues.

Complexity in the modeling process is not limited just to the length of the computer program and the number of model parameters that require calibration. Complex models require more complex data bases and with the growth of GIS systems and remotely sensed data, the fitting process has become more difficult. The model user must know how to deal with inconsistencies in data, as the inaccuracy of the results can depend as much on the lack of data quality as on the difficulty in optimizing the parameter values. Many users fail to review the data and then misunderstand the effect that poor data quality can have on the assumed optimum model.

1.1.4: Problem: Complexity and Model Constraints

Constraints also must be considered in model development. For example, calibrated parameter values should be constrained to be hydrologically rational. A calibrated model that yields goodness-of-fit criteria indicating sufficient prediction accuracy is not a good model if the parameter values are irrational. Parameters based on poor calibration technique will most likely not accurately represent the physical processes. Likewise, rational outputs from model components must be assessed to ensure the development of a physically rational model. Prediction accuracy capabilities may decrease when an irrational model is applied to other watersheds.

Highly complex models representing numerous physical processes may prove difficult to evaluate constraints on parameter values. For example, lack of adequate empirical data for complex physical processes may make it impossible to evaluate the hydrologic rationality of parameters. Likewise, it may be more difficult to ensure the rationality of each parameter value as the number of parameters increase with more complex models. On the contrary, simple models may be forced to adjust parameter values outside a rational range in order to account for missing processes in the model. Therefore, model complexity may influence the ability to ensure the rationality of the hydrologic processes being modeled.

1.1.5 Problem: Complexity and Calibration Criteria Selection

The selection of calibration criteria is an important step in the modeling process. Calibration criteria often include goodness-of-fit statistics such as the relative bias, the relative standard error, and the correlation coefficient. Graphical analyses can be used as calibration criteria to develop a visual representation of the model prediction capabilities. Also, fitting of extreme values, such as base flows or peak discharges in hydrologic modeling, can be included in the calibration criteria. The selection of calibration criteria is an important aspect of the modeling process and should be determined prior to beginning calibration to ensure consistency throughout the process.

The selection of calibration criteria adds complexity to model development and calibration. The use of different criteria in calibrating the same model may result in variations in the calibrated parameter values. This will result in different model outputs and, therefore, possibly different designs based on these different results. Likewise, calibration criteria such as best fit of extremes may limit the applicability of using the calibrated model for additional purposes. Also, the more calibration criteria used, the more complicated the steps of calibration become. Calibration criteria should be thoughtfully selected to ensure an efficient and effective calibration process.

1.1.6 Problem: Complexity and Model Calibration Fitting Methods

Fitting methods in calibration can vary including subjective optimization, objective optimization, numerical optimization, and more. Within these methods are additional factors such as the objective functions used to assess model prediction accuracy and the criteria used to determine further calibration changes. For example, parameter sensitivity may be used as a criterion for determining which parameter values should be adjusted in the calibration process.

Fitting methods can vary in complexity from simple graphical methods to complex trial-and-error, or subjective, fitting. The more complex models generally require more complex fitting methods. However, the more complex fitting methods require the user to have greater knowledge of the physical processes, statistical analysis, the modeling process, and the model being calibrated. Attempting to calibrate the model without this prerequisite knowledge can lead to a model that has irrational coefficients, less then optimum goodness of fit, and inaccurate sensitivities. Better knowledge of complex model fitting can avoid these problems.

1.1.7 Summary of Issues in Model Complexity

Each of the components in the modeling process is influenced by complexity. More complex models may require more scientific knowledge, greater rationality constraints, more complicated fitting processes, and complete data records. However, much debate exists regarding the importance of complexity of models in accurately representing the system. It is arguable that the more complex a model, the less sensitive the individual components become and, therefore, the less likely the prediction accuracy will increase. Likewise, the effects of issues within the data on the accuracy and goodness of fit is yet to be determined. Further knowledge of the effects of complexity on model development is required to ensure beneficial advancements in hydrologic modeling.

1.2 Goals and Objectives

The goal of this research was to analyze the issue of complexity in hydrologic modeling. Complexity can enter through each one of the five components of modeling identified above. Greater knowledge of issues related to complexity can improve modeling results. This goal was achieved through the following objectives:

1) To formulate and analyze a series of different model structures to study the relationship between model complexity and prediction accuracy
2) To assess the effects of data anomalies on prediction accuracy of hydrologic models
3) To assess the ability of randomly as well as systematically selected water quality grab samples to adequately represent a population

4) To demonstrate the use of model sensitivity in improving efficiency of subjective optimization
5) To show the effect of calibrating to optimally fit peak discharge rates or to optimally fit baseflows on overall prediction accuracy
6) To develop guidelines for improving all components of modeling of complex hydrologic models

Guidelines are needed to assist engineers and hydrologists in the model development process in order to produce accurate and efficient water quantity models.

1.3 Implications

The use of complex hydrologic and water quality models for problem solving and engineering design requires a greater accuracy of skills than that required of using simple models such as the Rational method or unit hydrograph models. Models such as HSPF that have hundreds of parameters that can be adjusted should not be used without a full appreciation of the vagaries that can result when using a complex model and complex data base. A user, even one with considerable experience with simple models, needs guidelines to follow when fitting model parameters and subsequently interpreting the model output. Fulfilling the above objectives should lead to guidelines related to proper use of complex models in solving complex problems that involve water quantity and quality related issues.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Much research has been done in the areas of model development and calibration for hydrologic modeling. Previous research addressing model structure, calibration fitting methods, model constraints, and data complexity are discussed in this chapter. Explanations of certain hydrologic processes are also discussed.

2.2 Model Structure

Rushton et al. (2006) developed a rainfall-runoff model that estimates groundwater recharge in various climates. Runoff and infiltration calculations take into account soil moisture deficit. Evapotransporation calculations take into account bare soil and crop type and actual evapotransporation values are determined based on soil properties and root depths. The model was tested for two different climates and proved to be reliable in estimating groundwater recharge.

Mandeville et al. (1970) tested a rainfall runoff model on the drier and more variable climate of the Ray Catchment at Grendon Underwood. The model incorporates simple runoff and routing descriptions as well as the effects of soil moisture on actual evaporation and transpiration in a climate in which soil moisture deficit occurs. The results show that runoff volume is better represented by a two or three independent parameter model than a “coaxial graphical correlation”. Also, the hydrograph is better forecasted when simply routing elements are included in the model. The authors state that optimization of interdependent parameters should be avoided to obtain stable optimized values.

Saxton (1983) emphasizes the importance of soil water when modeling hydrologic or agricultural systems. He reviews methods for simulating soil water for hydrologic and agricultural uses including mathematical representations and simulation models. He explains that input values must depend on climate, soil characteristics, and vegetation type to adequately represent the mass balance of soil water. Saxton developed SPAW, a soil water simulation model, and applied it to both corn and wheat in varying climates to demonstrate the capability of models to account for factors effecting input values and more accurately represent soil water hydrology.

Dawdy and O’Donnell (1965) discuss the two types of models: over-all models and “complete specification” models. Over-all models are lumped models consisting of components based on empirical relationships. The parameters and construction of components are adjusted so that the results are within a certain tolerance of known outputs based on known inputs. “Complete specification” models are less empirical and more based on relationships of physical properties. Dawdy and O’Donnel developed a model of catchment behavior based on both of these model types and utilized a computer program to optimize the parameters. The results show that computers are capable of optimizing parameters in an objective manner.

Holtan et all. (1967) developed a mathematical model to simulate the infiltration process. The model consisted of the rainfall-runoff process, soil moisture accounting, and evapotranspiration. The parameter values were based on available data and information regarding soil and vegetation.

Madramootoo and Broughton (1987) developed a deterministic model to simulate surface and subsurface flow from agricultural fields. The model simulated the most important hydraulic and hydrologic processes on an hourly basis during the growing season. These processes included rainfall, interception, depression storage, infiltration, evapotranspiration, drainage, and overland flow. The drainage consisted of both subsurface and surface methods. The model was applied to two hypothetical 20 ha agricultural fields. The results imply that subsurface drainage lowers the water table faster. Therefore, this drainage method provides more opportunity for infiltration and reduces flooding in agricultural fields.

Bennis and Crobeddu (2007) developed an improved rational hydrograph method for modeling runoff on small urban catchments. The improved method takes into account rainfall with varying intensity, pervious and impervious areas in determining the runoff coefficient, and calibration of parameters in sequence. The new model is a linear model that was calibrated for low and high intensity rainfall events in sequence. Six rainfall events were analyzed with the new model on two urban catchments. The results showed good agreement to actual runoff values and equivalent runoff values to a more complex non-linear reservoir model which was used as an alternative comparison.

2.3 Fitting Methods

Mein and Brown (1978) developed a statistical method to determine the sensitivity of parameters based on the response surface shape. The method applies to parameters optimized with a quadratic object function, such as the sum of squares. The method is applied to the Boughton Model, which is a watershed model. The results suggest that model results are not overly sensitive to individual parameter values. However, relationships between parameter values and watershed characteristics are imprecise. Therefore, it cannot be assumed that changing a parameter value will represent changes in the watershed characteristic.

Dawdy and O'Donnel (1965) developed a simple, 9 parameter model of the hydrologic cycle. The model was optimized using computing techniques that were developing at the time to decrease subjectivity. The results showed that computer techniques are a feasible approach to modeling the hydrologic cycle and eliminating subjectivity in the optimization of parameters.

Nash et al. (1970) discuss the principles involved in modeling river flow and provide suggestions. The model should reflect physical characteristics of the area being modeled if expected to be applied to other areas. The simpler the model is, the more optimal the parameters will be defined. Automatic optimization is suggested to eliminate subjectivity in fitting the model to data.

2.4 Model Constraints

Benaman and Shoemaker (2004) address the difficulty in calibrating complex models, such as TMDL models. In their research, complex models involve multiple parameters with a wide range of suggested values. They provide a method for reducing parameter ranges by taking into account site-specific data before calibration or sensitivity analyses are completed for the models. The steps include selecting parameters and determining parameter range, conducting an initial Monte Carlo analysis, conducting an interval-spaced sensitivity analyses, selecting a threshold, reducing the parameter range, and confirming the new parameter range with a final Monte Carlo analysis. The method was applied to the Soil and Water Assessment Tool Model in the Cannonsville Reservoir System Watershed. Both hydrology and sediment output were considered, but sediment proved early to be the most sensitive out put component to parameter changes. The method successfully reduced the range of sediment load outputs simulated from the model. This method can potentially reduce the calibration process for models used in developing TMDLs.

O'Connel et al. (1970) applied a simple river flow model to the Brosna Catchment at Ferbane and conducted multiple tests to optimize the six parameters. In each test, certain parameters were optimized, while others were set to a fixed value, some values eliminating the parameter completely. The results determined that it is easy to account for 80-85% of the initial variance with a model. To better obtain optimum values of parameters, avoiding interdependent parameters is suggested.

2.5 Data Complexity

Schilling and Fuchs (1986) explain the sources of inaccuracies involved in modeling, including input data error, simplification of processes in model, parameter errors, and numerical problems. They compare hydrograph results from simple models with a complex reference model to evaluate the sources of error in modeling. The models were spatially distributed. Simplifications of the reference model were conducted one at a time to determine the effects of each model component on the accuracy of the results. Based on the analyses, they determined that spatial resolution of the rainfall data has the greatest impact on the accuracy of hydrograph simulation. Inaccurate representation of spatial rainfall distribution also contributes to error in modeling runoff from real storms.

Simplification of an insensitive model component did not seriously affect the model accuracy, assuming that the model parameters are correct and the spatial rainfall distribution is realistically represented.

2.6 Hydrologic Processes

Allen et al. (1998) present a method to calculate evapotranspiration rates for agricultural purposes. The method is based on a reference crop evapotranspiration rate (ETo), which is defined as evapotranspiration from a hypothetical grass reference surface that is never short of water. ETo is only affected by climatic parameters. Allen et al. adjust ETo based on crop type and environmental conditions. A rate for crop evapotranspiration under standard conditions (ETc) is developed by adjusting ETo for a particular crop, still assuming the water supply is not lacking. ETc is calculated with the following equation:

$$ETc = ETo * Kc \qquad \text{Eq. (2-1)}$$

where Kc is the crop coefficient. Kc incorporates factors in addition to climate to the calculation of ETc. The value of Kc is based on the crop type and the time during the growing season. Kc is lowest during the initial stages of the growing season and increases linearly during the developing stage of the crop until the crop reaches a maximum height. The middle stage is assigned the maximum Kc value. In the final stage, the crop development has ceased and less water is required for transpiration. The land cover is at a maximum, which decreases evaporation capabilities. Therefore, the value of Kc decreases linearly until the harvest point. Allen et al. (1998) provide a table of Kc values based on the crop type and growing stage.

Kc can be divided into coefficients to represent the individual processes of evaporation and transpiration. These supplemental crop coefficients are known as the basal crop coefficient, Kcb, for transpiration, and the evaporation coefficient, Ke. The sum of Kcb and Ke equals Kc. The final adjustment is for non-standard conditions. Allen et al. represent the effects of environmental and water stresses on crop evapotranspiration by multiplying ETc by a water stress coefficient, Ks. The evapotranspiration values calculated by this method can be used as guidelines in determining the amount of water required for agricultural practices to compensate for evapotranspiration losses.

Lu et al. (2005) define the concept of potential evapotranspiration (PET) as the maximum rate of evapotranspiration that can occur from a short green crop, standing at a uniform height and completely shading the ground. The value of PET is based on an infinite supply of water and only limited by available energy. Numerous empirical relationships exist to estimate the potential evapotranspiration. Lu et al. (2005) introduce the Thornthwaite method (Thornthwaite 1948), which estimates potential evapotranspiration based on average monthly temperature and represents a 12 hour day and 30-day month:

$$PET = 1.6 * L_d \left(\frac{10T_j}{I} \right)^a \qquad \text{Eq. (2-2)}$$

where PET = monthly PET (cm), T_j = mean monthly temperature (degrees Celsius), L_d is the time from sunrise to sunset in multiples of 12 hours, and the exponent a is given by:

$$a = 0.49 + 0.0179 * I - 0.0000771 * I^2 + 0.000000675 * I^3, \qquad \text{Eq. (2-3)}$$

and I is the annual head index, given by:

$$\mathrm{I} = \sum_{i=1}^{12} \left(\frac{T_j}{5} \right)^{1.5}. \qquad \text{Eq. (2-4)}$$

CHAPTER 3

METHODOLOGY

The first task was to develop a multi-component model that reflected important processes of the hydrologic cycle. Each component was based on the continuity of mass: inflow minus outflow equals the change in storage. Each component is represented by a function, often linear, with one or two parameters to be fitted. The data required to fit the coefficients are limited to rainfall and runoff. Because data bases for daily rainfall and runoff depths are commonly available, a daily time increment, rather than hourly, was selected.

3.1 Model Development

3.1.1 Model Components

A complex model was developed to simulate the processes in the hydrologic cycle. The model simulates rainfall, interception, infiltration, surface runoff, evapotranspiration, interflow, and baseflow. The model was divided into five layers or zones: (1) interception layer; (2) surface layer; (3) root zone; (4) vadose zone; and (5) saturated zone. Figure 1 represents the hydrologic processes and zones simulated by the model. The model uses daily rainfall data as input and simulates the processes to generate daily runoff depth and discharges that can be compared to the measured runoff data for the watershed.

Developing a new model was an alternative to applying an existing model. This alternative enabled the model to have characteristics that are important to assessing the effects of complexity. Specifically, the model components were assembled such that gradual changes in structural complexity could be made. Additionally, the individual components were designed such that each had just one or two parameters that controlled its functioning. When a component was removed, then the change in the number of parameters would be minimal.

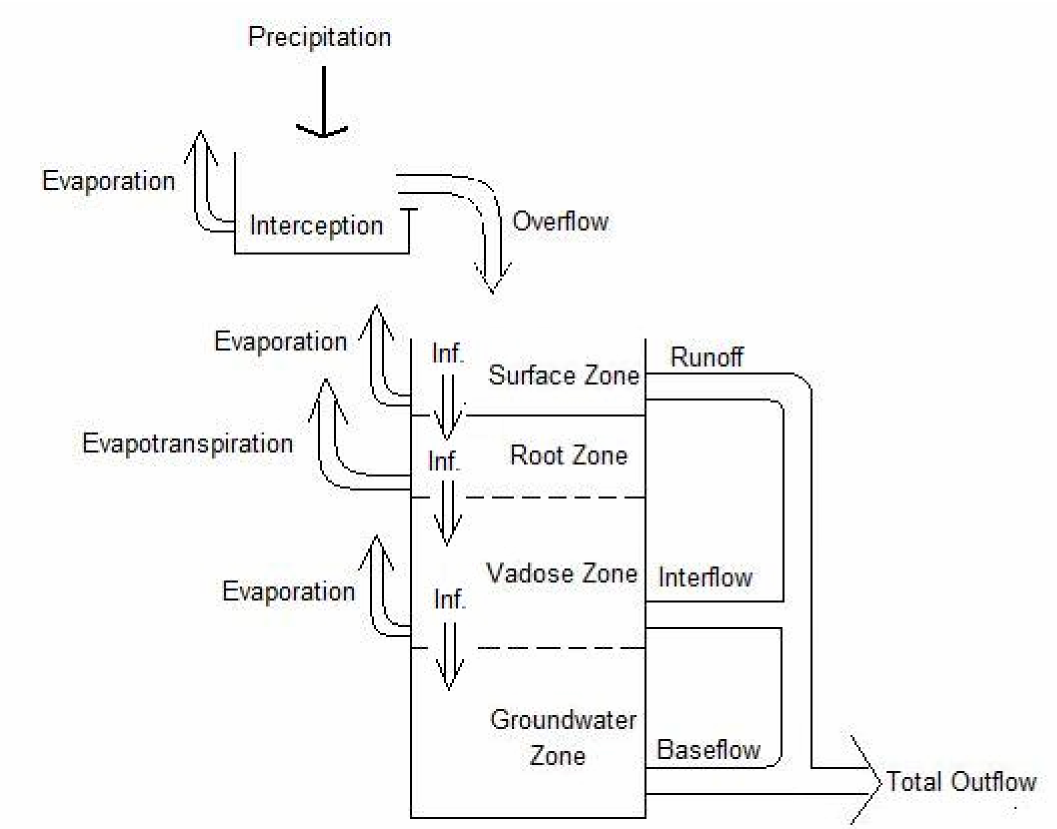

Figure 3.1. Flowchart of Hydrologic Processes and Zones Simulated by the Model

3.1.2 Rainfall

Rainfall data were collected from the National Oceanic and Atmosphere Association's National Climate Data Center (http://www.ncdc.noaa.gov/oa/ncdc.html).

The rainfall gauge used was located in Front Royal, Virginia, and is identified by the COOP ID number 443229. The rain gauge is located at 38 degrees and 54' north and 78 degrees and 11' west. Data from January 1, 2003, to December 31, 2006 was used. Four years of daily precipitation depths were used totaling to 1460 data values. The data were recorded in inches per day and converted to millimeters per day for the purpose of this research. The data had an average rainfall depth of 2.92 mm/day. The standard deviation equaled 7.59 mm. The maximum rainfall event equaled 61.45 mm.

3.1.3 Interception

In the event of precipitation, the interception layer is the first layer to intercept and store rainfall. The maximum amount of storage in the interception layer is a set value from Madramootoo and Broughton (1987). For this analysis a value equal to 0.5 mm was used. When the storage maximum of the interception layer is reached, the excess rainfall overflows and is added to the surface storage.

3.1.4 Surface Runoff

Once the maximum depth of interception is reached, the excess rainfall is added to the surface storage depth. Surface storage can be depleted through surface runoff, infiltration, or evaporation. Surface runoff, which is part of the total runoff, is a function of surface storage, landcover, and drainage area. The equation for surface runoff is:

$$QS = PSZ1 * SSZ^{PSZ2} \qquad \text{Eq. (3-1)}$$

where QS = surface runoff depth (mm), PSZ1 = calibrated surface runoff interception parameter, SSZ = surface storage depth (mm), and PSZ2 = calibrated surface runoff shape parameter. For each time period, Eq. 2 is used to compute the potential surface runoff. If the calculated depth of runoff exceeds the depth of water in surface storage, then the runoff depth equals the surface storage depth and storage is set to zero.

3.1.5 Infiltration

Infiltration is simulated with or without rainfall, assuming the water storage conditions in each layer are favorable. Water infiltrates from the surface layer to the root zone; the root zone to the vadose zone; and the vadose zone to the saturated zone. The infiltration rates vary with soil moisture in the appropriate zone. Infiltration from the surface zone into the root zone depends on the soil moisture of the root zone, the storage depth in the surface zone, and a calibrated infiltration parameter specific to the surface zone. The equation for computing infiltration depth into the root zone is:

$$ISZ = PISZ*SSZ*(1-SRZ/DR) \qquad \text{Eq. (3-2)}$$

where ISZ = infiltration into the root zone (mm), PISZ = infiltration parameter specific to the surface zone, SSZ = storage depth in the surface zone (mm), SRZ = storage depth in the root zone (mm), and DR = depth of the root zone (mm). The parameter PISZ is dimensionless. If the depth of infiltration calculated exceeds either the depth of water stored in the surface layer or the available depth in the root zone, then the depth of infiltration is set equal to the minimum of these two depths. The final computed depth is subtracted from the surface zone storage and added to the root zone storage.

Infiltration into the vadose zone depends on the soil moisture of the vadose zone and a maximum infiltration rate specific to the layer. The infiltration depth increases as the root zone storage increases and as the vadose zone soil moisture decreases. The equation for infiltration into the vadose zone is:

$$IRZ = PIRZ*(SRZ/DR)*(1–SVZ/DV) \qquad \text{Eq. (3-3)}$$

where IRZ = infiltration into the root zone (mm), PIRZ = calibrated infiltration parameter specific to root zone (mm), SRZ= storage depth in the root zone (mm), DR= depth of the root zone (mm), and SVZ = storage depth in the vadose zone (mm), and DV = depth of the vadose zone (mm). If the depth of infiltration calculated exceeds either the depth of water stored in the root zone or the available depth in the vadose zone, then the depth of infiltration is set equal to the minimum of these two depths. The computed depth is subtracted from the root zone storage and added to the root zone storage.

The infiltration into the saturated zone is directly related to the soil moisture of the vadose zone and the infiltration parameter for the vadose zone. The equation for infiltration into the saturated zone is:

$$IVZ = PIVZ*(SVZ/DV) \qquad \text{Eq. (3-4)}$$

where IVZ = infiltration into the saturated zone (mm), PIVZ = infiltration parameter specific to the saturated zone (mm), SVZ = depth of storage in the vadose zone (mm), and DV = depth of the vadose zone (mm). As the soil moisture increases, the infiltration depth increases. If the depth of infiltration calculated exceeds the depth of water stored in the vadose zone, then the depth of infiltration is set equal to the depth of water in the vadose zone.

3.1.6 Evaporation and Transpiration

Evapotranspiration occurs daily based on a maximum daily rate calculated throughout the year. The maximum amount of water that can be depleted daily through evapotranspiration depends on a maximum crop specific evapotranspiration rate (MET). MET is calculated based on the method derived by Allen et al. (2005) and explained in Chapter 2. MET is a function of the potential evapotranspiration (PET), which accounts for only climatic factors, and a crop based coefficient, which accounts for the effects of crop type on transpiration and evaporation.

The potential evapotranspiration (PET) depth, calculated using Eq. 2-2, is used to represent Allen et al.'s Reference Crop Evapotranspiration (ETo). ETo represents the maximum amount of evapotranspiration that can occur from a reference crop assuming that the water supply is sufficient. Temperature is the only climatic factor in calculating ETo for the purpose of this model. The average monthly temperatures for the state of Maryland were determined from the Environmental Data Service (1968). Monthly PET values were calculated and the following sinusoidal function was fit to represent the daily changes in PET (mm):

$$PET = PPET * (0.193 + 0.185*\sin(6.283*(ID + 250)/365))*10 \qquad \text{Eq. (3-5)}$$

where PPET = potential evapotranspiration parameter and ID = day of year.

The maximum daily evapotranspiration rate, MET was calculated by using 2-1 from Allen et al. (1998). For this research, ETc is referred to as MET and ETo is referred to as PET. The values specified for these terms were used in place of the terms identified by Allen et al. The crop coefficient for field corn was used, provided by Allen et al.

MET was then divided into a maximum crop specific evaporation rate (ME) and transpiration rate (MT). These maximum values represent the total daily amount of evaporation and transpiration that would occur for a specific crop, assuming water shortage never occurs, and sum to the MET. The division is based on the stage of the crop's growing season. As a crop matures, the area of land that is shaded by the crop increases. An increase in shaded area decreases the amount of evaporation. However, crop growth increases the amount of water that can be transpired. Therefore, transpiration is directly and evaporation is indirectly related to stage of development of the crop.

Weighted coefficients for evaporation and transpiration, WE and WT, respectively, were calculated based on the height of the crop, which is related to the maturity of the crop. As WE increases, WT decreases, with the sum always equal to 1. The respective weighted coefficient is multiplied by MET to calculate ME or MT, the daily maximum values of transpiration and evaporation, respectively.

The actual amount of water evaporated and transpired is updated as the model simulates the processes in each layer. If the actual amount for either process exceeds the maximum value allotted for that day, the process ceases. For example, if the actual amount of evaporation equals ME after evaporation is applied to the surface layer, the evaporation process will not occur in the remaining zones for that time period. The model updates the maximum allowable rates for evaporation and transpiration (ME and MT, respectively) by subtracting the actual depths that occur in each layer from ME and MT throughout the daily simulation.

The actual amount of evaporation that occurred in each layer was based on the water supply available, the daily potential evapotranspiration (PET), and the evaporation parameters specific to the layer. In the interception layer, the actual amount of evaporation is calculated by:

$$EI = PEXI * STI \qquad \text{Eq. (3-6)}$$

where EI = actual amount of evaporation from interception layer (mm), PEXI = calibrated evaporation parameter for the interception layer, and STI = storage depth in the interception layer (mm). Therefore, as the storage in the interception zone increased, the evaporation rate increased. If EI exceeds either the storage in the interception or the maximum allowable daily evaporation depth, then EI is made equal to the smaller of these two values. It is subtracted from the surface storage and added to the daily total.

In the surface zone, evaporation is simulated after runoff and infiltration. The actual amount of evaporation is calculated with the following equation:

$$ESZ = PESZ*TO*SSZ \qquad \text{Eq. (3-7)}$$

where ESZ = actual amount of evaporation from surface zone (mm), PESZ = calibrated evaporation parameter for the surface zone (Celsius $^{-1}$), TO = daily temperature (Celsius), and SSZ = storage depth in the interception layer. If the calculated depth of evaporation exceeds either the storage in the surface zone or the maximum allowable daily evaporation depth, then ESZ is made equal to the smaller of these two values and is subtracted from the surface zone storage.

The actual evaporation from the root zone is calculated by:

$$ERZ = PERZ * SRZ/DR \qquad \text{Eq. (3-8)}$$

where ERZ = actual amount of evaporation from root zone (mm), PERZ = calibrated evaporation parameter for the root zone (mm), SRZ = depth of storage in the root zone (mm), and DR = depth of the root zone (mm). PERZ acts as a maximum evaporation depth for the root zone and the depth of evaporation increases with an increase in the soil moisture in the root zone. If ERZ exceeds either the storage in the root zone or the maximum allowable daily evaporation depth, then ERZ is made equal to the smaller of these two values.

The vadose zone is the final zone in which evaporation occurs. As in the root zone, the evaporation parameter acts as a maximum rate and the actual evaporation increases with soil moisture. The actual amount of evaporation from the vadose zone was calculated by:

$$EVZ = PEVZ * SVZ/DV \qquad \text{Eq. (3-9)}$$

where EVZ = actual evaporation from the vadose zone (mm), PEVZ = calibrated evaporation parameter for the vadose zone (mm), SVZ = depth of storage in the vadose zone (mm), and DV = depth of the vadose zone (mm). If EVZ exceeds either the storage in the vadose zone or the maximum allowable daily evaporation depth, then EVZ is made equal to the smaller of these two values.

The model only simulates transpiration in the root zone. The actual amount of water transpired is based on the soil moisture of the root zone and the maximum amount of transpiration allotted for that time period. The actual depth of transpiration is calculated by:

$$TRZ = TXH * SRZ/DR \qquad \text{Eq. (3-10)}$$

where TRZ = actual depth of water transpired in the root zone (mm), TXH = maximum depth of transpiration allotted for that day (mm), SRZ = storage depth of water in the root zone (mm), and DR = depth of root zone (mm). The fraction SRZ/DR represents the soil moisture in the root zone; therefore, transpiration increases with soil moisture in the root zone. If TRZ exceeds either the storage in the root zone or the maximum allowable daily transpiration depth, then TRZ is made equal to the smaller of these two values.

3.1.7 Outflow

Groundwater is depleted from the vadose zone and saturated zone as interflow and baseflow, respectively. The amount of groundwater flow depleted from each layer is calculated based on Darcy's law given by the following equation:

$$Q=k*I*A \qquad \text{Eq. (3-11)}$$

where Q = discharge depth, k = hydraulic conductivity of layer, *i* = hydraulic gradient, and A = area of the watershed. The model developed for this research works with depths as opposed to volumes; therefore, the area component was placed with the storage depth for the appropriate zone. The hydraulic conductivity is identified as K_S and K_U for the saturated and unsaturated, or groundwater and vadose, zones, respectively. The hydraulic gradient is specific to each zone and is a calibrated parameter.

3.1.7.1 Interflow

Interflow is the daily outflow from the vadose zone. The following equation was used to calculate the daily interflow from the vadose zone:

$$QV = \text{SVZ*KU/(PQV1+PQV2*sin(6.283*(ID+61)/365))} \qquad \text{Eq. (3-12)}$$

where QV = depth of interflow (mm), SVZ = storage depth in the vadose zone (mm), KU = unsaturated hydraulic conductivity (mm/hour), PQV1 = interflow parameter (mm/hour), PQV2 = parameter representing the cyclical component of interflow (mm/hour), and ID = day of the year. The intent of the cyclical component was to allow the model to account for seasonality. In the event that seasonality is not a factor, PQV2 would have no effect on the prediction accuracy. The unsaturated hydraulic conductivity, K_U, varies with the soil moisture content as well as the soil type. Therefore, the hydraulic conductivity of the vadose zone is calculated based on the soil moisture of the vadose zone using the following equation from Todd and Mays (2005):

$$K_U = K_S * \left(\frac{SM - 0.2}{0.8} \right)^3 \qquad \text{Eq. (3-13)}$$

where K_U = unsaturated hydraulic conductivity (mm/hour), K_S = saturated conductivity discussed below (mm/hour), and SM = soil moisture in the vadose zone. The hydraulic gradient is represented by the parameter PQV1 and is calibrated in the model. A second parameter, PQV2, is added to Darcy's equation to include a seasonal effect. It was assumed that more outflow would occur in the winter because of less evaporation occurring. Therefore, PQV2 was multiplied by a sinusoidal function to vary the daily outflow throughout the year.

3.1.7.2 Baseflow

Baseflow occurs in the saturated zone. The following equation was used to calculate the outflow from the saturated zone:

$$QG = SGZ*KS/PQGZ \qquad \text{Eq. (3-14)}$$

where QG = baseflow (mm), SGZ = storage depth of the saturated zone (mm), KS = saturated hydraulic conductivity (mm/hour), and PQGZ = baseflow parameter (mm/hour). The hydraulic conductivity, K_S, of the saturated zone is a constant value based on the soil type. The values used were based on empirical data provided by Smith (2002). The hydraulic gradient is represented by the parameter PQGZ and is calibrated in the model.

3.2 Steps in the Mass Balance Process

The processes previously defined are simulated in a specific order by the model and the depth of water stored in each layer is updated continuously based on the mass balance equation: input minus output equals the change in storage. The processes in the interception layer are simulated first. During a rainfall event, precipitation is added to the interception storage depth, STI. Once the maximum interception storage depth is reached, the remaining rainfall is stored as excess rainfall. Evaporation is simulated next in the interception layer. The actual amount of evaporation is calculated and subtracted from the interception layer storage, STI. If the actual depth is greater than the maximum allowable evaporation depth allotted for that day, then the depth is set equal to the maximum allowable depth.

Next, the processes in the surface layer are simulated. Excess rainfall from the interception layer is added to the storage depth of the surface layer, SSZ. Then, the amount of infiltration and runoff are calculated and subtracted from the storage depth. The infiltration depth is added to the root zone depth, SRZ, and the runoff depth is added to the total outflow of the model. Finally, the depth of water lost to evaporation is calculated and subtracted from the storage depth, SSZ. If the actual depth is greater than the maximum allowable evaporation depth allotted for that day, then the depth is set equal to the maximum allowable depth and subtracted from storage.

The model then simulates the root zone processes. The depth of water that infiltrates into the vadose zone is calculated. The depth is subtracted from the storage depth of the root zone, SRZ, and added to the storage depth of the vadose zone, SVZ. Next, evapotranspiration is simulated. Evaporation is simulated first. The actual evaporation depth is calculated and subtracted from the root zone storage, SRZ. If the actual depth is greater than the maximum allowable evaporation depth, then actual the depth is set equal to the maximum allowable depth. Next, transpiration is simulated. The actual depth of transpired water is calculated and subtracted from the root zone storage depth, SRZ.

The processes in the vadose zone are then simulated. The depth of infiltration into the saturated zone is calculated. This depth is depleted from the vadose zone storage, SVZ, and added to the saturated zone storage, SGZ. Next, the depth of interflow is calculated. The interflow is subtracted from the vadose zone storage depth, SVZ, and added to the total outflow of the system. Finally, evaporation is simulated and depleted from the vadose zone storage, SVZ. If the actual depth is greater than the maximum allowable evaporation depth allotted, then the depth is set equal to the maximum allowable depth.

The baseflow process in the saturated zone is simulated. The depth of baseflow for the time period is calculated. The baseflow is subtracted from the depth of the saturated zone, SGZ, and added to the total outflow from the system.

The entire cycle is then repeated for the next day.

