ABSTRACT

Title of Dissertation: A STRATEGY FOR CALIBRATING THE HSPF MODEL

Angélica L. Gutiérrez-Magness, Doctor of Philosophy, 2005

Dissertation directed by: Professor Richard H. McCuen
Department of Civil and Environmental Engineering

The development of Total Maximum Daily Loads (TMDLs) and environmental policies rely on the application of mathematical models, both empiric and deterministic. The Hydrologic Simulation Program-FORTRAN (HSPF) is the most comprehensive model, and it is frequently applied in the development of TMDLs for nonpoint sources control. Despite the wide use of HSPF, a documented strategy for its calibration is not available. Furthermore, the most common calibration approach uses subjective fitting and focuses on the attainment of statistical goodness of fit, ignoring in many cases the rationality of the model.

The goal of this research was to develop a strategy for calibrating the HSPF model in combination with the model-independent-parameter estimator (PEST). PEST is an objective parameter estimator that should eliminate some of the subjectivity from the calibration process and reduce the repetitive effort associated with subjective fitting.

The strategy was established through a series of analyses, which included the development of a weighted multi-component objective function used as the criterion for calibration. The weights are a function of the flow components of the measured runoff.

The use of this new weighting procedure improves model and prediction accuracy. Methods of rainfall disaggregation and their effect on the prediction accuracy were studied. The results indicated that methods based on analyses of actual storm frequency data provided the most accurate daily-disaggregated values and thus, the best conditions to achieve accurate predictions with the HSPF. Analyses showed that the HSPF model requires a start-up period of about a year to allow the predicted discharges to become insensitive to erroneous estimates of the initial storages. The predictions during the start-up period should not be used for either calibration or the analysis of the goodness of fit. Analyses also showed that using HSPF as a lumped model can reduce the prediction accuracy of discharges from a watershed with an inhomogeneous land use distribution. The fulfillment of the research objectives provides a systematic procedure that improves the hydrologic calibration of the HSPF model.

A STRATEGY FOR CALIBRATING THE HSPF MODEL

By

Angélica Lucia Gutiérrez-Magness

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland at College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:

Professor Richard H. McCuen, Chair
Associate Professor Kaye L. Brubaker
Associate Professor Glenn E. Moglen
Professor Yaron M. Sternberg
Associate Professor Keith Herold

DEDICATION

This dissertation is dedicated to my family and especially to my husband Skip. Their patience, encouragement, and love helped me through the difficult times and made possible the culmination of my research. To my father who has been my inspiration.

ACKNOWLEDGEMENTS

I would like to acknowledge the guidance and friendship of Dr. Edward Kaplan during my Master studies at Stony Brook University. His challenging ideas and encouragement motivated my return to graduate school.

I would like to acknowledge and thank my advisor, Dr. Richard H. McCuen, for his guidance and commitment with my education. His passion for excellence took me beyond my expectations, helping me to be the engineer I am today.

I would also like to acknowledge Dr. John Doherty from Watermark Computing Inc. His technical guidance and invaluable assistance made great part of this work possible.

I would like to acknowledge Dr. Luis Alberto Jaramillo, who during my undergraduate studies in Colombia showed me the importance of environmental engineering to society.

I would also like to acknowledge the assistance of my dissertation committee for their guidance and recommendations, which are greatly appreciated.

Finally, I would like to acknowledge the opportunities and trust that Mr. Narendra Panday granted me. The technical challenges of the projects during my work at the Maryland Department of the Environment (MDE) were in part, the motivation of my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

HSPF – Hydrologic Simulation Program-FORTRAN

PEST – model-independent-parameter estimator

TMDL – Total Maximum Daily Loads

CBPO – Chesapeake Bay Program Office

EPA - Environmental Protection Agency

GIS – geographic information system

MDE – Maryland Department of the Environment

DNREC –

DSA – Digital Signal Analysis

ADC – Analog-to-digital Converter

DAC – Digital-to-analog Converter

USGS – United States Geological Survey

SWM – Stanford Watershed Model

SWMM – Storm Water Management Model

DWSM – Dynamic Watershed Simulation Model

KINEROS – KINematic Runoff and EROSion Model

CHAPTER 1

INTRODUCTION

1.1 RESEARCH NEED

The application of mathematical models is a common tool used to address environmental pollution problems. In many cases, the development of regulations to address issues of water quality pollution is supported by the results of mathematical models such as the Hydrologic Simulation Program – FORTRAN (HSPF), which simulates hydrologic and water quality processes. The documentation of the model applications is usually aimed to the description of the input data and to the analysis of the model results. However, documentation of the calibration strategy or a discussion of the effect of the data assumptions on the accuracy of the model predictions is not usually available.

Several states have made the decision to use the HSPF model for the estimation of pollutant loads to address some of the localized impairments throughout the United States. Specifically, the model results are used to develop policies that fulfill the requirements of the Federal Clean Water Act in the development of Total Maximum Daily Loads (TMDLs). A TMDL is an estimated value of the maximum amount of a given pollutant that a body of water can assimilate without violating water quality standards. Although accurate predictions of water quality are highly dependent on the accuracy of the predicted discharges and despite the intense use of the HSPF model, a calibration strategy has not being designed and documented.

Where a watershed model is used as a major part of the quantitative aspect of making pollution estimates within a region, it is important to have a systematic procedure for calibration. Many of the current calibration procedures focus on the attainment of high values of the correlation coefficient or other measurements of goodness of fit, but the procedures disregard fundamental elements of a reliable calibration. A model assessment based solely on goodness-of-fit statistics is inadequate because it ignores rationality as a criterion. High correlation can result from simple models and from complex models with irrational parameters. Model quality should not be based solely on the achievement of a level of explained variance, but on the understanding of the model structure, the rationality of the model, and on the effect of the model assumptions on the prediction accuracy.

Subjective calibration it is still a widely used approach regardless of the multiple problems that the methodology presents. Professed benefits of subjective calibration are related to the understanding of the physical processes that occur in the watershed through the trial-and-error method and the awareness of model limitations gained during the calibration process. However, the attainment of these benefits is highly dependent on the modeler's knowledge of hydrology, statistics, and modeling. In many cases, the modeler lacks knowledge in one or more of these areas, which can reduce the likelihood of finding the true statistically optimum parameter values. A calibration strategy can help offset the lack of knowledge on any one of the areas mentioned above by providing guidance on issues related to the fitting process. Given that the level of experience differs from modeler to modeler, it is difficult to expect that a subjective calibration of a large watershed model conducted by a group of individuals can provide a fair and equal treatment to the calibration.

The calibration of complex models, which includes continuous hydrograph models such as the HSPF, requires sophisticated calibration methods. Relatively new software or model-independent-parameter estimators facilitate the calibration of these models using a predetermined optimization function based on analytical methods such as weighted least squares. However, to ensure that the final parameter values reflect the hydrologic components (surface runoff, interflow, and baseflow) that the models are designed to represent, these model-independent-parameter estimators require complex objective functions. Little is known about the interaction of the HSPF model and the model-independent-parameter estimators or about the settings of the objective function.

The greatest benefits of applying parameter estimators to complex models is the potential for reproduction of the calibration results, the elimination of subjectivity from the calibration process, and the reduction of the repetitive work associated with subjective fitting. However, the parameter estimator is only the tool to expedite the time for calibration and not the underlying principle to achieve parameter calibration. The development of a calibration strategy for the hydrologic component of the HSPF that includes the use of a model-independent-parameter estimator is therefore necessary.

1.2 RESEARCH GOAL AND OBJECTIVES

Model calibration is sensitive to four factors: (1) the data base that describes the watershed; (2) the model, its complexity, structure, and constraints; (3) the objective function (s) used to define "best fit"; and (4) the constraints placed on the parameters. In assessing the quality of a calibration attempt, the process of fitting should be judged on the rationality of calibrated parameters and on the accuracy of the model predictions. A comprehensive calibration strategy must assess both inputs and outputs, where the inputs are not limited to the quality of the measured hydrologic data, but to a comprehensive examination of the hydrologic processes in the watershed and the representation of such processes in the model.

In order to develop a calibration strategy that in nature is replicable, it is necessary to use a model-independent-parameter estimator. This will allow for the replacement of the subjective calibration with a more objective procedure. The goal of this research is to develop and test a systematic procedure for calibrating the hydrologic component of the Hydrologic Simulation Program-FORTRAN (HSPF) in a way that provides reasonable assurance of the highest possible accuracy for regional studies at small scale and for TMDL development.

The successful development of a calibration strategy capable of accomplishing the objectives of this research will advance the state of the art in the broad field of model calibration. Although the strategy will be developed using the HSPF model, the procedure will be applicable to the calibration of complex hydrologic models. To meet this goal, the following specific objectives were analyzed:

1. To assess the importance of the spatial and temporal scale of rainfall input that drives the in-land and in-stream processes in applications developed with HSPF.
2. To develop an objective function for the model-independent-parameter estimator (PEST) that will be a function of the streamflow proportions.

3. To examine the sensitivity of the HSPF parameters used for the hydrologic calibration, including the assessment of calibrating monthly parameter values over the calibration of annual parameter values.
4. To examine the effect of error in the estimated initial HSPF soil storages on the prediction accuracy, especially for short record lengths.
5. To show the effect of using stationary land-use data on a watershed that is undergoing land use change on the prediction accuracy of the mean daily runoff.
6. To assess the effect of a nonspatial distribution of land use on the prediction accuracy with HSPF.
7. To provide a rational strategy for the calibration of the HSPF model.

1.3 IMPLICATIONS OF RESEARCH

The fulfillment of the research objectives will provide a systematic procedure that improves the hydrologic calibration of the HSPF model; presents a systematic method to determine the accuracy and quality of the predictions; and exposes some of the limitations in the current approach to modeling. A systematic calibration strategy should enable HSPF to be applied such that analyses by different users are more consistent and optimum accuracy is achieved given the data base. This will reduce problems associated with HSPF being used by those unfamiliar with basic principles of modeling.

With the use of a model-independent-parameter estimator for the development of the calibration strategy and for the calibration of the model, subjective calibrations can be replaced with more objective methods. The parameter estimator not only provides the means of calibration but the possibility of more in depth-analysis of the application. The time demanded by subjective calibrations can be better invested in the understanding of the model assumptions and their effect on the accuracy of the model predictions.

It is expected that in the end, managers using results from the application of the HSPF model will recognize the need of a calibration strategy as a way to reduce uncertainty in the model predictions. Furthermore, the recognition of uncertainty in the model predictions is expected to trigger a change in the current modeling approach by incorporating additional analysis of uncertainty as part of the standard modeling process. Fair policies and regulations need to be based on results from models that are the product of a systematic calibration strategy.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The literature review specific to this research is aimed to examine some of the most important publications related to the calibration of hydrologic models and basic documentation of both, the HSPF model (Bicknell et al., 1993) and the model-independent-parameter estimator PEST (Doherty, 2001). In addition, documentation describing some of the mathematical methods available in the parameter estimator PEST and used in the research, are presented.

2.2 EFFFECT OF THE SPATIAL AND TEMPORAL SCALE OF THE INPUT DATA ON CALIBRATION

Whether using subjective calibration or inverse modeling methods for calibration, the spatial resolution and temporal scale of the input data are factors in the prediction accuracy. These elements affect the level of association between the calibrated parameter values and the physical processes represented by the parameters. Holman-Dodds (1998) studied the effect of the scale of input precipitation in hydrologic model calibration by testing values of the infiltration capacity with three hydrologic models. The results indicated that smaller parameter values resulted when the sampling interval of precipitation increased, to compensate for lower rainfall intensities that resulted from the smoothing of the precipitation signal. In some cases however, the calibration did not completely compensate for the loss of temporal variability in the precipitation inputs.

Daily depths for example is the most common scale of precipitation, yet, when used as input data for applications with the HSPF model the data is disaggregated to hourly values. The disaggregation process introduces inaccuracies to the predictions, not only because of the change of the scale, but in estimating the accurate times when the rainfall occur. In a similar sense, the scale at which land use is aggregated based on strong relations, sometimes in systematic ways, within and among variables that are near to one another introduces inaccuracies to the predictions. For example, the data variation in spatial resolution results in erroneous forecasting of urban growth when using coarse data and influences the calibrated value of the parameters (Dietzel, 2003). The dilemma in the calibration of lumped models is not only to fit model parameters that ignore most of the spatial and temporal variability in the watershed, but to predict data that is temporally and spatially fixed and that reflects the response of distributed processes.

2.3 HYDROLOGIC SIMULATION PROGRAM - FORTRAN (HSPF)

The Hydrologic Simulation Program - FORTRAN (HSPF) model uses computer technology to simulate watershed hydrology and water quality in natural and man-made systems, and it is thought to be the most comprehensive management tool presently available. Its origin dates to 1959 with the Stanford Watershed Model (SWM) and the Hydrocomp model developed by Hydrocomp Inc. Although the HSPF model is not conceptually difficult to understand, it is highly parameterized, including hydrologic parameters for each analyzed subwatershed.

The HSPF model represents the environment by using elements that consist of nodes and zones. A node corresponds to a point in space, while a zone corresponds to a finite portion of a watershed and is characterized by storage. A channel reach is simulated as a one-dimensional element made of a single zone that is situated between two nodes. As for the modeling of the land phase of the hydrologic cycle, it is simulated through a third type of element called unit-segment. A unit-segment is a portion of land that is assumed to have uniform properties without nodes, but with a number of zones. Fixed rules that govern the grouping of zones and nodes to form elements are based on the similarity of characteristics such as soil type, slope, land use, and others, commonly determined through the use of GIS methodologies. In there a single parameter structure applies to all elements that are conceptually identical in a unit-segment. For example, in the case of unit-segments used to simulate in-land processes, variations between segments are represented only by variations in the values of parameters. The same applies to another element such as a reach, or any other finite element.

The HSPF Model simulates the fate and transport of pollutants over the entire hydrologic cycle. Two distinct sets of processes are represented in HSPF: (1) processes that determine the fate and transport of pollutants at the surface or in the subsurface of the land areas of a watershed, and (2) in-stream processes. The former will be referred to as land or watershed processes, while the latter as in-stream or river reach processes.

The representation of constituents can be done at various levels of detail, with the option of simulating them for both on-land and in-stream environments. These choices are made, in part, by specifying the modules that are used, and thus the choices establish the model structure used for any problem. In addition to the choice of modules, other types of information must be supplied for the HSPF calculations, including model parameters and time-series of input data. Time-series of input data include meteorological data, point sources, reservoir information, and other types of continuous data as needed for the application. Different modules can be used to represent the transport of constituents in pervious and impervious land (impervious land does not involve subsurface transport). Choices in the representation of the transport of constituents in river reaches are few, but the fate of in-stream constituents can also be made more or less complex by the choice of different modules.

2.3.1 Modeling of a Model (Land) Segment

Multiple land use types can be represented in a single HSPF model, each using different types of modules and different model parameters. A model segment is a subdivision of the simulated watershed, and it is commonly defined as an area with similar hydrologic characteristics. In terms of modeling, all processes are computed for the spatial unit (1-acre). To obtain information for a model segment, the number of acres of each particular unit type in the land segment are multiplied by the values (fluxes, loads, and other processes) computed for the corresponding spatial unit. Although the modeling is performed on a temporal basis, there are few applications where land use information changes with time.

$$ROVOL = (K_S * ROS + COKS * ROD) * DELTS \tag{2-1}$$

where K_S is a weighting factor $(0.00 \leq K_S \geq 0.99)$; $DELTS$ is the modeling interval in seconds; $COKS$ is the complement of $K_S (1 - K_S)$; ROS is the total rate of outflow at the start of the interval; and ROD is the total rate of demanded outflow at the end of the interval.

2.3.2 Modeling Hydrologic Processes

The modeling of the overland flow in pervious areas is performed with the PWATER module (Figure 2-1) through a linked set of theoretical and empirical mathematical functions. Land surface (PERLND) and soil processes of the hydrologic cycle including interception, surface detention, soil moisture storage, surface runoff, infiltration, interflow, evapotranspiration, percolation to groundwater, and groundwater outflow are simulated in the hydrologic representation within HSPF. In addition, energy (heat) balance calculations based on input meteorological data (for example, cloud cover, radiation, wind speed, air temperature, and evapotranspiration) are performed using the SNOW module that determine snow accumulation and melt.

The formulation of the processes that controls the water budget in the HSPF model includes parameters that can be set on an annual or monthly basis. For the analyses presented in this document and because of the uncertainty and the lack of monthly data for storages and rates of infiltration, evapotranspiration, and flow recession, and to simplify the assessment, it was decided to perform the tests using parameter values that do not vary monthly. The parameters included in the analyses are shown in Table 2-1.

Table 2-1. **Parameters that control the water budget in pervious areas**

Parameter	Units	Parameter description
LZSN	Inches	Lower zone nominal storage
INFILT	inches/hr	Index to the infiltration capacity of the soil
AGWRC	day^{-1}	Basic groundwater recession rate if KVARY is zero, and groundwater does not receives inflow.
NSUR	Complex	Manning's n for the assumed overland flow plane
INTFW	None	Interflow inflow parameter
IRC	day^{-1}	Interflow recession parameter
LZETP	None	Lower zone E-T parameter. It is an index to the density of deep- rooted vegetation
DEEPFR	None	Fraction of groundwater inflow that will enter deep (inactive)
BASETP	None	Fraction of remaining potential E-T that can be satisfied from baseflow (groundwater outflow), if enough is available.
AGWETP	None	Fraction of remaining potential E-T that can be satisfied from active groundwater storage if enough is available.
UZSN	Inches	Upper zone nominal storage

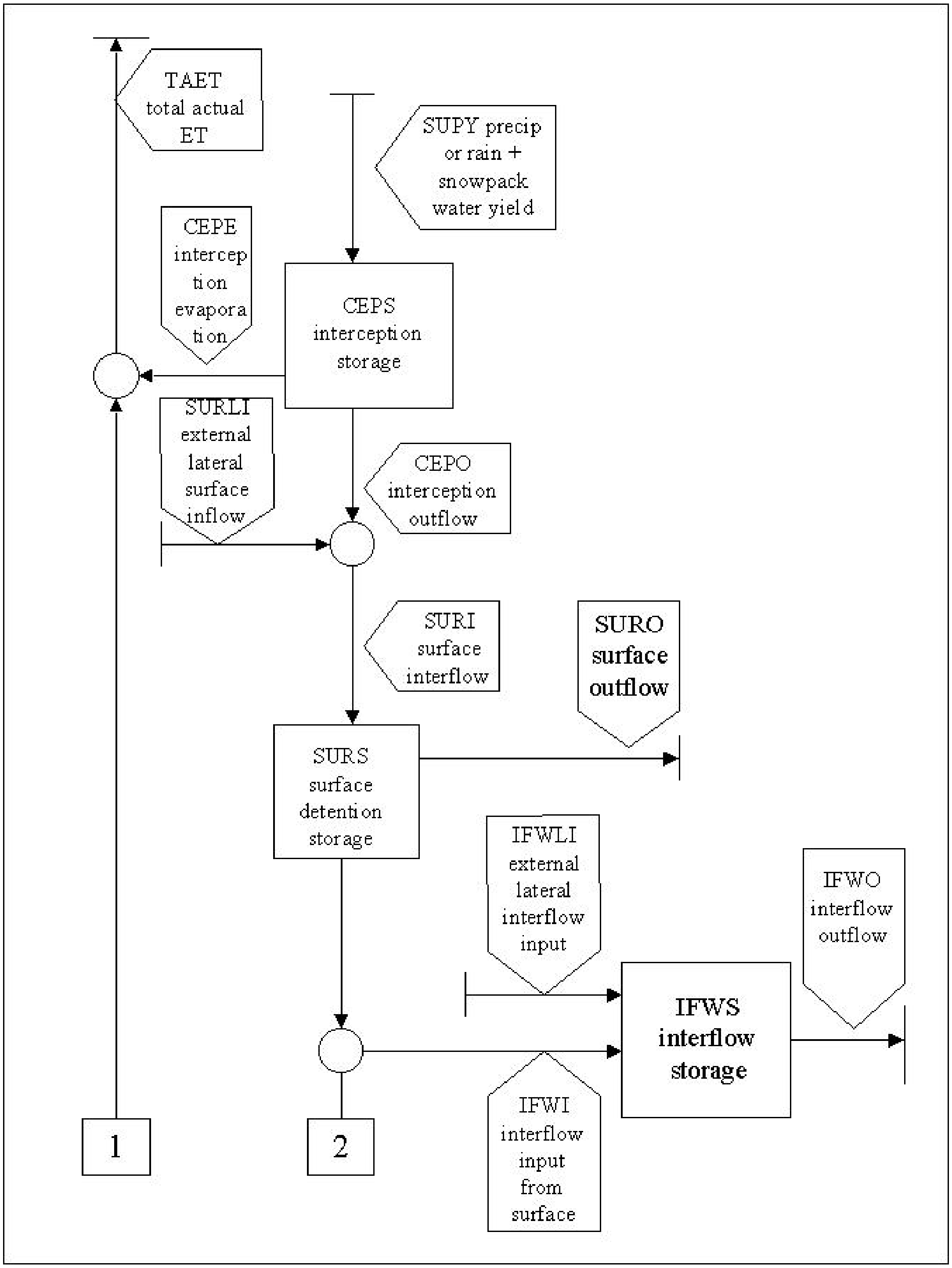
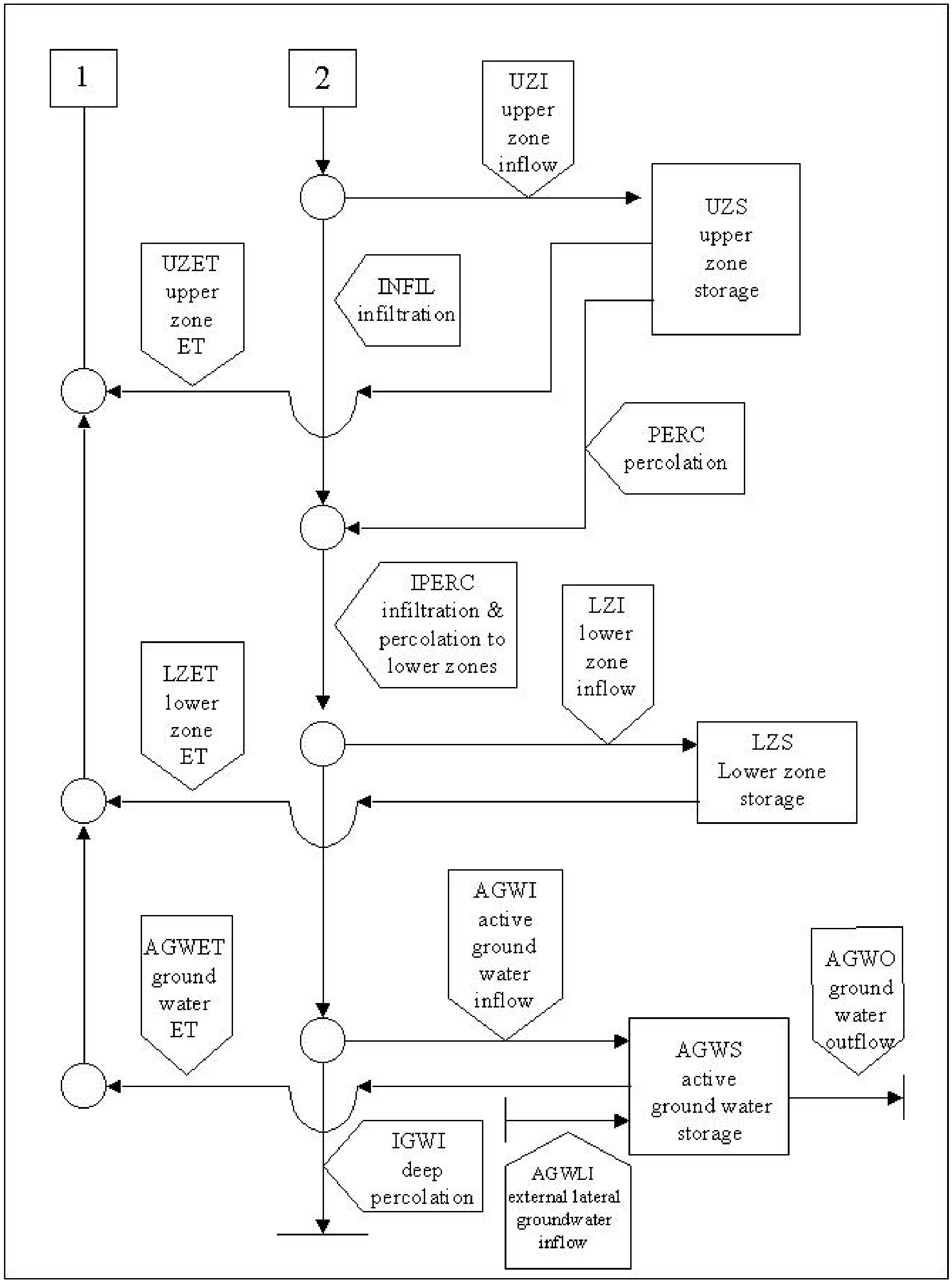


Figure 2-1. PWATER Module Diagram



Continuation Figure 2-1 PWATER Module Diagram

Source: HSPF manual

The modeling of the overland flow in impervious area is performed with the IWATER module (Figure 2-2) through a linked set of theoretical and empirical mathematical functions. Land surface (IMPLND) processes of the hydrologic cycle include retention storage, surface detention storage, surface runoff, and evaporation. Unlike in the pervious areas, there is no parameter regulating the rate of evaporation from the retention storage and the demand will draw upon all of the interception storage unless the demand is less than the interception storage.

The overland flow in the pervious and impervious area is treated as a turbulent flow process. It is simulated using the Chezy-Manning equation and an empirical expression, which relates outflow depth to the detention storage. The model contains two equations to calculate the rate of overland flow discharge. Eq. (2-2) is used when the rate of the overland flow is increasing. Eq. (2-3) is used when the surface is in equilibrium or receding.

$$SURO = \Delta 60 * SRC * (SURSM * (1.0 + 0.6(SURSM / SURSE)^3)^{1.67} \qquad \textbf{(2-2)}$$

$$SURO = \Delta 60 * SRC * (SURSM * 1.6)^{1.67} \qquad \textbf{(2-3)}$$

where $SURO$ is the surface outflow (in./interval), $\Delta 60$ makes the equations applicable to a range of time steps Δ (hr/interval), SRC is a routing variable, $SURSM$ is the mean surface detention storage over the time interval (in.), and $SURSE$ is the equilibrium surface detention storage for the current supply rate (in.). The equilibrium surface detention storage is calculated as follows:

$$SURSE = ((0.00982 * (NSUR * LSUR / \sqrt{SLSUR})^{0.6})SSUPR^{0.6} \qquad \textbf{(2-4)}$$

where $NSUR$ is the Manning's n for the overland flow plane, $LSUR$ is the length of the overland flow plane (ft), $SLSUR$ is the slope of the overland flow plane (ft/ft), and $SSUPR$ is the rate of precipitation to the overland flow surface. The routing variable SRC is calculated with the following equation:

$$SRC = 1020.0 * \sqrt{SLSUR} / (NSUR * LSUR) \quad \textbf{(2-5)}$$

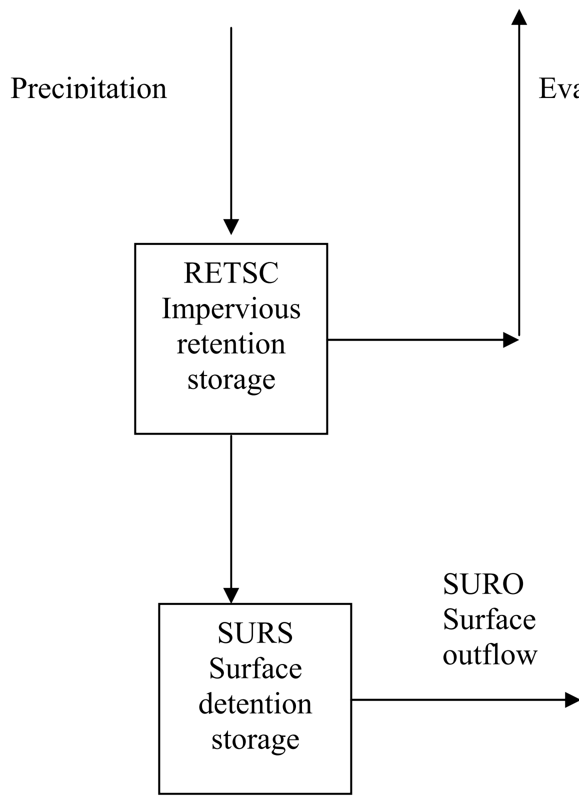


Figure 2-2. IWATER simplified module diagram

Source: HSPF manual

2.3.3 Physical Interpretation of Parameter AGWRC

Using the Master-Depletion-Curve Method (Eq. 2-6), which provides a model of flow from groundwater storage and information on the recession limb, the number of days for the groundwater to recede can be calculated for values of AGWRC between 0.1 and 0.999:

$$Q_t = Q_0 e^{-\frac{t}{k}} \tag{2-6}$$

where Q_t is the discharge at time t, Q_0 is the discharge at time t=0 , and k is the number of days for the water to recede. Eq. 2-6 can be rewritten as:

$$\frac{Q_t}{Q_o} = AGWRC \tag{2-7}$$

where $AGWRC$ is the daily recession constant of groundwater flow, that is, the ratio of current groundwater discharge to groundwater discharge 24 hours earlier. Replacing *AGWRC* in Eq. 2-7 and solving for *k,*

$$AGWRC = e^{-\frac{t}{k}} \Rightarrow \ln AGWRC = -\frac{t}{k} \tag{2-8}$$

For t=1, and solving for *k*:

$$k = \frac{1}{-\ln AGWRC} \tag{2- 9}$$

Table 2-2. Value of the parameter AGWRC and the corresponding number of days for the groundwater to recede.

AGWRC	Receding rate: $k(days)$
0.9990	999.5
0.9894	93.8
0.9800	49.5
0.9700	32.8
0.9690	31.7
0.9604	24.7
0.9500	19.5
0.9310	14.0
0.8000	4.5
0.7000	2.8
0.6000	2.0

The allowable minimum and maximum values recommended in the HSPF manual for the parameter AGWRC are between 0.001 and 0.999. This range of values represents the number of days for the groundwater to recede. For values between 0.001 and 0.970 the groundwater recession varies between hours to days, while for values greater than 0.970 and below 0.999, the groundwater recession varies between months to years. Increasing the value of AGWRC slows down the groundwater outflow and flattens the slope of the baseflow curve.

2.3.4 Modeling of a Reach

Within the HSPF program, the RCHRES module sections are used to simulate hydrology, sediment transport, water temperature, and water quality processes that result in the delivery of flow and loadings to a body of water, such as a bay, reservoir, or the ocean. Flow through a reach is assumed to be unidirectional. In the solution technique of normal advection, it is assumed that simulated constituents are uniformly dispersed throughout the waters of the RCHRES and move at the same horizontal velocity than the water. The inflow and outflow of materials is based on a mass balance.

The HSPF uses a convex routing method to move mass flow and mass within the reach (Eq. 2-1). Outflow may leave the reach through one of the five possible exits, and the processes that occur in the reach will be influenced by precipitation, evaporation, and other fluxes:

2.4 EVALUATION OF CONTINUOUS BASEFLOW SEPARATION TECHNIQUES

Baseflow separations procedures using digital filters are commonly applied to daily streamflow values where the low-amplitude, low frequencies are associated to the baseflow component and the high-amplitude, high frequencies are associated to the quickflow component. In areas where daily values are not available but monthly volumes are, the application of such methods is difficult because the short-term flow variability is lost in the monthly values. Hughes and Watkins (2003) investigated the effect of the long-term average baseflow responses from both daily and monthly volumes using an algorithm similar to that found in PEST. The results indicated that the applied digital filter was most effective when using short-time steps. At the monthly scale, the individual events were lost and in some cases, a single season looked like a single event. This indicated that by using monthly data, the baseflow could be underestimated and, thus, to generate a higher volume of baseflow the value of the baseflow recession constant (α) needed to be reduced. Two of the most common baseflow separation techniques are reviewed in this section.

2.4.1 Local-Minima Technique for Hydrograph Separation

The local minima technique is a series of simple rules for the separation of the baseflow portion from the total streamflow hydrograph. Several versions of this method exist including the version from the Institute of Hydrology, U.K. (1980), and the version developed by Pettyjohn and Henning (1979), implemented by the U.S. Geological Survey in the program HYSEP.

The technique implemented in PEST uses the daily streamflow values, and it is described as follows: First, a sampling interval with an odd number and a minimum value of 3 days is selected. Within each interval, a comparison between the streamflow for the centered day and the streamflows for the two adjacent days are made. If the value for the centered day is the lowest of the three values within the interval, then it is selected as a local minimum, otherwise, it is assumed that a local minimum is not contained within the interval. Then, the interval slides over to the next day and evaluates the new interval. The days selected as local minimum values are joined by a series of straight lines, and the values between the minima are obtained by linear interpolation. Because the days preceding the center of the first interval as well as the days after the center of the last interval are not assessed, the closest local minimum values are assigned to these non-evaluated days. In contrast to the method implemented in HYSEP where the size of the interval is a function of the drainage area, the method implemented in PEST gives the user the option to set the interval size for the calibration process and, thus, to select the best interval for the simulated watershed regardless of the watershed area.

2.4.2 Sliding-Interval Method for Hydrograph Separation

As the local-minima technique, this method is also a set of rules to determine the baseflow portion of the total streamflow hydrograph. In the PEST parameter estimator, the selection of the interval size is not set by the watershed area, rather it is selected by the user. The interval sizes are always odd numbers with a minimum value of 3 days. The total streamflow values are compared within the interval, and the lowest value within the interval is designated as the baseflow on the day in the middle of the interval. The interval then slides over to the next day, and the process is repeated for the new interval. Because the days preceding the center of the first interval are not evaluated, the value assigned to the center of the first interval is assigned to the non-evaluated days. A similar situation is experienced at the end of the series with the subsequent days to the center of the last interval not being evaluated; in this case, the last assigned value in the series is used for the non-evaluated days.

2.5 METHOD OF RAINFALL DISAGGREGATION

METCMP was developed by the USGS as part of a grant from EPA, and it is a compilation of methods to manipulate meteorological data on a temporal basis including computing daily solar radiation, pan evaporation, daily potential evaporation, distributing daily information into hourly values, etc. The software manipulates data (for example, fill-in missing values, disaggregate daily into hourly values, etc.) from a primary station based on information contained in databases from secondary stations.

2.5.1 To Fill-in Missing Precipitation Data

This procedure finds all missing values in a primary data set and fills them in, using user-specified secondary data of the same time step. If more than one secondary station is used, the fill-in value is computed as a weighted average by one of the following methods: 1) by the reciprocal of the distances to the primary station; 2) by the reciprocal of the squares of these distances; 3) by equal weights; 4) by user-defined weights. Missing values are filled in by the weighted adjusted sum of the secondary stations.

2.5.2 Disaggregate Daily Precipitation to Hourly Values

The program also distributes accumulated values in a primary data set using information from a secondary data set. In order to disaggregate daily values, the station may not contain missing periods. The values are distributed according to a secondary station whose total precipitation over the accumulated period is closest to the total precipitation value at the primary station. If none of the stations had data within a user-defined accumulated data tolerance (percent of missing values), the accumulated values in the primary data set are distributed according to a symmetric triangular distribution in which the maximum allocation of rainfall depth will be at noontime. The accumulated data tolerance (0 – 100%) represents the allowable range of daily totals from the hourly stations, expressed as a ratio of the accumulated precipitation to the daily value being distributed (for example, zero percent means that the daily total from the hourly stations must match the daily value to be distributed exactly). If the daily total for the hourly station being used is zero, but the daily station is nonzero, the data is distributed evenly over the day (flat). If none of the hourly stations has good quality data over the whole day, or if none has a daily total within a user-specified tolerance of the daily station total, the output is written with 23 hours of the accumulated data code, followed by the daily total.

2.6 MODEL-INDEPENDENT-PARAMETER ESTIMATOR (PEST)

The use of mathematical formulations to model physical processes are of special interest for the management of the natural resources in the world, yet the complexity of these applications may result in an incorrect or inadequate solution to the initial modeling objective. The fitting process in the HSPF is commonly made through subjective calibration, but automatic methods are now possible. The disadvantage of subjective calibration is that the process is not reproducible, it is time consuming, and the evaluation of potential calibration scenarios when varying model or data assumptions is very limited. In contrast when applying inverse modeling methods, the results are replicable and the process of adjusting parameter values is expedited. The replication of model results is one of the most important factors where parameter estimators such as PEST (Doherty, 2001) can reduce the amount of uncertainty. Although the calibration of the model parameters is not entirely objective because the calibration criteria are specified by the user, the tool makes the calibration a more controlled and replicable process.

PEST works with existing models as a tool for the interpretation of data, parameter calibration, and predictive analysis, yet PEST can exist independently of any particular model. For this study, PEST was used for parameter calibration. PEST uses the principle of the weighted least-squares (Carroll and Ruppert, 1988) applied to the objective function and an iterative process. The model parameters are adjusted using a nonlinear estimation technique known as the Gauss-Marquardt-Levenberg (Levenberg, 1944 and Marquardt, 1963) method and the single value decomposition. The Gauss-Marquardt-Levenberg technique is based on the linearization of the relation between model parameters and model predictions at the beginning of all iterations. The linearization is formulated as a Taylor expansion about the best parameter set. In the Taylor expansion, the derivatives of the predictions with respect to the parameter values are calculated to obtain a new set of parameter values that have a smaller error squared than at the most recent iteration. The new parameter values are tested by running the model again and by evaluating the improvement of the calibration as a reduction of the objective function. This process of parameter adjustment and evaluation of the objective function value is repeated until the objective function does not decrease. This set of parameters is assumed to be the optimized set.

The parameter estimator PEST contains mathematical expressions and methodologies such as signal analysis and hydrograph separation that can be used as part of the calibration criteria. Within the digital signal analysis, PEST has the capability to perform digital filtering only to continuous data and constant time step, for example, daily total flow, by using a Digital Control System, discussed later on in section 2.8. The purpose of filtering the actual and predicted total flow is to separate high from low frequencies and to remove the random component. Two commonly used methods for hydrograph separation are included in PEST and are discussed in previous sections: the Local-minimum and the Sliding-Interval methods.

Provided that PEST includes three methods to locate the optimum parameters it was of interest to determine the optimum method of calibration, when using the model-independent-parameter estimator PEST in combination with the HSPF model. The methods are: (1) the Gauss-Marquardt-Lambda method (M-L), (2) the single value decomposition method (SVD), and (3) the single-value decomposition method-Assist (SVD-A). The Gauss-Marquardt-Lambda (M-L) method is an iterative method to minimize the sum of squares of M functions in N variables. It requires the finite-difference approximation of the Jacobian matrix in all iteration and uses the Jacobian matrix to minimize the sum of squares in a local region of parameter space. This method uses the method of Steepest Descent for finding the nearest local minimum of the objective function. A problem that can be encountered with the M-L method is when the Jacobian matrix is singular or else numerically very close to singular, thus the Jacobian matrix cannot be transposed. For these cases, the singular value decomposition technique will provide a useful numerical answer, although, not necessarily the expected answer, as the SVD may be able to detect weak patterns in the data that may be associated with specific hydrological processes and parameters equations.

The SVD method yields linearly coupled patterns which maximize the explained cross-variance between two time-dependent data sets. SVD uses also the Jacobian matrix and because of its capability to detect weak signals in the data, it allows for the attainment of the true dimensionality of the data. The procedure in PEST is as follow: In a single and independent iteration to the calibration process, referred to as the iteration for the matrix-rank estimation, the first-order sensitivity of the parameters is calculated and the maximum number of singular values to be used in the solution is determined.

The SVD method calculates the first-order-sensitivity of the parameters at all iterations to determine the number of singular values to use in the solution. The SVD-A is an extension of the SVD method with the difference that the first-order-sensitivity of the parameters is calculated only once at the matrix-rank estimation iteration. However the modeler can decrease the number of singular values to find the solution rather than using the number determined by PEST during the matrix rank estimation.

When using the M-L method, the HSPF parameters are individually calibrated. When using the SVD or SVD-A method, PEST calculates the first-order-sensitivity of the HSPF parameters during the matrix-rank estimation iteration to determine the number of singular values to be used in the solution. At the end of the calibration, the process is reversed and the HSPF parameters as such, are computed.

2.7 METHOD OF LEAST SQUARES

The calibration of hydrologic models is based on the concept of curve fitting for n points, in which each point corresponds to a pair of values composed of the actual and predicted discharges. Although the principle is not always the most appropriate whenever a goodness-of-fit measure is needed, the method is widely applied to hydrologic problems. In a continuous case, a desired function $f(t)$ is to be approximated by an actual function $f^*(c,t)$ in which c is an adjustable parameter. Depending on the complexity of the problem, $f(t)$ such as that $f(t_j) \approx y_j$, $j = 1,....,n$, is used for predicting values of y.

The least squares method is used by PEST to evaluate the improvement of the optimization criteria and, thus, to evaluate the state of the calibration. Since $f^*(c,t)$ is an approximation of the desired function $f(t)$, the total squared error will be a function of c and is calculated as:

$$E^2_{(c)} = \int_{-\alpha}^{\alpha} \left[f(t) - f^*(c,t) \right]^2 dt \qquad \textbf{(2-10)}$$

In the case where the daily flows represent a discrete data set of a finite size of N samples, and if the parameter c can be optimized in such way that the squared error is the minimum, then the sum of the errors squared can be expressed as:

$$E^2_{(c)} = \sum_{n=0}^{N-1} \left[f_n - f^*(c,t_n) \right]^2 \qquad \textbf{(2-11)}$$

where $f(n)$ is a set of values of $f(t)$ $\left[f_o, f_1, \ldots, f_{N-1} \right]$.

In the case of calibrations using PEST, each term in Eq. 2-11 is multiplied by a weight to account for the relative accuracy of the measurement f_n (Eq. 2-11 assumes that the weights are 1.0). As demonstrated later in the analyses, these weights are an important factor in achieving a successful calibration using PEST. Taking into consideration that the calibration of the hydrology in the HSPF model is a function of more than one parameter, then Eq. 2-11 can be expressed as

$$f^*(c,t) \equiv f^*(t) = \sum_{m=0}^{M-1} c_m \phi_m(t) \qquad \textbf{(2-12)}$$

where $f^*(t)$ is a combination of a set of functions $\left[\phi_0, \phi_1, \ldots, \phi_{M-1} \right]$, M is the number of adjustable parameters $c_0, c_1, \ldots, c_{M-1}$. In this case, the minimum value for E^2 is obtained by solving the partial derivatives of E^2 with respect to the c_k coefficient and setting the derivatives equal to zero. Replacing Eq. 2-12 in Eq. 2-11 the total squared error is expressed as:

$$E^2(c_m) = \sum_{n=0}^{N-1}\left[f_n - \sum_{m=0}^{M-1} c_m \phi_m(t_n) \right] \qquad \textbf{(2-13)}$$

The partial derivatives can be expressed as a set of M linear equations:

$$\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} c_m \phi_{mn}\phi_{kn} = \sum_{n=0}^{N-1} f_n \phi_{kn} \qquad \textbf{with k=0,1,....M-1} \qquad \textbf{(2-14)}$$

Unfortunately, the number of solutions for the set $[c_m]$ that minimizes E^2 depends on the nature of the function set $[\phi_n]$ as well as the sample set f_n.

2.8 DIGITAL SIGNAL ANALYSIS (DSA)

The term digital signal analysis refers to the interpretation of signals produced by time-varying physical processes. A short review of digital signal analysis is included in this document as DSA can be used for hydrograph separation and because two digital filters are implemented in the model-independent-parameter estimator PEST: the Butterworth filter (Butterworth, 1930) and the Baseflow-separation filter (Nathan and McMahon, 1990).

The signal separation is made using a Digital Control System (Figure 2-3) that produces an analog signal with the same units as the input analog signal. In the case of the daily flows, the input and output units of the signals are in cubic feet per second (cfs). The signal is separated into deterministic and random components. It is assumed that the baseflow portion can be related to the deterministic component and that the quickflow portion can be related to the random component of the digital signal.

Analog signals are converted to digital form by an analog-to-digital converter (ADC) and are sent to the Digital Signal Processor. In PEST and within the processor, the digital form is filtered using either the Butterworth filter or the Baseflow-separation filter. Once the data have been filtered, the output signals from the Digital Signal Processor are converted to an analog control signal by a digital-to-analog converter (DAC). The analog filtered signal from the Baseflow-separation filter is the quickflow component. For this reason, in future references to this process, it will be referred to as the Quick-flow filter. The filtering process is done to both the measured and model predicted data so that the principle of least-squares can be applied.

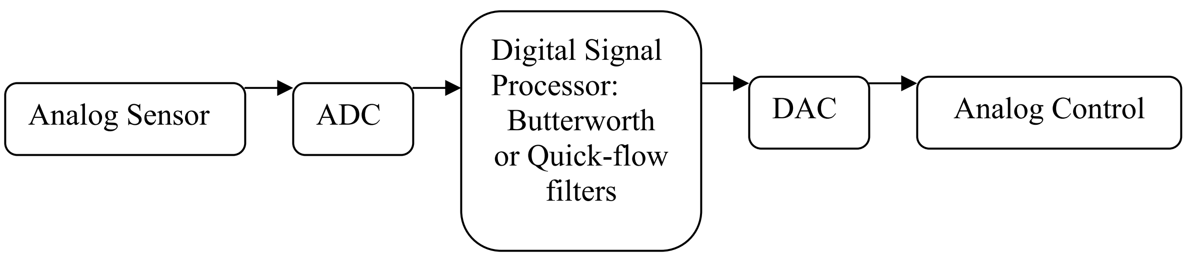


Figure 2-3. Digital Control System. Analog-to-Digital converter (ADC); Digital-to-Analog converter (DAC).

The assessment of the frequency-domain for the Butterworth filter was made through a nonrecursive algorithm in which the output signal is a function of the input signal, rather than a function of past computed values of the output signal. In contrast, the Quick-flow filter was implemented using a recursive algorithm and a zero phase shift. The phase shift is the angle of the transfer function in the s plane and it is described in section 2.10.

2.9 ANALOG AND DIGITAL FILTERING

A filter is defined as a system that operates on input functions. The purpose of the functions is to pass the spectral content of an input signal in a specified band of frequencies through the filter. Mathematical models of signals are classified as transient or steady-state, as shown in the voltage output of an audio oscillator (Figure 2-4). The transient part of the signals can be modeled as exponentially saturating and decaying sinusoids while the steady-state part can be modeled as a periodical sine function.

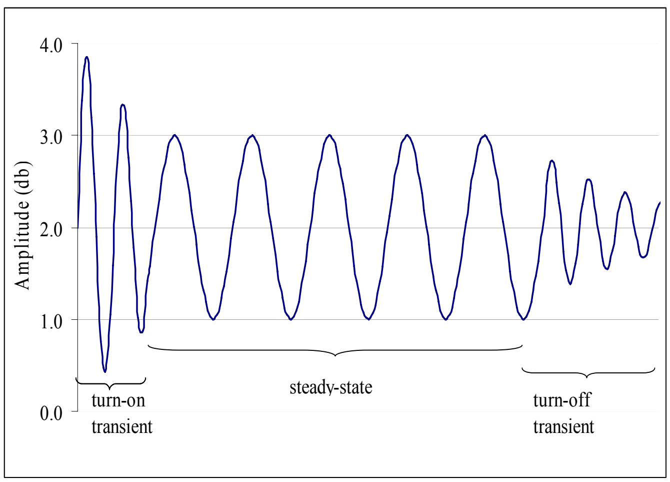


Figure 2-4. Output of an audio oscillator

The transient signals can be found at the beginning and/or end of a periodic signal as the turn-on transient section and the turn-off transient section. Within the steady-state signal, an ideal filter transmits frequencies in the specified pass-band only, without attenuation or phase shifting as observed in Figure 2-5. The ideal filter also presents brick-wall transitions between the pass-band and the stop-bands, for example, the brick-wall transitions in Figure 2-4 refers to a maximum and a minimum value in the output signal between of 3.0 and 1.0 db. Non-ideal filters, however, contain a transition region between the pass-band (values in the steady-state section) and the stop-bands (the values of the output signal immediately before and after the values in the steady state section), and the walls may not be flat (Figure 2-6) containing attenuations.

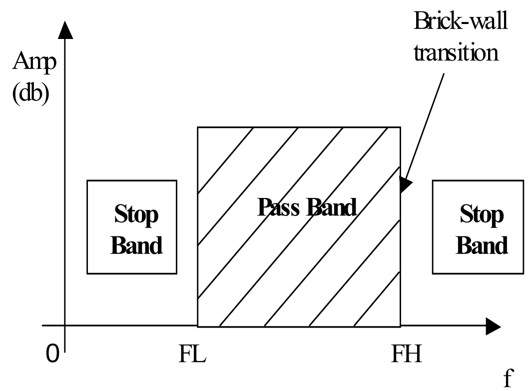


Figure 2-5. **Ideal Band-Pass Filter. FL is the lower cutoff frequency and FH is the upper cutoff frequency.**

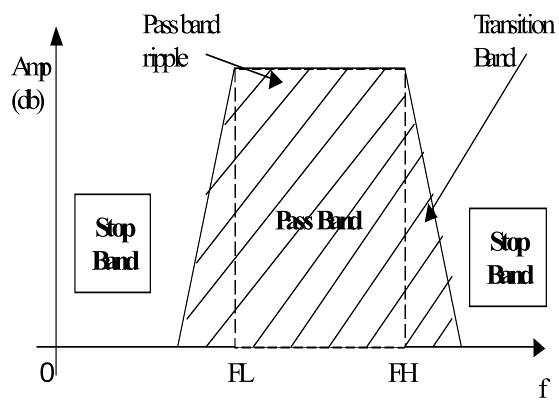


Figure 2-6. **Non-ideal Band-Pass Filter**

According to the location of the pass band in Figure 2-5, four basic ideal filters can be designed: (1) low-pass, (2) high-pass, (3) band-pass, and (4) band-reject. In the low-pass filters, the pass-band extends from 0 to FH, and the stop-band lies above FH.

In the high-pass filters, the pass-band extends above the FL while the stop-band goes from 0 to FL. In the band-pass filters, the pass-band extends between FL and FH, rejecting all signals outside of this range. Finally, the band-reject filter transmits all signals except those between FL and FH. These ideal filters can be implemented using equations that describe the various signal characteristics. Two of these approximations, the Butterworth filter and the Quick-flow filter, are included in PEST and thus, available for flow separation in hydrologic analyses.

2.10 BUTTERWORTH DIGITAL FILTER

In PEST, the Butterworth filter is used to separate the deterministic component of the runoff. Such signal is characterized by having maximum flatness in both the pass-band and the stop-band which provides the association with the baseflow component. The main goal in the design of a Butterworth filter is to approximate the ideal brick-wall transition between the pass-band and the stop-band frequency response, as observed in the ideal low-pass filter (Figure 2-7). The advantage of the Butterworth filter in relation to other alternative transfer functions is that it does not produce a pass-band ripple, thereby providing theoretically infinite attenuation as the frequency increases. However, the main limitation is that the slope of the brick-wall transition or the roll-off response is low.

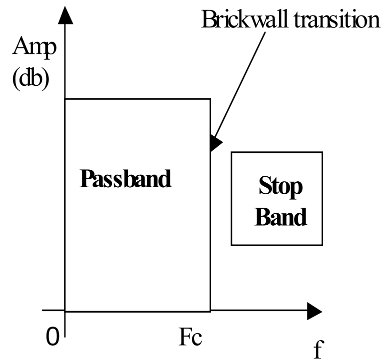


Figure 2-7. **Ideal Low-pass filter**

The analog transfer function relates the spectrum of the input signal to the spectrum of the corresponding output signal through polynomial equations. The transfer function is defined as the Laplace transform of the output divided by the Laplace transform of the input:

$$H_{(s)} = \frac{Y_{(s)}}{X_{(s)}} \quad \textbf{(2-15)}$$

where $Y_{(s)}$ is the Laplace transform (of the output $Y_{(t)}$, $X_{(s)}$ is the Laplace transform of the input $X_{(t)}$ and $s = j\omega$. The one-sided Laplace transform $X_{(t)}$ is defined by:

$$X_{(s)} = \ell^{-1}\left[X_{(t)}\right] = \int_0^{\infty} X_{(t)} e^{-st} dt \quad \textbf{(2-16)}$$

where $X_{(t)}$ is the input at time t, and s is the root of the polynomial in the denominator, also known as a pole.

In the case of the ideal low-pass filter, which is the basis for the Butterworth filter, the amplitude response or magnitude of the transfer function is:

$$H_{(s)} = \frac{1}{(s-s_1)(s-s_2)....(s-s_n)} \quad \textbf{(2-17)}$$

where $H_{(s)}$ is the amplitude response, n is the order of the transfer function, s_i is the location of the poles in the S plane, and n is the order of the filter and number of poles in the S plane. The poles are equally spaced only in the left half of the circle in the S plane because a pole in the right half plane causes instability, which means that the response to the transient input signal would increase rather than decay. Using Figure 2-8, the location of the poles can be defined as follows:

$$s_{(n)} = \cos\left[\pi \frac{2i+n-1}{2n}\right] + j\sin\left[\pi \frac{2i+n-1}{2n}\right] \quad \textbf{(2-18)}$$

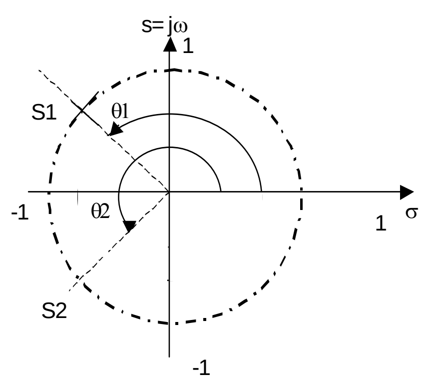


Figure 2-8. Pole locations in the S plane for a second-order Butterworth low-pass filter. Θ1= 1350 and Θ2= 2250

The bilinear transformation is used as a transformation from the s-plane to the z-plane, or from the analog filter design to the digital filter. The goal of this transformation is to improve the simulation of the rectangular passband shape by improving the power gain characteristic (Figure 2-9). The power gain $|H_{(s)}|^2$ is preferred over the amplitude ($H_{(s)}$) response to describe the characteristic of the filter. (Figure 2-9):

$$\left|H_{(s)}\right|^2 = 10\log_{10}\left|H_{(s)}\right|^2 \textbf{ Decibels} \tag{2-19}$$

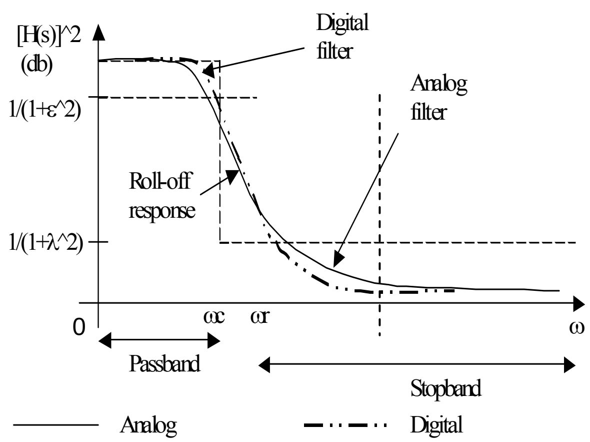


Figure 2-9. Characteristics of the Butterworth Low-pass filter

Theoretically, the response roll-off can be improved either by increasing the order of the filter or by increasing the roll-off rate. To increase the roll-off or the slope of the brick-wall transition, PEST provides the option of increasing the number of stages from 1 to 3, where 1 stage is 6 db/octave, 2 stages is 12 db/octave, and 3 stages is 18 db/octave. The positive aspect is that the roll-off response will better approximate the vertical characteristic of the brick-wall transition of the ideal low-pass filter. The down side of increasing the roll-off rate is that the phase of the output signal is shifted. Overall, this filter is not recommended to be used in hydrologic calibrations unless the implementation of the filter in PEST is changed to a zero phase shift. In terms of hydrology, this phase shift results in a poor correlation with the actual flow because the timing in which the high frequencies are filtered does not coincide with the timing in which the quickflow occurs in the actual flow data.

2.11 QUICK-FLOW FILTER

The Quick-flow filter available in PEST uses a recursive algorithm that is characterized by the use of past values of the output in the calculation of a present value. The algorithm includes delay elements that store a quantity at $t = (m - 1)T$ so that this same quantity is available at $t = mT$. However, the Quick-filter was implemented with a zero phase shift in which $m = 0$, which results in a better correlation between the output signal and the actual flow because the timing in which the high frequencies are filtered, coincide with the timing in which the quickflow occurs in the actual flow data.

The separation approach was initially reported by Nathan and McMahon (1990). The algorithm filters out the high frequencies passing only the low frequencies of the daily flows. The technique provides an automated, objective, and repeatable estimation of the stochastic component associated with the quickflow portion of the hydrograph. The concepts applied (for example, unit delays, stored numerical coefficient, etc.) and the properties (for example, transfer functions, impulse response etc.) of the non-recursive systems are also part of the recursive system. The linear algorithm of the contained Quick-flow filter in PEST is of the following form:

$$f_k = \alpha f_{k-1} + \frac{(1+\alpha)}{2}(y_k - y_{k-1}) \qquad \textbf{(2-20)}$$

where f_k is the filtered quick response signal at day k, y_k is the daily flow value for day k, and α is the filter parameter that affects the degree of attenuation and, thus, the predicted volume of quickflow. The values of α for the calculation of the quickflow have been experimentally calculated between 0.9 and 0.995; however, a trial-and-error analysis is recommended for each set of data.

Using this quickflow component, the baseflow portion of the hydrograph could be calculated as the difference between the total flow and the quickflow component; however, the exponential decay characteristic associated with the storage depletion of water in the soil is not present in this estimate of baseflow. To eliminate the negative values from the output signal, PEST provides the option of constraining the output to positive values.

The number of passes when using the Quick-flow filter controls the degree of smoothing and the phase distortion. PEST provides options of one or three passes. In the three-pass option, the reverse pass is done to nullify any phase distortion due to the forward pass of the filter.

CHAPTER 3

AN EVALUATION OF RAINFALL DISAGGREGATION METHODS

3.1 INTRODUCTION

Rainfall information is the driving force for many hydrologic calibrations, specifically the HSPF model (Bicknell et al. 1993) in which the inland and in-stream processes are driven by precipitation intensities; however, the density of rainfall stations that collect short-interval data is usually low, and when the data are not available at the necessary spatial and temporal scales, a common practice is to use information from distant stations as input to the HSPF. The uncertainty of the accuracy of this transferred information affects the predicted discharges and other values of other model processes. One of the implications of inaccurate precipitation data is that the modeling of processes such as erosion and sediment transport, in which the intensity of the rainfall is a primary factor in the accuracy of the predicted sediment, are poor or unreliable. Thus, the accuracy of predicted water quality components such as phosphorous, nitrogen, and other compounds attached to the sediments, would also be affected by the inaccuracy of the sediment modeling (Bergman et al. 2002).

In the case of sediment prediction, erosion is directly proportional to the rainfall intensity often to a power greater than 1.0. In this case, it may be expected that the use of mean daily rainfall intensities may lead to under estimation of sediment loads computed with models such as HSPF, while more accurate estimates of erosion can be expected with hourly rainfall intensities. Taking this into consideration, an important question arises: Which would provide greater accuracy, hourly intensities disaggregated from mean daily rainfall depths from a gage on-site or hourly intensities transferred from a gage outside of the watershed? Also, how does the accuracy of rainfall intensity estimates vary with the distance between the hourly gage and the watershed?

When hourly rainfall depths are needed as an input to a model, then measured daily total rainfall depths at the base gage, denoted as Y, can be directly disaggregated. As an alternative, measured hourly depths at a satellite gage, denoted as X, can be used to estimate the hourly values at the base station. A satellite gage is assumed to be a nearby station with measured hourly data and the source of information for the disaggregation. The following methods of disaggregation or transfer of rainfall data are alternatives for providing hourly estimates at sites where hourly data are not available: (1) the uniform disaggregation method, where measured daily depths are available at the base station Y only; (2) the weather pattern disaggregation method where measured daily depths are available at the base station Y and local meteorological distribution patterns are available for disaggregating the measured daily depths at the base gage Y; (3) the satellite transfer method where measured hourly data are available at a nearby satellite gage, with data not available at Y; and (4) the satellite ratio disaggregation method, where the measured daily data at the base site are disaggregated using hourly data from the satellite gage. Each of these cases will be assessed for accuracy. These four methods can be classified as either bivariate or univariate method, as well as either a transfer or disaggregation method (Table 3-1).

Table 3-1. Classification of transfer and disaggregation methods

Method	Univariate	Bivariate	Transfer	Disaggregation
Uniform Disaggregation	●			●
Weather Pattern Disaggregation	●			●
Satellite Transfer		●	●	
Satellite Ratio Disaggregation		●	●	●

Accurate rainfall data are important for model calibration. It is understood that rain at the stream gage does not drive the discharge, and that the proximity between the rain gage and the stream gage is less important than the proximity between the stream gage and a site representative of the watershed. However, if the location of the rain gage is assumed to be representative of the rainfall watershed, then as the distance between the rain gage and the stream gage increases, the rain gage data at the satellite station X is less able to represent hourly rainfall data at the base station Y. Bradley et al. (2002) showed that the spatial correlation of hourly rainfall decreased quickly with increases in the separation distance between stations. This is a source of error that can also affect the accuracy of calibrated model parameters, with the fitted parameters deviating from their true values solely because of a poor association between the measured on-site discharges and the measured rainfall transferred from a distant rain gage.

An objective of this analysis was to conduct systematic assessments using a reasonably dense network of measured rainfall data to assess the importance of the spatial and temporal scales of the rainfall. Methods for disaggregating daily rainfall are presented and evaluated. The accuracy of the disaggregation models can be assessed on the basis of their prediction of measured hourly rainfall depths. Socolofsky et al. (2001)

disaggregated daily rainfall estimates to obtain hourly values. Durrans et al. (1999) also disaggregated rainfall depths for use with complex models.

To achieve this objective, the following questions need answers: (1) Would the disaggregation of mean daily rainfall depths be more accurate than the transfer of hourly rainfall from a distant gage? (2) How does the accuracy of hourly rainfall vary with the distance between the satellite and base gages? and (3) Of alternative disaggregation methods, which provides acceptable accuracy?

3.2 DATA

To evaluate the study objectives, hourly rainfall records from 74 gages located within or near the Chesapeake Bay watershed were selected. The location (Figure 3-1) and distance between stations were determined using a Geographic Information System (GIS). To achieve a large sample size while minimizing time sampling error variation, data for a common period 1984-99 were used. The criterion of time sampling variation was used to reject data from a particular station when, within the testing period, more than 20% of the data were missing from the records. To make paired comparisons, a spatial limit of approximately 40 kilometers was arbitrarily set, which generated 87 station pairs. The minimum distance between stations was 1.9 kilometers, while the maximum distance was 41.4 kilometers. Ten years of hourly rainfall data for 87 pairs of stations were obtained for the period between 1984 and 1999. Although the length of the analyzed data for each pair was the same (10 years), the specific periods varied from pair to pair depending on the availability of the data.

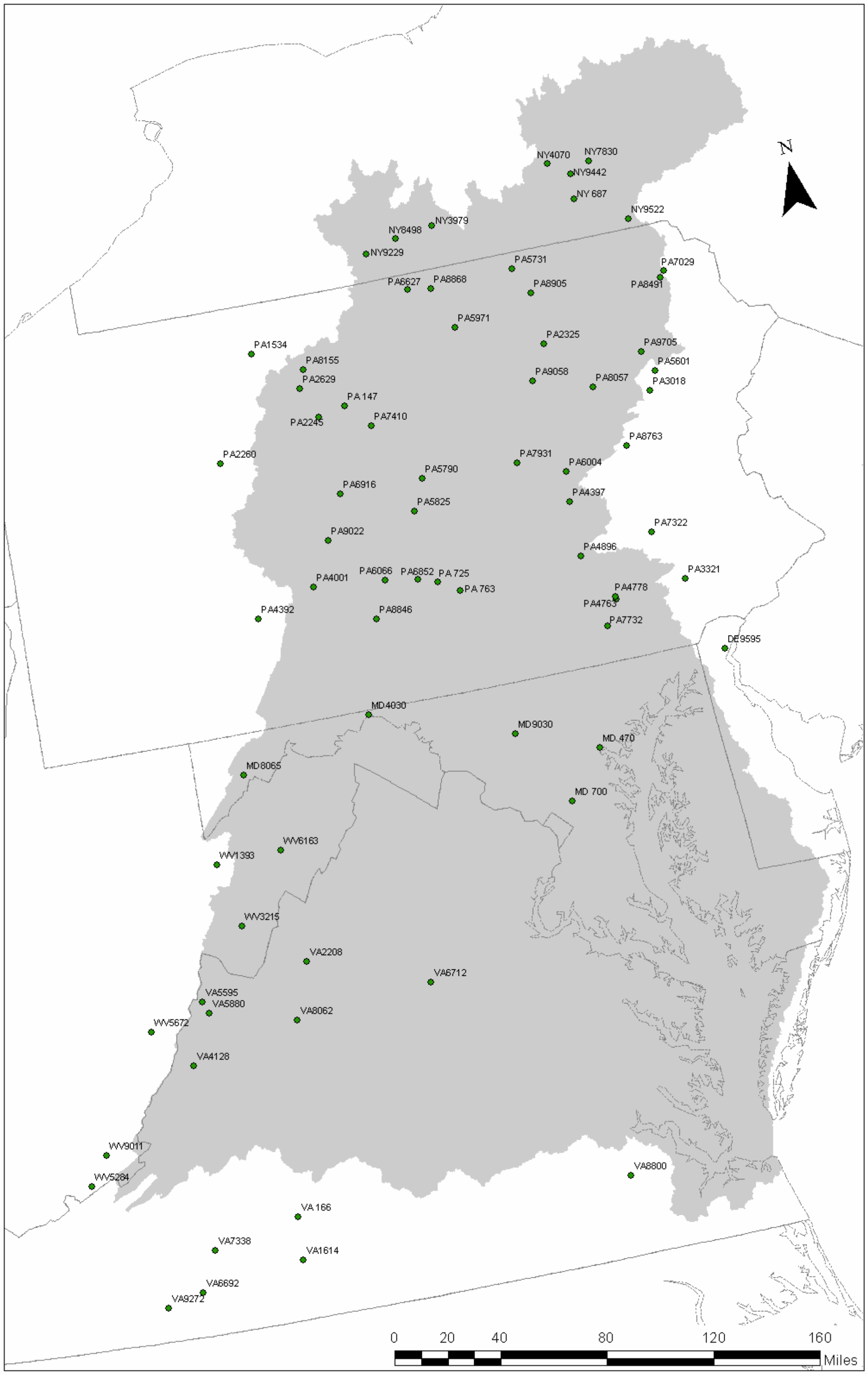


Figure 3-1. Chesapeake Bay watershed and NOAA stations

Storm cell movement was considered a potential factor in assessing the accuracy of transferring hourly rainfall information from one rain gage station to another. An analysis might show that the cross-correlation coefficient between hourly rainfall depths is higher for non-zero time lags. The exact time lag would reflect average storm cell movement. However, the correlation may be tempered by the lack of consistency in storm cell direction. To assess the potential for a significant temporal lag due to storm cell movement, data for station pairs located within 8 km of each other were used to develop cross correlations for lags of ±2 hr, ±1 hr, and 0 hr. The 8-km criterion yielded five station pairs. These cross correlations were computed for each day in the 10-yr record using the hourly rainfall depths. The proportion of statistically significant correlations was determined for each station pair. Critical values for the Pearson correlation (McCuen, 1985) coefficient for a 5% level of significance and degrees of freedom based on the sample size and the number of lags were used to find the percentage of days where the cross correlation was significant.

Figure 3-2 shows the percentage of days in the 10-yr period on which the cross correlation was statistically significant. The results indicate a lag time of zero would provide the best accuracy for the prediction of hourly rainfall depths. Four of the five analyses indicate that the cross correlation is greatest for lag 0. The one exception was for the station pair with a separation distance of 1.91 km, in which the cross correlation was higher for a lag time of +1 hr. However, for that station the percentage of days where the cross-correlation was significant for a zero lag time was essentially the same as for the 1 hour (0.262 vs. 0.293). Overall, the results indicate that storm cell movement in the region does not seem to be a factor in the accuracy of rain depth transfer, and the best transfer of rainfall depths will be for the same period. Therefore, all subsequent analyses were based on zero-lag computations.

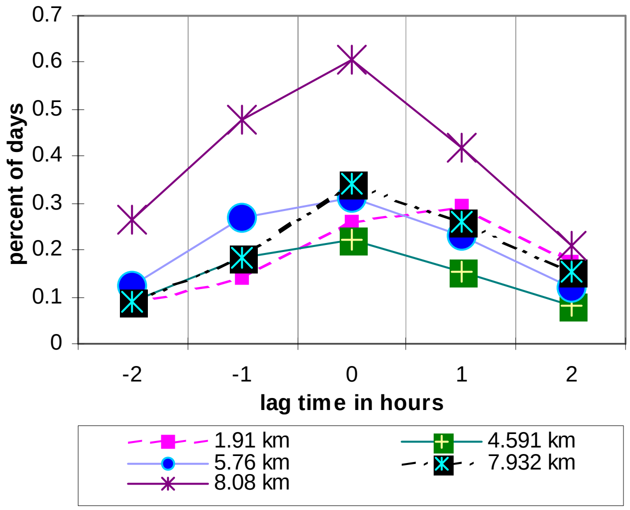


Figure 3-2. **Percentage of days on which the cross-correlation is significant at the 5% level using the Pearson test for station pairs of small separation distances**

3.3 MEASURES OF PREDICTION ACCURACY

To assess the accuracy of the alternative methods, both one-station (for example, univariate) and two-station (for example, bivariate) comparisons were made. For the univariate case the single rain gage was denoted as *Y*, with measured values indicated as Y and predicted values as $\hat{Y}$. Subscripts were used to indicate the time (hour, day, year). For bivariate analyses, the gage where predictions were made is denoted as *Y*, while the satellite gage from which data were transferred to make the predictions is denoted as gage *X*. Again, subscripts were used to define the specific time of a rainfall depth.

The accuracy of predictions was assessed using goodness-of-fit statistics. Both systematic and non-systematic error variations were calculated. Bias is a measure of the systematic error, which reflects consistent under prediction or consistent over prediction. Nonsystematic error variations measure the expected magnitude of errors about the true values. This is sometimes referred to as random variation, although it may actually be due to systematic variation associated with an uncontrolled variable. The goodness of fit-statistics (McCuen, 1993) used to assess accuracy include the bias $\bar{e}$, (Eq. 3-1); the standard error of estimate S_e, (Eq. 3-2); the modified standard error of the estimate S_{em}, (Eq. 3-3); the relative bias R_b, (Eq. 3-4); and the relative standard error R_e, (Eq. 3-5). These statistics are used in comparing the measured and disaggregated hourly precipitations. The statistics were calculated using hourly rainfall depths, both measured and predicted at gage *Y*.

.

$$\bar{e} = \frac{1}{N}\sum_{m=1}^{10}\sum_{j=1}^{365}\sum_{i=1}^{24}\left(\hat{Y}_{ijm} - Y_{ijm}\right) \qquad \textbf{if } Y_{ijm} > 0 \textbf{ or } \hat{Y}_{ijm} > 0 \tag{3-1}$$

in which *N* was the number of hours for the 10-year period with continuous rainfall data at both stations for which $Y_{ijm} > 0$ or $\hat{Y}_{ijm} > 0$; Y_{ijm} was the measured hourly rainfall value at gage *Y* on day *j,* which is assumed to be the true value; $\hat{Y}_{ijm}$ was the predicted hourly rainfall at gage *Y* on day *j,* with the subscript *i* indicating the hour at which the rainfall was measured; $\bar{e}$ was the bias in the hourly estimates.

The standard error of estimate between the predicted and measured hourly rainfall depths is a measure of the accuracy of a method and measures both systematic and random error. The S_e is computed by:

$$S_e = \sqrt{\frac{1}{v}\sum_{m=1}^{10}\sum_{j=1}^{365}\sum_{i=1}^{24}(\hat{Y}_{ijm} - Y_{ijm})^2} \qquad \textbf{if } Y_{ijm} > 0 \textbf{ or } \hat{Y}_{ijm} > 0 \tag{3-2}$$

where the degrees of freedom ($v = N - 1$). For the bivariate case (for example, two-gage methods), the method may introduce a bias, which will influence the computed standard errors. To remove the effect of the systematic error, a modified standard error of estimate S_{em}, was computed:

$$S_{em} = \sqrt{\frac{1}{v}\sum_{m=1}^{10}\sum_{j=1}^{365}\sum_{i=1}^{24}(\hat{Y}_{ijm} - \bar{e} - Y_{ijm})^2} \tag{3-3}$$

Although the standard error of Eq. (3-2) is a measure of accuracy, the modified standard error of Eq. (3-3) is a measure of the precision.

The goodness-of-fit statistics can be standardized to yield dimensionless indices. The relative bias, R_b, is calculated by dividing the bias of Eq. (3-1) by the mean measured hourly depth:

$$R_b = \frac{\bar{e}}{\bar{Y}} = \frac{\bar{e}}{\frac{1}{N}\sum_{m=1}^{10}\sum_{j=1}^{365}\sum_{i=1}^{24}Y_{ijm}} \qquad \textbf{for } Y_{ijm} > 0 \tag{3-4}$$

where N is the number of hours used in the comparison. The relative standard error is:

$$R_e = S_e / S_y \tag{3-5}$$

in which

$$S_y = \left[\frac{1}{N-1}\sum_{m=1}^{10}\sum_{j=1}^{365}\sum_{i=1}^{24}(\bar{Y}_{ijm} - Y_{ijm})^2\right]^{0.5} \qquad \textbf{for } Y_{ijm} > 0 \tag{3-6}$$

When comparing daily rainfall depths, the hourly summation in Eqs. (3-1) to (3-6) is omitted.

3.4 UNIFORM DISAGGREGATION METHOD

Where daily rainfall depths are available at a site and a satellite station that measures hourly rainfall is not nearby, it would be necessary to disaggregate daily values into hourly values. If hourly intensities are needed, then a systematic means of disaggregating the daily total is the necessary option. Measured daily rainfall depths (Y_{jm}) can be disaggregated into hourly estimates ($\hat{Y}_{ijm}$) by the following:

$$\hat{Y}_{ijm} = \frac{1}{24} Y_{jm} \tag{3-7}$$

where the weight of 1/24 divides the daily total depth into 24 equal hourly depths. This is a simple method that could fail because it ignores the fact that all storms are not of a 24-hour duration. However, it can serve as a bench markcase.

To assess the accuracy of disaggregating daily depths into hourly depths, the measured hourly values were aggregated, for example, totaled, to obtain a value that is assumed to be a measured daily depth. The aggregated daily values are denoted as Y_{jm}. Then the daily totals are uniformly disaggregated with Eq. (3-7) into a hourly estimates, $\hat{Y}_{ijm}$. This method eliminates any systematic error because the entire daily rainfall depth is distributed throughout the day. Therefore, the method is unbiased.

The accuracy of the disaggregation model was assessed by comparing the predicted hourly values $\hat{Y}_{ijm}$ and the actual measured hourly values Y_{ijm}. Although this comparison assumes that only the daily values were recorded, as would be the case for a daily recording gage, the hourly rainfalls were actually measured and are therefore available for making the hourly comparisons.

The bias or systematic error in the model predictions must be zero because the daily precipitation was uniformly distributed over 24 hours. Figure 3-3 shows the relative standard error of Eq. (3-5) for the 74 gages. The values range from about 91% to 98%, which indicates that a uniform separation of a measured daily value has nearly as much random scatter as when the mean hourly value is used as the estimate. These results indicate that the uniform disaggregation method is not an accurate predictor of hourly rainfall, which was the result expected because most of the smaller storms have durations much less than 24 hours. Additionally, this method would under predict the number of hours of zero rainfall as volumes for short duration storms would be spread over 24 hours.

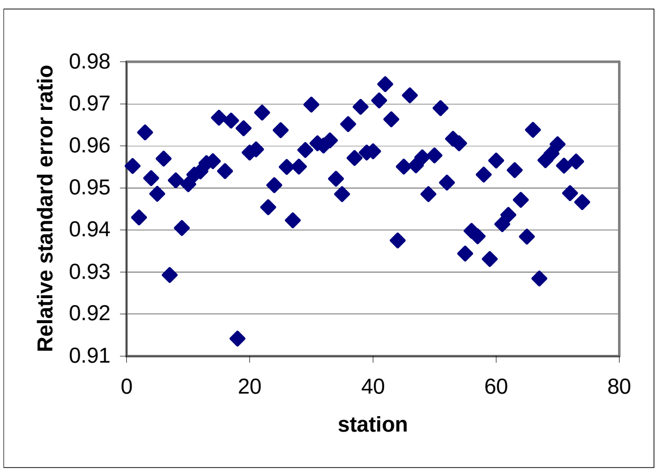


Figure 3-3. Relative standard error ratio for the uniform disaggregation of daily rainfall depths into hourly values.

3.5 UNIVARIATE WEATHER PATTERN DISAGGREGATION

Where meteorological studies have previously been performed to produce hourly meteorological distribution patterns, they can be used to disaggregate daily depths measured at the base gage Y. The USGS National Research Program in Denver (Hay et al., 2002) has developed such patterns for the Chesapeake Bay watershed, using the latitude (x), longitude (y), and elevation (z) of the climate stations as the independent variables in the method. Although the method focuses on the spatial distribution of point data to better represent watershed climate variability, the method also provides a weather pattern distribution for the disaggregation of daily depths into hourly rainfall values. The weather patterns were developed using large meteorological databases, and multiple linear regression (MLR) for each dependent climate variable, which included temperature and pressure. Seven weather patterns were developed for the Chesapeake Bay, with the allocation of each weather pattern varying from month to month. All of the patterns assume that the daily rainfall is distributed over 24 hours, which may be realistic for low-frequency rainfalls (for example, 2-yr and more extreme) but is probably unrealistic for the smaller storms. Acreman (1990) showed that, when considering all storm magnitudes, the durations were approximately exponentially distributed, with few events having durations greater than 15 hours. Figure 3-4 and Figure 3-5 show 2 of the 7 weather-derived allocation patterns for the month of May in the Chesapeake Bay watershed. The patterns vary from a nearly uniform pattern, such as, to a pattern characterized by considerable hour-to-hour variation (Figure 3-5). To apply the weather patterns for any one day, measured meteorological data such as pressure are used to decide which one of the seven weather patterns is the most appropriate for that day. An obvious disadvantage of this method is that, when a model is used in the forecast mode, the meteorological data would not be available. However, by comparing the results from this method to those of the uniform disaggregation method, the value of hour-to-hour variation can be assessed.

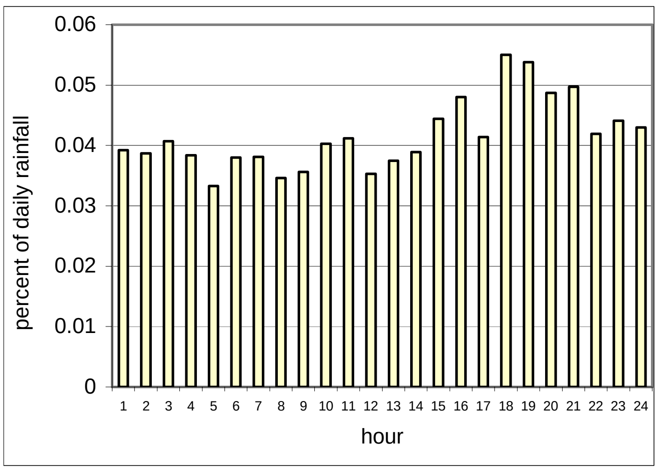


Figure 3-4. **Precipitation pattern distribution Type 1 for the month of May in the Chesapeake Bay watershed**

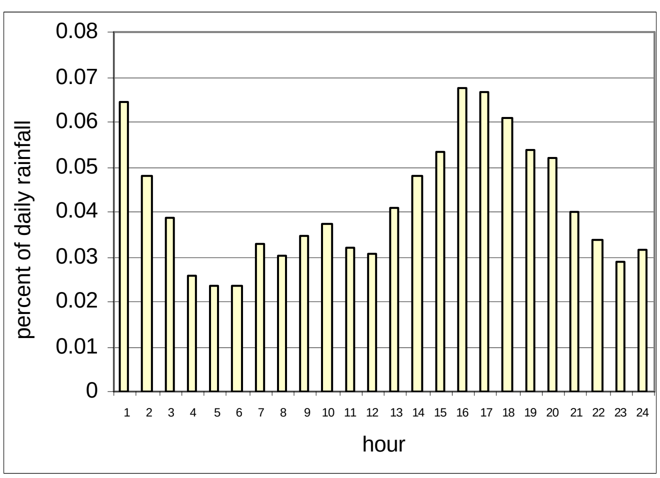


Figure 3-5. **Precipitation pattern distribution Type 2 for the month of May in the Chesapeake Bay watershed**

The climate patterns were used as the second method for disaggregation of daily rainfall depths. A predicted hourly depth is some fraction of the measured daily depth:

$$\hat{Y}_{ijm} = F_i Y_{jm} \quad \textbf{(3-8)}$$

in which F_i is the proportion of the daily total depth allocated to hour i. As with the uniform disaggregation method, the systematic error is zero because the entire daily rainfall is distributed throughout the day.

The accuracy of the disaggregation model was also assessed by comparing the predicted values $\hat{Y}_{ijm}$ with the actual measured values Y_{ijm}. The relative standard error of estimate was computed for each of the 74 stations. Since the method has a zero bias, the standard error measures both the precision and accuracy. The relative standard errors for the 74 stations ranged from 90% to 98% (Figure 3-6), which indicates poor accuracy. The values are not much better than the mean. The weather pattern derived estimates do not accurately compare to the measured values because many storms are less than 24 hours in duration, and therefore, distributing the daily total over a full 24 hours dampens the variation inherent to the actual hourly rainfall depths. These results are almost identical to the results from the uniform disaggregation method, which indicates that 24-hour distribution patterns are not reliable. Large storms are often of longer duration (Levy and McCuen, 1999), and thus the weather pattern disaggregation method might provide greater accuracy when used solely for longer duration rainfall data.

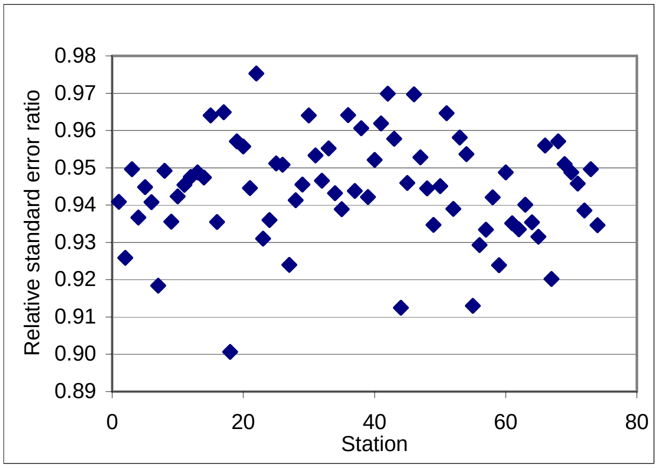


Figure 3-6. Relative standard error ratio for the disaggregation of daily rainfall depths into hourly values using weather patterns.

3.6 BIVARIATE SATELLITE TRANSFER METHOD

As disaggregation of measured daily depths does not seem to provide acceptable accuracy, then alternatives that transfer information from satellite gages need to be assessed. The simplest model would be to assume that the rainfalls at the two gages were identical. Thus, the predicted hourly depth at station Y equals the measured hourly value at gage X:

$$\hat{Y}_{ijm} = X_{ijm} \tag{3-9}$$

where X_{ijm} was the measured hourly rainfall depth at the satellite gage X for hour i, day j, and year m; and $\hat{Y}_{ijm}$ was the predicted hourly rainfall value for gage Y. The sample size to calculate the statistics for the hourly estimates was equal to the number of hours where rainfall was measured at either X or Y.

To assess the accuracy of this method, the predicted and measured hourly depths were compared for the days on which precipitation occurred at X or Y. The results indicated that the model of Eq. (3-9) introduced a systematic error into the predicted values, as the calculated relative biases were nonzero (Figure 3-7).

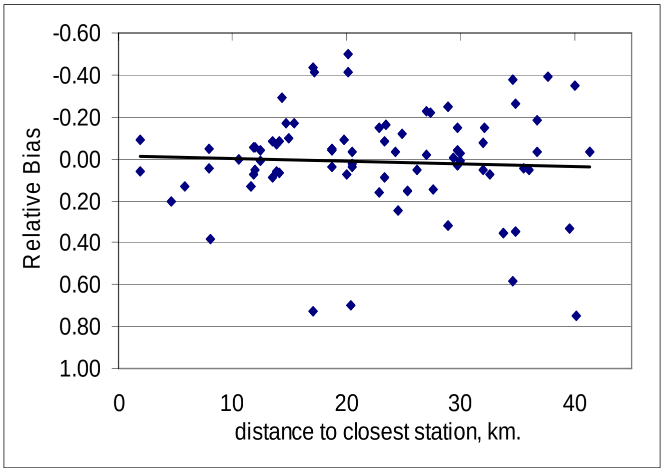


Figure 3-7. Relative biases for the estimation hourly values using the satellite transfer method as a function of the distance between the two closest stations.

The under-prediction is rational since measured precipitation would be lost for the days in which the precipitation was recorded at the gage Y but rainfall did not occurred at the satellite gage X. Conversely the over-prediction will occur on instances in which precipitation was recorded at the gage X but rainfall did not occurred at the gage Y.

Of the 87 pairs of stations, 36 had a relative bias greater than 10%, which is a significant loss or gain for any rainfall estimation. However, the relative bias for the remaining 51 pairs was less than 10% despite the distance of up to 41 km between stations. These results indicate that additional factors other than distance, influence the under or over prediction of rainfall depths.

The relative standard error ratios for hourly values estimated with the satellite transfer method are shown in Figure 3-8. The magnitudes indicate inaccurate predictions, as the ratios are greater than 1. In fact, the method provides less accuracy than the uniform and weather pattern disaggregation methods, even for gages that are in close proximity to each other. The analysis of Figure 3-8 show an increasing trend, which indicates that prediction accuracy decreases as the distance between gages increases.

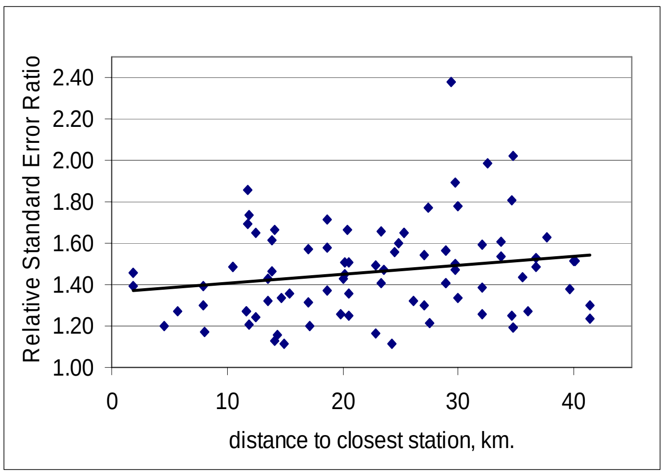


Figure 3-8. Relative standard error ratio for the estimation of hourly values using the satellite transfer method as a function of the distance between the two closes stations.

3.7 SATELLITE TRANSFER OF DAILY RAINFALL DEPTHS

Since some localities have daily measured rainfall available, but not hourly, it was of interest to assess the accuracy of the transposition of daily values. The transposition of storm totals would be a third option. McKay (1970) discusses the transposition of storm totals. The accuracy of estimating daily rainfall depths from a satellite station when measured rainfall depths are not available at a location within the watershed being modeled was also assessed using Eq. (3-9), but omitting the hourly subscript *i*. The sample size used in calculating the statistics was equal to the number of days on which rainfall occurred at either the base and satellite gages.

The relative bias was computed using the daily values for each of the 87 station pairs. As expected, the biases were the same as those computed using the hourly values because the data for the analysis were the same (Figure 3-9). The relative standard error however, was significantly better than that obtained in the satellite transfer method of hourly rainfall depths (Figure 3-10). Forty percent of the analyzed pairs had a relative standard error of less than 0.4, which suggests a reasonable level of accuracy in the transposed values; Daily rainfall totals can be transferred with considerably better accuracy than the transfer of hourly rainfall depths. The result also reveals a trend in the relation between the standard error ratio and the distance separating the gages, with the accuracy decreasing with increasing separation distance. Similarly, Bradley et al. (2002) showed a decreasing trend in storm accumulation accuracy with increases in separation distance.

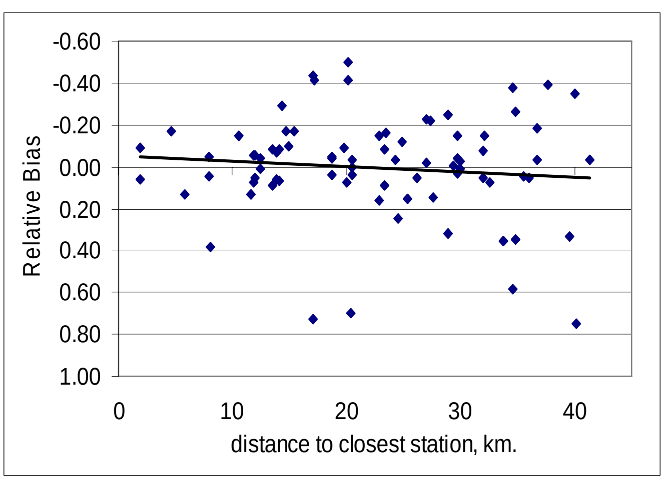


Figure 3-9. Relative biases for the estimation daily values using the satellite transfer method as a function of the distance between the two closest stations.

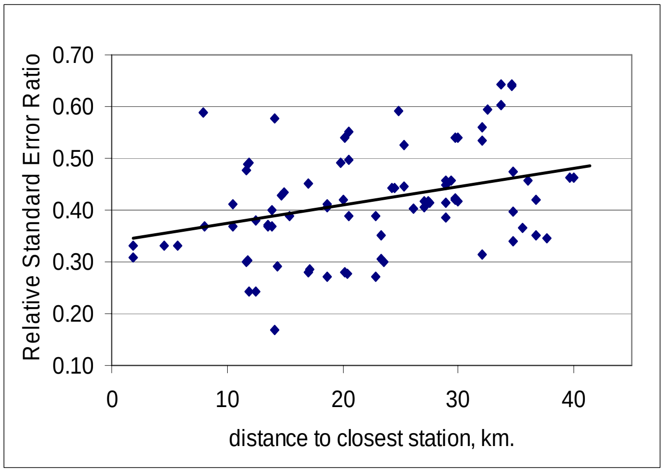


Figure 3-10. **Relative standard error ratio for the estimation of daily values using the satellite transfer method as a function of the distance between the two closes stations.**

3.8 BIVARIATE SATELLITE RATIO DISAGGREGATION

As an alternative to the univariate analyses of Eqs. (3-7) and (3-8), two-site analyses can be made. This is useful where daily rainfall depths have been recorded at a gage within the watershed, but hourly proportions from a satellite station outside the watershed can be used to distribute the measured daily depths into hourly values. Measured hourly rainfall depths (X_{ijm}) at satellite gage X are used to proportion daily rainfall depths measured within the watershed (Y_{jm}) using the proportion of the daily rainfall (X_{jm}) at X:

$$Y_{ijm} = \left(\frac{X_{ijm}}{X_{jm}} \right) Y_{jm} \tag{3-10}$$

To assess the accuracy when using Eq. (3-10) the value of X_{jm} is aggregated from the hourly values of X_{ijm}. In this analysis, the values of Y_{jm} are obtained by summations of the measured values of Y_{ijm} for any j and m:

$$\hat{Y}_{jm} = \sum_{i=1}^{24} Y_{ijm} \tag{3-11}$$

Then the goodness-of-fit statistics are obtained using Eqs. (3-1) to (3-6).

The ratio in the parentheses of Eq. (3-10) is the proportion of the daily total rainfall for gage X that occurred in hour i. These proportions were then multiplied by the aggregated daily value Y_{jm} to predict hourly depths at station Y. The accuracy of the disaggregation model can then be assessed by comparing the predicted values $\hat{Y}_{ijm}$ and the actual measured values Y_{ijm}.

The relative bias was computed for each station pair in order to assess the significance of the systematic variation. A negative bias resulted for all of the station pairs, which indicates that the predictions made with Eq. (3-10) systematically under estimate the actual amount of precipitation. The negative bias occurs for the following reason. If it did not rain at gage X on a given day when rainfall was recorded at gage Y, then the recorded daily rainfall cannot be disaggregated for that day, and Eq. (3-10) will predict zero rainfall for all of the hours of that day at Y. Conversely, rainfall that occurred at X but not at Y does not cancel this error since Eq. (3-10) correctly gives zero rainfall at Y. Hence, a systematic, negative bias is introduced. The relative biases of the station pairs indicate underprediction from 5% to 62% with a mean value of about 25%. A relative bias of - 62% would correspond to an under-estimation of rainfall of 15.2 in. per year, assuming that the mean annual precipitation was about 40 inches. Thus, the biases of the bivariate satellite ratio disaggregation method are hydrologically significant.

The relative biases of Figure 3-11 do not reveal a relation with separation-distance. The bias for station pairs in close proximity to each other is similar to that for gages separated by 30 to 40 kilometers or more. Thus, an underprediction of 25% can be expected regardless of separation distance.

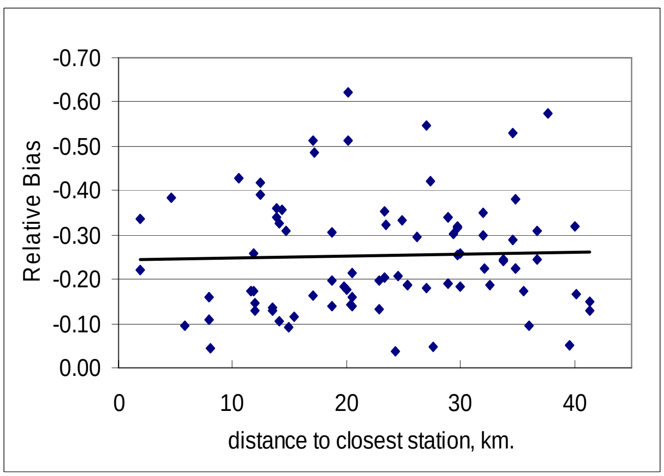


Figure 3-11. Relative bias for the distribution of daily rainfall depths using the bivariate satellite disaggregation approach

The relative standard errors were computed for each station pair and are plotted in Figure 3-12 as a function of separation distance. The hourly rainfall at station Y was poorly associated with the hourly rainfall depths at the satellite gage X as indicated by the high values of (R_e). Since all of the values except one exceed a ratio of 1, the method fails to improve prediction accuracy above that of the two-univariate disaggregation methods. That is, the additional information content of the satellite gage is not of significant value for the disaggregation of daily rainfall.

The relation between the standard error ratio and separation distance showed a slight increasing trend as a function of the distance between stations. However, given that the values are greater than 1, the trend is not important, as prediction accuracy is poor.

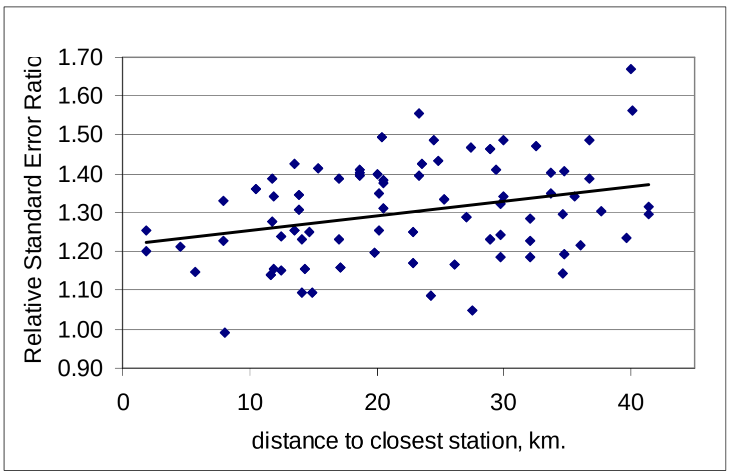


Figure 3-12. Relative standard error ratio for the distribution of daily rainfall depths using the bivariate satellite disaggregation approach

Since the satellite ratio disaggregation method is biased, the effect of the bias was removed from the standard error using the modified standard error of Eq. (3.1-3). This yields the precision of the predictions. A comparison between S_e and the modified standard error S_{em} for stations of up to 8 kilometers apart (Table 3-2) indicates that the systematic error introduced by the model is not a significant source of inaccuracy in the predictions of hourly rainfall. A similar observation goes for the comparison of the relative standard error ratio (R_e), and the modified relative standard error ratio (S_{em} / S_y). The contribution of the bias to the overall error is about 1%. This result is rational because the hourly bias is very small in comparison with the nonsystematic variation within the hourly values. In summary, the total bias on an annual basis is very significant, as underprediction is likely with this method; however, the nonsystematic error variation is independently quite significant.

Table 3-2. Comparison of standard error S_e and the modified standard error S_{em} for stations of up to 8 kilometers apart

Separation distance (Kilometers)	Station ID	Station ID	S_e	S_{em}	$R_e = S_e / S_y$	S_{em} / S_y
	X	Y				
1.91	364778	364763	0.1827	0.1818	1.2650	1.2587
1.91	364763	364778	0.1674	0.1657	1.2196	1.2072
4.59	368491	367029	0.2652	0.2645	1.0953	1.0925
4.59	367029	368491	0.1687	0.1641	1.2620	1.2276
5.76	368758	368763	0.1530	0.1526	1.1507	1.1477
7.93	445880	445595	0.1366	0.1362	1.3388	1.3349
7.93	445595	445880	0.1356	0.1347	1.2420	1.2338
8.08	180465	180470	0.0960	0.0960	0.9922	0.9919

In summary, the four methods of rainfall disaggregation tested in this study provided poor accuracy in the prediction of hourly rainfall data. A comparison of the univariate disaggregation methods (Table 3-3) indicated that the prediction accuracy using the weather patterns method was not better than using the uniform distribution, as the relative standard error ratio values were similar. Similar accuracy was also obtained between the two bivariate methods (Figure 3-13) with relative standard error ratio values greater than 1.0. In spite of the extensive database used in the analyses, the results were discouraging for those who require accurate estimates of hourly rainfall intensities. However, the results do provide a valuable indication of the potential bias and inaccuracy of rainfall depths disaggregated from daily values or transferred from a nearby hourly rain gage, and the potential implications of this inaccuracy on watershed model calibrations.

Table 3-3. Accuracy of the univariate rainfall disaggregation methods. S_e/S_y = relative standard error ratio

Method of Disaggregation			
		uniform	**weather patterns**
minimum	$\frac{S_e}{S_y}$	0.91	0.90
mean		0.95	0.94
max		0.98	0.98
effort		minimum	significant

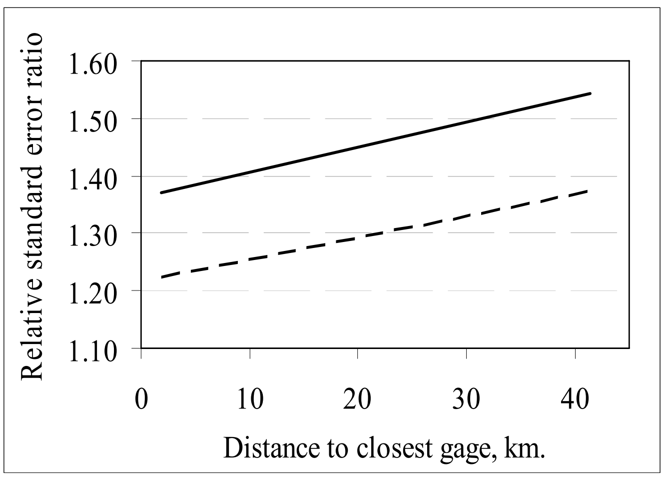


Figure 3-13. Accuracy of the bivariate rainfall disaggregation methods. Solid line = satellite transfer and dotted line = satellite ratio.

CHAPTER 4

EFFECT OF DAILY RAINFALL DISAGGREGATION ON PREDICTED DISCHARGES

4.1 INTRODUCTION

The required accuracy of disaggregated daily rainfall may depend on the sensitivity of the model to rainfall data. The results from the analyses of rainfall disaggregation in CHAPTER 3 indicated that all of the methods smooth the daily precipitation such that the natural intensity of the rainfall was lost. These results provide a valuable indication of the potential implications on the accuracy of predicted runoffs due to inaccuracy introduced by the disaggregation of daily rainfall depths. Thus, it was of interest to determine if the disaggregation of daily rainfall would have an effect on the accuracy of both the daily and the hourly HSPF predicted discharges, exhibited as a negative or positive relative bias or in a high value of the relative standard error ratio.

4.1.1 Data

The analyses were conducted using data from 8 hypothetical and forested watersheds each with a drainage area of 5 mi2 and with the flow distribution found in Table 4-3. The assumed actual watershed outflow was the sum of the HSPF predicted flow components SURO, IFWO, and AGWO and it will be referred to as the measured runoff. Actual hourly rainfall data between 1992 and 1999 were used to aggregate daily totals. The actual hourly rainfall data were used to produce the HSPF measured runoff.

4.1.2 Method of Analyses

Three methods of daily rainfall disaggregation were selected for the analysis: (1) a uniform distribution over 24 hours; (2) the Soil Conservation Service (SCS, 1986) 24-hour storm distribution; and (3) a depth-duration-dependent separation. For the SCS method, a type II storm distribution over a 24-hour period was selected. Both the SCS and the depth-duration methods were centered at 12:00 noon.

The disaggregation procedure began by aggregating actual hourly rainfall depths (Y_{ij}) into daily values (Y_j). Then the daily totals were disaggregated into hourly values ($\hat{Y}_{ij}$) using each of the three-disaggregation methods. The following expression was used in the Uniform disaggregation method:

$$\hat{Y}_{ij} = \frac{1}{24} Y_j \qquad \textbf{(4-1)}$$

where the weight of 1/24 divides the daily total depth into 24 equal hourly depths, i is the hour, and j is the day.

The SCS method analyzed rainfall-frequency data from the Weather Bureau's Rainfall Frequency Atlases (NWS, 1961) for areas less than 400 mi2, for durations up to 24 hr, and for frequencies from 1 yr to 100 yr to derive four dimensionless rainfall distributions. In the type II storm distribution, the peak intensity of the storm was assumed to occur at the center of the storm defined by increments of 6-min depth, and at about the middle of the 24-hr period. The storm is arranged as a continuous sequence of 6-min incremental depths representing the rainfall depth for that duration and frequency. For example, the maximum 6-min depth is subtracted from the maximum 12-min depth; the 12-min depth is subtracted from the maximum 18-min depth and so on to 24 hours. For this analysis, the SCS cumulative values at the hour, are shown in Table 4-1, columns 2 and 5. The SCS cumulative values were then used to derive the multiplication factors (M_i) to distribute the daily rainfall as follows: the SCS cumulative value for the first hour was assigned as the multiplication factor at hour 1; the multiplication factor at hour 2 was obtained by subtracting the hour 1 cumulative SCS value from the SCS value at hour 2; the multiplication factor at hour 3 was obtained by subtracting the hour 2 cumulative SCS value from the SCS value at hour 3 and so on. The multiplication factors are found in Table 4-1, columns 3 and 6 and are used to allocate the daily rainfall depth using the following expression:

$$\hat{Y}_{ij} = Y_j * M_i \qquad \textbf{(4-2)}$$

were M_i is the multiplication factor (See Table 4-1) for hour i.

Table 4-1. SCS cumulative, dimensionless one-day type II storm and multiplication factors to allocate daily rainfall depths to hourly values.