3.3 Subjective Optimization Calibration

For each version of the model, the parameters were calibrated using the subjective optimization process. A maximum of 14 parameters were available for calibration, depending on the model. The process involved multiple steps. First, initial parameter values were entered into the program. These estimates were developed based on average daily rainfall and runoff values from the calibration data and climatic and crop characteristics for the month of June. Developing reasonable initial estimates should minimize the number of required iterations to achieve the optimum. Second, the model output was analyzed based on criteria to determine the prediction accuracy of the model total outflow compared to the observed total outflow. Third, based on the analysis, changes were made to selected parameters. No more than two to four parameters were changed at a time, which should enable the sensitivities of the individual parameters to be qualitatively assessed. It was important that changes were made to parameters that were independent of each other so that independent effects could be assessed. When the model yielded results that best satisfied the criteria for accuracy, the calibration process ended and the respective parameters were used as the final model.

3.3.1 Goodness-of-Fit and Model Assessment

The calibration criteria were based on a combination of goodness-of-fit statistics and graphical analyses. Overall goodness-of-fit statistics including the correlation coefficient, standard error of estimate, relative standard error, average bias, and relative bias were analyzed. Calibration was conducted with the goal of increasing the correlation coefficient to increase the association between the observed and the predicted runoff of the model. This would suggest good prediction accuracy. Another calibration goal was to decrease the standard error and relative standard error. A relative standard error close to one suggests that the model does not provide greater prediction accuracy than using a mean value of the observed data to predict the runoff values. Therefore, minimizing the standard error and relative standard error through calibration will improve the accuracy of the model. Finally, the average and relative biases should optimally be near zero. Bias reveals whether the model is overpredicting or underpredicting the observed values. The greater the bias, the greater the need for continuing the calibration process. These statistics as a group portrayed the prediction accuracy of the entire model.

3.3.2 Peak Discharge Prediction Accuracy

The overall accuracy prediction of the model is based on the goodness-of-fit statistics. When necessary, however, additional criteria and graphical analyses were used to analyze individual problems in predictions and to determine which parameters would be changed to most effectively improve the goodness of fit. The accuracy of the predicted peak discharge rates was analyzed. The highest discharge for each month for the observed data and the corresponding computed peak for that same day were compared. The prediction accuracy of the computed peak discharges was analyzed based on the standard error and bias for each monthly peak value. The greater the standard error, the less accurate the prediction capability of the model and the greater the need for further calibration. Ideally, the model should be calibrated to avoid overall and local biases for predicted peak discharges. Biases suggest poor prediction accuracy.

In addition to the computed bias, biases were identified by plotting the observed peak values on the x-axis and the predicted peak values on the y-axis, when necessary. A model with perfect prediction accuracy would reveal a line at a 45 degree angle on the graph. Deviation from this line reveals the bias in prediction accuracy. If the plotted points lie above the 45 degree line, the model is overpredicting the peak discharge values. If the plotted points lie under the 45 degree angle line, the model is underpredicting the peak discharge value. The further that the points lie from the straight line, the more necessary it is to continue the calibration process to attain better prediction accuracy.

Model parameter changes are necessary to correct biases in peak discharge prediction. To correct the overprediction of peaks, surface infiltration parameters can be increased to limit the occurrence of surface runoff. Decreasing the surface runoff parameters can also improve positive biases of peak discharge rates. Increasing the evaporation will decrease the overall runoff and, therefore, also decrease the peak discharge rates. Decreasing the surface infiltration and evaporation parameters and increasing the surface runoff parameters will have the opposite effect, and, therefore, correct the underprediction of peak, or negative bias.

3.3.3 Low Flow Prediction Accuracy

Similar to the peak discharge rates, the lowest flow observed and predicted value for the same day for each month of record was identified and defined as the monthly base flow. The values were plotted against each other to identify overall or local biases, when necessary. The standard error and bias for each monthly base flow value was analyzed. Where poor standard error and significant bias existed, changes were made to improve the accuracy of prediction. For example, if a negative bias exists for low flows, the model is underpredicting compared to the observed low flow values. To correct this, the parameters influencing the groundwater outflow should be increased. Likewise, for overpredictions or a positive lowflow bias, the groundwater outflow parameters should be decreased. Other possible solutions include adjusting the evaporation or infiltration parameters accordingly.

3.3.4 Water Balance Prediction Accuracy

The water balance and the storages for each zone in the model were reported at the end of each calendar year simulated. The water balance is defined as the total runoff and evapotranspiration subtracted from the total rainfall for each year, as shown in the following equation:

$$WB = P - Q - ET \qquad \text{Eq. (3-15)}$$

where WB = water balance, P = total rainfall, Q = total runoff, and ET = total evapotranspiration. The storages for each zone refer to the depth of water in each zone. The storage in the surface layer refers to the water remaining after infiltration, evaporation, and runoff has occurred. The storage in the root zone is the depth of water remaining following infiltration and evapotranspiration from the root zone. The storage in the vadose zone is the remaining depth of water after infiltration, evaporation, and interflow has occurred. And the storage in the groundwater refers to the remaining depth of water following the occurrence of baseflow. It would be expected that the storages in each zone would remain fairly constant and the water balance would equal zero. This suggests that the rainfall each year is equal to the amount of water lost to evapotranspiration, surface runoff, and baseflow. The further the water balance deviated from zero and the greater the changes in storage, the less accurate the model, and the greater the need to continue calibrations.

A positive water balance suggests that the rainfall is exceeding the amount of evapotranspiration and runoff that occurs. To correct this, it is necessary to analyze the distribution of water throughout the hydrologic cycle. If the water leaves dominantly through evapotranspiration, it may be beneficial to increase the amount of runoff by changing surface runoff or groundwater outflow parameters. This will increase the amount of water leaving the surface and groundwater zones and improve the positive water balance. If the water leaves dominantly through surface runoff or groundwater, it would be beneficial to increase evapotranspiration parameters. This will allow more water to leave through evapotranspiration and improve the positive water balance. A negative water balance means that the amount of rainfall is less than the amount of runoff and evapotranspiration. Therefore, based on the water distribution throughout the hydrologic cycle, decreasing the evapotranspiration, runoff, or outflow parameters for the appropriate layer in the ground would correct the negative water balance.

3.3.5 Graphical Analysis

When necessary, a graphical analysis was conducted to visually represent the prediction accuracy. The daily values of the observed rainfall, observed runoff, and predicted runoff were plotted for each year. From the graphical analysis of the first year, the accuracy of the initial storage values can be determined. Poor predictions in the beginning of the data record reveal poor initial storage values. If this is the case, changes to the initial storage values in the interception, surface, root, vadose, or groundwater zone must be made. This can be considered an initial condition calibration in which the storages rather than the parameter values are adjusted. The interception storage would have the least effect on the prediction accuracy, as only evaporation occurs within the interception zone. If the model is over predicting in the first few years, it would be beneficial to decrease the storages. Likewise, if the model is undepredicting in the first few years, increasing the initial storages may be necessary.

The accuracy of the hydrograph recessions of the watershed being modeled is also apparent through graphical analysis. The recessions reveal how accurately the model represents the watershed. If the recessions are too steep, the model is simulating runoff too quickly following the peaks. Therefore, changes should be made to the infiltration, evapotranspiration, and runoff for the surface zone. Conversely, if the computed recessions are too flat, the model is not simulating runoff at a fast enough rate following the peaks and the opposite parameter changes should be made.

The surface zone parameters have the most influence on the peak discharges because peaks occur during or following a rainfall event. Surface runoff has the greatest impact on the total runoff during this time, because the storage in the surface layer is immediately increased from the rainfall. The remaining zones are increased at a slower rate following the simulation of infiltration through each zone as well as evaporation and transpiration. Therefore, the outflow from the groundwater occurs at a much steadier rate then the runoff from the surface layer.

By increasing the parameters for infiltration from the surface into the root zone, the peaks will be decreased because less water will be available in the surface to runoff during a rainstorm. Decreasing the infiltration parameters will therefore increase the runoff by allowing more water to remain in storage in the surface zone and, therefore, be available to become surface runoff. Increasing the surface runoff parameters will increase the peaks by allowing more water to leave the surface storage at a time. Likewise, decreasing the surface runoff parameters will decrease the peaks by limiting the amount of surface runoff that can occur.

Finally, increasing or decreasing the evapotranspiration parameters will decrease or increase, respectively, the amount of water available in the storages of the surface, root, and vadose zone and, therefore, influence the total amount of water available for runoff in either of the zones. Without the graphical analysis, the inaccuracies in recession prediction would be difficult to detect.

The timing and magnitude of the peaks is observed along with the recessions through graphical analysis. The program outputs the monthly observed peak and the predicted runoff corresponding to the time of the observed peak. However, for the remaining peaks, a graphical analysis is necessary to determine the prediction accuracy. Trends such as over or under prediction of peaks or error in the timing of the peaks can be detected through graphical analysis and corrected by altering parameters.

3.3.5 Optimization Guidelines

Based on the calibration criteria and graphical analyses, changes were made to selected parameters to improve the prediction accuracy. General rules were developed to follow during the calibration process to provide guidance for parameter adjustments:

(1) Problem: Biased water balance. First, analyze the water distribution throughout the zones in the model. Then:
- A. Negative Bias
  - i. If the water distribution is dominantly runoff from the surface layer
    - a. decrease the surface runoff
    - b. increase the infiltration into the root zone
    - c. increase evaporation from the surface
  - ii. If the water distribution is dominantly baseflow from the saturated zone
    - a. decrease the saturated zone outflow
    - b. decrease infiltration into the groundwater
    - c. increase evaporation from the root and vadose zone
  - iii. If the water distribution is dominantly evapotranspiration
    - a. decrease the evaporation from the zone with the greatest evaporation
    - b. decrease the transpiration from the root zone
- B. Positive Bias
  - i. If the water distribution is dominantly runoff
    - a. increase the surface runoff
    - b. decrease the infiltration into the root zone
    - c. decrease evaporation from the surface
  - ii. If the water distribution is dominantly baseflow
    - a. increase the saturated zone outflow
    - b. increase infiltration into the groundwater
    - c. decrease evaporation from the root and vadose zone

iii. If the water distribution is dominantly evapotranspiration
   a. increase the evaporation from the zone with the greatest evaporation
   b. increase the transpiration from the root zone

(2) Problem: Biased Peaks
  A. Overpredicted
    i. Decrease the surface runoff
    ii. Increase the surface evaporation
    iii. Increase the surface infiltration
    iv. Increase evaporation in root zone and vadose zone to increase available storage for infiltration
  B. Underpredicted
    i. Increase the surface runoff
    ii. Decrease the surface evaporation
    iii. Decrease the surface infiltration
    iv. Decrease evaporation in root zone and vadose zone to decrease available storage for infiltration

(3) Problem: Biased Baseflow
  A. Overpedicted
    i. Decrease the groundwater outflow
    ii. Increase evaporation from the root zone
    iii. Decrease infiltration into the groundwater
  B. Underpredicted
    i. Increase the groundwater outflow
    ii. Decrease evaporation from the root zone
    iii. Increase infiltration into the groundwater

(4) Problem: Poor accuracy in first year
  A. Overpredicting
    i. Decrease initial storage values
  B. Underpredicting
    i. Increase initial storage values

(5) Problem: Inaccurate Recessions
  A. Too fast
    i. Increase infiltration from surface
    ii. Decrease surface runoff
    iii. Increase surface evaporation
  B. Too slow
    i. Decrease infiltration from surface
    ii. Increase surface runoff
    iii. Decrease surface evaporation

Based on the analysis of the goodness-of-fit criteria and the graphical analysis, these guidelines can be followed to determine the most effective parameter changes in the model to increase the prediction accuracy. For example, if the model output reveals that the model is overpredicting the peaks, the parameter in the surface runoff equation, PSZ1, could be reduced. If the water balance is negative, the parameters for evaporation, PESZ, PERZ, or PEVZ can be adjusted appropriately. Such adjustments are made until the goodness-of-fit statistics and graphical analysis suggest that the model is capable of adequate predictions.

3.4 Model Complexity Reduction Process

Objective 1 aims to determine the affect of model structure complexity on prediction accuracy. The model previously described represents the most complex model in this study. To explore the effects of model structure complexity, the model was simplified 15 times to represent reductions in the complexity of model structure. Each new model was recalibrated and the prediction accuracies were compared. Reductions were made based on two criteria: (1) parameter sensitivities and (2) physical rationality. A sensitivity analysis was conducted for Model 1. Statistically speaking, sensitivities reflect the importance of a parameter in a model. For example, the most sensitive parameter is the most important parameter in model prediction and vice versa. The parameters were ranked from the most important to the least important parameter. These rankings were a key component in determining the simplifications of the model.

For each of the eliminations, the least important parameter was identified and eliminated. This implies that the value was set equal to zero, resulting in the elimination of the entire process, for the case of linear equations. For example, if the evaporation parameter for the root zone, PERZ, was eliminated and set equal to zero, the value of root zone evaporation would always equal zero. Therefore, the root zone evaporation process was essentially eliminated. The only exception occurred for exponential parameters, such as the surface runoff parameter PSZ2. The elimination of PSZ2 involved setting PSZ2 equal to one, resulting in a linear equation for the surface runoff. This eliminated PSZ2; however the equation still contained the parameter PSZ1, which explains why the entire process of surface runoff was not eliminated.

The second criterion for eliminations was physical rationality. In some cases, the elimination of the least sensitive parameter influenced other processes remaining in the model, making the parameter infeasible to eliminate. For example, at one point, eliminating the root zone infiltration parameter, PIVZ, would cease water from infiltrating into the groundwater zone. Without replenishing the groundwater zone through infiltration from the upper zones, the zone would be completely depleted through outflow. When such instances occurred, choices were made based on physical rationality. In some cases, zones were combined, such as the vadose zone and the groundwater zone, to enable the parameter elimination without isolating any zone from receiving water. In other cases, the second least important parameter was eliminated instead. Each of the eliminations, and the reasons supporting them, are explained as the steps of model complexity reduction are discussed herein.

3.5 Grab Sample Analysis Procedure

The goal of the grab sample data analysis was to determine the number of grab samples needed to provide accurate statistics that reflected the statistics of the entire record. This goal was achieved through the following four analyses:

- Compute statistical characteristics of an entire record of suspended sediment data, including the mean and probability distribution.

- Randomly eliminate specified percentages of the data record to simulate grab measurements and recompute the same statistical characteristics.
- Randomly eliminate data points that reflect measurements during storm events to represent grab samples and recompute the same statistical measures.
- Randomly eliminate data points that reflect measurements during low flows and recomputed the statistical characteristics.

These analyses will be used to indicate whether or not grab samples, both random and systematic, can provide statistics that represent a full record of data.

The observed data used in this study was provided by USGS and collected from the Rappahannock River in Remington, Virginia, identified as site 01446000 in the USGS records. The watershed was 619 square miles. Four years of data records, totaling to 1461 daily data values, were collected from the years 1989 to 1992. The water quality data consisted of suspended sediment loads (tons/day). The mean of the observed record equaled 271 and the standard deviation equaled 44.6. These statistical characteristics of the observed data record were compared to the random and systematic grab sample subsets analyzed in the study. The analysis was repeated using the concentrations from the same data record; however, significant differences in the results did not exist.

3.5.1 Analysis Procedure

A computer program was developed to randomly select values from the observed data record to represent the grab sample subset. The program performs statistical tests to determine whether or not the total observed record of data and the subsets have the same mean and are from the same distribution. For the statistical tests analyzed, the total observed record refers to the entire four year water quality record and the grab sample subsets refers to the random and systematic reductions to the entire data record.

The hypothesis of equal means was examined with three tests: (1) the one-sample Z-test; (2) the one-sample t-test; and (3) the two-sample t-test. Each test makes different assumptions regarding the sample and population. The one-sample Z-test assumes that the complete observed data record represents the entire population and the grab samples subset represents the sample. Accepting this hypothesis would suggest that the mean of the grab samples subset is representative of the mean of the complete observed data record. The variance of the observed data record is known and is assumed to be the population variance.

The one-sample t-test assumes that the population variance is unknown. Use of this test recognizes that the observed data of n years is an incomplete record of the entire population. Therefore, accepting the null hypothesis for the one-sample t-test assumes that the mean of the grab sample measurements can represent the true population mean of past, present, and future concentrations. This is the most continuously applied test of grab samples.

The two-sample t-test assumes that both the observed data record and the grab samples subset represent samples of the population and random variation exists in the complete data record as well as in the subset. Rejection of the null hypothesis would suggest that the two samples give different assessments of the population mean. In most case, this would imply that he mean of a small subset of grab samples is inaccurate

The Kolmogorov-Smirnov one-sample test was used to determine whether or not the complete observed data record and the grab sample subset represent the same probability distribution. The test is nonparametric, meaning a specific distribution is not assumed. Rejection of the null hypothesis might occur because the probability distribution was a poor assumption or because the parameters of the assumed population were wrong. Acceptance of the null hypothesis would suggest that probabilities based on the grab samples would be representative of the n-year record and, therefore, the population.

The results for each test were analyzed for both random and systematic data reductions. For the random reductions of the observed data record, approximately 50%, 25%, 10%, 5%, 2.5%, and 1% of the total record were selected to represent the grab sample record subset. The results were analyzed to evaluate the record length of the grab sample subset at which characteristics of the subset no longer represent those of the complete observed data record. In some historical records, grab sample measurements appear to have been made only during high flows, while with other records, measurements seem to have been made only during low flows. For these systematic reductions, two analyses were made. First, the observed daily record was separated into a subset of TSS values that were above the mean; values below the mean were discarded. Then a proportion (50%, 25%, etc.) of these values were selected at random to represent the grab sample subset. The six tests described previously were then applied using the grab sample subset and complete record of TSS values. Second, the observed daily record was separated into a subset of TSS values that were below the mean; values above the mean were discarded. Then a proportion of these values were selected at random to represent the grab sample subset. The six tests previously described were then applied using the grab sample subset and the complete record of TSS values. The results were analyzed to determine the effect of grab samples selected randomly and systematically.

3.5.2 Analysis Procedure of Grab Samples Using Systematic Elimination

Water quality samples are not always temporally collected at random. In some cases, the grab samples are collected during high flows, while in other cases, sampling occurs during low flows. To assess the effect of the proportion of grab samples relative to the record length on the accuracy of statistics, two analyses were conducted. First, an intermediate subset of the total observed data record was formed to contain only values above the mean flow. This subset represents the practice of collecting samples during storm events. Then the samples subset was compiled by randomly selecting values from the subset of high flows, with the proportions of 50%, 25%, 10%, 5%, 2.5%, and 1%.

Second, an intermediate subset of the total observed data record was formed to contain only values below the mean flow. This subset represented the policy of collecting samples in the absence of storm events. The samples subset was compiled by randomly selecting values from the subset of low flows, with the proportions of 50%, 25%, 10%, 5%, 2.5%, and 1%. In both analyses, the results were analyzed using the same statistical tests as in the random elimination analysis to determine whether the mean, variance, and distribution of the grab sample subsets represent the total observed data record. The results are discussed herein.

CHAPTER 4

MODEL CALIBRATION

Subjective optimization was used in fitting the parameters to the data set. Multiple fitting criteria (i.e., water balance, runoff bias, the correlation coefficient, the standard error ratio, and the relative bias) were used. Both annual values and 4-year averages were considered in optimizing the coefficients. A primary intent was to understand how prediction accuracy varied with model complexity.

4.1 Model 1

The 14-parameter model previously described is considered the optimum model for this research and identified as Model 1. It is the most complex models and is hypothesized to provide the most accurate predictions of runoff from the watershed. The final parameter values calibrated for Model 1 are shown in Table 4.1-1. The results of the initial calibration for initial storage values are shown in Table 4.1-2.

Table 4.1-1. Calibrated Parameter Values for Model One

PEXI	PXI (mm)	PSZ1	PSZ2	PISZ	PESZ (C^{-1})	PIRZ (mm)	PERZ (mm)	PQV1 (mm/hr)	PQV2 (mm/hr)	PIVZ (mm)	PEVZ	PQGZ (mm/hr)	PPET
0.5	0.5	0.15	0.9	0.2	150	15	7.5	0.004	0.00375	0.6	8	0.000024	3.9

Table 4.1-2. Initial Calibration Results of Initial Storage Values

STI (mm)	SSZ (mm)	SRZ (mm)	SVZ (mm)	SGZ (mm)
0.01	0.001	100	2450	2162.5

4.1.1 Prediction Accuracy of Model 1

The goodness of fit statistics that result from these parameters values in model one are shown in Table 4.1-2 and the graphical representation of the predicted and observed runoff for each year of data are shown in Figures 1:A-D in Appendix A. The relative biases for each individual year alternates from negative to positive, suggesting that the model does not consistently over or underpredict. The biases for the individual years range from negative 0.19 to positive 0.31, suggesting that in those years, the predicted values may be contain an error of -19 and 31 percent, respectively. The overall bias, however, is only 0.001, suggesting that on average, the model overpedicts by only 0.1 percent for all four years.

The water balances for years 1 and 3 have the greatest magnitude and are roughly equal in value but opposite in sign. This is most likely because year 1 and 3 are the wettest and driest years, respectively. The goal of the calibration was to have the individual year water balances as near to zero as possible, but ultimately to produce a final water balance near zero, which was attained.

The runoff bias for the entire four years equals 3 mm. This suggests that the model neither consistently under or overpredicts the runoff, because the total bias is near zero. The bias for the individual years range from -178 to 112 mm. This suggests the models inability to accurately adjust to the varying rainfall and storage values for each individual year. However, the final runoff bias is near zero, which suggests that the model adjusts by neither over or underpredicting and, therefore, an overall bias does not exist.

The R values for years 1, 3, and 4 suggest that the model explains 38, 50, and 33 percent of the variation for each year, respectively. In year 2, the model explains 10 percent of the variation. However, based on the cross correlation results for Year 2, the observed runoff is poorly correlated with the rainfall recorded, as it had the lowest correlation of the four years. A cross correlation analysis between the predicted runoff and observed rainfall revealed a higher correlation than the observed runoff by 0.2. Therefore, it can be assumed that the model is predicting consistently based on the rainfall, and Year 2 contains unexplainable observed runoff based on the rainfall data available. The same explanation is applicable to the poor Se/Sy value for year 2, as the correlation coefficient, R, is a function of Se/Sy.

Table 4.1-2. Goodness of Fit Statistics for Model 1

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	203	-178	0.62	0.8	-0.19
2	-57	95	0.32	0.96	0.17
3	-181	-26	0.71	0.72	-0.07
4	-27	112	0.58	0.83	0.31
Total	-63	3		0.78	0.001

4.1.2 Calibration Process of Model 1

The data used to calibrate the model was from a watershed in Front Royal, Virginia. The watershed is 11.2 square miles. The runoff and rainfall gauges were located 3 miles apart. The stream observed was Manassus Run. Both the rainfall and runoff data were provided as daily averages for the years 2003-2007. The rainfall and runoff data were provided by NOAA and USGS, respectively.

A cross correlation analysis between the rainfall and runoff revealed the following results for years 1 through 5, respectively: (1) 0.54, (2) 0.36, (3) 0.52, (4) 0.44, and (5)

0.2. Upon first calibrating the data, the goodness of fit statistics were considerably lower for the fifth year, regardless of the parameter values. Based on the poor cross correlation results, it was decided to discard the fifth year of data and calibrate with the remaining four years, each of which had a significantly higher cross correlation value.

Many calibration issues occurred during the subjective optimization process. The first year in the data used to calibrate the model experienced much more rainfall than the following three years. Year one received around 1500 mm of rainfall, whereas the second highest depth of rainfall was only 1145 mm. As a result, the calibrated parameters that best fit the final three years resulted in a large underprediction of the first year in total runoff. In order to correct this, the initial storage values for year one were adjusted. The vadose zone storage was increased. Outflow from the vadose zone is directly related to the storage in the vadose zone. Therefore, by increasing the initial depth of water in the vadose zone, more water will be released from groundwater to contribute to the total runoff. This change compensates for the high runoff observed on the watershed without effecting the calibration for the remaining years that did not receive high rainfall depths and, therefore, did not have high runoff depths.

A constant problem which is apparent in the figures is the underprediction of large peaks and the overprediction of smaller peaks. A compromise had to be found between these two problems, because fixing one would negatively affect the other. The peaks are a function of the storage in the surface layer as well as the parameters PSZ1 and PSZ2. These parameters were adjusted until the optimal results were found in which the bias was low and the water balance and runoff bias were close to zero. PSZ2 was particularly important, because it influences the shape of the runoff function, as opposed to just the magnitude as in PSZ1. Therefore, while the peak discharge prediction contains biases associated with the storm magnitude, the overall bias of the peaks is near zero resulting in the most optimal model.

The model generally overpredicted the baseflows for each year. However, decreasing the baseflow by adjusting the groundwater outflow parameters would affect the overall runoff bias by decreasing the total runoff and reducing the peaks. While this would correct the overprediction of the smaller peaks, it would increase the error in the larger peaks, which have a greater impact on the correlation coefficient value. Therefore, a compromise again had to be found between these different criteria. The infiltration parameter in the surface zone was decreased to solve the problem. This decreased the amount of water infiltrating into the ground and, therefore, the amount available for baseflow. The change also increased the peak discharge values because more water was available as surface runoff. This decreased the error in predicting the greater peaks and increased the error in the smaller peaks. The change in the peaks was to a certain extent balanced by the decreased baseflow. Likewise, the runoff bias was not affected greatly because the surface runoff was increasing while the baseflow was decreasing.

4.1.3 Sensitivity Analysis of Model 1

Parameter sensitivity is a measure of parameter importance. High relative sensitivities are associated with parameters that are important to prediction accuracy. Sensitivity is calculated as the change in the prediction value divided by the percent change in the parameter. A sensitivity analysis was conducted on each parameter in the full calibrated model to determine the first simplification in model complexity to be made. Each parameter value was individually decreased by 20% of the calibrated value.

The goodness-of-fit statistics were noted and compared. The relative bias was deemed the most important criterion. The water balance was used as the second most important criteria. The relative standard error and correlation coefficient were viewed as equally important and used as the third comparison. The runoff bias was considered the least important criterion. The relative sensitivities for each model parameter were compared for each of the goodness-of-fit criteria. For example, the relative sensitivities of PSZ1 and PSZ2 in terms of the relative bias were compared. The parameter with greatest relative sensitivity was considered the more important of the two parameters.

Table 1 of Appendix B compares the goodness-of-fit statistics for the calibrated Model 1 with the goodness-of-fit statistics that resulted from each parameter change. The parameters can be ranked as having high, medium, or low relative sensitivity. The parameters PSZ2, PPET, PSZ1, and PISZ have the highest sensitivities, changing the relative bias by 12%, 8%, 6%, and 4%, respectively. The parameters PQVZ, PQGZ, PIRZ, PEVZ, and PIVZ have moderate sensitivities, changing the relative bias by 2%, 2%, 1.5%, 0.8%, and 0.7%, respectively. The parameters PXI, PERZ, PEXI, PQVZ2, and PESZ have the lowest sensitivities with near zero percent changes in the relative bias.

The results showed the following general trends. The surface parameters were the most important parameters, with PSZ2 being the most influential on the prediction accuracy. PSZ2 most likely had the greatest significance because it is the only exponential parameter in the model. The infiltration parameters in each zone were of moderate significance, as they determine how the water is distributed between zones. However, the importance of the infiltration parameters decreased as the parameters represented lower zones in the system. The outflow parameters from the groundwater layers, PQGZ and PQV1, were of equal and moderate importance in prediction accuracy; however, both were significantly less than PSZ1 and PSZ2, suggesting that the surface layer contributes to the majority of the total runoff. With an exception for PPET and PEVZ, the evaporation parameters had little effect on the prediction accuracy of the model. The interception parameters, PEXI and PXI, proved to be the least important in affecting the prediction accuracy, as they contribute to such a small depth of water.

4.2 Model 2

As a result of the sensitivity analysis, modifications to the interception process were selected for Model 2. The interception parameters proved to have the least effect on the model prediction accuracy. Therefore, setting the value of parameters PEXI and PXI equal to zero was the first simplification from the optimal model. This implies that the interception process no longer occurs, and the rainfall is added directly to the surface storage in the surface zone. The rest of the model remained the same and calibration through subjective optimization was conducted for the 12 remaining parameters. Figure 4.2-1 shows the flow chart of the hydrologic processes and zones simulated by Model 2.

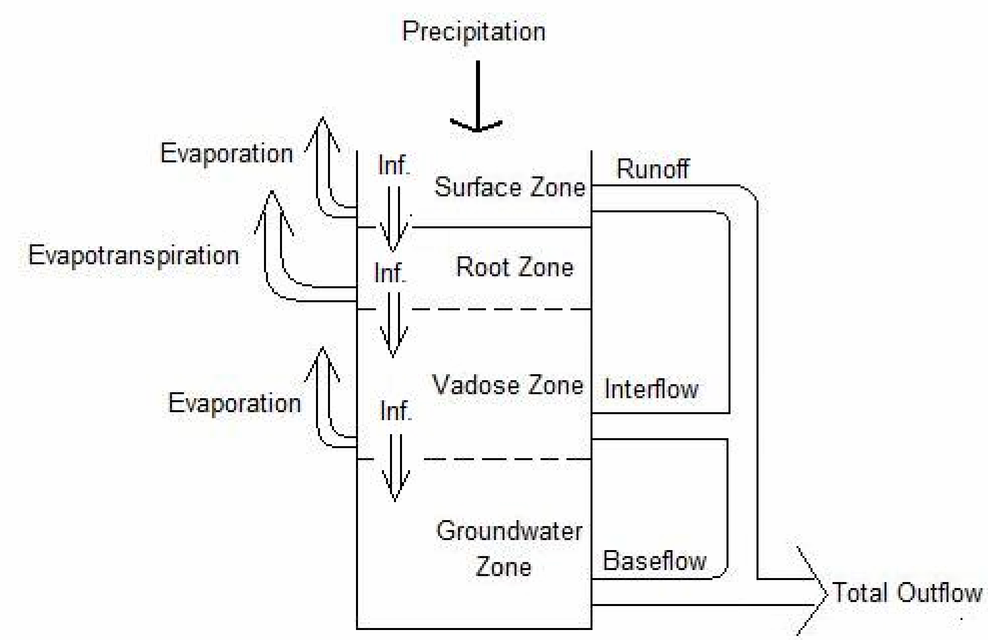

Figure 4.2-1. Flow Chart of Hydrologic Processes and Zones Simulated by Model 2

4.2.1 Calibration Process of Model 2

The calibration process for each model involved multiple steps. A minimal amount of parameters was changed in each step to ensure that the effects of each change were understood. Each calibration step was labeled with the model number and the appropriate letter representing the rank in the calibration process for that model.

4.2.1.1 Calibration Run 2A

The first calibration run for Model 2, labeled Calibration Run 2A, was conducted with the optimal parameters from Model 1. The results are shown in Table 4.2-1. Model 2 worsens the water balance by 4 mm, the runoff balance by 8 mm, and the relative bias by less than 1%. The relative standard error remains the same. These changes are not hydrologically meaningful; however, calibration runs were still made to attempt to improve the prediction accuracy.

Table 4.2-1. Goodness of Fit Statistics for Model 2A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	202	-176	0.62	0.8	-0.18
2	-59	97	0.31	0.96	0.18
3	-182	-24	0.71	0.71	-0.04
4	-27	114	0.57	0.83	0.27
Total	-67	11		0.78	0.004

4.2.1.2 Calibration Run 2B

To calibrate Model 2, the values of the most sensitive parameters were adjusted, based on the sensitivity analysis for Model 1. In Calibration Run 2B, the value of PSZ2 was changed from 0.9 to 0.925 in an attempt to improve the peak prediction problem reported in the Model 1 analysis. PSZ2 was selected because the sensitivity analysis determined that PSZ2 was the most significant of the 14 parameters. The results are shown in Table 4-2-2. While this change improved the prediction of larger peaks, the overall goodness-of-fit statistics worsened. Compared to Model 1, the water balance became more negative by roughly 41 mm; the runoff bias increased by 64 mm; the relative standard error increased by 2%; and the relative bias increased by almost 3%. Therefore, the optimal value of PSZ2 still remains at 0.9 as in Model 1.

Table 4-2-2. Goodness-of-Fit Criteria for Model 2B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	180	-150	0.63	0.79	-0.16
2	-66	108	0	1	0.2
3	-184	-17	0.71	0.71	-0.03
4	-34	126	0.43	0.92	0.3
Total	-104	67		0.8	0.03

4.2.1.3 Calibration Run 2C

The second most significant parameter, based on the sensitivity analysis, was PSZ1. Therefore, for Calibration Run 2C, the value of PSZ1 was adjusted next from 0.15

to 0.165 in an attempt to fix the peak values. Again, the prediction of larger peaks improved, but the overall goodness-of-fit criteria worsened, as shown in Table 4.2-3. Compared to Model 1, the water balance became more negative by 51 mm per year; the runoff bias increased by 78 mm; the relative standard error increased by 1%; and the relative bias increased by almost 1%. Therefore, as with PSZ2, the optimal value of PSZ1 remains at 0.15 as in Model 1.

Table 4.2-3. Goodness-of-Fit Criteria for Model 2C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	174	-144	0.63	0.79	-0.15
2	-68	112	0.17	1	0.2
3	-186	-14	0.71	0.72	-0.02
4	-34	127	0.45	0.9	0.3
Total	-114	81		0.79	0.03

4.2.1.4 Calibration Run 2D

Finally, for Calibration Run 2D, the value of PPET was adjusted from 3.9 to 3.95 in an attempt to correct the runoff bias by increasing evaporation slightly. PPET was the third most significant parameter based on the sensitivity analysis. The results are shown in Table 4.2-4. Compared to Model 1, the water balance became more negative by 14 mm; the runoff bias improved by roughly 3 mm and has a zero bias; the relative bias decreased by 0.08%; and the standard error remains the same. While the goal of attaining zero bias in total runoff was achieved through this calibration change, the change in goodness of fit is not hydrologically meaningful enough to declare that the new value of PPET is optimal. Therefore, the original values calibrated for Model 1 were determined to be the most optimal parameter values for Model 2.

Table 4.2-4. Goodness-of-Fit Statistics for Model 2D

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	197	-177	0.6	0.8	-0.19
2	-61	95	0.31	0.96	0.17
3	-184	-28	0.71	0.71	-0.05
4	-28	110	0.58	0.83	0.26
Total	-76	0		0.78	0.0002

4.2.2 Prediction Accuracy of Model 2

Based on this analysis, the optimal parameter values remained the same for Model 2 as in Model 1 and Table 4.1-1. The goodness-of-fit statistics for the calibrated Model 2 are shown in Table 4.2-1. The water balance, runoff bias, and relative bias worsen in comparison to Model One. However, they are not significantly different from those of Model 1, suggesting that neither Model 1 nor 2 is more accurate than the other. In Model 1, the maximum depth of water that the interception layer could store was 0.5 mm. Therefore, in comparison to model two, the surface layer, root zone, vadose zone, and saturated zone would only receive an extra 0.5 mm a day of rainfall. In comparison to a range of 1,000 to 1,500 mm a year, 0.5 mm is not a significant deduction to affect the prediction accuracy. This is reflected in the sensitivity analysis, since the interception parameters were proven to be the least significant of the 14 parameters. This implies that complexity based on the number of parameters in the model is only beneficial if the parameters included are of a certain level of sensitivity.