Hour	SCS cumulative	multiplication factor M_i	hour	SCS cumulative	multiplication factor M_i
1	0.0108	0.0108	13	0.7724	0.1092
2	0.0223	0.0115	14	0.8197	0.0473
3	0.0347	0.0124	15	0.8538	0.0341
4	0.0483	0.0136	16	0.8801	0.0263
5	0.0632	0.0149	17	0.9019	0.0218
6	0.0797	0.0165	18	0.9206	0.0187
7	0.0984	0.0187	19	0.9371	0.0165
8	0.1203	0.0219	20	0.9519	0.0148
9	0.1467	0.0264	21	0.9653	0.0134
10	0.1808	0.0341	22	0.9777	0.0124
11	0.2351	0.0543	23	0.9892	0.0115
NOON	0.6632	0.4281	24	1.0000	0.0108

The depth-duration method was based on analyses of actual storm frequency data (Kreeb, 2003) using 15 stations in Maryland. Kreeb (2003) analyzed rainfall records and determined the fraction of actual storms frequency according to five depth classes and seven duration classes. In this method the total depth of daily rainfall in the frequency table determines the number of hours in which the daily rainfall is disaggregated (Table 4-2). The number of hours varied with the total storm depth. For example, since 87% of storms with a depth less than 0.1 in. lasted no more than 1 hour, this was used as the duration for all of the storms with a depth less than 0.1 inch. For storms with a depth greater than 0.1 in. and less than 0.25 in., 24% had duration of 4 to 6 hours. Surrounding cells had significant fraction so 5-hr duration was used. Similar analyses were used for the other daily depths. The selected numbers of hours within each range to allocate the daily precipitation are: 1, 5, 9, 15 and 20 (Table 4-2). Each daily storm total was then disaggregated into equal parts and centered at 12 noon through the following expression:

$$\hat{Y}_{ij} = Y_j / N_v \quad \textbf{(4-3)}$$

were N_v is the number of hours as a function of the daily volume v.

Table 4-2. **Frequency of storms using 15 stations in Maryland and by depth of precipitation (in.). The percentage storms in each depth class is noted in parenthesis.**

Duration (hr)	Depth of precipitation for the day (in.)					
	0.01 - 0.10	**0.10 - 0.25**	**0.25 - 0.50**	**0.5 - 1.00**	**>1.00**	**sum**
1	0.2857 (87)	0.0214 (15)	0.0167 (8)	0.0043 (2)	0.0008 (1)	0.3289
2	0.0164 (5)	0.0257 (18)	0.0221 (10)	0.0089 (5)	0.0025 (2)	0.0756
3	0.0085 (3)	0.0223 (15)	0.0198 (9)	0.0083 (5)	0.0038 (3)	0.0627
4 – 6	0.0099 (3)	0.0351 (24)	0.0475 (22)	0.0221 (13)	0.0087 (6)	0.1234
7 – 12	0.0058 (2)	0.0337 (23)	0.0629 (30)	0.0528 (30)	0.0266 (19)	0.1818
13 – 24	0.0024 (0)	0.0070 (5)	0.0397 (19)	0.0611 (35)	0.0515 (37)	0.1617
24 <	0.0000 (0)	0.0009 (0)	0.0043 (2)	0.0172 (10)	0.0435 (32)	0.0659
Sum	0.3287 (100)	0.1461 (100)	0.2131 (100)	0.1747 (100)	0.1374 (100)	1.0000

Source: Kreeb (2003)

Table 4-3. **Flow distribution of the eight hypothetical watersheds**

Watershed	Percent of Baseflow	Percent of Interflow	Percent of Surface Runoff
1	81%	7%	12%
2	24%	66%	10%
3	58%	35%	7%
4	49%	22%	29%
5	35%	8%	57%
6	59%	27%	14%
7	70%	10%	20%
8	35%	32%	33%

4.1.3 Measures of Accuracy

The accuracy of the daily and hourly outflow predictions was measured through the goodness-of-fit statistics calculated for the 8 years of the period of record. Provided that instantaneous water quality sample are commonly matched to the measured discharge for the same hour, the accuracy of the hourly predictions were also evaluated on the seasonal watersheds. It was of interest to determine if the accuracy of the hourly predictions varied from season to season. The seasonal dataset was composed of the same three months for eight years, with a total of 24 months per dataset. Season #1 data were obtained from the months of January, February, and March; the months of April, May, and June were considered season # 2; July, August, and September are season # 3; and October, November and December are season # 4. The bias is:

$$\bar{e} = \frac{1}{N}\sum_{j=1}^{N}\left(\hat{Q}_j - Q_j\right) \tag{4-4}$$

were Q_j is the actual daily outflow for the day j, or if the analysis is for the accuracy of the hourly outflows then Q_j is the actual hourly outflow for the hour j. N is the number of days when the analysis is using the daily outflows, or the number of hours when the analysis is using the hourly outflows. The standard error is:

$$S_e = \sqrt{\frac{1}{\nu}\sum_{j=1}^{N}(\hat{Q}_j - Q_j)^2} \tag{4-5}$$

where $\bar{e}$ is the mean bias of the model predicted daily or hourly runoff; ν is the degrees of freedom ($\nu = N - 1$), and S_e is the standard error of estimate between the actual and predicted daily or hourly runoffs. The standardized bias (R_b) is:

$$R_b = \frac{\bar{e}}{\bar{Q}} = \frac{\bar{e}}{\frac{1}{N}\sum_{j=1}^{N} Q_j} \tag{4-6}$$

where $\overline{Q}$ is the mean of the actual daily or hourly runoff for the period of analysis. The relative standard error (R_e) is:

$$R_e = S_e / S_y \tag{4-7}$$

in which and S_y is the standard deviation of the actual daily or hourly runoff:

$$S_y = \left[\frac{1}{N-1} \sum_{j=1}^{m} (Q_j - \overline{Q})^2 \right]^{0.5} \tag{4-8}$$

4.2 EFFECT OF RAINFALL DISAGGREGATION ON THE ACCURACY OF PREDICTED HOURLY RUNOFF

The accuracy of the predicted hourly runoff is a function of various factors, including the accuracy of the disaggregated rainfall. When a model is designed to address water quality pollution, inaccurate predictions of runoff can reduce the accuracy of predicted pollutant concentrations. The accuracy of the predicted pollutant concentrations is a function of the accuracy of the predicted runoff.

Disaggregation methods that included regional or local information were expected to provide better predicted discharges than the 24-hr uniform disaggregation method. The accuracy of the predictions using the 24-hr uniform disaggregation method had the poorest accuracy. The results by season are presented in Table 4-4 through Table 4-7. For this particular method, the accuracy of the predictions was a function of the flow proportions in the watershed and the season of the year. For watersheds with predominant baseflow (watersheds 1, 3, 4, 6, and 7), the relative bias for the winter months varied between -0.895 and -1.184 (Table 4-4), while for the summer months it varied between -5.282 and -8.190 (Table 4-6). The poorest accuracy was observed during the summer months for all of the watersheds with predominant baseflow.

A similar case was observed with the nonsystematic variation S_e / S_y varying between -4.833 and -7.397 during the winter months and between -3.739 and -6.910 during the summer months for watersheds with predominant baseflow. However, when the portion of baseflow was less than 50% of the total watershed outflow (watersheds 2, 5, and 8), the relative standard error ratio varied between -0.740 and -0.930 for the winter months (Table 4-4) and between -0.823 and -0.885 for the summer months (Table 4-4). The poor accuracy of the predicted hourly outflow is explained as the method of disaggregation ignores the temporal variability and the change in intensity of actual storms. During the summer months, sporadic and more intense storms are difficult to predict with any disaggregation method than in the winter months with more predictable and consistent types of precipitation.

As in the case of the 24-hr Uniform method, it was expected that the accuracy of the predicted hourly outflow using the SCS method were poor because of the disaggregation over a 24-hour period. However, the results indicate that since the SCS method was derived using actual data, the accuracy of the predicted hourly discharges was significantly better than the predictions using the Uniform method. The relative bias varied between 0.005 and -0.129 with the least accurate predictions during the summer season (Table 4-6) and with a systematic underprediction during the fall season.

Table 4-4. Goodness-of-fit statistics of hourly discharge for the Winter months (January, February, and March).

Watershed	Relative bias			Relative standard error ratio		
	Uniform	SCS	Depth-Duration	Uniform	SCS	Depth-Duration
1	-1.167	-0.012	-0.118	7.397	1.848	0.944
2	0.024	-0.004	-0.000	0.745	1.868	0.737
3	-1.068	-0.006	0.023	6.998	2.147	0.787
4	-1.048	-0.015	0.001	4.833	1.685	0.924
5	0.047	0.005	-0.000	0.909	1.549	0.935
6	-1.184	0.023	0.016	4.893	1.663	0.840
7	-0.895	0.008	0.041	5.256	1.581	0.909
8	0.042	-0.004	0.007	0.905	1.690	0.909

Table 4-5. Goodness-of-fit statistics of hourly discharge for the Spring months (April, May, and June).

Watershed	Relative bias			Relative standard error ratio		
	Uniform	SCS	Depth-Duration	Uniform	SCS	Depth-Duration
1	-1.894	0.025	-0.191	12.152	1.133	0.886
2	-0.181	-0.058	-0.051	0.740	0.946	0.715
3	-2.237	-0.016	-0.101	12.756	0.953	0.791
4	-2.673	-0.031	-0.087	8.425	1.188	0.929
5	-0.235	-0.052	-0.045	0.966	1.311	0.992
6	-2.642	-0.029	-0.083	13.270	1.124	0.847
7	-2.880	-0.006	-0.111	16.786	1.160	0.930
8	-0.227	-0.041	-0.082	0.939	1.217	0.941

Table 4-6. **Goodness-of-fit statistics of hourly discharge for the Summer months (July, August, and September).**

Watershed	Relative bias			Relative standard error ratio		
	Uniform	SCS	Depth-Duration	Uniform	SCS	Depth-Duration
1	-6.119	-0.095	-0.680	6.618	0.662	0.993
2	-0.029	0.060	-0.011	0.823	1.333	0.812
3	-6.907	-0.080	-0.367	6.910	0.774	0.902
4	-5.282	0.074	-0.021	3.739	1.181	0.866
5	0.005	0.118	0.083	0.874	1.180	0.865
6	-8.190	-0.122	-0.083	6.743	0.706	0.847
7	-7.426	-0.129	-0.433	6.386	0.698	0.963
8	-0.228	-0.012	-0.122	0.885	1.112	0.845

Table 4-7. **Goodness-of-fit statistics of hourly discharge for the Fall months (October, November, and December).**

Watershed	Relative bias			Relative standard error ratio		
	Uniform	SCS	Depth-Duration	Uniform	SCS	Depth-Duration
1	-2.646	0.007	-0.118	7.619	1.207	0.944
2	0.047	0.021	-0.006	0.766	1.247	0.748
3	-2.444	0.034	0.012	8.048	1.462	0.813
4	-2.121	0.010	-0.008	4.408	1.343	0.872
5	0.082	0.021	0.002	0.889	1.399	0.922
6	-2.678	0.046	0.020	6.680	1.237	0.846
7	-2.466	0.015	0.007	5.803	1.272	0.861
8	0.046	0.020	0.013	0.904	1.369	0.901

The predicted hourly outflow using the disaggregated rainfall from the depth-duration method was expected to have the best accuracy of the methods tested; because the precipitation data used in the development of the depth-duration distribution was from the same region from where the disaggregated daily precipitation time series was recorded (Table 4-4 through Table 4-7). In addition, the precipitation data used in the development of the depth-duration method included a common period with the data to be disaggregated; yet, the disaggregated data were not included in the development of the distribution method.

For all seasons except for summer, the relative bias varied between 0.001 and -0.200; the variation during the summer season was between 0.800 and -0.680. The type of precipitation in the area may explain the disparity of values between the relative biases in the summer months and the relative biases for the remaining seasons. Thunderstorms are the most common type of precipitation during the summer months, with large amount of rainfall during short periods. These characteristics are not considered in the depth-duration method because as the volume of the storm increases, the number of hours in which the daily total is distributed also increases. The method uses annual average storm volumes and disregards the intensity of the precipitation during the different seasons. Thus, the lack of seasonality in the development of the method explains the poor prediction accuracy. For all of the methods, the summer months provided the poorest accuracy of predicted hourly outflow.

The values of the nonsystematic variation were consistent throughout the year. The lack of seasonality in the development of the disaggregation methods was observed in the poor accuracy of the watershed outflow predictions; in particular during the summer season. This suggests that if the seasonality is included during the development of a rainfall disaggregation method, the accuracy of the predicted outflows could improve. A comparison of the relative standard error ratio among the methods indicates that spatial variation was a factor in the accuracy of the disaggregated daily rainfall and thus, in the accuracy of the hourly predicted runoff. Although the accuracy of the depth-duration method was poor as indicated by the relative standard error ratio varying between 0.700 and 0.990, the method provided the best results as the distribution reflects more local precipitation patterns.

Table 4-8. Statistical summary of HSPF predictions of hourly runoff for the calibration period.

	Relative bias			Relative standard error ratio		
watershed	**uniform**	**SCS**	**depth-duration**	**Uniform**	**SCS**	**depth-duration**
1	-0.9539	-0.0055	-0.0802	2.0728	1.0319	3.0478
2	-0.0010	0.0008	-0.0091	1.1955	0.9448	1.0797
3	-0.9533	-0.0040	-0.0324	2.7284	0.9811	1.3836
4	-0.9519	-0.0038	-0.0186	1.8617	1.0352	1.5606
5	0.0064	0.0141	0.0036	1.8464	1.0900	1.5728
6	-0.9588	0.0074	-0.0266	10.7267	1.0171	1.7399
7	-0.9559	-0.0032	-0.0342	1.5196	1.0768	1.9755
8	-0.0162	-0.0043	-0.0137	1.7554	1.0466	1.5130

The seasonal analysis provided information about potential problems in the disaggregation methods due to the specific characteristics of the data on the seasonal basis. The results indicated that none of the methods provided accurate predictions as the disaggregation of daily rainfall into hourly values introduces additional noise and uncertainty to the predictions of hourly runoff. The volumes and the timing of the storms during the summer time are perhaps the most difficult aspects to replicate in the disaggregation of daily rainfall to hourly values, which suggest that an additional analysis including the seasonality of the rainfall should be performed. Provided that the attenuation effect of the channel on the predicted runoff was not studied, it is also suggested to investigate the effect of the channel in the accuracy of the predicted discharge.

4.3 EFFECT OF THE DISAGGREGATION OF DAILY RAINFALL ON THE PREDICTION ACCURACY OF DAILY RUNOFF

The values of the predicted daily runoff when using the HSPF model are usually obtained through the average of predicted hourly runoff over the 24-hr period. These calculations smooth the hourly fluctuations of the predicted runoff and inaccuracies in the daily disaggregated rainfall may not have a significant effect on the accuracy of the daily predicted runoff. To examine this hypothesis, analyses of the relative bias for the total runoff and the flow components were performed.

The relative bias of the predicted daily outflows (Table 4-9) was the same as the obtained in the hourly outflow predictions. This is simply because the hourly time-step of the calibrations. Provided that hourly rainfall is supplied as input, all the hydrologic processes in the model are computed on the hourly basis; the mean daily outflow is computed as the sum of the predicted hourly values.

The results of the predicted daily discharges shown in Table 4-9 indicate that the predictions of outflow when using a 24-hr uniform patter disaggregation provided the poorest accuracy. As in the results of the hourly predictions, the effect of the error in the precipitation was significantly evident in the underprediction of outflow for watersheds with a predominant baseflow component (watersheds 1, 3, 4, 6, and 7). This effect is due to the smoothing of the daily precipitation and the loss of the natural intensity of the rainfall. The relative bias for the SCS was similar to the relative bias obtained in the depth-duration methods and always below a 10% in magnitude.

The nonsystematic variation of the predicted daily outflow was larger when rainfall disaggregated with the 24-hr uniform method was used for the analysis. Again, the error varied as a function of the flow proportions in the watershed. In contrast, the relative standard error ratios for the predictions using the SCS or the depth-duration disaggregated precipitation were significantly lower than those obtained in the predictions when using the precipitation from the 24-hr uniform method. The accuracy of the prediction in the SCS and the depth-duration methods was moderate.

Table 4-9. Goodness-of-fit statistics of daily runoff for the 8 years of calibration.

	relative bias			relative standard error ratio		
watershed	uniform	SCS	depth-duration	uniform	SCS	depth-duration
1	-0.9539	-0.0055	-0.0802	3.1291	0.4412	0.4775
2	-0.0010	0.0008	-0.0091	0.4386	0.3873	0.3346
3	-0.9533	-0.0040	-0.0324	3.2932	0.4650	0.5504
4	-0.9519	-0.0038	-0.0186	2.9368	0.3476	0.3561
5	0.0064	0.0141	0.0036	0.5060	0.2844	0.3076
6	-0.9588	0.0074	-0.0266	10.6491	0.3447	0.5070
7	-0.9559	-0.0032	-0.0342	1.9373	0.3499	0.6584
8	-0.0162	-0.0043	-0.0137	0.5210	0.3454	0.3243

The results of the analyses suggest that the method of daily rainfall disaggregation is important in the accuracy of the HSPF predicted runoff. Furthermore, the results suggest that better accuracy of the predicted daily runoff may be attained when the method of disaggregation is based on analyses of actual storm frequency data and when seasonality is taken into consideration.

CHAPTER 5

SENSITIVITY ANALYSIS OF THE PREDICTED RUNOFF TO CHANGES IN THE PARAMETER VALUES

5.1 INTRODUCTION

A sensitivity analysis of HSPF is important because it provides information about the effect of change in the parameter values to the predicted discharges and because such knowledge can increase the efficiency and reliability of the calibration. Fitting the parameters of the HSPF model may be simple or complex depending on the user's selection of the type and number of parameters to be optimized. Although progress in computational techniques has reduced the time required for model calibration, these advances have not reduced the problem of parameter intercorrelation.

The analyses were focused on explaining the effects of parameter changes on the watershed outflow depth (inches). It was expected that the importance of a parameter was a function of the watershed characteristics and that the sensitivity would change in response to the meteorological conditions under which the analysis was performed. In addition, because the importance of the parameters varies with the separation of the flow components, the analysis was made for watersheds with relatively high and relative low baseflows.

5.1.1 Data and Method of Analyses

Actual hourly rainfall data and hypothetical watersheds were used to generate daily watershed outflow data so that the true parameter values that control the processes in the PERLAND module were known. The generated data were assumed to be from a forested watershed with a drainage area of 5 mi2. It is important to clarify that the analyses in this study are all related to the watershed water budget parameters and not to the channel transport parameters of the HSPF model controlled by the F-tables in the input files.

The formulation of the processes that controls the water budget in the HSPF model includes parameters that can be set on an annual or monthly basis. Because of the uncertainty and the lack of monthly data for storages and rates of infiltration, evapotranspiration, and flow recession, and to simplify the assessment, it was decided to perform the tests using parameter values that do not vary monthly. The parameters included in the analyses are shown in Table 5-1.

Table 5-1. Parameters that control the water budget in pervious areas

Parameter	Units	Parameter description
LZSN	Inches	Lower zone nominal storage
INFILT	inches/hr	Index to the infiltration capacity of the soil
AGWRC	day^{-1}	Basic groundwater recession rate if KVARY is zero, and groundwater does not receives inflow.
NSUR	Complex	Manning's n for the assumed overland flow plane
INTFW	None	Interflow inflow parameter
IRC	day^{-1}	Interflow recession parameter
LZETP	None	Lower zone E-T parameter. It is an index to the density of deep- rooted vegetation
DEEPFR	None	Fraction of groundwater inflow that will enter deep (inactive) groundwater, and, thus, be lost from the system
BASETP	None	Fraction of remaining potential E-T that can be satisfied from baseflow (groundwater outflow), if enough is available.
AGWETP	None	Fraction of remaining potential E-T that can be satisfied from active groundwater storage if enough is available.
UZSN	Inches	Upper zone nominal storage

The distributions of the watershed outflow (baseflow, interflow, and surface runoff) for the two hypothetical watersheds are 81%, 7%, and 12% for watershed 1, and 33%, 32%, and 35% for watershed 2. The sum of the baseflow, interflow, and surface runoff volumes is referred to as the watershed outflow, while the sum of the interflow and surface runoff is referred to as the quickflow.

Hypothetical watersheds were selected for these analyses so that the true values of the parameters would be known. In addition, the hourly rainfall used for the analyses were measured data, so that the sensitivity values would reflect actual storm sequences. Finally, the watershed outflow was generated without error variation so that that the bias and standard error would be known exactly, for example, they would be zero. Given these conditions, the relative bias, the relative standard error, and the standard error of the predictions were chosen as the criteria to reflect the sensitivity of the parameters.

The sensitivity analyses followed the following general procedure:

1. Parameter values within the recommended range indicated in the HSPF manual were assumed for the modeling of the hypothetical watersheds as the starting point for the analysis.
2. HSPF daily watershed outflows were computed for the surface, interflow, and groundwater layers of the hypothetical watershed; the sum of these components is referred to as the watershed outflow and is denoted as Q_j for day *j*. The sensitivity analyses were made by changing only the value of one parameter at a time, while the other parameters remained constant at their base values. Parameter AGWRC was only increased and decreased by 2% of its original value. Changes of 10% were made to each of the other parameters. The limit on AGWRC is necessary because the maximum value of AGWRC that HSPF allows is 0.999. Since the

initial value for the hypothetical watersheds was 0.97 (representing a receding rate of about 33 days), a value larger than 2% would result in a value of AGWRC greater than 1.00, which would be irrational. Overall, the initial values were selected based on the physical meaning of the parameter (see Table 5-1).

3. The importance of parameters will be measured using the bias and the standard error ratio. The bias reflects the systematic error while the standard error is a measure of the random error. If either of these values deviate from zero, the parameter is considered important, with greater importance associated with greater deviation. Goodness-of-fit statistics were calculated on an annual basis between the true (Q_j) and predicted ($\hat{Q}_j$) daily watershed outflows to determine the parameter sensitivity under the various climatological conditions. The bias is:

$$\bar{e} = \frac{1}{N}\sum_{j=1}^{365}\left(\hat{Q}_j - Q_j\right) \tag{5-1}$$

and the standard error is:

$$S_e = \sqrt{\frac{1}{\nu}\sum_{j=1}^{365}(\hat{Q}_j - Q_j)^2} \tag{5-2}$$

where $\bar{e}$ is the mean bias of the model predicted daily watershed outflow; ν is the degrees of freedom ($\nu = N - 1$); N is the number of days in the year; and S_e is the standard error of estimate between the true and predicted daily watershed outflows. The standardized bias (R_b) is:

$$R_b = \frac{\bar{e}}{\bar{Q}} = \frac{\bar{e}}{\frac{1}{N}\sum_{j=1}^{365}Q_j} \tag{5-3}$$

where $\overline{Q}$ is the annual mean of the true daily watershed outflow. The relative standard error (R_e) is:

$$R_e = S_e / S_y \tag{5-4}$$

in which and S_y is the standard deviation of the true daily watershed outflow:

$$S_y = \left[\frac{1}{N-1} \sum_{j=1}^{365} (Q_j - \overline{Q})^2 \right]^{0.5} \tag{5-5}$$

5.2 ANALYSES OF OUTFLOW SENSITIVITY

The sensitivity analyses of predicted runoff provide information on the importance of the parameters that represent the hydrologic processes in the HSPF. Knowledge of the potential importance of the parameters as a function of the flow components will lead to a better calibration approach and thus to more accurate predictions.

Provided that the measured runoff was generated under the assumption that the loss of water to deep percolation did not occur (DEEPFR = 0.0), a preliminary analysis to determine the effect of including the parameter DEEPFR in the overall analyses was made. The parameter DEEPFR represents the fraction of groundwater inflow that will enter deep (inactive) groundwater and, thus, be lost from the system. The results indicated that a value of 0.05 did not influence the accuracy of the predicted runoff. Based on these results it was decided not to include this parameter in the subsequent analyses. However, it is recommended to investigate the possibility of deep percolation prior to the design of any HSPF application.

5.2.1 Sensitivity of Flow Discharge to Changes in AGWRC

To determine the effect of the AGWRC nonlinearity on the relative bias and the relative standard error of the predictions a sensitivity analysis was performed. The results indicated that the parameter AGWRC was the most important of the analyzed parameters. The importance of the parameter AGWRC regardless of the baseflow or quickflow dominance in the watershed may be explained by the position of the parameter AGWRC in the flow diagram of water movement and storages, modeled in the PWATER section of the PERLND Application Module (Figure 2-1). If the percolation is zero, AGWRC controls the groundwater outflow (AGWO) to the stream. The other two exits of water to the stream are located above the ground water outflow, specifically in the interflow outflow (IFWO) and in the surface outflow (SURO) boxes. Thus, the accurate prediction of the baseflow component is greatly controlled by the accuracy of the parameter AGWRC.

It is important to investigate the nonlinearity of the parameters as it affects the sensitivity of the parameters and the interpretation of the optimized values. For example, AGWRC showed nonlinearity that related to its temporal meaning. The effect of parameter nonlinearity was evident from the relative biases (Table 5-2) and relative standard error ratios (Table 5-3) as they were lower when AGWRC was reduced by 2% than when it was increased by 2%. Although an increase of 2% for the parameter AGWRC (from 0.97 to 0.9894) represents 61 more days for the groundwater to recede (Table 2-2), the reduction of 2% in AGWRC (from 0.97 to 0.9506) represents 13 less days for the ground water to recede. The nonlinearity effect was also observed from the relative standard error ratios, as the standard error of the predicted runoff decreased when the parameter value was reduced.

The most significant effect occurs when the parameter value is increased (Figure 5-1) with under prediction of runoff. In Watershed 1 the average under prediction was 2.5% with a maximum value of 16% for the year 1992. In Watershed 2, the average under prediction was 1.4%, with a maximum value of 8% for the year 1999. These results were expected because Watershed 1 is baseflow dominated.

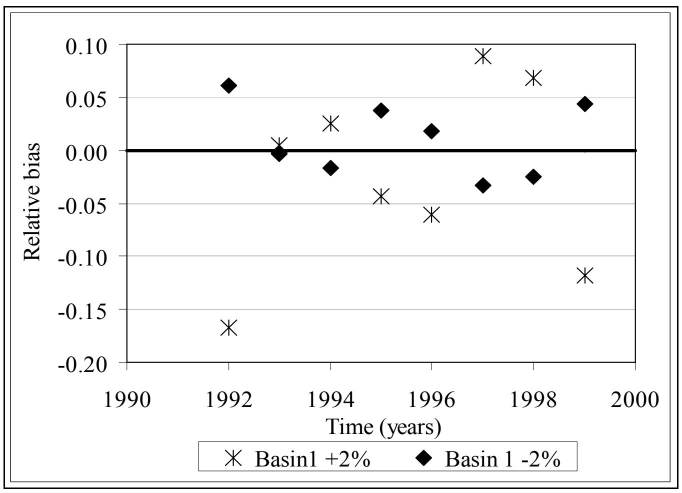


Figure 5-1. Effect of 2% change in AGWRC on the relative bias of the runoff

Table 5-2. Relative bias of the runoff after a +/- 2% change in parameter AGWRC. Initial value of AGWRC was 0.97 in both watersheds.

YEAR	ANNUAL	WATERSHED 1 - R_b		WATERSHED 2 - R_b	
	Precipitation	2%	-2%	2%	-2%
1992	38.4	-0.1671	0.0615	-0.0211	0.0077
1993	42.5	0.0046	-0.0000	-0.0048	0.0046
1994	43.1	0.0252	-0.0170	-0.0041	0.0018
1995	37.1	-0.0435	0.0379	-0.0794	0.0367
1996	53.8	-0.0611	0.0186	-0.0285	0.0057
1997	34.3	0.0883	-0.0332	0.0649	-0.0153
1998	33.7	0.0686	-0.0249	0.0447	-0.0173
1999	43.5	-0.1177	0.0433	-0.0807	0.0308
Mean	40.8	-0.0253	0.0103	-0.0136	0.0068

The nonlinearity effect was also observed in the relative standard error ratio (S_e / S_y), as the error of the predicted runoff decreased as the parameter value decreased (Figure 5-2 and Table 5-3). Since AGWRC is associated with baseflow, it was expected that larger changes in (S_e / S_y) would occur for the watershed with a predominant baseflow when the AGWRC was changed. The parameter AGWRC controls the amount of groundwater outflow to the stream; thus, the accuracy of the runoff for a watershed in which the baseflow is dominant should be more affected by changes in the AGWRC parameter than for watersheds where the quickflow is the dominant component.

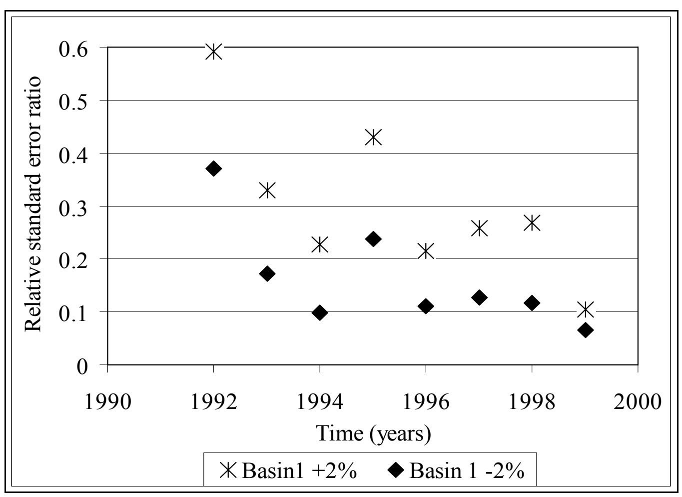


Figure 5-2. Effect of 2% change in AGWRC on the relative standard error ratio of the runoff for the watershed with predominant baseflow component.

Table 5-3. Relative standard error ratio (S_e / S_y) of the runoff after a +/- 2% change in parameter AGWRC. The initial value of AGWRC was 0.97 in both watersheds.

Year	Annual	Watershed 1 - (S_e / S_y)		Watershed 2 - (S_e / S_y)	
	Precipitation (in.)	2%	-2%	2%	-2%
1992	38.4	0.591	0.371	0.034	0.024
1993	42.5	0.330	0.173	0.019	0.011
1994	43.1	0.228	0.098	0.029	0.013
1995	37.1	0.431	0.237	0.091	0.046
1996	53.8	0.214	0.111	0.087	0.041
1997	34.3	0.257	0.126	0.117	0.055
1998	33.7	0.268	0.117	0.086	0.034
1999	43.5	0.104	0.066	0.047	0.028
Mean	40.8	0.303	0.162	0.064	0.032

5.2.2 Effect of Flow Proportions in Parameters Importance

The flow proportions that constitute the runoff were expected to have an effect on the importance of the parameters. Parameters that affect baseflow volumes would likely be more important on watersheds where baseflow was the predominant flow component. This was the case of parameter AGWRC for example, with a greater effect on Watershed 1 with 81% of baseflow than for Watershed 2 where baseflow was only 33%.

The results indicated that the sensitivity of the predicted discharges to changes in the parameters was a function of the flow proportions (baseflow, interflow, and surface flow) and the amount of precipitation during the year. Similarly, changes in parameters associated with surface flow were expected to have a greater effect on the predicted runoff for watersheds in which the quickflow was the predominant component. The parameter UZSN in this case, which is a surface storage parameter, would most likely have a greater effect on the accuracy of the runoff for Watershed 2 where the proportion of quickflow was 67%, than for Watershed 1 where 19% of the total flow was quickflow.

The changes in the relative bias were used as indicators of parameter importance. Although 2% change was used for AGWRC, 10% change was used for all other parameters, as they are much less important than AGWRC. The relative bias for all of the parameters and each year of record was computed using the daily pairs of actual and predicted outflows (Eq. 5-3) and are shown in Table 5-4 through Table 5-7. LZET was the second most important parameter, but much less important than AGWRC. The other parameters were much less important than LZET, but UZSN, INFILT, and LZSN were noticeably more important than BASETP, INTFW, NSUR, AGWETP, and IRC. In the quickflow-predominant watershed, for example, Watershed 2, AGWRC was also dominant, with UZSN less important and LZET even less important. The LZSN parameter showed minor importance, but the other six parameters did not show any effect.

In the baseflow-predominant watershed, the second most important parameter was LZET, which controls the evapotranspiration loss from the lower-zone storage and thus, is one of the mechanisms that control the volume of water in storage. Its importance may be explained by the amount of baseflow in this watershed (81%), its position in the model (Figure 2-1), and by the nominal capacity of the storage in comparison to other storages in the model. Overall, the nominal capacity of the lower-zone storage (LZSN) is one order of magnitude greater than the magnitudes of both the upper zone (UZSN) and the interception storages (IFWS) from where evapotranspiration is withdrawn. Thus, a change in the volume of water in the lower zone will have a greater effect on the total predicted outflow than changes in the upper zone or the interception storages.

The parameters in Watershed 2, which is the watershed with a predominant quickflow component, were also analyzed for relative importance. The parameter UZSN was ranked second in sensitivity behind AGWRC. This may be explained by the proportion of quickflow in the watershed (67%) and by the fact that UZSN is the only parameter in this soil layer that controls several processes (for example, evapotranspiration, and infiltration and percolation to lower zones). Actual evapotranspiration from this soil layer is based on the moisture in storage in relation to its nominal capacity, and it will occur only if the ratio of upper zone storage to nominal capacity (UZS/UZSN) is greater than 2.0.

Table 5-4. Relative bias of the runoff for Watershed 1 after a change of +10% in parameter values

YEAR	LZET	INFIL	LZSN	BASE	AGW	NSUR	INTF	IRC	UZSN
1992	-0.024	0.006	-0.038	-0.002	-0.001	-0.001	0.003	0.000	-0.028
1993	-0.018	0.001	-0.007	-0.001	-0.001	-0.001	-0.000	0.000	-0.010
1994	-0.013	0.005	-0.004	-0.001	-0.000	0.000	0.000	0.000	-0.006
1995	-0.059	0.008	0.012	-0.002	-0.001	-0.001	0.002	0.000	-0.013
1996	-0.013	-0.002	-0.007	-0.000	0.000	-0.001	0.000	0.000	-0.010
1997	-0.021	0.012	0.006	-0.001	-0.001	0.001	0.001	0.000	-0.000
1998	-0.015	0.001	-0.002	-0.001	-0.001	-0.000	-0.000	0.000	-0.006
1999	-0.060	0.015	0.004	-0.001	-0.001	-0.001	0.001	0.000	-0.017
Mean	-0.028	0.006	-0.005	-0.001	-0.001	-0.001	0.001	0.000	-0.011

Table 5-5. Relative bias of the runoff for Watershed 1after a change of –10% in parameter values

Year	LZET	INFIL	LZSN	BASE	AGW	NSUR	INTF	IRC	UZSN
1992	0.021	-0.006	0.037	0.002	0.001	-0.001	-0.004	0.000	0.033
1993	0.018	-0.001	0.007	0.001	0.000	-0.001	0.000	0.000	0.010
1994	0.013	-0.005	0.004	0.001	0.000	0.000	-0.000	0.000	0.006
1995	0.063	-0.008	-0.011	0.002	0.001	-0.001	-0.001	0.000	0.015
1996	0.013	0.002	0.008	0.000	0.000	-0.001	-0.000	0.000	0.011
1997	0.020	-0.012	-0.007	0.001	0.001	0.001	-0.001	0.000	0.001
1998	0.017	-0.002	0.004	0.001	0.001	-0.000	0.000	0.000	0.006
1999	0.064	-0.011	-0.001	0.001	0.001	-0.001	-0.001	0.000	0.020
Mean	0.029	-0.005	0.005	0.001	0.001	-0.001	-0.001	0.000	0.013

Table 5-6. Relative bias of the runoff for Watershed 2 after a change of +10% in parameter values

Year	LZET	INFIL	LZSN	BASE	AGW	NSUR	INTF	IRC	UZSN
1992	-0.001	-0.000	-0.003	-0.001	-0.000	-0.000	0.000	-0.001	-0.032
1993	-0.002	0.000	0.000	-0.000	-0.000	-0.001	0.000	0.000	-0.010
1994	0.000	0.002	0.002	-0.001	-0.000	-0.001	0.001	0.000	-0.007
1995	-0.042	-0.001	-0.016	-0.002	-0.001	-0.002	0.002	-0.000	-0.019
1996	-0.006	-0.001	-0.003	-0.000	-0.000	-0.001	0.000	0.000	-0.005
1997	-0.011	0.004	-0.006	-0.001	-0.001	-0.000	0.001	0.000	-0.011
1998	-0.002	0.003	-0.005	-0.001	-0.000	0.000	-0.000	0.000	-0.003
1999	-0.010	-0.002	-0.023	-0.001	-0.001	-0.002	0.001	0.000	-0.034
Mean	-0.009	0.001	-0.007	-0.001	-0.001	-0.001	0.001	0.000	-0.015

Table 5-7. Relative bias of the runoff for Watershed 2 after a change of -10% in parameter values

YEAR	LZET	INFIL	LZSN	BASE	AGW	NSUR	INTF	IRC	UZSN
1992	0.001	0.000	0.002	0.001	0.000	0.000	-0.000	0.000	0.033
1993	0.001	0.000	0.001	0.001	0.000	0.001	-0.000	-0.003	0.011
1994	0.000	-0.002	-0.002	0.001	0.000	0.001	-0.001	0.000	0.008
1995	0.039	0.002	0.013	0.002	0.001	0.002	-0.002	0.000	0.016
1996	0.008	0.001	0.003	0.000	0.000	0.000	-0.000	0.000	0.006
1997	0.014	-0.003	0.008	0.001	0.001	0.001	-0.000	0.000	0.013
1998	0.002	-0.003	0.006	0.001	0.000	-0.000	0.000	0.000	0.002
1999	0.012	0.002	0.024	0.001	0.001	0.002	-0.001	0.000	0.036
Mean	0.010	-0.001	0.007	0.001	0.001	0.001	-0.001	0.000	0.016

The effect of the flow proportions on the importance of the parameters was also observed in the calculated standard error ratio, as shown in Figure 5-3 (a) and Figure 5-

3 (b) for Watershed 1, and in Figure 5-3 (c) and Figure 5-3 (d) for Watershed 2. However, the order of parameter importance was different to that indicated using the results of the relative bias. For example in Watershed 1 with a predominant baseflow, the parameter INFILT was the second most important parameter as indicated by the relative standard error ratio in contrast with LZET, which was identified as the second most important parameter by the relative bias. Parameter LZET has a greater effect on the volume of the runoff rather than in the temporal distribution. In the case of the parameter LZET, by increasing or decreasing its value, the total volume of water leaving the watershed is decreased or increased leading to a positive or negative bias. However, in the case of the parameter INFILT, the main effect is the distribution of the water among the soil layers and not in the volume of the runoff, which could make a difference for pollutant transport. In Watershed 2, the parameter UZSN was classified as the second most important in both the standard error ratio and the relative bias analyses. In this case, the flow proportions in the watershed influence the effect of the parameter in both the volume and distribution of the runoff.

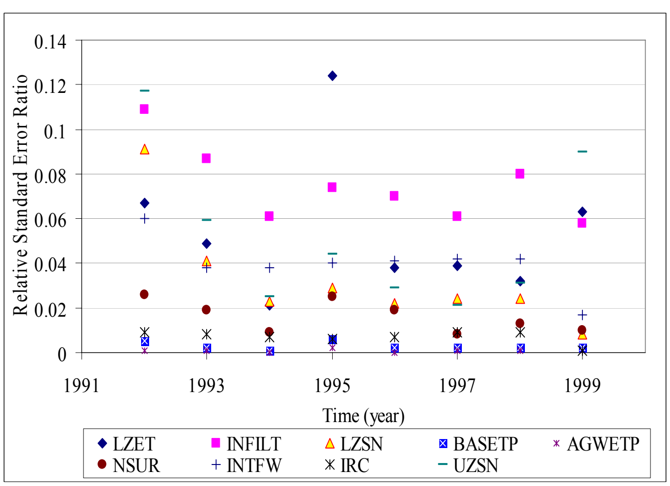


Figure 5-3. **Relative standard error ratio of the daily predicted runoff. Effect of change in parameter value. Watershed 1: (a) +10%; (b) – 10%. Watershed 2: (c) + 10%; (d) – 10%**

The results from the relative bias and the relative standard error ratio were used to classify the parameters within four groups (Table 5-8). Regardless of the flow proportions in the watershed and when the parameter were varied by 10%, a first group (Low) changed the runoff by less than 0.5%; in the second group (Medium) the change in the runoff was between 0.5% and 3%. For Watershed 1 and the third group (High), the runoff varied between 3% and 6%, while in Watershed 2 for the same group the variation of the outflow was between 3% and 4%. In a similar way but In this case, with a parameter variation of 2% (AGWRC) and for the group classified as Extreme, the runoff in Watershed 1 varied between 6% and 16% while in Watershed 2 the variation was only between 4% and 8%. These results suggest that the runoff will be more affected by changes and inaccuracies in the parameters that control the water budget in watersheds with predominant baseflow than by similar changes in watersheds with predominant quickflow.

Table 5-8. Relative importance of the parameters base on the maximum change of the runoff. For AGWRC the sensitivity in the runoff is due to a 2% change in the parameter value.

Parameter	Watershed 1	Watershed 2
AGWRC	6% < Extreme < 16%	4% < Extreme < 8%
LZETP	3% < High < 6%	3% < High < 4%
LZSN	3% < High < 6%	0.5% < Medium < 3%
UZSN	0.5% < Medium < 3%	3% < High < 4%
INFILT	0.5% < Medium < 3%	Low < 0.5%
INTFW	Low < 0.5%	Low < 0.5%
NSUR	Low < 0.5%	Low < 0.5%
IRC	Low < 0.5%	Low < 0.5%
BASETP	Low < 0.5%	Low < 0.5%
AGWETP	Low < 0.5%	Low < 0.5%

Regardless of the flow proportions in the watershed, the parameter importance varied from year to year. In fact, the effect can vary significantly and the variation is often associated with the variation in rainfall depths and temperature. The two largest values of the standard error ratio for the parameter AGWRC were observed for the years 1992 and 1995 (see Figure 5-2). During these two years, the rainfall amounts were low (38.4 and 37.1 in.), and the observed air temperatures and thus the annual potential evapotranspirations were higher in comparison to other years.

CHAPTER 6

EFFECT OF FLOW PROPORTIONS ON HSPF MODEL CALIBRATION EFFICIENCY

6.1 INTRODUCTION

As the state of technology has advanced, the interest in using complex models has increased. Satellite imagery and GIS have provided comprehensive databases that did not exist when the previous generation of hydrologic models was developed. More complex models require a larger array of inputs and more sophisticated methods of calibration. Advances in computer speed and storage capacity have made possible more powerful calibration methods, e.g., PEST (Doherty, 2001) and SCE (Yapo, 1996). While the basis for these methods (Wilde and Beightler, 1967) has existed for decades, the methods are now practical for use with continuous rainfall-runoff models. However, for these algorithms to converge to the global optimum solution, advances in formulating and quantifying components of the objective function are necessary.

Continuous hydrograph models, while having been used in the research community for more than a generation, e.g., HSPF and the Storm Water Management Model (SWMM), are now being used for planning, design, and watershed assessment. Unlike simple single-storm event models, e.g., The KINematic Runoff and EROSion model (KINEROS) and the Dynamic Watershed Simulation Model (DWSM), that can be calibrated with analytical least squares, continuous flow models need to use a multicomponent objective function in order to locate the global optimum solution. With continuous streamflow data, finding the optimum parameter values requires accurate fitting of multiple hydrologic criteria such as peak flows, baseflows, flow volumes, recession rates, storage volumes, evapotranspiration rates, and the autocorrelation of daily flows. The statistical objective function that defines best fit needs to have components that reflect each of the important hydrologic criteria. Finding the global optimum solution requires proper weighting of the individual components of the objective function. Having poor estimates of the objective function weights may prevent the calibration algorithm from reaching the global optimum, which means that the final parameter estimates will not accurately reflect the hydrologic processes that they represent. Decisions made using a model with erroneous parameter values can be faulty.

Algorithms that are used to calibrate continuous flow models, such as the HSPF package (Bicknell et al. 1993 and 2001), are able to use multicomponent objective functions, but the overall effectiveness of the calibration depends on the selection of accurate weights for the components of the objective function. To date, a reliable method for assigning weights to the components has not been reported. A method for obtaining estimates of the weights to apply to the components is presented in this document. In addition to improving prediction accuracy, a reliable method of assigning weights will reduce the time required for calibration. Accurate weights should improve decisions made with the model, such as the selection of a Total Maximum Daily Load (TMDL). A TMDL is a calculation of the maximum amount of a contaminant or pollutant that a waterbody can receive and still meet water quality standards. A TMDL also allocates pollutant loadings among point and nonpoint sources.

The theoretical and empirical mathematical functions used in the HSPF model include parameters that reflect specific processes of the hydrologic cycle. Identifying the best values of these parameters is a goal of calibration. However, the calibrated values of the parameters need to be rational in order for the model to be hydrologically accurate. The importance of the individual parameters will be a function of the hydrologic characteristics of the watershed and, therefore, their fitting should be guided by both mathematical and hydrologic criteria.

6.2 MULTIPLE OBJECTIVE CALIBRATION OF HSPF

The quality of a simulation of hydrologic processes can be represented by an objective function that reflects the accuracy of hydrologic criteria. Computerized optimization methods produce objective calibrations, reduce the time needed to find the optimum parameter values, and make the fitting process reproducible. The replication of the solution is perhaps one of the most important aspects from a legal standpoint when establishing TMDLs because in many cases, the model results are challenged in court. From the plaintiff's perspective, it seems unreasonable to comply with regulations set from the outcome of a subjective optimization, when systematic criteria can be used. Thus, a reasonable scenario is for a political jurisdiction to require the use of consistent optimization criteria for calibration.

The model-independent-parameter estimator PEST uses the Gauss-Marquardt–Levenberg (Marquardt, 1963) algorithm for minimizing the objective function (Φ), which has the option of using a weighting of multiple criteria. Each criterion is based on a sum of the squared differences between the model predictions of a hydrologic criterion and the corresponding measured values. The components of the objective function represent individual hydrologic objectives, such as matching peak flows or monthly volumes of runoff. For analyses, the quality of a calibration is measured by the final weighted sum of the components of the objective function.

6.3 OBJECTIVE FUNCTION COMPONENTS

For continuous watersheds models, the objective function must reflect the different physical processes that are inherent in a continuous hydrograph, including peak flows, recessions, and baseflows. This requires a multicomponent objective function, and the importance of the individual components should reflect the relative importance of the hydrologic processes inherent to the measured flows that make up the continuous hydrograph. A single-component objective function, such as the correlation coefficient or the Nash-Sutcliffe efficiency index (Nash and Sutcliffe, 1970), cannot be sufficiently sensitive to variation of all hydrologic processes. For example, model parameters that reflect baseflow rates need an objective function component that is sensitive to variation in those rates.

To test the hypothesis that flow proportion weighting is critical to achieving accurate parameter values, analyses were undertaken to develop relations. The analyses were made using data from eight hypothetical watersheds that had varying amounts of flow proportions (surface runoff, interflow, and baseflow). The HSPF generated flows were assumed to be actual data and are referred to as “measured”. The measured flows are the sum of the surface flow, quickflow, and baseflow components prior to entering the stream. Stream processes were not included in the analysis.

Six possible components, each of which represents a different hydrologic criterion, are presented below and were tested with the HSPF model. Each component of the objective function is a least squares computation, with weights applied to each of the components. The overall value (Φ) of the objective function is minimized using the following weighted sum:

$$\Phi = \sum_{i=1}^{m} w_i \Phi_i \tag{6-1}$$

where w_i is the weight applied to the i^{th} component Φ_i and m is the number of components used for a given calibration. The components of the objective function developed for the PEST algorithm were defined based on hydrologic criteria between the measured and predicted continuous hydrographs. The goal was to incorporate physical hydrologic concepts into the calibration process through the objective function. The total continuous hydrograph was separated into parts that reflect the different hydrologic processes that would be inherent to almost any continuous discharge hydrograph. These flow components are used in computing the components of the objective function. The following six components were formulated to reflect hydrologic criteria:

Daily Outflow Component: This component is a measure of the accuracy of the predicted daily outflows. The least squares calculation between the predicted and measured daily watershed outflows is:

$$\phi_1 = \sqrt{\sum_{i=1}^{n} (\hat{Q}_i - Q_i)^2} \tag{6-2}$$

where Q_i is the measured daily watershed outflow, $\hat{Q}_i$ is the HSPF predicted daily watershed outflow, the subscript i is the day, and n is the number of days of record.

Monthly Volumes Component - A model used for TMDL estimation must provide accurate runoff volumes. The underestimation of volumes can lead to negatively biased pollution loads. Therefore, it is important to accurately reproduce flow volumes. The second objective function component (Φ_2) reflects the accuracy of measured and predicted monthly volumes:

$$\phi_2 = \sqrt{\sum_{j=1}^{m}(\hat{V}_j - V_j)^2} \tag{6-3}$$

where V_j is the measured monthly outflow volume, $\hat{V}_j$ is the HSPF predicted monthly volume at the edge of the stream, the subscript j is the month, and m is the number of months of record.

Autoregression Component: The autoregressive nature of the flows relate to the accumulation of the precipitation within the watershed. The autoregressive properties of the runoff hydrograph reflect the effect of the storage properties of a watershed on the release rate of water from storage. The degree of autocorrelation would primarily depend on the smoothness of the baseflow. Continuous models such as HSPF include storage parameters that represent different aspects of watershed storage, and therefore, the objective function should include a component that is sensitive to storage. HSPF parameters, such as the upper zone nominal storage (UZSN) and the lower zone nominal storage (LZSN), control the release of water from the storages.

The degree to which flows are similar in adjacent time periods is modeled using a one-day-lag autoregressive equation as the basis of a component in the objective function. The autoregressive equation relates day t outflow to outflow for day $t-1$. A one-day time lag is individually applied to the measured and predicted daily flows using the following equations:

$$A_i = Log_{10}(Q_i + 0.001) - Log_{10}(Q_{i-1} + 0.001) \tag{6-4}$$

$$\hat{A}_i = Log_{10}(\hat{Q}_i + 0.001) - Log_{10}(\hat{Q}_{i-1} + 0.001) \tag{6-5}$$

where Q is the measured daily watershed outflow, $\hat{Q}$ is the HSPF predicted daily watershed outflow, and the subscripts i and $i-1$ refer to the day i and $i-1$, respectively. Then, the least squares calculation is based on the difference of the one-day lag variable A of the measured and predicted outflows:

$$\Phi_3 = \sqrt{\sum_{i=1}^{n-1}\left(\hat{A}_i - A_i\right)^2} \tag{6-6}$$

in which n is the simulated number of days and Φ_3 is the value of the autoregression component of the objective function of Eq. 6-1.

Quick-flow Filter Component: A continuous hydrograph is often thought to consist of three parts: direct runoff, interflow, and baseflow. These are represented in the HSPF model as three intermediate outflows that combine to form the total discharge. The sum of the direct runoff and the interflow is referred to as the quickflow. A version of PEST was modified to include filters that could be used to separate the discharge hydrograph into two parts: quickflow and baseflow. A Butterworth filter

(Butterworth, 1930) can be used to obtain a continuous baseflow time series, while the Quickflow filter (Nathan and McMahon, 1990) based on moving averages can be used to extract the quickflow time series.

The measured and HSPF predicted daily watershed outflows were separately filtered using the Quickflow filter. Two output signals associated with the quickflow portion were generated, one for the HSPF predicted discharge rates and one for the measured discharge rates. The Quickflow filter is:

$$q_i = \alpha q_{i-1} + \frac{(1+\alpha)}{2}(Q_i - Q_{i-1}) \quad \textbf{(6-7)}$$

where q_i is the quickflow discharge at time i; i is the time step index (days); Q_i is the watershed outflow at time i; and α is the scaling parameter that controls the volume of quickflow. Smakhtin and Watkins (1997) found that the optimal filter parameter α usually fluctuates in the range from 0.985 and 0.995 and recommended the value of 0.995 as being suitable for most of the daily baseflow and quickflow separations. Therefore, 0.995 was used herein. The filtering provided by Eq. 6-7 is sometimes referred to as a recursive filter (Shumway, 1988). The least squares quickflow component (Φ_4) was computed with the filter output signals:

$$\Phi_4 = \sqrt{\sum_{i=1}^{n} \left(\hat{q}_i - q_i\right)^2} \quad \textbf{(6-8)}$$

where $\hat{q}_i$ is the predicted quickflow discharge at time i, and q_i is the value for quickflow discharge at time i derived from the measured discharge time series.

Butterworth Filter Component: A Butterworth filter (Butterworth, 1930) with a high-pass band and a cutoff-frequency specified by the user (in $days^{-1}$) was applied to obtain two output signals, which are referred to as the predicted and measured surface runoff. The filter is described by:

$$s_i = L^{-1}\left[H_s * L(Q_i)\right] \tag{6-9}$$

where s_i is the filtered surface runoff output signal at time i; i is the time step index; Q_i is the total predicted or measured flow at time i; H_s is the transfer function of the system that relates the spectrum of the input signal to the spectrum of the corresponding output signal, and L is the Laplace transform. The least squares calculation for this component (Φ_5) between the two surface runoff time series is

$$\Phi_5 = \sqrt{\sum_{i=1}^{n}\left(\hat{s}_i - s_i\right)^2} \tag{6-10}$$

in which $\hat{s}_i$ is the HSPF predicted surface runoff at time i, and s_i is the measured surface runoff at time i.

Baseflow Separation Component: A second baseflow separation component was incorporated into PEST to separate the baseflow from the total runoff. The method of separation was a modification of the method included in HYSEP (Sloto and Crouse, 1996). In the modified version developed herein, the length of the interval is specified by the user. The component has two options: (1) the local minimum method and (2) the sliding-interval method. The hydrograph separation is applied to both the measured and HSPF predicted outflow to obtain two time series, with each representing the baseflow portion of the respective series. The least squares calculation for this component (Φ_6) is used as a measure of the accuracy of the baseflow separation:

$$\Phi_6 = \sqrt{\sum_{i=1}^{n}\left(\hat{b}_i - b_i\right)^2} \qquad \textbf{(6-11)}$$

where b_i is the measured daily baseflow obtained through the hydrograph separation, $\hat{b}_i$ is the predicted daily baseflow obtained through the hydrograph separation, the subscript i is the day, and n is the simulated number of days in the record.

6.3.1 Alternative Objective Functions

Two alternative objective functions were analyzed to determine if the individual components would generally have an effect on prediction accuracy and to determine if the Butterworth and Quickflow filters could specifically be used as methods for hydrograph separation instead of the sliding-interval method or the local-minimum method. The first objective function consisted of components 1, 2, 3, 4, and 5, while the second objective function included components 1, 2, 3, 4, and 6. Each of the components of the objective function is associated with a particular portion or portions of the total watershed outflow. The autoregression component, Eq. (6-6), is associated with watershed storage. The Butterworth filter component, Eq. (6-10), and the baseflow separation component, Eq. (6-11), are associated with baseflow. The Quickflow filter component, Eq. (6-8), is associated with the quickflow portion (i.e., surface runoff plus interflow).

6.4 IMPORTANCE OF FLOW PROPORTIONS TO CALIBRATION

The use of incorrect weights with Eq. (6-1) can prevent the model-independent-parameter estimator PEST from identifying the optimum hydrologic solution. Therefore, the first need was to develop a method of selecting weights in such a way that they reflect the hydrologic processes imbedded in the measured streamflows. Analyses were made using data from eight hypothetical watersheds with areas of 13 km2 (5 mi2) and a forested land use. The distributions of flow (surface runoff, interflow, and baseflow) for the eight watersheds were varied as shown in Table 6-1. The hypothetical watersheds were designed to show a wide range for each flow type. The eight discharge time series were produced using the HSPF model. Thus, the exact solutions including the true parameter values were known. The total computed outflow was a surrogate for measured data. The nine calibrated HSPF parameters were: AGWRC, BASETP, AGWET, NSUR, LZETP, LZSN, UZSN, INFILT, and INTFW. The HSPF parameter values for the eight hypothetical watersheds are given in Table 6-1.

The proportions of baseflow and quickflow in each of the eight watersheds were estimated using either the sliding-interval or the local-minimum method, whichever produced the better accuracy. The baseflow separation analyses were made using the sliding-interval method and the local-minimum method with intervals of 3, 5, 7, 9, 11, 13, 15, and 17 days. The relative bias and relative standard error ratio (McCuen, 2003) between the estimated and true flow proportions were compared and used to select the better of the two methods and the best interval. The relative bias$(\bar{e}/\bar{Y})$, which is the ratio of the bias $\bar{e}$ to the mean discharge $\bar{Y}$, is a measure of the systematic error of the predicted discharge rates and is computed by:

$$\frac{\bar{e}}{\bar{Y}} = \frac{1}{N\bar{Y}} \sum_{j=1}^{N} \left(\hat{Q}_j - Q_j \right) \quad \textbf{(6-12)}$$

where Q_j is the measured daily outflow for the day j, or if the analysis is for the accuracy of the hourly outflows, then Q_j is the measured hourly outflow for the hour j; and N is the number of days when the analysis is using the daily outflows, or the number of hours when the analysis is using the hourly outflows. The relative standard error (S_e / S_y) of estimate is:

$$\frac{S_e}{S_y} = \left[\sum_{j=1}^{N} (\hat{Q}_j - Q_j)^2 \Big/ \sum_{j=1}^{N} (Q_j - \bar{Q})^2 \right]^{0.5} \quad \textbf{(6-13)}$$

where the S_e / S_y is a measure of the nonsystematic error, with a value of zero indicating a perfect fit.

Table 6-1. Percentages of flow distribution for the eight hypothetical watersheds and true values of the parameters that control the water budget of pervious areas. Bf = percent of baseflow; If = percent of interflow; Sf = percent of surface runoff; and Qf = percent of quickflow.

Watershed	Bf	If	Sf	Qf	LZSN	INFILT	UZSN	INTFW
1	81	7	12	19	5	0.070	2.0	0.5
2	70	10	20	30	7	0.035	2.0	0.8
3	59	27	14	41	5	0.030	2.5	1.4
4	58	35	7	42	3	0.040	2.0	2.0
5	49	22	29	51	6	0.025	1.0	1.0
6	35	32	33	65	2	0.030	1.5	0.5
7	35	8	57	63	2	0.025	0.5	0.5
8	24	66	10	76	2	0.020	1.0	3.0

Note: For all of the watersheds AGWRC = 0.97, BASETP=0.01, AGWET=0.01, NSUR=0.08, and LZETP 0.08.

The baseflow proportion of the total flow influenced the selection of the hydrograph separation method. The results of the baseflow separation analyses indicated that the sliding-interval method is the more accurate for watersheds in which the baseflow proportion was greater than 50 percent of the watershed outflow (watersheds 1-4). For watersheds in which the baseflow proportion was less than 50 percent of the total flow (watersheds 5- 8), the local-minimum method provided the better accuracy. Figure 6-1 shows the relationship between the number of intervals and the percentage of baseflow for the separation method that yielded the better accuracy. The length of the interval decreased as the contribution of baseflow to the total discharge increased. These results were incorporated into the optimization process through the baseflow separation component.

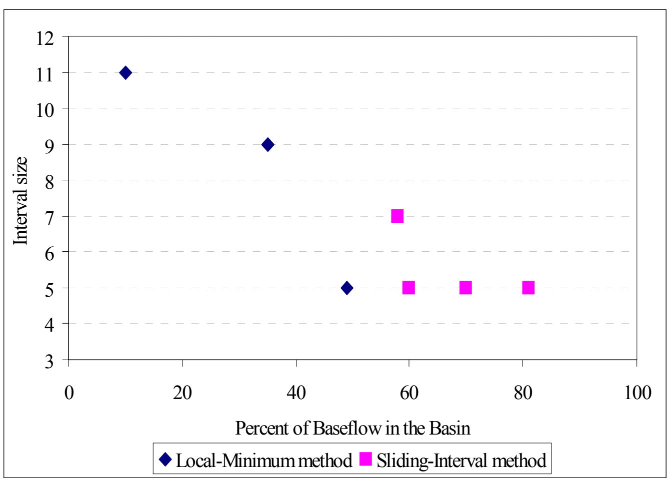


Figure 6-1. **Variation of the number of intervals with the percentage of baseflow for alternative hydrograph separation methods (Watersheds 6 and 7 had the same fraction of baseflow and the optimum number of intervals.)**

6.5 WATERSHED FLOW PROPORTIONS

The hourly variation of measured streamflow reflects the proportions of direct runoff, interflow, and baseflow. A streamflow series with a high percentage of baseflow should place more weight on the components of the objective function that relate to baseflow. Therefore, knowledge of the flow proportions should be helpful in initially setting the weights w_i of Eq. (6-1). It was hypothesized that the flow proportions influence the ability of PEST to find the optimum parameter values. The objectives of this analysis were to determine (1) whether or not the initial assignment of the objective function weights needed to be related to the flow proportions in the discharge time series and (2) whether or not estimates of the flow proportions could provide reasonable estimates of the weights. The first objective represents analysis, while the second objective is the synthesis.

The first objective function was assessed using analyses with both randomly assigned weights and weights assigned based on flow proportions. For the runs in which the optimal solution was reached, the contribution of the component to the objective function was graphed against the proportions of flow, i.e., baseflow and quickflow. The analyses indicated strong associations between the baseflow portion and the autoregression component and between the quickflow portion and the quickflow component (Figure 6-2). An association between the baseflow portion and the Butterworth filter component was less evident.

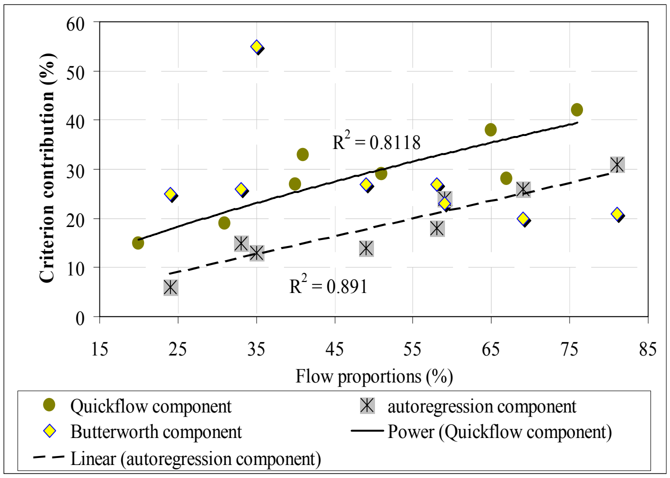


Figure 6-2. **Relation between the percent contribution of the component to the first objective function (Φ) versus the flow proportion, either quickflow or baseflow**

Relations between the flow proportions and the contributions of the components to the objective function (Φ_i) were fitted with the data. The following linear model relates the contribution of the autoregression component to the percentage of baseflow:

$$C_A = 0.364 P_b \tag{6-14}$$

in which C_A is the percent contribution of the autoregression component to the objective function and P_b is the percentage of the total runoff that appears as baseflow. Equation (6-14) had a correlation coefficient of 0.891.

As with the baseflow, a relation between the quickflow and the contribution of the Quickflow filter component to the objective function was established. A power model that provided $R^2 = 0.812$ was fitted to the data to relate the proportion of the total flow that is quickflow (P_q) to the Quickflow filter contribution (C_q):

$$C_q = 1.98 P_q^{0.6908} \tag{6-15}$$

To determine if the established relations of Eqs. (6-14) and (6-15) were effective, even when the initial parameter values were significantly different from the true values, calibrations were made for eight hypothetical watersheds. Random values were assigned to the initial parameter estimates, but the component contributions were set to the values established by Eqs. (6-14) and (6-15). The analyses using the HSPF model with PEST yielded the optimum parameter values regardless of the initial parameter estimates as long as the weights were set using the flow proportion models of Eqs. (6-14) and (6-15). However, the number of iterations required to reach the optimum increased as the initial parameter estimates deviated from the true parameter values. When the weights assigned for Eq. (6-1) did not accurately reflect the flow proportions, convergence to the true parameters was not assured.

6.5.1 Test of Objective Function # 1

The flow proportions of Eqs. (6-14) and (6-15) were then tested using actual data from two watersheds: (1) Bundicks Branch watershed located in the Atlantic Coastal Plain Physiographic Province in Delaware, with a drainage area of 16 km2 (6.25 mi2) and (2) the Little Falls at Blue Mount located in the Piedmont Physiographic Province in Maryland, with a drainage area of 135 km2 (52.9 mi2). The period of record for Bundicks Branch was from 08/19/1998 to 04/09/2000, and the total streamflow depth for the period of analysis was 19.7 cm (7.74 in.) or approximately 11.8 cm/yr (4.6 in./yr). The period of record for Little Falls was from 05/01/1992 to 12/31/1998, and the total streamflow depth for the period of analysis was 318 cm (125 in.), which is about 44.3 cm/yr (17.4 in./yr).

Prior to conducting the test, a hydrograph separation was applied to both watersheds to determine the proportions of baseflow and quickflow. The average baseflow during the simulated period was 80% for Bundicks Branch and 75% for Little Falls. The sliding-interval method was selected based on the results shown in Figure 1. The accuracy of the calibrations was assessed using the relative bias and relative standard error ratio of Eqs. (6-12) and (6-13).

Several factors were taken into consideration when evaluating the quality of the calibrations using the fitted models and actual data. The fitted flow proportion models of Eqs. (6-14) and (6-15) were developed assuming forested conditions (100%). Forest cover was the dominant land cover on both of the test watersheds, with the forested areas in Bundicks Branch and the Little Falls watersheds being 44% and 40%, respectively. Both watersheds included nine land uses, but only the parameters of the predominant land use (forest) were optimized. The parameters of the remaining land uses used the PEST option of tying the parameters of the nondominant land uses to those of the dominant land use.

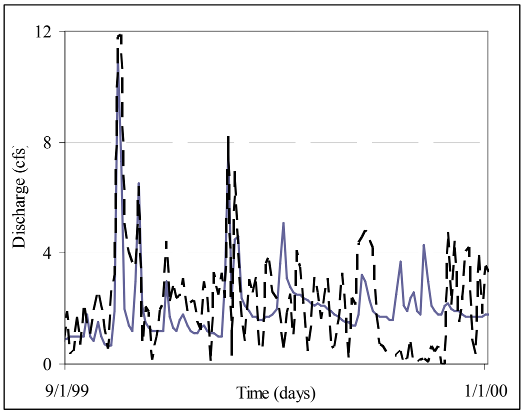
The analyses of actual data were undertaken to assess whether or not: (1) knowledge of flow proportions would affect the accuracy of HSPF calibrations, (2) alternative objective functions would influence calibration accuracy, and (3) the weights for the components of the objective function influenced calibration accuracy. The objective function weights were set in two ways, random assignment and assignment to reflect the flow proportions of the measured data. In the two cases, the same initial values of the parameters were used to ensure that differences in results would not occur because of the initial values of parameters. For the case of assignment based on flow proportions, Eqs. (6-14) and (6-15) were used. This comparative analysis tested whether or not knowledge of the flow proportions of the continuous discharge hydrograph would lead to improved values of the objective function and better estimates of the HSPF parameters.

6.5.2 Bundicks Branch Watershed: Objective Function # 1

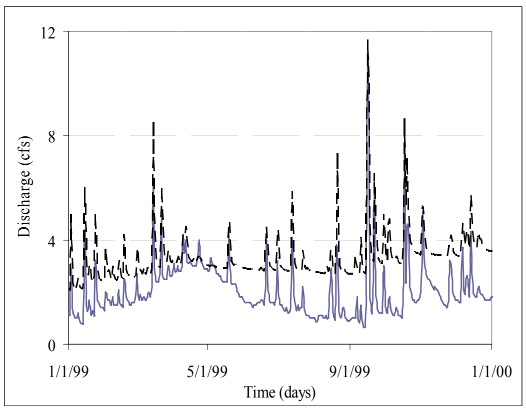
Although the record length of the measured daily discharge rates was short, such record lengths are a common problem in the development of TMDLs. The assigned weights were first set randomly, which corresponds to randomly assigned flow proportions. The initial percentage contributions of the terms of the objective function are given in Table 6-2. The accuracy of the predicted daily discharges was poor (see Table 6-3 and Figure 6-3(a)). For 1998, the optimized parameter values yielded a relative bias of 21% and a relative standard error ratio of 2.19. For 1999, the relative bias and relative standard error ratio were 7% and 1.90, respectively. Figure 6-3 shows consistent overprediction for 1999, which also occurred throughout the period of record.

Table 6-2. Initial contribution, as percentage of the total, of the individual components of the objective functions for objective functions 1 (OF1) and 2 (OF2) using random (RW and flow-proportion (FPW) weighting.

	Little Falls watershed				Bundicks Branch watershed			
	OF1		OF2		OF1		OF2	
Φ_1	RW	FPW	RW	FPW	RW	FPW	RW	FPW
Φ_1	27	25	3	13	25	20	4	9
Φ_2	0	6	53	3	11	11	9	4
Φ_3	3	35	29	41	34	30	17	38
Φ_4	57	20	12	25	14	8	64	24
Φ_5	13	14	--	--	16	31	--	--
Φ_6	--	--	3	18	--	--	6	25



(a)



(b)

Figure 6-3. Bundicks Branch watershed. Measured and predicted daily discharge during the calibration period using the first objective function (a) using arbitrary weights to the component contributions and (b) using the weights of Method 1.

For the second analysis, the component contributions, which are also given in Table 6-2, were set by Eqs. (6-14) and (6-15). The accuracy of the predicted discharges was significantly better than that obtained from the first analysis, as indicated by the goodness-of-fit statistics (Table 6-3). The relative biases for the two full years of record were 45% and 61% and the relative standard error ratios were 1.34 and 1.24. Although the overprediction is still significant, the reductions of the relative biases were significant in comparison to the values of the first analysis. Figure 6-3(b) shows the actual and predicted hydrographs. Figure 6-3(b) shows some improvement over Figure 6-3(a). This comparative analysis shows that knowledge and use of the flow proportions can lead to improved calibrations when applying HSPF and PEST to actual measured data.

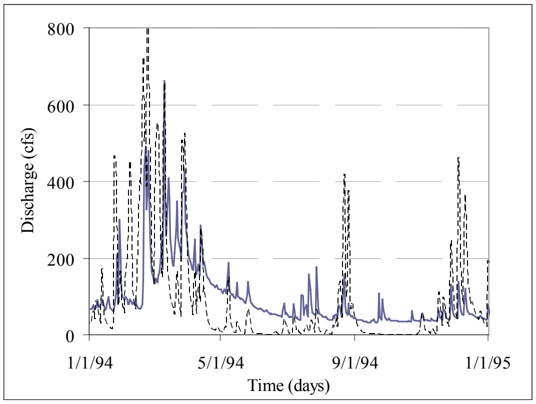
Table 6-3. Bundicks Branch watershed –Goodness-of-fit statistics using random and flow proportion weighting. (OF1 = objective function # 1; OF2 = objective function # 2; P = annual precipitation (cm); R_{bR} = relative bias for total runoff; S_e/S_y = standard error ratio for total runoff; R_{bB} = relative bias for baseflow).

Objective function	Year	P	Random			Flow-proportion		
			R_{bR}	S_e/S_y	R_{bB}	R_{bR}	S_e/S_y	R_{bB}
OF1	**1998**	40	0.21	2.19	0.32	0.45	1.34	0.45
	1999	126	0.07	1.90	0.43	0.61	1.24	0.35
	2000	17	-0.20	2.17	-0.16	0.21	1.14	0.42
OF2	**1998**	40	-0.39	1.92	-0.42	-0.09	1.03	0.15
	1999	126	0.35	2.05	-0.11	0.22	0.59	0.34
	2000	17	1.48	2.90	-0.34	0.23	1.42	0.32

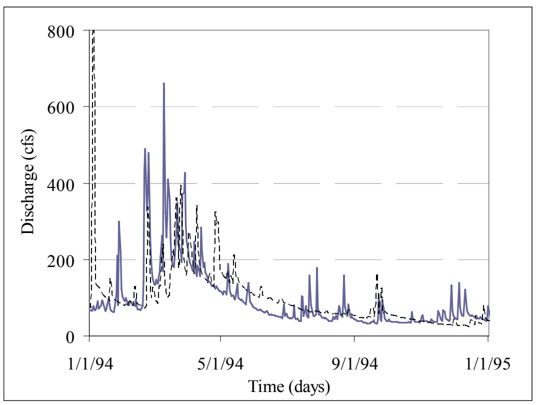
6.5.3 Little Falls Watershed: Objective # 1

The Little Falls database (1992 – 1998) includes a greater variation of hydrologic conditions and larger runoff depths than with the Bundicks Branch database. It was expected that with higher runoff, the predicted discharges would be less sensitive to the optimized parameters and perhaps more sensitive to parameters that control the transport of water in the stream channel. Sensitivity of the stream parameters was not included in this analysis. Two analyses were performed, one using arbitrary weights for the objective function components and the second using weights derived from the flow components.

For the case of random selection, the multicriterion algorithm was not successful because the weights did not reflect the flow proportions that made up the total measured streamflow. The goodness-of-fit statistics (Table 6-4) indicate that the continual underprediction of the baseflow was the most significant problem, with the poorest negative relative bias of baseflow being -83% for the year 1998. The average underprediction of baseflow for the record length was 61%, which is substantial. In addition to baseflow, the relative bias was computed for the total runoff. While the average was much better than for the baseflow, the relative bias was still significant for 4 of the 7 years. The largest relative bias for the total flow was -49% (underprediction) for the year 1998. Accuracy was also assessed with the relative standard error ratio of the total runoff, with an average of 1.39, which is considered poor. For the year 1994, which is shown in Figure 6-4(a), the underprediction of the total flow was –29%, with a relative standard error ratio of 1.7. The use of weights that did not reflect the flow proportions prevented PEST from approaching a more accurate solution.



(a)



(b)

Figure 6-4. **Little Falls watershed. Measured and predicted daily discharge during the calibration period using the first objective function (a) using arbitrary weights to the component contributions and (b) using the weights of Method 1.**

Table 6-4. **Little Falls watershed – Method 1: Statistical summary of HSPF predictions using random and flow proportion weighting. (P = annual precipitation (cm); R_{bR} = relative bias for total daily runoff; S_e/S_y = standard error ratio for total runoff; R_{bB} = relative bias for baseflow).**

		Random			Flow-proportion		
Year	P	R_{bR}	S_e/S_y	R_{bB}	R_{bR}	S_e/S_y	R_{bB}
1992	117	0.00	0.91	-0.61	0.06	0.57	0.48
1993	131	-0.02	1.22	-0.25	0.04	0.48	0.65
1994	85	-0.29	1.69	-0.72	-0.01	0.56	0.07
1995	121	0.15	0.97	-0.65	0.12	0.47	0.47
1996	139	-0.01	1.95	-0.54	0.05	0.66	0.23
1997	102	0.24	1.88	-0.68	0.24	0.51	0.35
1998	58	-0.49	1.07	-0.83	0.10	0.51	0.17
mean	108	-0.06	1.39	-0.61	0.09	0.54	0.34

In the second analysis, the weights w_i were based on the flow proportions and Eqs. (6-14) and (6-15). The goodness of fit can be visually assessed in Figure 6-4 (b), while the goodness-of-fit statistics are given in Table 6-4. The relative bias of the baseflow is large, with a mean relative bias of 34%. The annual values varied from 7% to 65%. These relative biases may be misleading because the mean baseflow is small, which may inflate the relative bias. The relative bias of the total flow is 8.6% with annual values varying from -1% to 24%. The mean relative standard error ratio of 54% is reasonably good, with the mean annual values showing a small range, i.e., 47% to 66%.

The important comparison is between the goodness-of-fit statistics for the random setting of the weights versus the flow-proportion weighting. Setting the weights using the flow proportions decreased the relative standard error ratio from a very poor value of 1.39 to a moderate value of 0.54. This drop is very significant and indicates the benefit of flow proportion weighting. Without weighting, HSPF provided very poor prediction of discharges, with the predicted values being highly biased and very imprecise. Flow weighting improved both the relative bias of the baseflow and precision of the total discharge, with the improvements being very substantial.

6.5.4 Test of Objective Function # 2

Two analyses were made to determine if the initial component contributions had an effect on the accuracy of the predicted daily discharges. In the first analysis, the weights that control the initial contributions of the function were randomly set, and the accuracy of the predicted daily discharges was evaluated. The goodness-of-fit statistics indicate that the accuracy of the daily discharges predicted by HSPF (Table 6-3) using randomly assigned weights was poor. For 1998, the optimized condition yielded a relative bias in the total flow of -0.393 and a standard error ratio of 1.92. For 1999, the relative bias and standard error ratio were 0.349 and 2.05, respectively. These statistics indicate a very poor fit, which is substantiated by Figure 6-6(a).

For the second analysis, the component contributions were set with the flow-proportion models of Figure 6-5 or Table 6-5. The accuracy of the predicted discharges was significantly better as indicated by the goodness-of-fit statistics (Table 6-3). The relative biases were -0.09 and 0.22 for 1998 and 1999, respectively; the standard error ratios were 1.038 and 0.59. While these results do not suggest high prediction accuracy, they are better than the corresponding values for the analysis when initial values were randomly selected without the aid of the equations of Table 6-5. These results indicate that knowledge and use of the flow proportions provide better prediction accuracy with actual data than when the flow proportions are not considered.

To evaluate whether or not the incorporation of the baseflow separation in the objective function was favorable to the improvement of the prediction accuracy, the results from Figure 6-3(b) and 6-6(b) were analyzed. Figure 6-6(b) shows the predicted and measured discharges for the second analysis. The results indicate considerable improvement in goodness of fit when compared to the fit shown in Figure 6-3(b), which was based on the first objective function. This shows that the selection of algorithms to represent the components of the objective function can influence the calibration accuracy.

The poor accuracy of the optimizations for both objective functions may be related to several factors including the short period of record for calibration, the uncertainty in the accuracy of the precipitation caused either by the sampling error or by the method used for the disaggregation of daily into hourly depths, the assumption of a fully forested watershed, the nature of lumped models in which the spatial variation of the land use is ignored within the simulated model segment, and the inability of the objective function components to filter the noise in the input data, among others.

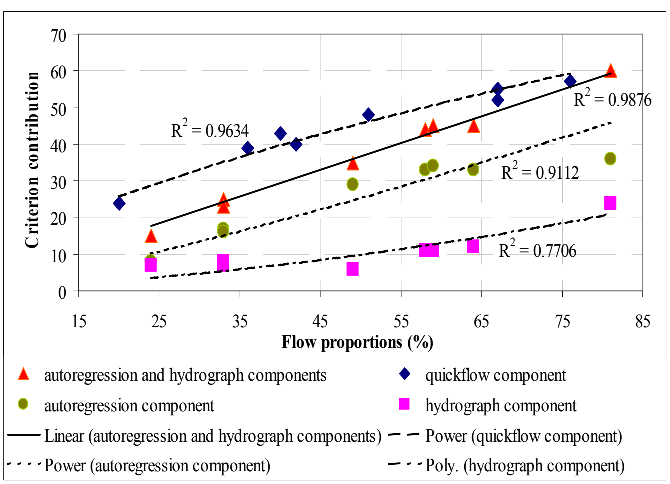


Figure 6-5. **Relation between the percentage contribution of the component to the second objective function (Φ) versus the flow proportion, either quickflow or baseflow.**

Table 6-5. Fitted model and correlation coefficient of multiple determination (R^2) using objective function # 2 to the components contribution and portions of flow in successful optimizations. $\hat{A}$, $\hat{b}$, $\hat{q}$, and $\hat{Y}$ are the contributions to the component of the objective function, X_b is the percentage of baseflow in the watersheds, and X_q is the percentage of quickflow in the watershed.

Objective function component	Model	R^2
Autoregression Φ_3	$\hat{A} = 0.1983 X_b^{1.2388}$	0.9112
Quickflow filter Φ_4	$\hat{q} = 3.9127 X_q^{0.6275}$	0.9634
Hydrograph separation Φ_6	$\hat{b} = 0.002 X_b^{2} + 0.0961 X_b$	0.7706

To test the usefulness of the relations of Table 6-5 and Figure 6-5, the measured data for the Bundicks Branch and Little Falls watersheds were analyzed. The objective of the verification was to assess whether or not accurate estimates of the weights of the objective function components would lead to improved goodness-of-fit statistics for the prediction of total discharge rates. More importantly, the comparison of results based on an alternative objective function could show whether or not fitting accuracy was influenced by the selection of objective function components.

6.5.5 Bundicks Branch Watershed: Objective Function # 2

Two analyses were made to determine if the initial component contributions had an effect on the accuracy of the predicted daily discharges. In the first analysis, the weights that control the initial contributions of the function were randomly set, and the accuracy of the predicted daily discharges was evaluated. The goodness-of-fit statistics indicate that the accuracy of the daily discharges predicted by HSPF (Table 6-3) using randomly assigned weights was poor. For 1998, the optimized condition yielded a relative bias in the total flow of -0.393 and a standard error ratio of 1.92. For 1999, the relative bias and standard error ratio were 0.349 and 2.05, respectively. These statistics indicate a very poor fit, which is substantiated by Figure 6-6(a).

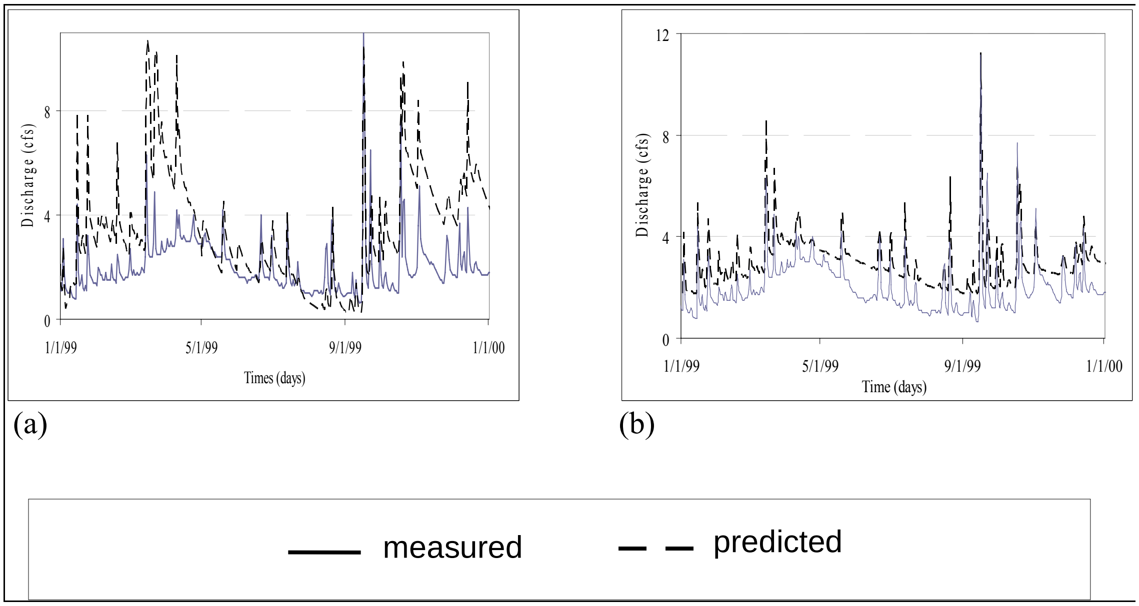


Figure 6-6. Bundicks Branch watershed. Measured and predicted daily discharge during the calibration period using the second objective function (a) using arbitrary weights to the component contributions and (b) using the weights of Method 2.

For the second analysis, the component contributions were set with the flow-proportion models of Figure 6-5 or Table 6-5. The accuracy of the predicted discharges was significantly better as indicated by the goodness-of-fit statistics (Table 6-3). The relative biases were -0.09 and 0.22 for 1998 and 1999, respectively; the standard error ratios were 1.038 and 0.59. While these results do not suggest high prediction accuracy, they are better than the corresponding values for the analysis when initial values were randomly selected without the aid of the equations of Table 6-5. These results indicate that knowledge and use of the flow proportions provide better prediction accuracy with actual data than when the flow proportions are not considered.

To evaluate whether or not the incorporation of the baseflow separation in the objective function was favorable to the improvement of the prediction accuracy, the results from Figure 6-3(b) and Figure 6-6 (b) were analyzed. Figure 6-6(b) shows the predicted and measured discharges for the second analysis. The results indicate considerable improvement in goodness of fit when compared to the fit shown in Figure 6-3 (b), which was based on the first objective function. This shows that the selection of algorithms to represent the components of the objective function can influence the calibration accuracy.

The poor accuracy of the optimizations for both objective functions may be related to several factors including the short period of record for calibration, the uncertainty in the accuracy of the precipitation caused either by the sampling error or by the method used for the disaggregation of daily into hourly depths, the assumption of a fully forested watershed, the nature of lumped models in which the spatial variation of the land use is ignored within the simulated model segment, and the inability of the objective function components to filter the noise in the input data, among others.

6.5.6 Little Falls Watershed: Objective Function # 2

Hydrograph separation was expected to improve the accuracy of the predicted daily discharges, especially the baseflow. The effect of the initial component contributions to the objective function was tested through two analyses. The components contribution were randomly set in the first analysis while in the second analysis the contributions were set according to the equations in Table 6-5. For both analyses, the accuracy of the predicted daily discharges was evaluated (Table 6-6). The results indicate that the accuracy of the predicted daily discharges when the weights for the objective function were not based on flow proportions was poor (

Figure 6-7(a)), as indicated by the goodness-of-fit statistics (Table 6-6). When the assigned weights were set so that the initial component contributions were those obtained from Table 6-5, the accuracy of the predicted discharges was significantly better (Table 6-6 and

Figure 6-7(b)) than in the first analysis. The average relative bias in the baseflow decreased from 54% to 18% while the average relative standard error ratio decreased from 143% to 52%. Both of these reductions are substantial.

Table 6-6. Little Falls watershed – Objective function # 2: Goodness-of-fit statistics for the calibration using random and flow proportion weighting. (P = annual precipitation (cm); R_{bR} = relative bias for total runoff; S_e/S_y = standard error ratio for total runoff; R_{bB} = relative bias for baseflow).

		Random			Flow-proportion		
Year	P	R_{bR}	S_e/S_y	R_{bB}	R_{bR}	S_e/S_y	R_{bB}
1992	117	0.15	1.10	-0.47	-0.15	0.51	-0.00
1993	131	-0.00	1.32	-0.19	-0.04	0.47	0.47
1994	85	-0.32	1.69	-0.62	-0.08	0.53	-0.00
1995	121	0.14	0.99	-0.58	0.03	0.45	0.27
1996	139	-0.00	2.03	-0.42	0.00	0.76	0.11
1997	102	0.22	1.86	-0.65	0.14	0.45	0.34
1998	58	-0.56	0.99	-0.82	-0.03	0.48	0.07
mean	108	-0.05	1.43	-0.54	-0.02	0.52	0.18

The accuracy of the predicted daily discharges was similar to the accuracy obtained when using the components of the first objective function as reflected in the values of the relative standard error ratios; however, a significant improvement was observed in the predictions of the baseflow. These results suggest that refinements in the components of the objective function are beneficial to the accuracy of the predictions and that despite the similar values in the nonsystematic error the second objective function is a better alternative for optimization.

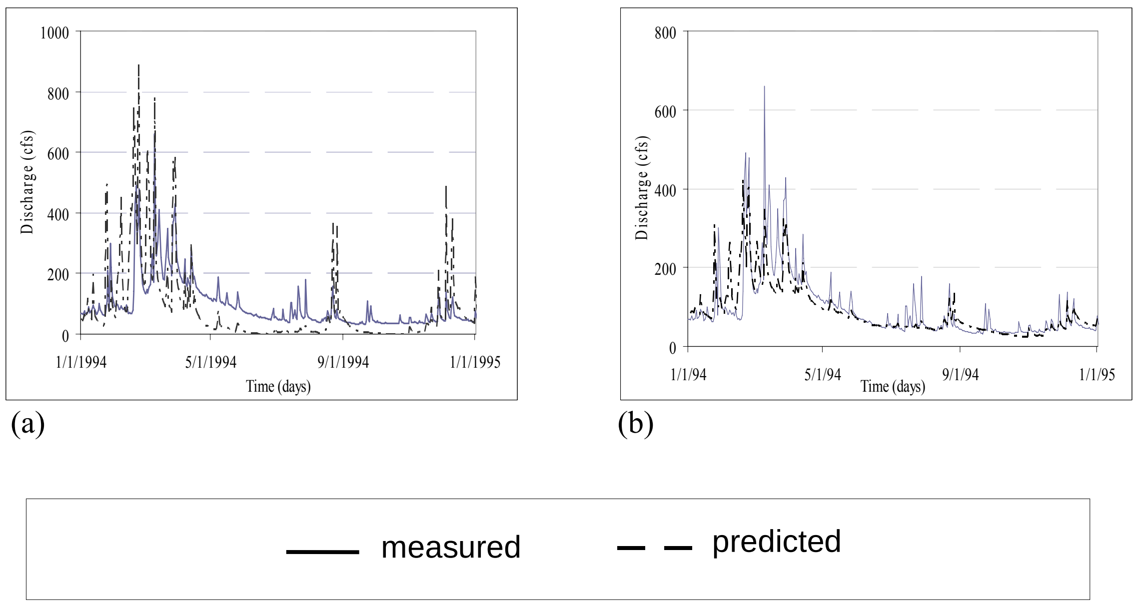


Figure 6-7. Little Falls watershed. Measured and predicted daily discharge during the calibration period using the second objective function (a) using arbitrary weights to the component contributions and (b) using the weights of Method 2.

CHAPTER 7

EFFECT OF CALIBRATING ANNUAL VS. MONTHLY PARAMETER VALUES

7.1 INTRODUCTION

The question of interest is: When the HSPF parameters are allowed to vary monthly, does the prediction accuracy improve? If prediction accuracy does not improve, then the HSPF option of monthly variation in the parameter values is not a useful mechanism for increasing accuracy. Knowledge of the variation of the monthly parameter can increase the efficiency and reliability of the hydrologic calibration and thus, the confidence in the accuracy of TMDLs. The goal of these calculations was to determine the effect of calibrating annual vs. monthly parameter values on the accuracy of the predicted discharges and on the rationality of the models. The analyses focused on comparisons of the predicted and measured daily discharges (cfs) from calibrations using annual parameter values and monthly parameter values.

7.1.1 Data and Calibration Criteria

Measured hourly rainfall data from the NOAA station 185934 and measured discharge data from the Little Falls at Blue Mount basin located in the Piedmont Physiographic Province of Maryland with a drainage area of 52.9 mi2 were used for the analysis. The period of record for Little Falls was between 04/01/1992 and 12/31/1998, and the total discharge depth for the period of analysis was 125 in. (318 cm), which is about 17.4 in./yr (44.3 cm/yr). The average baseflow during the calibration period was

75%, which was determined using the sliding-interval method. The 1997 land-use data produced by the Maryland Department of Planning (MDP, 1997) were used for the calibration. The land use data were grouped into three major categories: 40% of the area is covered by forest, 36% by agriculture, and 24% by pervious urban. For the analyses using annual parameters, the parameters that represent the hydrologic processes in these three land use categories were calibrated for a total of 27 parameters (9 per land use). Imperviousness was not included in the calibration as the impervious area in the watershed was less than 1%.

The calibration criteria were defined as a function of the flow proportions in the watershed and are discussed in section 3.4.5. The objective function includes the following components: (1) daily outflow, (2) monthly volumes, (3) autoregressive, (4) quick-flow filter, and (5) hydrograph separation. The relative bias and the relative standard error ratio were used as the measure of accuracy of the predicted daily discharge.

The goodness-of-fit statistics of the total predicted discharge, the 15 lowest independent baseflow values, and the 15 largest independent peak flow values were calculated for all of the analyses on the annual basis. The period of independence was set to 5 days. Because of precipitation data availability, the calculated goodness-of-fit for the 1992 does not includes the months of January, February, and March. The bias is:

$$\bar{e} = \frac{1}{N}\sum_{j=1}^{d}\left(\hat{Q}_j - Q_j\right) \tag{7-1}$$

and the standard error is:

$$S_e = \sqrt{\frac{1}{\nu}\sum_{j=1}^{d}(\hat{Q}_j - Q_j)^2} \quad \textbf{(7-2)}$$

where $\bar{e}$ is the mean bias of the model predicted daily runoff; ν is the degrees of freedom ($\nu = N - 1$); d is the number of days in the year; N is the total number of computed values; S_e is the standard error of the estimate between the measured and predicted daily discharge. The standardized bias for the annual calculations (R_b) is:

$$R_b = \frac{\bar{e}}{\overline{Q}} = \frac{\bar{e}}{\frac{1}{N}\sum_{j=1}^{d} Q_j} \quad \textbf{(7-3)}$$

where $\overline{Q}$ is the annual mean of the measured daily discharge. The relative standard error (R_e) is:

$$R_e = S_e / S_y \quad \textbf{(7-4)}$$

in which and S_y is the standard deviation of the measured daily discharge:

$$S_y = \left[\frac{1}{N-1}\sum_{j=1}^{365}(Q_j - \overline{Q})^2\right]^{0.5} \quad \textbf{(7-5)}$$

7.1.2 Method of Analyses

Two analyses were made. First, all parameters were calibrated as being constant throughout the year. Second, a combination of constant annual parameters and monthly parameters were used for calibration. For the second set of analyses AGWRC, INFILT, and LZSN were held constant throughout the year, but CEPS, NSUR, INTFW, IRC, and UZSN varied monthly.

The parameters included in the analyses are listed in Table 7-1. For all of the calibrations, some of the parameters were calibrated with constant annual values as indicated on Table 7-1, column 1; however, the constant parameters were replaced by their monthly equivalent for the calibration with varying parameters (see Table 7-1, column 2). Initial parameter values were selected within the recommended range in the HSPF manual and used in all the analyses. Overall, the initial values and the range of possible calibrated values were based on the physical meaning of the parameter.

Table 7-1. List of parameters calibrated annually and monthly; parameters varying monthly start with the letter V.

Annual	Monthly	Units	Parameter description
LZSN	LZSN	inches	Lower zone nominal storage
INFILT	INFILT	inches/hr	Index to the infiltration capacity of the soil
AGWRC	AGWRC	day^{-1}	Basic groundwater recession rate if KVARY is zero, and groundwater does not receives inflow
DEEPFR	DEEPFR	none	Fraction of groundwater inflow that will enter deep (inactive) groundwater, and, thus, be lost from the system
CEPS	VCSFG	inches	Interception storage capacity
NSUR	VNNFG	complex	Manning’s n for the assumed overland flow plane
INTFW	VIFWFG	none	Interflow inflow parameter
IRC	VIRCFG	day^{-1}	Interflow recession parameter
LZETP	VLEFG	none	Lower zone E-T parameter. It is an index to the density of deep- rooted vegetation
UZSN	VUZFG	inches	Upper zone nominal storage

The calibration of the parameter values should be bound to a range of feasible values that represent the hydrological processes. Upper and lower bound values and the initial value for the calibration are given in Table 7-2. For the calibrations where the parameters varied monthly a sinusoidal function was fitted to the data. In the case of the sinusoidal function the angle was bound between 1 and 360 degrees. The amplitude of the sinusoidal function (Table 7-3) was bound in such way that the calculated HSPF parameter were within the allowable HSPF range of values.

Table 7-2. Bounds and initial values by land use of the parameters that represent the hydrologic processes in the HSPF model

Bounds	LZSN (in.)	INFILT (in./hr)	AGWR (1/day)	UZSN (in.)	NSUR (complex)	INTFW	IRC (1/day)	LZETP	CEPS (in.)
lower	1.00	0.10	0.920	0.10	0.05	0.10	0.05	0.01	0.10
upper	30.00	0.50	0.999	9.99	0.30	5.00	0.90	0.30	9.99
Initial values for calibration									
forest	10.00	0.12	0.985	2.00	0.07	0.50	0.15	0.25	2.00
agricult.	8.00	0.10	0.975	2.00	0.09	0.50	0.10	0.20	2.00
urban	10.00	0.08	0.920	2.00	0.06	0.50	0.10	0.20	2.00

Table 7-3. Bounds of the sinusoidal function fitted to the HSPF parameters varying monthly

	UZSN	INTFW	LZETP	**CEPS (in.)**
	Amplitude (sinusoidal)			
Lower bound	0.1	0.1	0.01	0.1
Upper bound	2.5	3.0	0.3	1.8

Provided that PEST includes three methods to locate the optimum value of the parameters it was of interest to determine the optimum method of calibration, when using the model-independent-parameter estimator PEST in combination with the HSPF model. A first group of analyses using constant annual values were performed. For these analyses ten calibrations were made using the three fitting methods discussed in CHAPTER 2: (1) the Gauss-Marquardt-Lambda method (M-L), (2) the single value decomposition method (SVD), and (3) the single-value decomposition method-Assist (SVD-A). Within this group of analyses, three calibrations using the SVD method and 27 singular values were made, to test the effect of the truncation parameter “EIGTHRESH” on prediction accuracy. The truncation parameter EIGTHRESH is the ratio of lowest to largest eigenvalue at which, the truncation to determine the number of singular values used in the solution happens. The truncation parameter "EIGTHRESH" was varied to 1.E-4, 1.E-6, and 1.E-8. The optimum value of EIGHTHRESH was then used in the remaining seven calibrations: (1) Using the M-L method with 27 parameters, (2) using the SVD method with 27, 20, and 15 singular values, and (3) using the SVD-A method with 27, 20, and 15 singular values. When using the M-L method, the HSPF parameters are individually calibrated. When using the SVD-A method, PEST calculated the first-order-sensitivity of the HSPF parameters during the matrix-rank estimation iteration to determine the number of singular values to be used in the solution. At the end of the calibration, the process is reversed and the HSPF parameters as such, are computed.

For the second group of analyses (i.e., using a combination of constant annual parameters and parameters varying monthly), four calibrations were made using the M-L and the SVD-A methods: (1) M-L with 225 parameters, (2) SVD-A with 105 singular values, and (3) SVD-A with 54 singular values and fitting a sinusoidal function to the monthly HSPF parameters. It was expected that a seasonal trend could be observed in the parameters when fitting the sinusoidal function to the HSPF monthly parameters.

Calibrated parameter values from calibrations using constant annual parameters were used to set the variables in the sinusoidal function. For example, the optimum value of the UZSN annual parameter obtained in the SVD (20) was used as the initial value for the mean (lambda) in the sinusoidal function.

7.2 PREDICTION ACCURACY WHEN CALIBRATING ANNUAL PARAMETER VALUES

The use of annual parameter values may provide the necessary accuracy of predicted discharges and shorter times in the calibration process. The calibration of fewer parameters reduces the potential problems of intercorrelation among parameters and reduces the complexity of the calibration. However, it may be that the annual values do not reflect the seasonal changes in hydrologic processes such as the change in soil storages or the change in the rate of evaporation.

An issue related to the number of possible parameters to calibrate is that this number is not limited to the parameters that represent the hydrological processes in the HSPF model, as the number of parameters is multiplied by the number of simulated land uses. However, to reduce the complexity of the calibration in most cases, the calibration of parameter values for nondominant land uses are tied to the parameters of the dominant land uses. Although this practice may or may not provide sufficient prediction accuracy, it is supported by the lack of knowledge of specific parameter values for some of the simulated land uses.

7.2.1 Effect Threshold Value on Convergence

The value of the convergence threshold Eigthresh influenced the capability of PEST to reach the optimum. When using the SVD method with 27 singular values the results shown in Table 7-4 indicated that the accuracy of the predicted discharges was not significantly different when the Eigthresh value was set to 1.0E-4 or 1.0E-6. The relative biases were $R_{bR} = -0.014$ and $R_{bR} = -0.004$ for an Eigthresh value of 1.0E-4 and 1.0E-6, respectively; in contrast, the relative bias for an Eigthresh of 1.0E-8 was

$R_{bR} = -0.068$. These results suggest that the value of Eigthresh is a factor in the prediction accuracy and that an optimum needs to be determined for each specific set of data. The optimum value for the hydrological data used in this analysis was selected as 1.0E-6 because it provided the least biased model and the lowest value of the relative standard error ratio of the predicted discharges ($S_e/S_y = 0.532$).

Table 7-4. Statistical summary of HSPF predictions using the SVD method and 27 singular values for the calibration of annual parameters

Year	P	*Eigthresh*					
		1.0 E – 4		**1.0 E – 6**		**1.0 E -8**	
		R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
1992	117	-0.179	0.781	-0.127	0.751	-0.461	0.975
1993	131	0.006	0.349	-0.024	0.360	-0.127	0.413
1994	85	-0.059	0.625	-0.061	0.613	-0.111	0.603
1995	121	-0.020	0.425	0.005	0.446	-0.035	0.541
1996	139	0.028	0.567	0.049	0.559	0.043	0.546
1997	102	-0.035	0.332	-0.028	0.355	-0.059	0.377
1998	58	0.083	0.450	0.105	0.458	0.079	0.512
Average over calibration period	**108**	**-0.014**	**0.535**	**-0.004**	**0.532**	**-0.068**	**0.575**

7.2.2 Effect of the Calibration Method on the Objective Function

The final value of the objective function is important because it reflects the accuracy of the calibration and the quality of the parameters. However, this should not be the only criterion that is used as a measure of accuracy because the components of the objective function may not account for all of the hydrologic processes. Furthermore, the overall value of the objective function cannot be decomposed into the effects of the individual flow components; therefore other goodness-of-fit calculations are warranted.

The initial value of the objective functions for all of the calibrations using annual values was 1.47E5. The final values of the objective function were: 45352 using the M-L method; 47021, 46999, and 48676 using the SVD method with 27, 20, and 15 singular values, respectively; and 78689, 65721, and 62330 using the SVD-A method with 27, 20, and 15 singular values, respectively. Based on these values, the results suggest that the best calibrations were obtained with the M-L method, followed by the SVD method with 20 singular values. The least accurate results seem to be from the SVD-A method in which the number of singular values used in the solution are set by the user.

7.2.3 Accuracy of Predicted Total Runoff

The method of calibration was a factor in the prediction accuracy. The results indicated that the SVD method, in which the first order sensitivity of the parameters is calculated at all iterations and used to determine the number of singular values for the solution, provided the most accurate goodness-of-fit statistics (Table 7-5). The downside of the method in comparison to the SVD-A or the M-L methods, is the longer time to find the optimum parameter values. The accuracy of predicted total runoff was assessed annually (Table 7-5 and Table 7-6). The accuracy of the predicted daily discharge was calculated using Eqs. 5-1 through 5-5.

Regardless of the method or the number of singular values used for the solution, underprediction was observed in all calibrations. Although the least biased predictions where obtained with the SVD (27), with a $R_{bR} = -0.004$, this analysis had the third smallest final value of the objective function (47021); which suggests that additional criteria such as the accuracy of the predicted baseflow and peak flows will be necessary to determine which of the parameter fitting methods provided the best accuracy.

Table 7-5. Statistical summary of HSPF predictions during the calibration period using the M-L, SVD-A and SVD methods for the calibration of annual parameters (R_{bR} = relative bias for total discharge; S_e/S_y = standard error ratio for total discharge).

	M-L (27)		SVD-A (27)		SVD (27)	
Year	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
1992	-0.270	0.783	-0.177	0.753	-0.127	0.751
1993	-0.010	0.342	-0.076	0.378	-0.024	0.360
1994	-0.071	0.624	-0.092	0.610	-0.061	0.613
1995	-0.024	0.428	0.009	0.450	0.005	0.446
1996	0.039	0.527	0.033	0.566	0.049	0.559
1997	-0.010	0.335	-0.002	0.417	-0.028	0.355
1998	0.080	0.429	0.095	0.539	0.105	0.458
Average over calibration period	**-0.199**	**0.527**	**-0.174**	**0.554**	**-0.004**	**0.532**

Table 7-6. Statistical summary of HSPF predictions during the calibration period using the SVD-A and SVD methods for the calibration of annual parameters (R_{bR} = relative bias for total discharge; S_e/S_y = standard error ratio for total discharge).

	SVD-A (20)		SVD (20)		SVD-A (15)		SVD (15)	
Year	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
1992	-0.091	0.735	-0.274	0.763	-0.057	0.719	-0.172	0.681
1993	-0.198	0.438	-0.115	0.376	-0.147	0.431	-0.067	0.389
1994	-0.222	0.742	-0.138	0.625	-0.189	0.698	-0.075	0.610
1995	0.037	0.464	-0.089	0.473	0.053	0.495	0.011	0.417
1996	-0.029	0.574	-0.003	0.561	0.025	0.579	0.041	0.584
1997	-0.010	0.671	-0.051	0.342	-0.006	0.624	0.009	0.432
1998	0.121	0.546	0.033	0.418	0.1383	0.529	0.091	0.480
Average over calibration period	**-0.074**	**0.618**	**-0.084**	**0.527**	**-0.037**	**0.605**	**-0.016**	**0.556**

A joint analysis of the relative bias and the relative standard error ratio indicate that the most accurate predictions were provided by the SVD (27) with a relative bias of

$R_{bR} = -0.004$ and a relative standard error ratio. The SVD (20), the SVD (15), and the M-L (27) followed in accuracy. The least accurate predictions were provided by the SVD-A (20) method with $R_{bR} = -0.074$ and $S_e/S_y = 0.618$, followed by the SVD-A (15), and the SVD-A (27). While the M-L (27) method provided the smallest final objective function value, the prediction accuracy based on the relative bias and relative standard error ratio values was not significantly better than the values for the SVD-A (27). The accuracy of the M-L method however, was significantly better than the value for the SVD-A method with 20 and 15 components.