4.3 Model 3

Based on the initial sensitivity analysis of Model 1, the next simplification was to set the value of the second least important parameter equal to zero. The evaporation parameter from the root zone, PERZ, was chosen since the model accuracy was less sensitive to the evaporation parameters than parameters for other processes. Therefore, the evaporation process was eliminated and only transpiration and infiltration occurred in the root zone for Model 3. Model 3 was then calibrated based on the remaining eleven parameters.

4.3.1 Calibration Process of Model 3

4.3.1.1 Calibration Run 3A

Calibration Run 3A contained the optimal parameters from Models 1 and 2. The results are shown in Table 4.3-1. In comparison to the results from Model 1, Model 3 improved the water balance by 33 mm; worsened the runoff bias by 18 mm and the relative bias less than 1%; and produced an equal value for the relative standard error. The time series graphs, shown in (Appendix A) Figures A-2:A-D, that the model overpredicts small peaks and underpredicts larger peaks. This problem existed in the previous two models as well.

Table 4.3-1. Goodness-of-Fit Statistics for Model 3A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	215	-177	0.62	0.8	-0.19
2	-50	98	0.3	0.97	0.18
3	-177	-20	0.71	0.71	-0.03
4	-19	119	0.57	0.83	0.28
Total	-30	21		0.78	0.008

4.3.1.2 Calibration Run 3B

In an attempt to improve the prediction of larger peak discharges, the parameter, PSZ2, was increased from 0.9 to 0.95 for Calibration Run 3B. The results are shown in Table 4.3-2**.** While the change improved the accuracy of the larger peaks, it had negative effects on other goodness-of-fit criteria. The smaller peaks increased as well, which caused an increase in the runoff bias by 133 mm and worsened the overall relative bias by almost 5%, compared to the Model One. The increase in the runoff bias caused the water balance to become more negative, by 44 mm. The relative standard error worsened, resulting in a correlation coefficient of 0.0 for years 2 and 4.

Table 4.3-2. Goodness-of-Fit Criteria for Model 3B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	171	-125	0.63	0.79	-0.13
2	-64.5	121	0	1.06	0.22
3	-181	-5	0.71	0.72	-0.008
4	-33	145	0	1.02	0.34
Total	-107	136		0.82	0.05

4.3.1.3 Calibration Run 3C

In order to find a compromise between improving the greater peak discharge predictions and the overall goodness of fit, a smaller change in PSZ2 was tested in Calibration Run 3C at the value of 0.91. The results are shown in Table 4.3-3. Compared to Model 1, the relative standard error increased by 1%; the relative bias increased by 1.6%; the water balance became improved by 16 mm; and the runoff bias increased by 38 mm. The results from this calibration were better than the second and slightly worse than the first, although the difference between this calibration and the first calibration run were insignificant. Therefore, the initial parameter value of PSZ2, which equaled 0.9, is deemed optimal.

Table 4.3-3. Goodness-of-Fit Criteria for Model 3C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	206	-167	0.62	0.79	-0.18
2	-53	102	0.25	0.98	0.19
3	-178	-17	0.71	0.71	-0.03
4	-21	124	0.52	0.87	0.29
Total	-46	43		0.79	0.017

4.3.1.4 Calibration Run 3D

The next problem addressed through calibration was the water balance and runoff bias. The second parameter change involved PEVZ, the evaporation parameter from the vadose zone. In Model 3, evaporation does not occur in the root zone, meaning that more evaporation will occur from the vadose zone. However, the rate of evaporation in the vadose zone was less than the rate of evaporation from the root zone, as it is located at a greater depth in the ground, so the increase in evaporation from the vadose zone does not counteract the loss of evaporation from the root zone. Therefore, as shown in the results from the Calibration Run 3A in Table 4.3-1, the total evaporation decreases, causing an increase in runoff. However, since the evaporation decreased at a greater rate than the runoff increased, the water balance improved from the optimal results from Model 1.

Based on these findings, adjustments were made to PEVZ to attempt to balance the bias in the water balance. For Calibration Run 3D, PEVZ was decreased from 8 to 7.5 to decrease evaporation from the vadose zone and, therefore, correct the water balance. The results are shown in Table 4.3-4**.** Compared to the first run the water balance improved and the runoff bias worsened, each by an insignificant amount per year. The relative bias worsened by 0.3% and the relative standard error did not change. However, further decreasing the value of PEVZ in an attempt to correct the water balance only worsened the relative bias and the runoff bias.

Table 4.3-4. Goodness-of-Fit Criteria for Model 3D

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	216	-176	0.62	0.8	-0.19
2	-48	100	0.3	0.97	0.18
3	-174	-18	0.71	0.71	-0.03
4	-17	122	0.57	0.83	0.29

Total	-23	27		0.78	0.011

4.3.1.5 Calibration Run 3E

The next change, for Calibration Run 3E, was to increase PEVZ from 8 to 8.5, in an attempt to improve the runoff bias. The results are shown in Table 4-3-5**.** As expected, the water balance worsened from the first run in the Model 3 calibration, and the runoff improved slightly. However, the change in runoff bias was not significant enough to affect the relative bias or the relative standard error. Therefore, the negative effects on the water balance are more significant than the positive effects on the runoff bias, and further changes in PEVZ would not be beneficial to the overall goodness of fit.

Table 4.3-5. Goodness-of-Fit Criteria for Model 3E

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	215	-175	0.62	0.8	-0.18
2	-54	99	0.3	0.96	0.18
3	-181	-21	0.71	0.71	-0.04
4	-23	117	0.57	0.83	0.28
Total	-44	19		0.78	0.008

4.3.2 Prediction Accuracy of Model 3

Changes to the parameter values during the calibration process resulted in changes in the goodness of fit of the model. While parameter changes may improve individual goodness-of-fit criteria, the overall goodness of fit did not experience hydrologically meaningful improvements to suggest more optimal parameter values than the original values. Therefore, Model 3A is deemed the most optimal**.** In comparison to the results from Model 1, Model 3 improved the water balance by a 33; worsened the runoff bias by 18 mm and the relative bias less than 1%; and produced an equal value for the relative standard error. These changes, however, are not hydrologically significant.

Therefore, the process of evaporation in the root zone does not contribute to improving the prediction accuracy of the model and the simplified Model 3 and the most complicated model have equal prediction capabilities.

4.4 Model 4

Based on the sensitivity analysis, PQV2 was the next parameter to be set equal to zero. PQV2 was the seasonal component for the vadose zone which allowed more runoff in the winter than the summer, assuming that the root zone is frozen in the winter, and less evaporation occurs. Therefore, Model 4 was based on Model 3, with the PQV2 parameter removed. The remaining ten parameters in Model 4 were then calibrated.

4.4.1 Calibration Process of Model 4

4.4.1.1 Calibration Run 4A

Calibration Run 4A was conducted using the optimal parameters from the previous three models. The results are shown in Table 4.4-1.

Table 4.4-1. Goodness-of-Fit Statistics for Model 4A

Year	**WB (mm)**	**Runoff Bias (mm)**	**R**	**Se/Sy**	**e/y**
1	207	-169	0.59	0.82	-0.18
2	-51	98	0	1.007	0.18
3	-174	-23	0.69	0.73	-0.04
4	-14	115	0.51	0.87	0.27
Total	-32	20		0.806	0.008

Compared to Model 1, the water balanced improved by 31 mm; the runoff bias worsened by 17 mm; the relative standard error worsened by less than 3%; and the relative bias worsened by less than 15. Also, in Year 2, the correlation coefficient is now zero. However, it is important to consider that the correlation coefficient is based on the relative standard error. In year 2, the relative standard error increased by less than 5%

from Model 3 to Model 4. Based on a graphical analysis, a seasonal problem occurs in all four years. The model tends to overpredict runoff in the summer and fall and underpredict in the winter leading into the spring. This is a result of the removal of a seasonal parameter that would otherwise control this error. This trend, however, does not greatly affect the overall bias of the model. Based on the goodness-of-fit statistics, the elimination of PQV2 slightly worsens the overall goodness of fit of the model and further calibration may be necessary.

4.4.1.2 Calibration Run 4B

PQV2 represented seasonal changes in the groundwater outflow. For Calibration Run 4B, in an attempt to compensate for the elimination of a seasonal component in the model, PEVZ was increased from 8 to 10. Evaporation is the only process remaining in the program that has a seasonal component. It is based on the temperature, which varies with the seasons. Therefore, increasing PEVZ, the evaporation parameter in the vadose zone, may counteract the effects of eliminating the cyclical component PQV2 and, therefore, improve the goodness-of-fit statistics. The results are shown in Table 4.4-2. The water balance worsened from the Model 1 by 29 mm; the runoff bias increased by 20 mm; the relative standard error worsened by less than 3%, and the relative bias worsened by less than 1%. These values are not a significant change from the Calibration 4A results. Therefore, changing PEVZ does not improve the calibration of Model 4.

Table 4.4-2. Goodness-of-Fit Statistics for Model 4B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	194	-173	0.59	0.82	0.18
2	-66	88	0	1.002	0.16
3	-190	-38	0.69	0.73	-0.06
4	-30	99	0.52	0.86	0.23

Total	-92	23		0.804	-0.009

4.4.1.3 Calibration Run 4C

For Calibration Run 4C, changes were made based on the sensitivity analysis. PSZ2 was determined the most sensitive parameter and was, therefore, changed for Calibration 4C. PSZ2 was increased from 0.9 to 0.925 in an attempt to correct the continuous problem of underpredicting large peaks. The results are shown in Table 4.4-3. Compared to Model 1, the water balanced worsened by 6 mm; the runoff bias increased by 73 mm; the relative standard error increased by 4%; and the relative bias worsened almost 3%. Therefore, based on the goodness-of-fit criteria, Calibration Run 4C is not an improvement to Calibration Run 4A.

Table 4.4-3. Goodness-of-Fit Statistics Model 4C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	186	-144	0.59	0.81	-0.15
2	-58	109	0	1.05	0.2
3	-176	-16	0.69	0.73	-0.03
4	-21	127	0.34	0.95	0.3
Total	-69	76		0.82	0.03

4.4.1.4 Calibration Run 4D

For Calibration Run 4D, the next most significant parameter, PPET, was increased from 3.9 to 4. This was also an attempt to affect the seasonal problem and improve the runoff bias. The results are shown in Table 4.4-4. Compared to the Model 1, the water balance improved by 13 mm; the runoff bias improved by 3 mm; the relative standard error worsened by 3%; and the relative bias improved by less than 1%. Also, the correlation coefficient for Year 2 equals zero. Calibration Run 4D was deemed the most accurate calibration with goodness-of-fit statistics comparable to the Model 1, despite the zero correlation in Year 2.

Table 4.4-4. Goodness-of-Fit Statistics for Model 4D

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	200	-172	0.59	0.82	-0.18
2	-54	93	0	1.005	0.17
3	-178	-29	0.69	0.73	-0.05
4	-17	109	0.52	0.86	0.26
Total	-50	0		0.81	0.0001

4.5 Model 5

For Model 5, PESZ, the surface evaporation parameter, was set equal to zero; therefore, the surface evaporation process was eliminated. PESZ was the next most insignificant parameter based on the sensitivity analysis. This elimination means that evaporation only occurs in the vadose zone. The remaining nine parameters were calibrated for Model 5.

4.5.1 Calibration Process of Model 5

4.5.1.1 Calibration Run 5A

Calibration Run A for Model 5 was run using the optimal parameters calibrated for Model 4. The results are shown in Table 4.5-1. Compared to Model One, the water balance worsens by 132 mm; the runoff bias worsens by 235 mm; the relative standard error worsens by 3%; and the relative bias worsens by almost 10%. The significant change in prediction accuracy is the result of eliminating a process that is of importance to the model.

Table 4.5-1. Goodness-of-Fit Statistics for Model 5A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y

1	133	-87	0.6	0.81	-0.09
2	-94	151	0	1.03	0.27
3	-195	14	0.69	0.73	0.02
4	-38	161	0.43	0.91	0.38
Total	-195	238		0.81	0.1

4.5.1.2 Calibration Run 5B

For Calibration Run 5B, the main goal was to fix the runoff bias and water balance bias. The model overpredicts the runoff because evaporation is no longer occurring from the surface zone and the water that would normally evaporate is remaining in storage. Surface runoff is a function of storage. Therefore, runoff is occurring at higher rates as a result of the increased storage. To correct this, PISZ, the surface infiltration parameter, was increased as well as PPET, the overall evaporation parameter. This forces the excess surface water to be infiltrated and increasing the amount that will evaporate from the Vadose Zone. The results are shown in Table 4.5-2. Compared to the Model One, Calibration B improves the water balance by 34 mm; worsens the runoff bias by 5 mm; worsens the relative standard error by 2%; and worsens the relative bias by less than 8%. These goodness-of-fit statistics are an improvement from Calibration Run 5A.

Table 4.5-2. Goodness-of-Fit Statistics for Model 5B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	222	-204	0.57	0.83	-0.22
2	-61	84	0.29	0.97	0.15
3	-184	-38	0.68	0.74	-0.07
4	-13	97	0.59	0.81	0.23
Total	-37	-61		0.8	-0.02

4.5.1.3 Calibration Run 5C

Calibration Run 5B overcorrected the runoff bias resulting in a negative bias. Therefore, for Calibration Run 5C, PISZ was decreased to .275, a value three quarters of the way between the value in Calibration Runs 5A and 5B. The results are shown in Table 4.5-3. Compared to the Model One, the water balance worsens by 9 mm; the runoff bias is unchanged; the relative standard error worsens by 2%; and the relative bias is unchanged. These differences are not hydrologically significant, suggesting that Model Five is capable of attaining prediction accuracies at the same level as Model One.

Table 4.5-3. Goodness-of-Fit Statistics for Model 5C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	202	-180	0.58	0.82	-0.19
2	-68	97	0.25	0.98	0.18
3	-187	-29	0.69	0.74	-0.05
4	-19	109	0.57	0.83	0.26
Total	-72	-3		0.8	-0.001

4.6 Model 6

Based on the Model 1 sensitivity analysis, the least important of the remaining parameters is PIVZ, the infiltration parameter from the vadose zone to the saturated zone. Following PIVZ in order of decreasing importance are PIRZ, PEVZ, and PQV1 and PQGZ with equal importance. Of these five parameters, it would be physically irrational to eliminate PEVZ, as evaporation in the vadose is the only remaining evaporation process. Eliminating PIRZ would stop water from infiltrating beyond the root zone causing all rainfall to become surface runoff and the initial storage of the groundwater would eventually be lost through outflow. Likewise, eliminating PIVZ would stop groundwater from infiltrating into the groundwater zone, and eventually the groundwater zone would be eliminated as well. Therefore, for Model 6, the Saturated and Vadose Zone were combined into one zone labeled the Groundwater Zone, eliminating PIVZ and PQGZ. Water now infiltrates from the root zone into the groundwater zone. The water then is either evaporated or released as groundwater outflow from the Vadose Zone. Figure 6.6-1 shows the hydrologic processes and zones simulated by Model 6. The remaining seven parameters in Model 6 were calibrated.

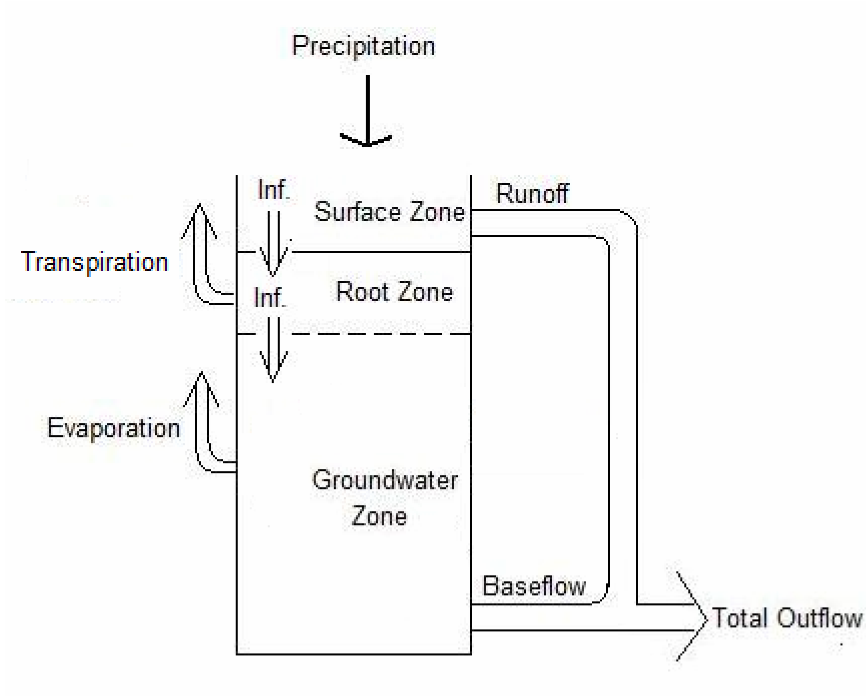

Figure 6.6-1. Flowchart of Hydrologic Processes and Zones Simulated by Model 6

4.6.1 Calibration Process of Model 6

4.6.1.1 Calibration Run 6A

The parameters calibrated for Model 5 were used for the calibration of Model 6A. The results are shown in Table 4.6-1. Compared to the Model 1, the water balance worsened by 1237 mm; the runoff bias worsened by 1093 mm; the relative standard error increased by 14%; and the relative bias increased by roughly 44%. Compared to Model 5, the water balance worsened by 1228 mm; the runoff bias worsened by 1096 mm; the relative standard error worsened by 12%; and the relative bias worsened by roughly 44%. Model 5 had required a change in the optimal parameter values, which suggested a less physically rational model. Likewise, Model 6 requires further calibration of the optimal parameter values from Model 5, suggesting an even less physically rational model.

Table 4.6-1. Goodness-of-Fit Statistics for Model 6A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-612	586	0	1.09	0.62
2	-325	321	0	1.05	0.58
3	-296	56	0.69	0.73	0.1
4	-66	136	0.55	0.84	0.32
Total	-1300	1099		0.92	0.44

The results for Calibration Run 6A show that the model is greatly overpredicting runoff which causes a negative water balance. The individual breakdown of the water process for each year shows that in comparison to the calibrated Model 5 breakdown, more water is released from the Groundwater zone through outflow, evaporation, and transpiration than was released from the Vadose and Saturated Zone in Model 5. The surface runoff remained unchanged. And as expected, water is not released from the Saturated Zone, as it was eliminated for Model 6.

The addition of the initial water storage in the saturated zone to the vadose zone is reflected in the positive runoff bias. The soil moisture in the Vadose Zone for Models One through Five and the Groundwater Zone for Model Six is calculated by dividing the storage depth by the depth of the respective zone. The initial storages for the Vadose Zone in Models One through Five produced an initial soil moisture of roughly 50%. By adding the Saturated Zone into the Vadose Zone, the ratio between the initial storage of water and the depth of the entire Groundwater Zone increases and, therefore, the initial soil moisture increases. Outflow and evaporation from the vadose zone, which is now the saturated zone, is a direct function of soil moisture. Therefore, more water will be released through these processes, contributing to the negative water bias and positive runoff bias.

4.6.1.2 Calibration Run 6B

To correct this, PQV1 was first drastically decreased from 0.004 to 0.00028 to decrease the outflow from the Groundwater Zone for Calibration Run 6B. The results are shown in Table 4.6-2. Compared to Calibration Run 6A, the parameter change improved the runoff bias by decreasing the outflow from the Groundwater Zone. The water balance improved by 1119 mm; runoff bias improved by 782 mm; the relative standard error improved by 11%; and the relative bias improved by 31%. The improvement, however, was not sufficient to produce goodness-of-fit statistics that are comparable to previous models. Also, the runoff bias was overcorrected and went from being positive to negative. Therefore, while the water balance is still negative, suggesting more water is released than precipitated, the model is now producing too little runoff.

Table 4.6-2. Goodness-of-Fit Statistics for Model 6B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	213	-269	0.54	0.85	-0.28
2	-84	5	0.25	0.97	0.01
3	-233	-105	0.67	0.75	-0.18
4	-78	52	0.59	0.82	0.12
Total	-181	-317		0.81	-0.13

4.6.1.3 Calibration Run 6C

For Calibration Run 6C, the negative water balance and now negative runoff bias were addressed. PPET, overall evaporation parameter, was decreased from 4.1 to 3.7. Theoretically, this will decrease the amount of evaporation that occurs, and improve the negative water balance. However, the increase storage remaining in the model will increase the runoff and outflow, as both are a function of storage. Therefore, the parameter change will also improve the negative runoff bias. The results are shown in Table 4.6-3. Compared to Calibration Run 6B, this change improved the water balance by 132 mm; improved the runoff bias by 46 mm; the relative standard error remained unchanged; and decreased the relative bias by almost 2%. While this is an improvement, further calibration is still necessary.

Table 4.6-3. Goodness-of-Fit Statistics for Model 6C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	256	-266	0.54	0.85	-0.28
2	-46	16	0.24	0.98	0.03
3	-201	-90	0.68	0.74	-0.15
4	-58	69	0.58	0.82	0.16
Total	-49	-271		0.81	-0.11

4.6.1.4 Calibration Run 6D

Despite the significant improvements in calibration Run 6C, the water balance and runoff are still negatively biased. Changing PPET improved the water balance at a faster rate than it improved the runoff bias. Therefore, for Calibration Run 6D, changes to both PPET and PQV1 were made. PPET was decreased to 3.35 and PQV1 was increased to 0.00031. The results are shown in Table 4.6-4. Compared to Calibration 6C, the water balance bias worsened by 23 mm; the runoff bias improved by 99 mm; the relative standard error remained unchanged; and the relative bias improved by 4%. Further calibration is still necessary.

Table 4.6-4. Goodness-of-Fit Statistics for Model 6D

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	289	-248	0.55	0.84	-0.26
2	-16	40	0.23	0.98	0.07
3	-168	-62	0.68	0.74	-0.11
4	-33	98	0.56	0.83	0.23
Total	72	-172		0.81	0.07

4.6.1.5 Calibration Run 6E

For Calibration Run 6E, PQV1 was increased to the value 0.0004 to increase the runoff from the Groundwater zone. Increasing the runoff will improve both the negative runoff bias and the positive water balance bias. The results are shown in Table 4.6-5. The change improved the water balance by 36 mm; improved the runoff bias by 114 mm; the relative standard error remained unchanged; and the relative bias was decreased by 5%. These goodness-of-fit statistics for Calibration E are comparable, although still show less accuracy than the previous models. However, one last parameter adjustment was conducted.

Table 4.6-5. Goodness-of-Fit Statistics for Model 6E

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	250	-209	0.57	0.83	-0.22
2	48	73	0.22	0.98	0.13
3	-191	-37	0.68	0.73	-0.06
4	-48	115	0.56	0.84	0.27
Total	-36	-58		0.81	-0.02

4.6.1.6 Calibration Run 6F

Calibration Run 6F was a final attempt to improve the water balance and runoff bias. By decreasing PPET to 3.25, Calibration Run 6F attempted to again decrease the amount of evaporation that occurs within the model and, therefore, improve the water balance. The newly negative runoff bias will also be improved, because as more storage remains in the system, more runoff is simulated. The results are shown in Table 4.6-6.

The parameter change for Calibration Run 6F improved the water balance by 29 mm; improved the runoff bias by 20 mm; the relative standard error remained unchanged; and the relative bias decreased by 0.5%. In comparison to the Optimal Model, Model 6 improves the water balance by 56 mm; increases the runoff bias by 35 mm; increases the relative standard error by 3%; and increases the relative bias by 1.4%. The water balance is a positive improvement from the Model 1 and the increase runoff and relative biases are insignificant changes. However, the 3% increase in the relative standard error shows the Model 6 has poorer prediction accuracy than the Optimal Model. Also, the parameter values are beginning to deviate from the original values in order to compensate for the processes that have been eliminated and attain a comparable goodness of fit. This migh lead to a model that is less physically rational and not applicable to outside data sets. These findings show that at a certain stage in the simplification of model complexity, the physical rationality and prediction accuracy of the model diminishes.

Table 4.6-6. Goodness-of-Fit Statistics for Model 6F

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	263	-208	0.57	0.83	-0.22
2	-36	78	0.21	0.98	0.14
3	-180	-30	0.68	0.73	-0.05
4	-40	124	0.55	0.84	0.29
Total	7	-38		0.81	-0.015

4.7 Model 7

Based on the sensitivity analysis, PIRZ, the infiltration parameter for the root zone, is the least significant of the remaining parameters. Therefore, for Model 7, PIRZ was eliminated. The root zone and groundwater zone were combined. Therefore, the remaining model has only two zones: the surface zone and the groundwater zone. Only runoff and infiltration occur in the surface layer and transpiration, evaporation, and outflow each occur in the groundwater zone. The remaining six parameters were calibrated. Figure 4.7-1 shows the hydrologic processes and zones simulated

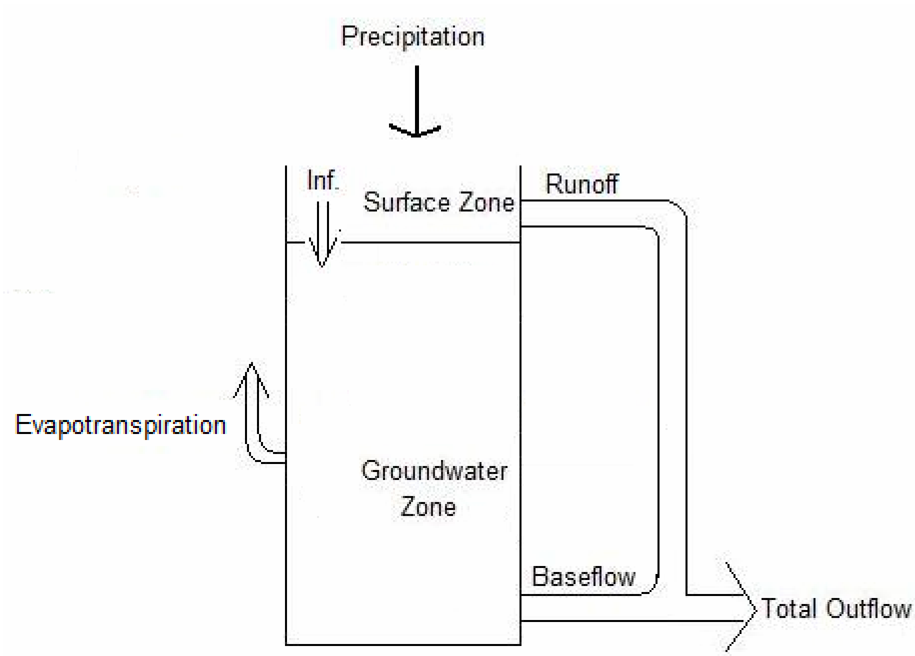

Figure 4.7-1. Flowchart of the Hydrologic Processes and Zones Simulated by Model 7

4.7.1 Calibration Process of Model 7

4.7.1.1 Calibration Run 7A

Calibration Run 7A, the first calibration run for Model 7, was conducted using the optimal parameters from Model 6. The results are shown in Table 4.7-1. Compared to the optimized results for Model 6, the amount of transpiration nearly doubled for each of the four years, while the evaporation rate remained the same. The surface runoff increased and the outflow from the vadose zone decreased for each of the four years.

Transpiration and evaporation are each of function of soil moisture. For Model 6, transpiration is a function of soil moisture in the root zone; whereas, in Model 7, transpiration is a function of soil moisture in the groundwater zone. Therefore, the increase in transpiration from Model 6 to Model 7 suggests that the soil moisture in the groundwater zone in Model 7 is greater than the soil moisture in the root zone in Model 6. Evaporation remained the same for both models, suggesting that the maximum depth of evaporation was achieved and, therefore, the change in soil moistures did not affect the total evaporation.

The infiltration rate from the surface zone in Model 6 was based on the root zone soil moisture; however, in Model 6, the infiltration rate is based on the groundwater zone soil moisture. As a result of the increase in soil moisture in the groundwater of Model 6 compared to the root zone of Model 7, the infiltration of surface water decreased from Model 6 to Model 7. This explains the increase in surface runoff, because more water is stored in the surface storage and converted to runoff.

In Model 6, the outflow from the vadose zone was directly related to the soil moisture of the vadose zone. In Model 7, the outflow from the groundwater is related to the soil moisture of the previous vadose and root zone combined, or the new groundwater zone. While the soil moisture increased when comparing the root zone of Model 6 to the groundwater zone of Model 7, the vadose zone soil moisture of Model 6 decreased in comparison to the groundwater zone of Model 7. Therefore, the outflow from the groundwater decreased from Models 6 to 7.

These changes in the distribution of the water throughout the model are reflected in the goodness-of-fit criteria shown in Table 4.7-1. Compared to Model 1, the water balance worsened by 563 mm; the runoff bias increased by 33 mm; the relative standard error increased by 4%; and the relative bias increased by 1.3%. The large change in the water balance is a result of the increase in transpiration. Further calibration of the remaining parameters for Model 7 is required.

Table 4.7-1. Goodness-of-Fit Statistics for Model 7A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	60	-140	0.58	0.82	-0.15
2	-207	77	0	1.02	0.14
3	-339	-61	0.69	0.73	-0.1
4	-140	88	0.45	0.9	0.21
Total	-626	-36		0.82	-0.014

4.7.1.2 Calibration Run 7B

To correct the largely negative water balance, the transpiration rate was addressed. The model does not contain a parameter that directly affects the transpiration rate. However, the parameter PPET controls the maximum amount of evapotranspiration that can occur each day. Therefore, for Calibration 7B, PPET was decreased from 3.25 to 2.5. The results are shown in Table 4.7-2. Compared to Calibration Run 7A, the water balance improved by 444 mm; the runoff bias worsened by 84 mm; the relative standard error increased by 1%; and the relative bias increased by 3.6%. The change in PPET greatly improved the water balance by decreasing the depth of water lost to transpiration. However, with more water available in storage, more water is available as runoff, resulting in the increase in runoff bias and relative bias. Further calibration is needed to correct the runoff bias.

Table 4.7-2. Goodness-of-Fit Statistics for Model 7B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	191	-127	0.58	0.82	-0.13
2	-87	114	0	1.04	0.21
3	-222	-12	0.69	0.73	-0.02
4	-64	145	0.36	0.94	0.34
Total	-182	120		0.83	0.05

4.7.1.3 Calibration Run 7C

Calibration Run 7B resulted in a negative water balance and positive runoff bias. Therefore, the runoff must be decreased, which would fix both the water balance and runoff bias. To decrease the runoff, the surface runoff parameter PISZ was decreased from 0.15 to 0.125. The results are shown in Table 4.7-3. Compared to Calibration Run 7B, the water balance improved by 149 mm; the runoff bias improved by 70 mm; the relative standard error improved by 1%; and the relative bias improved by 3%.

Compared to Model 1, Model 7 improves the water balance by 30 mm; increases the runoff bias by 47 mm; increases the relative standard error by 3%; and increases the relative bias by almost 2%. While the water balance and runoff bias show insignificant changes, the relative standard error and relative bias begin to show the decrease in prediction capabilities with the decrease in model complexity. Furthermore, the parameter values of Model 7 continue to deviate from the rational parameter values calibrated for Model 1.

Table 4.7-3. Goodness-of-Fit Statistics for Model 7 C

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	258	-196	0.56	0.84	-0.21
2	-51	73	0.2	0.99	0.13
3	-200	-41	0.68	0.74	-0.07
4	-40	113	0.57	0.83	0.27

Total	-33	-50		0.81	-0.02

4.7.2 Sensitivity Analysis of Model 7

With Model 7, the goodness-of-fit statistics began to show a change. Therefore, the sensitivities were expected to change as the parameter values influence the goodness of fit. Model 7 included roughly 50% of the parameters of the original model. Sensitivity analyses were conducted to determine the most important of the remaining six parameters in the model and, therefore, make rational future eliminations. While the goodness of fit of the model 7 predictions are still consistent with Model 1, the analysis was conducted to determine the effects of simplifications on the importance of the remaining six parameters compared to the original 14 parameters. Each of the remaining parameter values was decreased by 20% from the optimal parameters values of Model 7C. The goodness-of-fit criteria resulting from each parameter change are compared to the Model 7C results in Table 2 of Appendix B. The order of importance of the goodness-of-fit criteria is as follows: (1) relative bias; (2) water balance; (3) relative standard error; (4) runoff bias; and (5) rational coefficient.

The analysis showed similar results to sensitivity analysis from Model 1. The parameters are ranked as follows in regards to significance: (1) PSZ2; (2) PSZ1; (3) PQVZ; (4) PEVZ and PPET; and (5) PISZ. Each of these six parameters had high-to-moderate importance in model 1. Therefore, it was expected that all six were significant in Model 7, as well. However, the ranking and magnitude of significance changed slightly from Model 1 to Model 7. PSZ2 and PSZ1 continue to be highly sensitive parameters, changing the relative bias of the predictions by 55% and 40%, respectively. However, PPET has reduced slightly in significance and is ranked as moderate, with a

29% change in relative bias. PQV1, PEVZ, and PISZ also have moderate sensitivities, changing the relative bias by 31%, 30%, and 22%, respectively. These values are similar to the percent change caused by PPET; however, PPET changes the water balance by 1444 mm, whereas PQV1, PEVZ, and PISZ change the water balance by 1150 mm, 1231 mm, and 971 mm, respectively. In Model 1, PISZ was more sensitive than both PQVZ and PEVZ; however, in Model 7, the methods of outflow and evaporation have been reduced with the elimination of parameters, which caused these parameters to become more sensitive. Considering that the parameters with low sensitivity were eliminated in the first simplifications, it is expected that Model 7 does not include any low sensitivity parameters.