7.2.4 Accuracy of Predicted Baseflow and Peak Flow

As the overall goodness-of-fit statistics were inconclusive, an analysis of the distribution of the underpredicted and overpredicted flows in terms of baseflow and peak flows may be a useful tool for making a judgment. The relative bias was analyzed for the predictions of the 15 lowest independent baseflow values and the 15 largest independent peak flow values (Table 7-7). The average of the yearly values of the relative bias and the relative standard error were used to draw the conclusions about the best prediction accuracy of baseflow and peak flow. Overprediction of baseflow was observed in all of the methods except in the SVD (20), which had a slight underprediction (-0.038) and the SVD-A (27) with a significant underprediction (-0.320). The accuracy of the baseflow predictions was very poor for all of the calibrations using the SVD-A method (the average relative bias were -0.32, 0.291, and 0.269 for 27, 20, and 15 singular values, respectively. These results suggest that although the SVD-A method provided faster solutions, the accuracy of these solutions was not the best. When the SVD (20) method was used, the relative bias of the predicted daily runoff indicated significant underprediction ($R_{bR} = -0.084$); however, in terms of the baseflow and the peak flows, the SVD (20) had the most accurate predictions with values of $R_{bR} = -0.038$ and $R_{bR} = -0.189$, respectively.

Overprediction of baseflows and underprediction of peak flows are expected as any least squares calibration tends to provide mid-level predictions. The analysis of the sign of the biases of baseflow and peak flows suggest that the relative bias of the predicted total discharge (see Table 7-5 and Table 7-6) is actually an effect of the summation of the overpredicted baseflow and the underpredicted peak flow (Table 7-7 and Table 7- 8). For example, the least biased model using the total daily flow was the SVD (27) with a relative bias of $R_{bR} = -0.004$; however, the relative bias for the baseflow was $R_{bB} = 0.178$ and for the peak flow was $R_{bP} = -0.228$. In contrast, the SVD (20) was the third most biased model with a relative bias for the total daily flow of $R_{bR} = -0.084$, yet the relative bias for the baseflow was $R_{bB} = -0.038$ and for the peak flow was $R_{bP} = -0.189$.

Table 7-7. Relative biases (R_{bB}) using the 15 lowest daily flow values

	M-L	SVD			SVD-A		
Year	**27**	**27**	**20**	**15**	**27**	**20**	**15**
1992	-0.112	0.269	-0.270	0.065	-0.798	0.046	0.122
1993	0.422	0.461	0.123	0.363	-0.063	0.109	0.137
1994	0.008	0.025	-0.217	-0.052	-0.327	-0.043	-0.092
1995	0.135	0.246	0.045	0.291	-0.379	0.545	0.566
1996	0.234	0.218	0.025	0.156	-0.011	0.199	0.229
1997	0.090	0.013	0.082	0.254	-0.321	0.789	0.637
1998	0.032	0.016	-0.057	0.014	-0.338	0.394	0.282
Average over calibration period	**0.116**	**0.178**	**-0.038**	**0.156**	**-0.320**	**0.291**	**0.269**

Table 7- 8. Relative biases (R_{bP}) using the 15 largest daily flow values

	M-L	SVD			SVD-A		
Year	**27**	**27**	**20**	**15**	**27**	**20**	**15**
1992	-0.407	-0.479	-0.409	-0.340	-0.747	-0.326	-0.337
1993	-0.158	-0.173	-0.223	-0.256	-0.210	-0.183	-0.214
1994	-0.389	-0.396	-0.391	-0.397	-0.458	-0.484	-0.414
1995	-0.229	-0.247	-0.253	-0.199	-0.198	-0.200	-0.199
1996	-0.046	-0.021	0.038	-0.026	0.126	-0.038	-0.081
1997	-0.070	-0.111	-0.019	-0.015	-0.045	-0.188	-0.085
1998	-0.156	-0.171	-0.065	-0.126	-0.174	-0.188	-0.117
Average over calibration period	**-0.208**	**-0.228**	**-0.189**	**-0.194**	**-0.244**	**-0.230**	**-0.207**

The above analyses and comparisons show that for an small the number of parameters the method of calibration is an important factor in the accuracy of the predicted runoff. In addition, it is recommended to use the goodness-of-fit statistics of the baseflow and peak flow to evaluate the over or underprediction of the model, rather than to use the relative bias of the total runoff. The best accuracy was obtained with the

SVD method in which PEST internally and at all iterations determined the number of parameters to be used in the solution.

7.2.5 Rationality of the Calibrated Parameters

The rationality of the calibrated parameters is as important as the prediction accuracy reflected in the goodness-of-fit statistics. For all of the calibrations and for the predominant land use (forest), the parameter values of the important parameters were expected to calibrate to rational values. In addition, the final values of the important parameters were expected to be similar among the calibrations, but differences could occur because of variations in the method of calibration and the number of selected singular values for the solution.

The importance of the annual parameters was found to be a function of the hydrological characteristics of the watershed (section 5.2.2), i.e., parameters that represent processes of baseflow should be more important in basins with a predominant proportion of baseflow than in basins not dominated by baseflow. A hydrograph separation for the runoff during the period of calibration was performed for the runoff in the Little Falls basin. The results indicated that 75% of the total discharge in the Little Falls basin was baseflow. Thus, the parameter AGWRC was expected to be the most important parameter, followed by LZETP, LZSN, and INFILT.

It is important to note that the importance of the parameters in section 3-3 was determined using data for a hypothetical watershed and with a single forested land use. The land use proportions in the actual watershed were 40% forest, 36% agriculture, and 24% pervious urban. For the analyses using annual parameters, the parameters that represent the hydrologic processes in these three land use categories were calibrated for a total of 27 parameters (9 per land use). Therefore, it is expected that the proportion of land uses in the actual watershed will affect the importance of the parameters.

Table 7-9. **Final parameter values for the calibrations using single annual values in the forested land use using the M-L, SVD and SVD-A methods for calibration.**

Parameter	M-L	SVD			SVD-A		
	27	**27**	**20**	**15**	**27**	**20**	**15**
AGWRC (1/day)	0.982	0.977	0.985	0.977	0.988	0.920	0.920
LZETP (none)	0.300	0.300	0.300	0.050	0.300	0.300	0.300
LZSN (in.)	1.376	3.572	2.377	3.403	9.009	4.139	4.988
INFILT (in./hr)	0.024	0.260	0.500	0.260	0.182	0.500	0.500
UZSN (in.)	3.853	2.410	0.100	2.795	1.910	1.088	0.100
INTFW (none)	0.636	0.910	0.346	5.000	0.802	0.100	0.100
NSUR (complex)	0.300	0.298	0.050	0.300	0.050	0.300	0.300
IRC (1/day)	0.900	0.900	0.900	0.900	0.900	0.900	0.900
CEPS (in.)	0.100	0.117	1.895	0.209	2.460	0.100	0.100

The most accurate final parameter values were determined based on the goodness-of-fit statistics shown in Table 7-5 and Table 7-6. A comparison of the final parameter values for the forested land use (40% of the area) and among the four best calibrations (SVD (27), SVD (15), M-L (27) and SVD (20)) did not detect a significant difference for AGWRC, which was expected to be an important parameter (Table 7-9). The values of AGWRC in the least accurate calibrations (SVD-A (27, 20, and 15)) showed significant differences, which based on the goodness-of-fit statistics of these calibrations, suggest that the parameter never reached its optimum. Except for the SVD (15) method, the final value of the parameter LZETP was always at the upper bound of the feasible calibrating values. The range of values for LZSN was between 1.3 and 3.5 for the four best calibrations, and between 4.1 and 9.0 for the least accurate calibrations. INFILT calibrated to 0.26 in SVD (27) and SVD (15); however, it reached the maximum possible value in the SVD (20) calibration, which suggests problems in the calibration. The largest variation among the methods in the final parameter values was observed in Table 7-9 for the parameter LZSN.

Values of unimportant parameters (UZSN, INTFW, NSUR, IRC, and CEPS) may not approach their true values; however, the effect of unimportant parameters on the accuracy of the predicted runoff is generally small when compared to the effects of the important parameters. It was expected that the final values of unimportant parameters would be different among the calibrations, and that the deviation of the parameter value from the expected value would be larger than that for the more important parameters. The final value of the parameter IRC was always at the upper bound of the feasible calibrating values (0.90). A similar behavior was observed in the parameter CEPS with final parameter values at the lower bound for M-L (27) and SVD-A (20 and 15). Reasonable values of NSUR (0.05) for a forested area were only obtained in the SVD (20) and in the SVD-A (27); the NSUR values for the remaining calibrations were unrealistic and at the upper bound of the feasible calibrating value.

The final value of AGWRC and LZET for the remaining land uses (agriculture 35% and pervious urban 25% of the area), shown in Table 7-10 and Table 7-11 had significantly more variation when compared to the variation of the final parameter values obtained for the forested land use. The variation of LZS for the remaining land uses was between 1.3 and 10.8. The large variation among final parameter values is due to the unimportance of the parameters or to their inability to calibrate.

Table 7-10. **Final parameter values for the calibrations using single annual values in the agricultural land use using the M-L, SVD and SVD-A methods for calibration.**

	M-L	**SVD**			**SVD-A**		
Parameter	**27**	**27**	**20**	**15**	**27**	**20**	**15**
AGWRC (1/day)	0.982	0.993	0.967	0.991	0.990	0.999	0.999
LZETP (none)	0.300	0.168	0.054	0.209	0.300	0.300	0.300
LZSN (in.)	1.376	4.203	6.787	10.833	6.031	5.192	2.748
INFILT (in./hr)	0.024	0.100	0.100	0.111	0.100	0.100	0.100
UZSN (in.)	3.853	2.854	5.620	3.271	2.208	1.823	2.332
INTFW (none)	0.636	0.483	0.215	0.100	0.365	0.100	0.481
NSUR (complex)	0.300	0.149	0.300	0.300	0.300	0.300	0.300
IRC (1/day)	0.900	0.489	0.050	0.050	0.050	0.705	0.900
CEPS (in.)	0.100	0.766	0.100	0.658	1.887	0.100	0.100

Table 7-11. **Final parameter values for the calibrations using single annual values in the pervious urban land use using the M-L, SVD and SVD-A methods for calibration.**

	M-L	**SVD**			**SVD-A**		
Parameter	**27**	**27**	**20**	**15**	**27**	**20**	**15**
AGWRC (1/day)	0.982	0.986	0.999	0.989	0.982	0.999	0.999
LZETP (none)	0.300	0.300	0.300	0.300	0.010	0.300	0.135
LZSN (in.)	1.376	10.448	7.180	5.956	8.635	4.737	6.672
INFILT (in./hr)	0.024	0.050	0.078	0.051	0.123	0.050	0.050
UZSN (in.)	3.853	1.445	0.477	0.778	0.613	0.679	1.591
INTFW (none)	0.636	0.482	0.696	0.973	0.352	0.479	0.482
NSUR (complex)	0.300	0.300	0.300	0.069	0.040	0.300	0.300
IRC (1/day)	0.900	0.900	0.050	0.215	0.726	0.900	0.050
CEPS (in.)	0.100	0.100	0.100	1.442	1.389	0.100	0.100

7.3 PREDICTED RUNOFF ACCURACY WHEN CALIBRATING PARAMETERS VARYING MONTHLY

The seasonal changes in the soil and meteorological conditions may not be accurately reflected by temporally constant parameter values. To study this problem, the HSPF option of using monthly values of some parameters was tested. Increasing the number of parameters should enhance the amount of detail in the modeling process, and thus, increase the accuracy of the predicted runoff. However, problems of intercorrelation may be significant, which could dampen the benefit of adding parameters to the analysis When using the Marquardt-Lambda option in PEST, the matrix (JtQJ + λI) may be singular or near singular because of parameter intercorrelation (Doherty, 2001). As an alternative, calibrations using the Single Value Decomposition (SVD) of the Single Value Decomposition-Assist (SVD-A) methods in parallel, were compared to the results from the analysis of annual values using the SVD (20). The advantage of using the SVD and the SVD-A over the M-L method is to reduce the number of iterations by reducing the maximum number of singular values to be used in the solution. When using the SVD-A method, PEST calculates the first-order sensitivity of the HSPF parameters during the matrix-rank estimation iteration to determine the number of singular values to be used in all iterations. When using the SVD method, PEST calculates the first-order sensitivity of the parameters at all iterations and determines the number of singular values to be used per iteration. At the end of the calibration, the process is reversed and the HSPF parameters as such, are computed.

This second group of analyses used a combination of single annual and monthly values. The parameters AGWRC, INFILT, and LZS were always calibrated as annual values because this is the only option given by the HSPF model. The parameter NSUR (Manning's coefficient) was always calibrated using the single monthly values as a seasonal trend was not expected. Four calibrations were made for this analysis: (1) using the M-L method with 225 HSPF parameters to calibrate (6 parameters varying on the monthly basis and 3 parameters constant throughout the year for each of the three simulated land use categories), (2) using the SVD-A method and 105 singular values to calibrate the annual and monthly parameters, (3) using the SVD-A method and 57 singular values, and fitting a continuous sinusoidal function to the monthly parameters.

The purpose of fitting the continuous and discrete functions to the monthly parameter values was to provide a more rational parameter estimates through the emulation of the seasonality effect. It was uncertain as to how the upper and lower bounds imposed to the amplitude, the mean, and the angle of the sinusoidal function could affect the accuracy of the monthly parameter values. Bounds on the parameters of the sinusoidal function (Table 7-3) were imposed to limit the calculated HSPF monthly parameters to the allowable range of values in the HSPF model. The inability to define conditional parameters in PEST was also a factor. The monthly parameters were calculated using the following equation:

$$P_i = mean + A * \sin\left[((360/12) * i + angle_i)/57.2958\right] \qquad \textbf{(7-6)}$$

where P_i is the monthly parameter value for month i; $mean$ is the final parameter value determined in the SVD (20) calibration; A is the amplitude of the function; $angle$ is the calibrated angle for the sinusoidal function; and 57.2958 is the conversion factor from degrees to radians.

The sinusoidal function was not used for the calibration of the interflow recession rate represented by VIRCFG. The interflow recession rate was calibrated using the single annual value represented by the parameter IRC. The rationale for this decision was that, as the baseflow recession rate represented by AGWRC, the recession was not expected to have a significant variation from month to month and that it was a characteristic of the soil rather than a function of the hydrologic conditions.

7.3.1 Effect of calibration on the objective function value

The final value of the objective function provides information about the calibration process. A final objective function value that is significantly lower than the initial objective function value indicates that the accuracy of the predicted runoff was improved through the calibration of the parameter values; however, the accuracy of the predictions may still be poor. The initial parameter values were the same for the M-L (225) and the SVD (105) calibrations (Table 7-2), while the initial parameter values for the SVD2-A (57) were the final parameter values in the SVD (20); however, the weights for the calibration criteria were set so that the contribution of the components were the recommended in section 3.4.5.

The initial values for the objective functions were 1.67E5 for the M-L (225), 1.48E5 for the SVD-A (105), and 6.08E4 for the SVD-A (57). The final values of the objective function were 7.74E4 for the M-L (225), 3.51E4 for the SVD-A (105), and 3.75E4 for the SVD-A (57). These results indicate that the M-L (225) did not reach the optimum and that better accuracy of the predicted runoff was achieved by using annual values. Because the initial values of the objective function were different for the three calibrations, the accuracy of the predicted runoff was determined by the percent of which the objective function was reduced at the end of the calibration process. The results suggest that the SVD-A (105) with a 76% reduction provided the best accuracy, followed by the M-L (225) with a 54% reduction, and the SVD2-A (57) was least accurate with a 38% reduction. The most important difference among the calibrations was the computation time for the calibrations, mainly due to calculation reduction in the number of parameters and singular values used in the solution. While the time for the

SVD-A (57) and the SVD-A (105) was on the order of hours, the time for the M-L (225) was on the order of days.

7.3.2 Accuracy of predicted total runoff

The goodness-of-fit statistics are important when deciding on a method of calibration. The accuracy of the predictions in this case will be a function of the method and the number of parameters used in the solution. It was expected that, by increasing the number of parameters and by introducing a seasonality effect through the use of a sinusoidal function, the accuracy of the prediction would improve. The results (Table 7-12) indicate that for the M-L (225) and the SVD2-A (105) the accuracy of the calibrations was very poor with underpredictions of 15%. These values were significantly higher than those obtained from the best calibrations using annual values (Table 7-5 and Table 7-6) where the underprediction varied between -0.4% and -8%. For the SVD-A (57) using the sinusoidal functions, the accuracy of the predictions was slightly better ($R_{bR} = 0.006$) than the best calibration using annual values only SVD (20), where the relative bias was $R_{bR} = -0.084$. The values of the relative standard error ratio were higher for the M-L (225) and the SVD-A (105) than those resulting from calibrations using annual values (Table 7-5 and Table 7-6); this may be explained by the increase in the relative bias when calibrating monthly values. In contrast, the standard error ratio for the SVD2-A (57) and the SVD (20) was the same (0.525 and 0.527, respectively). None of the calibrations using monthly values or a sinusoidal function to reflect the seasonal changes in the parameter values improved the accuracy of the predictions.

Table 7-12. Statistical summary of the HSPF predictions using flow proportion weighting as the calibration criteria. (R_{bR} = relative bias for total discharge; S_e/S_y = standard error ratio for total discharge), and P = annual precipitation (cm).

		M-L (225)		SVD-A (105)		SVD-A (57)	
						Sinusoidal	
Year	P	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
1992	117	-0.629	1.119	-0.504	1.006	-0.231	0.758
1993	131	-0.198	0.405	-0.239	0.437	-0.019	0.370
1994	85	-0.166	0.616	-0.185	0.623	-0.056	0.619
1995	121	-0.141	0.470	-0.090	0.431	0.028	0.401
1996	139	-0.017	0.513	-0.091	0.409	0.020	0.534
1997	102	-0.063	0.322	0.040	0.330	-0.021	0.375
1998	58	-0.011	0.465	0.012	0.376	0.140	0.505
avg. over the period of record	**108**	**-0.155**	**0.533**	**-0.151**	**0.589**	**0.006**	**0.525**

The expectation was that the use of the monthly variation in the parameters could improve prediction accuracy as seasonality was observed in the measured data. In the case of the sinusoidal function the poor accuracy may be explained by the restriction imposed when setting the mean in equation (7-6) as the final value of the annual parameter found in the solution with SVD (20). At the very least, the monthly model should not perform more poorly than the annual model. In the case of the M-L (225) and the SVD2-A (105) not only the models yield less accurate predictions, but the models did not approach the accuracy attained with the models using annual parameters only. Causes for this can only be speculative, with possible reasons: (1) the error variation in the data is too great to separate out the effect of seasonal variation; (2) the objective function is insensitive to the monthly components; and (3) PEST cannot effectively calibrate with so many parameters. In the case of the model using sinusoids to reflect seasonal variation, the expectation of the worse case scenario was that the amplitude of the sinusoids could approach zero, which would leave the monthly variation constant and identical to the annual model. The magnitude of the relative bias value (0.006) as well as the relative standard error ratio (0.525) were almost identical to the magnitude of the relative bias of the annual SVD models (Table 7-12 and Table 7-5), suggesting that the data did not have a seasonal trend (Figure 7-1).

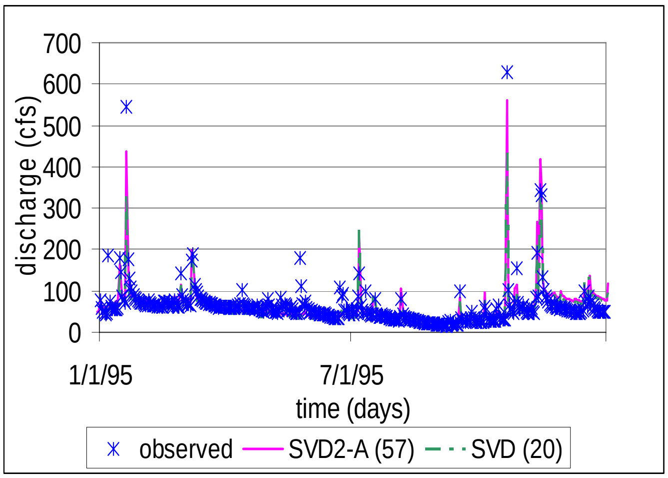


Figure 7-1. Measured and predicted runoff using the SVD-A (57) and the SVD (20) models

7.3.3 Accuracy of predicted baseflow and peak flow

When evaluating the accuracy of the predictions, it is important to use various criteria (i.e., the final value of the objective function, the goodness-of-fit statistics for a predefined number of the lowest and largest flows etc.). The purpose of using multiple criteria is to identify possibly misleading indications by the statistics of the total discharge because of the summation of underpredicted peak flows and overpredicted baseflows, or vice versa. The poor model performance reflected in the goodness-of-fit statistics of the total daily discharge suggests that the calibration of monthly values was not an important factor to improve the accuracy of the predictions; however, it was of interest to determine if the poor performance systematically affected the baseflow or the peak flow, or if instead, the added uncertainty is equally distributed between baseflow and peak flows. The accuracy of the predicted baseflow was poor for the models using monthly values or the sinusoidal function to derive the monthly values. The relative bias in these two models varied between -18% and 12%. In contrast, the relative bias for the baseflow in the SVD (20) was R_{bB} =-0.038, and for the SVD-A (57) was R_{bB} = 0.122. Similar results were observed with the accuracy of the peak flows where the relative bias for the SVD (20) was R_{bP} =-0.189 and for the SVD-A (57) was R_{bP} =-0.185.

Table 7-13. Relative biases using the 15 largest and 15 lowest predicted flow values. R_{bB} = relative bias for baseflow; R_{bP} = relative bias for peak flow; and P = annual precipitation (cm).

		M-L (225)		SVD-A (105)		SVD-A (57)	
						Sinusoidal	
Year	P	R_{bB}	R_{bP}	R_{bB}	R_{bP}	R_{bB}	R_{bP}
1992	117	-0.729	-0.675	-0.449	-0.660	-0.069	-0.449
1993	131	0.007	-0.269	-0.016	-0.328	0.509	-0.179
1994	85	-0.170	-0.435	-0.013	-0.466	0.000	-0.377
1995	121	-0.188	-0.288	0.220	-0.321	0.202	-0.143
1996	139	-0.024	0.011	0.002	-0.172	0.164	-0.025
1997	102	0.001	-0.118	0.451	-0.173	0.051	-0.074
1998	58	-0.172	-0.101	0.235	-0.239	-0.002	-0.049
avg. of annual values	**108**	**-0.182**	**-0.268**	**0.061**	**-0.337**	**0.122**	**-0.185**

It is not reasonable to expect that the calibration using the 57 singular values using monthly values could arrive to the same point in parameter space as the calibration using the 20 singular annual values because the SVD-A (57) is calibrating linear combinations of the HSPF parameters (singular values) and not the HSPF parameters as such. These results suggest that given the current limitations in PEST with the inability to set conditional bounds when fitting a sinusoidal or other function to the parameters varying monthly, the calibration of single annual values may provide the same accuracy and the time for calibration could be less. However, when strong seasonality is present, it is important to have a sensitivity analysis to determine what other factors such as the accurate prediction of snow in the HSPF, may affect the calibration process.

7.3.4 Importance of the parameters

It is known in calibration theory that (1) the value of the objective function reflects the values of the parameters and (2) the more important parameters, i.e., the more sensitive parameters, move to their true values more quickly than the less important parameters. The final values of the objective function for M-L (225), SVD-A (105), and SVD-A (57) are noticeably different (7.74E4, 3.51E4, and 3.75E4, respectively), which suggest quite different parameter values. However, the final values of the important parameters are expected to be similar in all calibrations regardless of the number of components.

Table 7-14. Final value of parameters calibrated as single annual values.

Parameter	FOREST			AGRICULTURE			PERVIOUS URBAN		
	M-L (225)	**SVD-A (105)**	**SVD-A (57)**	**M-L (225)**	**SVD-A (105)**	**SVD-A (57)**	**M-L (225)**	**SVD-A (105)**	**SVD-A (57)**
AGWRC	0.990	0.981	0.989	0.980	0.999	0.986	0.993	0.989	0.967
LZSN	7.410	4.928	3.572	6.820	4.899	5.398	9.499	7.210	3.140
INFILT	0.334	0.328	0.476	0.126	0.100	0.100	0.132	0.117	0.089
IRC	**monthly values**		0.300	**monthly**		0.300	**monthly values**		0.300

The final values for the most important parameter (AGWRC) in the predominant land use (Table 7-14) showed minimal variation in the number of days for the ground-water to recede; the recession rated varied between 50 days (AGWRC = 0.981) and 90 days (AGWRC = 0.990). These values were comparable to the value obtained in the SVD (20) with 0.985 (representing about 60 days of recession). However, based on the accuracy of the baseflow, judged by the relative bias value, indicated that, while the SVD-A (57) overpredicted the baseflow by 12%, the SVD (20) underpredicted by 3.8%, suggesting that a recession rate of 60 days is perhaps the most accurate value of AGWRC. For the agricultural and urban land uses the variation was significant; in the agricultural land AGWRC varied between 999 days (AGWRC= 0.999) and 50 days (AGWRC= 0.98), while in the pervious urban varied between 100 and 90 days.

The range of final values of LZSN was higher in the monthly models (Table 7-15) than in the annual SVD models (Table 7-6). The variation in the forested land use was significantly higher for the monthly models (3.6 - 7.4 inches) than the variation in the annual SVD models (2.4 – 3.6 inches). However, similar final values were obtained for the forested land use between the SVD-A (57) and the SVD (20) (3.57 vs. 2.4 inches, respectively) and the agricultural land use (5.4 vs. 6.8 inches, respectively). Values for the urban land use were significantly different between the SVD-A (57) and the SVD (20) (3.1 vs. 7.2 inches, respectively). The magnitude of the relative bias of the baseflow suggests that it is likely that the SVD (20) provided better values of LZSN (-0.038) than those obtained with the SVD-A (57) (0.122).

The final value of the parameter INFILT in the forested land use was similar for the M-L (225) and SVD-A (105) with 0.334 and 0.328 in./day, respectively, while the value in the SVD-A (57) was 0.476 in./day. The final values for the remaining land uses were similar and with an average value on 0.10 in/day. Given that the INFILT value in the SVD (20) was at the upper bound, the value of the SVD-A (57) is considered a better value.

Some of the HSPF parameters varied monthly, as opposed to being constant throughout the year. The more important parameters are likely to show consistent month-to-month trends, while the values of the less important parameters may show irrational trends. The interflow recession parameter VIRCFG was calibrated as monthly values in the M-L (225) and SVD2A- (105), but as single annual value (IRC) in the SVD2-A (57) where a sinusoidal function was fitted to determine the monthly values. The rationale for this decision was that the recession was not expected to be seasonal, thus, a sinusoidal function would not provide a reasonable representation of the recession process. The final values in the M-L (225) and the SVD2-A (105) (Table 7-15) showed a trend with the largest values during the winter months and the lowest values during the summer months, suggesting that the parameter is overall more important in winter months than in any other season, and that perhaps there is a seasonal behavior in the recession. This variation of slower interflow release during the winter months and faster interflow release during the summer months may be related to the type of rainfall during the winter and summer months, with more intense and shorter storms during the summer and more lengthy precipitation during the winter months. The annual values for all the land uses obtained in the SVD2-A (57) were 0.3 which represent an interflow recession rate of about 1 day. In contrast, the annual values in the SVD (20) (Table 7-4 through Table 7-6) were at the upper bound (0.9 for the forested land use) or at the lower bound (0.05 for the agricultural and urban land uses), which indicates problems in the SVD (20) calibration. An IRC value of 0.9 represents an interflow recession rate of about 14 days.

The annual parameter LZETP was calibrated with a single value in SVD-A (20), with the calibrated values being 0.3, 0.054, and 0.3 for the forested, agricultural and urban land uses, respectively. The parameter VLEFG represents the same processes as the LZETP but with month-to-month variation. The values of VLEFG obtained for the forested land use (673HTable 7-15) in the M-L – (225) varied between the lower and upper bounds with the lower values during the months of April-May and October-November; however, the random variation from month to month seems irrational. In contrast, the values for the forested land use from the SVD-A (105) were at the upper bound throughout the year, except for the months of January through March with a constant value of 0.25. The final values obtained with the sinusoidal model indicated that the amplitude of the curve was close to zero to produce values between 0.29 and 0.31 throughout the year. The results from the SVD-A (105) and the SVD-A (57) suggest that the monthly variation is not an important process in the calibration.

Table 7-15. Final value of annual and monthly varying parameters for the predominant (forest) land use

Month	VIFWF			VIRCFG			VLEFG		
	M-L (225)	SVD2-A (105)	SVD2-A (57)	M-L (225)	SVD2-A (105)	SVD2-A (57)	M-L (225)	SVD2-A (105)	SVD2-A (57)
Jan	0.87	1.71	0.31	0.42	0.90		0.25	0.25	0.31
Feb	0.53	0.10	0.31	0.17	0.90		0.25	0.25	0.31
Mar	0.51	0.10	0.31	0.49	0.90		0.25	0.25	0.31
Apr	0.21	0.10	0.31	0.10	0.90		0.03	0.30	0.31
May	0.36	0.38	0.30	0.44	0.28		0.02	0.30	0.30
Jun	0.32	1.10	0.30	0.54	0.05	0.3	0.28	0.30	0.30
Jul	0.40	0.10	0.29	0.43	0.38		0.11	0.30	0.30
Aug	0.51	010	0.29	0.29	0.76		0.30	0.30	0.29
Sep	0.29	0.30	0.29	0.22	0.34		0.18	0.30	0.29
Oct	0.43	0.10	0.29	0.45	0.58		0.01	0.30	0.29
Nov	0.18	0.10	0.30	0.20	0.90		0.03	0.30	0.30
Dec	0.87	0.10	0.30	0.31	0.90		0.25	0.30	0.30

VIFWF is the vector of monthly parameter values for the interflow inflow. The random variation of the month-to-month values obtained in the M-L (225) and the SVD-A (105) (Table 7-15) suggests that the monthly variation is important; however, the random variation indicates that the calibration process needs to be modified in order to obtain rational monthly values. The range of values in the SVD-A (57) (0.29 – 0.31) for the forested land use were similar to the value in the SVD (20) (0.346) calibration, suggesting that the sinusoidal curve is not the best option to reflect the seasonal variation of the parameter.

Table 7-16. Final value of monthly varying parameters for the predominant (forest) land use

Month	VCSFG			VNNFG			VUZFG		
	M-L (225)	SVD2-A (105)	SVD2-A (57)	M-L (225)	SVD2-A (105)	SVD2-A (57)	M-L (225)	SVD2-A (105)	SVD2-A (57)
Jan	2.91	2.37	1.90	0.05	0.05	0.30	0.98	0.55	2.51
Feb	2.99	3.35	1.91	0.05	0.05	0.05	1.43	1.93	2.51
Mar	2.25	2.98	1.91	0.30	0.05	0.30	1.24	0.62	2.51
Apr	1.37	0.10	1.91	0.22	0.05	0.05	1.83	1.73	2.51
May	1.72	0.10	1.90	0.05	0.05	0.05	1.17	3.78	2.50
Jun	1.68	4.61	1.90	0.08	0.05	0.30	2.75	2.47	2.50
Jul	1.82	1.19	1.89	0.16	0.30	0.05	1.86	2.53	2.49
Aug	1.68	1.70	1.89	0.14	0.04	0.04	1.98	3.12	2.49
Sep	1.34	0.21	1.89	0.21	0.04	0.04	1.82	1.26	2.49
Oct	1.90	1.84	1.89	0.30	0.30	0.30	1.93	2.89	2.49
Nov	1.97	1.35	1.90	0.30	0.05	0.30	2.18	1.94	2.50
Dec	2.25	1.15	1.90	0.30	0.05	0.05	1.46	0.47	2.50

The parameters CEPS and VCSFG represent the interception storage capacity on grass blades, leaves, branches, trunks, and stems of vegetation. The monthly values for the parameter VCSFG (Table 7-16) showed an irrational trend in the M-L (225), with lower values during the summer months and higher values during the winter months. Although the largest value in the SVD-A (105) was obtained in the month of June (4.61), the irrational variation throughout the year made these values inadequate as final values. In the SVD-A (225) the final value was the same as the assigned initial value, suggesting either insensitivity of the predicted discharge to the monthly variation of VCSFG or high intercorrelation with other parameters. The annual values in the SVD (20) (Table 7-4 through Table 7-6) were 1.89, 0.10 and 0.10 for the forested, agricultural and urban land uses, respectively.

The Manning's n for the assumed overland flow plane is represented by the parameters NSUR on the annual basis and VNNF on the monthly basis. In the analysis of the monthly values, it was expected that the lowest values for the forested land use would occur during the winter months when the under storage vegetation had decayed. The largest values would be expected during the late Summer and Fall seasons when the vegetation was most dense and leaves had fallen from the trees. The monthly values (Table 7-16) in the M-L (225) were rational following the expected trend. The final values in the SVD-A (105) were either at the lower or upper bound suggesting problems in the calibration. In the case of the SVD-A (57) were the monthly values were calibrated rather than fitting a sinusoidal curve, the month to month variation was irrational and from the two extremes of possible values suggesting also problems of high intercorrelation with other parameters.

VUZFG is the vector of monthly parameter values for upper zone storage. The month-to-month variation for the forested land use in the M-L (225) and the SVD-A (105) did not show a trend, but the variation was rational in comparison to the annual value obtained in the SVD (20) of 0.10, suggesting that the month-to-month scatter is important. The monthly values of the parameters in the agricultural (35%) and urban (25%) land uses generally did not show a significant trend, rather random variation from month-to-month in the M-L (225) and in the SVD (105) methods.

The importance of the parameters reflects the importance of the hydrologic process that the parameter represents, the variation of the processes throughout the year, the correct mathematical representation of the processes in the model, and to some extent, the land use characteristics in the watershed. The results indicate that mathematically, the accuracy does not necessarily improve as the number of singular values increases, and although the predictions of the SVD (57) and the SVD (20) had a similar accuracy, the final parameter values in the SVD (57) were generally more rational than the obtained as final parameter values in the SVD (20), suggesting that seasonality is important for the rationality of the model.

Perhaps a more rational formulation to calibrate the parameters varying monthly, not discussed in this analysis, could be the use of a discrete mass distribution function such as Poisson, which provides a single value per month. The issue with the sinusoidal function is that a value is calculated for all days of the year; yet, HSPF uses a single parameter value per month.

CHAPTER 8

EFFECT OF THE INITIAL STORAGE VALUES ON PREDICTION ACCURACY

8.1 INTRODUCTION

The estimates of the initial upper zone storage (UZS) and the initial lower-zone storage (LZS) influence the hydrologic performance of the HSPF model. Error in the initial storage values will be particularly important when the period of record used for the model calibration and TMDL development are on the order of one or two years because the prediction accuracy during the initial period of record may be adversely affected by error in estimates of the initial storages. Although the true initial storage conditions are unknown, the prediction accuracy will likely become poorer as the estimates of the initial storages deviate from their true values. The goal of these analyses was to assess the effects of erroneous initial estimates of the upper and lower-zone storages on the accuracy of the predicted discharges. The following analyses were used to assess this goal: (1) determine if the accuracy of the predicted discharges is affected by the proportions of flow in the watershed when the estimates of the initial storages are in error; (2) determine the effect of the initial storage estimates on the convergence time of the predicted discharges; (3) assess the effect of error in the initial storage estimates on the calibrated parameters and the capability to reach an optimum; and (4) evaluate the effect of rainfall conditions during the start-up period on the accuracy of the predicted discharges.

8.1.1 Data and Method of Analyses

Eight years of measured hourly rainfall data (01/01/1992 through 12/31/1999) and three hypothetical watersheds with a drainage area of 5 mi2. (Table 8-1) were used to evaluate the objectives. Although measurements of storages are never available on actual watersheds, the use of hypothetical data enables the evaluation of the estimates of the initial soil storages. The HSPF generated daily discharge using the hypothetical watersheds and assumed true discharge is referred to as the measured discharge. In order to reduce the effect of drainage area all discharges are reported in inches. The generated daily average discharges were 0.0432 in. for watershed 1, 0.036 in. for watershed 2, and 0.0414 in. for watershed 3.

Table 8-1. Flow proportion, as percentages for each of the hypothetical watersheds

Watershed	Baseflow	Interflow	Surface runoff	Quick flow
1	81	7	12	19
2	59	27	14	41
3	45	40	15	55

The following general procedure was used in all the analyses: First, values of the parameters that control the hydrologic processes in the watershed were selected for the analyses. When the analysis included calibration the selection included the lower and upper bound as well as the initial value of the parameter. These values were based on the physical meaning of the parameters and were set to the same values for all of the calibrations. The assumed true values of the state variables LZS and UZS in the tested watersheds were as follows: for watersheds 1 and 2, LZS was 5 in. and UZS was 2 in.; for watershed 3, LZS and UZS were both 3 in. Second, the estimate of each storage value (UZS and LZS) was varied systematically by ± 25, ± 50, ± 75, and ± 100% from its true value. Third, the relative bias (R_b) was used to assess the accuracy of the predicted flow components and the total predicted discharge rates using the following equation:

$$R_b = (Q_p - Q_a)/Q_a \quad \textbf{(8-1)}$$

where R_b is the relative bias, which is a measure of the systematic error variation between the actual and predicted discharges; Q_a is the actual discharge (in.); and Q_p is the predicted discharge (in.). The comparison of results across watersheds is possible by using the relative bias; in this calculation the effect of the initial storage on the predicted discharges is removed.

For the analysis to determine the time at which the predicted discharges are insensitive to the estimates of initial soil storages, referred to as the start-up period, a criterion associated with a significant discharge was selected. The start-up period is also known as the "model spin-up". Based on the notion that standard errors in regression models are generally between 20% and 40% of the true daily discharge, a 30% or greater difference between the actual and predicted discharges was selected as the criterion:

$$e_b \geq 0.3 * Q_a \quad \textbf{(8-2)}$$

where e_b is error bound and is equal to the difference between the predicted and actual discharges $(Q_p - Q_a)$, Q_a is the actual discharge (in.), and Q_p is the predicted discharge (in.).

For the analysis to assess whether or not the start-up periods associated with the predicted storages are similar to the start-up periods associated with predicted discharges, the relative bias of the predicted water holding capacity of the soil storage was calculated using the following equation:

$$R_e \leq (S_p - S_a)/S_a \tag{8-3}$$

where S_a is the actual soil storage (in.) and S_p is the predicted soil storage (in.).

In this case, criteria of 0.1 in. and 0.3 in. were selected to define the start-up period, time at which the predicted soil storage, referred to as the nominal storages, is not longer affected by inaccuracies in the initial estimates of the soil storages.

8.2 EFFECT OF INITIAL SOIL STORAGES ESTIMATES ON THE ACCURACY OF THE PREDICTED RUNOFF

Previous analyses of discharge sensitivity (Section 5.2) indicated that the sensitivity of the predicted discharges was a function of the flow proportions in the watershed. Therefore, it was of interest to determine if the estimates of the initial storages and the proportions of flow were a factor in the accuracy of the predicted discharges. The period over which the error in initial storages influences the computed discharges was also of importance. The relative bias was plotted vs. time at the following days: 30, 60, 90,120,150,180, 210, 270, and 300. Using this information, the time required for the predicted and actual discharges to converge were established. It was expected that the error between the predicted and actual discharges would decrease through time. Overprediction was expected when the estimates of the amount of water in the storage reflected by the values of UZS and LZS were overestimated, while underprediction was expected when UZS and LZS were underestimated.

In the analysis to assess the effect of UZS and LZS deviations on the accuracy of the predicted flow components the goodness-of-fit statistics of the total discharge, the baseflow and the peak flows were examined. The relative bias of the total discharge was calculated using Eq. (8-1). The systematic error of the baseflows and peak flows was based on the 20 lowest and the 20 largest predictions, respectively.

8.2.1 Effect of Flow Proportions and Deviations in LZS on Convergence

The effect of the error in the estimates of LZS was observed in the predicted discharges for all of the hypothetical watersheds, with large over and underpredictions at the start of the calibration and with lower error in the predictions as time passed. The largest underprediction of discharge for all of the watersheds and when LZS was underestimated was calculated for the predictions at day 60; these results suggest that before the convergence process between the actual and predicted discharges is initiated, a time delay of about 60 days is needed for the error in LZS to propagate to the system. The relative biases of the predicted discharges are given in Table 8-2 through Table 8-5

Table 8-2. Relative bias of the predicted discharges for deviations of -25 and +25 % in LZS

Watershed	% deviation LZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-25	-0.18	-0.19	-0.10	-0.09	-0.09	-0.08	-0.12	-0.15	-0.11	-0.08
2		-0.13	-0.23	-0.12	-0.10	-0.11	-0.10	-0.13	-0.18	-0.12	-0.11
3		-0.09	-0.10	-0.05	-0.04	-0.06	-0.03	-0.04	-0.06	-0.07	-0.06
1	+25	0.13	0.12	0.06	0.05	0.07	0.07	0.10	0.17	0.15	0.09
2		0.06	0.10	0.03	0.01	0.09	0.04	0.13	0.25	0.18	0.14
3		0.05	0.05	-0.01	-0.03	0.06	-0.03	0.00	0.04	0.10	0.09

Table 8-3. Relative bias of the predicted discharges for deviations of -50 and +50 % in LZS

Watershed	% deviation LZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-50	-0.29	-0.40	-0.24	-0.21	-0.19	-0.18	-0.23	-0.28	-0.19	-0.14
2		-0.21	-0.41	-0.32	-0.26	-0.25	-0.23	-0.27	-0.36	-0.22	-0.21
3		-0.16	-0.23	-0.15	-0.11	-0.14	-0.09	-0.11	-0.14	-0.12	-0.12
1	+50	0.21	0.18	0.07	0.06	0.13	0.10	0.17	0.31	0.31	0.17
2		0.09	0.14	-0.01	-0.06	0.16	0.02	0.17	0.42	0.39	0.23
3		0.07	0.08	-0.08	-0.14	0.10	-0.13	-0.05	0.06	0.22	0.19

Table 8-4. Relative bias of the predicted discharges for deviations of -75 and +75 % in LZS

Watershed	% deviation LZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-75	-0.36	-0.57	-0.36	-0.33	-0.30	-0.28	-0.34	-0.41	-0.27	-0.22
2		-0.26	-0.54	-0.51	-0.45	-0.42	-0.42	-0.44	-0.55	-0.31	-0.30
3		-0.20	-0.33	-0.26	-0.21	-0.21	-0.17	-0.19	-0.23	-0.17	-0.17
1	+75	0.23	0.18	0.04	0.02	0.16	0.09	0.20	0.42	0.46	0.21
2		0.09	0.13	-0.08	-0.17	0.21	-0.05	0.16	0.51	0.57	0.18
3		0.06	0.07	-0.15	-0.27	0.14	-0.25	-0.12	0.06	0.35	0.21

Table 8-5. Relative bias of the predicted discharges for deviations of -100 and +100 % in LZS

Watershed	% deviation LZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-100	-0.40	-0.69	-0.48	-0.44	-0.41	-0.39	-0.46	-0.54	-0.35	-0.29
2		-0.29	-0.63	-0.65	-0.61	-0.58	-0.63	-0.60	-0.73	-0.39	-0.39
3		-0.24	-0.42	-0.37	-0.31	-0.30	-0.28	-0.27	-0.32	-0.22	-0.22
1	+100	0.24	0.15	-0.00	-0.04	0.17	0.05	0.17	0.46	0.60	0.14
2		0.07	0.09	-0.15	-0.26	0.24	-0.13	0.10	0.50	0.74	-0.04
3		0.05	0.04	-0.20	-0.35	0.18	-0.19	-0.14	0.03	0.45	0.08

When comparing the results for a particular watershed, watershed 1 for example, the underestimation of LZS at day 60, produced underpredictions reflected in the values of the relative bias of -0.19, -0.40, -0.57, and -0.69 for errors in LZS of -25%, -50%, -75%, and -100%, respectively. By day 300, the errors had been reduced to -0.08, -0.14, -0.22, and -0.29, respectively. These results indicated that the magnitude of the relative bias increases as the deviation in the LZS increases and that the relative bias decreases with time, regardless of the flow proportions in the watershed. The effect of deviations in LZS was relatively larger for watersheds with predominant baseflow (watersheds 1 and 2) when compared to watersheds with predominant quickflow (watershed 3); this effect was observed in the magnitude and the slope of the relative bias shown by the fitted trend lines in Figure 8-1.

When interpreting a trend fitted to the calculated relative bias, the effect of low values in the denominator of Eq. 8-1 must be taken into consideration as they may influence the direction and slope of the fitted trend. The effect of low runoff on the relative bias was observed on days 60 (03/01/92) and 240 (08/30/92) when low discharges in the denominator produced high relative bias.

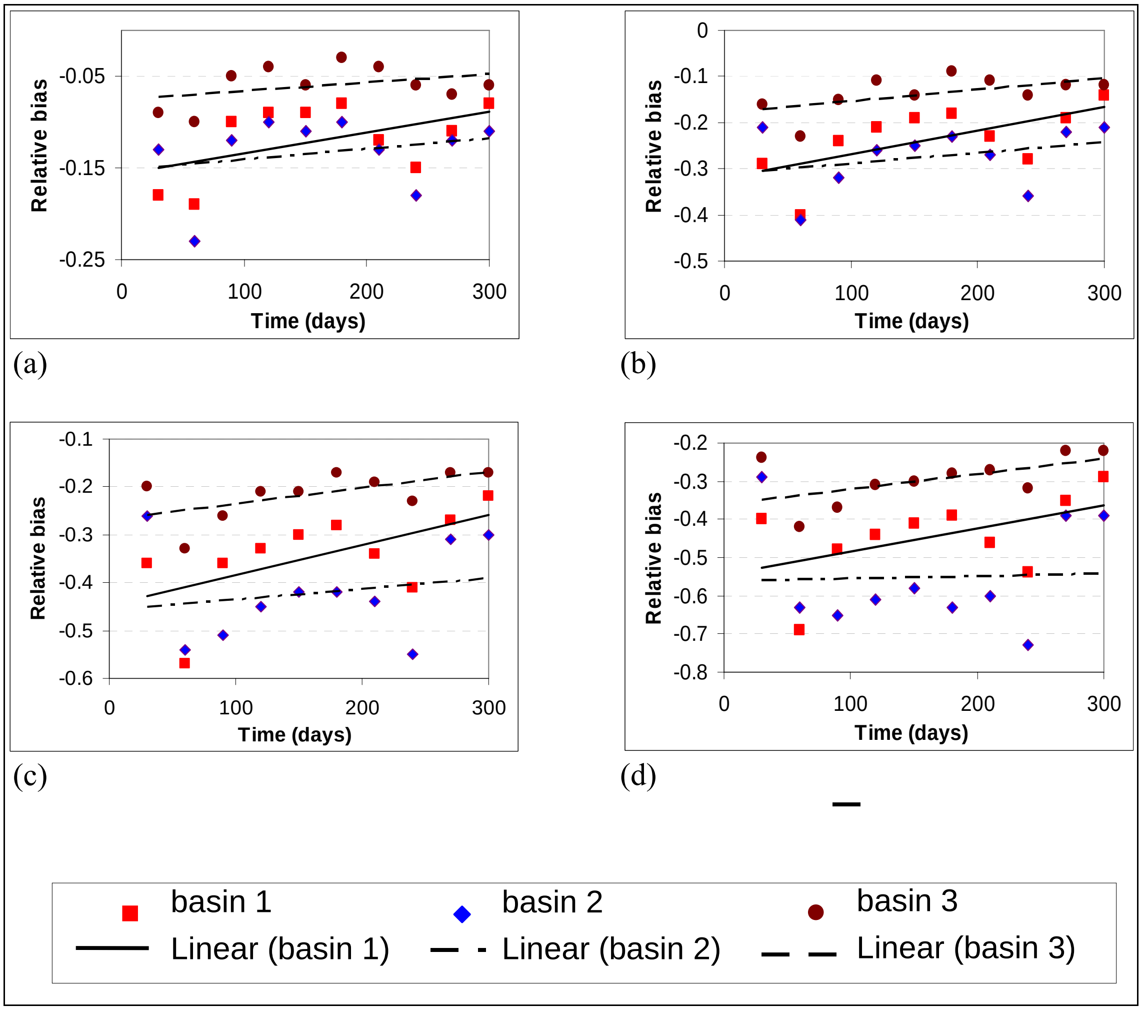


Figure 8-1. **Error in the predicted discharge caused by a negative deviation in LZS. (a)=25%; (b)=50%; (c)=75%; and (d)=100%**

When overestimating LZS the direction and slope of the linear trend was also influenced by very low monthly precipitation and thus, very low discharges. This outcome is depicted in Figure 8-2 in which low discharges in the denominator of the relative bias for days 240 and 270 significantly affects the direction and slope of the fitted trend, diverging rather than converging to zero at day 300. The monthly precipitation depths in both cases were about 2.2 in. (5.5 cm). However, if the data points from days 240 and 270 are ignored so that the effect of low precipitation is removed from the figure the linear trend converges to zero and reflects only the error in the initial LZS value.

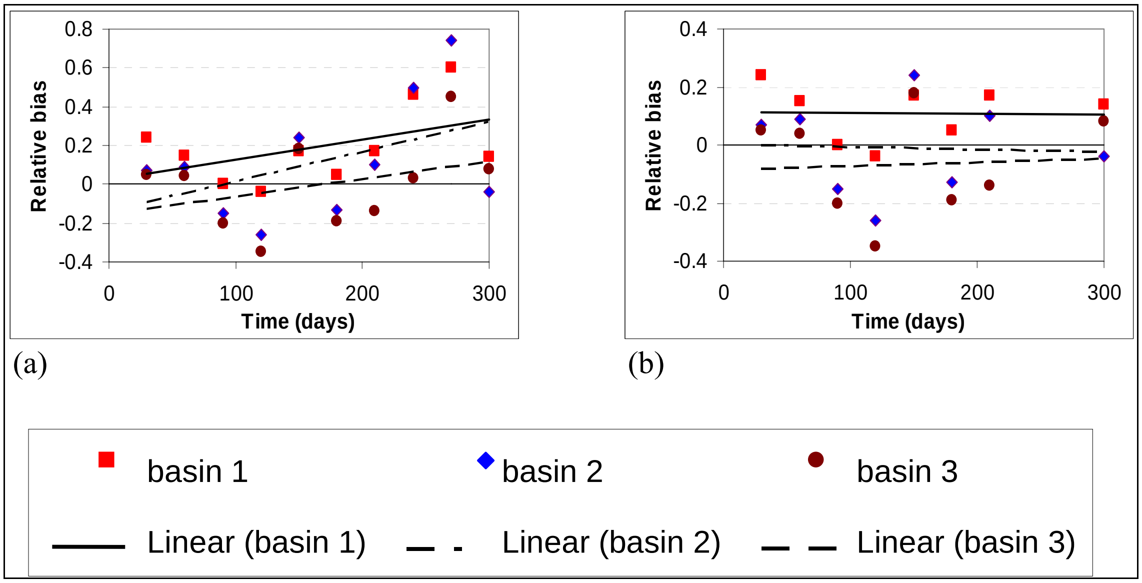


Figure 8-2. **Relative bias and modified linear trend in the predicted discharge caused by a + 100% deviation. (a) data from 300 days and (b) data from days 240 and 270 were removed to fit the linear trend so that the effect of low precipitation was removed.**

Although the effect of underestimating LZS was always underpredicted discharges (Figure 8-1), the effect of overestimating LZS was a fluctuation between over and underpredictions for all of the watersheds (Figure 8-2). This fluctuation in the relative bias is possibly caused by a seasonality effect, which may have an important implication to the proper calibration of HSPF.

8.2.2 Effect of Flow Proportions and Deviations in UZS on Convergence

Erroneous estimates of the upper zone storage (UZS) adversely affected the accuracy of the predicted discharges. Underprediction was always observed for underestimations of UZS and overpredicted discharges were always observed for overestimations of UZS. However, the magnitude of the error was lower than the error caused by erroneous estimates of LZS. As with LZS, a time delay for the estimates of UZS to affect the predicted discharges was also observed. At day 90, the largest error in the predicted discharge was computed when underestimating UZS (Table 8-6 through Table 8-9), and at day 60 when overestimating UZS. An alternative explanation for this phenomenon may lie in the dates when the application starts and thus, with possible frozen ground conditions which will cause low infiltration during the months of January and February. Consequently, to determine if flow proportions would have an effect on the trend of the predicted discharges when the initial UZS deviated from its true value, the calculated ratios for days 30 and 60 were not included in the analysis (Figure 8-3). Regressions were fitted to reduce the effect of the errors due to seasonality, thus making the effect of the flow proportions more noticeable.

Table 8-6. Relative bias of the predicted discharges for deviations of -25% and +25 % in UZS

Watershed	% deviation UZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-25	-0.00	-0.00	-0.06	-0.07	-0.05	-0.04	-0.04	-0.05	-0.03	-0.02
2		-0.00	-0.01	-0.09	-0.09	-0.06	-0.05	-0.04	-0.06	-0.03	-0.03
3		-0.00	-0.04	-0.12	-0.11	-0.08	-0.08	-0.07	-0.07	-0.04	-0.04
1	+25	0.03	0.02	0.08	0.07	0.05	0.04	0.03	0.04	0.02	0.02
2		0.00	0.03	0.10	0.08	0.05	0.05	0.04	0.06	0.03	0.03
3		0.04	0.11	0.12	0.09	0.07	0.07	0.05	0.06	0.04	0.04

Table 8-7. **Relative bias of the predicted discharges for deviations of -50% and +50 % in UZS**

Watershed	% deviation UZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-50	-0.00	-0.01	-0.08	-0.11	-0.09	-0.07	-0.09	-0.11	-0.07	-0.05
2		-0.00	-0.02	-0.16	-0.17	-0.13	-0.11	-0.09	-0.12	-0.07	-0.06
3		-0.00	-0.05	-0.21	-0.22	-0.17	-0.17	-0.15	-0.16	-0.08	-0.08
1	+50	0.16	0.10	0.13	0.11	0.08	0.06	0.05	0.07	0.05	0.03
2		0.03	0.10	0.20	0.14	0.10	0.08	0.08	0.10	0.06	0.05
3		0.17	0.29	0.23	0.16	0.13	0.12	0.09	0.11	0.07	0.07

Table 8-8. **Relative bias of the predicted discharges for deviations of -75% and +75 % in UZS**

Watershed	% deviation UZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-75	-0.00	-0.01	-0.09	-0.12	-0.11	-0.10	-0.15	-0.18	-0.12	-0.09
2		-0.00	-0.02	-0.21	-0.25	-0.20	-0.16	-0.15	-0.19	-0.10	-0.09
3		-0.00	-0.06	-0.26	-0.31	-0.26	-0.26	-0.24	-0.26	-0.14	-0.13
1	+75	0.32	0.17	0.17	0.14	0.10	0.08	0.08	0.11	0.08	0.05
2		0.14	0.22	0.27	0.19	0.13	0.11	0.11	0.15	0.09	0.07
3		0.39	0.49	0.31	0.22	0.18	0.15	0.13	0.15	0.10	0.10

Table 8-9. **Relative bias of the predicted discharges for deviations of -100% and +100 % in UZS**

Watershed	% deviation UZS	Number of days within the period of record									
		30	60	90	120	150	180	210	240	270	300
1	-100	-0.00	-0.01	-0.09	-0.13	-0.13	-0.12	-0.20	-0.24	-0.17	-0.13
2		-0.00	-0.02	-0.24	-0.30	-0.25	-0.21	-0.21	-0.29	-0.13	-0.13
3		-0.00	-0.06	-0.30	-0.38	-0.35	-0.33	-0.31	-0.39	-0.17	-0.20
1	+100	0.50	0.24	0.20	0.17	0.12	0.09	0.09	0.13	0.10	0.06
2		0.29	0.34	0.33	0.23	0.17	0.13	0.13	0.21	0.09	0.10
3		0.64	0.68	0.39	0.27	0.22	0.19	0.16	0.21	0.11	0.13

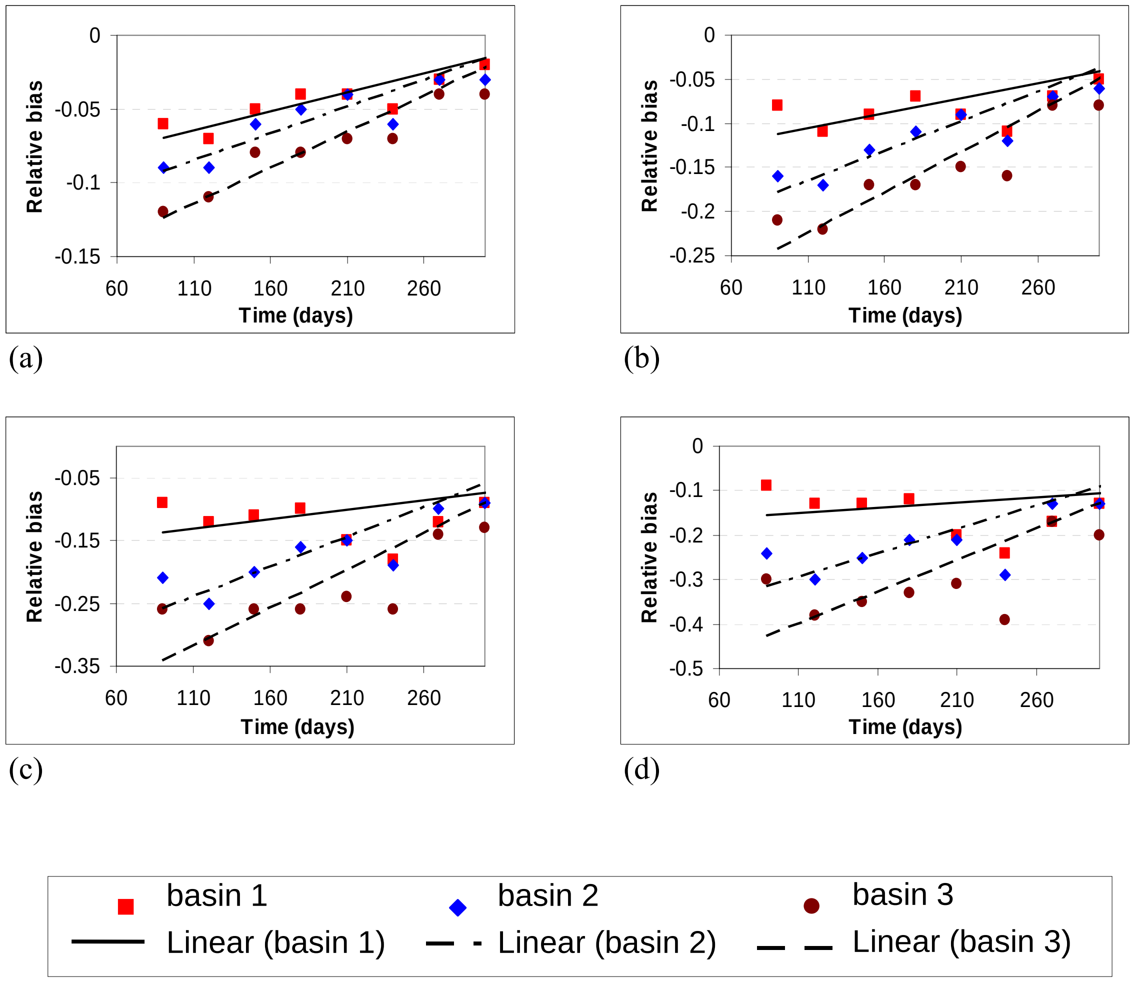


Figure 8-3. Error in the predicted discharge caused by a negative deviation in UZS. (a)=25%; (b)=50%; (c)=75%; and (d)=100%

The flow proportions influenced the accuracy of the predicted discharges; the largest over and underpredictions were observed in the watershed with predominant quickflow (55% for watershed 3). For example when overestimating UZS by 100% and at day 90, the relative bias for watershed 3 is 0.39 while the relative bias for watershed 1 is 0.20 (Table 8-9). An additional observation is related to the reduction of the error through time. By day 300, the relative bias in all watersheds was of similar value

(Figure 8-3 and Figure 8-4). These results again, show the convergence of the predicted discharges towards the actual values as time passes. As with LZS, the relative error in the predicted discharges increased as the deviation in UZS increased, but over time, the fitted trend to the predictions converged to zero.

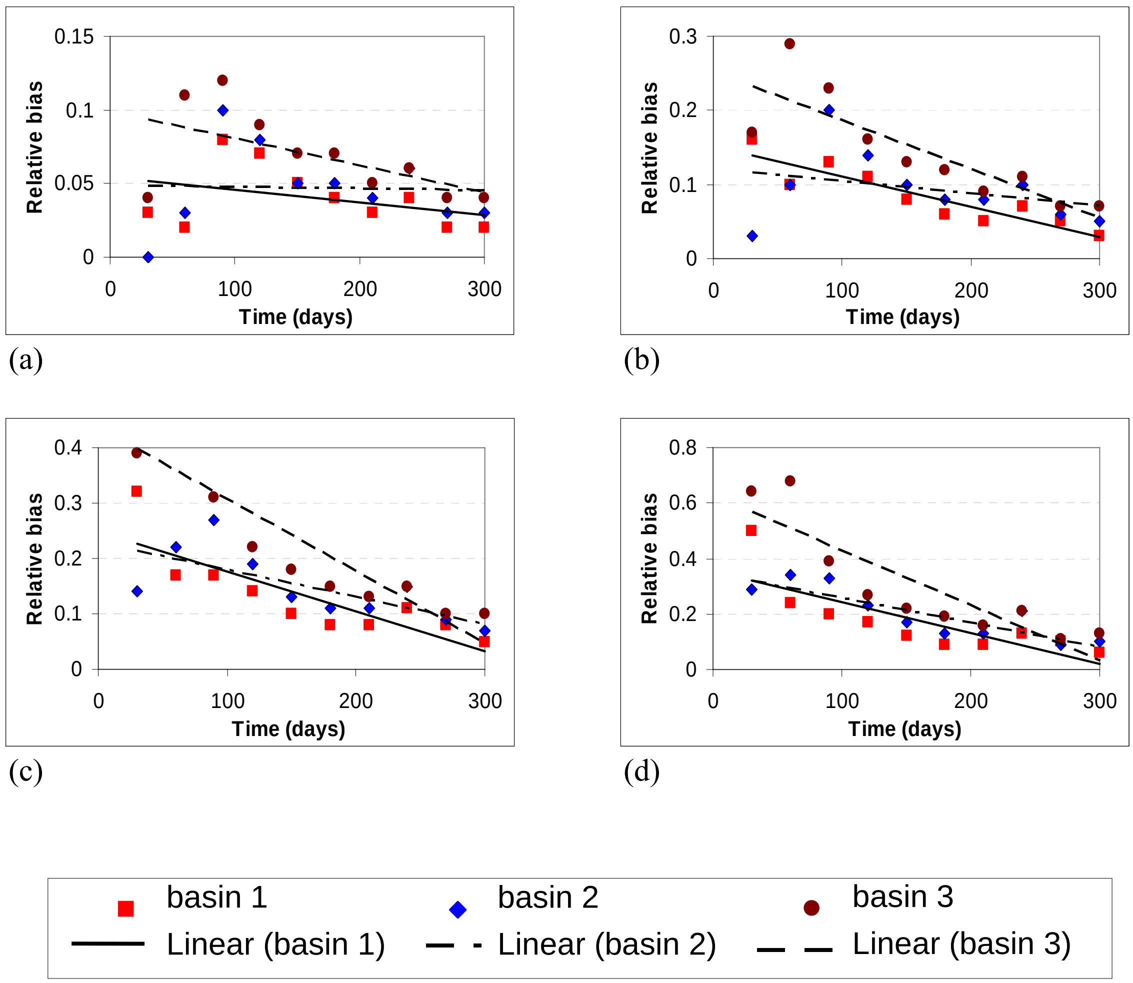


Figure 8-4. **Error in the predicted discharge caused by a positive deviation in UZS. (a)=25%; (b)=50%; (c)=75%; and (d)=100%**

8.2.3 Effect of Initial Storage Values on the Systematic error of the Predicted Flow Components

To investigate the effect of the estimates of the initial storages during the start-up period on the accuracy of the predicted runoff, an analysis of the goodness-of-fit statistics was performed. Determining the sensitivity of the predicted flow to deviations in the initial storages may contribute to better approaches for calibration and to have a better understanding of the implications of the error on the prediction accuracy. The accuracy of total discharges estimates is sometimes difficult to evaluate based solely on the relative bias of the daily discharge. The hypothetical watersheds 1 and 2 (Table 8-1) were used in this analysis. For the watershed with larger proportion of baseflow, it was expected that deviations in either LZS or UZS would be mostly observed in the accuracy of the predicted baseflow. In contrast, an atypical behavior regarding the over or underprediction was expected for watershed 2, because all the hydrologic processes would have a similar importance because the flow proportions (Table 8-1).

Table 8-10. Relative bias of the total predicted discharge for the first year of the period of record

Over and underestimations of LZS and UZS (%)	LZS		UZN	
	Watershed 1	Watershed 2	Watershed 1	Watershed 2
100	0.266	0.279	0.176	0.201
75	0.220	0.232	0.131	0.149
50	0.161	0.170	0.086	0.098
25	0.088	0.093	0.042	0.049
0	0	0	0	0
-25	-0.098	-0.108	-0.037	-0.047
-50	-0.195	-0.217	-0.065	-0.091
-75	-0.286	-0.323	-0.086	-0.131
-100	-0.367	-0.415	-0.103	-0.164

The results of the relative biases of the total predicted discharges (Table 8-10) and for the first 365 days of the period of record, suggest that for basins with predominant baseflow, the over or underestimation of LZS will be the most important factor when determining error in the water mass balance. Similar results are obtained from the analysis of the relative standard error ratio (Table 8-11). However, the relative standard error ratio when error in LZS and UZS exist suggests that the best accuracy of the predicted discharge is attained when both state variables (LZS and UZS) are underpredicted.

Table 8-11. Relative standard error ratio of the total predicted discharge for the first year of the period of record

Deviation from true parameter value (%)	LZS		UZN	
	Basin 1	Basin 2	Basin 1	Basin 2
100	1.160	0.881	0.438	0.302
75	0.753	0.590	0.345	0.239
50	0.441	0.349	0.255	0.171
25	0.198	0.159	0.160	0.091
0	0	0	0	0
-25	0.192	0.137	0.154	0.096
-50	0.386	0.268	0.241	0.185
-75	0.552	0.389	0.286	0.255
-100	0.686	0.486	0.312	0.305

The results in Table 8-12 indicated that when of underestimating LZS and UZS the most significant effect on the predicted baseflow was underprediction. The largest values were obtained for underestimations of LZS with $R_{bB} = -0.608$, and $R_{bB} = -0.426$ for watersheds 2 and 1, respectively. Conversely, overestimation of LZS and UZS produced overprediction of baseflow. The largest values were obtained for overestimation of LZS with $R_{bB} = 0.224$, and $R_{bB} = 0.192$ for watershed 1 and 2, respectively. In terms of baseflow, it seems that overestimation of LZS and UZS would produce a lower relative bias in the predicted baseflow component during the first year of the period of record.

Table 8-12. Relative bias (R_{bB}) of the 20 lowest predicted discharges (baseflow) for the first year of the period of record

Over and underestimations of LZS and UZS (%)	LZS		UZN	
	Watershed 1	Watershed 2	Watershed 1	Watershed 2
100	0.224	0.192	0.130	0.147
75	0.233	0.157	0.099	0.117
50	0.191	0.175	0.066	0.084
25	0.116	0.120	0.032	0.045
0	0	0	0	0
-25	-0.109	-0.127	-0.037	-0.050
-50	-0.214	-0.268	-0.079	-0.104
-75	-0.321	-0.440	-0.123	-0.164
-100	-0.426	-0.608	-0.165	-0.222

The over and underprediction of peak flows (Table 8-13) was similar in both watersheds when LZS and UZS deviated from their true value. This may be explained by the fact that the peak discharges are more affected by variations in parameters that control surface runoff from impervious area (not included in the analysis) than by variations in the initial storage parameters of pervious areas. However, predictions of peak flows seem to have a lower systematic error when underestimating LZS and UZS, which suggest that underestimation of the initial storages could provide the lowest relative error.

Table 8-13. **Relative bias (R_{bP}) of the 20 largest predicted daily discharges (peak flow) for the first year of the period of record**

Over and underestimations of LZS and UZS (%)	LZS		UZN	
	Watershed 1	Watershed 2	Watershed 1	Watershed 2
100	0.736	0.720	0.216	0.227
75	0.485	0.501	0.172	0.178
50	0.283	0.303	0.127	0.126
25	0.124	0.138	0.076	0.065
0	0	0	0	0
-25	-0.109	-0.114	-0.070	-0.063
-50	-0.217	-0.219	-0.116	-0.120
-75	-0.310	-0.311	-0.146	-0.167
-100	-0.385	-0.388	-0.167	-0.205

8.3 EFFECT OF LZS AND UZS DEVIATIONS ON THE START-UP PERIOD OF PREDICTED RUNOFF

The time to which the predicted discharges become insensitive to the estimates of the initial storages is of interest when evaluating prediction accuracy, as error in the estimates propagates to the predicted runoff. The length of time to the point of insensitivity is of importance because in many cases this period can be a large part of a short record length used for calibration. In these cases, an alternative approach may be needed to provide the time for the model to offset the error in the initial estimates.

Table 8-14. **Number of days for the predicted flows to be independent of the estimate of the initial storages**

Watershed	Parameter deviation (%)							
	-25	-50	-75	-100	+25	+50	+75	+100
	Parameter LZS							
1	0	249	283	291	86	346	347	418
2	0	283	301	328	283	345	371	429
3	0	79	283	283	0	285	345	346
	Parameter UZS							
1	86	86	86	87	86	86	86	87
2	0	87	113	248	0	86	88	112
3	67	88	145	283	67	85	93	283

The HSPF model required a start-up period to allow the predicted discharges to become insensitive to erroneous estimates of the initial storages. It was expected that based on the selected criterion of Eq. (8-2), the time for the predicted discharges to become insensitive to the estimates of initial soil storage would increase as the deviation in the estimates of initial storage increased. The results shown in 699HTable 8-14 indicated that the time for the model to offset the erroneous initial soil storage was slightly longer (less than 6 months) for overestimations than underestimations of LZS (Figure 8-5(a)). This difference can be explained by the HSPF algorithm that determines the amount of infiltrated and percolated water into the lower zone. The lower-zone storage ratio of LZS/ LZSN determines the fraction of direct infiltration plus percolation that enters the lower-zone storage through the following equations. When the ratio (LZS/LZSN) is 1.0 or greater than 1.0 (overestimation of LZS):

$$LZFRAC = 1.0 - (LZS / LZSN) * (1.0/(1.0 + INDX))\wedge INDX \quad \textbf{(8-4)}$$

When the ratio (LZS/LZSN) is less than 1.0 (underestimation of LZS):

$$LZFRAC = (1.0/(1.0 + INDX))\wedge INDX \quad \textbf{(8-5)}$$

INDX is defined as:

$$INDX = 1.5 * ABS((LZS / LZSN) - 1.0) + 1.0 \quad \textbf{(8-6)}$$

where LZFRAC is the fraction of infiltration plus percolation entering the lower-zone storage and ABS is a function for determining the absolute value. When the ratio of LZS/LZSN is greater than 1.0, a minimum amount of water is allowed to enter the lower-zone storage (Bicknell et al, 2001). In this case, the additional water that otherwise would be in the lower-zone storage is kept in storages located in the upper zones of the system. Thus, the error propagates to the upper-zone storages. In contrast, when the ratio of LZS/LZSN is lower than 1.0, the lower-zone storage is affected by a deficit of water not flowing downwards. With the error constrained to the lower-zone storage, a shorter time for the nominal storage to reach its true value is observed. In regard to the effect of the flow proportions on the start-up period, the results shown in

Figure 8-5 indicated that, when the deviations in LZS are greater than 25% (which it should be assumed to be the normal case), the start-up period was similar for the three hypothetical watersheds, suggesting that the flow proportion is not an important factor for the start-up period, and that in any case, a start-up period of about a year should be allowed for calibration purposes.

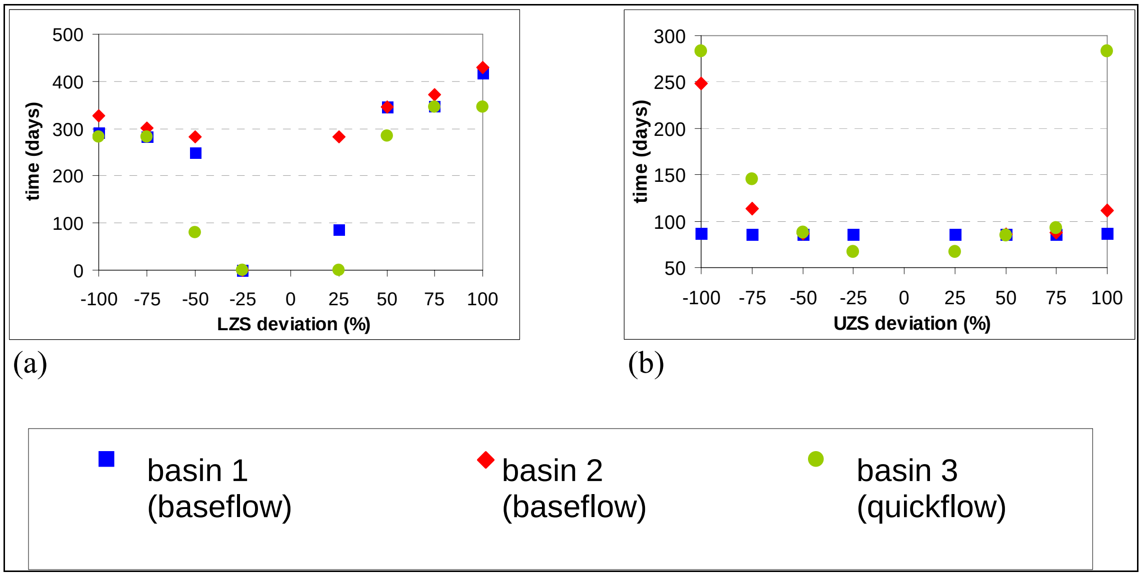


Figure 8-5. Number of days in which the predicted discharge is within 30% error of the actual daily discharge. Error caused by a deviation in (a) the initial lower-zone storage (LZS) and (b) the initial upper zone storage (UZS).

Erroneous estimates of the initial upper zone storages (UZS) significantly increased the start-up period for watersheds with predominant quickflow when the deviations in UZS estimates were greater than 75% (Table 8-14 and Figure 8-5). It may be that for deviations in UZS less than 75% the start-up periods were affected by the threshold of the selected criterion in Eq. 7-2 (a 30% or greater difference between the actual and predicted discharges). Although the differences between the predicted and actual discharges had different values for the different lower UZS deviations, the differences were not large enough to be greater than the threshold. In contrast, the start-up period was a function of the flow proportions for deviations of UZS greater than 75%. These results clearly suggest that the estimate of the initial upper zone storage is an important value which must be considered when the available record for calibration is short (less than 1 year). The ideal scenario should allow for a start-up period of about 1 year and to calibrate only with subsequent predictions.

8.4 EFFECT OF LZS AND UZS DEVIATION ON THE START-UP PERIOD OF THE PREDICTED STORAGES

The amount of water in the soil at the beginning of a discharge record is generally unknown, but it certainly influences the record of measured discharges. If the estimates of the watershed storages for the initial conditions are incorrect, then the computed depletion of water from the simulated storages will not correspond to the measured values. Nevertheless, the simulated amounts of water in the system will eventually converge to the actual state, after which the effect of the errors in the predicted storages will be negligible. It was of interest to assess whether or not the start-up periods associated with the storages were similar to the start-up periods associated with discharges. Although measurements of storages are never available on actual watersheds, the use of hypothetical data enables the start-up period for storages to be evaluated.

In sections 8.2.1 and 8.2.2 it was demonstrated that the predicted discharges may experience the effect of error in the initial storages for about a year from the start of the period of record. To determine the effect on the start-up period of the predicted storages caused by deviations in LZS and UZS, data from the hypothetical watersheds 1 and 2 (Table 8-1 were used in the analysis. Although in both watersheds the assumed true initial value for LZS was 5.00, the daily average lower-zone storage for watershed 1 was 6.63 in. and for watershed 2 was 6.16 in.. The daily average upper zone storage for watershed 1 and 2 was 2.61 in. and 3.52 in., respectively. The convergence criteria to determine when the predicted nominal storages were insensitive to the error of the soil storage estimates were set to 0.1 in. and 0.3 in.

Table 8-15. Start-up period (days) based on the relative error (UZS)

% deviation	Watershed 1		Watershed 2	
	Convergence criteria			
	0.1 in.	**0.3 in.**	**0.1 in.**	**0.3 in.**
-100	206	156	181	85
-75	183	91	115	79
-50	157	85	86	56
-25	87	0	73	0
0	0	0	0	0
+25	85	0	66	0
+50	86	45	80	55
+75	87	57	85	57
+100	87	66	85	65

Table 8-16. Start-up period (days) based on the relative error (LZS)

% deviation	Watershed 1		Watershed 2	
	Convergence criteria			
	0.1 in.	**0.3 in.**	**0.1 in.**	**0.3 in.**
-100	284	78	317	114
-75	269	66	292	79
-50	250	46	273	56
-25	70	0	86	0
0	0	0	0	0
+25	86	0	269	0
+50	307	57	333	250
+75	347	269	376	287
+100	418	308	451	345

The flow proportions were not a significant factor for the start-up period when UZS or LZS were overestimated, Table 8-15 and Table 8-16, respectively. For example, using the convergence criteria of 0.1 and 0.3, for a positive 100% deviation in LZS the start-up periods in watershed 1 were 408 and 308 days while for watershed 2 were 451 and 345 days. Similarly, for a positive deviation of 75% in UZS the start-up periods were 87 and 85 days for watershed 1 and 2, respectively, when using the 0.1 criterion. When LZS and UZS are underestimated, the start-up period variation was slightly larger between watersheds. This effect may be explained by the lower and assumed true value of UZS (2.0 in.) in comparison to the assumed true value of LZS (5.0 in.). Overestimation of LZS produced longer start-up periods than underestimations of LZS (Figure 8-6) and underestimations of UZS produced longer start-up periods than overestimations than UZS (Figure 8-7). In this respect, it seems it would be better to over predict the initial UZS and to underpredict the initial LZS.

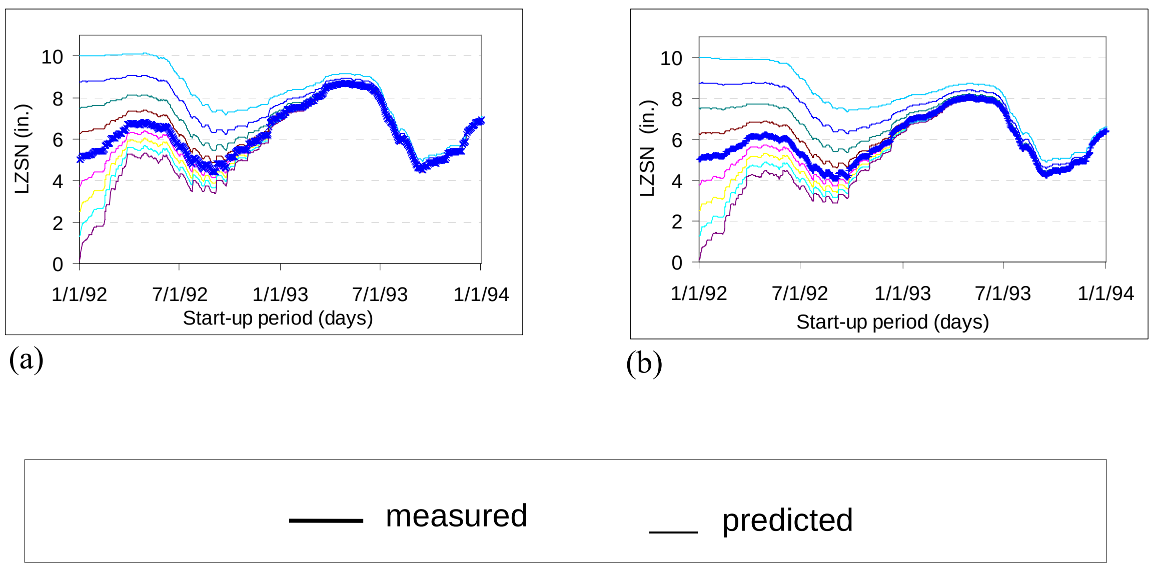


Figure 8-6. **Effect of LZS estimates on the predicted lower zone storage (LZSN). (a) watershed # 1 and (b) watershed # 2.**

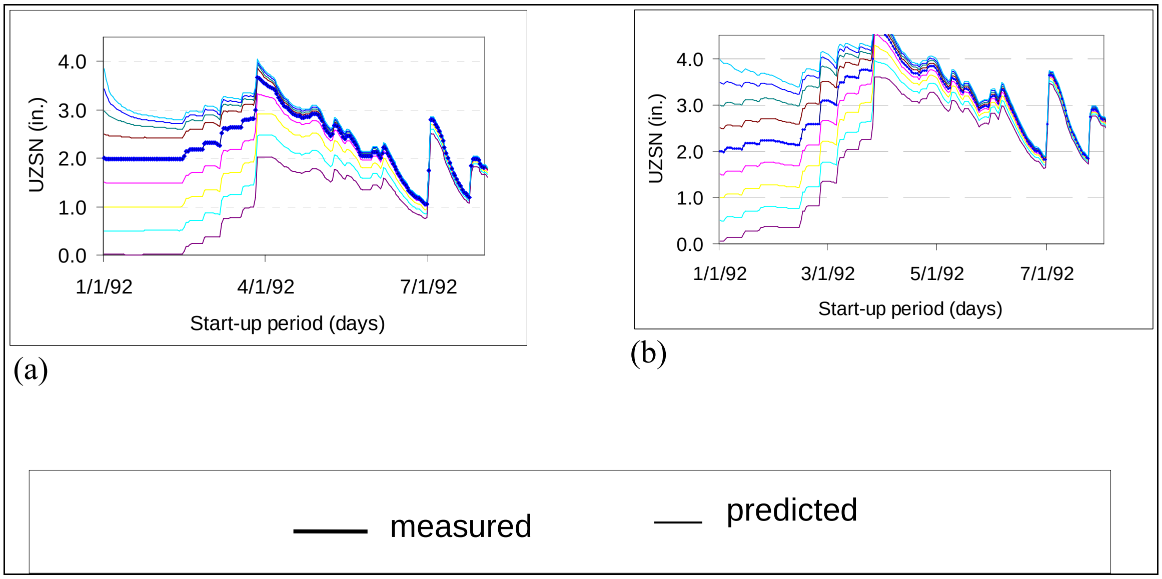


Figure 8-7. Effect of UZS estimates on the predicted upper zone storage (UZSN). (a) watershed # 1 and (b) watershed # 2.

The results in this section coincide with the results in section 8.3 where it was concluded that the flow proportions did not have a significant effect in the start-up period of the predicted discharges. However, it is clear that the first year of data should be designated as the start-up period and predictions from this period should be disregarded for calibration or to compute the goodness-of-fit statistics of the predictions.

8.5 EFFECT OF INITIAL STORAGE ESTIMATES ON CALIBRATION ACCURACY

It is a common practice to calibrate using the entire record, including the start-up period. As the start-up period distorted the predicted discharges and storages, it seems reasonable to expect that the optimized parameter values could be erroneous if the start-up period of record were used in the calibration. If the errors can adversely affect the parameter values, then the calibration should only be based on the record after the start-up period. Using the calibration criteria described in section 6.3.1, the objective of the following analyses was to determine if deviations in the state variables UZS and LZS had an effect on the accuracy of the optimized parameters and if the length of the record and the climatological conditions during the selected period for calibration were important factors for the parameters to reach their true values. Analyses were made using watershed 1 with length of records of 1 and 8 years and the values of the initial storages UZS and LZS were individually increased by 50% of their true values. In analyses in which a calibration was made, the initial values of the parameters did not equal the true values. It was expected that an unimportant parameter would not reach its true value, but that important parameters would optimize regardless of the record of length.

8.5.1 Effect of a Start-up Period on the Prediction Accuracy

Determining accurate values of the initial soil storages in hydrologic modeling may be a matter good fortune; however, preventing the effects of such speculations in the calibration process may have positive rewards that will be reflected not only in accurate predictions, but also in more reliable parameter values. The purpose of this analysis was to determine if by excluding the predicted discharges during the start-up period more accurate predicted discharges and more accurate parameter values could be attained. By removing these predictions from the calibration and analysis, it was expected to remove the noise and error caused by inaccurate initial storages. Two independent analyses were made: one with UZS starting at the true value and LZS off by 50%; in the second analysis UZS was off by 50% and LZS was the true value. The analysis was made using watershed 1 (Table 8-1) and 8 years of actual climatological data (01/01/1992 through 12/31/1999). The start-up period was assumed to be the year 1992.

Table 8-17. Statistical summary of HSPF predicted discharge for calibrations including and excluding the start-up period (1992)

LZS: 50% deviation				UZS: 50% deviation			
1992-1999		1993-1999		1992-1999		1993-1999	
R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
0.103	0.275	-0.009	0.077	0.106	0.277	-0.03	0.163

Table 8-18. Statistical summary of HSPF annual predicted discharges for calibrations including and excluding the start-up period (1992)

		LZS: 50% deviation				UZS: 50% deviation			
		1992-1999		**1993-1999**		**1992-1999**		**1993-1999**	
Year	P	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
1992	98	-0.104	0.582	0.150	0.450	-0.130	0.627	-0.081	0.292
1993	108	0.008	0.418	0.004	0.098	0.021	0.417	-0.033	0.159
1994	110	0.082	0.202	-0.007	0.079	0.086	0.201	-0.030	0.267
1995	94	0.296	0.649	-0.019	0.088	0.308	0.645	-0.032	0.213
1996	137	0.080	0.253	-0.012	0.142	0.081	0.254	-0.038	0.185
1997	87	0.163	0.246	-0.005	0.093	0.169	0.275	-0.014	0.143
1998	85	0.111	0.202	-0.019	0.104	0.111	0.197	-0.027	0.168
1999	111	0.395	0.297	-0.024	0.021	0.396	0.297	-0.055	0.057

The effect of error in the initial storages can be removed by excluding the start-up period predictions from the period of record. Better accuracy was attained when the predictions from the start-up period were removed from the calibration process. The improvement of the total predicted discharges was observed throughout the calibration (Table 8-17) and from year-to-year (Table 8-18). A more significant improvement in the accuracy of the predicted discharges was obtained when the inaccuracies caused by incorrect estimates of LZS were removed from the calibration process, than when the inaccuracies due to deviations in UZS were removed. This may be explained by the calibration criteria that emphasize the calibration of baseflow rather than the calibration of quickflow, or because in this case, baseflow was the predominant flow component in the watershed.

Using the results from Table 8-18 and a 50% deviation in LZS, the magnitude of the relative bias during the period of calibration (1992-1999) varied between 0.8% and 39.5%, when predictions from the start-up period (1992) were included in the calibration process. In contrast, when the start-up period was excluded, the magnitude of the relative bias during the period of calibration (1993-1999) varied between 0.4% and 15.0% for years of calibration (1993 - 1999). Similar results were obtained with the relative standard error ratios which were also reduced when the start-up period was excluded from the calibration. The values were reduced from a range between 0.202 and 0.649 to a range of values between 0.021 and 0.450.

When UZS was deviated by 50% of its true value and when predictions from the start-up period (1992) were included in the calibration the magnitude of the relative bias varied between 2.1% and 39.6% (Table 8-18). When the start-up period was excluded, the absolute value of the bias varied between 1.4% and 8.1%. The relative standard error ratio also decreased from a range between 0.197 and 0.627 when the start-up period was included in the optimization to a range of values between 0.057 and 0.292 when the start-up period was excluded. The reduction of the systematic variation caused by error in the estimation of the initial storages reduced the values of the relative standard error ratio. This improvement in accuracy was significant.

Table 8-19. Final parameter values from the LZS and UZS independent analyses, when the start-up period (1992) was included and excluded from the optimizations

		AGWR	LZETP	LZSN	UZSN	INFILT	INTFW	NSUR	IRC
True Value		0.970	0.30	5.00	2.00	0.07	0.50	0.08	0.30
Initial		0.925	0.10	16.00	8.00	0.21	2.00	0.20	0.80
Year		Final parameter values							
1993–1999	+50% LZS	0.969	0.34	5.97	1.66	0.08	0.53	0.05	0.01
1992-1999	+50% LZS	0.974	0.01	10.63	3.25	0.11	0.44	0.05	0.01
1993–1999	+50% UZS	0.969	0.31	6.08	2.06	0.05	0.40	0.50	0.01
1992-1999	+50% UZS	0.975	0.01	9.22	3.25	0.12	0.43	0.05	0.01

The final parameter values (Table 8-19) of the independent calibrations of LZS and UZS were closer to the true values when the predictions from the start-up period were removed from the calibration. Not only AGWRC optimized, but other important parameters such as the nominal storages (LZSN and UZSN) and the lower zone evapotranspiration parameter (LZETP) calibrated to values near the true values. In addition, relatively unimportant parameters such as INFILT and INTFW were also optimized. IRC consistently ended at the lower bound of possible values and never calibrated. Except for when the start-up period was removed from the calibration and UZS was deviated from its true value, NSUR also consistently ended at the lower bound of possible values and never calibrated. When the start-up period was excluded, the parameter moved in the opposite direction from the true value ending at the upper bound of possible values (0.5). This unexpected behavior may be explained by the intercorrelation with other parameters such as INTFW or IRC which also control processes in the upper soil layer. Although the sources of error in the calibration of parameters for real watersheds are considerably large and diverse, the elimination of error caused by erroneous estimates of the initial soil storages reduces some of the uncertainty in the final parameter values.

8.5.2 Effect of Precipitation on the Importance of the Start-Up Period

In the previous section, the importance of the start-up-period was demonstrated. However, it was of interest to determine if the importance of the start-up period was a function of the precipitation depths during the year. If precipitation was a factor, then the findings of this analysis could provide some guidance in applications with short period of record. In many cases, HSPF applications are developed using records limited to as little as one year of data. In these cases, providing the start-up period would be impossible, unless some type of data patching, using the same period of record twice could provide for the start-up period. In this case, discontinuity could be a factor for prediction accuracy.

Short record lengths of one year were used for the calibrations. Two sets of years with similar annual depths of precipitation were selected: the first set included years where the annual rainfall was near average while the second set included years with below average rainfall. The two sets are as follows: (1) 1989, 1990, 1993, 1994, and 1999 with 114.7 cm, 112.1 cm, 108.8 cm, 110.0 cm, and 110.7 cm, respectively, and (2) 1988, 1991, 1997, and 1998 with 82.4 cm, 77.7 cm, 85.9 cm, and 85.5 cm, respectively. Two independent sets of calibrations were made: the first, with a 50% deviation to LZS and the second, with a 50% deviation to UZS. The mean and the standard deviation of the actual discharge ($\overline{X}$ and S_x) and the predicted discharge ($\overline{Y}$

and s_y) are shown in 719HTable 8-20. The relative bias and the relative standard error ratio were used as the criteria for measuring the prediction accuracy.

Table 8-20. **Statistical summary of HSPF predictions when using a 50% deviation in the initial lower (LZS) and upper (UZS) zone storage values. (P = Annual precipitation (cm); $\overline{X}$, $\overline{Y}$, mean of measured and predicted discharges, respectively (10E-2 m^3/sc); s_x and s_y standard deviation of measured and predicted. discharges, respectively (10E-2 m^3/sc); R_{bR} = relative bias of the daily predicted discharge; S_e/S_y = standard error ratio of the daily predicted discharge).**

Storage	Year	P	$\overline{X}$	s_x	$\overline{Y}$	s_y	R_{bR}	S_e/S_y
LZS	1989	114.7	15.91	22.64	16.50	21.22	0.037	0.226
	1990	112.1	11.64	7.03	12.38	6.59	0.062	0.456
	1993	108.0	14.64	15.14	15.13	15.16	0.034	0.174
	1994	110.0	16.37	24.60	16.04	24.60	-0.106	0.729
	1999	111.0	15.69	40.81	16.90	40.72	0.078	0.169
	1986	82.4	5.95	8.06	8.25	5.92	0.023	0.114
	1991	77.7	6.48	6.51	6.65	5.18	0.027	0.142
	1997	87.0	7.48	7.47	10.01	7.22	0.011	0.157
	1998	85.0	14.61	20.29	13.45	20.55	-0.079	0.531
UZS	1989	114.7	15.91	22.60	16.45	20.81	0.034	0.246
	1990	112.1	11.64	7.03	12.69	5.98	0.088	0.418
	1993	108.0	14.64	15.14	14.64	15.11	0.001	0.081
	1994	110.0	16.37	24.60	13.61	24.61	-0.163	0.730
	1999	111.0	15.69	40.81	16.33	40.81	0.094	0.095
	1986	82.4	5.95	8.06	8.31	5.71	0.030	0.122
	1991	77.7	6.48	6.51	6.56	6.37	0.014	0.178
	1997	87.0	7.48	7.47	10.34	7.45	0.044	0.162
	1998	85.0	14.61	20.29	13.69	20.29	-0.063	0.524

The relative bias is an indicator of the systematic error variation of the model predictions. The results shown in Table 8-20 suggest that the annual precipitation is not a factor in the prediction accuracy. The relative bias of the predictions (Table 8-20) using similar annual precipitation was moderate for years with average annual rainfall (below 10% overprediction) and reasonably good for years with below average rainfall depths (below 5% overprediction). However, for years in which extremely high precipitation was recorded for the Jan-Apr period (1994 and 1998), the prediction accuracy was poor. For calibrations with a 50% deviations in the initial LZS, the overpredictions varied between 1.1% and 7.8%, except in the years 1994 and 1998 in which the underpredictions were -10.6% and -7.9%, respectively (see Table 8-20). For calibrations with a 50% deviations in the initial UZS, the overpredictions varied between 0.1% and 9.4%, except again for 1994 and 1998 with underpredictions of -16.3% and -6.3%, respectively. The inaccuracies for 1994 and 1998 were related to the irregular distribution of rainfall (Table 8-21) throughout the year: Very high rainfall depths during the Jan-Apr period (50.4 and 50.8 cm, respectively), and very low rainfall throughout the rest of the year.

The trends in biasedness are rational as the mean value of the predicted discharges is increased by the presence of very high precipitation events. The increment in the mean annual discharge introduced error into the calibration process, as least squares calibration tends to calibrate towards mean values. Nevertheless, the effect of outliers on prediction accuracy was more severe during the Jan-Apr period than during any other season because of the combined effect with the error in the initial storages. These results suggest that caution must be taken when calibrating to conditions that include very high precipitation events during the first months of the record. The combined effect of outliers and error in the initial storages significantly affect prediction accuracy. In these cases, the extension of the period of record should be necessary although the potential issues of discontinuity should be investigated.

The relative standard error ratios (Table 8-20) showed considerable variation for similar annual rainfalls which suggest that factors other than the annual rainfall depth may contribute to the variation of accuracy under similar depths of annual precipitation.

For calibrations using average annual precipitation depths and deviations of 50% in the initial lower-zone storage (LZS) the relative standard error ratio (S_e/S_y) varied between 0.169 and 0.785. The variation of S_e/S_y for calibrations using the low annual precipitations and with a 50% deviation in the initial LZS was between 0.114 and 0.531. Moderate values were obtained for 1990 (0.456 and 0.418 for a 50% deviation in LZS and UZS, respectively). Good accuracy (values less than 0.300) was obtained for the years 1986, 1989, 1991, 1993, 1997, and 1999 in which the error in the initial storages was the only effect during the Jan-Apr period.

For years with very high precipitation during the Jan-Apr period (1994 and 1998) the combined effect of outliers and the error in the initial storages adversely affected the prediction accuracy, as reflected in the calculated values of the relative standard error ratio. For 1994 and 1998 the values of S_e/S_y were 0.729 and 0.530 for a 50% deviation in LZS, and 0.730 and 0.524 for a 50% deviation in UZS, respectively. Therefore, the total annual precipitation does not appear to be a factor, rather it is the distribution of rainfall during the year.

A comparison of the standard error ratios of Table 8-20 and the monthly rainfall depths in Table 8-21 suggests that poor accuracy was associated with years with high rainfall during the early part of the year (Jan-Apr). The total rainfalls for Jan-Apr are 44.7, 32.5, 44.0, 50.4, and 32.4 cm for 1989, 1990, 1993, 1994, and 1999, respectively. Thus the largest standard error ratio, which occurred in 1994, occurred in the year when the Jan-Apr rainfall was largest. The same trend was evident for the four below average rainfall years (1986, 1991, 1997, and 1998). The standard error ratio was largest in 1998, which was the year with the largest Jan-Apr total rainfall. Thus, it was of interest to evaluate the effect of extreme rainfall events on prediction accuracy and a possible solution to increase the chances of calibration.

Table 8-21. Distribution of monthly rainfall depths (cm)

	1989	1990	1993	1994	1999	1986	1991	1997	1998
Jan	6.5	10.3	6.9	11.7	11.4	6.0	8.3	7.1	14.4
Feb	7.0	5.0	7.2	10.3	6.5	5.3	2.0	5.6	15.6
Mar	20.1	6.4	20.6	21.9	8.7	13.0	12.4	12.9	14.1
Apr	11.1	10.8	9.3	6.5	5.8	4.7	4.9	5.8	6.7
May	7.3	13.2	9.1	7.7	4.4	5.7	3.4	5.3	8.6
Jun	6.8	6.2	7.7	7.2	5.2	4.6	2.0	2.9	8.2
Jul	5.4	11.7	4.3	11.5	5.2	4.1	6.3	3.2	3.6
Aug	9.1	16.9	6.5	8.7	15.5	8.5	6.4	10.7	2.3
Sept	12.7	3.3	10.4	10.0	29.2	5.2	9.3	3.7	3.2
Oct	8.2	9.3	7.7	4.6	6.3	7.0	6.5	8.7	2.7
Nov	9.2	6.0	7.8	5.0	5.0	13.1	5.1	14.8	2.9
Dec	11.4	13.0	11.3	4.9	7.5	5.2	11.1	5.2	3.2
Total	**114.8**	**112.1**	**108.8**	**110.0**	**110.7**	**82.4**	**77.7**	**85.9**	**85.5**
Mean	9.6	9.3	9.1	9.2	9.2	6.9	6.5	7.2	7.1

8.5.3 Effect of Outliers on Prediction Accuracy

Associations between poor prediction accuracy (Table 8-20) and high monthly depths of precipitation (Table 8-21) were established in the previous section. However, since the variations in accuracy were not entirely explained by the anomalies in the monthly precipitation, characteristics of the observed discharge, specifically the presence of outliers, must be investigated. The most significant implication of outliers is their potential to distort the calibration process. In this case, it is best to address the problem of outliers through a with-vs.-without sensitivity analysis. Analyses that use values reduced to reasonable amounts are compared with the analyses that use the measured data with the extreme events.

Table 8-22. **Annual precipitation (P cm), S_e/S_y = standard error ratio of the daily predicted discharge, and the five largest measured discharges (10E-2 m3/sc) by year. The month in which the discharge occurred is shown in parenthesis.**

YEAR	P	**Largest daily discharges**				
1994	110.0	302.08 (3)	175.11 (3)	152.12 (3)	144.30 (3)	90.98 (3)
1998	85.0	171.66 (3)	166.87 (3)	120.83 (3)	104.57 (2)	67.14 (3)
1990	112.1	81.47 (5)	68.95 (12)	44.17 (4)	38.40 (12)	36.73 (5)
1989	114.7	281.16 (12)	207.25 (3)	102.34 (3)	95.23 (4)	84.13 (5)
1993	108.0	117.83 (12)	96.11 (3)	89.14 (3)	68.64 (4)	68.44 (5)
1999	111.0	776.33 (9)	74.25 (12)	55.73 (9)	54.54 (3)	46.81 (9)
1997	87.0	78.18 (3)	29.76 (3)	26.42 (3)	24.98 (3)	24.75 (3)
1991	77.7	81.61 (3)	21.55 (1)	21.10 (3)	19.34 (3)	18.09 (3)
1986	82.4	47.12 (3)	27.85 (3)	26.59 (3)	22.09 (3)	21.41 (3)

To determine if the outliers were an important factor in the prediction accuracy, the calibrations (1994, 1998, and 1990) with the three largest relative standard error ratios (0.729, 0.531, and 0.456) in (Table 8-20) were selected for additional analysis. These new calibrations used modified records of precipitation and discharge. Specifically, the values of the larger flows and their corresponding rainfalls were lowered so they were not extreme events. The with-and-without-outlier comparison may help identify the reason for the poor goodness-of-fit statistics in some years. The modification procedure was as follows: The five largest measured discharges in each year were designated as extreme events and, therefore, as potential outliers (Table 8-22). The hypothesis was that the reduction of the outliers to lower values could significantly improve the accuracy of the calibration. The measured and the modified daily discharges are shown in Table 8-23. The reduction of the outliers was made for the day in which the outlier was observed by reducing the precipitation amount and the measured daily discharge such that they were still larger values for the year but not extreme values. For cases in which the modified discharge was less than the two adjacent values, the lowest value of the two was assigned. The hourly precipitation values greater than 0.13 cm (0.05 in.) and the daily discharge in the measured discharge record were reduced in half. For the cases in which the modified precipitation value was less than 0.13 cm, a default value of 0.13 cm was assigned. The measured and the modified monthly precipitation depths are shown in Table 8-24. In 1990, the first extreme event occurred in April, with others in May and December. In 1994, all of the extreme events occurred in March. In 1998, all of the extreme events occurred in February and March.

Table 8-23. Measured daily discharge (10E-2 m3/sc) by year. Modified daily values in parenthesis.

YEAR	Largest daily discharges				
1994	302.08 (35.5)	175.11 (78.2)	152.12 (68.0)	144.30 (64.7)	90.98 (40.5)
1998	171.66 (77.0)	166.87 (74.9)	120.83 (54.2)	104.57 (47.0)	67.14 (38.1)
1990	81.47 (49.3)	68.95 (31.0)	44.17 (19.8)	38.40 (19.3)	36.73 (17.8)

Table 8-24. Measured monthly rainfall depths (cm). Modified monthly amounts in parenthesis.

Month	1990	1994	1998
Jan	10.3	11.7	14.4
Feb	5.0	10.3	(13.6) 15.6
Mar	6.4	(15.7) 21.9	(11.0) 14.1
Apr	(4.2)10.8	6.5	6.7
May	(4.9)13.2	7.7	8.6
Nov	6.0	5.0	2.9
Dec	(4.9)13.0	4.9	3.2

The prediction accuracy (Table 8-25) was significantly improved for the predictions using 1994 and 1998 data, given that all of the outliers were located within the Jan-March period. For these two years the improvements in accuracy were significant, between 30% and 40 % of the relative standard relative ratio and decreases in the absolute values of the biases. However, although the relative bias for predictions in the year 1990 were reduced from 6% to -0.3% when the LZS had a deviation of 50%, and from 8.8% to -1.2% when a UZS had a 50% deviation, the values of the relative standard error ratio increased slightly (0.456 to 0.462 for LZS and 0.418 to 0.442 for UZS). Although these are not significant increases, they do suggest that outliers in the latter part of the year have less of an impact on accuracy than do the extreme events that occur during the early part of the year when the errors in the initial storages are more significant. The outlying events interact with the overestimation of initial storages. The largest discharge value in 1990 was a relative low value (0.8147 m3/sec (Table 8-22)) when compared to the largest discharge in 1994 (3.021 m3/sec) or 1998 (1.718 m3/sec). In this case, other precipitation factors such as above average precipitation during the summer months or the fluctuation of monthly rainfall from above to below average may be influencing the accuracy of the results. In general, the results suggest that the detection of outliers is important as outliers can adversely affect the prediction accuracy and that, a sensitivity analysis should be undertaken to evaluate the effect of modified outliers on the accuracy of the predicted discharges.

Table 8-25. Statistical summary of HSPF predictions when using a 50% deviation in the initial lower (LZS) and upper (UZS) zone storage values. P = Annual precipitation (cm); R_{bR}, relative bias with and without outliers of the predicted daily discharge, respectively; S_e/S_y, relative standard error ratio with and without outliers of the predicted daily discharge.

			With outliers		Outliers adjusted	
Storage	Year	P	R_{bR}	S_e/S_y	R_{bR}	S_e/S_y
LZS	1990	112.1	0.062	0.456	-0.003	0.462
	1994	110.0	-0.106	0.729	-0.042	0.496
	1998	85.0	-0.079	0.531	-0.039	0.318
UZS	1990	112.1	0.088	0.418	-0.012	0.442
	1994	110.0	-0.163	0.730	-0.049	0.487
	1998	85.0	-0.063	0.524	-0.036	0.311

The need to have reasonable estimates of initial soil storages is unquestionable, as the analyses showed that the prediction accuracy will likely become poorer as the estimates of the initial storages deviate from their true values. However, measurements of storages are never available on actual watersheds. From the modeling perspective, the solution is to provide a start-up time in the application to guarantee that erroneous estimates of state variables representing the initial conditions in the system, will not be a factor in the calibration or in the prediction accuracy.