As model parameters were eliminated, the prediction accuracy became more sensitive to the remaining parameters. This is apparent by comparing the results from the Model 1 sensitivity analysis in Table 4.1-2 with the results from the Model 7 sensitivity analysis in Table 4.7-3. The percent change for each goodness-of-fit component increased in magnitude from the Model 1 sensitivity analysis to the Model 7 sensitivity analysis. This is not apparent when looking at the total percent change because the time span of four years balances out the biases and reduces the impact of the changes; however, analysis of the individual years reveals the increase in magnitude of the percent change for each goodness-of-fit criterion. For example, for a 20% change in PSZ2, the total relative bias for Models 1 and 7 changed by -6767% and -2756%, respectively. This would suggest that parameters in Model 1 are more sensitive than Model 7. However, for Model 1, a 20% decrease in PSZ2 resulted in an 84%, -60%, 175%, and 58% change in the relative bias for years 1, 2, 3, and 4, respectively. In Model 7, a 20% decrease in

PSZ2 resulted in a 162%, -377%, 543%, and -163% change in the relative bias for years 1, 2, 3, and 4, respectively. It is apparent that the overall relative bias is misleading and the individual years reveal the greater change in Model 7. Therefore, the parameter sensitivity increased from Model 1 to 7 as the number of parameters in the model decreased.

4.8 Model 8

For Model 8, the least significant of the remaining parameters in Model 7, PQV1, was set equal to zero, eliminating the process of interflow from the vadose zone. PQV1 controlled the outflow from the groundwater zone. Therefore, Model 8 only portrays surface runoff, infiltration into the groundwater, and evapotranspiration from the groundwater. Figure 4.8-1 shows the hydrologic processes and zones simulated by Model 8.

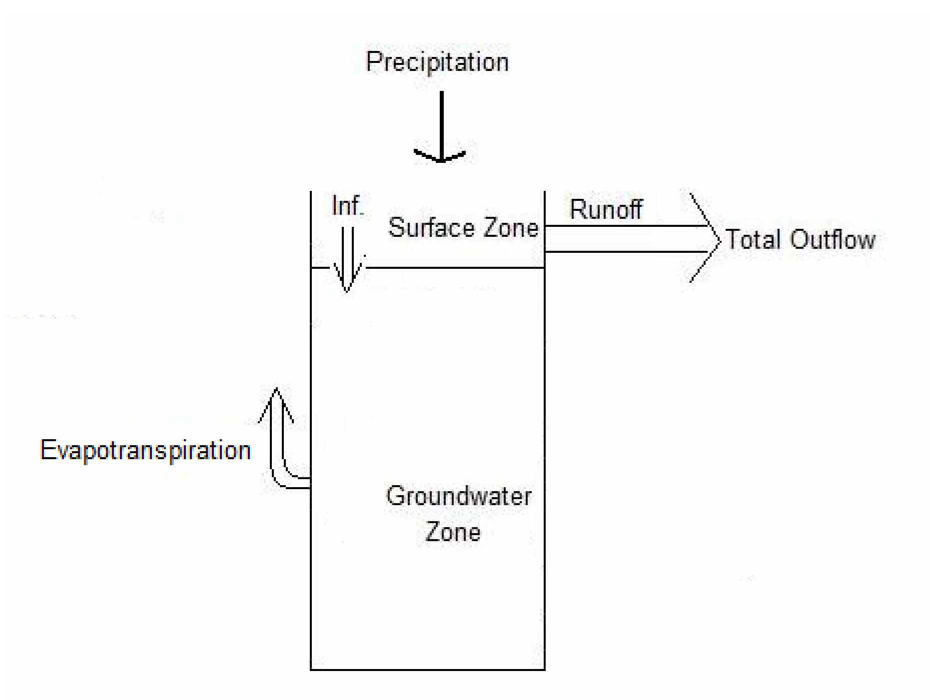

Figure 4.8-1. Flowchart of the Hydrologic Processes Simulated by Model 8

4.8.1 Calibration Process of Model 8

4.8.1.1 Calibration Run 8A

For Calibration Run 8A, the model was run using the parameters calibrated for Model 7D. The results are shown in Table 4.8-1. The elimination of PQV1 alters the distribution of water throughout the model. Eliminating the outflow from the groundwater increases the storage in the groundwater zone and, therefore, decreases the depth of water capable of infiltrating into the groundwater. With decreased infiltration, the surface storage increases, causing an increase in runoff from the surface. An increase in groundwater storage would also increase evapotranspiration, but evaporation is already at a maximum depth for each day. The transpiration, however, increased slightly.

Despite these changes in the distribution of water throughout the system, the overall runoff decreased because of the loss of groundwater outflow. This is apparent based on the negative runoff bias of 506 mm, from a runoff bias of -3 mm from Model 7D. The positive water balance bias also reflects the under prediction of runoff. Therefore, for Calibration Run 8B, parameter values were adjusted to increase the surface runoff.

Table 4.8-1. Goodness-of-Fit Statistics for Model 8A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	383	-325	0.48	0.88	-0.34
2	71	-59	0.12	0.997	-0.11
3	-108	-152	0.65	0.76	-0.26
4	23	30	0.58	0.82	0.07
Total	368	-506		0.83	-0.2

4.8.1.2 Calibration Run 8B

PSZ1 and PSZ2 are the surface parameters that control surface runoff. PSZ1 influences the magnitude of the runoff function, while PSZ2 controls the shape. By manipulating the values through multiple calibrations, it became apparent that increasing PSZ1 improved the runoff bias, but worsened the relative standard error. Decreasing PSZ2 improved the relative standard error but worsened the runoff bias. Therefore, both parameters were adjusted until a compromise between the two criteria was established. The results are shown in Table 4.8-2.

Compared to Model 7D, the water balance improved by 29 mm or about 7 mm/yr; however, the values for individual years are comparable. The runoff bias worsened by 52 mm or 13 mm/yr; however, again, the individual years were comparable. The relative standard error increased by 2%, while the relative bias worsened by 2%. While these changes suggest that the reduction of complexity provides equal goodness of fit, the parameter values were changed greatly: PSZ2 decreases from 0.9 to 0.65 and PSZ1 increased from 0.15 to 0.375. This suggests that the model parameter values are deviating from physically rational values and may affect the ability to apply the model to further data sets.

Table 4.8-2. Goodness-of-fit Statistics for Model 8B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	253	-191	0.52	0.86	-0.2
2	-30	52	0.14	0.99	0.09
3	-183	-60	0.66	0.76	-0.1
4	-36	108	0.52	0.86	0.25
Total	4	-92		0.83	-0.04

4.9 Further Reductions in Complexity

Further parameter reductions in complexity were conducted through three different approaches. The first approach, titled the sensitivity track, was based solely on the sensitivity analysis and ignores to some extent on physical model structure. In the second and third approach, titled the sensitivity processes track and processes track, respectively, eliminations were made with the purpose of comparing the importance of sensitivity and physical rationality. Physical rationality was represented by maintaining the three hydrologic processes: runoff, infiltration, and evaporation, as long as possible. Sensitivity was represented by keeping the parameter PSZ2 in the model, the most important parameter, as long as possible.

For the sensitivity track, parameter eliminations were made in the following order: (1) PSZ1; (2) PPET; (3) PEVZ and PISZ, titled Models 9, 10, and 11, respectively. While PPET is more significant than PEVZ and PISZ, eliminating PEVZ before PPET would be inefficient, as PEVZ is the only means of evaporating water. PPET influences only the magnitude of the function representing the maximum amount of evapotranspiration possible as opposed to the actual amount of water evaporated. Likewise, PISZ is more significant than PEVZ, but eliminating one at a time would be irrational because without a means to remove water from the groundwater zone, infiltrating water into the groundwater will only increase storage. The final model contains only the parameter PSZ2 and the surface runoff process.

In comparing sensitivity with the processes, both approaches eliminated PPET first, titled Model 12. The sensitivity processes track consisted of eliminations in the following order: (1) PSZ2 and (2) PEVZ and PISZ, titled Models 13 and 14, respectively.

The processes track consisted of eliminations in the following order: (1) PEVZ and PISZ and (2) PSZ2, titled Models 15 and 16, respectively. The final model for both approaches contains only the surface runoff parameter PSZ1. Figure 1 in Appendix C outlines the different elimination tracks. The results for all three approaches are discussed herein.

4.10 Sensitivity Track

4.10.1 Model 9

The next parameter was eliminated based on the sensitivity analysis of Model 7 was PSZ1, the surface runoff parameter that controlled the magnitude of the surface runoff equation. Because the surface runoff equation contains to parameters, PSZ1 was set equal to one to enable the surface runoff process to remain but be represented by a simpler equation. For Calibration Run 9A, the model was run using the optimized parameters from Model 8B. The results are shown in Table 4.10.1-1.

4.10.1.1 Calibration Run 9A

Eliminating PSZ1 increased the amount of surface runoff and created a negative water balance. Eliminating PSZ1 is the equivalent to setting PSZ1 equal to 1, which represents a large increase from 0.375, the optimized value from the previous model, i.e., 8D. Therefore, eliminating PSZ1 caused an increase of 1895 mm in surface runoff bias compared to Model 8B. The relative bias increased by 0.75 and the relative standard error increased by 2.27, which resulted in a correlation coefficient equal to zero for each of the four years.

Table 4.10.1-1. Goodness-of-Fit Statistics for Model 9A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-405	484	0	2.4	0.51

2	-469	536	0	3.9	0.98
3	-529	366	0	2.5	0.62
4	-426	600	0	4.9	1.41
Total	-1829	1987		3.1	0.79

4.10.1.2 Calibration Run 9B

To correct the runoff bias in Model 9A, the surface parameter PSZ2 was decreased from a value of 0.65 to 0.35. The results are shown in Table 4.10.1-2. Compared to Model 8B, the water balance worsened by 66 mm; the runoff bias improved by 65 mm; the relative standard error worsened by 4%; and the relative bias improved by 3%. However, the relative bias for the individual years worsened overall. To attain these results, however, the parameter PSZ2 was decreased by almost 50%, suggesting further deviation from physically rational parameter values.

Table 4.10.1-2. Goodness-of-Fit Statistics for Model 9B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	220	-158	0.43	0.91	-0.17
2	-72	98	0.17	0.99	0.18
3	-204	-35	0.57	0.82	-0.06
4	-45	122	0.42	0.91	0.29
Total	-100	27		0.87	0.01

4.10.2 Model 10

Based on the sensitivity analysis, PPET is the second most important parameter. However, the remaining less significant parameters, PEVZ and PISZ cannot be eliminated before PPET. Without PEVZ, the evaporation parameter from the groundwater, PPET lacks a purpose. Likewise, without PISZ, water cannot infiltrate into the groundwater to be evaporated. Therefore, for Model 10, the evaporation parameter PPET was eliminated. The elimination required PPET to be set equal to one because the concept of the maximum daily evaporation, PET, still exists, only the ability to adjust the magnitude, represented by PPET, was eliminated. The model was run based on the optimized parameter from Model 9B. The results are shown in Table 4.10.2-1.

4.10.2.1Calibration Run 10A

The elimination of PPET had little effect on the runoff bias, relative standard error, relative bias, or correlation coefficient. However, the bias in the water balance increased significantly because the maximum amount of allowable evapotranspiration decreased without the parameter PPET available to increase the mean of the function. With a large decrease in evaporation and insignificant compensation by the surface runoff, the rainfall is greater than the combined evaporation and runoff which caused a positive water balance bias.

Table 4.10.2-1. Goodness-of-Fit Statistics for Model 10A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	506	152	0.43	0.91	-0.16
2	243	115	0	1	0.21
3	144	-6	0.57	0.83	-0.01
4	246	164	0.3	0.96	0.39
Total	1139	121		0.88	0.05

4.10.2.2 Calibration Run 10B

To fix the water balance of Model 10A, evaporation or runoff must be increased. However, the runoff bias is already positive; which implied either that the predicted runoff must decrease to correct the runoff bias. Also, increasing PEVZ, the groundwater evaporation parameter, did not effect on the amount of evaporation simulated, suggesting that evaporation is already at a daily maximum. Therefore, the only option for calibration was to improve the runoff bias, as the water balance will always be positive with the maximum evaporation already attained. PSZ2 was decreased from 0.35 to 0.335 to decrease the total runoff and correct the bias. PISZ was increased from 0.275 to 0.3 to increase the infiltration of rainwater into the groundwater zone and, therefore, decrease runoff. The results are shown in Table 4.10.2-2. Compared to Model 9B, the runoff bias improved by 24 mm; the water balance bias worsened by 1156 mm; the relative standard error worsened by 1%; and the relative bias worsened by less than 1%, although the individual years worsened overall.

Table 4.10.2-2. Goodness-of-Fit Statistics for Model 10B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	548	-194	0.4	0.92	-0.2
2	271	86	0.19	0.98	0.16
3	167	-32	0.56	0.83	-0.05
4	270	138	0.38	0.93	0.32
Total	1256	-3		0.88	-0.001

4.10.3 Model 11

For Model 11, two of the three remaining parameters were eliminated: PEVZ and PISZ. It was necessary to eliminate both simultaneously because without infiltration, water can not enter the groundwater zone to evaporate. Likewise, without evaporation from the groundwater, infiltrated water can not be removed from the groundwater zone. Therefore, each parameter was set equal to zero, eliminating the evaporation and infiltration processes. This elimination creates a one-parameter model that simulates only precipitation and surface runoff with the parameter PSZ2.

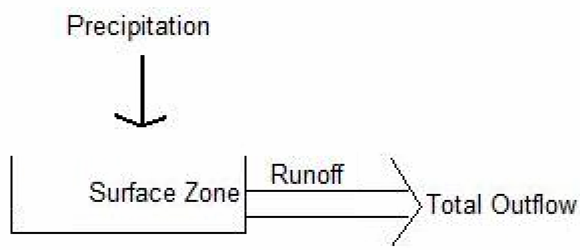

Figure 4.10-1. Flowchart of Hydrologic Processes and Zones Simulated by Model 11.

4.10.3.1 Calibration Run 11A

Model 11A was first run using the parameters calibrated from Model 10B. The results are shown in Table 4.10.3-1. Compared to Model 10, the water balance worsened by 887 mm; the runoff bias worsened by 2220 mm; the relative standard error worsened by 38%; and the relative bias worsened by 89%. This is because the process of infiltration and evaporation were completely eliminated, leaving only surface runoff as the outlet for precipitation. The value of remaining parameter, PSZ2, must be greatly adjusted to account for these changes.

Table 4.10.2-1. Goodness-of-Fit Statistics for Model 11A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-53	523	0	2.6	0.55
2	-110	608	0	1.5	1.1
3	-120	428	0	1.08	0.73
4	-86	663	0	1.75	1.56
Total	-369	2223		1.26	0.89

4.10.3.2 Calibration Run 11B

To improve the goodness-of-fit statistics, PSZ2 was decreased to reduce the surface runoff simulated. The calibrated value of PSZ2 decreased from 0.335 to 0.075. The results are shown in Table 4.10.3-2. Compared to Model 10B, the total runoff bias worsened by 43mm; the water balance bias worsened by 643 mm; the relative standard error worsened by 13%, with all correlation coefficients equaling zero; and the relative bias increased by less than 2%.

While the overall runoff bias and relative bias did not change significantly from Model 10B, the individual years experience drastic changes in both of these statistics. The runoff biases range from -382 mm to 220 mm for Model 11B. For Model 10B, the runoff biases range from -194 mm to 138 mm. This suggests a decrease in accuracy and is reflected in the poor relative standard error of the model. Likewise, the relative bias for individual years ranged from decreasing by 4% to increasing by 20% compared to Model 10B. Therefore, while the overall goodness-of-fit statistics may not reflect the biases that exist, the statistics for the individual years reveal the poor prediction accuracy of this one-parameter model.

Table 4.10.3-2. Goodness-of-Fit Statistics for Model 11B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	853	-382	0	1.08	-0.4
2	430	68	0	1.01	0.12
3	259	49	0	1.00	-0.08
4	358	220	0	1.08	0.52
Total	1899	-46		1.01	-0.02

4.11 Sensitivity vs. Hydrologic Processes

4.11.1 Model 12

For Model 12, the parameter PPET was eliminated in an attempt to maintain the rationality of the surface runoff equation. The elimination involved PPET being set equal to one to maintain the equation for the maximum daily evaporation rate, but remove the ability to adjust the magnitude of this rate. Then, the remaining model contained four parameters: PISZ, PEVZ, PSZ2, and PSZ1. With these parameters, the surface runoff, infiltration, and groundwater evaporation processes were simulated. The four parameters were calibrated and results compared to Model 8B.

4.11.1.1 Calibration Run 12A

Model 12A was run using the calibrated parameters from Model 8B. The results are shown in Table 4.11-1. Eliminating PPET limits the ability to adjust the magnitude of the function for daily allowable evapotranspiration. Evapotranspiration is still a cyclical function; however, the mean of the function can no longer be increased or decreased. This affects the water balance greatly as is apparent by Table 4.11-1. The elimination, however, has an insignificant effect on the runoff bias and remaining goodness-of-fit statistics. Compared to Model 8B, the water balance bias worsened by 1224 mm; the runoff bias improved by 66 mm; the relative standard error worsened by 1%; and the relative bias improved by less than 1%. Further calibrations were conducted to improve the goodness-of-fit statistics.

Table 4.11-1. Goodness-of-Fit Statistics for Model 12A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	538	-184	0.52	0.86	-0.19
2	282	74	0	1.01	0.13
3	161	-25	0.65	0.76	-0.04
4	247	161	0.39	0.93	0.38
Total	1228	26		0.84	0.01

4.11.1.2 Calibration Run 12B

The water balance is positive because the combined depth of runoff and evapotranspiration is less than the precipitation depth. However, the runoff bias is positive, suggesting that increasing the runoff simulated is unnecessary. Increasing the only remaining evaporation parameter, PEVZ, does not increase the amount of evaporation that occurs, suggesting that the model has reached the daily maximum allowable evaporation rate. Therefore, the only further improvement to the model is to decrease the runoff bias.

The runoff parameters, PSZ1 and PSZ2, were changed from 0.375 to 0.41 and 0.65 to 0.615, respectively. The results are shown in Table 4.11-2. Compared to Model 8B, the runoff bias improved by 81 mm; the water balance worsened by 1246 mm; the relative standard error worsened by 1%; and the relative bias improved by almost 4%, although the individual years worsened overall.

Table 4.11-2. Goodness-of-Fit Statistics for Model 12B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	546	-192	0.51	0.86	-0.2
2	286	70	0.1	0.999	0.13
3	164	-29	0.65	0.77	-0.05
4	253	154	0.44	0.9	0.36
Total	1250	3		0.84	0.001

4.12 Sensitivity Processes

4.12.1 Model 13

For Model 13, the parameter PSZ2 was eliminated by setting the value equal to one. PPSZ2 is the most sensitive parameter; however, the goal of approach 2 is to maintain the three hydrologic processes: runoff, infiltration, and evaporation, for as long as possible. This enables the comparison between the importance of sensitivity and rationality in prediction accuracy. The three-parameter model was calibrated and compared to Model 12B. Figure 4.12-1 shows the hydrologic processes and zones simulated by Model 13.

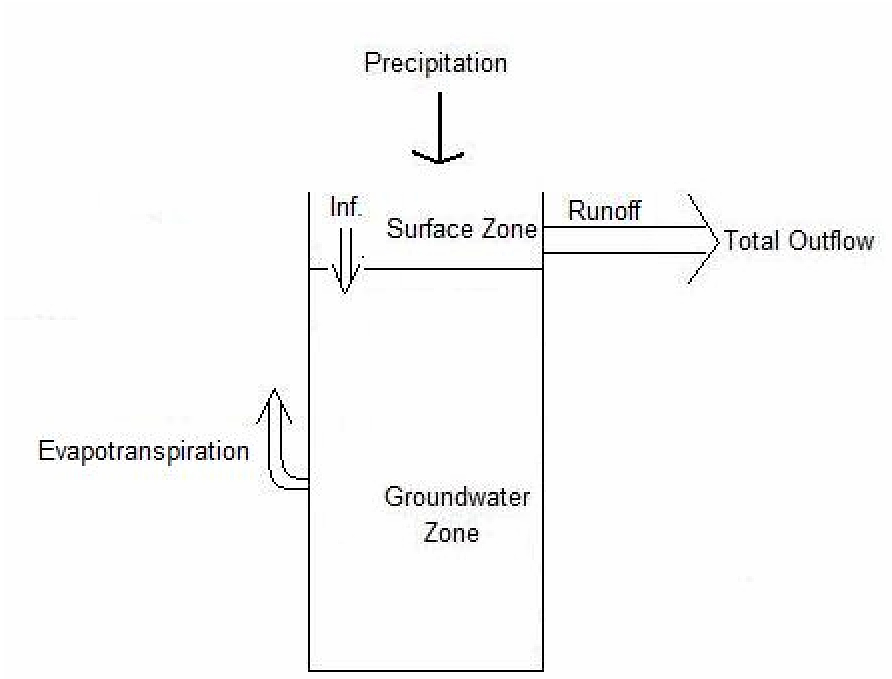

Figure 4.12-1. Flowchart of Hydrologic Processes and Zones Simulated by Model 13.

4.12.1.1 Calibration Run 13A

Model 13A was first run with the calibrated parameters from Model 8B. The results are shown in Table 4.12.1-1. Eliminating PSZ2 decreased the rationality and flexibility of the surface runoff equation. It is also the most sensitive parameter in the model. Therefore, eliminating PSZ2 caused runoff to increase greatly, as the exponent is essentially increasing from a value of 0.615 to 1 through elimination. This is apparent through the goodness-of-fit statistics. Compared to Model 8B, the water balance bias improved by 1155 mm; the runoff bias worsened by 1412 mm; the relative standard error worsened by 1.21%; and the relative bias worsened by almost 56%. Further adjustment was conducted to improve the prediction accuracy and compensate for the elimination of PSZ2.

Table 4.12.1-1. Goodness-of-Fit Statistics for Model 13A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	64	295	0	1.53	0.31
2	-29	398	0	2.62	0.72
3	-89	246	0	1.61	0.42
4	-42	476	0	3.42	1.12
Total	-95	1415		2.05	0.56

4.12.1.2 Calibration Run 13B

To decrease the runoff, the remaining surface runoff parameter, PSZ1, was decreased from 0.41 to 0.125. The results are shown in Table 4.12.1-2. Compared to Model 8B the water balance worsened by 10 m; runoff bias worsened by 8 mm; the relative standard error worsened by 8%; and the relative bias worsened by 0.005%. The correlation coefficient values for years 2 and 4 now equal zero as a result of 17% and 36% increase in the relative standard error for the respective years compared to Model 12B. Therefore, while the overall runoff bias and relative bias did not change much, the relative standard errors reflect a reduction in prediction accuracy as a result of eliminating a highly sensitive parameter. Also, the value of parameter PSZ1 decreased by almost 75%. This suggests that the parameter values may have deviated from physically rational values, resulting in a physically irrational model.

Table 4.12.1-2. Goodness-of-Fit Statistics for Model 13B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	539	-185	0.56	0.83	-0.19
2	299	57	0	1.17	0.1
3	173	-28	0.65	0.76	-0.07
4	230	178	0	1.26	0.42

Total	1240	11		0.92	0.005

4.12.3 Model 14

The evaporation and infiltration parameters, PEVZ and PISZ, were eliminated for the final model in Track 2. Both parameters were set equal to zero to remove the processes of evaporation and infiltration. The remaining one-parameter model contained the surface parameter, PSZ1. The model simulated only rainfall and surface runoff. Model 14 was calibrated and compared to Model 13B. Figure 4.12-1 shows the hydrologic processes and zones simulated by Model 14.

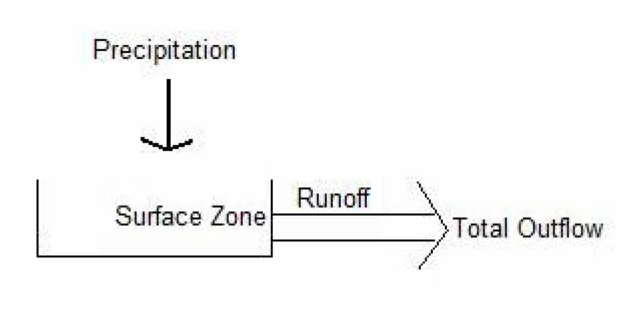

Figure 4.12-1. Flowchart of the Hydrologic Processes and Zones Simulated by Model 14

4.12.3.1 Calibration Run 14A

Eliminating the groundwater zone and processes left only surface runoff. This eliminated the entire physical processes of infiltration and evapotranspiration. Surface storage increased as the water is incapable of infiltrating. Surface runoff is a function of surface storage, resulting in an increase in surface runoff. This is shown in the goodness-of-fit statistics in Table 4.12.3-1. Compared to Model 13B, the water balance improved by 870 mm; the runoff bias worsened by 2214 mm; the relative standard error worsened by 46%; and the relative bias worsened by almost 89%. The correlation coefficients for all four years equaled zero. The remaining model parameter, PSZ1, was adjusted to improve the prediction accuracy.

Table 4.12.3-1. Goodness-of-Fit Statistics for Model 14A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-73	543	0	1.12	0.57
2	-97	596	0	1.78	1.08
3	-116	424	0	1.06	0.72
4	-84	661	0	2.17	1.56
Total	-370	2225		1.38	0.89

4.12.3.2 Calibration Run 14B

The water balance is impossible to correct, as the model is already at the maximum allowable daily evapotranspiration rate. Therefore, the runoff bias is the only improvement possible. PSZ1 was decreased from of 0.125 to 0.0012. The results are shown in Table 4.12.3-2. Compared to Model 13B, the water balance bias increased by 544 mm; the runoff bias worsened by 69 mm; the relative standard error worsened by 19%; and the relative bias worsened by 3%. Analyzing the goodness of fit for the individual years reflects even poorer prediction accuracy when compared with Model 13B. The runoff biases range from -660 mm to 442 mm for Model 14B; whereas for Model 13B, runoff biases only range from -185 to 178. Likewise the relative bias for Model 14B ranges from -0.60 to 1.04; for Model 13B, the relative standard bias only ranges from -0.19 to 0.42. Therefore, while the total biases may provide deceivingly accurate predictions, the individual years reflect the poor prediction accuracy of Model 14B.

Table 4.12.3-2. Goodness-of-Fit Statistics for Model 14B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y

1	1130	-660	0	1.23	-0.69
2	419	80	0	1.04	0.15
3	100	208	0	1.04	0.36
4	135	442	0	1.26	1.04
Total	1784	70		1.11	0.03

4.13 Hydrologic Processes

4.13.1 Model 15

The third approach focuses on maintaining the most sensitive parameter, PSZ2 as long as possible. Therefore, for Model 15, the groundwater infiltration and evaporation parameters, PISZ and PEVZ, respectively, were eliminated. Both parameters were set equal to zero to eliminate the process of infiltration and evaporation. The remaining model consisted of two parameters, PSZ1 and PSZ2, and simulated only rainfall and surface runoff. Model 15 was calibrated and compared to the results from Model 12B. Figure 4.13-1 shows the hydrologic zones and processes simulated by Model 15.

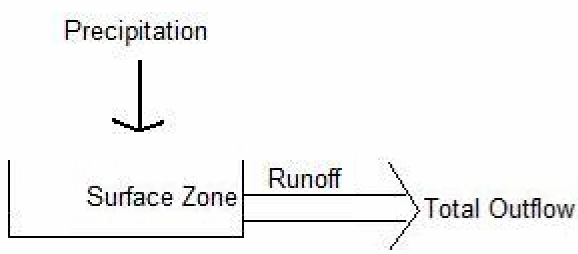

Figure 4.13-1. Flowchart of Hydrologic Processes and Zones Simulated by Model 15

4.13.1.1 Calibration Run 15A

Model 15 was run using the calibrated parameter values from Model 12B. The results are shown in Table 4.13.1-1. Eliminating the groundwater processes of infiltration and evaporation results in an increase in runoff, because runoff is the only outlet for the rainfall. As a result, the runoff bias increases greatly, which effects the remaining goodness-of-fit criteria. Compared to Model 12B, the runoff bias increased by 2224 mm; the water balance improved by 877 mm; the relative standard error increased by 43%; and the relative bias worsened by almost 89%. The correlation coefficient equals zero for all four years. Further calibration is necessary to improve the goodness of fit of the model.

Table 4.13.1-1. Goodness-of-Fit Statistics for Model 15A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-71	541	0	1.11	0.57
2	-101	599	0	1.6	1.09
3	-116	424	0	1.03	0.72
4	-86	663	0	1.84	1.56
Total	-373	2227		1.27	0.89

4.13.1.2 Calibration Run 15B

To improve the runoff bias, the runoff parameters, PSZ1 and PSZ2, were changed to 0.725 and 0.125, respectively. The results are shown in Table 4.13.1-2. Compared to Model 12B, the runoff bias increased by 25 mm; the water balance increased by 577 mm; the relative standard error increased by 18%; and the relative standard error increased by almost 1%. The correlation coefficients are now zero for each of the four years. The goodness-of-fit statistics for the individual years reveal much poorer prediction accuracy. The individual year runoff bias for Model 15B ranges from -396 to 255; whereas, for Model 12B, the runoff bias ranges from -192 to 154. Likewise, the relative bias ranges from -0.41 to 0.6 for Model 15B and only -0.2 to 0.36 for Model 12B. Therefore, while the overall goodness-of-fit may not seem to decrease much, the accuracy within the years has worsened with the elimination of the infiltration and evaporation processes.

Table 4.13.1-2. Goodness-of-Fit Statistics for Model 15B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	864	-396	0	1.08	-0.41
2	411	87	0	1.02	0.16
3	229	79	0	1.01	0.13
4	322	255	0	1.1	0.6
Total	1827	28		1.02	0.01

4.13.2 Model 16

The final model of the third approach eliminates PSZ2, the exponent parameter for the surface runoff equation. PSZ2 was set equal to one. The remaining model only consists of the surface parameter PSZ1 and only simulates rainfall and runoff, which is the same as Model 14. It represents a linear reservoir model. Model 16 was calibrated and compared to Model 15B.

4.13.2.1 Calibration Run 16A

Model 16 was first run with the calibrated parameters from Model 15B. The results are shown in Table 4.13.2-1. Compared to Model 15B, the water balance improved by 1,443; the runoff bias worsened by 2,210 mm; the relative standard error worsened by 236%; and the relative bias worsened by almost 89%. The surface storage is increasing because rainfall can not infiltrate into the groundwater. Surface runoff is a function of surface storage; therefore, an increase in surface storage causes an increase in runoff resulting in the drastic increase in positive runoff bias. Further calibration is necessary to correct the runoff bias.

Table 4.13.2-1. Goodness-of-Fit Statistics for Model 16A

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	-83	553	0	2.57	0.58

2	-97	595	0	4.28	1.08
3	-112	420	0	2.72	0.72
4	-92	670	0	3.38	0.89
Total	-384	2238		3.38	0.89

4.13.2.2 Calibration Run 16B

To correct the runoff bias, the runoff parameter, PSZ1, was decreased from 0.75 to 0.0012. The results are shown in Table 4.13.2-2. Compared to Model 15B, the runoff bias increased by 42 mm; the water balance decreased by 43 mm; the relative standard error increased by 9%; and the relative bias increased by 2%. However, the goodness-of-fit statistics for the individual years show much poorer prediction accuracy. For Model 16B, the runoff bias ranges from -660 to 442; whereas in Model 14B, the bias ranges from -396 to 255 mm. Likewise, the relative bias ranges from -0.69 to 1.04 for Model 16B and from -0.41 to 0.6 for Model 15B. Therefore, while the overall goodness-of-fit may not suggest a great reduction in prediction accuracy by eliminating the most sensitive parameter, the individual years reflect the inaccuracy. Also, the value of PISZ decreased significantly, suggesting a deviation from the rational parameter value calibrated in Model 1. This affects the physical rationality of the model.

Table 4.13.2-2. Goodness-of-Fit Statistics for Model 16B

Year	WB (mm)	Runoff Bias (mm)	R	Se/Sy	e/y
1	1130	-660	0	1.23	-0.69
2	419	80	0	1.04	0.15
3	100	208	0	1.04	0.36
4	135	442	0	1.26	1.04
Total	1784	70		1.11	0.03

CHAPTER 5

ANALYSES AND RESULTS

Once the model was calibrated, it could be used to study other factors about complexity. The following issues were studied: the effect of the lack of correlation between rainfall and runoff, the effect of calibrating to provide a reasonable fit to all flows versus fitting to optimize predictions of selected events such as high flows or low flows, the value of sensitivity analyses in calibration, the importance of assessing the hydrologic rationality of the optimized parameters, and the ability of incomplete data sets, such as water quality grab samples, to represent the population. The results for each analysis as well as the calibration of the models varying complexity are discussed herein.

5.1 Results of Model Structure Complexity Simplifications

Objective One aimed to examine the relationship between model structure complexity and prediction accuracy. Through the calibration of the 16 models, each with varying complexity, it was apparent that model structure complexity does effect goodness of fit. The goodness-of-fit statistics shown in Chapter 4 for each calibrated model reflected the change in prediction accuracy with decreasing complexity.

Based on the goodness-of-fit statistics for each of the models, it is apparent that the runoff bias was near zero regardless of the model structure complexity. However, the relative standard error did show significant changes. Figure 5.1-1 shows the change in the relative standard error based on the rank of eliminations in the Sensitivity I and the Sensitivity II track. The Sensitivity I track consisted of a final model containing only one parameter, PSZ1, the scale parameter of the runoff equation. The Sensitivity II track consisted of a final model containing only one parameter, PSZ2, the shape or exponent parameter of the runoff equation. Based on the graph, it is apparent that the simplifications ranging from 1 to 8 eliminations for both tracks do not show as significant change in the relative standard error. However, beyond 8 eliminations, significant changes occurred. It is important to not that the final model for the Sensitivity II track has poorer prediction accuracy than that of the Sensitivity I track. This suggests that the shape parameter is more important to prediction accuracy than the scale parameter.

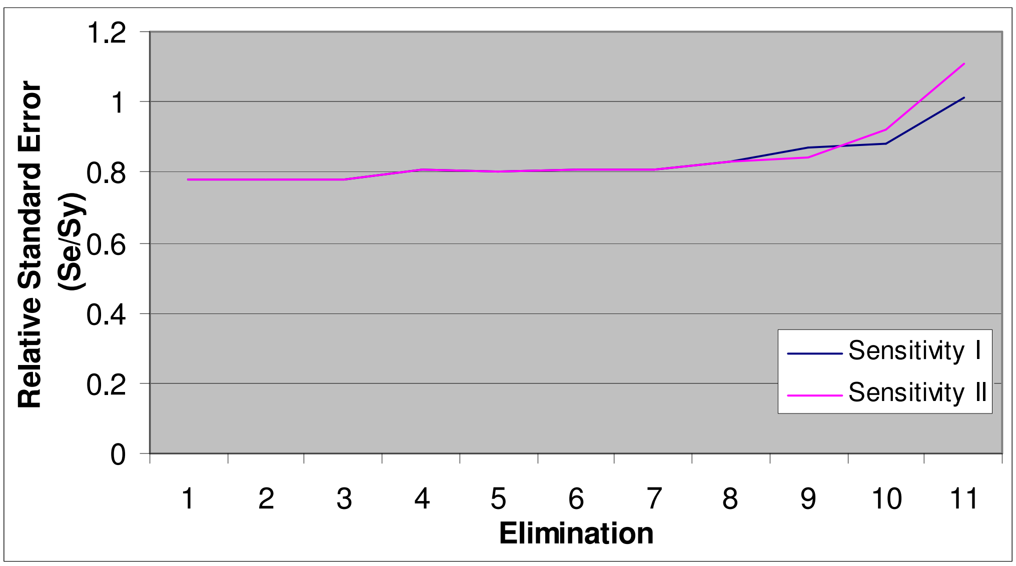

Figure 5.1-1. Change in the Relative Standard Error based on Rank of Eliminations for the Sensitivity I and the Sensitivity II Paths.

Figure 5.1-2 shows the change in the relative standard error based on the rank of eliminations in the Sensitivity II and Hydrologic Processes or Physical Rationality paths. As in Figure 5.1-1, the first 8 eliminations do not reveal a significant change in prediction accuracy. However, beyond 8 eliminations, significant changes occur. The tenth elimination reveals the difference in maintaining the most sensitive parameter in the model, PSZ2, versus the main physical processes, evaporation and infiltration. It is apparent from the graph that eliminating the hydrologic processes causes a greater increase in the relative standard error, implying poorer accuracy in predictions. This suggests that it is better to maintain important physical processes than the most important parameter for prediction accuracy.

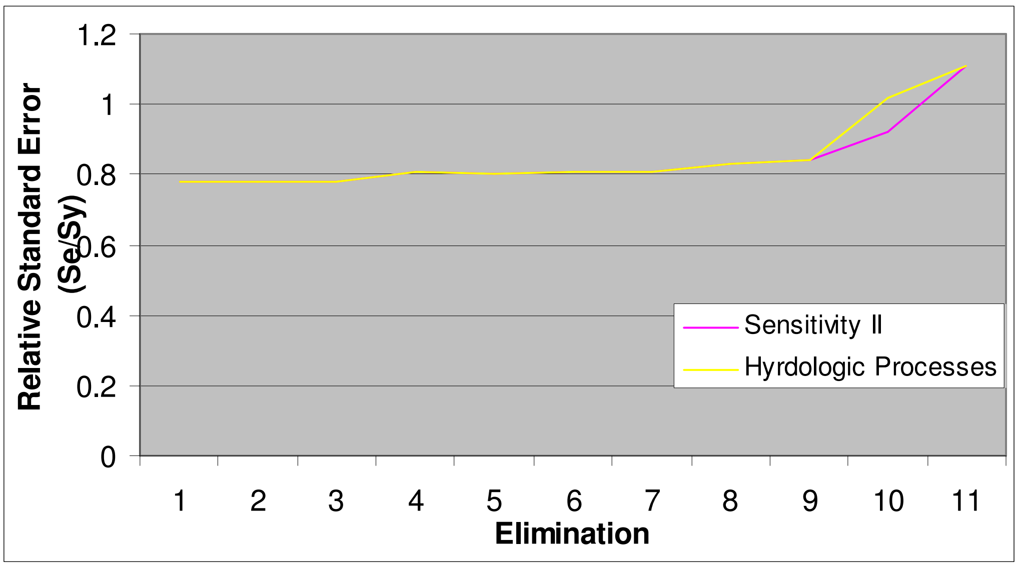

Figure 5.1-2. Change in the Relative Standard Error based on Rank of Eliminations for the Sensitivity II and the Hydrologic Processes Paths.

Table 5.1-1 shows the parameter values for each of the calibrated models. It is apparent that changes were insignificant for the beginning eliminations, but increased as the complexity of the model structure decreased. This is because models containing fewer parameters most likely contain parameters with higher sensitivities than models with more parameters. Therefore, as model complexity decreased, the importance of the remaining parameters increased as well. Therefore, further eliminations required adjustments to the values of remaining parameters to compensate for the elimination of sensitive parameters. It is also important to note that the need for parameter value changes may result in values deviating from physically rational values. Knowledge of parameter rationality is important to ensure the model does not provide irrational results.

Table 5.1-1. Parameter Values for Models 1 through 16.

Model	PEXI	PXI	PSZ1	PSZ2	PISZ	PESZ	PIRZ	PERZ	PQV1	PQV2	PIVZ	PEVZ	PQGZ	PPET
1	0.5	0.5	0.15	0.9	0.2	150	15	7.5	0.004	0.00375	0.6	8	0.000024	3.9
2			0.15	0.9	0.2	150	15	7.5	0.004	0.00375	0.6	8	0.000024	3.9
3			0.15	0.9	0.2	150	15		0.004	0.00375	0.6	9	0.000024	3.8
4			0.15	0.9	0.2	150	15		0.004		0.6	8	0.000024	4
5			0.15	0.9	0.275		15		0.004		0.6	8	0.000024	4.1
6			0.15	0.9	0.275		15		0.0004			8	0.000024	3.25
7			0.125	0.9	0.275				0.0004			8	0.000024	2.5
8			0.375	0.65	0.275							8		2.5
9				0.35	0.275							8		2.5
10				0.335	0.3							8		
11				0.075										
12			0.41	0.615	0.275							8		2.5
13			0.125		0.275							8		
14			0.0012											
15			0.725	0.125										
16			0.0012											

5.2 Effect of Rainfall Misalignment

Issues outside of model development can play an important role in the prediction accuracy capabilities of a model. In some cases, the complexity is not the cause of poor prediction accuracy, but errors existing outside of the model. For example, the process of collecting data used for calibration can contain human error. Also, choosing data recorded from rainfall and runoff gauges located some distance apart may result in time series of rainfall and runoff that is not correlated. Data issues must be considered in order to fully understand the prediction capabilities of a model.

While the reductions in model complexity in this research reveal decreases in goodness of fit, the relative standard error, correlation coefficients, and relative bias for even Model 1 do not reveal good prediction accuracy. The values for the relative standard error and relative bias are as high as 0.96 and 0.31, respectively, and as low as

0.32 for the correlation coefficient. An analysis was needed to assess the potential effect of poor cross correlation influencing the inability of the model to achieve high prediction accuracy. The data were analyzed to explain the poor overall goodness of fit despite the calibration process and high model complexity.

In basic correlation analysis, large data values can exert a large influence on a computed statistics. For example, the correlation coefficients and relative standard error are greatly influenced by sample points of high magnitude. A single large predicted or measured peak discharge can distort goodness-of-fit statistics. Because of the poor goodness of fit and obvious differences between computed and measured peak discharge rates, rainfall and runoff on days of high discharges were analyzed individually. Upon comparing the rainfall records with the runoff records, it was apparent that some large rainfall depths occurred on days different from the corresponding peak runoffs. For example, a rainfall might occur the day before or the day after the measured peak runoff. An example of this is shown in Figure 5.2-1.

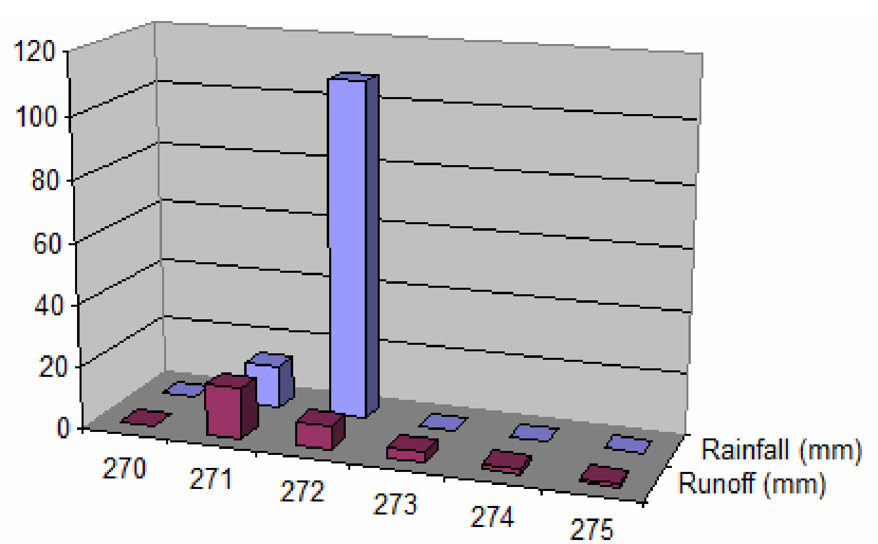

Figure 5.2-1. Graph of Misaligned Rainfall and Runoff Peak.

Rainfall and runoff misalignment is most likely the result of the distance between the locations of the rain and runoff gauge and the path and velocity of the storm. The peak of the storm may cause measured rainfall to occur before the storm moves over the watershed, or the rainfall may occur over the watershed before the storm moves over the rain gauge. This can introduce a source of variation into the data for which the model is not able to compensate. Additionally, a storm event may occur in one portion of the watershed, but not at the rain gauge. Likewise, the storm event may occur late in the evening on the watershed and reach the gauge in early morning, resulting in misaligned records, with the rainfall appearing to occur on the day before the storm runoff. As a result, the model will predict a peak runoff the day after the storm event, causing poor prediction accuracy.

To evaluate the potential effect of the problem, 12 misaligned rainfall events throughout the four years of data, i.e., 0.8% of the data record and 25% of all the peaks, were shifted to occur on the day of the corresponding runoff. Figure 5.2-2 graphs the adjustment of the rainfall and runoff that was shown in Figure 5.2-1. In all cases, the day of the rainfall was moved, not the runoff record. The rainfall depths switched so that rainfall was not lost. The changes are documented in Table 5.2-1. The total rainfall volume remained unchanged. The runoff and adjusted rainfall time series were then used as input to the model using the optimum calibrated parameter values. The results are shown in Table 5.2-2.

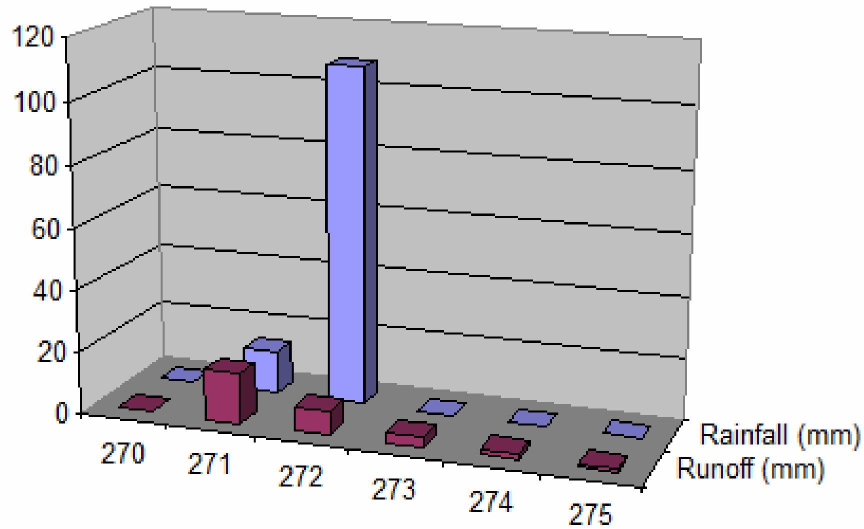

Figure 5.2-2. Adjusted Rainfall and Runoff from Figure 5.1-2.

The goodness-of-fit statistics with the altered rainfall series were compared with the statistics for the calibrated Model 1. The relative standard error decreased by 5%, while the individual years experienced decreases ranging from 1% to 13%. The correlation coefficients for the individual years increased by 1%, 25%, 4%, and 9%. The relative bias increased overall by 0.1%, which is an insignificant change. However, the individual years remained unchanged for years 1 and 2, but decreased by 2% and 5% for years 3 and 4, respectively. The water balance and runoff bias remained unchanged, which would be expected considering the total rainfall volume was not affected.

In addition to the goodness-of-fit statistics, the peak prediction accuracy increased as a result of the altered rainfall data. Peak prediction accuracy is measured by calculating the difference between the measured and predicted peaks for each month. A total bias was calculated for the four years. Figures 5.2-1 and 5.2-2 show the graphs of the measured peaks versus the predicted peaks for Model 1 with unaltered rainfall data and altered rainfall data, respectively. Figure 5.2-1 reveals that Model 1 underestimates the peak discharge rates causing a negative bias. This is apparent because the majority of the values fall below the 45-degree line. The graph for the altered rainfall data better distributes the values both above and below the 45-degree line, which indicates that the peaks are better predicted and neither a positive nor a negative bias exists. These observations are supported by the model output, which reveals a decrease in peak bias by almost 25%, from -2.6 mm to -2.0 mm.

The results support the hypothesis that the misaligned rainfall and runoff observations can introduce significant inaccuracy in overall predictions, regardless of the prediction capabilities of the model. Altering less than 1% of the rainfall data, or 25% of the peaks, resulted in significant improvements in the goodness-of-fit statistics. As a result of these findings, it is important to consider data collection when assessing the accuracy of a model. It may be useful to incorporate readings from more than one rain gauge, by taking an average value, in order to account for a lack of uniformity of rain events or poor timing of events over the watershed.

Table 5.2-1. Observed and Adjusted Rainfall Data for Data Complexity Analysis

	Observed			Adjusted		
Year	**Day**	**Rainfall (mm)**	**Runoff (mm)**	**Day**	**Rainfall (mm)**	**Runoff (mm)**
1	172	17.780	9.025	172	5.080	9.025
	173	5.080	10.121	173	17.780	10.121
	323	1.016	7.844	323	45.720	7.844
	325	45.720	5.820	325	1.016	5.820
2	37	13.462	8.434	37	20.320	8.434
	38	20.320	4.723	38	13.462	4.723
	102	38.608	4.554	102	13.716	4.554
	103	13.716	11.217	103	38.608	11.217
	121	5.080	1.687	121	0.000	1.687
	122	25.400	5.988	122	5.080	5.988
	123	0.000	8.772	123	25.400	8.772
	271	13.716	16.362	271	109.220	16.362
	272	109.220	7.506	272	13.716	7.506
	331	27.178	0.675	331	0.000	0.675
	332	0.000	5.482	332	27.178	5.482
3	91	3.810	3.880	91	0.000	3.880
	92	43.180	11.386	92	3.810	11.386
	93	0.000	12.483	93	43.180	12.483
	333	6.604	17.290	333	61.976	17.290
	334	61.976	13.579	334	6.604	13.579
4	18	3.048	3.964	18	10.922	3.964
	19	10.922	3.289	19	3.048	3.289
	244	134.112	2.952	244	5.334	2.952
	245	5.334	10.543	245	134.112	10.543
	320	0.000	8.603	320	40.640	8.603
	321	40.640	5.820	321	0.000	5.820

Table 5.2-2. Goodness-of-Fit Statistics for the Adjusted Rainfall Data

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	203	-178	0.63	0.79	-0.19
2	-58	95	0.57	0.83	0.17
3	-182	-26	0.75	0.67	-0.05
4	-26	112	0.67	0.76	0.26
Total	-63	3		0.73	0.002

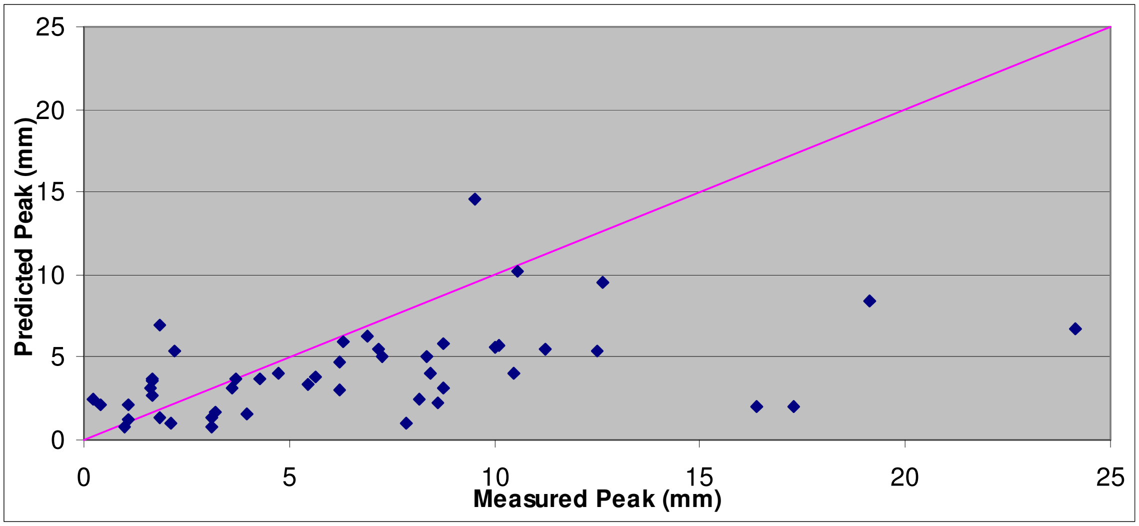

Figure 5.2-1. Peak Accuracy for Model 1

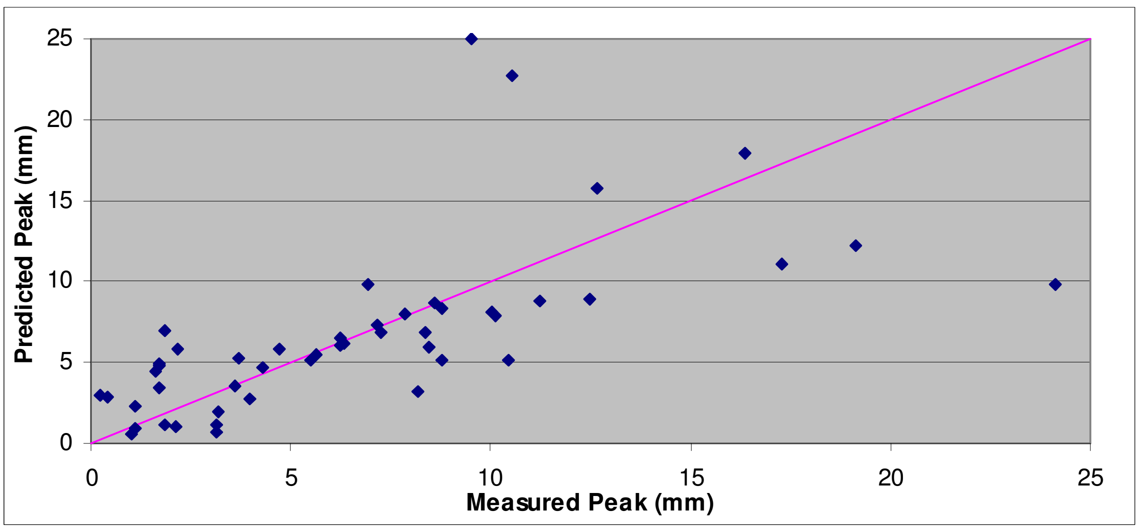

Figure 5.2-2. Peak Accuracy for Model 1 with Altered Rainfall Data

5.3 Optimization for Design Criteria

Modeling is beneficial for engineering design purposes such as flood control as well as for water quality issues to determine pollutant loads. These models often have required specific modeling criteria. For example, accurate peak prediction is required for flood control, whereas for water quality estimates, it is important to accurately predict base flows so as to avoid biased predictions of pollutant concentrations. These separate criteria are necessary to develop flood management designs as well as to address water quality regulations. However, calibrating a model to accurately fit one criterion may have negative effects on the overall accuracy of the model and provide misleading results. Additionally, a model calibrated with an emphasis on fitting either peaks or baseflows will likely not yield a model that accurately reflects watershed processes not associated with the peaks or baseflows. Parameters that reflect the other watershed processes will be disturbed.

To determine the effects of modeling for a specific design criteria, a comparison was conducted using two different criteria to calibrate Model 1, the 14-parameter model. Model 1 was recalibrated first, to eliminate the overall peak prediction bias and second, to eliminate the base flow bias in order to determine the effects of focusing on the main design criteria rather than the overall goodness-of-fit. The results for the specialty calibrations were compared to accuracy when calibrating to optimize the overall goodness of fit (see Table 5.3-1). The results will indicate the importance of considering all aspects of the calibration process and multiple criteria in determining parameter values that provide the best prediction accuracy. The hypothesis being tested is that calibrating to provide accurate estimates of either peak flows or base flow will distort the other parameters and then the ability to reflect all physical processes. The calibration results for the two criteria are discussed and compared herein.

Table 5.3-1. Goodness of Fit Statistics for Model One

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	203	-178	0.62	0.8	-0.19
2	-57	95	0.32	0.96	0.17
3	-181	-26	0.71	0.72	-0.07
4	-27	112	0.58	0.83	0.31
Total	-63	3		0.78	0.001

5.3.1 Calibration for Peak Discharge Prediction Accuracy

The model was first calibrated to accurately fit the monthly peak flows. The model identifies the greatest runoff flow of each month and provides a comparison of the measured and the predicted for the day. Peak prediction accuracy was analyzed based on the bias between the predicted and measured flow volume. The goal was to minimize the yearly average peak prediction bias and attain an overall bias for the four-year period equal to zero. The parameter values for the model optimized for peak flows are shown in Table 5.3-2. The surface parameters PSZ1 and PSZ2 were increased by 5% and 20%, respectively, in order to optimally fit the peak flows. The peak prediction biases are shown in Table 5.3-3 for both Model 1 and the model optimized for the peak discharge. The average peak prediction bias was reduced by 71%, 57%, 78%, and 60 % for years 1, 2, 3, and 4, respectively. The overall peak prediction bias for all four years was reduced by 77%, from -2.6 mm to -0.06 mm, through the optimization process, proving to be a better fit for the peak discharge rates.

Table 5.3-2. Parameter Values for Peak Discharge Optimization

PEXI	PXI (mm)	PSZ1	PSZ2	PISZ	PESZ (C^{-1})	PIRZ (mm)	PERZ (mm)	PQV1 (mm/hr)	PQV2 (mm/hr)	PIVZ (mm)	PEVZ	PQGZ (mm/hr)	PPET

0.5	0.5	0.2	1.08	0.2	150	15	7.5	0.004	0.00375	0.6	8	0.000024	3.9

Table 5.3-3. Peak Prediction Bias for Model 1 and the Model Optimized for Peaks

		Peak Prediction Bias				
Model		**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Average**
Model 1	High	2.03	5.07	2.26	5	
	Low	-17.36	-14.39	-5.24	-6.31	
	Difference	19.39	19.46	7.5	11.31	
	Average	-4.19	-2.92	-3.25	-0.05	-2.603
Optimized for Peaks	High	3.91	8.23	8.15	24.16	
	Low	-11.83	-13.57	-15.3	-6.56	
	Difference	15.74	21.8	23.45	30.72	
	Average	-1.21	-1.27	-0.71	2.97	-0.055

Figures 5.3-1 and 5.3-2 show the measured peaks versus the predicted peaks for Model 1 and the model calibrated for the peak flows, respectively. It is apparent through observation of Figure 5.3-1 that Model 1 underestimates the peaks fairly consistently. More than 80% of the observed peaks are underestimated. This is supported by the average bias equal to -2.603 mm. However, the graph of the model optimized to fit the peak flows (see Figure 5.3-2) distributes the values both above and below the 45-degree line, suggesting that neither a positive nor negative bias exists. This is also supported by the -0.05 mm bias calculated by the model. Both figures reveal errors in prediction; however, the model optimized for peak flows distributes the errors to avoid bias in predictions of peaks.

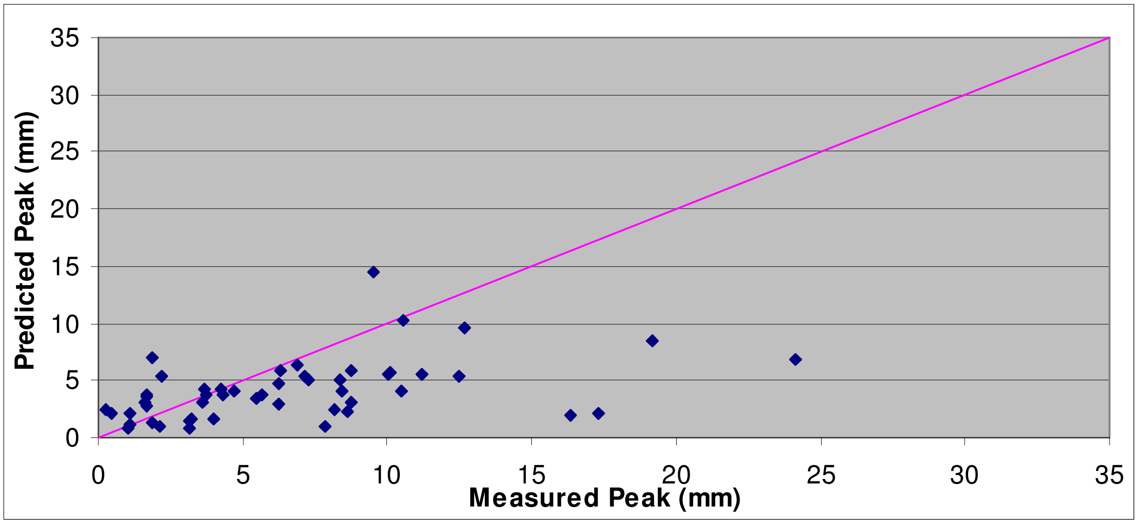

Figure 5.3-1. Peak Prediction Accuracy for Model 1

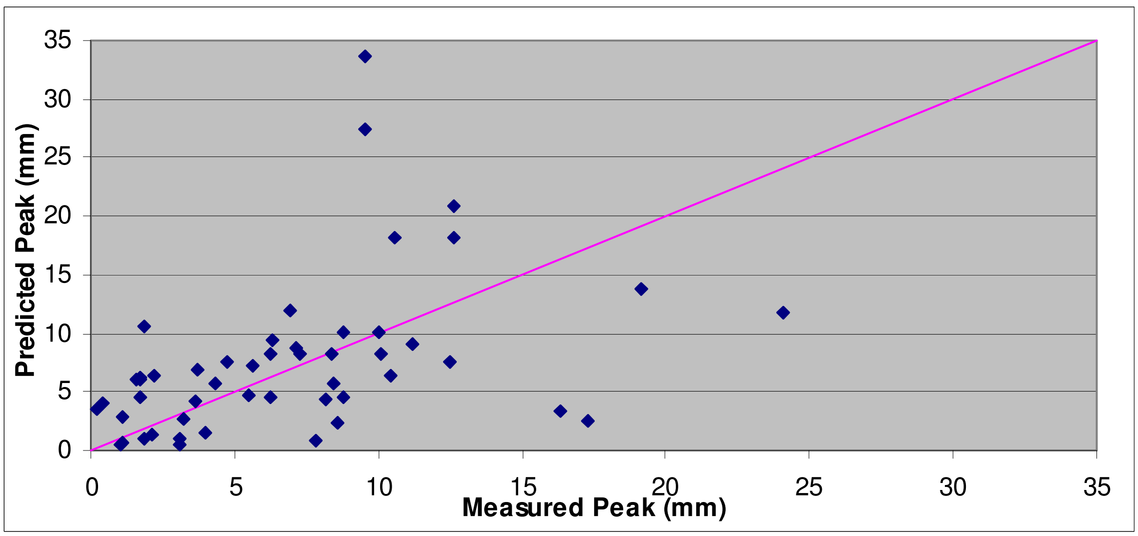

Figure 5.3-2. Peak Prediction Accuracy for Peak Optimization Model

The goodness-of-fit statistics for the model optimized for the peaks are shown in Table 5.3-4. Based on the goodness-of-fit statistics, the overall relative standard error increased by 35%; the relative bias increased by almost 18%; the runoff bias worsened by 452 mm; and the water balance worsened by 363 mm compared to Model 1. The correlation coefficients for years 2 and 4 were reduced to zero while years 1 and 3 decreased by 14% and 19%, respectively.

Table 5.3-4. Goodness-of-Fit Criteria for the Model Optimized for Peaks

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	41	13	0.48	0.89	0.014
2	-119	187	0	1.5	0.34
3	-203	37	0.52	0.87	0.06
4	-85	217	0	1.85	0.51
Total	-366	455		1.13	0.18

These results show that optimizing the model parameters to provide a best fit of the peak discharge rates can cause poor accuracy of the remaining prediction criteria. Therefore, while an in experienced modeler may choose to focus on calibrating for accurate peak predictions to meet the demands of engineering purposes such as storm water management design, calibrating for only one criterion will most likely result in poor overall prediction accuracy.

5.3.2 Calibration for Base Flows Prediction Accuracy

The model was then calibrated to provide accurate predictions of the base flows. Similar to the peak flow analysis, measured base flows were identified monthly and compared to the predicted base flow for the respective day. The accuracy of prediction was measured based on the bias between the predicted and observed flows. The average bias for each year as well as the total bias for the four years of data were calculated. The parameter values for the model optimized for base flows are shown in Table 5.3-5. The outflow parameters PQGZ for the groundwater zone and PQV1 for the vadose zone were decreased by 85% and 99%, respectively, to optimally fit the peaks. The results are shown in Table 5.3-6. Optimizing for the base flows resulted in a decrease of 80%, 78%, 80%, and 35% for the Year 1, 2, 3, and 4 biases, respectively. The average bias for the four years of data was decreased by 98%, from 0.363 to -0.007 compared to Model 1.

Table 5.3-5. Parameter Values for Model Optimized for Base Flows

PEXI	PXI (mm)	PSZ1	PSZ2	PISZ	PESZ (C^{-1})	PIRZ (mm)	PERZ (mm)	PQV1 (mm/hr)	PQV2 (mm/hr)	PIVZ (mm)	PEVZ	PQGZ (mm/hr)	PPET
0.5	0.5	0.2	0.9	0.2	150	15	7.5	0.00005	0.00375	0.6	8	0.0000036	3.9

Table 5.3-6. Base Flow Prediction Results
for Model 1 and the Model Optimized for Base Flow

		Base Flow Prediction Bias				
Model		**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Average**
Model 1	High	4.110	0.920	1.560	0.660	
	Low	-0.640	0.000	-0.920	-0.130	
	Difference	4.750	0.920	2.480	0.790	
	Average	0.460	0.380	0.430	0.180	0.363
Optimized for Base Flow	High	4.520	0.413	1.782	0.571	
	Low	-1.046	-0.382	-0.854	-0.420	
	Difference	5.566	0.795	2.636	0.991	
	Average	0.090	-0.085	0.086	-0.117	-0.007

Figures 5.3-3 and 5.3-4 graph the measured and predicted base flows for Model 1 and the model calibrated to accurately predict base flows. The graph for Model 1 reveals a positive bias in base flow prediction, as about 80% of the values are located above the 45-degree line. The model calibrated to accurately predict the base flows does not contain an overall bias, but instead distributes the values more evenly around the 45-degree line. Therefore, the model calibrated for base flows effectively reduces the bias in base flow prediction.

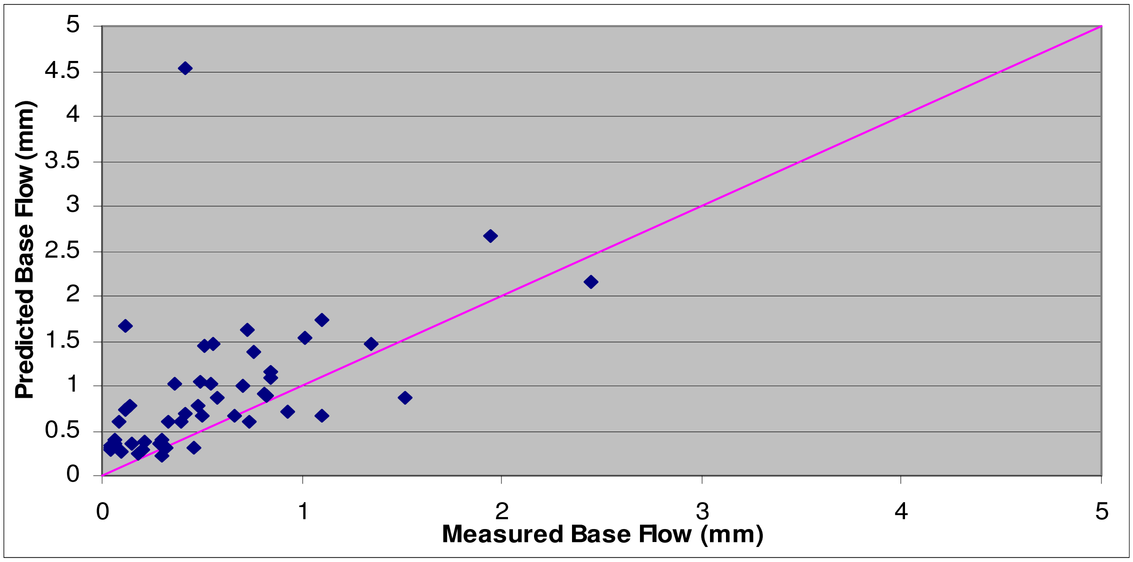

Figure 5.3-3. Base Flow Prediction Accuracy for Model 1

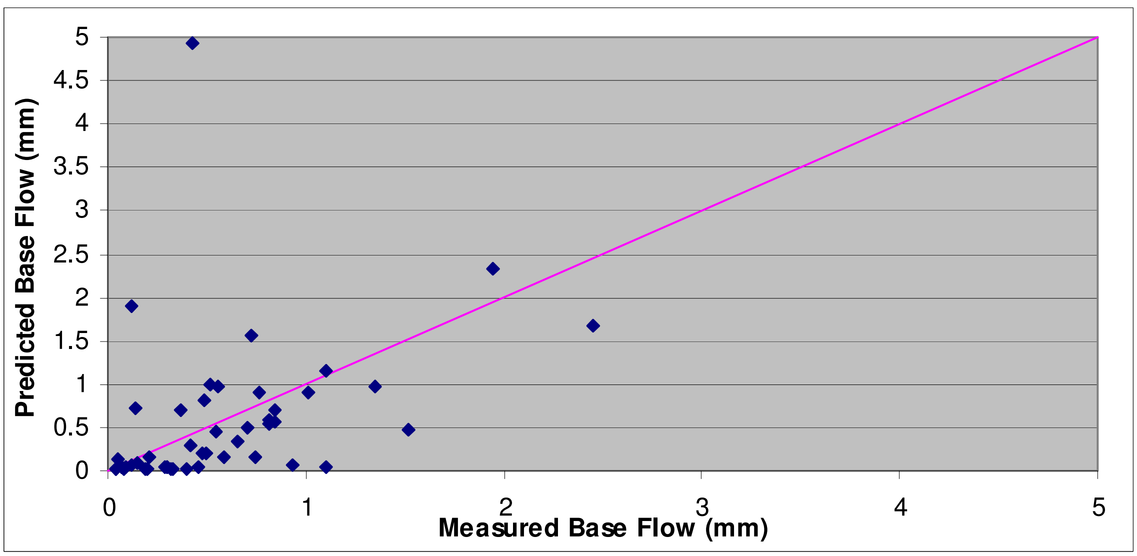

Figure 5.3-4. Base Flow Prediction Accuracy for the Model Optimized for Base Flows

While prediction of the base flows improved, the remaining goodness-of-fit criteria worsened as a result of calibrating for the base flows. The goodness-of-fit results are shown in Table 5.3-5. Compared to Model 1, calibrating to fit the base flows increased the water balance bias by 188 mm; increased the runoff bias by 283 mm; increased the relative standard error by 8%; and increased the relative bias by almost 11%. The correlation coefficients for years 2 and 4 were reduced to zero and years 1 and

3 decreased by 6% and 4%, respectively. Therefore, by calibrating for base flow prediction accuracy alone, the remaining goodness-of-fit criteria worsen compared to Model 1, resulting in overall poor prediction accuracy.