CHAPTER 9

EFFECT OF STATIONARY LAND USE ON PREDICTION ACCURACY

9.1 INTRODUCTION

The accuracy of predicted daily discharges using the HSPF model is affected by several factors, including the accuracy of the land use database and the nontemporal variation of the land-use, which may introduce error into the predictions when the watershed is undergoing land use change. Specific problems related to the accuracy of the land use databases and the difficulty in applying high-spatial resolution satellite image data for analysis of large urban areas for example, have been studied (Herold et al, 2002). The results of the study showed that problems due to the spectral and spatial complexity of urban areas may cause confusion between different roof types, roads and bare soil when the classification of the satellite image data is performed. In HSPF applications to real watersheds, these inaccuracies in the land use data base classification are transferred to the HSPF discharge predictions. To isolate these inaccuracies in the current analysis, it was decided to use hypothetical land use data.

The effects of watershed urbanization on streams are well documented (Leopold, 1968; Hammer, 1972; Hollis, 1975; Arnold et al., 1982). They include extensive changes in watershed hydrologic regime and channel morphology. Failure to account for these changes in HSPF applications and due to the use of stationary land use, may introduce inaccuracies in the model predictions. The effect of failure to account for temporal nonstationarity of land use with the HSPF can be evaluated. The standard version of the HSPF uses stationary land use data; however, a modified version of HSPF is maintained by the EPA-Chesapeake Bay Program Office (CBPO) in which linear interpolation between land-use databases from the years 1987, 1997, and 2002 is performed to input the land use data as a time series. However, the performance requirements of the modified HSPF are larger than those provided by a personal computer.

The purpose of this analysis was to assess the effect of assuming stationary land use for a watershed that has actually undergone land use change on the accuracy of the predicted discharges. To simulate the land use variation the HSPF data sequences that reflect nonstationarity were generated. Individual years of runoff were generated such that the land use was constant within a year, but changed from year to year in a systematic manner. The individual years were then combined to form a multiyear sequence. This discharge time series, referred to as the measured discharge, would reflect nonstationarity, but the discharges predicted by HSPF would be based on an assumed stationary land use sequence.

9.1.1 Data and Method of Analysis

Analyses were made using data from a hypothetical watershed that experienced uniformly increasing urbanization. The hourly rainfall data used for the analyses were measured data so that the assumed true daily discharges would reflect actual storm sequences. The watershed area for the hypothetical watershed was 1200 acres (1.88 mi2), and the assumed initial separation of the outflow was 70% baseflow, 27% interflow, and 3% surface runoff. The increment of urbanization would decrease the proportion of baseflow and increase the proportion of surface flow. The measured discharge was generated so that the true parameter values and the assumed true discharge were known. Given these conditions, the relative bias and the relative standard error of the predictions were chosen as the criteria to reflect the accuracy of the predicted discharges.

A 14-yr hypothetical land use time series was developed in which 37.5% of the forested area was assumed to be converted to pervious urban. Figure 9-1 shows the annual acreage of forest and pervious urban. The number of acres for each particular land use remained constant throughout each year but it changed from one year to the next (Table 9-1). Although 14 years of data seem short for a hydrologic analysis to assess the effect of using nonstationary land use on predicted discharges, this period is considered large in the context of continuous modeling. Fourteen runs were made to assemble the time series of measured daily discharges using the following procedure: First, the measured daily discharges for year 1 were obtained from HSPF in which the stationary land use was for year 1; similarly, the measured daily discharges for year 2 were computed using the stationary land use was of year 2. This procedure was followed for each of the 14 years. Then the 14 years were combined into a single 14-yr record. Second, five 14-yr analyses were made, each using the stationary land use from years 1, 4, 8, 10, and 12. Third, the goodness-of-fit statistics were calculated between the measured and predicted discharges for each year of the calibrations, and the accuracy of the predicted daily discharges was evaluated.

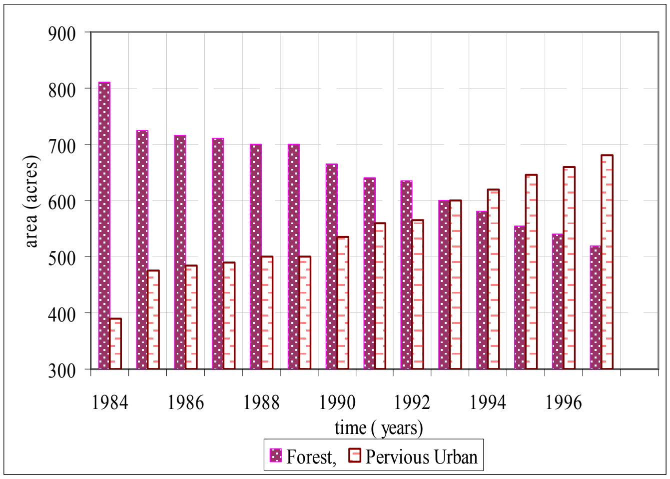


Figure 9-1. **Assumed nonstationary land use for a hypothetical watershed experiencing uniformly increasing urbanization.**

Table 9-1. **Assumed number of acres by land use and year in the hypothetical watershed.**

Year	forest	Pervious urban
1984	810	390
1985	725	475
1986	715	485
1987	710	490
1988	700	500
1989	700	500
1990	665	535
1991	640	560
1992	635	565
1993	600	600
1994	580	620
1995	555	645
1996	540	660
1997	520	680

9.1.2 Measures of Prediction Accuracy

To assess the accuracy of the predicted runoffs, two goodness-of-fit statistics were computed at the end of each calibration yearly, using the measured and predicted daily runoff. The relative bias (R_b), which is the ratio of the bias $\overline{e}$ to the mean discharge $\overline{Y}$, is a systematic error of the predicted discharge rates and is computed by:

$$R_b = \frac{1}{N\overline{Y}} \sum_{j=1}^{N} \left(\hat{Q}_j - Q_j \right) \qquad \textbf{(9-1)}$$

where Q_j is the measured daily outflow for the day j, or if the analysis is for the accuracy of the hourly outflows, then Q_j is the measured hourly outflow for the hour j; N is the number of days when the analysis is using the daily outflows, or the number of hours when the analysis is using the hourly outflows. The relative standard error (S_e/S_y) of estimate is:

$$S_e/S_y = \left[\sum_{j=1}^{N} (\hat{Q}_j - Q_j)^2 \bigg/ \sum_{j=1}^{N} (Q_j - \overline{Q})^2 \right]^{0.5} \qquad \textbf{(9-2)}$$

where the S_e/S_y is a measure of the nonsystematic error, with a value of zero indicating a perfect fit.

9.2 ACCURACY OF THE DAILY PREDICTED DISCHARGE

The characteristics of the predicted discharges are a function of several factors that includes the accuracy of the land use data. From hydrologic studies it is well known that, as a watershed undergoes urbanization, the duration of high flows decreases but the magnitude of the discharge hydrograph increases. This is responsible for problems such as higher erosion rates and reductions in groundwater recharge. Ignoring the change on the discharge hydrograph with the use of stationary land use, inaccuracies will be observed on the predicted discharges and on the predicted nutrient and sediment concentrations as the error in the predicted discharge is transferred. The goodness-of-fit statistics for the year 1984 are presented to illustrate the effect of inaccurate land use data during the start-up period.

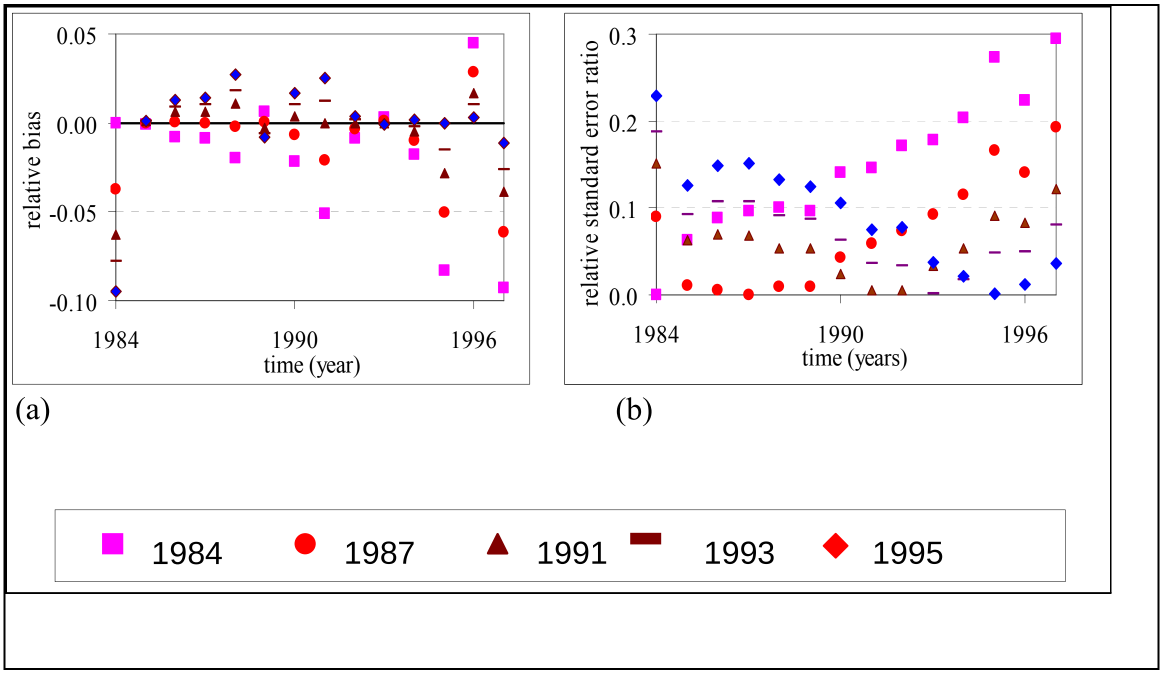


Figure 9-2. **Goodness-of-fit statistics of the predicted discharge when using stationary data. (a) relative bias; (b) relative standard error ratio**

9.2.1 Analysis of the Systematic Error Variation

The analysis of the systematic error is important as it provides information of the effect of error when stationary land use is applied for the calibration in the annual mass balance. In addition, the trend of the systematic error provides an indication of the annual percent change in the discharge hydrograph caused by urbanization. Two factors must be considered when interpreting the fitted trend: First, the effect of the stationary land use and second, the effect of very high or very low precipitation on the magnitude of the predicted discharges. Very large flows occurred in 3 of the 14 years and very small flows occurred in 4 of the 14 years (Table 9-2); these flows will be referred to as outliers. The years in which outliers were found are: 1984 (41.8 cfs), 1989 (1.06 cfs), 1993 (43.96 cfs and 0.88 cfs), 1994 (38.4 cfs and 1.07 cfs), and 1996 (1.41 cfs).

Table 9-2. Daily average (cfs), and the maximum and minimum annual discharge (cfs) for the period of record. (Difference between the daily average, and the maximum and minimum discharges for each year are shown in parenthesis). P = annual precipitation (cm).

		Discharge (cfs)		
Year	P	**Daily average**	**Maximum**	**Minimum**
1984	100.8	1.62	41.82 (40.20)	0.38 (1.24)
1985	98.9	1.52	32.54 (31.02)	0.51 (1.01)
1986	92.7	1.57	23.89 (22.32)	0.50 (1.07)
1987	97.8	1.93	19.02 (17.09)	0.62 (1.31)
1988	89.0	1.67	10.57 (8.90)	0.60 (1.07)
1989	126.9	2.67	24.94 (22.27)	1.06 (1.61)
1990	112.1	2.28	14.46 (12.18)	1.02 (1.26)
1991	77.7	1.58	21.77 (20.19)	0.57 (1.01)
1992	98.2	1.63	17.89 (16.26)	0.73 (0.90)
1993	114.7	2.67	43.96 (41.29)	0.88 (1.79)
1994	109.1	2.64	38.41 (35.77)	1.07 (1.57)
1995	93.1	1.54	15.00 (13.46)	0.68 (0.86)
1996	142.7	3.35	33.18 (29.83)	1.41 (1.94)
1997	82.4	2.11	16.22 (14.11)	0.95 (1.16)
Average	102.7			

In general, the systematic errors for the total discharge Figure 9-2 were marginal to moderate (below a 10% relative bias); however, the fitted trend was affected by the presence of outliers. The outliers were expected to modify the trend of the systematic error; however, the effect was stronger for years with a single outlier (low or high), than for years with combined low and high outliers (Table 9-2). For example, the most adverse effect in the trend of the systematic error (Figure 9-2 and Table 9-3) was expected for the year 1993, in which the largest difference between the maximum and the daily average was found (41.29 cfs); however, the presence of a low outlier (0.88 cfs) and the above average precipitation 56.8 in. (142.7 cm) counteracted its effect in the trend of the systematic error. On the other hand, the most adverse effect occurred in 1984 (the start-up period) in which only a high outlier was found, as measured by the second largest difference between the maximum and the daily average (40.20 cfs) and where the precipitation amount was average 40.3 in. (100.8 cm.)

Table 9-3. Relative bias of the predicted daily discharge (R_{bR}) when using stationary land use.

	Year of source of the stationary land use				
Year	1984	1987	1991	1993	1995
1984	**0.000**	-0.0373	-0.063	-0.078	-0.095
1985	-0.001	0.0001	0.001	0.001	0.001
1986	-0.008	0.0004	0.006	0.009	0.013
1987	-0.009	**0.0000**	0.006	0.010	0.014
1988	-0.020	-0.0019	0.011	0.018	0.027
1989	0.006	0.0006	-0.003	-0.006	-0.008
1990	-0.022	-0.0069	0.004	0.010	0.017
1991	-0.051	-0.0214	**-0.000**	0.012	0.025
1992	-0.009	-0.0037	-0.000	0.002	0.004
1993	0.003	0.0014	0.000	**-0.000**	-0.001
1994	-0.018	-0.0103	-0.005	-0.002	0.002
1995	-0.083	-0.0503	-0.028	-0.015	**0.000**
1996	0.045	0.0283	0.017	0.010	0.003
1997	-0.093	-0.0612	-0.039	-0.026	-0.011
Average	**-0.019**	**-0.012**	**-0.007**	**-0.004**	**-0.001**

The year 1996 was also of interest since the data show an effect unlike the remaining years (Figure 9-2 (a)). Specifically, the daily average discharges are overpredicted because of the largest annual record of precipitation and the largest difference (1.94 cfs) between the low outlier and the daily average discharge (Table 9-2). Thus, the variation in the systematic error shown in Figure 9-2 (a) is rational. As the number of years between the year during which the land use was measured and the year for which predictions are made increases, the magnitude of the systematic error also increases.

The results of the total predicted discharge suggest that the systematic error (Table 9-3) caused by the use of stationary land use on the accuracy of the predictions may be moderate (below a 5% error in magnitude) for ± 5 years from the year of origin of the stationary land use. However, these results may be misleading as the overprediction of baseflow is being compensated with underpredictions of quickflow or vice versa and thus, the changes in the total predicted volume of water are not significant. The start-up period (discussed in section 8.5.1) should not be included in the 5-yr range. In addition, it is important to be cautious in the interpretation of these results because of the moderate rate of urbanization. The results may be different and the errors may be larger when the rate and the density of urbanization increase.

A unique relation between the amount of annual precipitation and the systematic variation was not clearly observed (Table 9-3); however, when consecutive years of dry weather conditions occurred (1985-1988), the relative bias was marginal in comparison to values obtained for years in which the weather conditions were mixed (1994, 1995, and 1996). These results were expected as the discharge is reduced under dry weather conditions.

9.2.2 Analysis of Nonsystematic Error Variation

The nonsystematic variation of the model predictions can be assessed using the relative standard error ratio (S_e/S_y). As expected, the values of S_e/S_y (Figure 9-2(b)) were zero in years corresponding to the year in which the land use was stationary;

however, as the number of years between the stationary land-use data and the year of the predictions increased, the values of S_e/S_y also increased. The largest values (Table 9-4) were calculated at near the opposite end of the stationary land-use databases, for example, using the 1984 land-use database, $S_e/S_y = 0.295$ for predictions in the year 1997. Unlike the relative bias in which a clearly dominant trend in the data was not evident, a clear trend was observed for the nonsystematic error as shown in Figure 9-2(b). This result suggests that the nonsystematic error is more sensitive to the effect of stationary land use than the systematic error.

As with the relative bias, the magnitudes of the relative standard error ratio shown in Figure 9-2(b) seem small (less than 0.10) for predictions within ± 5 years from the year of origin of the stationary land use. In contrast, the error beyond the 5-yr range was as much as 0.29. The overall effect of the stationary land-use on the accuracy of the predicted discharges seems of moderate importance since the calculated values of S_e/S_y were below 0.3.

Table 9-4. **Relative standard error ratio of the predicted daily discharges using stationary land use.**

	Year of source of the stationary land use				
Year	**1984**	**1987**	**1991**	**1993**	**1995**
1984	**0.000**	0.090	0.152	0.188	0.229
1985	0.063	0.011	0.063	0.092	0.126
1986	0.088	0.005	0.070	0.107	0.149
1987	0.097	**0.000**	0.068	0.107	0.151
1988	0.100	0.009	0.054	0.091	0.132
1989	0.097	0.009	0.053	0.087	0.125
1990	0.140	0.043	0.024	0.063	0.106
1991	0.146	0.059	**0.005**	0.036	0.075
1992	0.171	0.073	0.005	0.034	0.078
1993	0.178	0.093	0.034	**0.002**	0.038
1994	0.204	0.115	0.053	0.018	0.022
1995	0.273	0.166	0.091	0.048	0.002
1996	0.223	0.140	0.083	0.050	0.012
1997	0.295	0.193	0.122	0.081	0.036
Average	**0.148**	**0.072**	**0.063**	**0.072**	**0.092**

The effect of stationary land use on the accuracy of predicted discharges is only one of the multiple sources of error within a calibration. The year-to-year analysis, excluding the start-up period (1984) and the effect of outliers in the calculated error, suggests that, the best predictions were achieved when the land use conditions used to make predictions was at the center of the record. For the analysis, the year 1991 would be the most central year of the 14-yr period. Using the 1991 stationary land-use data, the predictions yielded the lowest mean relative standard error ratio of $S_e / S_y = 0.063$ (Table 9-4). The poorest accuracy was obtained for calibrations using stationary land-use data from the extreme years most distant from the center or the record. Although the results suggest that the stationary land use may not be a very important factor in the accuracy of the predictions, the results may be remarkably different if the forested land use is transformed into a more impervious urban area.

9.3 ACCURACY OF THE BASEFLOW AND PEAK FLOW COMPONENTS

The assessment of the accuracy of predicted flow components is probably as important as the assessment of the predicted total discharge accuracy. It was expected that the increment of urbanization would decrease the proportion of baseflow and increase the proportion of surface flow as shown in Figure 9-3. Given that this was a watershed with a predominant baseflow component, it was expected that the effect of stationary land use was more noticeable in the baseflow than in the peak flow predictions. The relative bias of the 20 lowest predicted discharges (Table 9-5), referred to as baseflow and the relative bias of the 20 largest predicted discharges, referred to as peak flow (Table 9-6) were calculated. The average of the yearly value of the relative bias and the relative standard error were used to draw the conclusions about the best accuracy of the predictions.

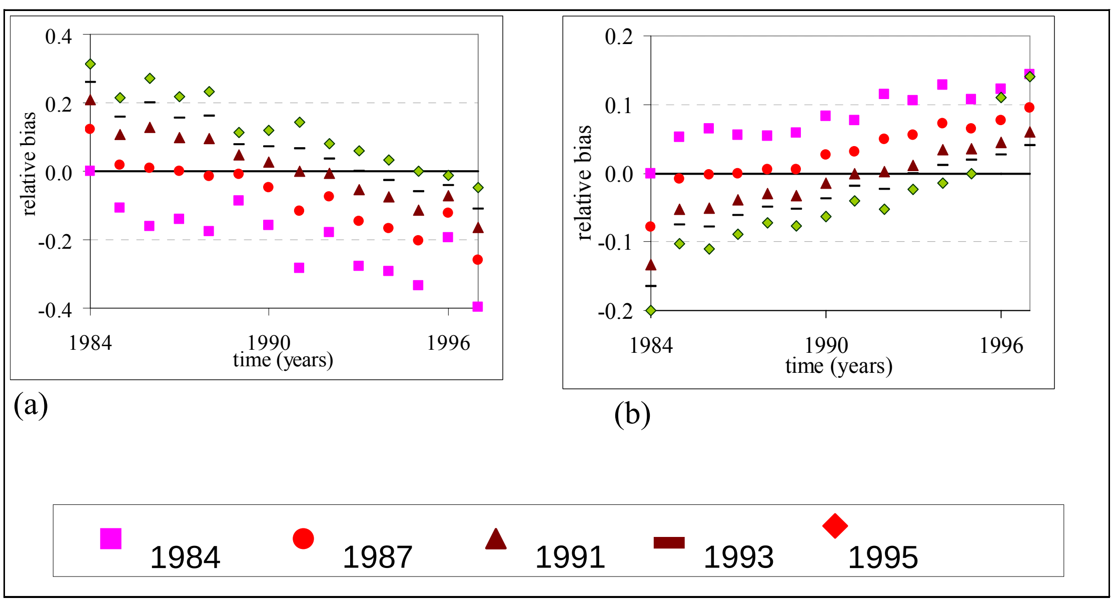


Figure 9-3. **Relative bias for (a) the predicted baseflow component and (b) the predicted peak flow component when using stationary land-use data**

Stationary land use was found to affect the trend of the systematic error of both, baseflow and peak flow. It is important to note the difference in the sample size used to calculate the goodness-of-fit statistics of the total discharge and the flow components. While the sample size of calculations for the daily discharge was of 365, the sample size to determine the systematic error in the baseflow and peak flows was of 20. Thus, it was not expected to observe a direct relation between the systematic error of the total discharge and the systematic error of the flow components.

Table 9-5. Relative biases using the 20 lowest predicted discharges (baseflow)

	Year of source of the stationary land use				
Year	1984	1987	1991	1993	1995
1984	0.000	0.123	0.209	0.259	0.314
1985	-0.107	0.019	0.107	0.158	0.214
1986	-0.161	0.009	0.128	0.200	0.272
1987	-0.141	0.000	0.099	0.155	0.219
1988	-0.177	-0.016	0.097	0.161	0.234
1989	-0.086	-0.008	0.047	0.078	0.114
1990	-0.157	-0.049	0.027	0.071	0.119
1991	-0.283	-0.117	0.000	0.067	0.142
1992	-0.178	-0.076	-0.005	0.036	0.081
1993	-0.278	-0.146	-0.053	0.000	0.060
1994	-0.293	-0.166	-0.076	-0.026	0.032
1995	-0.335	-0.204	-0.112	-0.059	0.000
1996	-0.195	-0.123	-0.072	-0.043	-0.011
1997	-0.397	-0.260	-0.164	-0.109	-0.048
Average	**-0.199**	**-0.072**	**0.017**	**0.068**	**0.124**

The systematic errors of the baseflow predictions were high (up to 40% relative bias) and the trend of the error increased as the distance between the stationary land-use data and the year of the predictions increased. A 40% error is very significant from the hydrologic perspective. The pattern of the predicted baseflow when using individual stationary land-use data (Figure 9-3(a)) follows the pattern of the forest land in Figure 9-1. This trend corroborates the findings from other hydrologic studies showing that when the watershed undergoes urbanization the hydrologic regime changes (Leopold, 1968), with a decrease in baseflow. Conversely, the trend of the systematic error of the peak flows (Figure 9-3(b)) followed the trend of urban land in Figure 9-1; increasing as a function of the urbanization. Because of the predominant baseflow in the hypothetical watershed, the magnitude of the largest error in the predicted peak flow was lower (20%) than the error in baseflow (40%), however, a 20% error is significant from the hydrologic perspective. As in the analysis of the total predicted flow, the best predictions were achieved when the land use data used to make predictions were at the center of the record. For the analysis, the year 1991 would be the most central year of the 14-yr period.

Table 9-6. Relative biases using the 20 largest predicted discharges (peak-flow)

	Year of source of the stationary land use				
Year	1984	1987	1991	1993	1995
1984	0.000	-0.078	-0.133	-0.165	-0.200
1985	0.052	-0.009	-0.052	-0.076	-0.103
1986	0.065	-0.003	-0.051	-0.079	-0.110
1987	0.055	0.000	-0.039	-0.062	-0.089
1988	0.054	0.005	-0.030	-0.050	-0.072
1989	0.059	0.005	-0.032	-0.053	-0.077
1990	0.083	0.026	-0.014	-0.037	-0.063
1991	0.077	0.031	-0.001	-0.019	-0.040
1992	0.115	0.049	0.003	-0.023	-0.053
1993	0.106	0.055	0.012	-0.000	-0.023
1994	0.129	0.073	0.034	0.011	-0.014
1995	0.107	0.065	0.036	0.019	0.000
1996	0.123	0.077	0.045	0.027	0.111
1997	0.144	0.095	0.060	0.040	0.141
Average	**0.084**	**0.028**	**-0.012**	**-0.033**	**-0.042**

Although the effect of nonstationary land use may seem of moderate importance to prediction accuracy when evaluating the total predicted runoff, the analyses of baseflow and peak flow establish that hydrologically significant inaccuracies are introduced to the discharge predictions when using stationary land use. For calibrations using the standard version of the HSPF which assumes stationary land use, it is recommended to choose land use data from the middle of the period of record. If the effect of land use change in the watershed is expected to be significant, then a sensitivity study should be made to assess the potential effect of using stationary land use on the accuracy of the predicted discharges.

CHAPTER 10

EFFECT OF LAND USE NONSPATIAL DISTRIBUTION ON PREDICTION ACCURACY

10.1 INTRODUCTION

The spatial distribution of land use within a watershed influences the characteristics of the measured discharge. In an attempt to model the hydrologic response of a watershed, models are based on simplifying assumptions. For example, some applications of the HSPF model involve nonhomogeneous watersheds in which the modeling of the channel transport is omitted. This prevents from being routed in a channel and evaluating cases that involve tidal influence. Because the HSPF model includes only unidirectional water movement, the modeling of channel routing in tidal influenced channels is not possible. In these cases, the processes that occur in the channel are ignored and the watershed model is based on the assumption that the discharges from the land are directly discharge to the receiving body of water.

A model that does not allow for land use spatial variation is referred to as a lumped model. The HSPF is applied as a lumped model because it calculates runoff from a 1-unit area of land and multiplies it by the number of land units for that land use. Then, the runoff is summed over all land uses to compute the total depth of watershed runoff. When the spatial distribution of land use is nonhomogeneous, the use of a lumped model may lead to inaccurate predictions.

The modeling of a watershed with HSPF generally involves subdividing the watershed into subareas, which are referred to as model segments. A model segment is a subdivision of the larger watershed, and it is commonly defined as an area with similar hydrologic characteristics. When channel routing is not used, then the parameter values from calibrated model segments are used as the parameter values for uncalibrated subwatersheds. The similarity of the land use proportions or the similarity of soil types between the uncalibrated and the calibrated subwatersheds is commonly used to decide whether or not parameters values can be transferred. However, differences in the spatial distribution of land use and the effect of transferring parameters for a calibrated subwatershed to an uncalibrated subwatershed are in most cases not considered. In reality, the spatial distribution may be just as important to achieving accurate discharge predictions as the land use proportions.

When the land use in a watershed is spatially nonhomogeneous, calibration of a spatially lumped model may distort the calibrated parameters. Ultimately, this would introduce error variation, both systematic and nonsystematic, into predictions. Analyses were made to assess the effect of spatial nonhomogeneity on the accuracy of the HSPF predicted discharge rates. Specifically, the goal was to determine if a nonspatial land use distribution is an important criterion for deciding whether or not to transfer parameters. Two objectives were formulated to assess this goal: (1) to determine if calibration of a lumped model with discharge from a watershed with significant channel processes causes inaccurate predictions and (2) to assess if calibration of a lumped model with discharge from a watershed with a spatial nonhomogeneity of land use introduces inaccuracy in the calibrated parameters and model predictions.

10.1.1 Data and Methods

The effect of spatial nonhomogeneity of land use on the accuracy of the predicted discharges would be difficult to assess using actual data because the degree of nonhomogeneity could not be controlled. However, the use of hypothetical data allows control of the generated data and enables the assessment of the potential effects of spatial nonhomogeneity of land use on the accuracy of model predictions. Fourteen years of measured precipitation were used to reflect actual storm sequences in the generated runoff.

A version of a watershed model (McCuen and Snyder, 1986), referred to as the SUBOPT model, was used to simulate daily discharge from a watershed without a channel system. The generated discharge was assumed to be from a hypothetical 20 mi2 forested watershed with a stationary land use; it will be referred to as the lumped-measured runoff.

To simulate the daily discharge from a watershed with distributed land use, a modified version of the SUBOPT model was used to allow for the systematic spatial variation of land use. The revised model used the convex method for routing discharge through channels. The generated daily discharge was assumed to be from a hypothetical watershed divided into four subareas of 5 mi2 each; a single stationary land use was assumed for the individual subareas. Two classes of land use were simulated with the distributed SUBOPT model: forest (F) and urban (U). The generated discharge will be referred to as the distributed-measured runoff. The components of the generated discharge are shown in Table 10-1.

To attain the first objective, a homogeneous forested watershed was assumed and two 14-yr discharge time series were generated: (1) using the lumped version of SUBOPT and (2) using the distributed version of SUBOPT. To generate both the lumped-measured discharge and the distributed-measured runoff, the SUBOPT parameter PINF was set to 0.95 for the forested area. PINF represents the proportion of rainfall that infiltrates, while the remaining proportion (1-PINF) represents the proportion of rainfall that is direct surface runoff.

For the second objective, the modified SUBOPT model was used to generate daily discharge from a distributed watershed composed of 50% forest (F) and 50% urban (U) areas. The parameter PINF was set to 0.95 for the forested area and to 0.75 for the urban area. The urban and forested categories were assigned to the four sequential subareas to generate the distributed-measured discharge as follows: (1) UUFF, with the urban land at the outlet of the watershed, and (2) FFUU, with a forested land at the outlet of the watershed. The components of the generated discharge are shown in Table 10-1.

The accuracy of the individual calibrations was evaluated for the 30 lowest and 30 largest discharges per year, and for all of the discharges throughout the period of calibration. For all of the calibrations the first year of record was designated as the start-up period, and therefore, the discharges for that period were removed from the calibration process. The accuracy of the HSPF predictions was assessed using the relative bias (R_b) calculated using Eq. 9-1, the relative standard error ratio S_e/S_y using Eq. 9-2, and the rationality of the final parameter values for the 13 years of record.

10.1.2 Hydrograph Separation of the Measured Runoff

The proportions of baseflow and quickflow in the total discharge are a function of various factors including the type and distribution of the land use and the surface area covered by each land use in the watershed. Having estimates of the separation of the total flow prior to calibration was important because the flow proportions were used to set the weights w_i of Eq. 6-1. These weights influence the capability of PEST to find the optimum parameter values. Sloto and Crouse (1996) provided methods of separating quickflow and baseflow from a continuous hydrograph. These methods were incorporated into a program to model the baseflow separation analyses. The program has options for baseflow separation using both the local-minimum method (section 2.4.1) and the sliding-interval method (section 2.4.2), with intervals of 3, 5, 7, 9, 11, 13, 15, and 17 days. The program provides measures of separation accuracy. Selection of the discharge separation method and the number of intervals used to make a separation was based on the highest accuracy for a given calibration.

Table 10-1. Method for the separation of baseflow and quickflow and proportions of the separated runoff.

SUBOPT measured runoff	Land use distribution	Land use @ outlet	Method for baseflow separation	Interval	Baseflow (%)	Quickflow (%)
Lumped	FFFF	Forest	Sliding-interval	7	80	20
Distributed	FFFF	Forest	Sliding-interval	7	90	10
Distributed	UUFF	Urban	Sliding-interval	7	80	20
Distributed	FFUU	Forest	Sliding-interval	7	80	20

10.2 EFFECT OF CHANNEL ROUTING OMISSION ON PREDICTION ACCURACY

The presence of a channel in a watershed influences the characteristics of the runoff. The time of concentration, which is defined as the time required for a particle of water to flow from the uppermost location to the outlet of the watershed, varies as a function of the size and relief of the watershed, as well as the length of the channel. For example, the time of concentration between two similar watersheds will be shorter for the watershed with a dominant channel system than for a watershed with minimal channel processes.

This analysis evaluates whether or not the parameters of the HSPF lumped model are distorted when the lumped model is used to fit a measured flood series generated from a spatially distributed watershed. When a hydrologic model is formulated, it is important not to ignore the modeling of channel transport, as error can be introduced in the model predictions. The following results demonstrate the effect of using a lumped model to represent a watershed that has important channel processes.

10.2.1 Analysis of the Mean Daily Discharge

A comparison of the measured and predicted mean daily discharge provides knowledge on the accuracy of the predicted water balance. In some HSPF applications, the modeling of channel transport is avoided if measured discharge data are not available for calibration or if a portion of the channel is tidal influenced. Two cases were investigated (1) where channel-transport processes that occur in the actual watershed are not considered by the model and (2) where channel transport processes are not important in the actual watershed. For both cases, the predicted mean daily discharge was lower than the mean measured discharge (Table 10-2). However, the underprediction was significantly less for the case where the channel transport processes do not occur in the actual watershed (0.065 and 0.057 for the lumped-measured discharge and the predicted runoff, respectively). The effect of omitting the modeling of the channel process is observed in the underprediction of the runoff.

Table 10-2. **Statistical summary of HSPF predictions assuming 100% forested FFFF watershed, with SUBOPT generated discharge assuming a DISTRIBUTED and LUMPED SUBOPT model. The statistics based on the HSPF predictions are: Mean ($\overline{Y}$), standard deviation (S_Y), relative bias (R_b), the relative standard error ratio (S_e/S_y) of the predicted runoff, and the average annual relative bias using the 30 lowest (R_{bB}) and the 30 largest (R_{bP}) predicted runoffs per year.**

SUBOPT RUNOFF	$\overline{Y}$	S_Y	R_b	S_e/S_y	R_{bB}	R_{bP}
DISTRIBUTED $\overline{X} = 0.093$ $S_X = 0.023$	0.067	0.025	-0.269	1.217	-0.325	-0.265
LUMPED $\overline{X} = 0.065$ $S_X = 0.026$	0.057	0.026	-0.115	0.808	0.077	-0.272

10.2.2 Analysis of the Standard Deviation of the Mean Daily Discharge

The magnitude of the standard deviation is frequently associated with the accuracy of the predicted quickflow attenuation that occurs as the water travels from upslope to downslope areas. Ideally, the variation of the predicted discharges will match that of the measured discharges. Underprediction of the variation would suggests that the predictions are insensitive to variations in the parameters representing the processes related to the quickflow, while overprediction of the variation would suggests that the predictions are overly sensitive to variations in these parameters. For both cases, the results indicated that the HSPF accurately models discharge for both the distributed-measured (0.025) and lumped-measured (0.026) discharge (Table 10-2). As these two standard deviations are nearly identical, it seems that the model accurately reflects the variation of discharge rates.

10.2.3 Analysis of the Relative Bias

The analysis of the relative bias is an important indicator as it measures the systematic error of predictions. However, when using the total runoff, a near-zero relative bias may not be a reliable indication of accuracy as this may be the effect of the summation of overpredicted baseflow and underpredicted quickflow. Consequently, the accuracy of the HSPF predicted discharge was evaluated using the relative bias of the total predicted discharge (R_b), the relative bias of the 30 lowest predicted discharges (assumed to represent the baseflow) (R_{bB}), and the 30 largest predicted discharges (assumed to represent the quickflow) (R_{bP}) per year. The period of independence between the 30 selected flows was set to 5 days. The results presented in Table 10-2 are the average of the annual values over 13 years of the simulated discharges.

The omission of the channel transport processes when modeling a forested watershed with HSPF yielded significant underprediction of the total runoff. The underprediction of the total discharge was 27%. In contrast, when the channel transport was not an important factor in the lumped-measured runoff, the HSPF underprediction of the total runoff, although significant, was only 11%. These values are rational. When a model ignores channel processes, the time required for water to flow from the uppermost part of the watershed to the outlet is greater than for a watershed with channel processes modeled. The slower velocities allow for greater predicted evapotranspiration and thus, greater underprediction of the discharge.

HSPF was used as a lumped model and fitted to the SUBOPT data from a distributed watershed with channel routing. HSPF underpredicted both baseflows by about 32% and peak flows by about 26%. When HSPF was applied to the SUBOPT data for a lumped condition, the relative bias of baseflows was improved, with an overprediction of about 8%; however, the peak flows were still underpredicted by about 27%. The underprediction of peak flows may be related to an inadequate component in the objective function calibrating the peak flows.

Although underprediction is expected whenever the channel routing is ignored for a watershed with distributed land use, the level of underprediction may vary as a function of the physical and hydrogeologic characteristics of the watershed, e.g., land use, slopes, soil types, and others. It is important to note that the method of calibration may also be a factor in the underprediction of the total runoff. The effect of underpredicting stormflow when using PEST may be due to the nature of the least squares method, which tends to calibrate towards the mean values of the total runoff.

10.2.4 Analysis of the Relative Standard Error Ratio

The relative standard error ratio of the predicted discharge is a measure of the nonsystematic error in the predictions; yet, its value is also influenced by the systematic error variation. The values (1.217 and 0.808 for the distributed-measured and lumped-measured runoff, respectively) of the relative standard error ratio were poor in both cases (Table 10-2). The higher value of the relative standard error ratio for the prediction of the distributed-measured discharge suggests that error is introduced into the predictions when channel routing is not performed with the HSPF model. However, the poor values in both calibrations suggest that additional factors other than the modeling of the channel transport influence the prediction accuracy. These additional factors may be related to the components of the objective function and the need for further refinement, the nature of the least squares which tends to calibrate toward the mean values and thus incompatible with the calibration of extreme events, or the lumped nature of the HSPF model.

10.2.5 Analysis of the Parameter Values

An analysis of the final parameter values was made to assess the rationality of the model and to measure the effect on the calibrated parameters when the channel system was ignored. The results of the analysis were expected to support the hypothesis that the parameters of the HSPF lumped model may be distorted when a lumped model is used to fit a flood series from a spatially distributed watershed. The distortion of the parameters was expected to be more noticeable in parameters that control the volume of the discharge rather than in parameters that control the rates at which the discharge was released.

From the sensitivity analysis of a forested watershed in CHAPTER 5, the most important parameters are the basic ground-water recession rate (AGWRC), soil storage parameters (LZSN and UZSN), and lower zone evapotranspiration (LZETP). These were selected to establish the rationality of the analyses. Important parameters were expected to calibrate to different values.

Table 10-3. **Final parameter values of the HSPF model when calibrating to a SUBOPT distributed-measured and a lumped-measured discharge of a forested area (FFFF)**

HSPF Parameter	HSPF Parameters							
	AGWR	LZSN	UZSN	INFIL	NSUR	INTFW	IRC	LZETP
Distributed	0.985	4.689	0.100	0.500	0.300	1.135	0.900	0.010
Lumped	0.984	7.716	0.821	0.283	0.821	0.749	0.813	0.163

The basic groundwater recession rate (AGWRC) reflects the number of days for the groundwater to recede. From the analysis of the physical interpretation of the parameter AGWRC in section 2.3.3 (see Table 2-2), its effect is nonlinear, especially for the larger values near 1.0. An AGWR value of 0.970 represents about 37 days while a value of 0.999 represents 999 days for the water to recede.

In both analyses, distributed and lumped, the recession coefficient calibrated to almost the same value, which suggests that this parameter is not affected by the omission of the channel transport in a forested watershed. These results are rational as the AGWRC coefficient controls the rate at which the water is released from the soil nominal storages and is less sensitive to the volume of water in the system.

The importance of the nominal storages is a function of the watershed flow proportions (section 5.2.2). The calibrated values of the nominal-storage parameters were lower for the distributed model than for the lumped model. These lower values are rational because the model attempts to reduce the potential of evapotranspiration in order to accurately predict the discharge. When the channel processes are ignored by omitting the modeling of the channel transport, the time of concentration increases, and the potential for evapotranspiration also increases because the water stays in the watersheds for a longer period. Thus, omitting channel transport causes lower values of the soil nominal storages.

The calibrated annual values of the parameter LZETP for the forested land use were 0.01 and 0.163 for the distributed and lumped runoff, respectively. LZETP represents an index to the density of deep-rooted vegetation. Thus, the calibrated value for the lumped-measured discharge seems more rational for the forested land. When the channel transport is omitted from the modeling process, the value of the parameter LZETP remains at the lower bound to compensate for the increment in the time of concentration and thus, a longer time for the water to evaporate.

10.3 EFFECT OF NONHOMOGENEITY OF LAND USE ON PREDICTION ACCURACY

Knowledge on the effect of the potential error caused by the simplification of the spatial distribution on the predicted discharge may lead to a better interpretation of the results and a better understanding of the model limitations. This analysis studied the lumped nature of the HSPF model and its effect on the accuracy of predicted discharges. The data generated with SUBOPT was assumed to be from a distributed watershed of 50% forest and 50% urban cover, but with a nonhomogeneous spatial distribution of land use. The SUBOPT generated discharge was considered the measured flow series. Flows for the following two spatial land use distributions were generated: (1) forest, forest, urban, and urban (FFUU) and (2) urban, urban, forest, and forest (UUFF). The land uses of the subareas are listed in sequence from the outlet to the headwaters of the watershed.

When using HSPF, the forest cover was always simulated as a 100% pervious area (P), while the urban cover was modeled as some proportion of pervious and impervious cover (PI). The proportions of the urban category were IPPP (25% impervious and 75% pervious) and IIPP (50% impervious and 50% pervious). Note that the order shown, IPPP or IIPP, does not imply anything about the location of the land use within the watershed because HSPF is being applied as a lumped model.

Table 10-4. Statistical summary of HSPF predictions for a 50% forested and 50% urban watershed, when using SUBOPT generated discharge of a DISTRIBUTED watershed (UUFF and FFUU). Mean of the measured and predicted discharges, respectively ($\overline{X}$ and $\overline{Y}$), standard deviation of the measured and predicted discharge, respectively (S_x and S_Y), relative bias (R_b), and relative standard error ratio (S_e/S_y) of the predicted runoff. Average annual relative bias using the 30 lowest (R_{bB}) and the 30 largest (R_{bP}) predicted runoffs per year. F = HSPF forest areas and U = HSPF urban areas.

Land use spatial distribution	Impervious area in the urban category (%)	Predicted					
		$\overline{Y}$	S_Y	R_b	R_{bB}	R_{bP}	S_e/S_y
UUFF $\overline{X}=0.097$ $S_X=0.033$	25	0.067	0.041	-0.317	-0.421	-0.136	1.213
	50	0.064	0.063	-0.347	-0.475	0.098	1.821
FFUU $\overline{X}=0.097$ $S_X=0.035$	25	0.066	0.041	-0.327	-0.434	-0.466	1.353
	50	0.063	0.061	-0.349	-0.419	-0.552	1.981

10.3.1 Analysis of the Mean Daily Discharge

The mean of the predicted discharge is an indicator of the accuracy of the predicted water volumes. While the measured mean daily discharge (Table 10-4) was the same for all of the land use distributions (0.097 in./day), the predicted mean values varied slightly among the analyses. As in the analysis of the effect of neglecting channel transport (section 10.2), the mean of the predicted discharges was lower than the mean of the measured discharge. Because all the means are essentially identical (Table 10-4), the nonspatial variation is not a factor in the accuracy of the predicted mean of the daily runoff.

10.3.2 Analysis of the Standard Deviation of the Predicted Daily Discharge

The standard deviations of the discharge are better indicators of the effects of the spatial distribution of land use. The standard deviations of the measured discharge (Table 10-4) were slightly different for the two spatial distributions of land use (0.033 and 0.035 for the land use distribution UUFF and FFUU, respectively). These are essentially identical, but the standard deviation for FFUU may be slightly larger because routing of the urban-area discharges may increase the likelihood of the peaks of the urban-area discharges matching in time the peaks of the forested-area discharges. In contrast, the standard deviations of the predicted discharge were higher than the standard deviations of the measured discharge for all of the cases (Table 10-4). For example for the UUFF land use spatial distribution with 25% and 50% imperviousness, the standard deviation of the measured discharge was 0.033 while for the predicted discharge were 0.041 and 0.063, respectively.

Three comparisons can be made, UUFF vs. FFUU, 25% urban vs. 50% urban, and predicted vs. measured runoff. In all cases, the standard deviations of the predicted are greater than those of the measured runoff. This is likely the result of the lumped nature of HSPF. The lumped model will increase the variation in the runoff because discharge from the lower portion of the watershed will not be released from the watershed earlier than that from the upper portion, as it was in the distributed model. The standard deviations of the predicted discharges are very different for the 25% and

50% imperviousness, 0.041 for 25% and 0.62 for 50%. The larger value for the higher degree of imperviousness reflects the flashiness of urban runoff. The standard deviations for UUFF and FFUU are very similar because HSPF is lumped and does not distinguish between the locations of the impervious area.

When using the same amount of imperviousness in the HSPF i.e., 25% (IPPP), to simulate the SUBOPT generated discharge for the two land use distributions (UUFF and FFUU), the value of the standard deviation for the predicted discharge was the same (0.041). These results suggest that the percent of imperviousness used to model the watershed is an important factor on prediction accuracy, even if the nonspatial land use distribution cannot be modeled.

10.3.3 Analysis of the Relative Bias

The systematic error of the model is measured by the relative bias (R_b) of the predicted discharges (Table 10-4).Three measures of the relative bias were computed: the total runoff, the baseflows, and peak flows. Each of these biases were assessed for two HSPF imperviousness category (25% and 50%) and for two SUBOPT land use distributions (FFUU and UUFF). The baseflow biases were all very large, with underpredictions of more than 40%. HSPF consistently underpredicted baseflow regardless of land use because it lacked a channel system. Long-term low flows depended on releases from the UZSN and LZSN, which are generally much lower than the SUBOPT low flows in the channel. Thus, the use of a lumped model to model a distributed watershed will distort computed discharge rates, and likely the parameter values.

The effect of using a lumped model on a distributed watershed is also in evidence from the biases of the peak discharges. The biases of the peak flows are much smaller for UUFF than for FFUU. When distributed data from a watershed with the urban land use near the outlet is modeled with a lumped model, then biases in computed peaks will be much less than for the case where the urban land is in the upper reaches of the watershed. For the results given in Table 10-4, the biases in the peaks averaged -2% for UUFF but -51% for FFUU. Thus, the effect of spatial nonhomogeneity is an important factor in calibrating HSPF, and peak discharges and parameters can be distorted.

10.3.4 Analysis of the Relative Standard Error Ratio

The analysis of the relative standard error ratio is important as it provides information of the nonsystematic error of the predictions and due to the lumped nature of the HSPF, and its inability to model the spatial distribution of land use. The results indicated that the accuracy of all of the predictions was poor as the calculated S_e/S_y were all greater than 1.0 (Table 10-4). However, lower values of S_e/S_y were calculated for the FFUU and UUFF distributions with lower amount of impervious areas (25%) than for distributions with 50% imperviousness. These results suggest that the amount of imperviousness is a factor in prediction accuracy when using HSPF and that better accuracy should be expected when the percent of imperviousness is low.

The effect of using a lumped model on a distributed watershed is also evident from the relative standard error ratios. When comparing UUFF and FFUU with the same amount of imperviousness, the value of the relative standard error ratio is slightly lower for UUFF than for FFUU (1.213 and 1.353, respectively). The lumped model does not distinguish the locations of the impervious area when modeling different distributions of land use. Thus, for the FFUU distribution with the impervious area at the uppermost area of the watershed, the later release of the discharge from the uppermost areas is not modeled. Thus, it is rational that better accuracy is attained when the impervious area is located at the outlet of the watershed.

CHAPTER 11

A CALIBRATION STRATEGY

11.1 INTRODUCTION

Potential solutions to environmental problems are commonly provided through the interpretation of results from mathematical models. Yet, the successful solution will be a function of the reliability of the model predictions, which, in turn, depend on the model accurately representing the system and the modeler's knowledge of the calibration process. The goal of a model calibration is to reduce the uncertainty of the results. Extensive work has been performed to produce HSPF, but little effort has been given to address the issue of HSPF calibration. The calibration strategy specific to the HSPF hydrologic component has been limited to the application of the Expert system software-HXPEXP (Lumb et al, 1994), which provides a set of rules for curve fitting. However, the process of adjusting parameter values for curve fitting is now expedited using automatic calibration methods such as those included in PEST. It is important to note that, although these automatic methods have significantly reduce the time for curve fitting, simultaneously, they have made knowledge of calibration a more critical requirement.

A calibration strategy involves an understanding of the model, the fitting method, and errors that can be introduced through the input data. For example, the selection of hourly-measured precipitation data from a distant gage versus hourly-disaggregated rainfall data from a nearby gage will have significant implications in the final calibration. Therefore, a calibration strategy should be thought of as a comprehensive approach to representing a system, and not just as a set of rules to be applied when curve fitting.

Given the enormous amount of subjectivity from the beginning to the end of the calibration process, the use of a systematic calibration strategy is recommended. The development of a calibration strategy will provide the knowledge that is necessary to produce calibrations that will incorporate hydrologic information within the mathematical computations. Conceptually, a systematic procedure that includes the use of a parameter estimator should be the most effective approach for calibrating the model. However, a number of concerns must be addressed including the interaction between the parameter estimator and the calibrated model, the capability for the modeler to incorporate the hydrologic information into the parameter estimator, and recognition that irrational models can also be the product of successfully-minimized objective functions.

11.2 ELEMENTS OF MODELING

A hydrologic model is, in general, composed of four basic elements: (1) the model, which includes equations that represent the hydrologic processes and the parameters that numerically define the processes at a location; (2) database tables, which numerically describe the study area; (3) the objective function, which defines the degree of agreement between the model predictions and the database tables; and (4) the constraints on the model algorithms, on the input data, and on the objective function.

These elements are the foundation for a calibration strategy. The importance of a calibration strategy lies in the assistance that it provides to the less experienced modeler to produce sensible results from the calibration.

11.2.1 Equations and Constraints of the Model Representing the Hydrologic Processes

To correctly apply a model requires an understanding of the elements of modeling. Simplification of the hydrologic processes represented by the model, data transformations, and assumptions of model linearity in the objective function will influence model predictions. Thus, caution should be exercised when interpreting model predictions. Even if the underlying theory behind the model equations is well described and understood, the challenge for the modeler is to determine the implications of such spatial or temporal simplifications in the calibration process.

For a model to be of practical value, simplifications made when formulating the model and contained in the equations that represent the hydrologic processes are necessary and expected. However, such simplifications will affect the prediction accuracy and the accuracy of the calibrated parameters. For example, the lumped nature of the HSPF enables the model to be applied for the modeling of large areas. Yet, the lack of either the modeling of water transport among the subareas within a model segment or the nonspatial distribution of the simulated land uses will restrict the application of the model to a number of cases.

1. It is recommended that, the model algorithms be evaluated to ensure their appropriateness to address the application under consideration.

For example, a model that emphasizes land surface processes may not be appropriate for use on a channel-process dominated watershed.

11.2.2 Data that Numerically Describe the Study Area and the Model Parameters

Monitoring programs commonly target their budget to the collection of precipitation data, land use data, and discharge. Less importance is generally placed on the collection of local supplementary data that may be important in calibrating individual parameters that represent the hydrologic processes, for example, infiltration rates, rates of evapotranspiration, baseflow recession rates, etc. The large-scale data intended to numerically describe the study area, but it is only part of the data that would be needed to fully describe the relation between the model and the watershed being modeled.

2. It is recommended that, the model algorithms be evaluated to ensure their appropriateness to address the application under consideration.

One approach that can be used to define a possible range of parameter values used as the target for calibration is to gather related information from other hydrologic studies. The accuracy of such data will influence the accuracy of the calibration. The lack of spatial and temporal variability in the available supplementary data imposes constraints to the development of the spatial and temporal distribution of parameters. Knowledge on the potential distributions of the parameters can be use as guidance to set bounds on the parameters during calibration and to evaluate the rationality of the calibrated parameters.

Data records are often short and may not cover the range of conditions that will be experienced in the future. For example, a 10-year record of rainfall may not include a period of drought on a major flood-producing rainfall. The calibrated parameters will reflect this lack of variability.

3. It is recommended that, the characteristics of the data used for calibration be assessed prior to calibration, with the assessment aimed at deciding the extent to which it contains the variability expected in the future.