Table 5.3-5. Goodness-of-Fit Results for Baseflow Optimization

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	288	-251	0.56	0.84	-0.26
2	51	-6	0	1.1	-0.01
3	-100	-104	0.67	0.76	-0.18
4	12	74	0	1.06	0.17
Total	251	-286		0.86	-0.11

5.4 The Importance of Parametric Sensitivity

Knowing the sensitivity of model parameters is central to model calibration. It is hypothesized that sensitive parameters cause greater variation in the degree of model fitting than the less sensitive parameters. Insensitive parameters may not approach a value that reasonably reflects a value appropriate for the hydrologic process. It is difficult to support the idea that the value of an insensitive parameter reflects the hydrologic effect of that parameter. For example, if parameter X is insensitive, then variation of the parameter about the optimized value can not be expected to change the goodness of fit and secondly, we will not know that the optimized value reflects the physical processes.

Complex models contain multiple parameters, each with varying levels of importance. The sensitivity of the model output to changes in the parameter value is an indication of its importance. Less sensitive parameters have less influence on the goodness of fit as opposed to highly sensitive parameters. As stated previously in the sensitivity analysis of the 14-parameter model, the model predictions are most sensitive to the surface parameters and less sensitive to the parameters from the interception, vadose, and saturated zones. An analysis was conducted to show the effects of varying parameters with high sensitivity versus low sensitivity.

The sensitivity analysis of the calibrated Model 1 showed that the surface storage parameter PSZ2 was sensitive while PIVZ was relatively insensitive. Therefore, these were chosen as the parameters used to demonstrate the importance of understanding parametric sensitivities. The surface parameter, PSZ2, was selected as the high sensitivity parameter in the analysis. The infiltration parameter, PIVZ, from the vadose zone was chosen as the low sensitivity parameter. Each of the parameter values were changed to a near zero value of 0.1 and to 100% of the calibrated value. The parameter values were adjusted to 0.1 rather than 0 because the goal was not to eliminate the parameter from the model, but to determine the effects of varying the values. The results were compared to the goodness-of-fit statistics of Model 1, shown in Table 5.4-1.

Table 5.4-1. Goodness of Fit Statistics for Model One

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	203	-178	0.62	0.8	-0.19
2	-57	95	0.32	0.96	0.17
3	-181	-26	0.71	0.72	-0.07
4	-27	112	0.58	0.83	0.31
Total	-63	3		0.78	0.001

5.4.1 High Parameter Sensitivity

The optimal value of PSZ2 was calibrated as 0.9 in the development of Model 1. Thus, the extreme values tested were 0.1 to 1.8. The results for the high sensitivity parameter adjustments are shown in Tables 5.4-2 and 5.4-3. Table 5.4-2 shows the results for the 100% increase in PSZ2 and Table 5.4-3 shows the results for the near zero value of PSZ2. Increasing the value by 100% caused a 363 mm increase in the water balance bias; a 744 mm increase in the runoff bias; a 25% increase in the relative standard error, and almost a 3% increase in the relative bias. The correlation coefficients for years 1, 3, and 4 were reduced to 0 and year 2 was reduced by 11%. Reducing the value to be near zero increased the water balance bias by 1568 mm; increased the runoff bias by 2321 mm; increased the relative standard error by 374%; and increased the relative bias by almost 93%. The correlation coefficients for each of the four years were reduced to zero.

Table 5.4-2. Goodness-of-Fit Statistics for Sensitive Parameter with a 100% increase from Calibrated Value

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	548	-579	0	1.17	-0.6
2	21	-41	0.21	0.99	-0.08
3	-174	-104	0	1.01	-0.18
4	31	-23	0	1.05	-0.05
Total	426	-747		1.03	-0.3

Table 5.4-3. Goodness-of-Fit Statistics for Sensitive Parameter with Near Zero Value.

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	-487	643	0	3.57	0.68
2	390	591	0	5.83	1.08
3	-376	417	0	3.71	0.71
4	-317	673	0	7.33	1.58
Total	-1571	2324		4.52	0.93

5.4.2 Low Parameter Sensitivity

In the calibration of Model 1, the optimum value of PIVZ was 0.6. In this analysis, the parameter is varied to a low value of 0.1 and a high value of 1.2. The results for the adjustments to the low sensitivity parameter, PIVZ, are shown in Tables 5.4-4 and 5.4-5. Table 5.4-4shows the goodness-of-fit statistics for the 100% increase in PIVZ and Table 5.4-5 shows the goodness-of-fit statistics for the near zero value of PIVZ. Increasing PIVZ by 100% increased the water balance bias by 87 mm; increased the runoff bias by 77 mm; did not change the relative standard error; and increased the relative bias by almost 3%. Decreasing PIVZ to nearly zero decreased the water balance bias by 29 mm; increased the runoff bias by 75 mm; did not change the relative standard error; and increased the relative bias by almost 3%. For both changes, the correlation coefficients were increased by a maximum of 1%.

Table 5.4-4. Goodness-of-Fit Statistics for Sensitive Parameter with a 100% increase from Calibrated Value

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	195	-170	0.62	0.97	-0.18
2	-81	117	0.31	0.97	0.21
3	-210	-1	0.71	0.71	-0.002
4	-53	134	0.57	0.83	0.31
Total	-150	80		0.78	0.03

Table 5.4-5. Goodness-of-Fit Statistics for Sensitive Parameter with Near Zero Value

Year	WB	Runoff Bias	R	Se/Sy	e/y
1	213	-187	0.61	0.8	-0.2
2	-31	71	0.32	0.96	0.13
3	-151	-52	0.71	0.72	-0.09
4	2	90	0.58	0.83	0.21
Total	32	-78		0.78	-0.03

5.4.3 Discussion of Results

The analysis revealed that adjusting insensitive parameters has only a minor effect on the goodness-of-fit statistics. Changes in the sensitive parameter, however, had a significant impact on each of the goodness-of-fit criteria. While the difference in the effect on the goodness-of-fit statistics would be expected from varying values of parameters with different significance levels, it is important to note the minimal degree of change that occurred within the low sensitivity analysis. Sensitive parameters will converge to an optimal value because the goodness-of-fit statistics will change significantly throughout the calibration process until the optimal value is selected. However, low sensitivity parameters may never converge to an optimal value, because they can take on a wide range of values within the calibration process without affecting the results. Therefore, more complex models may not always contain physically rational values for low sensitivity parameters, unless the calibrator is aware in advance of rational values of the parameters and purposely assigns these values to parameters with low sensitivity.

Likewise, less sensitive parameters may not be needed in a model as they don't contribute to improvements in model prediction accuracy. Parameters with low sensitivity may also reflect on the importance of the process itself, suggesting that the process is not important and can be removed from the model. It is important to represent all of the important physical processes, but including parameters with low sensitivities may only increase difficulties in calibration and create a model with more parameters but less rational values. Likewise, two different calibrators may end up with different calibrated values for parameters as a result of the low sensitivity parameters. Calibrators should consider the importance of parameter values before finalizing a model and take into account the range of rational values of insensitive parameters.

5.5 Hydrologic Rationality of Parameters

An analysis was conducted to assess the hydrologic meaning of the fitted parameter values. Since Model 1 was considered the most representative of the physical processes, it was chosen for the analysis. The hydrologic meaning of each parameter value for Model 1 was analyzed using yearly average storage values and depths for each zone determined. The average storage depths are as follows: interception storage of 0.22 mm; surface zone storage of 0.587 mm; root zone storage of 166 mm; vadose zone storage of 2425 mm; and groundwater zone storage of 2180 mm. The average zone depths are as follows: root zone depth of 1350 mm; vadose zone depth of 6470 mm; and groundwater zone depth of 2181 mm. For the model components that take into account the time of year or crop characteristics, values for the month of May were used, assuming that May represents moderate temperatures and the beginning to middle of the growing season. The formulas that represented each physical process in each zone were analyzed based on the average storages and depths and the calibrated parameter values to determine whether the model is producing hydrologically rational outputs.

5.5.1 Interception Zone

The model allows a maximum interception of 0.5 mm, which is a rational maximum daily interception depth based on Madramootoo and Broughton (1987). Based on the average interception storage, 0.28 mm of rainfall is intercepted. The evaporation process in the interception zone is represented by the following equation from Chapter 3:

$$EI = PEXI * STI \qquad \text{Eq. (5-1)}$$

where EI = actual amount of evaporation from interception layer (mm), PEXI = calibrated evaporation parameter for the interception layer, and STI = storage depth in the interception layer (mm). For model 1, PEXI equals 0.5; therefore, 50% of the storage in the interception zone to be released through evaporation. Based on the average storage depth, 0.14 mm is evaporated from the interception zone on an average daily basis, which equals less than 5% of the maximum allowable evaporation for the middle of May, the time period chosen for this analysis. This is rational, because interception is a small component in the hydrologic process and would not be expected to contribute greatly to the total evaporation.

5.5.2 Surface Zone

The surface zone simulates surface runoff, infiltration, and evaporation. The interception parameter, PSZ1, equals 0.15 and the exponential parameter, PSZ2, equals 0.9 for the calibrated Model 1. As explained in Chapter 3, the following equation defines surface runoff:

$$QS = PSZ1 * SSZ^{PSZ2} \qquad \text{Eq. (5-2)}$$

where QS = surface runoff depth (mm), PSZ1 = calibrated surface runoff interception parameter, SSZ = surface storage depth (mm), and PSZ2 = calibrated surface runoff shape parameter. Since the value of PSZ2 is near 1, it was assumed that the equation is essentially linear, and 15% of the surface storage is released through runoff. This is comparable to values from the Rational method. The Rational coefficient for cultivated land can range from 0.08 to 0.31 depending on the slope of the land (McCuen 2005). This implies that runoff values that range from 8% to 31% of rainfall are reasonable. The model allows roughly 15% of surface zone storage to runoff, which is within the range of rationality according to the Rational method.

The infiltration model component uses the surface storage, the complement of the soil moisture in the root zone, and the infiltration parameter, PISZ. The infiltration depth is calculated with the following equation from chapter 3:

$$ISZ = PISZ*SSZ*(1-SRZ/DR) \qquad \text{Eq. (5-3)}$$

where ISZ = infiltration into the root zone (mm), PISZ = infiltration parameter specific to the surface zone, SSZ = storage depth in the surface zone (mm), SRZ = the storage depth in the root zone (mm), and DV = depth of the root zone (mm). The average soil moisture in the root zone is roughly 12.3%. The value of PISZ is 0.2. Therefore, on average, the model allows roughly 17.5% of the surface storage to infiltrate. This value can vary with changes in the root zone soil moisture.

The evaporation from the surface zone is based on the surface storage, the parameter PESZ, and a temperature component and represented by the following equation from chapter 3:

$$ESZ = PESZ*TO*SSZ \qquad \text{Eq. (5-4)}$$

where ESZ = actual amount of evaporation from surface zone (mm), PESZ = calibrated evaporation parameter for the surface zone (Celsius $^{-1}$), TO = daily temperature (Celsius), and SSZ = storage depth in the interception layer. For the month of May, the temperature component equals roughly 15. The evaporation parameter, PESZ, equals 0.19. Therefore, when multiplied by the temperature component, 285% of the surface storage will be evaporated. This explains why PESZ is an insensitive component, considering if storage is available it will be depleted through evaporation. Regardless of the insignificance of PESZ, it is rational that the surface zone storage would be depleted by the end of the day. Therefore, after a rain event, the surface storage contributes significantly to the total evaporation. However, in the absence of rain events, the storage is empty and does not contribute to the total evaporation.

5.5.3 Root Zone

The infiltration from the root zone to the vadose zone is a function of soil moisture in the root zone, the complement of the soil moisture in the vadose zone, and the parameter PIRZ and is represented by the following equation from chapter 3:

$$IRZ = PIRZ*(SRZ/DR)*(1-SVZ/DV) \qquad \text{Eq. (5-5)}$$

where IRZ = infiltration into the root zone (mm), PIRZ = calibrated infiltration parameter specific to root zone (mm), SRZ= storage depth in the root zone (mm), DR= depth of the root zone (mm), SVZ = storage depth in the vadose zone (mm), and DV = depth of the vadose zone (mm). The parameter PIRZ acts as a maximum allowable infiltration rate, and the value is reduced based on the soil moisture fractions. PIRZ is calibrated to equal 15, suggesting that if the root zone is 100% saturated and the vadose zone is empty, 15 mm will be infiltrated. The average soil moisture in the root zone and vadose zone equal 12.3% and 37.5%, respectively. Therefore, 1.15 mm is infiltrated on average based on Eq. (5-5), totaling to roughly 421 mm per year.

The evaporation from the root zone is a function of the soil moisture and the parameter PERZ and is represented by the following function from chapter 3:

$$EVZ = PERZ * SVZ/DV \qquad \text{Eq. (5-6)}$$

where EVZ = actual evaporation from the vadose zone (mm), PERZ = calibrated evaporation parameter for the vadose zone (mm), SVZ = storage depth in the vadose zone, and DV = depth of the vadose zone. Similar to the infiltration, PERZ acts as a maximum evaporation rate. At a calibrated value of 0.93 and with an average soil moisture of 12.3%, only 0.11 mm is evaporated from the root zone on an average daily basis. While this is a small contribution to the total daily evaporation, the root zone is the only component in the model in which transpiration can occur. Therefore, it is rational to assume that transpiration is a greater outlet than evaporation in the root zone, justifying the small depth of evaporation that results from PERZ.

5.5.4 Vadose Zone

The outflow from the vadose zone is a function of the vadose zone storage, the unsaturated hydraulic conductivity, and the parameters PQV1 and PQV2. The daily depth of outflow is calculated with the following equation from chapter 3:

$$QV = \mathrm{SVZ*KU/(PQV1+PQV2*sin(6.283*(ID+61)/365))} \qquad \text{Eq. (5-7)}$$

where QV = depth of interflow (mm), SVZ = storage depth in the vadose zone (mm), KU = unsaturated hydraulic conductivity (mm/hour), PQV1 = interflow parameter (mm/hour), PQV2 = parameter representing the cyclical component of interflow (mm/hour), and ID = day of the year. The unsaturated hydraulic conductivity was estimated to equal 0.04 based on the average soil moisture calculated. The maximum value would occur when the sine function component is equal to one. Based on the calibrated values of PQV1 and PQV2, which equal 0.004 and 0.00375, respectively, the maximum average outflow from the vadose zone is 0.03% of the storage, or 0.73 mm, equaling roughly 266.5 mm per year.

The infiltration from the vadose zone into the groundwater is a function of the parameter PIVZ and the soil moisture in the vadose zone and is calculated from the following equation from Chapter 3:

$$\mathrm{IVZ = PIVZ*(SVZ/DV)} \qquad \text{Eq. (5-8)}$$

where IVZ = infiltration into the saturated zone (mm), PIVZ = infiltration parameter specific to the saturated zone (mm), and SVZ = storage depth in the vadose zone (mm), and DV = depth of the vadose zone (mm). As the soil moisture increases, the infiltration depth increases. PIVZ is calibrated to equal 0.6. Based on the average soil moisture, an average of 0.22 mm is infiltrated into the groundwater daily, totaling to 82 mm per year.

A low infiltration depth is rational considering the groundwater zone is saturated and not extracting water at a fast rate from the vadose zone.

The evaporation from the vadose zone is a function of the soil moisture in the vadose zone and the parameter PEVZ. The depth of evaporation is calculated with the following equation from chapter 3:

$$EVZ = PERZ * SVZ/DV \qquad \text{Eq. (5-9)}$$

where EVZ = actual evaporation from the vadose zone (mm), PERZ = calibrated evaporation parameter for the vadose zone (mm), SVZ = storage depth in the vadose zone (mm), and DV = depth of the vadose zone (mm). PEVZ acts as a maximum evaporation rate and has a calibrated value equal to 8 mm/day. The average soil moisture for the vadose zone equals 37.5%, suggesting that on average, 3 mm is evaporated from the vadose zone. Considering the maximum daily evaporation rate equals 3.2 mm, and the previous zones contributed only a small amount of storage to evaporation, the evaporation from the vadose zone can be considered rational. Therefore, the vadose zone is the main source of evaporation in the model which explains the higher sensitivity of PEVZ compared to other evaporation parameters.

5.6 Confirmation

A fifth year of data, which was from the same runoff and rain gauge stations used to calibrate the 16 models, was used to verify the prediction capabilities of the model. Each of the models was run with the parameters calibrated for the specific model and the fifth year of data. The goal of the confirmation process was to determine whether or not the model can provide accurate estimates of runoff beyond the range of the data used for calibration. The results are shown in Table 5.6-1.

5.6.1 Runoff Bias

The runoff bias for the confirmation data is comparable to the individual runoff biases for the four-year data. For Model 1, the runoff bias for the confirmation data equals 170 mm, while the bias for the individual years within the four-year data ranges from -178 to 95. For Model 7, the runoff bias for the confirmation data equals 157 mm, while the bias for the four-year data ranges from -196 to 113 mm. For Model 16, the runoff bias for the confirmation data equals -121, while the bias for the four-year data ranges from -178 to 112 mm. Therefore, the confirmation data results in a runoff bias within the range of the four-year data, and often times an improvement to the biases in the four-year data.

The total runoff bias for Models 1, 7, and 16 was 3, 50, and 70 mm, respectively, for the four-year data. Compared to the results for the confirmation data, Model 1 and 16 have less bias while Model 7 has an equal bias. The poorer runoff bias can be explained by the lack of additional years to offset the runoff biases in the confirmation data. Therefore, based on both the range and total runoff biases, the confirmation data proves that the model is capable of predicting runoff at a similar level of accuracy for additional data.

5.6.2 Water Balance

The water balance is similar for the confirmation data record and the four-year data record. Each of the water balances in the confirmation data set is within the range of the water balances for an individual in the four year data set. For Model 1, the water balance ranges from -181 to 203 mm for the four-year data and equals -231 mm for the confirmation data. For Model 7, the water balance ranges from -200 to 258 mm for the four-year data and equals -255 mm for the verification data. For Model 16, the water balance ranges from 100 to 1130 mm for the four-year data and equals 550 mm for the confirmation data. The total water balance for the four-year data is incomparable, considering it is the sum of four years of data and, therefore, would realistically equal four times the value of the confirmation data. Both data sets have a positive change in the water balance at Model 10, the point at which PPET is eliminated. However, the four year data set experiences a greater change because it is the culmination of four years rather than one. Therefore, the based on the results for the water balance, confirmation data set is representative of the results for the models calibrated with four years of data, suggesting the model can be extended to other data sets.

5.6.3 Relative Standard Error and Correlation Coefficient

The relative standard error is consistently worse for the confirmation data compared to the results from the four years of data. The total relative standard error for the four year data set ranges from 10 to 45% better than the confirmation data set. The relative standard errors for the fifth year of data all exceeded one, resulting in correlation coefficients equal to zero. Both sets of data experienced similar trends in the relative standard error for Models 1 through 8. For both data sets, the relative standard errors were fairly constant in magnitude for Models 1 through 3, increased in the relative standard error for Model 4, remained constant in magnitude until Model 7, and increased again for Model 8. The results suggest that the model simplifications had the same effect on the goodness of fit for both sets of data. Models 9 through 16 do not follow a decreasing trend, but do not contain significant differences, with relative standard errors ranging from 1.17 to 1.33.

5.6.4 Relative Bias

The relative bias varies from 0.0001 to 0.04 for the four year model and does not follow a pattern based on complexity simplifications. The confirmation model relative bias ranges from 0.45 to 0.87 and also does not follow a pattern based on complexity simplifications. While the relative bias for the confirmation data is much greater in magnitude than the total relative bias for four years, the results for the individual years within the four years of data are closer in magnitude. For example, the relative bias of individual years for models run with the four year data set had a range of -0.69 mm to 1.04 for the less complex models. The additional years were able to compensate for the greater range of bias and result in an overall relative bias near zero. Therefore, while the bias in the confirmation model is worse than those of the models run with four years of data, it can be explained to an extent by the lack of additional years to compensate for the negative and positive biases as well as the smaller average runoff value.

5.6.5 Confirmation Analysis

While the confirmation showed that the fifth year of data produced results following similar trends and with values in the range of the results from the four years of data, the prediction accuracy was not as good as in the calibrated models. This may be explained by the following observations. First, the increase in magnitude for the relative bias in the confirmation data can be explained by the difference in average runoff observed between the four years of data and the fifth year of data. The four years of data had an average of 1.72 mm/day of runoff whereas the confirmation data only had an average of 0.73 mm/day. Therefore, the average value less than 1 inflates the relative bias resulting in greater values than occurred with the four years of data.

The decrease in prediction accuracy may also be the result of a lower runoff-to-rainfall ratio in the confirmation data than in the preceding four years. The runoff-to-rainfall ration ranged from 0.39 to 0.63 for the four years of calibration data but was 0.33 for the confirmation data. Therefore, the ratio is lowest in the confirmation data implying a smaller portion of runoff was produced than expected using the calibrated model. These factors explain the tendency of the calibrated model to overpredict with the confirmation data. The overprediction is more evident from Figure 5.3-1, which reveals the large oveprediction of runoff for Model 1 with the confirmation data.

Based on graphical analysis, data anomalies may exist between the rainfall and observed runoff. Figure 5.6-1 shows the predicted and measured runoff, while Figure 5.6-2 shows the daily rainfall for the confirmation year. The graphs suggest that the watershed was nonresponsive to the rainfall that occurred from days 191 to 300. However, the model does respond and predicts runoff during this time period. As discussed in this research, error in the runoff gauge or misalignments between the rainfall and runoff may contribute to poor prediction accuracy. Therefore, the decrease in accuracy in year five may be the result of data anomalies rather than the inability to apply the calibrated model to other rainfall and runoff data.

Table 5.6-1. Verification Results for Year 5

Model	WB	Runoff Bias	R	Se/Sy	e/y
1	-231	170	0	1.1	0.64
2	-233	174	0	1.1	0.65
3	-227	173	0	1.1	0.65
4	-255	168	0	1.14	0.63
5	-238	189	0	1.15	0.71
6	-218	174	0	1.16	0.65
7	-255	157	0	1.13	0.59
8	-270	172	0	1.23	0.64
9	-327	231	0	1.33	0.87
10	64	214	0	1.3	0.8
11	203	225	0	1.23	0.84
12	100	178	0	1.23	0.67
13	147	131	0	1.29	0.49
14	550	-121	0	1.17	-0.45
15	229	199	0	1.21	0.75
16	550	-121	0	1.17	-0.45

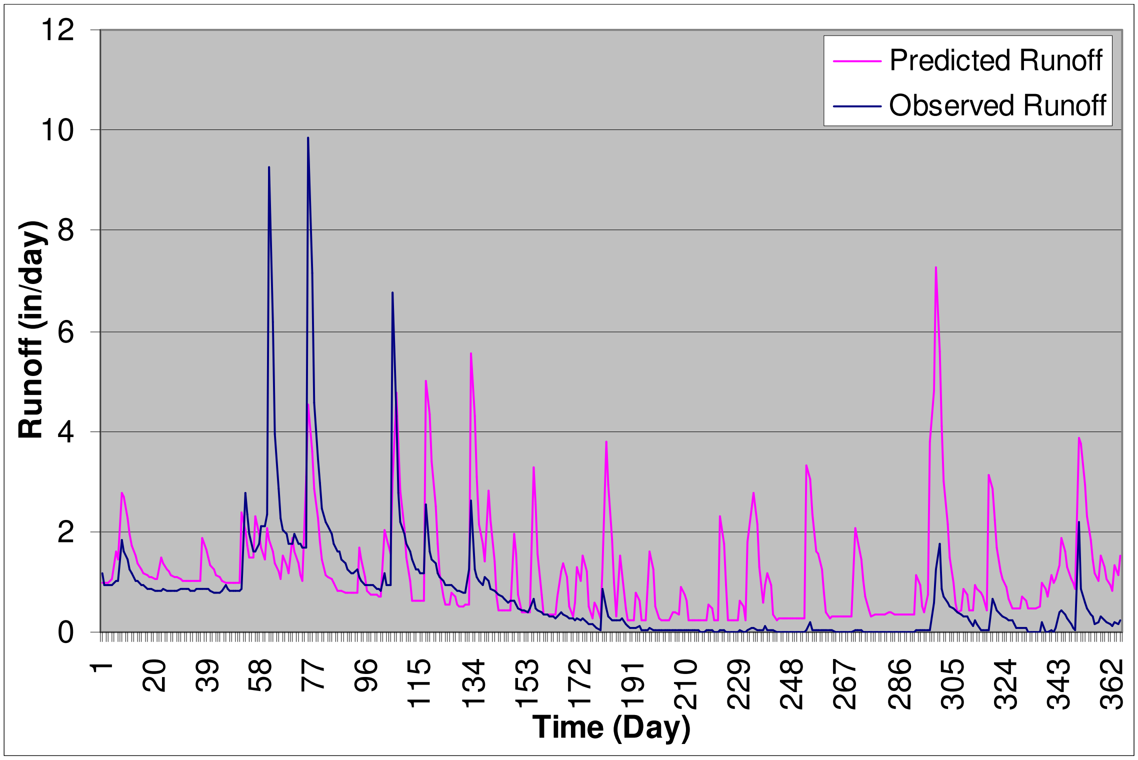

Figure 5.6-1. Runoff vs. Time for Year 5 Verification Data and Model 1.

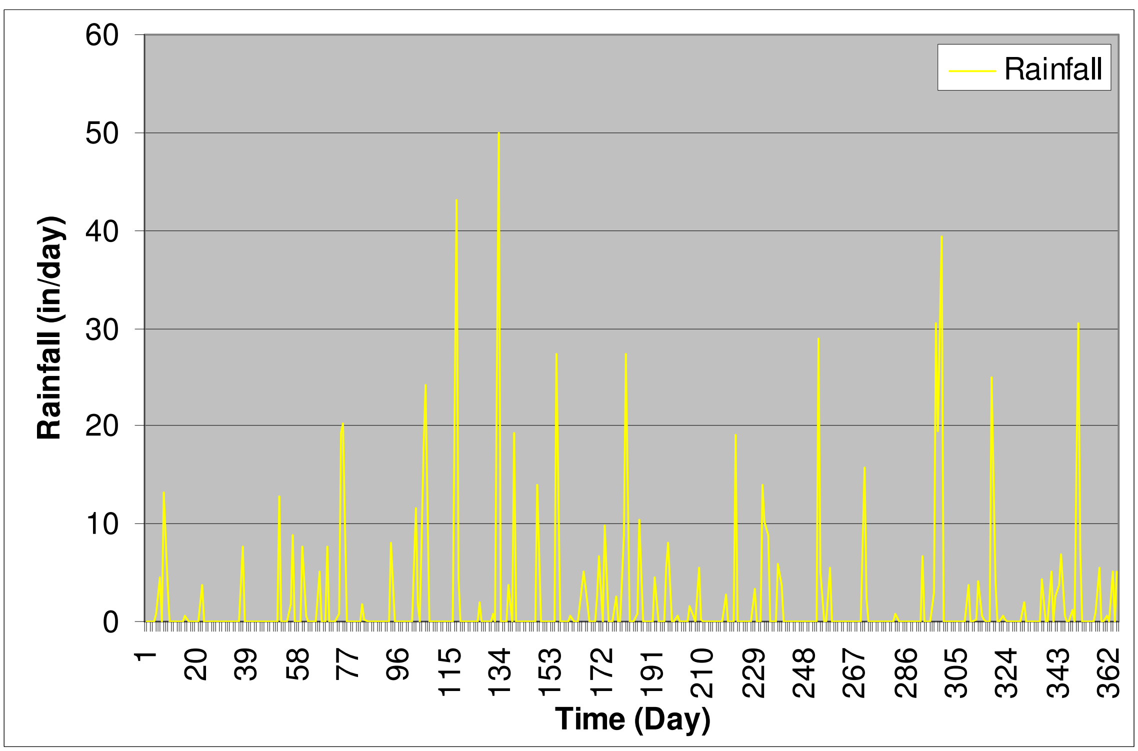

Figure 5.6-2. Rainfall vs. Time for Year 5 Verification Data

5.7 Assessment of Data Complexity in Water Quality Data

Data complexity is an important issue in hydrologic modeling. Processes in nature are characterized by variation, whether it is cyclical variation caused by seasonal changes, secular variation caused by temporal changes introduced by land development, or random variation caused by unknown factors. Sample data that are insufficient to characterize such variation will cause unrepresentative results when used in modeling. For example, poorly representative data will result in poorly calibrated model parameters, especially affecting the rationality of parameter values. Poorly calibrated parameters will reduce the prediction accuracy of the model. They also affect the ability to apply a model to additional systems for prediction because the values pertain to the unrepresentative data rather than the actual physical characteristics of the system. It is important that the complexities of data be understood and considered in hydrologic modeling in order to ensure rationality in calibration and accuracy of predictions.

Water quality data are inherently complex with considerable variation. However, measured water quality data are often only available as grab samples, which may not be representative of the entire population of data. Grab sample measurements are often made at irregular intervals rather than on a systematic temporal basis. Small samples are the norm. Thus, data sets may lack the total variation that would be inherent to longer records, and with small samples, the variation may not contain extreme variation. Thus, statistics based on grab samples may not reflect those of the population from which the grab samples were taken.

If a water quality record of n years of daily data was available, this would represent a sample from the population. This sample of 365*n values would have characteristics that are likely representative of the population of all possible values of the pollutant: past, present, and future. Variations in rainfall and runoff characteristics introduce variances into the n-year record of water quality measurements that cause it to deviate from the population, which is unknown.

When a record of daily values is not measured and instead grab samples are collected, the m grab samples represent a sample of a sample. Since the population is not known, but grab sample statistics would be used to represent characteristics of the population, it is necessary to examine how grab sample statistics would reflect the corresponding statistics of a complete n-year sample. From this analysis, the representativeness of grab sample records can be assessed.

While mean concentrations on loads computed from grab samples are of general interest, the distribution of the water quality parameters should be of central importance. Values in the tails of the distribution are important as are probability estimates, such as 95% exceedences. Therefore, evaluations of the ability of grab samples to provide accurate characteristics of underlying probability distribution are likely to be of greater value than the computation of mean values.

The goal of the grab sample data analysis was to determine the number of grab samples needed to provide accurate statistics that reflected the statistics of the entire record. This goal was achieved through the following four analyses which were discussed in Chapter 3:

- Compute statistical characteristics of an entire record of suspended data, including mean and probability distribution

- Randomly eliminate specified percentages of the data record to provide a sample of grab measurements and recomputed the same statistical characteristics.
- Randomly eliminate data points that reflect measurements during storm events to represent grab samples and recomputed the same statistical measures.
- Randomly eliminate data points that reflect measurements during low flows and recomputed the statistical characteristics.

These analyses were conducted using the procedure from Chapter 3 and the results are discussed herein.

5.7.1 Results for Random Elimination of Samples

The results for the analysis for the random selection of grab samples are shown in Table 5.7-1 for the 5% level of significance. Column 1 identifies the statistical test being made; Column 2 identifies the approximate percentage of the total observed record selected to represent the grab sample measurements; Column 3 identifies the sample size of the grab sample subset; Column 4 shows the calculated value for the subset for each statistical test; and Column 5 reveals whether the null hypothesis for each test is accepted or rejected. Results for the 1% and 10% levels of significance are given in Table 1 Appendix D.

The results for the one-sample Z-test, one-sample t-test, and two-sample t-test on the means indicate that the mean of the grab sample subset can represent the mean of the total observed record, regardless of the percentage of the total observed record eliminated. The only exception occurred for the one-sample t-test on the sample subset with only 1% of the observed data. The rejection of the null hypothesis occurred because the standard deviation of the subset equaled 7.4, whereas the standard deviation of the entire observed data record equaled 44.6. The denominator of the t-test statistic contains the standard deviation of the subset of the total observed record; therefore, a small standard deviation yields a large value of the test statistic. Acceptance of the test depends on the calculated test statistic being less than the critical test statistic. Whereas the two-sample t-test accounts for the variation in both samples, the one-sample t-test only uses the variation of the smaller subset. Therefore, rejection would be more likely, and it occurred for the 1% subset size.

The extremely low standard deviation in the 1% analysis is a result of both the random selection of data points and the small sample size used. The standard deviation is affected by both the sample size and random selection of data points. The smaller the sample size, the less probability of selecting very high valued points from a population made up of dominantly low values. However, if a large grab sample value happens to be selected, especially in a small sample, the standard deviation would be large because the small number of small values is not sufficient to offset the one extreme value. If a large sample value is not selected, as in the 1% anlaysis, the sample standard deviation will be relatively small. In such cases, the standard deviation of the grab sample subset will not reflect the mean of the total observed record.

The Kolmogorov-Smirnoff test results reveal that the grab sample subset is not representative of the total observed data set for eliminations with 50%, 25%, 10%, and 5% of the data remaining. For smaller samples containing 2.5% and 1% of the data, however, the test accepted the null hypothesis. This is most likely due to sampling variation within the sample selections. Also, a very large difference is required in order to reject the null hypothesis for small samples. Therefore, it is assumed that the poor representation of the distribution by 50% of the total observed data record implies that further reduction is unlikely to improve the distribution representation.

5.7.2 Grab Samples from below Mean Value

The results for the subsets of the observed values below the mean are shown in Table 5.7-2 for the 5% level of significance. Results for the levels of significance of 1% and 10% are shown in Table 2 in Appendix D. The results show that for the subset in which only 50% of the observed values below the mean were selected, the mean differs significantly from the mean of the total observed data record for each of the one-sample Z-test, one-sample t-test, and two-sample t-test at the 1%, 5%, and 10% levels of significance. For the remaining smaller subsets, the null hypothesis of equal means was rejected for the one sample t-test regardless of the sample size or level of significance. However, the null hypothesis was accepted for the Z-test and the two sample t-test for samples containing less than 50% of the data points below the mean. This is because the test statistics for the Z-test and two sample t-test contain the population standard deviation in the denominator. The population standard deviation, or the standard deviation of the total observed record, is much greater than that of the sample, which only contains a percentage of data points below the mean. Therefore, it is expected that both tests would have lower calculated test statistics, which are more likely to be accepted, than the one-sample t-test statistic. However, it is irrational to suggest that the subset containing less than 50% of the observed data below the mean better represents the subset containing 50% of the observed data below the mean. Therefore, the results show that grab samples collected in the absence of storms events are a poor representation of the mean of the total grab sample data. Exceptions to the rejection of this null hypothesis for smaller samples result from the decreasing power of statistical tests as the sample size is reduced.

The Kolmogorov-Smirnoff test reveals that the grab sample subset of values below the mean does not follow the same distribution as the total observed data record for the sample size containing 50%, 25%, 10%, and 5% of the total grab samples below the mean. Similar to the variance results, this is expected because the grab sample subset does not represent the spread of the data that occurs in the entire observed record. However, the results suggest that for the subsets containing 2.5% and 1% of the values below the mean, the distributions are the same. This is due to the small sample size, which results in a greater critical value and, therefore, greater likelihood of accepting the null hypothesis. Therefore, the Kolmogorov-Smirnoff is less powerful for smaller sample sizes. Therefore, the acceptance of the null hypothesis is ignored for the smaller subsets, and it is concluded that grab samples taken systematically in the absence of storm events will poorly represent the distribution of the total water quality record.

5.7.3 Grab Samples from above the Mean

The results for the selection of observed record values only collected during storm events, i.e., above the mean, are shown in Table 5.7-3 for the 5% level of significance. Results for the levels of significance of 1% and 10% are shown in Table 3 in Appendix D. The analysis of the mean values shows that the null hypothesis is rejected for the grab sample subset representing 50% of the observed data above the mean, regardless of the level of significance or test statistic used. However, the null hypothesis was accepted for the subset containing 25% of the values above the mean for the 2-sample t-test; the subset containing 10% of the values above the mean for the 1-sample t-test; the subset containing 5% of the values above the mean for the 1-sample Z-test and 2-sample t-test; and for all tests containing only 2.5% and 1% of the observed record above the mean. Further analysis showed that the standard deviations of the subsets varied greatly as a result of the random elimination of data. Values ranged from being significantly greater to significantly smaller then the total observed data record for the standard deviations for each of the random eliminations. This in combination with the effects of smaller sample sizes on the power of statistical tests explains the inconsistent acceptance and rejection of the null hypothesis as the sample size decreased. Therefore, as with the analysis conducted for values below the mean, it is assumed that the results for the 50% sample size are the most accurate in determining the extent to which the smaller grab sample subsets represent the total observed data record.

The results for the Kolmogorov-Smirnoff test show that the data selected from values above the mean do not follow the same distribution as the total observed data record regardless of the sample size or the level of significance. Considering the rejection of the null hypothesis for the mean, this result would be expected. The majority of the data points in the total grab sample lie below the mean. Therefore, a sample of data points taken above the mean will most likely follow a much different cumulative distribution than the entire observed data record.

5.7.4 Analysis of Data Complexity of Grab Samples

The results from the grab sample analysis in which data were randomly eliminated revealed that the mean could be accurately represented from smaller sample sizes; however, the distribution of the total grab samples is poorly represented by even

50% of the entire observed data record. The results from the grab sample analysis of values both above and below the mean revealed that the mean and distribution are poor representations of the total observed data record. This analysis represents grab samples collected consistently either during or in the absence of storm events. The random and systematic analyses revealed that randomness inherent in data as well as the selection processes influences the capability of representing the total sample and may cause misleading results. For example, smaller samples may be more representative than larger samples if the few data points collected represent the range of the population data. This does not imply, however, that smaller samples are more reliable than larger samples.

Based on these results, it is apparent that grab sample data should be used with caution for modeling purposes. Neither grab samples collected in the absence of storm events nor during storm events are representative of the total sample. Grab samples collected at random are representative of only the mean of the data points. Therefore, use of grab sample data for calibration purposes in modeling will most likely cause results that are misrepresentative of the actual system. The sample size of available data and the effect of missing data on prediction capabilities should be considered.

Table 5.7-1. Statistical Analysis for
Grab Samples Selected through Random Elimination
at a 5% Level of Significance

Test	% of Pop	Sample Size	Calculated Value	Decision
One-Sample Z-Test	50	689	0.57602	Accept
	25	355	-0.46843	Accept
	10	138	0.02148	Accept
	5	66	-0.16182	Accept
	2.5	45	-0.04097	Accept
	1	19	-0.54072	Accept
One-Sample t-	50	689	0.34701	Accept
	25	355	-0.69151	Accept

Test	10	138	0.03126	Accept
	5	66	-0.36866	Accept
	2.5	45	-0.10303	Accept
	1	19	-19.51606	Reject
Two-Sample t-Test	50	689	0.322488	Accept
	25	355	-0.39146	Accept
	10	138	0.020046	Accept
	5	66	-0.157488	Accept
	2.5	45	-0.040239	Accept
	1	19	-0.537198	Accept
KS-1 Test	50	689	0.93336	Reject
	25	355	0.906867	Reject
	10	138	0.713786	Reject
	5	66	0.39591	Reject
	2.5	45	0.137453	Accept
	1	19	0.217082	Accept

Table 5.7-2. Statistical Analysis for Grab Samples below the Mean Value Selected through Random Elimination at a 5% Level of Significance

Test	% of Pop	Sample Size	Calculated Values	Decision
One-Sample Z-Test	50	636	-4.173	Reject
	25	299	-2.346	Accept
	10	110	-1.348	Accept
	5	72	-1.053	Accept
	2.5	24	-0.6	Accept
	1	10	-0.402	Accept
One-Sample t-Test	50	636	-150.843	Reject
	25	299	-80.093	Reject
	10	110	-62.385	Reject
	5	72	-46	Reject
	2.5	24	-38.343	Reject
	1	10	-30.277	Reject
Two-Sample t-Test	50	636	-3.136	Reject
	25	299	-2.092	Accept
	10	110	-1.296	Accept
	5	72	-1.027	Accept
	2.5	24	-0.595	Accept

	1	10	-0.401	Accept
KS-1 Test	50	636	1.031	Reject
	25	299	0.951	Reject
	10	110	0.694	Reject
	5	72	0.481	Reject
	2.5	24	0.175	Accept
	1	10	0.183	Accept

Table 5.7-3. Statistical Analysis for Grab Samples above the Mean Value Selected through Random Elimination at a 5% Level of Significance

Test	% of Pop	Sample Size	Calculated Values	Decision
One-Sample Z-Test	50	75	10.084	Reject
	25	38	2.61	Reject
	10	13	3.918	Reject
	5	10	0.504	Accept
	2.5	3	1.531	Accept
	1	3	1.335	Accept
One-Sample t-Test	50	75	2.714	Reject
	25	38	4.883	Reject
	10	13	1.315	Accept
	5	10	3.806	Reject
	2.5	3	1.239	Accept
	1	3	1.252	Accept
Two-Sample t-Test	50	75	7.613	Reject
	25	38	2.567	Reject
	10	13	3.767	Reject
	5	10	0.502	Accept
	2.5	3	1.528	Accept

	1	3	1.333	Accept
KS-1 Test	50	75	0.904	Reject
	25	38	0.906	Reject
	10	13	0.904	Reject
	5	10	0.912	Reject
	2.5	3	0.92	Reject
	1	3	0.91	Reject

CHAPTER 6

GUIDELINES FOR EFFICIENT MODELING

Guidelines regarding data analysis, calibration, goodness of fit, and model development were established. Each guideline is based on the results from the analyses within this research. Following the data analysis guidelines will lead to a better understanding of model results affected by poor data quality. Consideration of calibration guidelines will improve the efficiency of the calibration process. Acknowledgement of the goodness of fit guidelines will improve the prediction accuracy resulting from the calibration process. And implementing the model development guidelines will improve the ability of a model to represent a system at an appropriate level of complexity.

6.1 GUIDELINES ON DATA ANALYSIS

Guideline: Data should be analyzed for anomalies before used for calibration or verification. Poor model prediction accuracy may be caused by data complexities rather than an inaccurate model. The goodness-of-fit statistics may suggest the model is inaccurate when, in fact, problems with the data base may actually be the cause.

In the analysis reported here, poor goodness of fit was influenced by poor correlation between the rain and runoff observations during storm events. Through analysis of the data, misaligned runoff and rainfall events were identified and altered, which resulted in a significant increase in prediction accuracy for the model. While it is unrealistic to alter data in model development, this analysis draws attention to the effects of data anomalies on the calibration and prediction capabilities of the model.

The following practices are suggested to minimize errors resulting from data anomalies.

a) Make *a priori* analyses of the data to identify anomalies that might possibly lead to poor goodness of fit. For example, conduct a cross-correlation analysis to determine the relation between the input data and output data used in calibration.

b) Consider the distance between the rainfall and runoff gauge when selecting data for model calibration to assess the potential for poor goodness of fit.

c) To avoid the effects of non-uniform rainfall events over a watershed, a weighted average of data records from multiple rain gauges near the watershed could be used.

d) To avoid the effects of non-uniform rainfall events over a watershed, use radar rainfall data.

Implementing this guideline is important for models with specific data characteristics. For example, a model that is being applied to different watersheds within which different rainfall and runoff gauges apply will be affected by the data used. The goodness of fit attained at one watershed may not be applicable to another watershed if the data are less significantly correlated. Model users must be aware of the effects of different data on the prediction capabilities of the model.

When calibrating with data characterized by high variance of flows, the prediction accuracy will be affected greatly by poorly correlated data. For example, rainfall events can range from a 2-year to a 100-year to a 500-year storm event. The depth of rainfall that occurs in each storm event varies significantly, resulting in a wide range of runoff values depending on the rainfall event. Therefore, as shown through this research, if the rainfall and runoff values are poorly aligned, the high peaks will be inaccurately predicted and result in poor goodness of fit. For example, if a 500-year storm is observed by a rainfall gauge on Day 1, but the storm does not affect the outflow until Day 2, the data is misaligned. The magnitude of the storm will cause a great difference between the predicted runoff and the observed runoff value for Day 1, as the model will respond to the storm on the same day it occurred. This will result in poor peak predictions. Therefore, it is important that data characterized by high variance values in particular be analyzed for data anomalies.

Models calibrated with short data records will be affected by the data issues addressed in this guideline. Short records contain few moderate to low flows, which are unable to compensate for any misaligned peak values. Therefore, the more accurate low flows cannot mask the poor predictions and produce reasonable goodness-of-fit statistics. Without analyzing the data first, the shorter data records will most likely result in poor prediction accuracy.

Guideline: When data samples are used to calibrate complex models, analyses should be conducted to determine whether or not the available data statistically represents the actual population before using the data in hydrologic modeling.

Water quality data are rarely available as a complete data set, i.e., collected at sampling interval needed. Grab samples are commonly used to calibrate models such as HSPF. However, grab samples are not always representative of the population and should be used with recognition of the representativeness of the data. Likewise, the procedure within which grab samples are collected may influence the ability of samples to represent the population. The following practices are suggested based on the grab sample analyses conducted within this research:

1) Grab samples should be used for the purpose of representing the long term mean value of a population rather than the distribution. As shown in the random grab sample subset, the mean values for the subset and total observed record were not significantly different, while the distribution differed significantly. Therefore, use of grab samples to model the distribution of the total yearly record would result in poor accuracy of predictions as well as irrational parameters.
2) Compare the statistical characteristics of the corresponding discharges with those of the entire yearly record to determine whether water quality grab sampling was conducted systematically or randomly. The systematic analysis determined that grab samples taken during or in the absence of storm events poorly represented the mean, variance, and distribution of the total observed record. Therefore, by analyzing the discharges corresponding to the grab samples with the yearly record of streamflows, the method in which the grab samples were selected can be determined. Discharges consistently above or below the mean of the yearly record would represent sampling strategically conducted during or in the absence of storms. If this is the case, the grab samples should be discarded for any purpose in modeling.

Implementing this guideline will improve the prediction capabilities of modeling by increasing the accuracy of data used in calibration. The more representative the available water quality data is of the population, the more likely the model results will the process being modeled. This is useful particularly with water quality data sets that often include only a few grab samples per year. Parameters that are fitted such that model predictions agree with grab sample measurements can be inaccurate if the grab samples themselves were not representative of the population from which the data were sampled. Inputting a small sample into a model that uses daily rainfall and streamflow data will require assumptions to be made regarding the missing data. Specifically, the statistical characteristics of the grab samples are assumed to accurately reflect the characteristics of the processes. Also, the goodness-of-fit statistics used to determine the prediction accuracy will be less reliable considering fewer data points are available to compare with the model outputs. If only a small sample of data is available, limitations on the model prediction accuracy can result from the inability of the sample to represent the population data.

6.2 GUIDELINES ON CALIBRATION

Guideline: Knowledge of parameter sensitivity is essential to efficiently and effectively calibrate a model. As shown in the parametric sensitivity exercise, a change in a highly sensitivity parameter has a much greater effect on the overall goodness of fit than a change of equal magnitude of a less sensitive parameter. Therefore, a lack of knowledge of parameter sensitivities may result in a calibration strategy where insensitive parameters are altered. This will produce little change in the prediction accuracy from that based on inaccurate initial parameter estimates. Iterations where little improvement in overall accuracy occurs because insensitive parameters are being changed can discourage continuing the calibration process. This may lead to drastic and unnecessary changes in parameter values in order to attain results that could be attained with more subtle changes to sensitive parameters. A lack of awareness of sensitivities may lead to assigning physically irrational values to insensitive parameters in order to acquire the desired goodness of fit. Also, realistic potential changes to sensitive parameters may be ignored.

This guideline is especially important in models with a large amount of parameters. The greater the number of parameters, the more likely the model is to contain insensitive parameters. Therefore, the probability of manipulating insensitive parameters is high. This would be prevented if a sensitivity analysis had been conducted prior to the start of calibration. Those calibrating highly complex models that contain many parameters should be aware of the importance of understanding parameter sensitivities to avoid irrational parameter values and inefficient calibration.

<u>Guideline</u>: When calibrating a model, emphasis should be placed on optimizing the parameters so that they reflect the physical processes that they represent (rather than calibrating to reflect only part of the data and the corresponding physical processes). Optimizing a model to meet specific design criteria may result in less than optimum goodness of fit and parameter values. Measured data reflect all of the physical processes; therefore, the calibration criteria selected should reflect all of the physical processes and not processes that are specific to selected parts of the data such as peak flows or low flows. As one part of this research, the model was calibrated using three objectives: (1) the best overall fit; (2) the best fit of the peak flows; and (3) the best fit of the low flows. The data base reflected all of the physical processes, not those specific just to peak flows or to low flows. The results showed that optimizing to meet one specific criterion caused poor goodness-of-fit statistics in regards to other criteria. For example, while calibration to ensure accurate fitting of peak flows, the overall water balance and runoff bias significantly worsened, as did the relative standard error and relative bias. Likewise, calibrating to get good prediction of the low flows resulted in poor prediction of the high flows.

In addition to the effect on goodness-of-fit criteria, model parameters based on the specific criterion such as fitting peak flows may be distorted from rational values for the physical processes represented by the model. The distortion occurs because the parameters attempt to compensate for the emphasis placed on the parameters principally responsible for the criterion of interest. For example, in the calibration to get unbiased estimates of the peak discharge rates, the parameter PSZ2 was modified. This required distortion of parameter PSZ1 to compensate for the emphasis on the peaks.

The following practices are suggested to avoid negative effects of these calibration issues:

1) If a model is calibrated to provide good predictions of one criterion, e.g., peaks, the model should not be used to predict other criteria, such as daily or low flows.
2) Consider the physical rationality of parameter values when emphasizing predictions of a single criterion.
3) Consider the physical rationality of model outputs when calibrating for a single criterion (i.e., the effects of accurate peak discharge on the total runoff).

4) With more complex objective functions, parameters could possibly be individually calibrated for low or high flows in order to better fit all ranges of flows.

Consideration of this guideline is important for models with a variety of users. A model calibrated with a specific goal may contain parameter values that do not apply to additional criteria. Lack of knowledge of initial calibration goals could result in misuse of the model. For example, if HSPF was calibrated to specifically predict peak volumes and then a user unknowingly used the calibrated parameters to trace nitrogen values through the water cycle, the predicted loads could be inaccurate because the parameters do not reflect the total range of flows. Specifically, the low flows and baseflows will be inaccurate resulting in poorly predicted concentrations. Therefore, models with a potential wide range of uses should be calibrated for overall goodness of fit rather than with specific goals, unless users are warned in advance of the calibration criteria.

Guideline: Examine the rationality of model parameters as rationality depends on complexity of the model. As the number of model parameters decreases, the values of the model parameters remaining deviate from rational values in order to compensate for the processes and parameters not included in the model. This is apparent by the parameter value changes that occurred as the model complexity decreased. While the prediction accuracy may remain the same, the model may not be representing the processes of the hydrologic cycle rationally. Therefore, it is unlikely that the model can be applied to other watersheds, because the parameter values are only representative of the data used for calibration. To avoid this issue, it is beneficial to identify a range of rational parameter values before beginning the calibration process. Assessing the hydrologic rationality of the model outputs, as was done in this research, is also beneficial in identifying irrational parameter values,

6.3 GUIDELINES FOR ASSESSING GOODNESS OF FIT

Guideline: As a model with an overall bias of zero may still have significant local biases, time series plots and goodness-of-fit statistics should be used to ensure that fitting did not lead to local biases. Local biases occur when a model overpredicts consistently during one segment of a time series and underpredicts consistently in another. The overall bias may be good, however, upon closer examination, the model contains biases. For example, suppose a model component with a cyclical trend is represented by the mean value of the component. The local biases that occur at the maximum and minimum of the function will balance each other, resulting in an overall bias of zero.

Likewise, a model calibrated with multiple years of data may overpredict some years and underpredict others. This example occurred throughout this research. As the model complexity was simplified, the range of runoff bias within the individual years increased while the overall bias remained unchanged. This does not imply that the model has acceptable goodness of fit. However, it would go undetected unless the calibrator makes additional analyses. This guideline is significant when multiple years of data are used in calibrating a model. It is difficult to fully understand the prediction accuracy for individual years over a large time span when the goodness-of-fit statistics represent only the overall results. The following suggestions are made to identify local biases:

1) Analyze model results over smaller time increments, i.e., monthly or yearly, as well as over the entire data record in order to identify local biases that exist within portions of the data.
2) Conduct a graphical analysis to attain a visual representation of the prediction of the data and identify local biases such as in years of high or low rainfall.

Guideline: Goodness-of-fit measures that separately reflect bias and accuracy should be used. An unbiased model may not necessarily yield the greatest accuracy. Model bias reflects systematic error variation. Biases alone do not necessarily reflect the actual prediction accuracy of the data, as accuracy involves both bias and precision, or systematic and nonsystematic error variation, respectively. A model could greatly overpredict for parts of the data record and underpredict for other parts but still maintain good overall unbiasedness, because the errors compensate for each other. This is apparent as the model complexity was decreased. The range of yearly biases increased, while the overall biases remained near zero. It is important to not only focus on the overall bias, but the overall prediction accuracy of the model.

Guideline: In addition to the overall prediction accuracy, the accuracy of subsets of the data, e.g., individual years of a multi-year data base, should be assessed and used in guiding the calibration. Statistics that reflect the overall prediction accuracy may be misleading when the biases and accuracy for individual years are highly variable and suggest poor goodness of fit. The simplification in model complexity conducted through this research showed little change in the overall goodness-of-fit statistics until only a few parameters remained. However, analyzing the goodness-of-fit statistics for the individual years revealed a decrease in accuracy and an increase in the range of biases for the individual years. The poor statistics for individual years occurred despite a consistent overall average bias of near zero for the water balance and runoff. Therefore, it is important that calibrators analyze the statistics for individual years to ensure that the overall accuracy of prediction is not misleading.

The same level of overall prediction accuracy can be achieved from several different combinations of parameters. This is sometimes referred to as “The Non Uniqueness Problem). The differences in the parameter values produce models that reflect the physical processes differently. Any one combination of parameter values can, therefore, yield different predictions for a given rainfall pattern such as the annual time series. Additionally, different sets of parameter values will yield different estimates of sensitivities. As the parameters are changed, the effect on predictions for each year will be different. In addition to changes on annual goodness-of-fit statistics, the different sets of parameters will influence other subsections of the data base, such as peaks or low flows. The different values will place different emphasis in the processes responsible for peaks and low flows.

This may be an issue for models that contain a large number of parameters and high complexity, such as the HSPF model. It may difficult when calibrating a model to determine the individual effects of each parameter on the individual years, and it may be more convenient to only focus on the overall goodness-of-fit. However, a thorough calibration should involve examining the individual years and determining if the accuracy of any one year can be improved without decreasing the overall goodness of fit.

6.4 GUIDELINES ON MODEL DEVELOPMENT

Guideline: In developing a model, all important physical processes should be represented. Maintaining physical processes is more important then parameter sensitivity; however, if all physical processes are represented, model eliminations should be based on parameter sensitivities. Parametric sensitivities have more meaning when the model includes components to represent all important physical processes. Comparison of the goodness-of-fit statistics for Models 13 and 15 shows that parameter sensitivity is not always the most important aspect of calibration. Model 13 included parameters for surface runoff, surface infiltration, and evaporation from the groundwater zone. Therefore, Model 13 represented three of the main hydrologic processes. Model 15 represented only one physical process, surface runoff, and included two parameters, PSZ2 and PSZ1. PSZ2 was consistently the most sensitive parameter throughout all stages of complexity. The analyses showed that maintaining the physical processes provided better prediction accuracy than including the most sensitive parameter. This suggests that even the most sensitive parameter can not compensate for the effect of maintaining physical rationality within a model. The parameters of the three physical processes were able to compensate for the variation associated with PSZ2. The following practices are suggested based on this finding.

1. Identify parameter sensitivities before calibrating a model.
2. Model complexity should be sufficient by including all important physical processes. This is necessary to ensure physical rationality of model outputs.

It is important to consider this guideline when simplicity is an important criterion of model development. The significance of including all physical processes in order to attain accurate results may not be recognized, which can lead to an inferior model. A compromise between simplicity of the processes modeled and the acceptable level of goodness of fit should be considered in model development.

Guideline: Where effects are of interest, model complexity must be sufficient to represent relevant physical processes. If a prediction is the only requirement of a model, then a relatively simple model may be adequate. However, if an estimate of an effect is needed, such as the effect of an infiltration rate on the amount of groundwater flow or evaporation, then a more complex, process simulating model is needed. The goal of modeling is to fit a function or functions to reflect the variation that exists within the data. The total variation in data is composed of two types of variation: (1) explained or systematic and (2) unexplained or unsystematic. Explained variation can be classified as secular, periodic, or cyclical and can be fit with one or more functions. The function or functions remove the explained variation from the total variation, leaving only unexplained variation. Unexplained variation is the result of physical occurrences that cannot be measured. Multiple functions better explain the individual processes and remove explained variation, however, the unexplained variation still exists which contributes to bias and inaccuracies. Therefore, regardless of the complexity of the functions contained in the model, unexplained variation will still exist. The same goodness of fit can be achieved with a simpler model that contains one main process as with a more complex model with components representing multiple processes.

In simple models, however, individual effects may be distorted because the calibration places an unbalanced measure of importance on the component principally responsible for the effect. For example, as the model complexity decreased in this research, the model was still capable of accurately predicting the total daily runoff. However, beginning in Model 8, surface runoff was the only contributor to the total runoff. The goodness-of-fit statistics indicated that predictions remained relatively accurate because the parameters still in the model were able to adjust for the missing parameters and functions. However, the rationality of the amount of surface runoff was most likely unrealistic, as groundwater is a large contributor to total runoff in reality.

This is an important guideline to consider when modeling for purposes beyond simply prediction accuracy. For example, in water quality modeling, the goal might be to trace chemicals or pollutants throughout the hydrologic cycle. If the water quantity components of the model are of low complexity, then the model will not accurately reflect the movement of either water or dissolved chemicals through the system. Therefore, if the goal is only a final prediction, then a simple model may be adequate. However, more complex goals require more complex models to reflect the physical reality of the system and the effects on the individual processes.

Guideline: Model components should be structured to reflect the physical process that they represent. This guideline is illustrated in the comparison of Model 11 with Models 14 and 16. The runoff process is commonly represented as a nonlinear function. For example, the SCS method represents the runoff depth as a squared function of the precipitation and initial abstraction. All of the USGS peak discharge equations use a nonlinear power-model form (Jennings et al. 1994). The slope-area method of discharge estimation represents the peak discharge as a function of the square root of the energy gradient of the area of interest. Therefore, it may be physically irrational to represent runoff with a linear function. Model 11 includes only one parameter, PSZ1. Models 14 and 16 are both one-parameter models that contain only PSZ2. The parameter PSZ1 forms a linear function, whereas the use of parameter PSZ2 forms a power function, both based on surface storage, to represent the runoff process. The results show that representing the runoff process with a linear function rather than with a nonlinear function caused an increase in runoff bias as well as relative bias. Not only did the overall bias increase, but the range of runoff biases for the individual years increased significantly when compared with the nonlinear model. The measured data reflects the nonlinear watershed processed; therefore, the nonlinear model yielded more accurate results than the simpler linear model.

The structure of model components is important to consider when developing complex models to represent multiple hydrologic processes. It is often simpler to utilize linear functions in order to minimize the total number of parameters to be calibrated. Minimizing the number of parameters is a valid objective. For example, the HSPF model, which reflects a wide range of physical processes, uses a number of linear functions. The model assumes that the large number of linear functions will mimic the effect of nonlinearity inherent to the data. However, consideration must be made in regards to the effects of simplifying inherently nonlinear functions on the prediction biases. A compromise between the simplicity of the model and the acceptable prediction biases that result from poorly represented physical processes should be considered.

When a model requires specific goals in prediction, poorly structured model components will result in poor prediction for the individual effects that may be needed to attain these goals. For example, if the goal is prediction of peak discharge rates and the model uses a linear function to predict runoff, the overall bias may remain accurate while the peaks are poorly predicted. Therefore, stormwater management will be poorly designed to adequately control peak discharges as a result of the poorly predicted peaks. Accurately structured model components will improve the portrayal of specific effects and allow the model to be used for specific design or research goals.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

The goal of this research was to explore the effects of model and data complexity on representing the physical processes of a watershed. This goal was accomplished through the following objectives: (1) to formulate and analyze a series of different model structures to study the relationship between model complexity and prediction accuracy; (2) to assess the affects of data anomalies and incomplete data sets on prediction accuracy of hydrologic models; (3) to demonstrate the use of model sensitivity in improving the efficiency of subjective optimization; (4) to show the effect of calibrating to optimally fit peak discharge rates or to optimally fit baseflows on overall prediction accuracy; and (5) to develop guidelines for improving all components of modeling of complex hydrologic systems. The results will improve the existing state of model development and calibration by enhancing the knowledge of data complexity, model complexity, and calibration.

The complexity of modeling varies with respect to model structure, calibration data, goodness-of-fit criteria, model constraints, and calibration fitting method. The model structure is complicated by the number of processes modeled as well as the formula structure representing the processes. More complex models require greater data bases, often containing inconsistencies that complicate the modeling process. Model constraints, such as the physical rationality of parameter values, make it difficult to rationally calibrate models that contain a large number of parameters. The complexity of the calibration criteria influences the efficiency of the calibration process as well as the applicability of the model for other purposes. The more complicated the model, the more sophisticated the fitting method must be, requiring greater user knowledge. These factors of model complexity were explored through the model simplifications and analyses within this research.

To achieve these objectives, a 14-parameter model was developed. The model represented the following hydrologic processes: (1) rainfall, (2) interception, (3) surface runoff, (4) interflow, (5) groundwater flow, (6) infiltration, and (7) evapotranspiration. The model included five layers: the interception, surface, root, vadose, and groundwater zones. These formed the most complex model, which was calibrated by subjective optimization using observed rainfall and runoff data.

The model was subjected to 15 different simplifications. The simplifications were made based on parameter sensitivities as well as the rationality of eliminating processes in a certain order. Each simplified model was calibrated with subjective optimization using the same observed rainfall and runoff data. The goodness-of-fit statistics were compared. Significant changes were identified and guidelines were developed regarding the effects of the simplifications in model complexity on the prediction capabilities of the model.

Additional analyses were conducted to explore other effects of complexity on the prediction capabilities of the model. The effects of data anomalies and incomplete data sets in hydrologic modeling were established. The effect of optimizing for specific design criteria on overall goodness of fit was demonstrated, and the role of parameter sensitivity in the calibration process was explained. Guidelines were developed based on these findings.

The guidelines provided as a result of this research can improve current state of model development, calibration, and use. They can be applied to multiple areas of modeling and types of models. Modelers attempting to attain simple models must be aware of the sacrifices in goodness of fit that will occur. Likewise, increasing the complexity of a model may not improve its goodness of fit and the calibration efficiency may not improve. Guidelines were provided through this research as to the effects of model complexity on the prediction capabilities of a model.

Regardless of the complexity of the model structure, data complexity can limit the prediction capabilities. Ignorance to the effect of data anomalies may result in poor model development and a frustrating calibration process. Likewise, incomplete data sets or data sets selected systematically are most likely not representative of the distribution of the population data and may provide model results that are not reflective of reality. Suggestions were made through this research to identify and avoid data anomalies as well as incomplete and systematically selected data sets.

Inexperienced calibrators will benefit from the guidelines regarding the calibration process and calibration criteria. Multi-parameter models will be more easily calibrated with knowledge of parameter sensitivity. Calibrators concerned with the individual effects of a process, such as in tracing nutrients through the hydrologic cycle, will benefit from the suggestions provided within to ensure rational parameter values and good overall prediction accuracy. Models with a variety of uses will be more applicable if not calibrated for specific design goals, as shown through this research.

7.2 Recommendations

While this research has improved the current state of model development and calibration, advancements in the field are still needed. The following are recommended research areas that will further progress the field of modeling and calibrating.

7.2.1 Data Analyses

1. Compare the effect of varying the distance between rainfall and runoff gauges for the calibration data.
2. Analyze the effects of increasing the data record length data on sensitivities of parameters and goodness of fit of the model.
   i) Include measured data for temperature, wind speed, radiation, etc. to better represent evaporation.
   ii) Determine the effects of modeling different soil types. This research only modeled clay; however, the model is capable of representing other soil types.
3. Determine if the data type (i.e., suspended solids, nutrients, etc.) affects the ability of incomplete or systematically selected data samples to represent the population data.

7.2.2 Model Structure

1. Analyze the effects of using more complicated functions to represent the hydrologic processes on the sensitivity of parameters and the goodness of fit of the model.
   i) Represent infiltration with a non-linear function such as Horton’s equation.
   ii) Represent evaporation with a more physically based function such as Penman’s equation.

2. Extend the analysis to include a water quality model. Determine the effects of the complexity of water quantity on the accuracy of water quality results.
3. Develop a more accurate portrayal of groundwater flow and Darcy's law
    i) Actual measurements of hydraulic gradient
    ii) Effect of K-values on results. K can take on multiple values.

7.2.3 Model Calibration

1. Calibrate components individually and compare to overall calibration
    i) Calibrate with available evaporation data to have more accurate parameters
    ii) Groundwater flow data
2. Determine if the sensitivity of individual components is based on the complexity of the functions used to represent them.
3. Determine the applicability of the model to other watersheds, i.e., different crop type, soil type, rainfall and runoff data.
    i) Does the model require recalibration?
    ii) Does the applicability vary based on the model complexity?

7.2.4 Parameter Sensitivity

1. Determine what makes certain parameters more sensitive than others.
2. Remove sensitive parameters and determine how insensitive parameters react.
3. Develop a method of approximately the standard errors of the parameters.

APPENDICES

Appendix A: Graphical Analyses of the Hydrologic Model

Figure A-1a. Graph of predicted and observed runoff versus time for year 1 of calibrated Model 1.

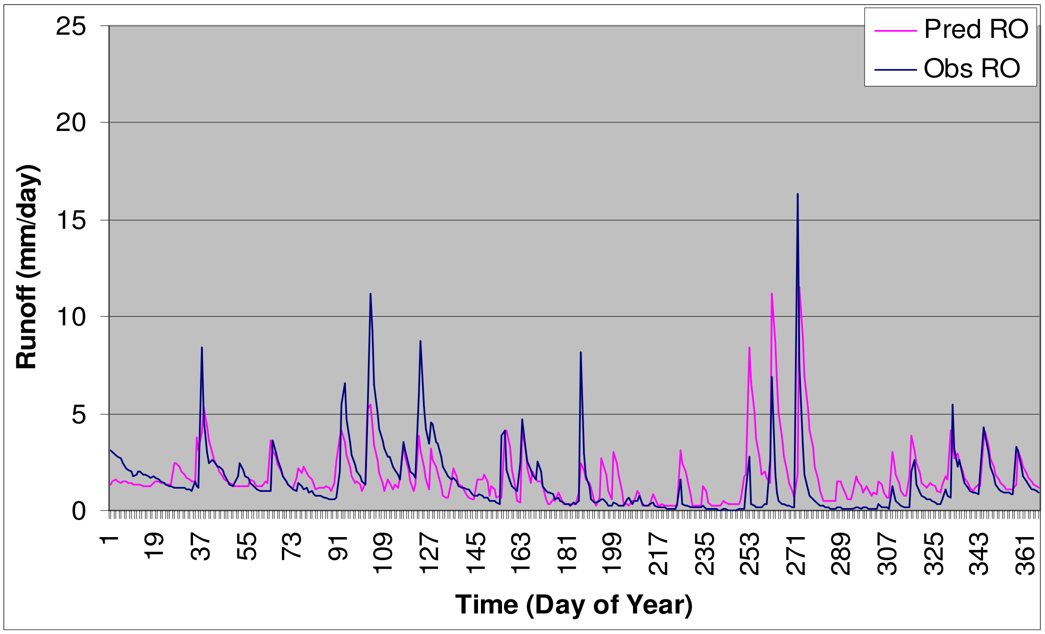

Figure A-1b. Graph of predicted and observed runoff versus time for year 2 of calibrated Model 1.

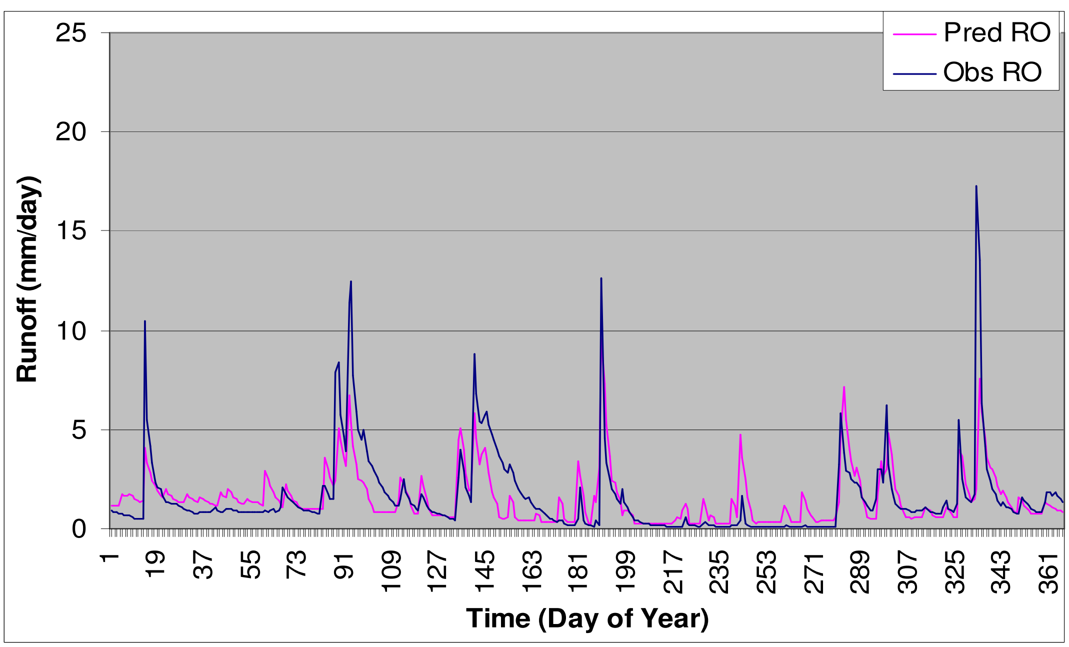

Figure A-1c. Graph of predicted and observed runoff versus time for year 3 of calibrated Model 1.

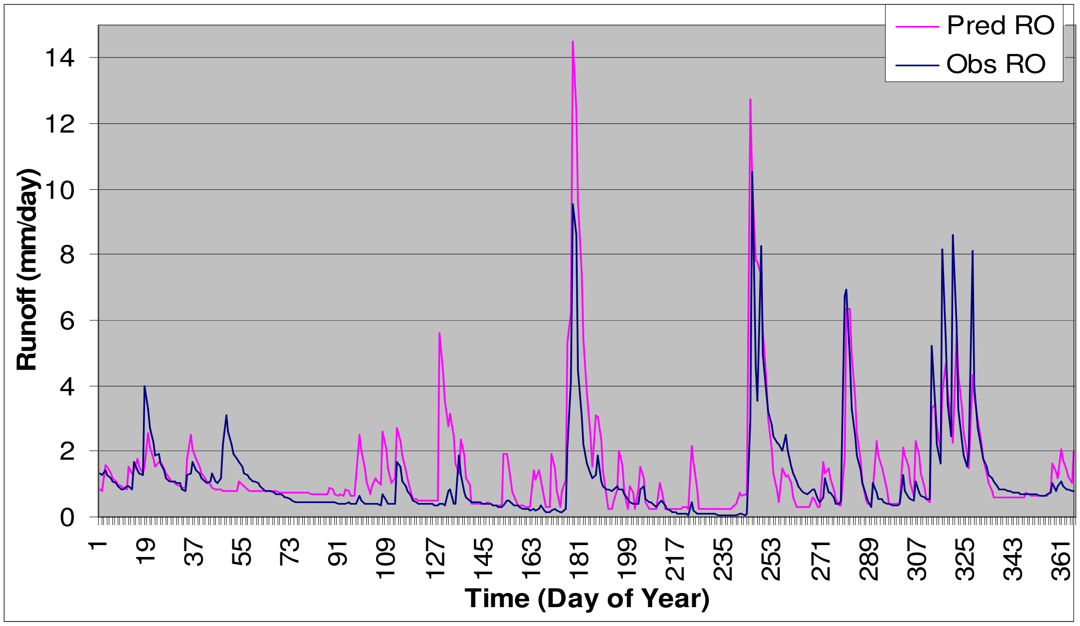

Figure A-1d. Graph of predicted and observed runoff versus time for year 4 of calibrated Model 1.

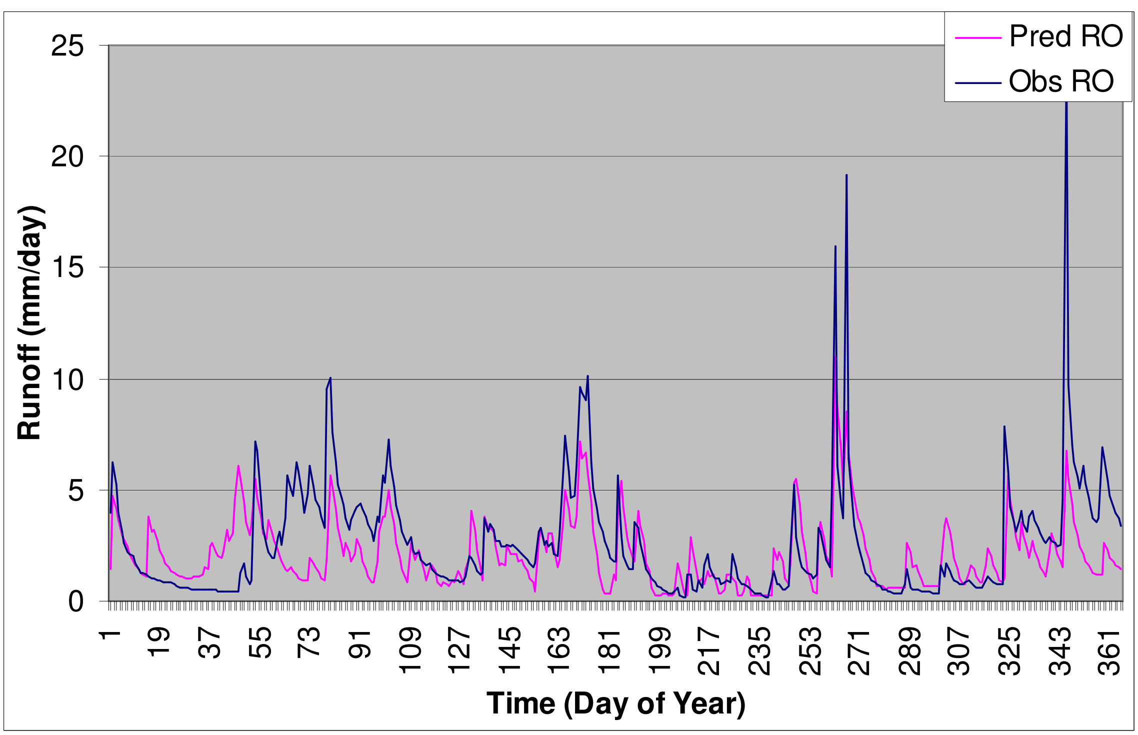

Figure A-2a. Graph of predicted and observed runoff versus time for year 1 of Model 3A.

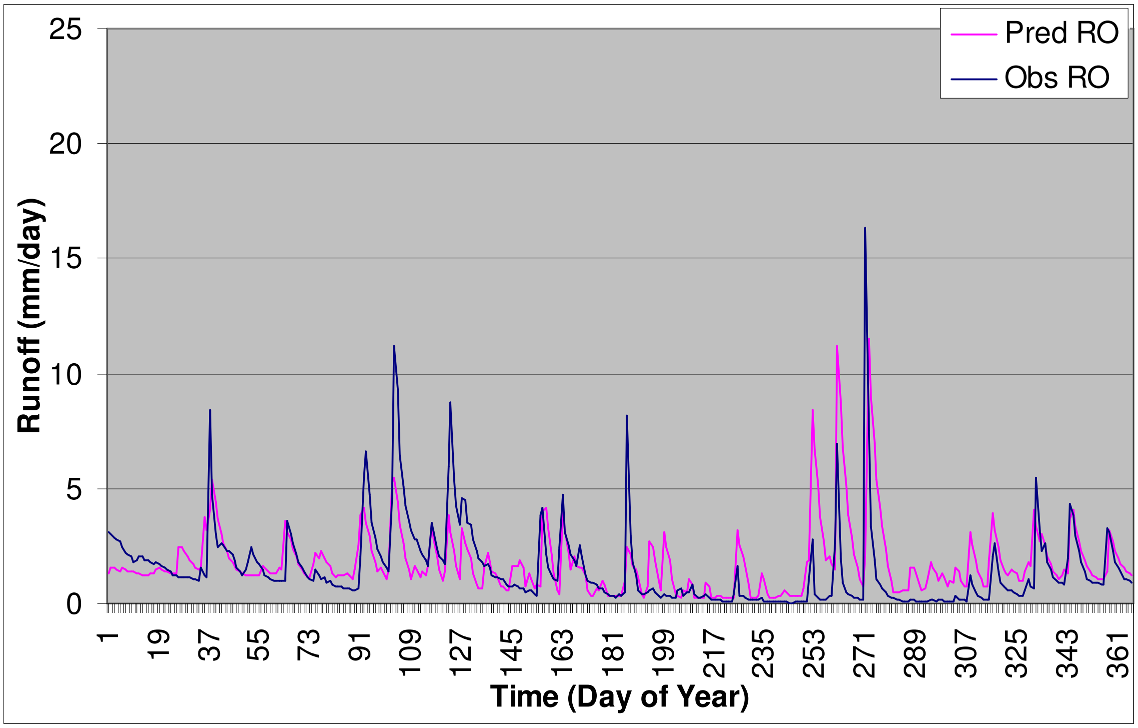

Figure A-2b. Graph of predicted and observed runoff versus time for year 2 of Model 3A.

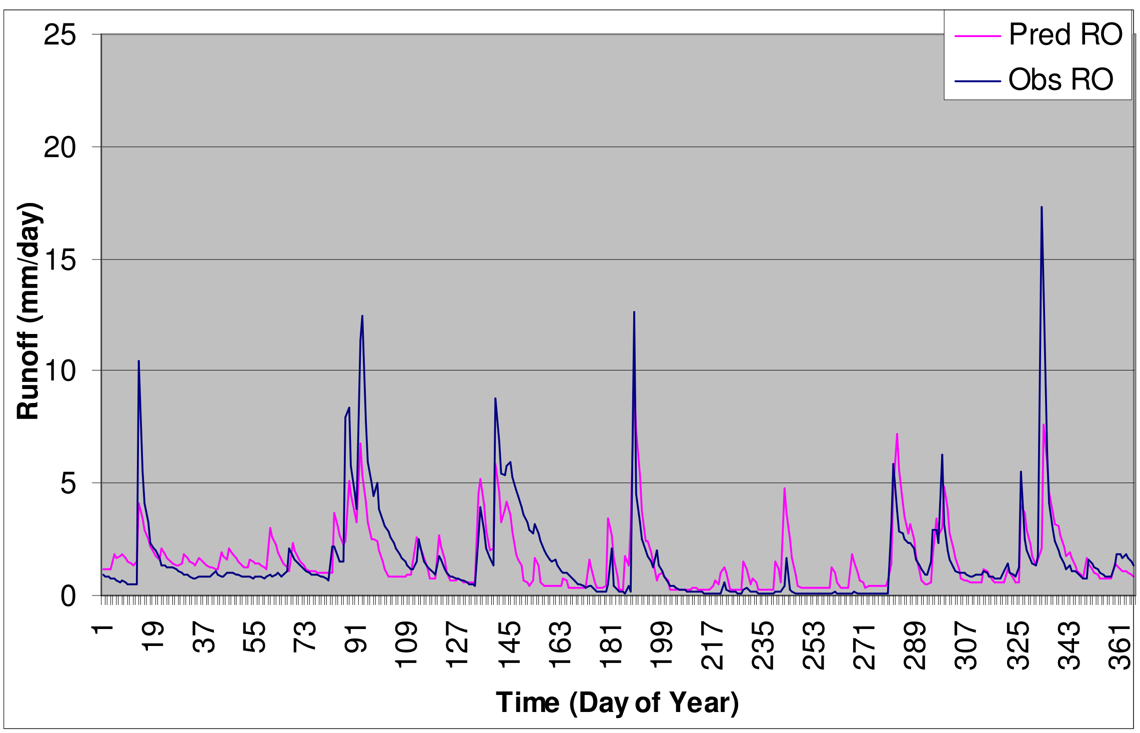

Figure 2C. Graph of predicted and observed runoff versus time for year 3 of Model 3A.

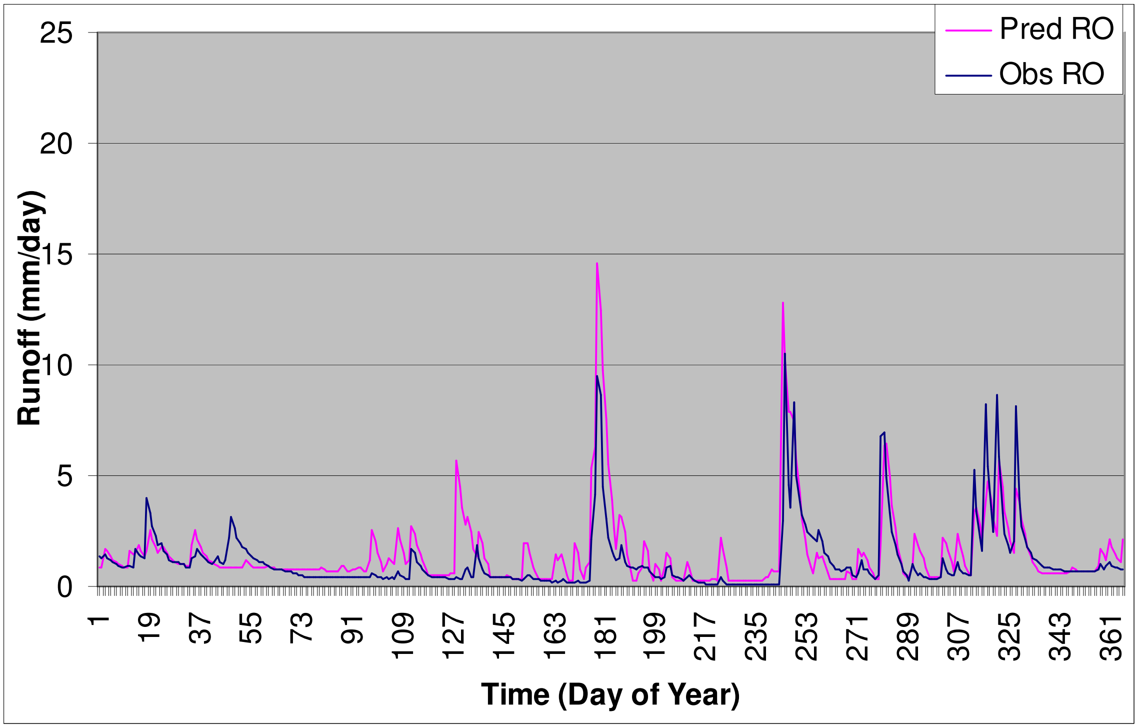

Figure A-2d. Graph of predicted and observed runoff versus time for year 4 of Model 3A.

Appendix B: Sensitivity Analyses

Table 1. Sensitivity Analysis of Model 1 based on a 20% decrease in each parameter value. The table shows the values for each goodness-of-fit criterion for the calibrated Model 1 and the changed values for each parameter change.

	Year	WB	Runoff Bias	R	Se/Sy	e/y
Model 1	1	203	-177	0.62	0.8	-0.19
	2	-59	96	0.32	0.96	0.175
	3	-183	-26	0.71	0.72	-0.04
	4	-29	111	0.58	0.83	0.26
	Total	-67	4		0.78	0.0018
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PSZ1	1	261	-245	0.58	0.83	-0.26
	2	-40	67	0.45	0.91	0.12
	3	-177	-44	0.7	0.73	-0.08
	4	-17	84	0.72	0.71	0.2
	Total	28	-138		0.76	-0.06
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PSZ2	1	338	-334	0.43	0.92	-0.35
	2	-22	38	0.52	0.87	0.07
	3	-174	-63	0.61	0.81	-0.11
	4	1.5	46	0.74	0.69	0.11
	Total	143	-312		0.81	-0.12
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PPET	1	267	-153	0.63	0.79	-0.16
	2	-21	145	0.27	0.98	0.26
	3	-148	34	0.72	0.71	0.06
	4	-12	171	0.55	0.85	0.4
	Total	87	197		0.78	0.08
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PEVZ	1	208	-176	0.62	0.8	-0.19
	2	-46	101	0.31	0.97	0.18
	3	-170	-16	0.71	0.71	-0.03
	4	-16	124	0.58	0.83	0.29
	Total	-23	33		0.78	0.01
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PQGZ	1	218	-192	0.62	0.8	-0.2
	2	-47	84	0.33	0.96	0.15
	3	-171	-38	0.71	0.72	-0.06
	4	-18	99	0.59	0.82	0.23
	Total	-18	-47		0.78	-0.019
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PEXI	1	203	-177	0.62	0.8	-0.19
	2	-59	96	0.32	0.96	0.175
	3	-183	-26	0.71	0.72	-0.04
	4	-29	111	0.58	0.83	0.26
	Total	-68	4		0.78	0.0018

Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PISZ	1	162	-128	0.63	0.79	-0.13
	2	-69	117	0.19	0.998	0.212
	3	-186	-10	0.72	0.71	-0.017
	4	-38	133	0.5	0.88	0.31
	Total	-132	112		0.79	0.045
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PESZ	1	203	-177	0.62	0.8	-0.19
	2	-59	96	0.32	0.96	0.175
	3	-183	-26	0.71	0.715	-0.04
	4	-29	111	0.58	0.83	0.26
	Total	-68	5		0.78	0.0019
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PIRZ	1	197	-183	0.61	0.8	-0.19
	2	-61	82	0.31	0.97	0.15
	3	-187	-37	0.71	0.715	-0.06
	4	-33	103	0.58	0.83	0.24
	Total	-84	-33		0.78	-0.013
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PERZ	1	203	-178	0.62	0.8	-0.19
	2	-57	95	0.32	0.96	0.17
	3	-182	-26	0.71	0.72	-0.044
	4	-27	112	0.58	0.83	0.26
	Total	-63	3		0.78	0.0013
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PIVZ	1	201	-175	0.62	0.8	-0.19
	2	-65	101	0.32	0.97	0.18
	3	-189	-20	0.71	0.715	-0.034
	4	-35	116	0.58	0.83	0.27
	Total	-88	22		0.78	0.009
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PXI	1	203	-177	0.62	0.8	-0.19
	2	-59	97	0.32	0.96	0.175
	3	-183	-25	0.71	0.72	-0.04
	4	-29	112	0.58	0.83	0.26
	Total	-68	6		0.78	0.0024
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PQVZ	1	231	-205	0.61	0.81	-0.22
	2	-41	76	0.33	0.96	0.14
	3	-178	-35	0.71	0.72	-0.06
	4	-29	108	0.59	0.82	0.25
	Total	-15	-56		0.78	-0.02
Parameter	Year	WB	Runoff Bias	R	Se/Sy	e/y
PQVZ2	1	201	-176	0.61	0.8	-0.19
	2	-59	96	0.3	0.97	0.175
	3	-182	-26	0.71	0.72	-0.04
	4	-28	110	0.57	0.83	0.26
	Total	-68	5		0.78	0.0018

Table 2. Sensitivity Analysis of Model 7 based on a 20% decrease in each parameter value. The table shows the values for each goodness-of-fit criterion for the calibrated Model 1 and the changed values for each parameter change.

Run	Year	WB	Runoff Bias	R	Se/Sy	e/y
Model 7 Calibration C	1	258	-196	0.56	0.84	-0.21
	2	-51	73	0.2	0.99	0.13
	3	-200	-41	0.68	0.74	-0.07
	4	-40	113	0.57	0.83	0.27
	Total	-33	-50		0.81	-0.02
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PSZ1	1	687	-519	0.14	0.995	-0.55
	2	312	-196	0.21	0.98	-0.36
	3	101	-261	0.55	0.84	-0.45
	4	204	-74	0.7	0.72	-0.17
	Total	1304	-1050		0.88	-0.42
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PSZ2	1	783	-619	0	1.11	-0.65
	2	371	-265	0	1.01	-0.48
	3	145	-316	0.37	0.93	-0.54
	4	257	-139	0.63	0.78	-0.33
	Total	1556	-1339		0.97	0.533
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PPET	1	672	-440	0.37	0.94	-0.46
	2	340	-134	0.19	0.99	-0.24
	3	157	-203	0.62	0.79	-0.35
	4	242	-7	0.65	0.77	0.02
	Total	1411	-784		0.85	-0.31
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PEVZ	1	635.0	-441.0	0.4	0.9	-0.5
	2	287.0	-138.0	0.2	1.0	-0.3
	3	88.0	-211.0	0.6	0.8	-0.4
	4	187.0	-21.0	0.7	0.8	-0.1
	Total	1198.0	-810.0		0.9	-0.3
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PISZ	1	545	-370	0.45	0.9	-0.39
	2	219	-87	0.16	0.99	-0.16
	3	35	-170	0.64	0.77	-0.29
	4	138	19	0.61	0.8	0.04
	Total	938	-608		0.84	-0.24
Parameter	**Year**	**WB**	**Runoff Bias**	**R**	**Se/Sy**	**e/y**
PGVZ	1	614	-442	0.36	0.94	-0.47
	2	265	-141	0.2	0.99	-0.26
	3	68	-216	0.61	0.8	-0.37
	4	170	-28	0.65	0.76	-0.07
	Total	1117	-827		0.85	-0.33

Appendix C: Model Simplifications

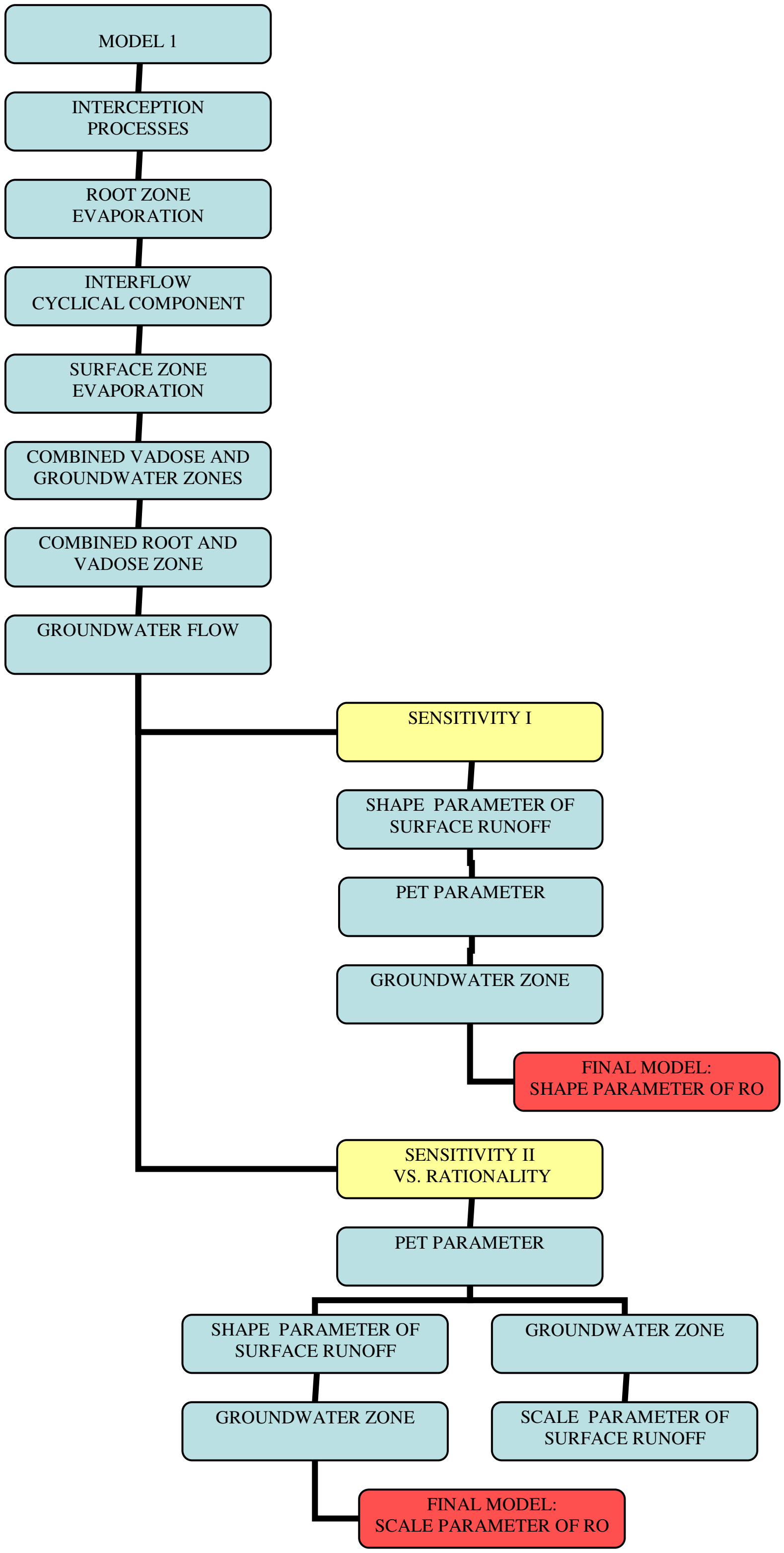

Figure C-1. Model Parameter Eliminations

Appendix D: Grab Sample Analyses Tables

Table D-1. Grab Sample Analysis for Random Elimination

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
50	One-Sample Z	1%	2.58	0.57602	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	0.34701	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	0.322488	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.062098054	0.93336	Reject
		5%	0.05		Reject
		10%	0.05		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
25	One-Sample Z	1%	2.58	-0.46843	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-0.69151	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	-0.39146	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.09	0.906867	Reject
		5%	0.07		Reject
		10%	0.06		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
10	One-Sample Z	1%	2.58	0.02148	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	0.03126	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	0.020046	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.14	0.713786	Reject
		5%	0.12		Reject
		10%	0.10		Reject

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
5	One-Sample Z	1%	2.58	-0.16182	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-0.36866	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	-0.157488	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.20	0.39591	Reject
		5%	0.17		Reject
		10%	0.15		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
2.5	One-Sample Z	1%	2.58	-0.04097	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-0.10303	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	-0.040239	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.24	0.137453	Accept
		5%	0.20		Accept
		10%	0.18		Accept
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
1	One-Sample Z	1%	2.58	-0.54072	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-19.51606	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-0.537198	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.37	0.217082	Accept
		5%	0.31		Accept
		10%	0.28		Accept

Table D-2. Grab Sample Analysis Results for Data Points above the Mean Value

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
50	One-Sample Z	1%	2.58	10.084	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	One-Sample t	1%	2.58	2.714	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	7.613	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	KS-1	1%	0.19	0.904	Reject
		5%	0.16		Reject
		10%	0.15		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
25	One-Sample Z	1%	2.58	2.61	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	One-Sample t	1%	2.58	4.883	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	2.567	Accept
		5%	1.96		Reject
		10%	1.64		Reject
	KS-1	1%	0.26	0.906	Reject
		5%	0.22		Reject
		10%	0.20		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
10	One-Sample Z	1%	2.58	3.918	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	One-Sample t	1%	2.58	1.315	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	3.767	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	KS-1	1%	0.43	0.904	Reject
		5%	0.34		Reject
		10%	0.30		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
5	One-Sample Z	1%	2.58	0.504	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	3.806	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	0.502	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.49	0.912	Reject
		5%	0.41		Reject
		10%	0.37		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
2.5	One-Sample Z	1%	2.58	1.531	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	1.239	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	1.528	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.83	0.92	Reject
		5%	0.71		Reject
		10%	0.64		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
1	One-Sample Z	1%	2.58	1.335	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	1.252	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	2-Sample t	1%	2.58	1.333	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.83	0.91	Reject
		5%	0.71		Reject
		10%	0.64		Reject

Table D-3. Grab Sample Results for Values below the Mean Value

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
50	One-Sample Z	1%	2.58	-4.173	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	One-Sample t	1%	2.58	-150.843	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-3.136	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	KS-1	1%	0.06	1.031	Reject
		5%	0.05		Reject
		10%	0.05		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
25	One-Sample Z	1%	2.58	-2.346	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-80.093	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-2.092	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.09	0.951	Reject
		5%	0.08		Reject
		10%	0.07		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
10	One-Sample Z	1%	2.58	-1.348	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-62.385	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-1.296	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.16	0.694	Reject
		5%	0.13		Reject
		10%	0.12		Reject

% of Population	Test	Level of Sig.	Critical Value	Calculated Value	Decision
5	One-Sample Z	1%	2.58	-1.053	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-46	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-1.027	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.19	0.481	Reject
		5%	0.16		Reject
		10%	0.14		Reject
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
2.5	One-Sample Z	1%	2.58	-0.6	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-38.343	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-0.595	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.24	0.175	Accept
		5%	0.20		Accept
		10%	0.18		Accept
% of Population	**Test**	**Level of Sig.**	**Critical Value**	**Calculated Value**	**Decision**
1	One-Sample Z	1%	2.58	-0.402	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	One-Sample t	1%	2.58	-30.277	Reject
		5%	1.96		Reject
		10%	1.64		Reject
	2-Sample t	1%	2.58	-0.401	Accept
		5%	1.96		Accept
		10%	1.64		Accept
	KS-1	1%	0.32	0.183	Accept
		5%	0.27		Accept
		10%	0.24		Accept

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D., and M. Smith (1998). Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Food and Agriculture Organization of the United States, Rome.

Benaman, J. and C.A. Shoemaker (2004). Methodology for Analyzing Ranges of Uncertain Model Parameters and Their Impact on Total Maximum Daily Load Process. Journal of Environmental Engineering 130(6):648-656.

Bennis, S. and E. Crobeddu (2007). New Runoff Simulation Model for Small Urban Catchments. Journal of Hydrologic Engineering 12(5):540-544.

Dawdy, D.R. and T. O'Donnell (1965). Mathematical Model of Catchment Behavior. Journal of the Hydraulics Division 91(HY4):123-137.

Environmental Data Service (1968). Climatic Atlas of the United States. U.S. Government Printing Office, Washington, D.C.

Holtan, H.N., England, C.B., and W.H. Allen, Jr. (1967). Hydrologic Capacities of Soils in Watershed Engineering. International Hydrology Symposium Fort Collins, CO, September:218-226.

Jenings, M.E., Thomas, Jr., W.O., and H.C. Riggs (1994). Nationwide Summary of U.S. Geological Survey Regional Regression Equations for Estimating Magnitude and Frequency of Floods for Ungaged Sites, 1993. WRI Report 94-4002. Reston, VA.

Lu, Sun, McNulty, and Amatya (2005). A comparison of Siz Potential Evapotranspiration Methods for Regional Use in the Southeastern United States. Journal of American Water Resources Association, 41(3):621-633.

Madramootoo, C.S. and R.S. Broughton (1987). A Computer Simulation Model of surface and Subsurface Flows From Agricultural Areas. Canadian Water Resources Journal 12(1):20-43.

Mandeville, A.N., O'Connell, P.E., Sutcliffe, J.V., and J.E. Nash (1970). River Flow Forecasting Through Conceptual Models Part III- The Ray Catchment at Grendon Underwood. Journal of Hydrology 11:109-128.

McCuen, R.H. (2005). Hydrologic Analysis and Design: Third Edition. Prentice Hall, Upper Saddle River.

Mein, R.G. and B.M. Brown (1978). Sensitivity of Optimized Parameters in Watershed Models. Water Resources Research 114(2):299-303.

Nash, J.E., and J.V. Sutcliffe (1970). River Flow Forecasting through Conceptual Models Part 1-A Discussion of Principles. Journal of Hydrology 10:282-290.

O'Connel, P.E., Nash, J.E., and J.P. Farrell (1970). River Flow Forecasting Through Conceptual Models Part II - The Brosna Catchment at Ferbane. Journal of Hydrology 10:317-329.

Risser, D. W. and Siwiec, S.F., 1996, Water-Quality Assessment of the Lower Susquehanna River Basin, Pennsylvania and Maryland: Environmental Setting: U.S. Geological Survey Water Resources Investigations Report 94-4245, 70 pages.

Rushton, K.R., Eilers, V.H.M, and R.C. Carter (2006). Improved Soil Moisture Balance Methodology for Recharge Estimation. Journal of Hydrology 318:379-399.

Saxton, K.E. (1983). Soil Water Hydrology: Simulation for Water Balance Computations. New Approaches in Water Balance Computations (Proceedings of the Hamburg Workshop, August 1983). IAHS Publ. no. 148

Schilling, W. ad L. Fuchs (1986). Errors in Stormwater Modeling- A Quantitative Assessment. Journal of Hydraulic Engineering 112(2):111-123.

Thomann, R.V. and Mueller, J.A. (1987). Principle of Surface Water Quality Modeling and Control. Harper Row, Cambridge.

Todd, D.K., and L.W. Mays (2005). Groundwater Hydrology: Third Edition. John Wiley & Sons, Inc., Hoboken.