For example, a model to be used for estimating water pollution concentrations during low flow conditions should have periods of baseflows that are expected in the future. The mean baseflow should be calculated on seasonal and monthly bases to identify possible trends with respect to times at which pollution is observed. Similarly, if flood peaks are of primary interest because of concerns of erosion in the simulated watershed, then the data records should include storms with high rainfall volumes and large peak discharge rates. Summary statistics should be compiled on the data base that determined the concern, and compared to the statistics of the data base that correspond to conditions used in the modeling exercise.

11.2.3 Objective Function that Controls the Calibration Process

The number, structure, and weights of the components of the objective function influence the accuracy of the model predictions. In PEST, for example, the multicriterion objective function consists of individual components that are multiplied by weights. The components should reflect the important components of flow such as baseflow or recession rates and the weights should be used as constants that scale the components. The weights carry valuable information when their respective components are hydrologically related to the specific characteristics of the flow.

4. It is recommended to set the weights, so that the contributions of the components to the objective function reflect the hydrologic characteristics of the watershed.

For example, the PEST objective function should assign greater weight to a baseflow-oriented component if runoff from the watershed is predominantly baseflow.

The analysis of flow proportions of the measured runoff is particularly important when using a parameter estimator, as it enables the modeler to set proper weights for the objective function components.

5. It is recommended to perform a hydrograph separation of the measured discharge, so that the proportions of baseflow and quickflow can be used to set the component contributions of the objective function.

Knowledge of the flow proportions in the calibrated watershed provides a mechanism to ensure that the calibration is not a merely mathematical process but a procedure driven by hydrologic principles.

Research on the use of parameter estimators to calibrate hydrologic models has been focused on the improvement of the gradient-based approaches or on the inclusion of more complex algorithms to reduce the time required to reach the minimum value of the objective function. However, multicomponent objective functions are not widely used even though they have the potential to significantly increase parameter accuracy. Furthermore, minimal research has been undertaken to identify an effective set of objective function components for the calibration of hydrologic processes. The use of improper components in an objective function may constrain the calibration process. For example, if the investigation is related to high concentrations of ammonia during the summer season and it is believed that these high concentrations are related to the low runoff rates during the summer months, then the objective function should include specific components such as a hydrograph separation for the calibration of the baseflow. On the other hand, if the investigation is related to erosion due to intense storms, then the objective function should include specific components for the calibration of stormflows.

6. It is recommended that, the components of the objective function be selected to reflect the important hydrologic criteria and the goals of the specific model application.

The use of an existing objective function should be avoided unless it can be shown that it is totally relevant to the issues being addressed by the analysis.

11.3 EVALUATION OF CALIBRATION ACCURACY

The assessment of a model calibration to replicate field data is a critical step in the modeling process. Visualizations and graphical comparisons of model output are excellent ways to start the assessment; however, more quantitative methods are needed to assess model accuracy. Goodness-of-fit statistics such as the relative bias, the relative standard error ratio, or the coefficient of determination are commonly used to determine prediction accuracy.

The coefficient of determination is an important statistic as it is a measure of prediction accuracy. However, as with any statistic, it is important to understand exactly what the coefficient reveals about prediction accuracy. Important information may be overlooked if too much reliance is placed on the interpretation of the calculated correlation coefficient. In linear models, for example, the sample size and the level of significance influence the sample coefficient and its interpretation. However, in nonlinear models or in models in which time is a variable, the coefficient of determination should be only one of many statistics used for accuracy evaluation. The selection of the goodness-of-fit statistics is important, as each has its limitations and none provides a complete quantification of model accuracy.

7. It is recommended to apply various means of goodness of fit to all parts of the model predictions, especially, to predictions relevant to the issues that underlie the modeling effort.

For example, if baseflow is especially relevant to the project objectives, then goodness-of-fit statistics should be used to measure the capability of the model to predict baseflow discharge rates.

In addition to the use of goodness-of-fit statistics, it is important to determine the rationality of the model using the calibrated parameter values, as they indicate the importance that the model places on the specific hydrologic processes of the watershed being modeled.

8. It is recommended to evaluate the rationality of the model using the calibrated model parameters, with the assessment aimed at establishing a rational relation between the calibrated parameters and the watershed conditions.

For example, AGWRC should reflect the recession rate that is appropriate for the type of soil and land use that exist in the watershed. The values of soil storage parameter should rationally reflect actual soil storages. Similarly, parameters that characterize infiltration and evaporation rates should reflect rates typical for the region, including temporal variations of the processes.

11.3.1 Goodness-of-Fit Statistics

The selection of the goodness-of-fit statistics is important, as each statistic measures specific aspects of prediction accuracy. Failure to understand the effect that a statistical index assesses may lead to both an inaccurate indication of model goodness of fit and a misreading of the quality of the calibration. When time is a factor, important assumptions on which commonly used statistical procedures are based are frequently violated, including normality, randomness, and independence. However, the use of statistics is accepted as valid approximations when evaluating prediction accuracy. Thus, knowledge of and caution in the interpretation of the goodness-of-fit statistics may provide models better able to represent the hydrologic processes.

Bias measures the systematic error in the model predictions and indicates if the model incorrectly adds or depletes water from storages over the period of calibration.

9. It is recommended that, the bias be calculated to assess the extent of under or overprediction on both seasonal and annual bases.

For example, the annual bias may have a value near zero, which would indicate an accurate annual mass balance. However, a low value may be the summation of large errors that occur seasonally. With the calculation of seasonal errors, the modeler may target specific processes, parameters, or data that occur during the problematic period to improve the accuracy of the overall prediction. Such biases are considered local biases over the duration of the time series.

When a regional hydrologic model is developed, it is important to define rational criteria for transferring parameter values from calibrated to uncalibrated watersheds. In addition to providing information on the systematic error of a calibration, the relative bias (Eq. 5-3) will allow for a comparison of accuracy between calibrations. The comparison of the relative bias from calibrations of different subareas can be used to determine if local factors influence the accuracy of model predictions.

> **10.** It is recommended that, the relative bias be calculated when defining criteria for parameter transference from calibrated to uncalibrated watersheds.

For example, the spatial distribution of land use may be an important factor. If the relative bias was significantly different for two neighboring watersheds of similar size, similar slopes, and land-use proportions, but different land-use-spatial distributions, then this last aspect should be taken into consideration when transferring parameters. In fact, it may suggest that the parameters should not be transferred to the other watershed.

The relative bias can also provide information on the accuracy of a mass balance, specifically to the accurate predictions of the baseflow and quickflow components. However in this case, it is important to include the evaluation of the bias and the mean of the flow components to determine the rationality of the relative bias of the flow component and the rationality of the relative bias of the total flow.

> **11.** It is recommended that, separate evaluations of the bias and the relative bias be made for the prediction accuracy of baseflow and peak flows.

For example, the relative bias of baseflow discharges may be large solely because of a low mean baseflow, suggesting that unless the model is intended to address a particular problem that occurs during baseflow, the focus of attaining accurate predictions should be directed to the accurate calibration of the quickflow component.

Although the bias and the relative bias are valuable indicators of prediction accuracy, these statistics cannot measure other important elements of goodness of fit. For example, they are insensitive to the accuracy of model predictions related to the timing at which the storms occur. In this case, the standard error of the estimate could be used to determine the precision of the predicted runoff. This goodness-of-fit statistic is computed from the variance of the predicted discharges, and it indicates the precision with which the HSPF model estimates the value of the dependent variable.

In regression models, the standard error of the estimate is used to compute confidence intervals of the parameters or prediction intervals. However, in models where time is involved, the calculation of confidence intervals using the standard error of the estimate is commonly avoided as important properties underlying the rationality of the statistic are violated. For example, the effective sample size may be significantly smaller than the number of measured discharges because of the serial correlation between temporally adjacent discharge values. Nevertheless, the standard error of the estimate can be used in conjunction with the standard deviation of the measured data to compute the relative standard error ratio that measures the nonsystematic error variation.

The standard deviation is a measure of the variation of the data, and it is an important statistic when defining criteria for transferring parameter values from calibrated to uncalibrated watersheds. When the standard deviation of the predicted discharges is much smaller than that of the measured values, it indicates that the model does not adequately reflect the hydrologic processes of the watershed. If the variation of the predicted discharges is much greater than that of the measured discharges, then the functional form of the model coefficients are likely inappropriate for the modeled watershed.

12. It is recommended that, the standard deviations of the predicted and measured discharges, both total flow and flow components be computed, and reasons for differences identified.

For example, the lumped nature of the HSPF model may be a constraint to predict similar standard deviations for the predicted and measured runoff for certain nonhomogeneous watersheds. In this case, the user should question the application of the model and not necessarily the accuracy of the model predictions.

The relative standard error ratio is a more useful single measure of the prediction capability of a model. This goodness-of-fit statistic is computed by dividing the standard error of estimate by the standard deviation of the measured data. This standardized statistic measures prediction accuracy, specifically the nonsystematic error of the predictions.

13. It is recommended that, the relative standard error ratio of the predicted discharges be computed and reasons for its high or low value identified.

For example, in analyses of sensitivity, this goodness-of-fit statistic may be used as an indicator of parameter importance. When changes are applied to individual parameters and the relative standard error ratio for predicted discharges is calculated, higher values of the statistic would suggest greater parameter importance. In other cases, the user may want to test the effect of using data from two nearby sites measuring precipitation on prediction accuracy. The calibration with the lower relative standard error will suggests that the precipitation data from this site better represents the conditions that occur in the watershed. Conversely, if the value of the relative standard error ratio is similar in both calibrations, then the user may conclude that the prediction accuracy are insensitive to the characteristics of the precipitation contained in both datasets.

11.3.2 Rationality of the Important Calibrated Parameters

The rationality of the calibrated parameters is an important indication of accuracy in model calibration. The goal in any model calibration is to reduce the uncertainty of the predictions and to estimate parameter values that can be directly related to the physical properties of the system and that correctly reproduce the water balance and flow components of runoff. Although model parameters can be conceptually related to the physical processes, fitted values are subject to considerable uncertainty because of interactions between physical processes and the inability of the model to mimic the transformation of rainfall to runoff. In addition, important parameters are expected to approach their true values while unimportant parameters may take any value during the calibration process.

14. It is recommended that, a sensitivity analysis be made after the model is calibrated, so that the important parameters can be identified.

Parameters identified as important in the sensitivity analysis should be assessed for rationality. However, it may not be possible to assess the rationality of the less important parameters.

An influential factor in the importance of the parameters is the climatic condition during the period of record. The importance of the parameters can vary significantly from season to season and the variation is often associated with the variation in climatic conditions such as rainfall depths and temperature. When snow is an important element of the hydrologic cycle, the parameters that control the calibration of snow may be more important during the winter months than the parameters that control the hydrologic processes during the rest of the year.

> **15.** It is recommended that, when snow is an important element of the hydrologic cycle, a sensitivity analysis be made so that the importance of the parameters that control the prediction of snow can be identified.

For example, if important parameters that represent hydrologic processes related to snow were not calibrated, then inaccurate snow depths could be predicted by the HSPF model. The form of precipitation (snow or liquid) is determined by the HSPF model based on meteorologic data. These inaccuracies would then be reflected in inaccurate runoff predictions and inaccurate parameters that control the calibration of the hydrologic processes. In such cases, the calibration of parameters that control the modeling of snow may be necessary to guarantee the correct form of precipitation. Data such as snow depths or the depth of a snow pack may be used for the calibration of these parameters.

The rationality of the parameter AGWRC, which represents the number of days for groundwater to recede, can be determined through a comparison of the recession rates computed from the measured and the predicted runoff. The allowable minimum and maximum values recommended in the HSPF manual for the parameter AGWRC are between 0.001 and 0.999. For values between 0.001 and 0.970 the groundwater recession varies between hours to days, while for values greater than 0.970 and below 0.999, the groundwater recession varies between months to years (Table 2-**2**).

16. It is recommended that prior to the calibration and through examination of the measured runoff, to determine a rational range of values for the parameter AGWRC.

For example, the number of days between the largest values of the baseflow during the spring time and the lowest value of baseflow during the summer time can be calculated. This information may be used not only to set the upper and lower bounds of AGWRC during the calibration, but also to determine the rationality of the calibrated parameter. The similarity of the estimated and calibrated values should be an indication of parameter rationality.

Although determining the rationality of the soil nominal storage parameters (LZSN and UZSN) is difficult because of the enormous variation in the soil, a simple calculation using the soil porosity can be performed to assess parameter rationality. The calibrated values of LZSN and UZSN represent the water holding capacity of the soil storages, defined in terms of nominal capacities rather than absolute. From local studies, the modeler may estimate a depth of soil; the depth may be to a confined layer, a ground-water table, or to a depth that will release water to baseflow. This estimated depth may be multiplied by an assumed value of the soil porosity in the area to provide potential values of the parameter values (UZSN or LZSN). This information will be compared to the calibrated parameters.

17. It is recommended to determine prior to the calibration, a set of potential values of the parameters UZSN and LZSN using estimated soil porosity, and multiplying this value by an estimated depth to a confined layer.

For example, if the estimated depth of soil is 50 in. and the estimated average porosity is 0.3 (30%), then the value of the parameter LZSN should be near 50*0.3= 15 in.

11.4 SELECTION OF DATA

The preparation of the data is perhaps one of the most important phases of modeling. The accuracy of model predictions will depend in part on having accurate information as input and the extent to which this information complies with the model constraints. Four elements are of special importance for a successful hydrologic calibration and are individually addressed in this section: (1) the collection of data related to the model parameters from the literature and other regional studies, (2) measured discharge data, (3) precipitation data, and (4) land use data.

11.4.1 Data Collection from Literature and Regional Studies

When using a parameter estimator or using subjective calibration, information about the distributions of the parameters that represent the hydrologic processes in the studied watershed is essential to a successful calibration. Local soil scientists should be consulted for information about field tests or estimations that they consider appropriate in the area. This information may provide a range of possible values or an estimate of the distribution of the parameter to which the parameter could be fitted.

18. It is recommended to request information from a soil scientist knowledgeable of the area where HSPF is applied, about possible distribution of parameters representing the hydrologic processes, and use this information to set the lower and upper bounds of the parameter.

When calibrations are compared, the significance of differences between calibrated values of the same parameter can also be determined if information on the potential distribution of the parameter is available. This practice will reduce the potential for irrationality in the calibrated parameter values.

The reliability of parameter values that are transferred from other watershed studies to a new, ungaged watershed will depend, in part, on the similarity between the site of interest and the regional analyses. Of special importance, the spatial distribution of land use in the two watersheds must be similar. It is not just the proportions of the land use in a nonhomogeneous watershed but the spatial location of the different land uses.

19. It is recommended that, the spatial distribution of the land use be included as criteria for transferring parameter values, with the assessment aimed at establishing the extent of similarity between the watersheds of interest.

For example, runoff from a watershed with impervious cover near the outlet will be flashier than runoff from a watershed with the same amount of impervious cover that is located near the watershed divide. The difference in the dispersion of runoff will be observed in the calibrated parameters. Calibrated parameter values for a watershed with flashy runoff will likely suggest low infiltration and low values of retention storage. Conversely, runoff from a watershed that is forested at the outlet would be quite different than that for the case where the impervious land use is near the outlet. If the parameters from a watershed with an impervious area near the outlet were to be transferred to a watershed with similar land-use proportions, but with a forested area located at the outlet of the watershed, predicted runoff rates would quite likely be erroneous.

Prediction accuracy will also depend on the similarity in climatic conditions between the regional watersheds and the watershed of interest. Rainfall, snowfall, and evapotranspiration amounts are important factors in water balance models such as HSPF. Therefore, where climatic conditions show considerable spatial variation, regionalized parameters should be carefully screened to ensure applicability.

20. It is recommended that, when selecting parameter values from regional studies, more importance be assigned to those where the site of the study is similar to the watershed being calibrated.

Parameter values obtained from other regions may be of little value because meteorologic and hydrologic conditions, as well as variations in the geologic setting, are likely very different. Their use would likely introduce bias into predicted discharges.

11.4.2 Precipitation Data

The HSPF algorithms that represent the physical processes of the hydrologic cycle make the discharge predictions extremely sensitive to the quality of the input precipitation. The selection of the precipitation data for the HSPF model is an important consideration because prediction accuracy depends on the spatial and temporal characteristics of the available data.

21. It is recommended that, the precipitation data be selected on the basis that the data represent the climatic conditions experienced in the watershed during the period of calibration.

Measured rainfall depths within the watershed, close to the watershed centroid, and at an hourly time step would be the ideal input. However, the modeler may face the decision of choosing between data from a gage that collects daily rainfall near the watershed outlet and data from a gage measuring hourly rainfall from a more distant place. Both options have both downfalls and benefits. When daily data are selected, the modeler will need to disaggregate the data, which adds uncertainty to the calibration; by choosing the hourly data, the modeler may have to deal with storms or dry periods in the rainfall data that were not experienced in the watershed and, therefore, not reflected in the runoff. Thus, the rainfall will appear mismatched to the runoff data.

Common sense advocates for the use of rainfall data at a small time scale measured at a gage located within the watershed, or at least from a nearby location. The question is: Would data from an hourly gage 2 miles away from the outlet be better than data from a daily gage 1 mile away?

22. It is recommended that, a comparison of the time sequences of rainfall and runoff be performed to ensure a reasonable degree of cross correlation between rainfall and runoff.

If the storm periods in the hourly rainfall show reasonable agreement with storm periods in the runoff data, then it may be reasonable to use the hourly data. If the cross correlation is poor, then the similarity between the daily rainfall and aggregated daily runoff should be checked for cross correlation.

The argument favoring the use of the hourly measured rainfall data over the disaggregated daily rainfall may be based on the assumption that the hourly information captures the timing and the intensity of actual storms near the outlet. This is important because the parameter values reflect processes that function on a small time step. The timing of the storms is the most difficult aspect to replicate, as indicated by the analyses of rainfall disaggregation. Accurate intensities and timing of storms will allow for a better transformation of rainfall to runoff within the model.

23. It is recommended that, cross-correlation analyses between rainfall and runoff should be made prior to calibration, with the intent of identifying both time-offset errors in the similarity of rainfall and streamflow patterns.

The concern of using disaggregated daily rainfall depths is related to the appropriateness of the method of disaggregation that significantly influences the accuracy of the predicted runoff.

The rationale to use disaggregated daily values rather than hourly measured data may be based on the similarity of regional characteristics between the parameters that represent the hydrologic processes and the meteorologic conditions contained in the disaggregated rainfall depths.

24. It is recommended that, the disaggregation of daily rainfall depths be based on storm frequency analyses of regional and local data, avoiding disaggregation over a 24-hr period.

For example, monthly or seasonal analyses of regional precipitation data, i.e., using a depth-duration-dependent method, may provide sufficient information to establish the time, the volume, and the duration of average storms. This short time step ensures that the characteristics of short and intense storms that occur during the summer months, and that the characteristics of longer but less intense storms during the winter months, will be transferred to the disaggregated values.

The method used to disaggregate daily rainfall depths into hourly depths can significantly influence the accuracy of the predicted runoff. From analyses on the effect of rainfall disaggregation on the accuracy of the predicted runoff (see CHAPTER 4), the depth-duration-dependent method provided the highest accuracy.

25. It is recommended that, the disaggregation of daily rainfall depths be based on storm frequency analyses, rather than disaggregating using the rainfall proportions from a single gage.

Regional information will make use of the broad-spectrum of meteorologic conditions that occur in a region.

Extreme rainfall events or outliers that occur during the first year of the calibration and inaccuracies in the estimation of initial soil storages adversely affect the calibration process and hinder the calibration of the parameters. The presence of outliers during the first months of calibration was noted to have an adverse effect on the calibration process when the start-up period was not included in the period of record (see CHAPTER 8). From these analyses, the start-up period was designated as the first year of period of record, so that the period of record was always one year longer than the period of calibration.

26. It is recommended to provide a start-up period to all HSPF applications to reduce the uncertainty of the predictions.

Although the presence of outliers affects the overall calibration process, the combination of error in the estimates of the soil storages and the presence of outliers during the first months of the period of record hampers the calibration of the parameters.

If the length of record of the calibration data is sufficiently long, then the first year should be used as the start-up period, and the period of record should be deferred to start of the second year of the available data. The predicted discharges from the start-up period should not be included in the calibration process or in the evaluation of the model performance. Conversely, if the period of record is short but the meteorologic data for the year prior to the calibration are available, then the period of record should be extended so that the year prior to the calibration can be used as the start-up period. When the meteorologic data for the year prior to the calibration are not available, one alternative, not investigated in this report, would be to place a copy of the available meteorologic record at the start of the record. This would create a record of length 2_n from an actual record of length n. Then the copy of the actual record of length n serves as the start-up time. The predictions during the start-up time would not be used for either calibration or the analysis of the goodness of fit. A potential problem with this approach would be the discontinuity that would exist at the end of the start-up period.

11.4.3 Land Use Data

Land use data is an important factor in the accuracy of the HSPF calibrations. The standard version of the HSPF model assumes stationary land use. If a watershed is undergoing land use change with time, a limit is imposed on the length of the record.

Use of the standard version for longer record lengths can significantly reduce prediction accuracy, with the extent of inaccuracy depending on the nature of the change in land use distribution.

When using stationary land use data for a watershed that is undergoing land use change, the acreage values for years other than the year of the land use source data will lead to errors in the predicted discharges. For example, a steady decline of forested area in a watershed will introduce a systematic error when stationary land use data is used. The inaccurate number of acres of the simulated land uses would affect the predicted runoff for years other than the year of the land use source. The error would increase as the length of time between the year of the land use source data and the year of the predictions is increased.

> **27.** It is recommended to assess the land use change for the period of calibration, with the analysis aimed at determining the potential to under or overpredict watershed runoff, and the potential distortion of the calibrated parameters due to the use of stationary land use.

For example, the analyses in CHAPTER 9 indicated that the systematic error in the predictions of ± 5 years from the year of the land use source was less than 30% when the land use was changed from a forested to a pervious urban area. However, the results may be remarkably different if the forested land use is transformed into an impervious urban area or if the rate of land-use change is greater.

11.4.4 Measured Discharge

Graphical analyses of the measured runoff will assist the modeler in the detection of seasonal trends. This information may be used to determine an appropriate number of parameters to calibrate. The calibration of monthly parameters may increase the detail in the simulated hydrologic process, and thus, increase the accuracy of the predicted discharges. However, caution may also be suggested, as intercorrelation among parameters increases with increases in the number of parameters. Also, the increase in the number of parameters may increase the prediction accuracy, although the increase may not be significant and may also lead to irrational parameters.

28. It is recommended that, the number of parameters to calibrate be based on the rationality of the simulated processes rather than the need for curve fitting.

For example, the calibration of the monthly interception storage capacity for the urban land use increases the possibility for curve fitting, but calibrated parameters with significant variation throughout the year may not represent the process being modeled.

Systematic variations in measured discharge time series will significantly influence the selection of single annual values or parameters varying monthly for calibration. Discharge records with little variation would place emphasis on parameter calibrated as single annual values. Any systematic variations should be recognized prior to fitting.

29. It is recommended that, the measured discharge time series be plotted to determine trends or seasonality in the runoff.

For example, if a graphical assessment shows that seasonality is an important characteristic of the discharge time series, then a discrete mass function (such as Poisson) could be fitted to the monthly HSPF parameters, such as those that control ET, to reflect seasonality. This should improve goodness of fit and yield more rational parameter values.

CHAPTER 12

CONCLUSIONS

12.1 INTRODUCTION

The successful development of a calibration strategy capable of accomplishing the objectives of this research was provided through the several analyses. The calibration strategy was developed using data and model assumptions commonly used by the HSPF modeling community and through the hydrologic and mathematic interpretation of the results. The use of a model-independent parameter estimator facilitated the development of the calibration strategy and allowed for the replication of results and the removal of subjectivity from the calibration approach. Throughout the interpretation of the analyses, it was demonstrated that the mathematical evaluation of the model is only part of the calibration process, and that knowledge on the basic concepts of hydrology, modeling, and statistics are necessary for the accurate application of the model. The contribution of this research lies in the concurrence of these concepts into the hydrologic interpretation of the modeling, the demonstration of potential problems caused by assumptions of the model, and the implications of such assumptions in management decisions. When using subjective calibration several factors influence the uncertainty in the calibrations: (1) the quality and availability of the input data; (2) the mathematical formulations can sometimes be extremely complex in describing a single process, which causes parameter intercorrelation; (3) the formulations are typically applied to very large areas, while calculations are made for one acre in the HSPF model; and (4) the fitting of model parameters using a subjective calibration yields parameter values that are not reproducible and subject to the bias of the modeler.

12.2 RAINFALL DISAGGREGATION

Accurate estimates of sediment and pollutant loads depend both on accurate rainfall intensities for short time increments and on a calibrated model. The ideal case is where measured rainfall, runoff, and pollutant concentrations are available on site at a short time increment, such as an hour. Hourly rainfall data are unlikely to be measured on site, so it would be necessary to transfer the hourly depths from a nearby gage in order to calibrate the model, and then use the model to predict erosion rates and pollutant concentrations. The intent was to assess the extent to which daily measured rainfall could be disaggregated, or hourly values transferred from a gaged site to an ungaged site. The alternative case of transferring hourly rainfall data from a nearby gage solely for predicting was not evaluated, as this is generally recognized as being inaccurate but often the only recourse.

In spite of the extensive database used in the analyses, the results were discouraging for those who require accurate estimates of hourly rainfall intensities. However, the results do provide a valuable indication of the potential bias and inaccuracy of rainfall depths disaggregated from daily values or transferred from a nearby hourly rain gage, and the potential implications of this inaccuracy on watershed model calibrations. The lack of representativeness of point rainfall was also shown by Huff and Neill (1957) even for larger time intervals of weekly and monthly data. The cross correlation of gages within 8 kilometers of each other did not reveal a consistent non-zero time lag in the measured hourly depths. Thus, a lag of 0 was used in making all comparisons. The methods that were used to disaggregate measured daily depths at a base station (Y) yield poor accuracy, even though they provide unbiased estimates. The two univariate methods used herein, i.e., uniform and weather pattern distributions, most likely fail because most storm events are less than 24 hours and the daily record does not give an indication when during the storm duration that the rain occurred. Thus, a new disaggregation method that would allow for distributions of shorter duration may yield more accurate estimates of hourly rainfall depths. The duration of storms could possibly be accurately predicted; however, the start time for the storm cannot be predicted, which contributes to the poor accuracy of the results.

The transfer of hourly information from a nearby satellite gage also failed to provide accurate estimates. Conversely, the transfer of daily rainfall information from a nearby satellite provided considerably greater accuracy, which indicates that daily rainfall totals can be transferred with a greater degree of confidence than hourly values. The transfer of hourly depths was biased, with some biases being highly significant from a hydrologic standpoint. It lacked precision, with the overall accuracy being poor. In fact, univariate disaggregation appeared to provide better results than the two-station transfer methods, but the results were still poor.

The accuracy of daily disaggregation using the univariate method of weather patterns was not better than using the uniform distribution as the relative standard error ratios were almost the same. Both methods smoothed the daily precipitation, thus losing the natural intensity of rainfall. This smoothing would lead to inaccurate predictions of sediment loads in watershed modeling, as sediment loads are highly dependent on rainfall intensity. The smoothing would also produce inaccurate parameter values in any calibration.

The results from the bivariate satellite methods showed a trend of decreasing prediction accuracy as the distance between the stations increased. The advantage of this method was the preservation of rainfall intensity, which is important for erosion and sediment transport modeling. When comparing the bivariate with the univariate methods, the former showed less accuracy than either the weather pattern or the uniform distributions.

12.3 THE EFFECT OF RAINFALL DISAGGREGATION ON PREDICTED DISCHARGES

Accurate predicted discharges in hydrologic models are influenced by the accurate input of rainfall depths. However, accurate values of hourly rainfall are important, only if the structure of the model using the precipitation data and making predictions of runoff is sensitive to the error in the rainfall data. Although many factors influence the accuracy of the predicted runoff in the HSPF, the accuracy of the rainfall depths, the temporal variability of the rainfall, and the rainfall intensity are the essence of accurate runoff predictions. The error of the predicted runoff was greater for hourly predictions than for the predictions of mean daily flow. Three methods of daily rainfall disaggregation were analyzed: (1) a uniform distribution over 24 hours; (2) the SCS 24-hour storm distribution; and (3) a depth-duration-dependent separation. The results indicated that of the three tested methods, the SCS and the depth-duration-dependent disaggregation provided similar accuracy when the predictions were evaluated at the hourly time step. However, when the predictions were evaluated using daily predictions, the depth-duration method provided better accuracy that the SCS method.

The poorest accuracy of hourly runoff was obtained with rainfall data from the 24-hr uniform disaggregation method. The uniform disaggregation smoothed the daily rainfall, which was reflected in lower storm runoff and higher baseflow volumes. During the summer months, sporadic and more intense storms are difficult to predict with a uniform disaggregation method, while in the winter months runoff are more predictable because of the patterns of precipitation.

The accuracy of the predicted runoff improved with the disaggregated precipitation from the SCS method. The results indicated that by using disaggregated rainfall from a method that takes into consideration the duration of the storms, the accuracy of the predicted runoff improved. However, the prediction accuracy varied among seasons. The predictions of runoff during the winter season were the least accurate, followed by the predictions during the fall and spring seasons, and finally the best accuracy was obtained during the summer season. The poor accuracy during the winter months is explained by the lack of seasonality of the SCS method.

In contrast, the accuracy of the hourly predicted runoff with the disaggregated rainfall from the depth-duration method was significantly better that the predictions from any other method; simply because the method was based on storm frequency analysis from the same region from where the disaggregated daily precipitation time series was recorded. The disaggregated data were not included in the development of the distribution method. For all of the seasons except for summer, the relative bias varied between –0.001 and 0.200; the variation during the summer season was between

–0.8 and 0.68. The disparity of values between the relative bias in the summer predictions and the relative bias in the remaining seasons may be explained by the type of precipitation in the area during the summer time. During the summer, thunderstorms are the most common type of precipitation characterized by high intensity and very short periods. These characteristics were lost with the use of annual-average storms used in the depth-duration method. Thus, the lack of seasonality in the development of the method explains the poor prediction accuracy.

A comparison of the relative standard error ratio among the methods indicates that spatial variation is important in the disaggregation and that none of the methods provided accurate predictions, as the disaggregation of daily rainfall into hourly values introduces additional noise and uncertainty to the predictions of hourly runoffs. Although the accuracy of the depth-duration method was poor, as indicated by the relative standard error ratio varying between 0.7 and 0.99, the method provided the best results as the distribution reflects local precipitation patterns.

Except for the predictions using the 24-hr uniform disaggregation, the accuracy of the predicted daily discharges when using the SCS or the depth-duration method was significantly better than the results obtained with the analysis of hourly data. The relative bias for the SCS method was similar to the relative bias obtained in the depth-duration methods and always below a 10% in magnitude. The nonsystematic variation of the predicted daily runoff for predictions using the SCS or the depth-duration disaggregated precipitation was also significantly lower the obtained when using rainfall from the 24-hr uniform method. The accuracy of the prediction in the SCS and the depth-duration methods was moderate.

12.4 THE EFFECT OF FLOW PROPORTIONS ON THE HSPF MODEL CALIBRATION EFFICIENCY

As modeling becomes a larger part of any watershed analysis, models will need to be more complex in structure. Improvements in model complexity will enable a wider array of problems to be addressed with potentially greater accuracy. However, more complex model structures require more sophisticated calibration strategies. Attempts to use traditional model calibration methods may potentially lead to inaccurate calibration solutions and, therefore, subsequent erroneous decisions.

New hydro-environmental problems such as the establishment of TMDLs require more sophisticated models, such as HSPF. Unfortunately, the methods of calibrating the more complex models have not kept pace with model development. The analyses presented herein, specifically for HSPF, provide a method to improve the process of calibrating complex models, possibly ensuring parameter values that more accurately reflect the watershed processes being modeled. One improvement to the calibration process was the development of a multi-criterion objective function and the systematic approach to set the weights to each of the objective function components when using a parameter estimator. Although the use of parameter estimators adds complexity to the calibration, it simultaneously can increase the accuracy of the predictions.

The results clearly indicated (1) that setting the initial weights of the individual components of the objective function was an important factor in the success of calibration and (2) that the likelihood of reaching an optimum solution was increased by selecting the initial weights using knowledge of the flow proportions. To select initial weights for the components of the objective function a separation of the total flow into baseflow and quickflow hydrographs was performed. The corresponding flow proportions were then used to set the objective function weights, using equations provided herein. The analyses provided information on the effect and importance of incorporating hydrologic information into the calibration process. The objective function components must reflect the hydrologic components inherent to the discharge record.

12.5 EFFECT OF THE INITIAL SOIL STORAGE ESTIMATES ON PREDICTION ACCURACY

Previous analyses of discharge sensitivity indicated that the sensitivity of the predicted discharges was a function of the flow proportions in a watershed. Therefore, it was of interest to determine if the accuracy of the predicted discharges was also a function of the flow proportions when the estimates of the initial storages were in error. Although the true initial storage conditions are always unknown, the use of hypothetical data provided information on the model response when poor estimates of the initial storages deviated from their true values. Three hypothetical watersheds with different baseflow and quickflow proportions were used for the analyses.

When the initial soil storages were deviated from their true value, the predicted nominal storages experienced large errors at the start of the period of record and as the deviation in the initial storage parameters increased, the initial error in the predicted storages also increased; however, the errors in the predictions decreased as the time passed. These errors were transferred to the predicted discharges. The analysis also indicated that for predictions during the first year of the period of record the flow proportions influenced the accuracy of the predicted nominal storages and discharges. However, that after the first year, the proportions of flow were not an important factor in prediction accuracy.

The effect of LZS erroneous estimates was relatively larger for watersheds with predominant baseflow when compared to the effect in predominant quickflow watersheds; this effect was reflected in the magnitude and the slope of the relative bias evaluated during the first year of the period of record. The effect of error in the estimates of the upper zone storage (UZS) was observed in inaccurate predicted nominal storages and discharges; yet, the magnitude of the error was lower than the magnitude of the error caused by erroneous estimates of LZS. Overprediction was always observed for overestimations of UZS and underprediction was always observed for underestimations of UZS.

The need to have reasonable estimates of the initial soil storages is unquestionable; however, this information is not available for real watersheds. The need to an alternative approach to overcome the problems caused by erroneous estimates of the initial soil storages in the prediction accuracy was therefore needed. It was of interest to determine the time to which the predicted discharges became insensitive to the estimates of the initial soil storages. The results are presented in the following section.

12.6 EFFECT OF THE START-UP PERIOD ON PREDICTION ACCURACY

Accurate estimates of the initial soil storages are important because errors in the estimates propagate to the predicted discharges; however, at some point in time the computed flows will be independent of the initial storage estimates. This period was referred to as the “start-up period”. The length of time to the point of insensitivity was of interest because in many cases this period can be a large part of a short record length used for calibration. The results indicated that the time for the predictions to converge was larger for erroneous estimates of LZS than for erroneous estimates of UZS. In addition, the time for convergence was slightly larger when LZS was overestimated than when LZS was underestimated. However, the start-up period was similar for the three hypothetical watersheds, suggesting that the flow proportion was not an important factor for the start-up period. The predicted discharges experienced the effect of error in the initial storages for about a year from the start of the period of record.

The presence of outliers was an additional factor that influenced prediction accuracy. The most severe effect of outliers on prediction accuracy was observed when their occurrence was during the initial part of the record because of the combined effect with erroneous estimates of the initial storages. The implication of outliers when using parameter estimators is the difficulty to calibrate because of the incompatibility between the method of calibration (least squares) and the variation of data that includes outliers. The outliers were addressed through a with- vs.-without sensitivity analysis in which the extreme values of daily precipitation and daily discharges were reduced to magnitudes more like the values observed in the overall data. Once the extreme discharges were reduced, the mean daily discharge was reduced and better calibrations were achieved. This is rational as the mean value of the predicted discharges is increased by the presence of very high precipitation events. The results of the analysis indicated that the alternative to overcome problems related to erroneous estimates of state variables representing initial conditions on prediction accuracy is to provide a start-up period of about a year. The predictions during the start-up period should not be included in the calibration process or when evaluating the accuracy of the model.

12.7 ON THE EFFECT OF LAND USE NONSTATIONARITY ON PREDICTION ACCURACY

The accuracy of modeled daily discharges using the HSPF model is affected by the nontemporal variation of the land-use data. These errors are caused by the assumption that the land use in the watershed remains constant for the duration of the period of record. Hydrologic studies (Dunne and Leopold, 1978) have demonstrated that, extensive changes in the hydrologic regime and the channel morphology takes place as the watershed undergoes urbanization. The duration of quickflow decreases while their magnitude increases. Consequently, baseflows are likely to decrease. Such changes in land use patterns reduce infiltration which leads to the reduction of baseflow and groundwater recharge.

If a watershed undergoes land use change during the period of record, then this change will be reflected in the measured discharge record. However, if a constant land use is assumed when calibrating HSPF, biased predictions are likely to occur as the analysis does not reflect the changes of the hydrologic regime. The problem is compounded because the HSPF parameters are held constant over the record length. If the assumption is that the parameters represent the lumped physical processes being modeled, then as the land use changes, the parameters should change to reflect the changes in the characteristics of the modeled land uses. The physical processes are changing in a temporally varying watershed, but the model does not allow the parameters to vary with the processes.

The results of this analysis indicated that for a watershed undergoing urbanization during a 14-yr period and when using land use data from years at the beginning or end of the period, the magnitude of the relative bias in the baseflow component was of up to 40%. The maximum error though, was reduced to 20% when using land use data from the middle of the calibration period. The magnitude of the relative bias for the quickflow during the 14-yr period was of about 5% when using land use data from the middle of the period. These results will vary for other HSPF applications as a function of the rate and density of urbanization. Controlling the inaccuracy in every source of data error will contribute to a reduction in the uncertainty of the final parameter values and, thus, a reduction in the uncertainty of the predictions.

12.8 ON THE EFFECT OF LAND USE NONSPATIAL DISTRIBUTION ON PREDICTION ACCURACY

The accuracy of the predicted discharges and the rationality of the calibrated parameters were affected by the omission of the channel transport processes and by the land use nonspatial distribution. The omission of the channel routing yielded significant underprediction of the predicted discharge because the time required for water to flow from the uppermost part of the watershed to the outlet is greater than for a watershed with channel processes modeled. The slower velocities allow for greater predicted evapotranspiration and thus, greater underprediction mainly of the baseflow component.

Underprediction was also observed in the predicted peak flows, however the underprediction of stormflow when using PEST may be due to the nature of the least squares method, which tends to calibrate towards the mean values of the total runoff.

The calibrated parameters were also affected by the omission of channel routing. For example, values of the nominal-storage parameters were lower for the distributed model than for the lumped model. When the channel processes are ignored, the time of concentration increases, and the potential for evapotranspiration also increases because the water stays in the watersheds for a longer period. These lower values are rational because the model attempts to reduce the potential of evapotranspiration in order to accurately predict the discharge.

The nonspatial land use distribution was determined to be an important criterion for deciding whether or not to transfer parameters. The calculated biases of the peak discharges were much smaller for UUFF than for FFUU. When distributed data from a watershed with the urban land use near the outlet is modeled with a lumped model, then biases in computed peaks will be much less than for the case where the urban land is in the upper reaches of the watershed. The standard deviations of the predicted discharge were for all cases, higher than the standard deviation of the measured discharge (Table 10-4). The lumped model increases the variation in the runoff because discharge from the lower portion of the watershed is not released from the watershed earlier than that from the upper portion, as it is in the distributed model.

CHAPTER 13

RECOMMENDATIONS FOR FUTURE RESEARCH

13.1 INTRODUCTION

The recommendations suggested in this chapter are extensions of the topics addressed in the dissertation. The analysis and recommendations provided in this document are intended to be the beginning of additional research to further understand the calibration process of the HSPF model. The use of the parameter estimator (PEST) is recommended, but the user should take the time to understand the interactions between HSPF and PEST and the implications, and constraints of using automatic calibration methods in order to achieve useful results. The use of PEST does not guarantee calibrated or even rational models, only expand the possibility to evaluate the effects of subjective decisions during the calibration process and facilitate the replication of model results. The capability of results replication is perhaps one of the most important benefits of the application of parameter estimators.

HSPF is frequently used to estimate pollution concentrations, not just flow rates and volumes. The analyses made as part of this research concentrated on water quantity issues, not the water quality issues. This was done because accurate estimates of pollutant concentrations depend on having accurate water quantity estimates. Having accurate estimates of water quantity is a prerequisite to accurate water quality estimates. The suggested analyses for future research are expected to increase the understanding of the HSPF model capabilities as well as the model limitations, so that regulations based on more objective models can be established in the future.

13.2 RAINFALL DISAGGREGATION METHODS

Accurate estimates of rainfall are important because rainfall is the driving force of the HSPF runoff predictions. Problems of data availability are not only related to the few number of measuring gages, but to the temporal scale at which rainfall is collected. Hourly rainfall is the most common temporal scale used in calibrations with the HSPF model, but daily data is the most common scale of measured rainfall. Thus, the temporal rainfall disaggregation from daily to hourly amounts needed to be investigated.

The analyses on the effect of rainfall disaggregation on the accuracy of the predicted discharges indicated that methods of disaggregation based on analyses of storm frequency provided the best results. However, the accuracy varied between poor for hourly predictions, and moderate for daily predictions. To some extent, these results were explained by the data used in the development of the storm-frequency fractions for the depth-duration-dependent method. Given that the monthly rainfall data in the analyzed region were somewhat uniform, seasonal variation was not considered in the development of the methodology (Kreeb, 2003).

The suggested study would establish if by determining the monthly and seasonally storm-frequency fractions, rather than annually, the prediction accuracy could increase. The duration and classification according to the fraction of rainfall could also be modified. The number of hours varied with the total storm depth method of rainfall disaggregation may benefit the accuracy of the predicted discharge. This possibility merits further investigation. Once the depth-duration-dependent method is applied to disaggregate the daily rainfall data and the calibration is performed, the accuracy of the HSPF predictions can be evaluated through the rationality of the calibrated parameters and using the goodness-of-fit statistics for all flows, baseflow, and peak flows, annually or seasonally calculated.

13.3 DEVELOPMENT OF AN OBJECTIVE FUNCTION

Research on the application of parameter estimators to hydrologic models has been focused on the improvement of the gradient-based approaches and on the inclusion of more complex algorithms to speed the time in which the minimum value of the objective function is reached. However, little investigation has been done in regard to the components of the objective function.

The use of a parameter estimator only provides the means of amounts of information for analysis that otherwise would be impossible to obtain and the warranty that the calibration results are replicable at any point. However, when using a parameter estimator the quality and rationality of the calibrations are a function of the amount and detail of the hydrologic information provided in the objective function. The current objective function can be enhanced through the addition of component/s related to the calibration of the storm runoff and through the incorporation of better formulations for the calibration of parameters varying monthly. Moderate to accurate predictions of baseflow are attained with the hydrograph separation and the autoregression components of the current objective function. However, the accuracy of the predicted quickflow using the Quickflow filter is poor. The evaluation of the Quickflow filter formulation or the addition of a component that provides information such as storm volumes and their effect on the calibration should be further investigated.

13.4 GROUNDWATER RESIDENCE TIME

Although this issue was not addressed in the analyses of the dissertation, it is important to mention it and to suggest further investigation to determine the effect of the current model formulation on the accuracy of the hydrologic model predictions. The residence time may be an important process in calibrating the water budget and water quality. For example 10-20 years is the estimated residence time of water to work through soils and aquifers, and into the waterways in the Chesapeake Bay watershed (Bachman et al, 1998; Focazio et al, 1998). A parameter that represents this delay does not exist in the current model formulation of the HSPF as it assumes that the release of groundwater through the baseflow is almost immediately after the rainfall occurs. Disregarding the residence time of groundwater in the current model formulation may have unknown effects related to the rationality of the calibrated parameters or even worse, to erroneous predictions of pollutant concentrations.

13.5 VALIDATION METHODS FOR MODEL PREDICTIONS

The validation of the HSPF model is not a common practice in the modeling community. The lack of model validation may be due to the amount and complexity of the data needed for the model, so that the available data is used for the calibration process. The implication of this practice is that the forecasting prediction accuracy is not really known. The forecasting prediction accuracy of pollutant load reductions in many government programs are justified and supported by the goodness-of-fit statistics obtained during the model calibration. From the theory that underlies confidence intervals, it is known that the prediction accuracy is always poorer than the calibration accuracy. (McCuen, 2005) indicates that goodness-of-fit statistics that accompany model calibration may not be good indicators of prediction accuracy. Methods of model validation such as the split-sample testing and jackknifing methods should be investigated to evaluate the forecasting prediction accuracy of the HSPF.

REFERENCES

Acreman, M.C. (1990). A Simple Stochastic Model of Hourly Rainfall for Farnborough, England. Hydrological Sciences Journal. 35(2): 119 - 148

Arnold, C.L., Boison, P.J., and Patton, P.C. (1982). Sawmill Brook: An example of Rapid Geomorphic Change Related to Urbanization. J. of Geology 90: 155 - 166

Bachman, L.J., Lindsey, B.D., Brakebill, J.B., and Powars, D.S. (1998). Ground-water discharge and base-flow nitrate loads of nontidal streams, and their relation to a hydrogeomorphic classification of the Chesapeake Bay watershed, Middle Atlantic Coast: U.S. Geological Survey Water-Resources Investigations Report 98-4059, 71 p.

Bergman M.J., Green W., and Donnangelo L.J. (2002). Calibration of Storm Loads in the South Prong Watershed, Florida, Using WATERSHEDS/HSPF, J. of the American Water Resources Association. 38(5): 1423-1436.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L. Jr., Donigian, A.S., and Johanson, R.C. (1997). *Hydrologic Simulation Program – FORTRAN. User's Manual for Release 11*: U.S. Environmental Protection Agency, National Research Laboratory, Athens, GA., EPA/600/R-97/080, 755 p.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Jobes, T.H., and Donigian, A.S. (2001). *HSPF User's Manual*. Aqua Terra Consultants, Mountain View, CA.

Bradley, A.A., Peters-Lidard, C., Nelson, B.R., Smith, J.A., and Young, C.B. (2002). Raingage Network Design Using NEXRAD Precipitation Estimates, J. Amer. Water Resources Assoc**.** 38(5): 1393 – 1408.

Butterworth, S. (1930). "On the Theory of Filter Amplifiers." Wireless Engineer, vol. 7, 536 – 541.

Carroll, R.J. and Ruppert D. (1988). "*Transformation and weighting in Regression",* Chapman and Hall, New York.

Dietzel, C. (2003). *Spatial Differences in Multi-resolution Urban Modeling.* Department of Geography, University of California at Santa Barbara, Santa Barbara, CA.

Doherty J. (2001). *PEST-ASP User's Manual*, Watermark Numerical Computing, Australia.

Dunne, T. and L.B. Leopold, (1978). *Water in Environmental Planning*, W. H. Freeman and Co., San Francisco.

Durrans, S.R., Burian, S.J., Nix, S.J., Hajji, A., Pitt, R.E., Fan, C., and Field, R. (1999). Polynomial-based Disaggregation of Hourly Rainfall for Continuous Hydrologic Simulation, J. of the American Water Resources Association, 35(5): 1213-1221.

Focazio, M.J., Plummer, L.N., Bohlke, J.K., Busenberg, Eurybiades, Bachman, L.J., and Powars, D.S. (1998). Preliminary estimates of residence times and apparent ages of ground water in the Chesapeake Bay watershed, and water-quality data from a survey of springs: U.S. Geological Survey Water-Resources Investigations Report 97-4225, 75 p.

Hammer, T.R. (1972). Stream Channel Enlargement Due to Urbanization, Water Resources Research. 8(6): 1530 – 1540.

Hay L.E., Clark M.P., Wilby R.L., Gutowski W.J., Leavesley G.H., Pan Z., Arritt R.W., and Takle E.S. (2002). Use of Regional Climate Model Output for Hydrologic Simulations. Journal of Hydrometeorology**.** 3: 571-590.

Herold, M., Scepan, J., Müller, A., and Günther, S. (2002). Object-Oriented mapping and analysis of urban land use/cover using IKONOS data. Proceedings of 22nd EARSEL Symposium “Geoinformation for European-wide integration, Prague, Czech Republick.

Holman-Dodds, J.K., Bradley, A.A., and Sturdevant-Rees P.L. (1999). Effect of Temporal Sampling of Precipitation on Hydrologic Model Calibration, Journal of Geophysical Research-Atmospheres, 104 (D16), 19645-19654.

Hollis, G.E. (1975). The effect of Urbanization on Floods of Different Recurrence Interval, Water Resources Research. 66: 84-88

Huff, F.A. and Neill, J.C. (1957). Areal Representativeness of Point Rainfall, Trans. AGU**.** 38(3): 341-345.

Kreeb, L.B. and McCuen, R.H. (2003). *Hydrologic Efficiency and Design Sensitivity of Bioretention Facilities*. Report, Department of Civil Engineering, University of Maryland, College Park, MD.

Levenberg, K. (1944). *A Method for the Solution of Certain Problems in Least Squares*. Quart. Appl. Math. 2, 164-168.

Leopold, L.B. (1968). *The Hydrologic Effects of Urban Land Use: Hydrology for Urban Land Planning – A Guidebook of the Hydrologic Effects of Urban Land Use*. USGS Circular 554.

Levy, B. and McCuen, R.H. (1999). Assessment of Storm Duration for Hydrologic Design. J. Hydrologic Engineering, ASCE, 4(3): 209-213.

Lumb, A.M. and Kittle, J.L., Jr., (1995). *Users manual for METCMP, a computer program for the interactive computation of meteorologic time series*, unpublished U.S. Geological Survey Report, 88 p.

Lumb, A.M., McCammon, R.B., and Kittle, J.L., Jr. (1994). *Users Manual for an expert system (HSPEXP) for Calibration of the Hydrologic Simulation Program-fortran*. U.S. Geological Survey Water-Resources Investigations Report 94-4168, 102 p.

Marquardt, D. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM J. Appl. Math. 11, 431-441.

McCuen, R.H. (1985). Statistical Methods for Engineers. Prentice-Hall Inc, Englewood Cliffs, NJ.

McCuen, R.H. and Snyder, W.M. (1986). Hydrologic Modeling: Statistical Methods and Applications. Prentice Hall, Inc Englewood Cliffs, NJ.

McCuen, R.H. (1993). Microcomputer Application in Statistical Hydrology**.** Prentice-Hall, Inc., Upper Saddle River, NJ.

McCuen, R.H., (2003). *Modeling Hydrologic Change,* CRC Press, Boca Raton, FL.

McCuen, R.H., (2005). Accuracy Assessment of Peak Discharge Models. Journal of Hydrologic Engineering, Vol. 10., No. 1. pp 16 – 22.

McKay, G.A. (1970). “Precipitation”, Chap. 2 of Handbook on the Principles of Hydrology (D.M. Gray, ed.), Water Information Center, Huntington, NY.

Nathan, R.J. and T.A. McMahon, (1990). Evaluation of Automated Base Flow and Recession Analyses. Water Resources Research, Vol. 26., No. 7, pp 1465-1473.

National Weather Service. (1961). "Rainfall Atlas of the United States, 30-minute to 24-hour Durations, 1 to 100-year Return Periods," Technical Publication 40, Department of Commerce, National Weather Service, Washington, DC.

Nash, J.E. and Sutcliffe, J.V. (1970). River flow forecasting through the conceptual models, 1: a discussion of principles. Journal of Hydrology 10 (3): 282-290

Service Conservation Service, (1986). Urban Hydrology for Small Watersheds, *Technical Release TR-55*, Soil Conservation Service, Hydrology Unit.

Shumway, R.H. (1988). *Applied Statistical Time Series Analysis*, Prentice-Hall, Upper Saddle River, N.J.

Sloto, R.A. and Crouse, M.Y. (1996). *HYSEP: A computer program for streamflow hydrograph separation and analysis*: U.S. Geological Survey Water-Resources Investigations Report 96-4040, 46 p.

Smakhtin, V.Y. and Watkins, D.A. (1997). "Low Flow Estimation in South Africa". *Report* No. 494/1/97, Water Research Commission: Pretoria, South Africa.

Socolofsky, S., Adams, E., and Entekhabi D. (2001). Disaggregation of Daily Rainfall for Continuous Watershed Modeling., J. Hydrologic Engineering, ASCE. 6(4): 300-309.

Wilde, D.J. and Beightler, C.S. (1967). *Foundations of Optimization*. Prentice-Hall, Upper Saddle River, NJ.

Yapo, P.O. (1996). *A multiobjective global optimization algorithm with application to calibration of hydrologic models***.** Ph.D. Dissertation, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ.