Estimating Uncertainty in HSPF based Water Quality Model: Application of Monte-Carlo Based Techniques

Anurag Mishra

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biological Systems Engineering

Brian L. Benham
Dan L. Gallagher
Kenneth H. Reckhow
Eric P. Smith
Mary Leigh Wolfe

July 28, 2011
Blacksburg, Virginia

Keywords: *water quality modeling, TMDL, fecal coliform, HSPF, uncertainty analysis, Monte-Carlo, Bayesian techniques, GLUE, MCMC, two-phase Monte Carlo analysis*

Estimating Uncertainty in HSPF based Water Quality Model: Application of Monte-Carlo Based Techniques

Anurag Mishra

Abstract:

To propose a methodology for the uncertainty estimation in water quality modeling as related to TMDL development, four Monte Carlo (MC) based techniques—single-phase MC, two-phase MC, Generalized Likelihood Uncertainty Estimation (GLUE), and Markov Chain Monte Carlo (MCMC) —were applied to a Hydrological Simulation Program–FORTRAN (HSPF) model developed for the Mossy Creek bacterial TMDL in Virginia. Predictive uncertainty in percent violations of instantaneous fecal coliform concentration criteria for the prediction period under two TMDL pollutant allocation scenarios was estimated. The average percent violations of the applicable water quality criteria were less than 2% for all the evaluated techniques. Single-phase MC reported greater uncertainty in percent violations than the two-phase MC for one of the allocation scenarios. With the two-phase MC, it is computationally expensive to sample the complete parameter space, and with increased simulations, the estimates of single and two-phase MC may be similar. Two-phase MC reported significantly greater effect of knowledge uncertainty than stochastic variability on uncertainty estimates. Single and two-phase MC require manual model calibration as opposed to GLUE and MCMC that provide a framework to obtain posterior or calibrated parameter distributions based on a comparison between observed and simulated data and prior parameter distributions. Uncertainty estimates using GLUE and MCMC were similar when GLUE was applied following the log-transformation of observed and simulated FC concentrations. GLUE provides flexibility in selecting any model goodness of fit criteria for calculating the likelihood function and does not make any assumption about the distribution of residuals, but this flexibility is also a controversial aspect of GLUE. MCMC has a robust formulation that utilizes a statistical likelihood function, and requires normal distribution of model errors. However, MCMC is computationally expensive to apply in a watershed modeling application compared to GLUE. Overall, GLUE is the preferred approach among all the evaluated uncertainty estimation techniques, for the application of watershed modeling as related to bacterial TMDL development. However, the application of GLUE in watershed-scale water quality modeling requires further research to evaluate the effect of different likelihood functions, and different parameter set acceptance/rejection criteria.

Acknowledgments

I would like to thank my major advisor Dr. Brian L. Benham for his continued support and guidance throughout my PhD program. Dr. Benham gave me great latitude in pursuing the research area of my interest. He had been extremely patient during my learning phase, my rough times, and has been very instrumental in shaping my career.

I would also like to thank my committee members Dr. Dan Gallagher, Dr. Kenneth Reckhow, Dr. Eric Smith and Dr. Mary Leigh Wolfe for accepting to serve on my committee and providing valued input and feedback. I am glad to have gotten in touch with Dr. Scotland Leman during my research. Dr. Leman taught me several concepts related to statistical analysis and Markov Chain Monte Carlo. He helped me in learning MATLAB programming as well.

I am extremely grateful to the staff at Center for Watershed studies for helping me during my research and Graduate Research Assistantship. I am especially indebted to Dr. Rebecca Zeckoski for teaching me several important details of watershed modeling, and programming in Visual Basic. Kevin Brannan and Gene Yagow were always there to answer my questions, no matter how busy they were, no matter how simple were the questions. I want to thank Denton Yoder for helping me in understanding basic concepts related to programming and database management. I want to thank my employer AQUA TERRA Consultants in their continued encouragement, and support.

Finally, I want to thank my wife Vineeta for her love, and constant support during my Graduate studies, and my parents for their patience, support and well wishes during my studies.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Under section 303(d) of the 1972 Clean Water Act, states, territories, and authorized tribes are required to develop a list of “impaired” waters. According to the U.S. Environmental Protection Agency (USEPA), over 40% of the assessed waters in the United States (some 60,000 individual river or stream segments, lakes, and estuaries) are impaired, primarily because of nonpoint source pollution (USEPA, 2009). The states, territories and authorized tribes are required to develop a Total Maximum Daily Load (TMDL) for these impaired waters. A TMDL specifies the reductions in the pollutant sources that will bring the impaired waters into compliance with the water quality standards. Mathematically, a TMDL is written as

$$TMDL = \sum WLA + \sum LA + MOS \qquad 1.1$$

Where, $\sum WLA$ = waste load allocation (point sources)

$\sum LA$ = load allocation (non-point sources)

MOS = margin of safety

Developing a TMDL often includes modeling the processes that contribute to the impairment with the application of a water quality simulation modeling software. The margin of safety (MOS) is often included in the TMDL calculation to account for the inherent uncertainty present in a natural system. Uncertainty is always present when modeling a natural system (Morgan and Henrion, 1990). However, typically no formal calculation is performed to quantify this uncertainty, as there is limited science-based guidance available to estimate the amount of uncertainty that is associated with modeling the processes dealing with water quality.

In 2001, the USEPA (2001) estimated the annual average cost for TMDL development to be $63-69 million per year for the next 15 years. The report also estimated that the cost of TMDL development and implementation could exceed $1 billion per year. High concentrations of pathogen indicator organisms (e.g. fecal coliforms, *E. coli*) are currently the leading cause of impairments, and responsible for 14% of the identified impairments nationwide (USEPA, 2009). In spite of the significant costs associated with developing bacterial TMDLs, there has been little attempt to quantify the uncertainty associated with modeling that is often conducted when developing this type of TMDL.

Without a measure of uncertainty in model predictions, one cannot assess the probability of achieving applicable water quality criteria, nor assess the risk associated with not achieving those criteria. Currently, when developing a TMDL, the MOS (when explicitly defined) is often an arbitrarily set percentage of the TMDL. The additional information about modeling uncertainty would provide decision makers and stakeholders with additional knowledge allowing them to make a more informed judgment when choosing a MOS and when comparing pollutant load allocation scenarios. Beven (1993) described inclusion of uncertainty analysis in the modeling process as “intellectual honesty”, which becomes imperative if significant public resources are at stake in the process.

Most water quality simulation modeling software used for TMDL development are a combination of process-based and empirical models that do not include detailed uncertainty analysis capabilities. The ‘Hydrological Simulation Program–FORTRAN (HSPF)’ model (Bicknell et al., 2005), which is supported by USEPA as part of a larger modeling package ‘Better Assessment Science Integrating Point and Nonpoint Sources (BASINS)’- is frequently used for developing bacterial impairment TMDLs. HSPF can simulate hydrology and various water quality constituents like sediment, indicator bacteria (IB), nitrates, phosphorus etc. in watersheds of varying size (Bicknell et al., 2005). HSPF outputs a deterministic time series of hydrology and water quality constituents without quantifying uncertainty.

Uncertainty in model predictions can be estimated using two categories of methods, namely Monte Carlo methods and first-order variance propagation (Beck, 1987; Summers et al., 1993). First-order methods assume linear models that limit their usability with respect to complex modeling software like HSPF (Summers et al., 1993). Paul et al. (2004) conducted a first-order analysis to estimate the contribution of sensitive parameters to the fraction of variance of simulated peak in-stream fecal coliform (FC) concentration in a watershed modeled with HSPF. They inferred that small uncertainties in selected water quality parameters could result in large uncertainties in the prediction of in-stream FC concentration. Paul et al. (2004) recommended the use of Monte Carlo based analysis to evaluate uncertainty in bacteria modeling using HSPF.

Monte Carlo simulation is a method that involves performing repeated simulations of the model in question using randomly selected parameter values from predetermined input parameter probability distributions. The process is repeated for a number of iterations sufficient to converge on an estimate of the probability distribution of output variables (Gardner and O’Neill, 1983). Monte Carlo simulations can be used to estimate uncertainty in water quality simulation modeling along with other Monte Carlo based methods that include,-two-phase Monte Carlo simulation

(Helton, 1994; McIntosh et al., 1994), Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992), Bayesian Monte Carlo (BMC) (Dilks et al., 1992), and Markov Chain Monte Carlo (MCMC) (Kuczera and Parent, 1998).

There have been some applications of Monte Carlo based techniques to estimate uncertainty in hydrologic modeling (e.g., Balin, 2004; Beven and Binley, 1992; Benaman and Shoemaker, 2002; Donigian et al., 2007; Hession et. al, 1996; Makowski et al., 2002; Stow et al., 2007), however, there are very few applications of these techniques to estimate uncertainty in water quality simulation modeling and TMDL development. It has been argued that the presence of too many competing methods for assessing uncertainty makes it difficult for modelers to select a method and interpret the results (Pappenberger and Beven, 1996). The research reported herein attempts to compare different Monte Carlo based techniques that can be used to estimate uncertainty in water quality modeling related to TMDL development.

1.1 Goals and Objectives:

The goal of this research was to evaluate selected uncertainty estimation techniques when applied to a HSPF model used to develop a bacterial impairment TMDL. To accomplish this goal, specific objectives of this research were to:

1. compare the applicability of single- and two-phase Monte Carlo in estimating uncertainty in HSPF-based water quality simulation modeling for TMDL development. (Chapter 3)
2. assess the impact of using log-transformed in-stream fecal coliform concentrations on predictive uncertainty when using the Generalized Likelihood Uncertainty Estimation (GLUE) technique with HSPF-based water quality simulation model for TMDL development. (Chapter 4)
3. evaluate the applicability of GLUE and Markov Chain Monte Carlo (MCMC) in estimating uncertainty in the water quality simulation modeling when using HSPF for bacterial TMDL development. (Chapter 5)

1.2 Dissertation Organization

Chapter one of this dissertation introduces the general concept of uncertainty analysis in water quality modeling. The second chapter provides a detailed literature review of water quality modeling and uncertainty analysis techniques. Chapters three, four, and five are specific to each of the study's three objectives. These chapters were developed as papers in the format accepted for the *Transactions of the American Society of Agricultural and Biological Engineers* and are intended to stand alone. There is some limited repetition in these chapters. The sixth chapter is an overall conclusion chapter.

References:

Balin, D. 2004. Hydrological Behaviour through Experimental and Modelling Approaches. Application to the Haute-Mentue Catchment. School of Acrhitecture, Civil and Environmental Engineering. Lausanne, Switzerland. Ecole Polytechnique Fédérale de Lausanne.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Benaman, J., and C.A. Shoemaker. 2002. Sensitivity and uncertainty analysis of a distributed watershed model for the TMDL process. National TMDL Science and Policy 2002 Speciality Conference, Water Environment Federation.

Beven, K. 1993. Prophecy, Reality and uncertainty in distributed hydrological modeling. *Adv. Water Resources.* 16(1): 41-51.

Beven, K.. and A. Binley (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr. T.H. Jobes, and A.S. Donigian, Jr. 2005. *HSPF Version 12.2 User's Manual.* AQUA TERRA Consultants. Mountain View, CA.

Dilks, D.W., R.P. Canale, and P.G. Meier. 1992. Development of Bayesian Monte Carlo Techniques for Water Quality Model Uncertainty. *Ecological Modeling.* 62(1-3): 149-162.

Donigian, A.S., and J.T. Love. 2007. The Housatonic River Watershed Model: Model Application and Uncertainty Analyses. 7th International IWA Symposium on System Analysis and Integrated Assessment in Water Management, May 7-9, 2007. Washington, DC. WATERMATEX Proceedings on CD-ROM.

Gardner, R.H., and R.V. O'Neil. 1983. Parameter Uncertainty and Model Predictions: A Review of Monte Carlo Results. In *Uncertainty and Forecasting of Water Quality, eds.* M.B. Beck, and G. Van Straten, 345 -257. Berlin, Germany: Springer-Verlag.

Hession, W.C., D.E. Storm, and C.T. Haan. 1996. Two-phase uncertainty analysis: an example using universal soil loss equation. *Trans. ASAE.* 39(4): 1309-1319.

Helton, J.C. 1994. Treatment of uncertainty in performance assessment for complex systems. *Risk Analysis*. 14(4): 483-511.

Kuczera, G., and E. Parent. 1998. Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: the Metropolis Algorithm. *Journal of Hydrology. 211: 69-85.*

Makowski, D., D. Wallach, and M. Tremblay. 2002. Using a Bayesian approach to parameter estimation; Comparison of GLUE and MCMC methods. *Agronomie*. 22: 191-203.

MacIntosh, D.L., G.W. Suter., and F.O. Hoffman. 1994. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. *Risk Analysis*. 14(4): 405-419.

Morgan, M.G., and N. Henrion. 1990. *Uncertainty*. Cambridge University Press, New York City, NY.

Pappenberger, F. and K.J. Beven. 2006. Ignorance is Bliss: Or seven reasons Not to Use Uncertainty Analysis. *Water Resources. Research.* 42(5). W05302, doi:10.1029/2005WR004820.

Paul, S., P.K. Haan, M.D. Matlock, S. Mukhtar, and S.D. Pillai. 2004. Analysis of the HSPF water quality parameter uncertainty in predicting peak in-stream fecal coliform concentrations. Trans. ASAE. 47(1): 69-78.

Stow, C.A., K.H. Reckhow, S.S. Qian, E.C. Lamon, G.B. Arhonditsis, M.E. Borsuk, and D. Seo. 2007. Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. *J. Amer. Water Resources Ass.* 43(6): 1499-1507.

Summers, J.K., H.T. Wilson, and J. Kou. 1993. A method for quantifying the prediction uncertainties associated with water quality models. *Ecological Modeling.* 65: 161-176

USEPA. 2001. The National Cost of the Total Maximum Daily Load Program (Draft Report). Office of Water, United States Environmental Protection Agency. Washington D.C.

USEPA. 2009. National Section 303(d) list fact sheet. United States Environmental Protection Agency. Available from: URL: http://oaspub.epa.gov/waters/national_rept.control.

Chapter 2. Literature Review

The following sections review the literature pertinent to TMDLs, water quality simulation modeling, and uncertainty analysis.

2.1 Water Quality and Total Maximum Daily Load (TMDL)

According to the U.S. Environmental Protection Agency (USEPA), over 40% of the assessed waters in the United States (some 60,000 individual river or stream segments, lakes, and estuaries) are impaired, primarily because of nonpoint source pollution (USEPA, 2009). Pathogens, typically represented by a fecal indicator bacteria (IB), are one of the leading causes of water quality impairments in the US and about 14% of the assessed river length (~150,362 km), and 3.3% of the assessed lakes, reservoir and ponds (~2347 square km) in the US are impaired due to excessive IB (USEPA, 2002). Elevated concentrations of IB are responsible for over 30% of the identified impairments in the state of Virginia (USEPA, 2011). Fecal coliform (FC) and enterococci (primarily for marine waters) are common IB, although *E. Coli* (EC) are used more frequently as an IB when assessing flowing, fresh water.

The 1972 Clean Water Act requires total maximum daily loads (TMDLs) be developed for impaired water bodies. A TMDL is the maximum amount of pollutant a waterbody can receive and still meet its intended use (Benham, 2002). Mathematically, a TMDL is represented as

$$TMDL = \sum WLA + \sum LA + MOS \tag{2.1}$$

where,

ΣWLA = waste load allocations (point sources), and

ΣLA = load allocations (non-point sources),

MOS = margin of safety

2.2 Water Quality Modeling and TMDL

Developing a TMDL often involves a process where the contributions of different pollution sources are quantified and linked to the water quality of the impaired waterbody. In this process, the allowable pollutant load is partitioned among the considered sources. Water quality simulation models are often used to link pollutant sources to water quality. These models use mathematical relationships to represent the fate and transport of pollutants from the source to the waterbody. Once developed, calibrated and validated, a water quality simulation model can be used to determine the needed pollutant reductions to achieve the TMDL, and assess the impact of various pollutant control management strategies.

Water quality simulation modeling software can be empirical or a combination of empirical and process-based. Watershed-scale hydrology and water quality simulation modeling software that have been used for TMDL development include Hydrological Simulation Program–FORTRAN (HSPF) (Bicknell et al., 2005), Soil and Water Assessment Tool (SWAT) (Neitsch et al., 2005), Agricultural Non-Point Source Model (AGNPS) (Young et al., 1987), and the Annualized AGNPS (AnnAGNPS) (Bingner and Theurer, 2001). HSPF has been widely used for developing IB impairment TMDLs (eg. Benham et al., 2005; VADCR, 2003; Yagow, 2001). SWAT has also been used in developing IB impairment TMDLs, but fewer TMDLs have been developed using SWAT compared to HSPF. Further, the microbial component of SWAT has not yet been validated against the measured data at watershed scale (Im et al., 2004).

2.2.1 Modeling Bacteria as a Water Quality Constituent

As mentioned earlier, pathogens are the second most widespread cause of water quality impairment in the United States. Pathogens however are difficult to identify and count, and are typically quantified in terms of a fecal IB (Rosen, 2000). Presence of IB means that the pathogenic organism may be present. Water quality criteria, therefore, typically specify concentration of a specific IB species.

Indicator bacteria fate and transport involves several processes that include, but are not limited to, manure production by animals, transport within the water column and groundwater, die-off, and regrowth. The type of animal and diet affects the production of IB. Manure management practices, hydrology, and sediment transport all affect IB transport. Soil moisture, temperature, pH, solar radiation, and time affects IB regrowth and die-off in and on the soil, and in the water.

Water quality modeling software that have been used to model IB fate and transport include the Agriculture Runoff Management II: Animal Waste Version model (ARM II) (Overcash et al., 1983), the Utah State Model (UTAH) (Springer et al., 1983), the MWASTE model (Moore et al., 1989), the COLI model (Walker et al., 1990), SWAT (Neitsch et al., 2005), and HSPF (Bicknell et al., 2005). These models utilize different levels of complexity when simulating IB fate and transport processes. HSPF has been widely applied to develop IB impairment TMDLs, (Benham et al., 2004; Yagow et al., 2001), and was used for the research reported here.

2.2.2 Hydrological Simulation Program-FORTRAN (HSPF)

HSPF is a continuous, watershed-scale modeling software that simulates hydrology and water quality processes (Bicknell et al., 2005). HSPF works on a mass balance approach, where water and water quality constituents are routed through appropriate pathways. HSPF uses several modules and sub-modules to simulate various processes.

In HSPF, the watershed is divided into subwatersheds that are modeled using modules that simulate processes in, or on pervious land areas (PERLND), impervious land areas (IMPLND), and reaches and reservoirs (RCHRES). The sub-modules PWATER and IWATER simulate runoff from PERLND and IMPLND, respectively; and the sub-module HYDR routes water through the RCHRES. Soluble water quality constituents are simulated using the PQUAL sub-module on PERLND, IQUAL on IMPLND, and GQUAL in RCHRES. When using HSPF to develop TMDLs for IB impairment, FC is typically simulated as a planktonic (dissolved) constituent. Although FC can move with the water like a dissolved pollutant, it can also be adsorbed to sediments, both on the land surface and in the waterbody (Yagow et al., 2001). At present, there are insufficient data to parameterize HSPF to allow the user to simulate FC as anything but a planktonic constituent.

When modeling FC with HSPF, the modeler can specify different buildup/washoff relationships for pervious and impervious areas. FC accumulation rate is input by the user for each PERLND and IMPLND segment as a constant input (ACQOP) or varying monthly (MON-ACCUM). Users generally calculate this input rate outside of HSPF using tools such as Bacteria Source Load Calculator (BSLC) (Zeckoski et al., 2005). FC can also be input directly into the streams using input time series. FC concentration in groundwater and interflow can be input to HSPF as a constant or monthly variable.

Die-off of FC on the land-surface is represented indirectly by providing an asymptotic limit of bacteria build up (SQOLIM-PERLND) in HSPF. Users input the asymptotic bacterial build-up limit value in HSPF for each PERLND and IMPLND. Generally, this limit is calculated assuming a first order die-off rate relationship. This asymptotic limit can vary by PERLND, IMPLND, and month. Release of bacteria in overland runoff is controlled by a user-defined parameter (WSQOP) that specifies the amount of runoff needed to wash off 90% of FC. The released FC is modeled in overland runoff, in streams, and in groundwater as a suspended or planktonic constituent. In-stream die-off is modeled using a temperature-dependent first order relationship (Chick's law).

2.3 Uncertainty in Water Quality Modeling

Models are a simplification of reality. With different models, varying levels of simplification are employed (Morgan and Henrion, 1992). Simplification introduces uncertainty in model output. Beven (1989) observed that the equations used in physically based models are good descriptors of processes that occur in a well-defined, spatially homogenous, and structurally stationary model watershed, but are less accurate when describing similar processes occurring in complex, multi-dimensional, heterogeneous, temporally variable “real” watersheds.

Beck (1987) expressed that uncertainty in water quality modeling is pervasive. The sources of uncertainty in water quality modeling can be broadly classified into knowledge uncertainty and stochastic variability. Knowledge uncertainty results when a modeler does not have complete knowledge about the modeled system or the parameters representing the system. Knowledge uncertainty is a property of the analyst conducting the study and the data available (Helton, 1994). Stochastic variability is a property of the system being modeled and arises because a system can behave in many different ways, as is expected in natural systems (Helton, 1994).

In TMDL development, a margin of safety is often included to account for the inherent uncertainty in water quality modeling. However, typically no formal calculation is performed to estimate the model or modeling uncertainty. As the cost of developing and implementing a TMDL is projected to be more than $1 billion per year (USEPA, 2001), it is important that the modelers quantify the uncertainty that is present in modeling estimates.

2.4 Estimating Uncertainty in Water Quality Modeling

Rigorous uncertainty analysis in water quality modeling is rare (Stow et al., 2007). Existence of too many competing methods of conducting uncertainty analysis and interpreting the results is considered a hindrance in rigorous water quality modeling uncertainty analysis (Pappenberger and Beven, 2006). One of the few attempts at estimating the propagation of parameter uncertainty when predicting in-stream FC concentration modeled using HSPF was conducted by Paul et al. (2004). They used first-order variance analysis to estimate the fraction of variance in simulated peak in-stream FC concentration that could be attributed to input parameters. They interpreted that small uncertainties in model input parameters can result in large uncertainties in predicted FC concentration. They realized the limitations of first-order variance methodology in uncertainty analysis and recommended using Monte Carlo based methods to estimate predictive uncertainty in in-stream FC concentration. Donigian and Love (2007) used Monte Carlo simulations to estimate uncertainty in hydrology and sediment modeling in a HSPF model developed for Housatonic river watershed.

Stow et al. (2007) used various Monte-Carlo techniques in combination with a simple Streeter-Phelps dissolved oxygen model to estimate input parameters and predictive uncertainty. Stow and his colleagues concluded that as a model becomes more complex, with more and more parameters, it becomes increasingly difficult to effectively and efficiently sample appropriate parameter space for uncertainty analysis. In the following sections, different techniques for propagating parameter uncertainty are discussed.

2.4.1 First order approximation

In first order approximation (FOA), variance of the output, $Var(O)$ is estimated as

$$Var(O) = \sum_{i=1}^{N} S_i^2 Var(P_i) \tag{2.2}$$

Where, S_i is the absolute sensitivity of the model output with respect to the parameter P_i and N is the number of sensitive parameters. The fraction of the total variance of the output, F_i can be attributed to a particular input parameter as

$$F_i = \frac{S_i^2 Var(P_i)}{\sum_{i=1}^{N} S_i^2 Var(P_i)} \tag{2.3}$$

FOA is computationally simpler to apply than other uncertainty estimations techniques and, therefore, has been widely used for uncertainty analysis (Tyagi and Haan, 2001). FOA assumes that the model has linear functional relations, small coefficients of variations of sensitive parameters, and near normal parameter distributions (Tyagi and Haan, 2001). In hydrologic and water quality modeling, these assumptions are rarely satisfied. Despite the shortcomings, FOA has been used by researchers to obtain information about various models and their parameters, and the effects on uncertainty in the model output. However, researchers also recommended using Monte Carlo simulation for uncertainty analysis. In this research, FOA was not investigated as an uncertainty quantification technique.

2.4.2 Monte Carlo Simulation

In a Monte Carlo (MC) simulation, repeated runs of the model in question are executed using randomly selected input parameter values. The parameter values are chosen randomly for each simulation from a predetermined parameter-specific probability distribution. The process is repeated for a number of runs sufficient to converge on an estimate of the probability distribution of output variables (Gardner and O'Neill, 1983). As a contrast to deterministic modeling where only a single set of input parameter values is used to simulate water quality output, MC simulations can be used to estimate the water quality output resulting from a set of parameters derived from predefined parameter distributions.

The predefined parameter probability distributions reflect parameter uncertainty. Parameter distributions can be obtained from a review of the pertinent literature, historical data, professional judgment, or other uncertainty estimation techniques like Generalized Likelihood Uncertainty Estimation (GLUE), Bayesian Monte Carlo (BMC), and Markov Chain Monte Carlo (MCMC), which are discussed later. Depending upon the model and existing knowledge of the parameters, the modeler may need to provide the covariance among the parameters to sample the parameter values effectively. As discussed later, the parameter distributions may reflect the covariance implicitly if they are obtained from techniques like GLUE and MCMC.

2.4.3 Two-phase Monte Carlo simulation

A two-phase Monte Carlo approach can be used to propagate and analyze stochastic variability and knowledge uncertainty separately in hydrology and water quality models (Hession et al., 1996). In a two-phase MC (TPMC) procedure, model parameters are classified as either knowledge uncertain or stochastically variable. Parameters about which knowledge is limited, or there is insufficient field data available to estimate their values, are considered knowledge uncertain. Values for these parameters are typically obtained through a model-calibration process. Stochastic parameters are those that vary spatially and/or temporally. Information about these parameters (typically a probability distribution) is generally estimated using available data and/or best professional judgment. Some parameters may be classified as both, knowledge uncertain and stochastically variable.

Suppose a model has sensitive parameters ***a***, ***b***, ***x***, and ***y*** of which the parameters ***a*** and ***b*** are knowledge uncertain and the parameters ***x*** and ***y*** are stochastically variable. To perform a TPMC analysis, *i* sets of knowledge uncertain parameters are generated by randomly sampling from predefined parameter probability distributions (figure 2.1). For each set of ***a*** and ***b*** parameter values, a set of *n* random values are generated for the stochastic parameters, ***x*** and ***y***, from their respective predefined distributions (figure 2.1). The model is run for the ***n*** stochastically variable parameter values and the output is plotted as a cumulative distribution function (CDF). The CDF defines the probability of a given output (Helton, 1994). Each CDF represents the output distribution due to stochastic variability. Similar CDFs are generated for the *i* sets of knowledge uncertain parameters. The resulting family of CDF curves describes both the knowledge uncertainly and stochastic variability.

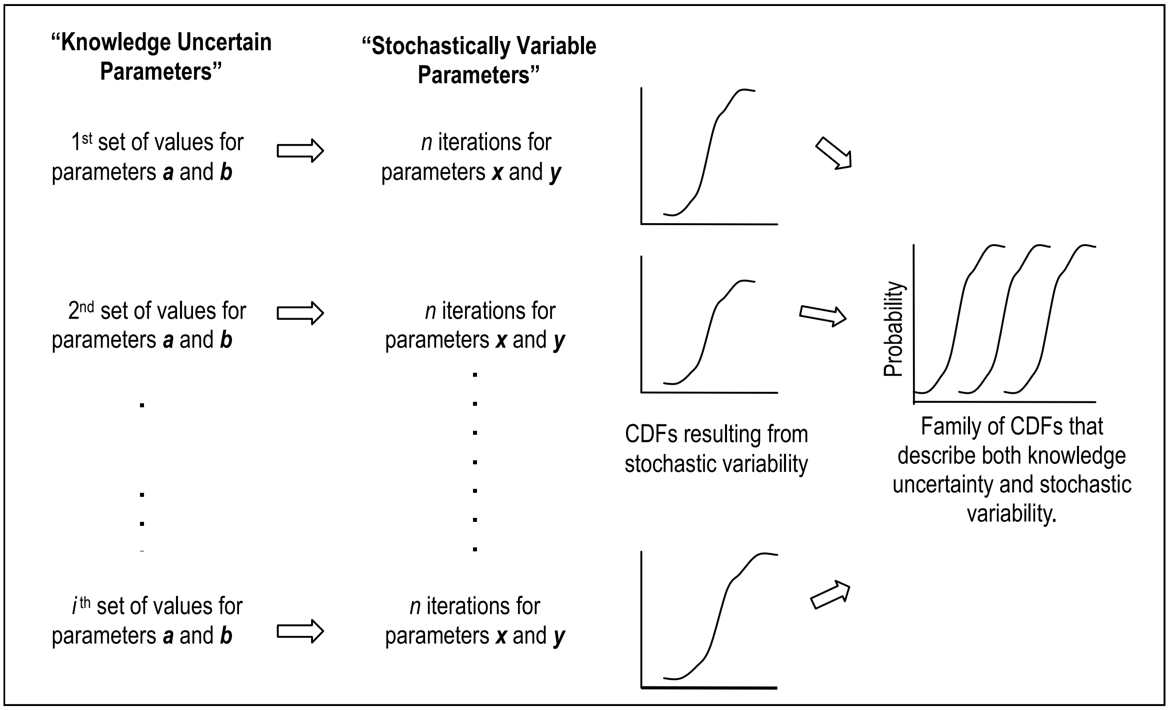


Figure 2.1 A two-phase Monte Carlo analysis to illustrate the effect of knowledge uncertainty and stochastic variability (adapted from Hession et al., 1996). Used under fair use guidelines, 2011

2.4.4 Regionalized Sensitivity Analysis

Regionalized Sensitivity Analysis (RSA) or Generalized Sensitivity Analysis (GSA) is a MC sampling approach to evaluate the sensitivity of model parameters suggested by Hornberger and Spear (1981). RSA can be used for selecting future sampling parameter distributions. To conduct RSA, the modeler must first define the range of key response variables as behavioral (within an acceptable/reasonable range) or non-behavioral (outside an acceptable range). The modeler then samples parameters from a set of predefined parameter distributions that are called "prior distributions." These prior distributions reflect the knowledge of the modeler about the parameters that define the system. Parameter samples generated from prior distributions are used in the model to simulate key response variables.

The parameter sets generating responses in the acceptable behavioral range are accepted and the remaining outputs are rejected. The CDFs of parameter sets that generated both behavioral and non-behavioral response variables are compared. If the two CDFs are significantly different, then the key response variables are sensitive to the parameters, and vice-versa. This exercise is conducted to identify critical uncertainties in the present knowledge of the system that can be used to better plan future research. In this research, RSA was not investigated; instead, Generalized Likelihood Uncertainty Estimation (GLUE) that is a successor of RSA was investigated.

2.4.5 Generalized Likelihood Uncertainty Estimation (GLUE)

Generalized Likelihood Uncertainty Estimation (GLUE) is a successor of RSA proposed by Beven and Binley (1992). The basic premise of GLUE is that there is not a single optimum set of parameters for a hydrologic model. Instead, there are multiple sets of parameters that acceptably represent a hydrologic model – a phenomena known as "equifinality."

In the GLUE approach, MC simulation is performed by generating different sets of parameters from prior distributions. In the majority of previous GLUE applications found in the literature, the prior distributions of parameters were uniform (Beven, 2001). A likelihood weight is assigned to all the parameter sets depending upon their ability to be a simulator of the system. When using GLUE, the likelihood term can be evaluated using any "goodness of fit" criterion that is used to compare observed and simulated response variables (Stow et al., 2007). This likelihood definition is different from the statistical definition of "likelihood function" and is a controversial aspect of GLUE (Stedinger et al., 2008).

In GLUE, the likelihood can be calculated using several different methods and can take into account one or more response variables. Beven and Binley (1992) illustrated several ways to calculate likelihood values. The likelihood values can be based on a single or multiple observed responses as illustrated in equations 2.4 and 2.5, respectively (for a more detailed list, refer to Beven and Binley (1992)).

$$L_e = (\sigma_e^2)^{-N} \tag{2.4}$$

Where,

$$\sigma_e^2 = 1/n\left(\sum_{i=1}^{n}(Y_i - Q_i)^2\right)$$

L_e = likelihood value,

σ_e^2 = variance of the residuals or mean square error

n = number of data points

Y_i = observed data point

Q_i = simulated data point

N = shaping parameter, chosen by the user.

$$L_m = \left\{ \sum_{j=1}^{m} \frac{(W_j)}{\sigma_{ej}^2} \right\}^N \qquad 2.5$$

where, L_m = likelihood function based on m observed responses

W_j = weight of response variable j

σ_{ej} = error variance of j^{th} response variable

When applying GLUE, a parameter vector θ is generated for each model run, and each model run results in a likelihood value. All the parameter sets having likelihood values that are in the acceptable range are retained for consideration. The likelihood values of the retained parameters sets are normalized so that the sum of likelihood values is unity. The normalized likelihood values can be treated as the probabilistic weighting function for the predicted variables and can be used to assess the uncertainty associated with the predictions. A distribution function of the predicted output may be calculated by plotting the predicted values against the likelihood of each prediction. Defining the uncertainty limits as the 5^{th} and 95^{th} percentile of the cumulative likelihood distribution yields a 90% probability interval.

Likelihood weights can be used to update the prior parameter distribution of input parameters using Bayesian equation (equation 2.6) (Fisher, 1922). The distribution of parameters resulting due to updating with available data is called the "*posterior distribution*".

$$L_p(\theta \mid y) = L_y(\theta \mid y) L_o(\theta) \qquad 2.6$$

where, $L_O(\theta)$ = prior distribution of parameters

$L_y(\theta|y)$ = calculated likelihood function of the parameter sets, and

$L_p(\theta|y)$ = posterior likelihood distribution of parameter sets

As evident from the equation 2.6, posterior distributions are a result of modeler's prior knowledge about the system and the observed data. Beven and Binley (1992) showed that the uncertainty is reduced when the likelihoods are updated with new observations. However, they also note that uncertainty cannot decline indefinitely, and may increase, as the hydrological parameters are stochastic in nature and behave differently in different storm events in the same watershed. This property limits the possibility of finding one optimum parameter set for a watershed model. Freer at al. (1996) applied GLUE to a simple hydrologic model, TOPMODEL to evaluate predictive uncertainty using different likelihood measures. They also demonstrated that using additional data to update the likelihood function could help to constrain the uncertainty bound of model prediction.

The GLUE approach has been widely used to conduct uncertainty analysis using different hydrologic modeling software (Balin, 2004; Beven and Binley, 1992; Freer et al., 1996). There have been a few attempts to quantify uncertainty in water quality modeling using GLUE (e.g. Benaman and Shoemaker, 2002; Setegn el al., 2009; Stow et al., 2007; Zheng and Keller, 2007) and GLUE has been suggested as a viable approach to estimate uncertainty in TMDLs (Stow et al., 2007). The application of GLUE for estimating uncertainty in a watershed model developed using HSPF or for TMDL development is practically non-existent.

2.4.6 Bayesian Monte Carlo Uncertainty Analysis

Bayesian Monte Carlo (BMC) uncertainty analysis (Dilks et al., 1992) is also a successor of the RSA technique illustrated by Hornberger and Spear (1981). However, as opposed to RSA, BMC does not categorize model outputs as acceptable or non-acceptable. In BMC, the likelihood function of each parameter set is used to weight the parameter set. The parameter sets with greater likelihood values have greater weight than the parameter sets with lower likelihood value. In this approach, the model assumes an error (ε) such as

$$Y = g[x,\theta] + \varepsilon \qquad 2.7$$

Where,

Y = response variable

g = model that is a function of state variable *x*, and input parameter *θ*

ε = model error that is normally distributed

The likelihood function for each parameter set in BMC is defined as (Dilks et al., 1992).

$$L(\theta \mid Y) = 1 \Big/ \sqrt{2\pi\sigma}\, \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \left(\frac{\varepsilon_i}{\sigma} \right)^2 \right] \qquad 2.8$$

where,

L(θ/Y) = likelihood function,

ε_i = error term at the individual data point, i,

n = number of observed data points, and

σ = standard deviation of the data error.

The likelihood varies as the function of data error, number of data points, and the standard deviation of data error. With high standard deviation, the value of likelihood remains constant over a wide range of data error, however, with low standard deviation, the likelihood value decreases as the model error increases. For multiple state variables, likelihood can be calculated as

$$L(\theta \mid e_1, e_2 .....e_n) = \prod_{i=1}^{n}\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi\sigma_j}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{e_{ij}}{\sigma_j}\right)^2\right] \qquad 2.9$$

Dilks et al. (1992) applied the BMC technique to a Grand River dissolved oxygen model in Michigan. When the resulting parameter posterior distributions were used, uncertainty decreased significantly. Two of the applications of BMC include analysis of estimates for managing Lake Erie levels (Venkatesh and Hobbs, 1999), and a biochemical oxygen demand (BOD) decay model (Qian et al., 2003). No applications of BMC with HSPF have been published.

2.4.7 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a commonly used technique for Bayesian inference by statisticians (Kass et al., 1998). It can also be termed as a special case of BMC. The MCMC method generates samples of parameter values from the posterior distribution by constructing a Markov Chain that has the posterior distribution as its equilibrium distribution (Robert and Casella, 2004).

Metropolis et al. (1953) proposed an algorithm to build a Markov chain. An important step in building a Markov chain is the choice of a statistical likelihood function. The statistical likelihood function in MCMC is similar to the likelihood function used in BMC. For *n* observations, as is the case with time-series output, the likelihood function is given by

$$L(\theta \mid Y) = \frac{1}{(\sqrt{2\pi})^n\sigma^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - Q_i)^2\right) \qquad 2.10$$

Where, σ = variance of residuals, Y_i = *i* th observed data point, and Q_i = *i* th simulated data point

This equation assumes that the residuals between observed and simulated values, or the errors, are normally distributed. To build a Markov chain where a new parameter value is sampled using the previous value, a jump specification is required. The new parameter vector

(θ_{new}) is sampled near the previous parameter vector (θ_{old}) using the symmetric probability distribution or $\pi(\theta_{old}|\theta_{new}) = \pi(\theta_{new}|\theta_{old})$. This symmetric distribution is centered on the last accepted parameter value by the relationship $\theta_{new}|\theta_{old} = N(\theta_{old}, s \cdot I)$, where s is the variance scaling factor and I is the identity matrix. The variance scaling factor affects the movement of Markov chain towards equilibrium. A high variance scaling factor might lead to slow chain movement and a very small variance scaling factor can result in haphazard parameter chain movement in all the possible parameter space. Although there is guidance to estimate the scaling factor, it is typically obtained by trial and error (Gelman et al., 2000).

Once the new parameter set is obtained, it is either accepted or rejected. This step is central point to the Metropolis algorithm. Acceptance or rejection of the new parameter set is determined by the ratio of the posterior probability density functions from the new and old parameter sets (equation 2.11).

$$r = \frac{\pi(\theta_{new} \mid Y)}{\pi(\theta_{old} \mid Y)} \tag{2.11}$$

The Metropolis algorithm rule is used to accept or reject a new parameter:

If $\boldsymbol{r} > 1$, accept the new parameter set

If $\boldsymbol{r} < 1$, generate a random number $\boldsymbol{u}$ from a uniform distribution [0,1]

If $\boldsymbol{r} > \boldsymbol{u}$, accept the new parameter set

If $\boldsymbol{r} < \boldsymbol{u}$, reject the new parameter set

There are several special cases of the Metropolis algorithm, including Metropolis-Hastings Algorithm (Hastings, 1970), Gibbs algorithm (Geman and Geman, 1984), and Metropolis within Gibbs Algorithm (Gelfand and Smith, 1990). These algorithms provide different ways to sample new parameter values from the old parameters. The MCMC approach has been used by many researchers to estimate posterior distributions and uncertainty with various hydrologic modeling software (Balin, 2004; Kuczera and Parent, 1998; Makowski et al., 2002; Marshall et al., 2004), and it has been suggested as a viable approach for estimating uncertainty in water quality modeling (Stow et al., 2007). However, application of the MCMC approach to estimate uncertainty in a watershed-scale water quality model has been limited. In this research, MCMC was used as a Bayesian technique to estimate uncertainty in a water quality model developed using HSPF.

2.5 Summary

Water quality modeling is often central to TMDL development and other similar watershed management efforts. It is widely recognized that the added information about the inherent modeling uncertainty in water quality modeling can aid stakeholders and decision makers in making more informed watershed management decisions. Stakeholders and decision makers can use uncertainty information to help decide among different water quality management plans and/or to direct planning efforts towards specific pollution sources.

With the advent of faster computers, Monte Carlo methods have gained popularity as viable techniques for uncertainty analysis. However, the presence of many competing uncertainty estimation techniques makes it difficult to conduct the uncertainty analysis and interpret the results. Bayesian uncertainty estimation techniques have been shown to be particularly useful since they allow the model parameters to be updated as the new data becomes available. Stow et al. (2007) compared several of these Monte Carlo based techniques using a simple Streeter-Phelps model. They suggested MCMC as a viable technique to conduct uncertainty analysis for complex watershed models. Qian et al. (2003) compared MCMC and BMC on a simple BOD decay model and suggested MCMC as a better uncertainty estimations approach for higher dimensional models. Most of these techniques have, however, rarely been used on watershed-scale water quality models that are often used in developing watershed management plans, like HSPF. The research presented here performs uncertainty analysis associated with HSPF modeling of FC in a small watershed in Virginia using single-phase MC simulation, two-phase MC, GLUE, and MCMC. These techniques and their results were compared with each other and suggestions were made for future research and applications of these techniques.

References:

Balin, D. 2004. Hydrological Behaviour through Experimental and Modelling Approaches. Application to the Haute-Mentue Catchment. School of Acrhitecture, Civil and Environmental Engineering. Lausanne, Switzerland. Ecole Polytechnique Fédérale de Lausanne.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Benaman, J., and C.A. Shoemaker. 2002. Sensitivity and uncertainty analysis of a distributed watershed model for the TMDL process. National TMDL Science and Policy 2002 Speciality Conference, Water Environment Federation.

Benham, B.L., C. Baffaut, R.W. Zeckoski, K.R. Mankin, Y.A. Pachepsky, A.M. Sadeghi, K.M. Brannan, M.L. Soupir, and M.J. Habersack. 2006. Modeling bacteria fate and transport in watersheds to support TMDLs. *Tran. ASABE.* 49(4): 987-1002.

Benham, B.L., K. Branna, K. Christophel, T. Dillaha, L. Henry, S. Mostaghimi, R. Wagner, J. Wynn, G. Yagow, and R. Zeckoski. 2004. Total maximum daily load development for Mossy Creek and Long Glade Run: Bacteria and general standard (Benthic) impairments. Richmond, Va.: Virginia Department of Environmental Quality. Available at http://www.deq.state.va.us/tmdl/homepage.html. Assessed Jan 25, 2005.

Benham, B.L., K. Brannan, T. Dillaha, S. Mostaghimi, G. Yagow. 2002. TMDLs (Total Maximum Daily Loads) – Terms and Definitions. *Virginia Cooperative Extension.* Pub No. 442-550.

Beven, K.J. 2001. *Rainfall-Runoff Modeling – The Primer,* John Wiley and Sons, New York, 360p.

Beven, K. 1993. Prophecy, Reality and uncertainty in distributed hydrological modeling. *Adv. Water Resources.* 16(1): 41-51.

Beven, K.. and A. Binley (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Beven, K. 1989. Changing Ideas in Hydrology – The case of Physically Based Models. *Journal of Hydrology.* 105(1-2): 157-172.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr. T.H. Jobes, and A.S. Donigian, Jr. 2005. *HSPF Version 12.2 User's Manual.* AQUA TERRA Consultants. Mountain View, CA.

Bingner, R.L., and F.D. Theurer. 2001. AnnAGNPS Technical Processes: Documentation Version 2. Unpublished report. Oxford, Miss.: USDA-ARS National Sedimentation Laboratory.

Dilks, D.W., R.P. Canale, and P.G. Meier. 1992. Development of Bayesian Monte Carlo Techniques for Water Quality Model Uncertainty. *Ecological Modeling.* 62(1-3): 149-162.

Donigian, A.S., and J.T. Love. 2007. The Housatonic River Watershed Model: Model Application and Uncertainty Analyses. 7th International IWA Symposium on System Analysis and Integrated Assessment in Water Management, May 7-9, 2007. Washington, DC. WATERMATEX Proceedings on CD-ROM.

Fisher, R.A. 1922. On the Mathematic Foundation of Theoretical Statistics. *Phil. Trans. Roy. Soc. London, A.* 222: 309-368.

Freer, J., K. Beven, and B. Ambroise. 1996. Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach. *Water Res. Res.* 32(7): 2161-2173.

Gardner, R.H., and R.V. O'Neil. 1983. Parameter Uncertainty and Model Predictions: A Review of Monte Carlo Results. In *Uncertainty and Forecasting of Water Quality, eds.* M.B. Beck, and G. Van Straten, 345 -257. Berlin, Germany: Springer-Verlag.

Gelfand, A.E., and A.F.M. Smith. 1990. Sampling Based Approaches to Calculating Marginal Densities. *J. Amer. Statistical Ass.85:* 398-409.

Gelman, A., J.B. Carlin, H.S. Stern, and R.D.B. 2000. Bayesian Data Analysis. Boca Raton, London, New York, Washington D.C., Chapman&Hall/CRC.

Gelman A., and D. Rubin. 1992. Inference from Iterative Simulation using Multiple Sequences. Statistical Science. 7: 457-511.

Geman, S., and D. Geman. 1984. Stochastic Relaxation, Gibbs' distribution and Bayesian restoration of images. IEEE trans. PAMI 6: 721-741.

Hastings, W.K. 1970. Monte Carlo methods using Markov chains and their applications. *Biometrika.*57: 97-109.

Helton, J.C. 1994. Treatment of uncertainty in performance assessment for complex systems. *Risk Analysis*. 14(4): 483-511.

Hession, W.C., D.E. Storm, and C.T. Haan. 1996. Two-phase uncertainty analysis: an example using universal soil loss equation. *Trans. ASAE.* 39(4): 1309-1319.

Hornberger, G.M., and R.C. Spear. 1981. An approach to the preliminary analysis of environmental systems. *Journal of Env. Mgmt (12):* 7-18.

Im, Sangjun, K.M. Brannan, S.M. Mostaghimi, and Jaepil Cho. 2004. Simulating Fecal Coliform Bacteria Loading from an Urbanizing Watershed. *Journal of Environmental Science and Health.* A39(3): 663-679.

Kass, R.E., B.P. Carlin, A. Gelman, R.M. Neal. 1998. Markov Chain Monte Carlo in Practice. A roundtable Discussion. *The American Statistician.* 52(2): 93-100.

Kuczera, G., and E. Parent. 1998. Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: the Metropolis Algorithm. *Journal of Hydrology. 211: 69-85.*

Makowski, D., D. Wallach, and M. Tremblay. 2002. Using a Bayesian approach to parameter estimation; Comparison of GLUE and MCMC methods. *Agronomie*. 22: 191-203.

Marshall, L. D. Nott, and A. Sharma. 2004. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. Water Res. Research. 40(2). doi:10.1029/2003WR002378, 2004.

MacIntosh, D.L., G.W. Suter., and F.O. Hoffman. 1994. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. *Risk Analysis*. 14(4): 405-419.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. J. Chemical Phys. 21(6): 1087-1092.

Moore, J.A., J.D. Smythe, E.S. Baker, J.R. Miner, and D.C. Moffitt. 1989. Modeling Bacteria Movement in Livestock Manure Systems. Transactions of the ASAE, 32(3): 1049-1053.

Morgan, M.G., and M. Henrion. 1992. *Uncertainty.* Cambridge University Press, Cambridge, UK.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2005. Soil and Water Assessment Tool Theoretical Documentation. Version 2005. Grassland Soil and Water Research Laboratory, Agriculture Research Service. Temple, TX.

Overcash, M.R., K.R. Reddy, and R. Khaleel. 1983. Chemical processes and transport of animal waste pollutants. In *Agricultural Management and Water Quality,* 109-125. Ames, Iowa: Iowa State University Press.

Pappenberger, F. and K.J. Beven. 2006. Ignorance is Bliss: Or seven reasons Not to Use Uncertainty Analysis. *Water Res. Res.* 42(5). W05302, doi:10.1029/2005WR004820.

Paul, S., P.K. Haan, M.D. Matlock, S. Mukhtar, and S.D. Pillai. 2004. Analysis of the HSPF water quality parameter uncertainty in predicting peak in-stream fecal coliform concentrations. Trans. ASAE. 47(1): 69-78.

Qian, S.S., C.A. Stow, and M.E. Borsuk. 2003. On Monte Carlo methods for Bayesian inference. Ecol. Modeling. 159(2-3): 269-277.

Robert, C.P., and G. Casella. 2004. *Monte Carlo Statistical Methods.* Springer-Verlag, New York City, NY.

Rosen, B.H. 2000. *Waterborne Pathogens in Agriculture Watersheds.* NRCS, Watershed Science Institute, Burlington, VT.

Setegn, S.G., R. Shrinivasan, A.M. Melesse, B. Dargahi. 2009. SWAT model application and uncertainty analysis in the Lake Tana Basin, Ethiopia. *Hydological processes.* 24(3): 357-367

Springer, E.P., G.F. Gillford, M.P. Windham, R. Thelin, M. Kress. 1983. Fecal coliform release studies and development of a preliminary nonpoint source transport model for indicator bacteria. Logan, Utah: Utah State University, Utah Water Research Laboratory.

Stedinger, J.R., R.M. Vogel, S.U. Lee, and R. Betchelder. 2008. Appraisal of the Generalized Likelihood Uncertainty Estimation (GLUE) Method. *Water Res. Res.* 44. W00B06, doi:10.1029/2008WR006822.

Stow, C.A., K.H. Reckhow, S.S. Qian, E.C. Lamon, G.B. Arhonditsis, M.E. Borsuk, and D. Seo. 2007. Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. *J. Amer. Water Resources Ass.* 43(6): 1499-1507.

Tyagi, A., and C.T. Haan. 2001. Uncertainty Analysis using Corrected First Order Approximation method. *Water Resources Research.* 37(6): 1847-1858.

USEPA (United States Environmental Protection Agency). 2011. Virginia Water Quality Assessment Report. Available: http://iaspub.epa.gov/waters10/attains_state.control?p_state=VA (Accessed 11 Feb 2011)

USEPA. 2009. National Section 303(d) list fact sheet. [cited 2009 Nov]. Available from: URL: http://oaspub.epa.gov/waters/national_rept.control.

USEPA. 2002. *2000 National Water Quality Inventory*. USEPA Office of Water. Available at: http://www.epa.gov/305b/2000report Accessed May, 2004.

USEPA. 2001. The National Cost of the Total Maximum Daily Load Program (Draft Report). Office of Water, United States Environmental Protection Agency. Washington D.C.

VADCR. 2003. Bacteria TMDLs for Abrams Creek and Upper and Lower Opequon Creek Located in Frederick and Clarke County, Virgnia. Virginia Department of Environmenatal Quality, and Virginia Department of Conservation and Recreation. Richmond, VA, USA.

Venkatesh, B.N., and B.F. Hobbs. 1999. Analyzing investments for Managing Lake Erie Levels Under Climate Uncertainty. *Water Resources Research.* 35(5): 1671-1684.

Walker, S.E., S.M. Mostaghimi, T.A. Dillaha, F.E. Woeste. 1990. Modeling animal waste management practices: Impacts on bacteria levels in runoff from agricultural lands. *Trans. ASAE* 33(3): 807-817.

Yagow, G.W., T.A. Dillaha., S.M. Mostaghimi, K.M. Brannan, C.D. Heatwole, and M.L. Wolfe. 2001. ASAE Annual International Meeting, Sacramento, CA. Paper – 01-2066.

Young, R.A., C.A. Onstad., D.D. Bosch, and W.P. Anderson. 1987. AGNPS, Agricultural Nonpoint Source Pollution Model: A Watershed Analytic Tool. Conservation Research Report 35. Washington, D.C.,: USDA.

Zeckoski, R. W., B.L. Benham, S. B. Shah, M.L. Wolfe, K.M. Brannan, M. Al-Smadi, T.A. Dillaha, S. Mostaghimi, C.D. Heatwole. 2005. BSLC: A tool for bacteria source characterization for watershed management. *Applied Engg. In Agriculture.*21(5): 879-889

Zheng, Y., and A.A. Keller. 2007. Uncertainty Assessment in Watershed-Scale Water Quality Modeling and Management:1. Framework and Application of Generalized Likelihood Uncertainty Estimation (GLUE) Approach. *Water Res. Res.* 43. doi: 10.1029/2006WR005345

Chapter 3. Evaluation of the applicability of single-phase and two-phase Monte Carlo analysis to estimate uncertainty in HSPF based water quality modeling.

Abstract. *Single-phase and two-phase Monte Carlo (MC) simulations were performed to estimate overall predictive uncertainty in violations of in-stream fecal coliform (FC) concentration for a Hydrological Simulation Program–FORTRAN (HSPF) model developed for Mossy Creek Total Maximum Daily Load (TMDL) in Virginia. Additionally, two-phase MC was also used to partition the effects of knowledge uncertainty and stochastic variability on output uncertainty. The two techniques were used in conjunction with two FC pollutant allocation scenarios presented in the Mossy Creek bacterial TMDL. The scenarios differed in the reductions specified from cattle directly depositing FC in the stream, and FC loadings from cropland. As estimated by the two techniques, the instantaneous FC criterion was violated less than one percent of the time (on daily basis) during the prediction period for both the allocation scenarios. However, the violations increased as high as 8% (two-phase MC) and 14% (single-phase MC) when 97.5% quantile of the output FC concentration was plotted for the scenario allowing a greater amount of direct deposit of FC in the stream. The two-phase MC results illustrated that cattle direct deposit of FC is a greater source of knowledge uncertainty than cropland FC loadings. Decision makers can use the results of an assessment like this to choose their level of confidence in achieving a water quality standard, selecting among the scenarios or prioritizing implementation efforts. Among single- and two-phase MC simulation, single-phase Monte Carlo is more computationally efficient while two-phase MC simulation can provide additional information about the effect of knowledge uncertainty and stochastic variability. With respect to watershed-scale water quality modeling, a satisfactory and unambiguous model parameter categorization may be difficult to achieve limiting the applicability of two-phase Monte Carlo simulation in these kinds of applications.*

Keywords. Water quality modeling, HSPF, Monte Carlo, two-phase Monte Carlo, TMDL, indicator bacteria, fecal coliform.

Introduction

Water quality models are often used to develop total maximum daily loads (TMDLs). A TMDL quantifies the amount of a given pollutant a waterbody can receive and still meet water quality standards. It includes pollution from permitted point sources, nonpoint, and natural background sources and a margin of safety (Benham et al., 2002). Mathematically, a TMDL can be represented as

$$TMDL = \sum WLA + \sum LA + MOS \qquad 3.1$$

where,

ΣWLA = waste load allocation (point sources), and

ΣLA = load allocation (non-point sources),

MOS = margin of safety

A margin of safety is included within a TMDL to account for the inherent uncertainty present in determining the TMDL. Uncertainty, which is always present when simulating a natural system, is a result of limited knowledge of the process being modeled and inherent stochastic (spatial or temporal) variability within that system (Beck, 1987; Suter et al., 1987). Without some measure of the uncertainty, one cannot accurately assess the probability of achieving a given water quality criteria or the risk of violating it. Although needed, there is limited science-based guidance available on how to estimate the amount of uncertainty associated with a TMDL.

In 2001, the U.S. Environmental Protection Agency (USEPA) estimated the annual average cost of developing TMDLs to be $63-69 million per year for the next fifteen years and the cost of implementing TMDLs to be between $1 and 3.4 billion per year for the next decade (USEPA, 2001). Pathogens, typically represented by a surrogate indicator bacteria (IB), being the second most widespread cause of water quality impairments (USEPA, 2006) will be responsible for a significant share of this expense. In spite of the potentially significant costs associated with developing bacterial impairment TMDLs, there have been only few attempts (e.g. Stow et al., 2007) to quantify the uncertainty associated with the modeling often performed to develop these TMDLs.

Most water quality modeling software currently used when developing TMDLs includes modules that are empirical or a mix of empirical and process-based. These software do not typically include uncertainty analysis capabilities. Hydrological Simulation Program–FORTRAN (HSPF) is a continuous-simulation model that simulates various hydrologic and water quality processes (Bicknell et al., 2005). It is a lumped parameter, watershed scale model, and produces a deterministic time-series of hydrology and water quality. HSPF has been used to develop a significant number of IB impairment TMDLs in Virginia (e.g. Benham et al., 2005; VADCR, 2003; Yagow, 2001).

Uncertainty in model predictions can be estimated using two categories of methods, namely Monte Carlo methods and first-order variance propagation (Beck, 1987; Summers et al., 1993). First-order methods assume linear models, which limit their usability with respect to a complex model like HSPF (Summers et al., 1993). Paul et al. (2004) conducted a first-order analysis to estimate the contribution of sensitive parameters to the fraction of variance in simulated peak in-stream fecal coliform (FC) concentrations (a common IB) in a watershed modeled with HSPF. They inferred that small uncertainties in selected water quality parameters could result in large uncertainties in the prediction of in-stream FC concentration. Paul et al. (2004) realized the limitations of first order variance methods for uncertainty analysis and recommended Monte Carlo based analysis to evaluate uncertainty in IB modeling using HSPF.

Monte Carlo simulation is a method in which repeated simulations of the model in question are performed using randomly selected input parameter sets. Parameter values are selected from parameter-specific probability distributions. The process is repeated for a number of simulations (iterations) sufficient to converge on an estimate of the probability distribution of output variables (Gardner and O'Neill, 1983). The results from these iterations can be aggregated to obtain relevant statistics about model output. Donigian and Love (2007) used Monte Carlo simulations to estimate uncertainty in hydrology and sediment modeling in a HSPF model developed for Housatonic river watershed.

Two-phase Monte Carlo analysis is a MC-based technique that propagates stochastic variability and knowledge uncertainty separately based on the methodology proposed by Helton (1994) and MacIntosh et al. (1994). Stochastic variability is the property of a natural system and can be further divided into spatial and temporal variability. Knowledge uncertainty is due to incomplete understanding of the system being modeled and can also be termed as subjective uncertainty.

Separating knowledge uncertainty and stochastic variability is important to draw useful insights into the model (Helton, 1994; MacIntosh et al., 1994). Knowledge uncertainty can be reduced by collecting more information about the system and hence it can be used as an indicator of the beneficial effects of collecting additional data (Hession et al., 1996). Stochastic variability normally cannot be reduced as it is the natural property of the system, but it can be quantified.

Information about uncertainty in water quality modeling can be used by decision makers and stakeholders to choose their level of confidence in achieving a particular water quality standard and the associated pollutant reductions needed to achieve that confidence level. Thus, an understanding about the source and amount of uncertainty is needed to effectively compare TMDL pollutant allocation scenarios. For this study, a two-phase MC analysis was used to independently evaluate the knowledge uncertainty and stochastic variability associated with predicted in-stream IB concentrations from a watershed model that was developed using HSPF. We also conducted simple Monte Carlo analysis on the same watershed model to compare the two uncertainty estimation techniques and evaluate their applicability.

3.1 Materials and Methods

3.1.1 Monte Carlo Simulation

In a simple or single-phase Monte Carlo (MC) simulation, repeated model runs are performed using parameter values that are randomly selected from a predetermined probability distribution for each simulation. The predetermined parameter-specific probability distributions used in MC are reflective of parameter uncertainty. Parameter distributions can be obtained from a review of the pertinent literature, existing data, best professional judgment, or other uncertainty estimation techniques like Generalized Likelihood Uncertainty Estimation (GLUE), Bayesian Monte Carlo (BMC), and Markov Chain Monte Carlo (MCMC). The application of these other uncertainty estimation techniques is beyond the scope of the research reported here. Depending upon the model and existing knowledge about the parameters, the modeler may also need to provide the covariance among the parameters to sample the parameter values effectively. The parameter distributions that are obtained using some techniques (like GLUE, and BMC) account for covariance implicitly. In this application, the parameters were assumed independent and covariance relationship was not provided.

3.1.2 Two phase Monte Carlo simulation

In a two-phase MC procedure (TPMC), model parameters are classified as either knowledge uncertain or stochastically variable. Parameters about which knowledge is limited, or there are insufficient field data available to estimate their values, are considered knowledge uncertain. Values for these parameters are typically obtained through a model-calibration process. Stochastic parameters are those that vary spatially and/or temporally. Information about these parameters (typically a probability distribution) is generally estimated using the data available. Some parameters may be classified as both, knowledge uncertain and stochastically variable.

Suppose a model has parameters ***a***, ***b***, ***x***, and ***y*** of which the parameters ***a*** and ***b*** are knowledge uncertain and parameters ***x*** and ***y*** are stochastically variable. To perform a TPMC analysis, *i* sets of knowledge uncertain parameters are generated by randomly sampling from a predefined parameter probability distribution (figure 3.1). For each set of ***a*** and ***b*** values, a set of

n random values are generated for the stochastic parameters, ***x*** and ***y***, from the predefined distributions for each parameter. The model is run for the n stochastically variable parameter values and the output is plotted as a cumulative distribution function (CDF). The CDF defines the probability of a given output (Helton, 1994). Each CDF represents the output distribution due to stochastic variability. Similar CDFs are generated for the i sets of knowledge uncertain parameter random values. The resulting family of CDF curves describes both the knowledge uncertainly and stochastic variability.

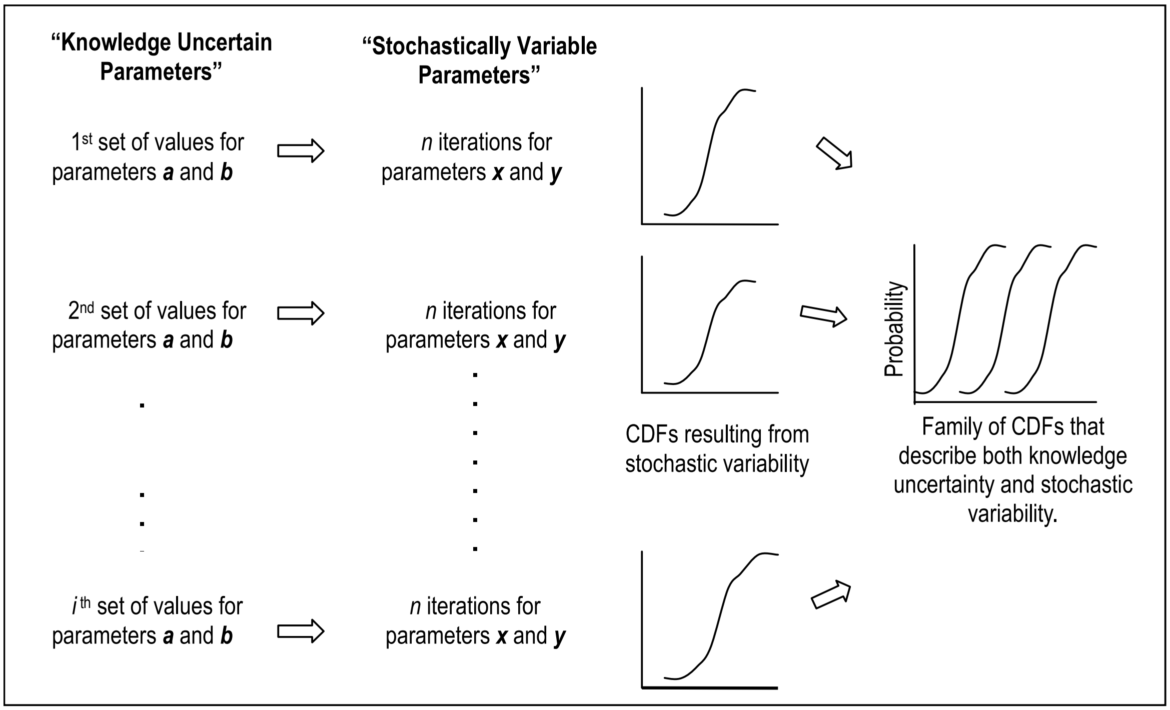


Figure 3.1 A two-phase Monte Carlo analysis to illustrate the effect of knowledge uncertainty and stochastic variability (adapted from Hession et al., 1996). Used under fair use guidelines, 2011

3.1.3 Modeling

3.1.3.1 Study Area

Mossy Creek, located in Rockingham and Augusta counties in Virginia (figure 3.2), was selected for this research. Mossy Creek was listed as impaired in 1996 due to violations of the instantaneous FC criterion and a TMDL was developed for Mossy Creek by the Department of Biological Systems Engineering (BSE) at Virginia Tech (Benham et al., 2004). The Mossy Creek watershed (4076 ha) is characterized as a rolling valley with the Blue Ridge Mountains to the east and the Appalachian Mountains to the west. The predominant land uses in Mossy Creek watershed are forest, pasture, and croplands. The primary sources of FC identified in the Mossy

Creek TMDL were direct deposition of feces in the stream by cattle (cattle loitering and defecating in the stream), and runoff from pastures where grazing animals defecate.

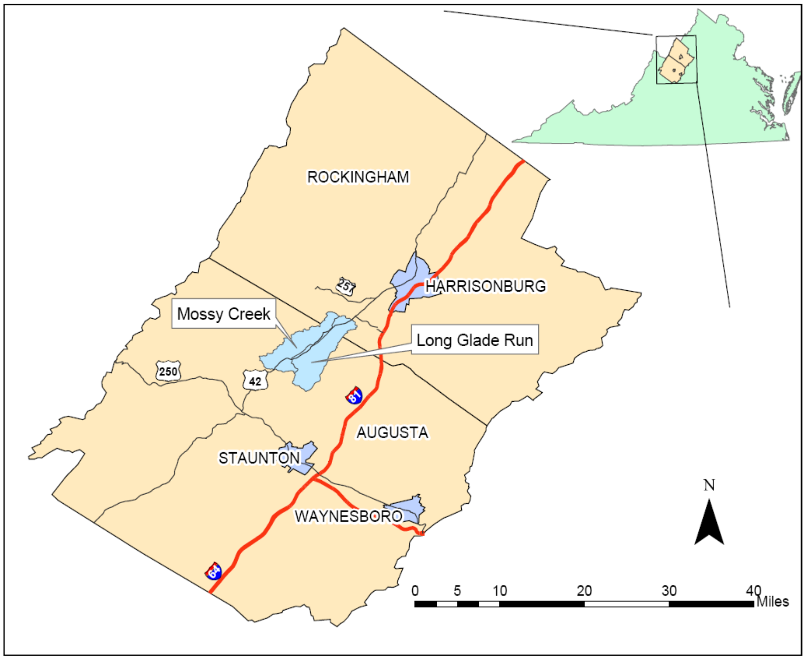


Figure 3.2 Mossy Creek Watershed (Benham et al., 2004). Used under fair use guidelines, 2011

Mossy Creek was monitored monthly by the Virginia Department of Environmental Quality (DEQ) between July 1992 and March 2003 for FC concentration and other selected water quality constituents at the station ID 1BMSS001.35 located near the outlet of the Mossy Creek watershed. BSE monitored Mossy Creek semi-monthly between February 1998 and December 2001 for selected water quality constituents including FC concentration near the DEQ site (Site QMA in figure 3.3). Daily flow data were also collected from May 1998 to December 2002 at the same site.

3.1.3.2 Mossy Creek Watershed Model

HSPF was used to develop the Mossy Creek bacterial impairment TMDL (Benham et. al., 2004). Mossy Creek was divided into eight subwatersheds for modeling and land use identification purposes (figure 3.3, table 3-1). Other data required by the model included rainfall, FC loading from cattle and wildlife, inflows from springs, solar radiation, and temperature as time series. TMDL modeling data development and acquisitions are described in the Mossy Creek TMDL (Benham et. al., 2004).

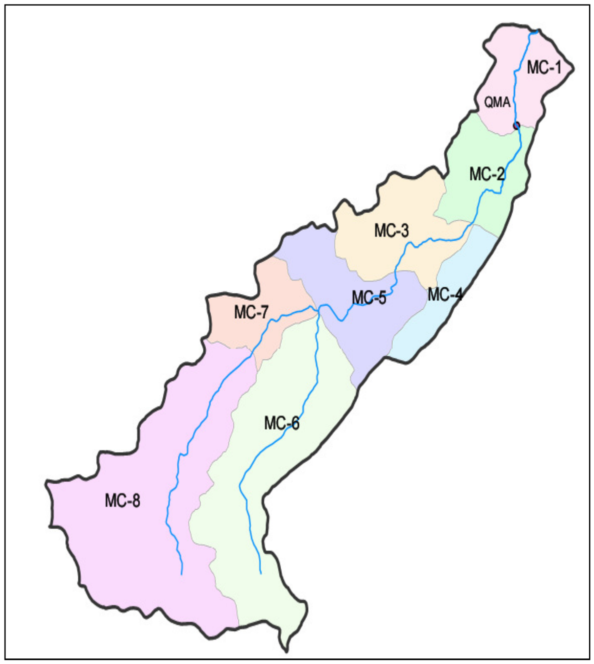


Figure 3.3 Mossy Creek watershed and its subwatersheds (Benham et al., 2004)
Used under fair use guidelines, 2011

Table 3-1 Land use distribution of Mossy Creek watershed (Benham et al., 2004)

Land use	Area (ha)	Percent of total area (%)
Forest	1025.1	25.15
Cropland	556.0	13.64
Pasture	2347.6	57.59
Farmstead	55.0	1.35
Low Density Residential	87.0	2.13
High Density Residential	3.6	0.09
Loafing Lot	1.6	0.04

Used under fair use guidelines, 2011

3.1.3.3 Hydrologic Parameters

Simulation of FC by HSPF requires information about several hydrologic and water quality parameters. BASINS Technical Note 6 (USEPA, 2000) describes the hydrologic parameters and provides typical values and possible limits for all the hydrologic parameters. Typically, the values of these parameters are refined through model calibration. Al-Abed and Whiteley (2002), and Lawson (2003) listed several hydrological parameters that are typically calibrated and thus considered sensitive. This subset of sensitive parameters as described in subsequent sections was used in the MC and TPMC analysis. For the TPMC analysis, the parameters that could be estimated using GIS and field data were considered as stochastic (stochastically variable) and the parameters that were estimated only through calibration were considered knowledge uncertain.

3.1.3.4 Stochastic Parameters

Index to mean infiltration rate (INFILT) was the only model parameter considered to be stochastic as it can be estimated using GIS data for soil and land use. To estimate the distribution of INFILT for each land use, the Mossy Creek watershed was divided into 30 m by 30 m cells. A value of INFILT was assigned to each cell according to the land use and soil type based on guidance from BASINS technical note 6 (USEPA, 2000). A histogram of INFILT values of cells within each land use were then plotted to estimate land use-specific INFILT probability distributions. The INFILT histograms suggested a triangular distribution for all land uses except loafing lot (table 3-2). We assigned a uniform distribution to the INFILT parameter for the loafing lot land use, as it was a small area within the watershed. The limits of the distribution were set as the range of the observed loafing lot INFILT.

Table 3-2 Distribution of stochastically variable parameter INFILT (index to mean infiltration rate, in/hr) by land use.

Land use	Distribution
Forest	Triangular (0.05, 0.1, 1)†
Pasture	Triangular (0.04, 0.09, 0.9)
High Density Residential	Triangular (0.01, 0.01, 0.1)
Cropland	Triangular (0.03, 0.17. 0.24)
Farmstead	Triangular (0.03, 0.15, 0.23)
Low Density Residential	Triangular (0.03, 0.17, 0.26)
Loafing Lot	Uniform (0.15, 0.23)‡

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively.
‡Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

3.1.3.5 Knowledge Uncertain Parameters

The hydrologic parameters that were considered knowledge uncertain are listed in table 3-3 and table 3-4. Table 3-3 contains the parameters that were not varied according to the land use or time of the year. All the parameters listed in table 3-3, except interflow recession coefficient (IRC), were assigned a uniform distribution. Lower and upper limits for the distributions correspond to the typical minimum and maximum limits for these parameters from BASINS technical note 6 (USEPA, 2000). A uniform distribution was the most obvious choice for most parameters, as no additional information is available about these parameters in the BASINS technical note (USEPA, 2000). For IRC, while the typical upper limit is 0.7, it is also chosen as the starting value for calibration process, and can be taken to be the most probable value.

Therefore, IRC was assigned a triangular distribution with lower limit, mode and upper limit as 0.5, 0.7 and 0.7, respectively.

Table 3-3 Distribution of knowledge uncertain hydrology parameters for all land uses

Parameter	Parameter Description	Type of Distribution
LZSN (inches)	Lower zone nominal soil moisture storage	Uniform (3,8)†
AGWRC	Groundwater recession rate	Uniform (0.92, 0.99)
DEEPFR	The fraction of infiltrating water lost to deep aquifers	Uniform (0.0, 0.2)
BASETP	Evapotranspiration by riparian vegetation as active groundwater enters streambed	Uniform (0.0, 0.05)
AGWETP	Fraction of model segment that is subject to direct evaporation from groundwater storage	Uniform (0.0, 0.05)
IRC	Interflow recession coefficient	Triangular (0.5, 0.7, 0.7)‡
INTFW	Coefficient that determines the amount of water which enters the ground from surface detention and becomes interflow	Uniform (1.0, 3.0)

† Numbers in parentheses show lower and upper limit of the uniform distribution, respectively. ‡Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively.

Knowledge uncertain parameters that were varied according to land use and time of year are listed in table 3-4. Again, upper and lower parameter distribution limits were assigned using BASINS technical note 6 (USEPA, 2000). The parameter distribution limits for the month of January are shown in table 3-4. Parameter values for the other months were calculated by multiplying the January distribution by a monthly adjustment factor that was generated based on similar TMDLs developed by BSE, and expert opinion.

Table 3-4 Distribution of hydrologic parameters which vary according to the land use and time of year, for the month of January.

Parameter	Land use	Distribution
UZSN (inches) Nominal upper zone soil moisture storage	Forest	Uniform (0.2, 0.3)†
	Cropland	Uniform (0.06, 0.1)
	Pasture	Uniform (0.06, 0.1)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.06, 0.1)
CEPSC (inches) Interception Storage Capacity	Forest	Uniform (0.05, 0.075)
	Cropland	Uniform (0.05, 0.075)
	Pasture	Uniform (0.05, 0.075)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.05, 0.075)
LZETP Index to lower zone evapotranspiration	Forest	Uniform (0.1, 0.2)
	Cropland	Uniform (0.1, 0.2)
	Pasture	Uniform (0.1, 0.2)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.1, 0.2)

†Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

3.1.3.6 Water Quality Parameters

Simulation of in-stream FC concentrations with HSPF requires the estimation of daily FC loading rates to the land surface (ACQOP), the asymptotic limit of accumulation of FC on the land surface (SQOLIM), and the FC loading rate directly deposited in the streams (direct deposit time series). The daily loading rates and asymptotic limits can be input as tables for monthly varying values, MON-ACCUM, and MON-SQOLIM, respectively. Additional water quality parameters that must be supplied include the IB wash-off potential (WSQOP), first-order decay rate for IB in the waterbody (FSTDEC) and the FSTDEC temperature correction coefficient (THFST). For the Mossy Creek TMDL, FC loading rates for pervious (PERLND) and impervious (IMPLND) land areas, and the FC direct deposit loads were calculated using the Bacteria Source Load Calculator (BSLC) (Zeckoski et al., 2005).

The FC loading rates depend upon several factors including species-specific feces production rates, species-specific fecal densities, die-off rates, animal density, and the fraction of time livestock are confined (Zeckoski et al., 2005). As cited in various TMDL reports and in the literature (ASAE, 2003; Geldrich, 1978; Yagow, 2001) the IB production rates in colony forming units per day (cfu/day) of dairy cattle, beef cattle and poultry can vary by several orders of magnitude. According to the Mossy Creek bacteria TMDL report, dairy cattle, beef cattle, and poultry are responsible for more than 94% of FC production in the watershed (Benham et al., 2004). Therefore, we hypothesized that the uncertainty in production rates of dairy cattle, beef cattle, and poultry are likely to have the greatest impact on FC concentration uncertainty in the Mossy Creek simulation. To incorporate uncertainty in the FC application rates for the appropriate land uses, the loading rates for pervious land areas were assigned a log-triangle distribution. The mode of ACQOP values were the average ACQOP values obtained using the BSLC, and the lower and upper limit of the ACQOP distribution were determined by multiplying the distribution mode by 0.1 and 10, respectively.

The application of manure to cropland varies by month, therefore, after defining a FC cropland loading distribution for January (table 3-5), the remaining months were adjusted using a monthly adjustment factor. This factor was developed using the trend of FC accumulation values obtained by the BSLC for each month. Deterministic values for ACQOP calculated using the BSLC were used for the other land uses – forest, low and high density residential, farmstead, and impervious areas.

Table 3-5 Summary of water quality parameters which have been reported as sensitive and are typically calibrated in hydrologic modeling.

Parameter	Land use	Type of Distribution
ACQOP-PERLND (cfu day^{-1}) (Accumulation of fecal coliform on pervious land per day)	Pasture Loafing Lot Cropland (January)	Log-triangle (1×10^9, 1×10^{10}, 1×10^{11})† ‡ Log-triangle (1.12×10^{11}, 1.12×10^{12}, 1.12×10^{13}) Log-triangle (2×10^6, 2×10^7, 2×10^8)
SQOLIM adjustment Factor (Factor which is multiplied to ACCUM values to obtain SQOLIM	All	Uniform (2.5, 11.5)*
SQOLIM-PERLND (Maximum accumulation of FC on pervious land)	All	ACQOP-PERLND (for each land use) x SQOLIM adjustment Factor
WSQOP-PERLND (Rate of surface runoff that will remove 90% of stored bacteria from pervious land surface)	All	Uniform (0.5, 2.4)
FSTDEC (day^{-1}) (First order decay rate of bacteria)	All	Triangular (0.12, 1.1, 2.52)

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡Log-triangle distribution implies that the logarithm of lower limit, mode and higher limit follows a triangular distribution. *Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

The SQOLIM parameter is typically calculated by multiplying ACQOP values by a factor of nine. This SQOLIM adjustment factor is based on the assumption that the die-off coefficient for FC on pervious land surface is 0.051 day^{-1} (base 10) (Zeckoski et al., 2005). Crane and Moore (1986) reviewed several studies and reported a bacteria die-off rate ranging from 0.04 - 0.20 day^{-1} (base 10), which translates to the SQOLIM adjustment factor of 2.5 to 11.5. The SQOLIM adjustment factor was assigned a uniform distribution between 2.5 and 11.5, and it was used to calculate the SQOLIM values for each land use.

There is no guidance available on estimating parameters WSQOP, FSTDEC and THFST when simulating FC as a water quality constituent. The values of these parameters are generally adapted from previous studies and further calibrated. Thus, these parameters were also considered knowledge uncertain. A review of values used in previous FC TMDLs shows a range of 0.5 to 2.4 for WSQOP (Lawson, 2003). Based on these reported values, a uniform distribution between 0.5 and 2.4 was used for WSQOP for all land uses. In a review by Bowie et al. (1985), FC die-off rates (FSTDEC) ranging from 0.12 to 2.52 day^{-1} were reported for various streams. The average of the reported values was 1.1 day^{-1}. A similar FSTDEC value was used in several TMDLs developed in Virginia, so a triangular distribution was assigned to FSTDEC with a mode of 1.1 and limits of 0.12 and 2.52 day^{-1}. For this study, we assumed that any uncertainty in THFST would be masked by the uncertainty in FSTDEC. Hence, a deterministic value of 1.07 was used for THFST.

FC directly deposited in a waterbody is input into HSPF as an hourly time series, and for this study, we used the BSLC-generated time series for the Mossy Creek TMDL. The sources responsible for FC direct deposit in Mossy Creek were cattle, wildlife, straight pipes, and one permitted point source. The FC production by wildlife and humans (straight pipes) were estimated to be less than 1% and 2%, respectively, of the total direct deposit FC. The FC discharge from the point source was considered negligible. As a result, uncertainty in direct deposit FC was assumed to be primarily due to cattle. To be consistent with the other bacteria load distributions, the cattle direct deposit load distribution was assumed to be log-triangular. To obtain this distribution the cattle direct deposit time-series was multiplied by a factor that has a log-triangular distribution with a mode of 1 and limits of 0.1 and 10.

3.1.3.7 Hydrologic Calibration and Validation

The hydrologic calibration period was 1 September 1998 to 31 December 1999, and the validation period was 1 January 2000 to 30 September 2002, the same periods used for the Mossy Creek bacterial TMDL (Benham et al., 2004). To calibrate the model, two-phase MC simulation was conducted with 50 knowledge uncertain and 30 stochastic iterations resulting in a total of 1500 HSPF simulations. The flow volume was output as watershed inches day^{-1} and averaged for all simulations. The average daily flow volume time series was used to calculate several calibration sufficiency statistics as guided by the 'Expert System for HSPF' (HSPEXP) (Lumb, 1994). HSPEXP user manual provides guidance for parameter adjustment to achieve satisfactory calibration statistics. The guidance was used to increase or decrease the limits of the parameter distributions during calibration. Care was taken not to violate the maximum possible parameters limit values (if possible, typical limits were not violated) suggested in the BASINS technical note 6 (USEPA, 2000). The process was repeated until satisfactory HSPEXP statistics were obtained for the calibration period. After calibration, the two-phase MC was conducted with 300 knowledge uncertain and 40 stochastic iterations (12000 HSPF simulations) for calibration and validation periods. The number of stochastically variable parameters, INFILT for each land use, was far less than the knowledge uncertain parameters, and therefore greater number of simulations was required to sample the parameter space of knowledge uncertain parameters. The selection of 300 and 40 was considered as a good balance to sample the knowledge uncertain and stochastically variable parameters effectively while still keeping the total number of simulations manageable. The output from all 12000 simulations was used to obtain an average flow volume time series. This average flow volume time series was used to calculate HSPEXP statistics (table 3-6). The parameter distributions obtained following calibration (table 3-7) were used for the MC and TPMC simulations.

Table 3-6 Summary statistics for the hydrologic calibration and validation period

Calibration Sufficiency Statistics	Default criteria for percent error	Calculated percent errors for calibration period (%)	Calculated percent error for validation period (%)
Total Volume	±10.0	-1.1	0.3
50% Lowest flows	±10.0	-4.2	6.3
10% Highest flows	±15.0	3.4	-6.3
Storm peaks	±20.0	6.6	14.7
Seasonal volume error	±30.0	10.8	22.6
Summer storm volume error	±50.0	-9.9	12.3

Table 3-7 Parameter distribution of hydrologic parameters following model calibration

Parameter Name	Distribution
LZSN-Forest	Uniform (3,6)
LZSN-Cropland	Uniform (3,6)
LZSN-Pasture	Uniform (3,6)
LZSN-Farmstead	Uniform (3,6)
LZSN-LDR	Uniform (3,6)
LZSN-HDR	Uniform (3,6)
LZSN-Loafing Lot	Uniform (3,6)
INFILT-Forest	Triangular (0.4, 0.7, 1)
INFILT-Cropland	Triangular (0.2, 0.3, 0.4)
INFILT-Pasture	Triangular (0.4, 0.65, 0.9)
INFILT-Farmstead	Triangular (0.1, 0.18, 0.3)
INFILT-LDR	Triangular (0.1, 0.18, 0.3)
INFILT-HDRs	Triangular (0.03, 0.07, 0.1)
INFILT-Loafing Lot	Uniform (0.2, 0.4)
DEEPFR (all land uses)	Uniform (0, 0.05)
BASETP (all land uses)	Uniform (0, 0.05)
AGWETP (all land uses)	Uniform (0, 0.02)
INTFW (all land uses)	Uniform (1.8, 3.8)
IRC (all land uses)	Triangular (0.5, 0.7, 0.7)
MON INTERCEP-Forest-January*	Uniform (0.05, 0.075)
UZSN-Forest-January*	Uniform (0.1, 0.2)
LZETP-Forest-January*	Uniform (0.1, 0.2)

*The values of remaining months and land uses were obtained by multiplying a pre-determined factor to these values.

3.1.3.8 Water Quality Calibration

For water quality calibration, a TPMC simulation consisting of 300 knowledge uncertain and 40 stochastic iterations (12000 HSPF simulations) was conducted for the period of 1 October 1998 to 31 December 2001. The output from each HSPF run included daily maximum, minimum, and average FC concentration time series. These values for each day were averaged for all 12000 simulations and plotted against the observed data (figure 3.4). Since the data were observed by collecting a grab sample once a day, it cannot be expected that the simulated average FC concentration will exactly match the observed data. However, it is reasonable to assume that the observed data will fall between the maximum and minimum simulated values for a specific day. For Mossy Creek watershed, 72.2% of observed data fell between the average maximum and minimum FC concentrations. The FC observed data violated the single-sample

FC criterion 60% of the time, while the average simulated FC concentration violated the single-sample FC criterion 77.2%. The model was assumed to be sufficiently calibrated for FC simulations as more than half of the observed data was in the band of average maximum and minimum FC concentrations. As is often the case when modeling FC there were insufficient observed data to permit water quality model validation.

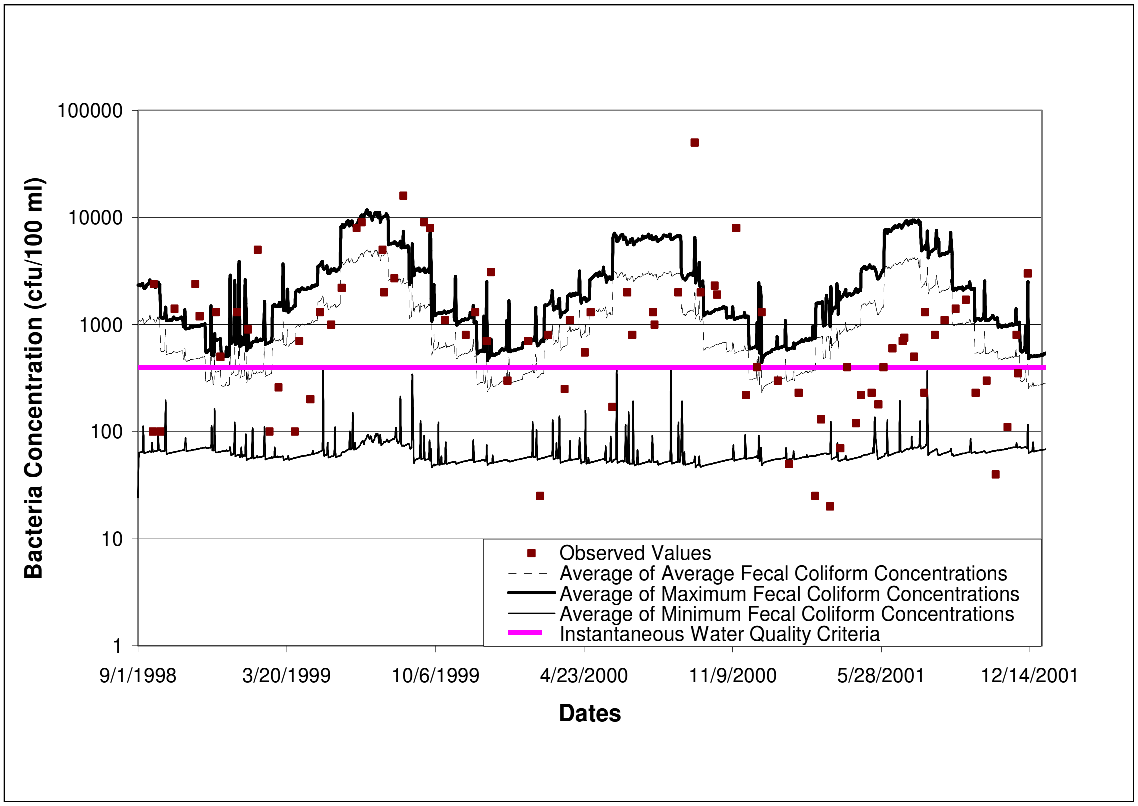


Figure 3.4 Observed and simulated fecal coliform concentrations at the water quality observation station

3.1.3.9 TMDL Pollutant Allocation Scenarios

A TMDL pollutant allocation scenario allocates the pollutant load among different sources and hence suggests the amount of reduction in pollutant loading from each source needed to meet the applicable water quality criteria. To simulate the Mossy Creek TMDL allocation scenarios in HSPF, a reduction factor was applied to the pollutant load from each source. For example, to simulate 94% reduction in cattle direct deposit of FC load, the pollutant load from cattle direct deposit was multiplied by 0.06. The Mossy Creek TMDL listed several pollutant allocation scenarios, with the two preferred allocation scenarios, shown in table 3-8.

Table 3-8 TMDL pollutant allocation scenarios for Mossy Creek TMDL resulting in no violations (Benham et al., 2004)

	Required source-specific fecal coliform load reductions (%)						
TMDL Allocation Scenario	Cattle Direct Deposit	Cropland	Pasture	Loafing Lot	Wildlife Direct Deposit	Straight Pipes	All residential pervious land segments
S1	99	90	98	100	30	100	95
S2	94	95	98	100	0	100	95

Used under fair use guidelines, 2011

Both TMDL allocation scenarios require 100% reduction in FC loading from illegal straight pipes discharging waste directly from homes. The major difference between the scenarios is the reduction in cattle direct deposit, wildlife direct deposit and loadings from cropland. Due to the low production of FC by wildlife, uncertainty in wildlife direct deposit was not considered in this study. A period of three and a half years (1218 days) that represents a range of hydrological events in Mossy Creek was selected to simulate the in-stream FC concentration under the two allocation scenarios.

3.2 Results and Discussion

3.2.1 Single-phase Monte Carlo Simulation

A daily average in-stream FC concentration time series was calculated using the 12000 daily average FC concentrations produced by the all Mossy Creek model iterations. Quantile time series (2.5% and 97.5%; 10% and 90%) were also calculated using the same model output. The average and quantile time series were plotted for the two TMDL allocation scenarios (figure 3.5). Whenever the predicted average FC concentration was greater than the instantaneous FC criterion of 400 cfu/100ml, it was considered a violation incident for the day. The percent of violations for each time series was calculated by dividing the number of violations by the number of days in the prediction period (1218 days) (table 3-9).

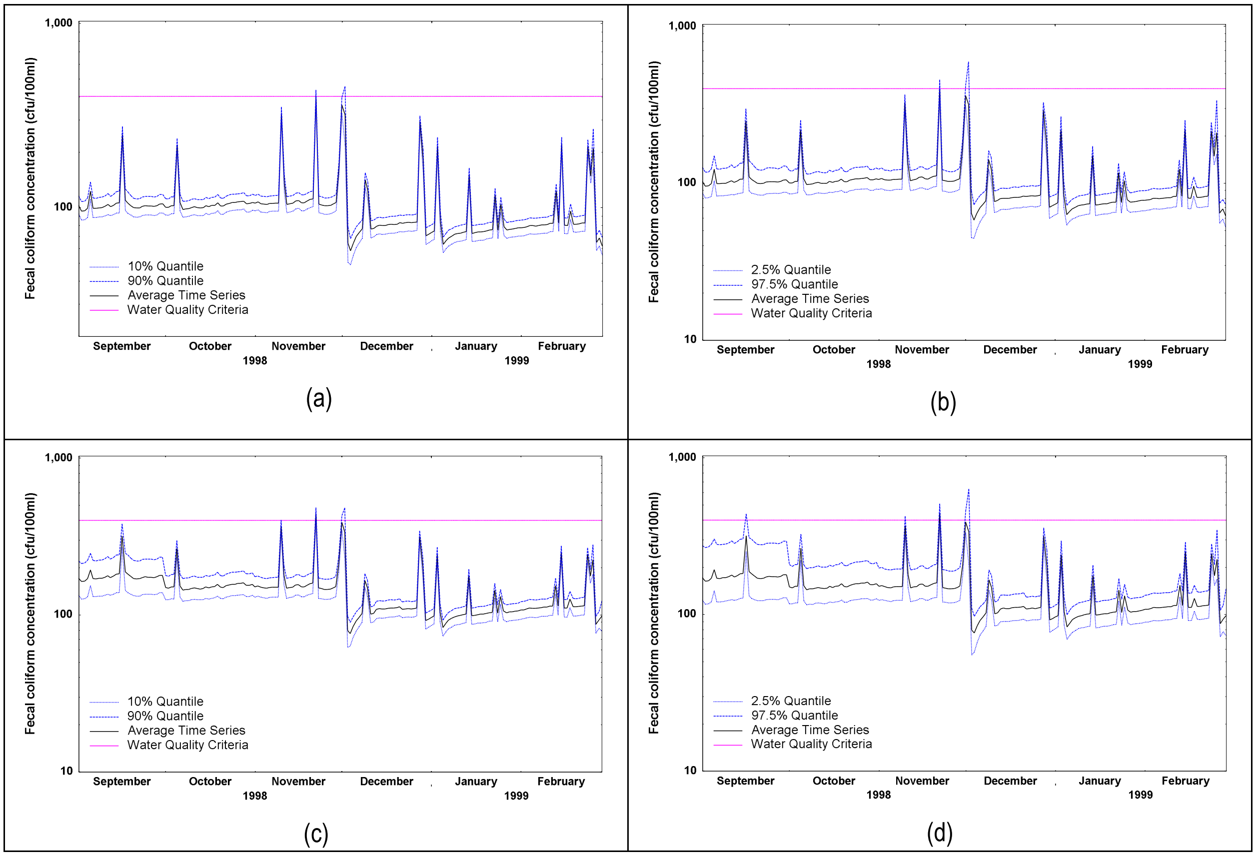


Figure 3.5 For TMDL allocation scenario S1, (a) 80% probability interval, and (b) 95% probability interval; for TMDL allocation scenario S2, (c) 80% probability interval, and (d) 95% probability interval. Representative plots show results for first six months of the simulation period.

Table 3-9 Percent of violations of single-sample fecal coliform criteria for the two TMDL allocation scenarios during the prediction period.

TMDL Allocation Scenario	Single-phase Monte Carlo simulation			Two-phase Monte Carlo simulation		
	Average of daily average time series	80% Probability interval	95% Probability interval	Average of daily average time series	80% Probability interval	95% Probability interval
S1	0.7	(0.2, 1.2)†	(0.1, 1.5)	0.7	(0.2, 1.2)	(0.1, 1.5)
S2	0.8	(0.3, 3.5)	(0.2, 14.6)	0.8	(0.4, 2.1)	(0.2, 7.8)

† Numbers in parentheses show the percent of violation incidences over a period of 1218 days by the respective time series for the probability interval.

Figure 3-5 illustrates that the 95% probability intervals -frames (b) and (d) are wider than the 80% probability intervals -frames (a) and (c) for both allocation scenarios; stated another way, as one seeks greater confidence in the predicted results, uncertainty increases. Similarly, the number of violations is greater at the 95% probability interval compared to 80% (table 3-9). The percent violation by the average time series is similar for the two TMDL allocation scenarios. However, the percent violations are different for S1 and S2 for the 80 and 95% probability intervals; the S2 scenario exhibits a greater uncertainty compared to S1, especially on the upper bound. The upper bounds of violations in the 95% probability interval increased as much as ten times for S2 compared to S1. Direct deposit from cattle is reduced 99% in S1 compared to only 94% in S2 (table 3-8), whereas loading from cropland is reduced in S1 by 90% compared to a 95% reduction in S2. These results illustrate that the FC direct deposit is a greater source of uncertainty than cropland FC loadings as the input uncertainty in the two sources was similar (log-triangle distribution, spread over two orders of magnitude). These insights into the behavior of different pollutant sources and the acceptable level of confidence in water quality prediction can help stakeholders and decision makers prioritize one allocation scenario over another.

3.2.2 Two-phase Monte Carlo Simulation

The TPMC simulation was conducted to evaluate the effects of knowledge uncertainty and stochastic variability separately on the FC water quality criterion violations and estimation the overall uncertainty. Using the FC output from all the simulations, the cumulative probability of number of violations for each knowledge uncertain iteration was calculated. For example, table 3-10 shows the cumulative probability for the number of violations for a selected knowledge uncertain simulation (40 HSPF iterations) for the S1 allocation scenario. The cumulative probabilities of the violations were plotted as a cumulative distribution function (CDF) for each knowledge uncertain HSPF iteration. Each CDF shows the probability of the number of violation incidences. The complete TPMC simulation included 300 knowledge-uncertain iterations yielding 300 CDFs. The maximum number of violation incidences for any HSPF run was 38 for Scenario S1 (figure 3.6) and 283 for Scenario S2 (figure 3.7).

Table 3-10 Example of cumulative probability for numbers of single-sample fecal coliform criterion violations for a given knowledge uncertain simulation

Number of single-sample FC criterion violations incidences	Cumulative Probability
0	0.13
1	0.28
2	0.60
5	1.00
10	1.00

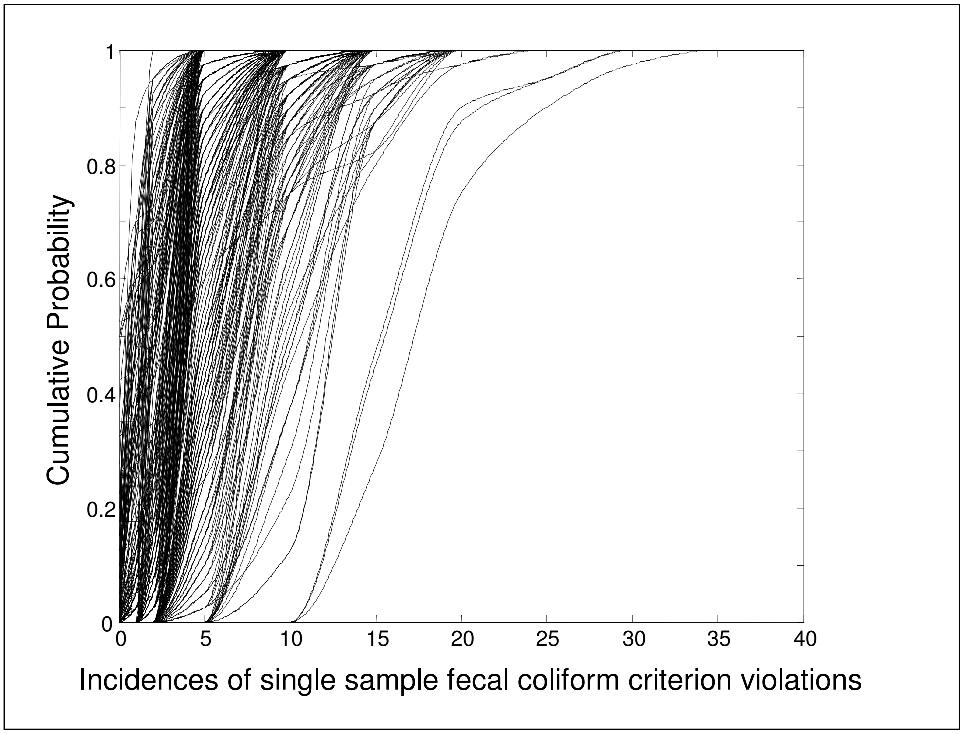


Figure 3.6 Distribution of cumulative distribution functions (CDF) resulting due to knowledge uncertainty for TMDL allocation scenario S1. Each individual CDF is a result of stochastic variability.

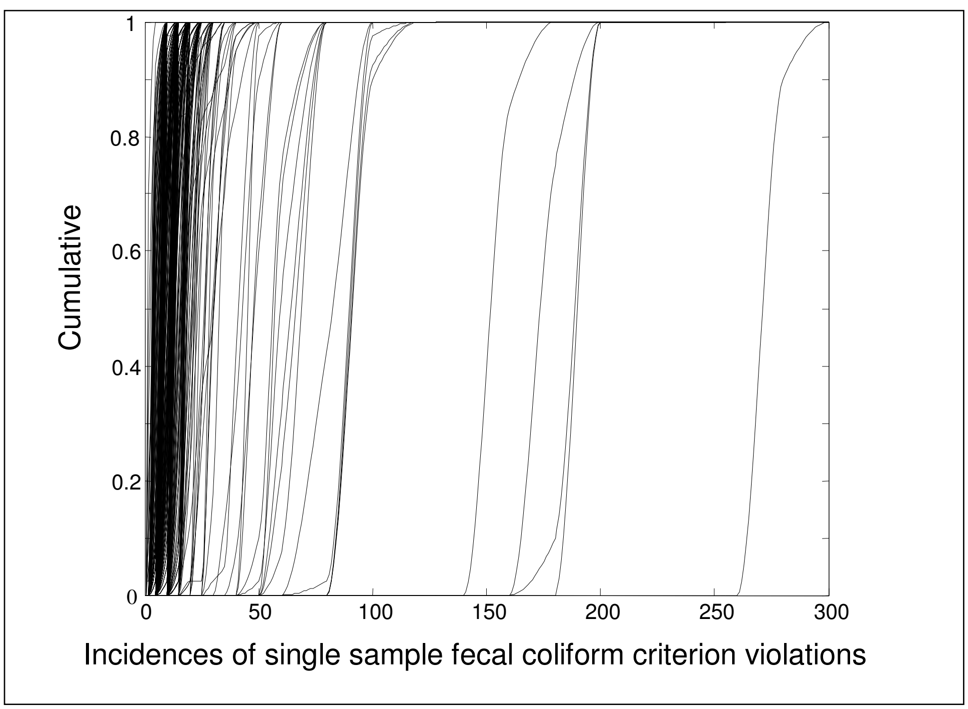


Figure 3.7 Distribution of cumulative distribution functions (CDF) resulting due to knowledge uncertainty for TMDL allocation scenario S2. Each individual CDF is a result of stochastic variability.

The collection of CDF curves illustrates the effect of knowledge uncertainty and stochastic variability on the number of violations. A vertical CDF would illustrate no stochastic variability effect while overlapping CDFs illustrate no knowledge uncertainty effect. Visual observation of curves suggests that the CDF curves of water quality criterion violations are nearly vertical for both TMDL allocation scenarios illustrating very little stochastic variability. It is evident from the spread of CDFs that compared to the S1 allocation scenario (figure 3.6), the S2 allocation scenario exhibited a greater effect of knowledge uncertainty (figure 3.7). The median of knowledge uncertain simulations was used to plot a CDF for each allocation scenario (figure 3.8) and conduct a Kolmogorov-Smirnov (KS) test. The KS test resulted in a p-value of 0.0 suggesting that the two datasets are significantly different.

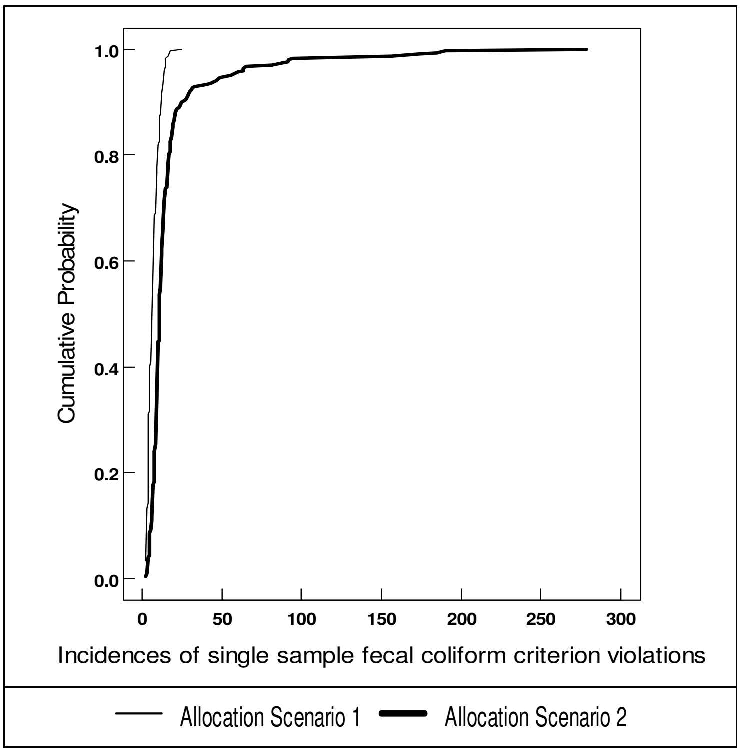


Figure 3.8 Comparison of TMDL allocation Scenarios by plotting the CDF of median of family of CDFs obtained from Two-phase Monte Carlo Simulation

An analysis of variance (ANOVA) of the two scenarios showed that there is a significant effect of knowledge uncertainty on number of violations for both allocation scenarios, but there is no significant effect of stochastic variability. Since there is a significant effect of knowledge uncertainty on model output, modelers are advised to focus future efforts in collecting more information about the knowledge uncertain parameters to reduce uncertainty in model output. In the present research, this result is probably an artifact of assigning only one parameter as stochastically variable. A different categorization of parameters could affect these results.

To illustrate the uncertainty in predicted in-stream FC concentrations at the watershed outlet, different quantiles and average of the average daily FC concentration were plotted (figure 3.9). Percent violations of single-sample instantaneous FC criterion were calculated as explained for single-phase MC earlier (table 3-9). As illustrated in table 3-9, the average time series percent violations are similar for the two TMDL allocation scenarios. However, the percent violations are different for S1 and S2 for the 80 and 95% probability intervals; the S2 scenario exhibits a greater uncertainty compared to S1, especially with respect to the upper bound. The average time series percent violations are the same for the simple MC and the TPMC, for both allocation scenarios. As both techniques used the same input parameter distributions, agreement among the average time series was expected. The range of percent violations for the S1 scenario is similar for simple MC and TPMC, but the range is smaller for S2 for TPMC. In other words, the estimated uncertainty is lower in TPMC compared to simple MC.

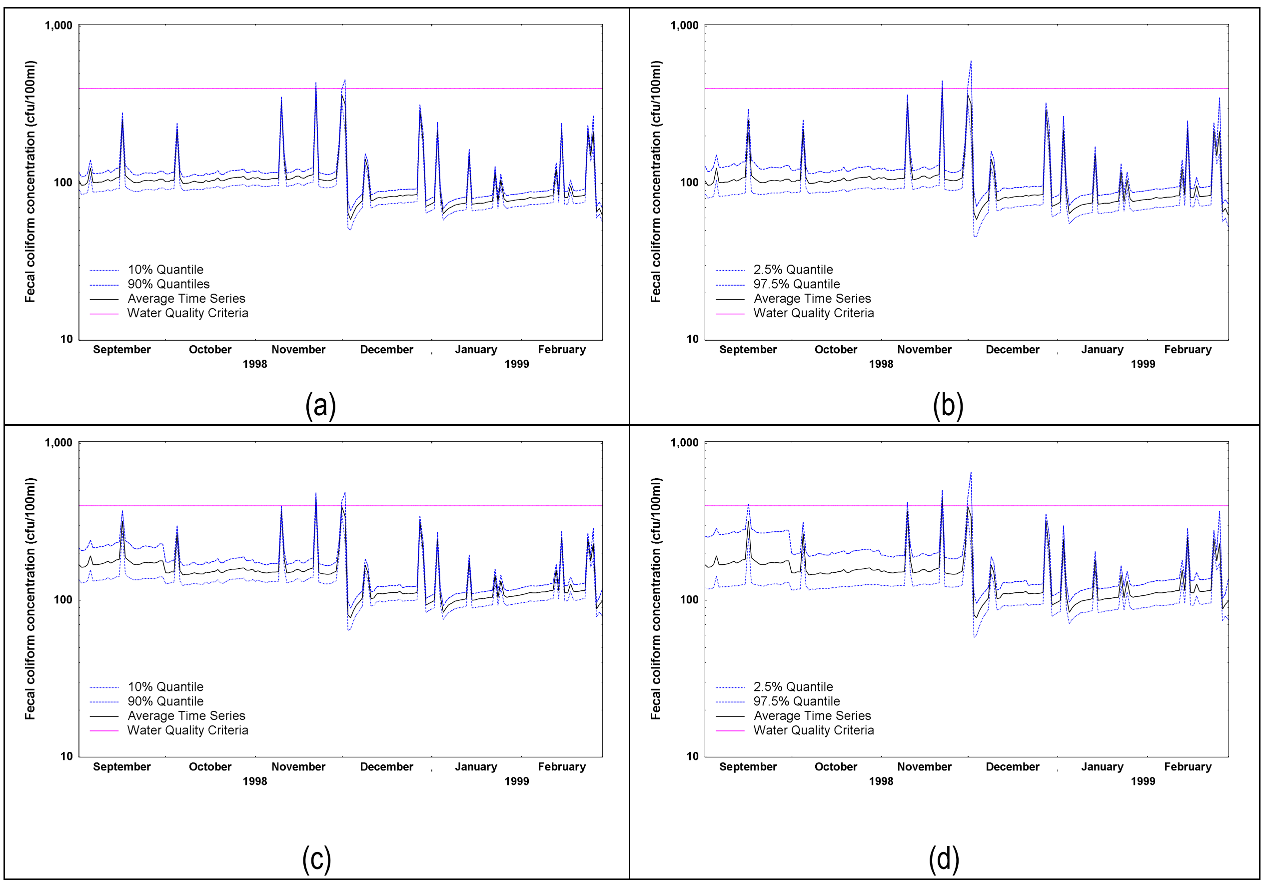


Figure 3.9 For TMDL allocation scenario S1, (a) 80% probability interval, and (b) 95% probability interval; for TMDL allocation scenario S2, (c) 80% probability interval, and (d) 95% probability interval. Representative plots show results for first six months of simulation period.

3.3 Summary and Conclusion

We presented two techniques for estimating uncertainty in FC criterion violations when using a HSPF based water quality model – a simple single phase MC approach and a TPMC approach. This study used the data collected and modeling files developed for the Mossy Creek, VA bacterial impairment TMDL (Benham et al., 2004). The two techniques reported similar percentage of water quality violations for the two alternative TMDL pollutant allocation scenarios. The two techniques also reported similar uncertainty estimates for the S1 allocation scenario that allowed only 1% of FC loading from cattle direct deposit but 10% of loading from cropland runoff. However, TPMC reported lower uncertainty for the S2 allocation scenario that allowed 6% of FC loading from cattle direct deposit and 5% loading from cropland runoff. The primary reason for TPMC to report lower uncertainty was that there were likely insufficient iterations to simulate the complete parameter space. In a MC simulation, the modeler's aim is to conduct sufficient iterations to closely match the simulated distribution and actual distribution of the parameters. In the TPMC analysis, the parameters were varied at most 300 times, as compared to 12000 times in single-phase MC. The TPMC may yield similar results to single-phase MC if the iterations of knowledge uncertain and stochastic variability were increased.

The TPMC results illustrated that there was a significantly greater effect of knowledge uncertainty than stochastic variability on the model output. This information can be used by the modeler in focusing resources on collecting more information about the knowledge uncertain parameters. This result is likely an artifact of assigning only one parameter as stochastically variable. This categorization of a parameter as knowledge uncertain or stochastically variable can be highly subjective. When using a lumped parameter watershed-scale water quality model, a single parameter is assigned to a large area for a simulation spanning multiple years, and it is difficult to define a parameter as fitting in one or the other category. Although some parameters may be defined as both, knowledge uncertain and stochastically variable, it is difficult to estimate a probability distribution for those parameters. Although, two-phase MC can be used as an effective tool to obtain more information about parameter behavior, a single-phase MC uncertainty analysis approach with a sufficiently large number of iterations may be a better choice for estimating uncertainty when compared with a more complex, more subjective TPMC approach.

Stakeholders and decision makers can base their water quality management decisions upon the results of assessments like that presented here. Ability to assess the uncertainty in pollutant allocation scenarios allows the stakeholders to make a more informed decision about the selection of a scenario. Generally, the allocation scenario that meets water quality standard with greater confidence is more expensive to implement, as it would need extensive pollution control measures. Such a scenario might be preferable where regulations are strict or the ecosystem is fragile, i.e., the risk of a water quality violation has potentially greater consequences. A cheaper allocation scenario with greater uncertainty might be preferred where the ecosystem is not fragile, there are conflicts of interests, or the funding for watershed management program is limited. Estimating uncertainty could also help in prioritizing implementation of control measures. Decision makers may choose to control the pollutant sources that are responsible for greater uncertainty ahead of others. For example, in this research, although the input uncertainty in IB loadings from cropland and IB direct deposit in waterbodies was similar, direct deposit was responsible for greater uncertainty than runoff from croplands.

The research reported here demonstrated that the single-phase Monte Carlo simulation technique is very useful in estimating the uncertainty in model output, and can be successfully used with complex modeling software like HSPF. In estimating total uncertainty, this technique is more useful than two-phase MC because of its computational efficiency, and unambiguous parameter categorization. However, the simple MC and two-phase MC techniques do not provide any assistance in calibrating the model. The utility of these techniques can be vastly improved if it is used in conjunction with other procedures that can be used to estimate model parameter distributions (e.g. Generalized Likelihood Uncertainty Estimation (GLUE), Bayesian Monte Carlo, and Markov Chain Monte Carlo (MCMC) Simulation). GLUE and MCMC are discussed in the following chapters.

References:

Al-Abed, N.A., and H.R. Whiteley. 2002. Calibration of the Hydrologic Simulation Program Fortran (HSPF) model using automatic calibration and geographical information systems. *Hydrological Processes.* 16(16). 3169-3188.

American Society of Agricultural Engineers Standards. 2003. D384.1: Manure production and characteristics. St. Joseph, Mich.: ASAE.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Benham, B. L., K. M. Brannan, G. Yagow, R. W. Zeckoski, T. A. Dillaha, S. Mostaghimi, and J. W. Wynn. 2005. Development of Bacteria and Benthic Total Maximum Daily Loads: A Case Study, Linville Creek, Virginia. Journal of Environmental Quality. 34: 1860-1872.

Benham, B., K. Branna, K. Christophel, T. Dillaha, L. Henry, S. Mostaghimi, R. Wagner, J. Wynn, G. Yagow, and R. Zeckoski. 2004. Total maximum daily load development for Mossy Creek and Long Glade Run: Bacteria and general standard (Benthic) impairments. Richmond, Va.: Virginia Department of Environmental Quality. Available at http://www.deq.state.va.us/tmdl/homepage.html. Assessed Jan 25, 2005.

Benham, B.L., K. Brannan, T. Dillaha, S. Mostaghimi, G. Yagow. 2002. TMDLs (Total Maximum Daily Loads) – Terms and Definitions. *Virginia Cooperative Extension.* Pub No. 442-550.

Beven, K.. and A. Binley 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr. T.H. Jobes, and A.S. Donigian, Jr. 2005. *HSPF Version 12.2 User's Manual.* AQUA TERRA Consultants. Mountain View, CA.

Bowie, G.L., W.B. Mills, D.B. Porcella, C.L. Campbell, J.R. Pagenkopf, G.L. Rupp, K.M. Johnson, P.W.H. Chan, and S.A. Ghereini. 1985. *Rates, Constants, and Kinetic Formulations in Surface Water Quality Modeling (Second Edition)* Tetra Tech Inc. Lafeyette, CA.

Crane, S. R., and J.A. Moore. 1986. Modeling enteric bacterial die-off: a review. *Water, Air, and Soil Pollution. 27(3-4): 411-439.*

Donigian, A.S., and J.T. Love. 2007. The Housatonic River Watershed Model: Model Application and Uncertainty Analyses. 7th International IWA Symposium on System Analysis and Integrated Assessment in Water Management, May 7-9, 2007. Washington, DC. WATERMATEX Proceedings on CD-ROM.

Gardner, R.H. and R.V.O'Neil. 1983. Parameter Uncertainty and model predictions: A review of Monte Carlo results. In *Uncertainty and Forecasting of Water Quality, eds.* M.B. Beck and G.Van Straten, 345-257. Berlin, Grmany: Springer-Verlag.

Geldrich, E.E. 1978. Bacterial productions and indicator concepts in feces, sewage, stormwater, and solid wastes. In *Indicators of Viruses in Water and Food*, ed. G. Berg. Ann Arbor, Mich.: Ann Arbor Science.

Helton, J.C. 1994. Treatment of uncertainty in performance assessment for complex systems. *Risk Analysis*. 14(4): 483-511.

Hession, W.C., D.E. Storm, and C.T. Haan. 1996. Two-phase uncertainty analysis: an example using universal soil loss equation. *Trans. ASAE.* 39(4): 1309-1319.

Lawson, L.G. 2003. HSPF model calibration and verification for bacteria TMDLs. Richmond, Va.: Water Division, VADDEQS.

Lumb, A.M., R.B. McCammon, and J.L. Kittle, Jr. 1994. Users manual for an expert systems (HSPEXP) for calibration of the Hydrologic Simulation Program-Fortran. Reston, Va: USGS, Water Services Investigations Report.

MacIntosh, D.L., G.W. Suter., and F.O. Hoffman. 1994. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. *Risk Analysis*. 14(4): 405-419.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2005. Soil and Water Assessment Tool Theoretical Documentation. Version 2005. Grassland Soil and Water Research Laboratory, Agriculture Research Service. Temple, TX.

Paul, S., P.K. Haan, M.D. Matlock, S. Mukhtar, and S.D. Pillai. 2004. Analysis of the HSPF water quality parameter uncertainty in predicting peak in-stream fecal coliform concentrations. Trans. ASAE. 47(1): 69-78.

Summers, J.K., H.T. Wilson, and J. Kou. 1993. A method for quantifying the prediction uncertainties associated with water quality models. *Ecological Modeling.* 65: 161-176

Suter, G.W. II, L.W. Barthouse and R.V. O'Neill. 1987. Treatment of risk in environmental impact assessment. *Environmental Mgmt.* 11(3): 295-303.

USEPA 2000. BASINS Technical Note 6. United States Environmental Protection Agency. Washington D.C. Office of Water.

USEPA. 2001. The National Cost of the Total Maximum Daily Load Program (Draft Report). Office of Water, United States Environmental Protection Agency. Washington D.C.

USEPA. 2006. *National Section 303(d) List Fact Sheet.* United States Environmental Protection Agency. Available at http://oaspub.epa.gov/waters/national_rept.control. Accessed Mar 20, 2006.

USEPA. 2009. *National Section 303(d) list fact sheet.* United States Environmental Protection Agency. Available from: URL: http://oaspub.epa.gov/waters/national_rept.control

VADCR. 2003. Bacteria TMDLs for Abrams Creek and Upper and Lower Opequon Creek Located in Frederick and Clarke County, Virgnia. Virginia Department of Environmenatal Quality, and Virginia Department of Conservation and Recreation. Richmond, VA, USA.

Yagow, G. 2001. Fecal coliform TMDL mountain run watershed, Culpeper County, Virginia. Biological Systems Engineering, Virginia Tech, Blacksburg, Va. Available at: http://www.deq.virginia.gov/tmdl. Accessed September 10, 2006.

Zeckoski, R. W., B.L. Benham, S. B. Shah, M.L. Wolfe, K.M. Brannan, M. Al-Smadi, T.A. Dillaha, S. Mostaghimi, C.D. Heatwole. 2005. BSLC: A tool for bacteria source characterization for watershed management. *Applied Engg. In Agriculture.*21(5): 879-889

Chapter 4. Evaluation of the applicability of using log-transformed in-stream indicator bacteria concentrations to calculate a likelihood function for estimating uncertainty using the Generalized Likelihood Uncertainty Estimation (GLUE) technique with an HSPF model.

Abstract. *Generalized Likelihood Uncertainty Estimation (GLUE) was used to estimate the posterior distributions of input parameters for a Hydrological Simulation Program – FORTRAN (HSPF) model used to develop the Mossy Creek bacterial Total Maximum Daily Load (TMDL) in Virginia. The posterior parameter distributions were used to estimate uncertainty in the violations of in-stream fecal coliform (FC) concentration criteria for two pollutant allocation scenarios presented in the TMDL. The TMDL allocation scenarios differed in the reductions specified from cattle directly depositing FC in the stream and cropland runoff loadings. The instantaneous FC criterion was violated less than 2% of the time (on a daily basis) over the prediction period for both the allocation scenarios. The results illustrated that direct deposit is a greater source of uncertainty in in-stream FC criteria violations compared to cropland runoff loads. As the simulated and observed FC concentrations can vary by orders of magnitude, the impact of a log transformation of the FC concentrations on uncertainty estimates was assessed by evaluating the GLUE likelihood function using both log-transformed and non log-transformed FC concentration data. When the FC concentration was log-transformed, the range of percent violations reduced from 0.2 – 22.7 to 0.1 – 3.1, at the 95% probability interval for one of the scenarios. The results underscore the importance of log-transforming the FC concentration, and how the likelihood function selection can affect the uncertainty estimates with GLUE application. Decision makers can use the results of an assessment like this in choosing their confidence level in achieving a water quality standard, selecting among the TMDL allocation scenarios, and prioritizing implementation efforts.*

Keywords. GLUE, Generalized Likelihood Uncertainty Estimation, HSPF, fecal coliform, uncertainty analysis.

Introduction

A Total Maximum Daily Load (TMDL) quantifies the amount of a given pollutant a waterbody can receive and still meet applicable water quality standards (Benham et al., 2002). Water quality modeling is often used to develop the TMDLs. However, while modeling a natural system, uncertainty is always present and is a result of limited knowledge of the system being modeled or stochastic variability within that system (Beck, 1987; Suter et al., 1987). Typically, no formal calculation is performed to estimate the uncertainty in predicted water quality; instead a margin of safety (MOS) is included in the TMDL calculation to account for the inherent uncertainty. In 2001, the U.S. Environmental Protection Agency (USEPA) estimated the annual average cost of developing TMDLs to be $63-69 million per year for the next fifteen years and the cost of implementing TMDLs to be between $1 and 3.4 billion per year for the next decade (USEPA, 2001). Pathogen impairments, typically represented by a surrogate indicator bacteria (IB), are the second most widespread cause of water quality impairments (USEPA, 2006) and represent a significant share of the estimated expense.

Most water quality modeling software currently used when developing TMDLs includes modules that are process-based, empirical or a mixture. These software do not typically include detailed uncertainty analysis capabilities. The Hydrological Simulation Program–FORTRAN (HSPF) is a continuous simulation model that simulates various hydrological and water quality processes (Bicknell, 2005), and has been widely used to develop IB impairment TMDLs (e.g., Benham et al., 2004; VADCR, 2003; Benham et al., 2003; Yagow, 2001). HSPF is a lumped parameter, watershed scale model, and produces a deterministic time-series of hydrology and water quality. Despite its widespread usage, there have been few applications of HSPF that included a detailed uncertainty analysis. A simple single phase Monte Carlo and a two-phase Monte Carlo (TPMC) approach to estimate uncertainty in predicted fecal coliform (FC) concentration was illustrated in Chapter 3 of this dissertation using an application of HSPF.

Another Monte Carlo based approach to estimate uncertainty in hydrologic modeling is the “Generalized Likelihood Uncertainty Estimation” (GLUE) approach, as proposed by Beven and Binley (1992). GLUE is based on the premise that there is not one set of model parameters that represents a “true” parameter set for a system. Instead, in the GLUE approach, several parameter sets are generated using MC simulation, and an assessment is made as to whether an input parameter set has a likelihood of being an acceptable simulator of the system.

The MC simulation is performed by generating different sets of model input parameters sampled from what are called “prior distributions.” The prior distributions are based on the knowledge about the system being modeled, the modeler’s experience, and relevant literature. The performance of each parameter set sampled from the prior distributions is assessed by comparing model output with observed data. This comparison is done using likelihood functions.

Likelihood function formulation is central to the GLUE approach. Likelihood can be one of many "goodness of fit" measures that are used to compare observed and simulated response variables (Stow et al., 2007). This likelihood definition differs from the statistical definition of "likelihood function" and is a controversial aspect of GLUE (Stedinger et al., 2008). Based on the likelihood assessment, acceptable parameter values are used to compute what are called "posterior" input parameter distributions of using the Bayesian equation. The posterior distribution parameter sets can then be sampled for subsequent Monte Carlo simulations. Output from those simulations can be used to validate the model or to estimate the predictive uncertainty.

GLUE has been suggested as a viable approach to estimate uncertainty when developing TMDLs (Stow et al., 2007). The GLUE approach has been used widely to conduct uncertainty analysis for a range of hydrologic models (Beven and Binley, 1992; Freer et al., 1996; Balin, 2004). However, there have been few attempts to use GLUE to quantify uncertainty of watershed-scale water quality modeling applications (Setegn el al., 2009; Zheng and Keller, 2007; Benaman and Shoemaker, 2002).

In this research, GLUE was used to estimate uncertainty in the violations of in-stream fecal coliform (FC) concentration criteria for two pollutant allocation scenarios presented in the Mossy Creek bacterial TMDL that was developed using an HSPF model (Benham et al., 2004). The observed and simulated FC data were log-transformed to assess the impacts of this transformation on uncertainty estimates. The null hypothesis was that log-transformation of the FC data in GLUE application does not affect the uncertainty estimates of the FC criterion violations. This is important as likelihood formulation and assessment is central to any GLUE application, and log-transformation will affect the likelihood calculation and therefore uncertainty estimation. The objective of this research was to evaluate the effect of using log-transformed FC concentrations to calculate a likelihood function for estimating uncertainty using the GLUE technique with HSPF.

4.1 Materials and Methods

4.1.1 Study Area

Mossy Creek, located in Rockingham and Augusta counties in Virginia (figure 4.1), was selected for this research. Mossy Creek was listed as impaired in 1996 due to violations of the instantaneous FC criterion and a TMDL was developed for Mossy Creek by the Department of Biological Systems Engineering (BSE) at Virginia Tech (Benham et al., 2004). The Mossy Creek watershed (4076 ha) is characterized as a rolling valley with Blue Ridge Mountains to the east and the Appalachian Mountains to the west. The predominant land uses in Mossy Creek watershed are forest, pasture, and croplands. The primary sources of FC identified in the Mossy Creek TMDL were direct deposition of feces in the stream by cattle (cattle loitering and defecating in the stream), and runoff from pastures where grazing animals defecate.

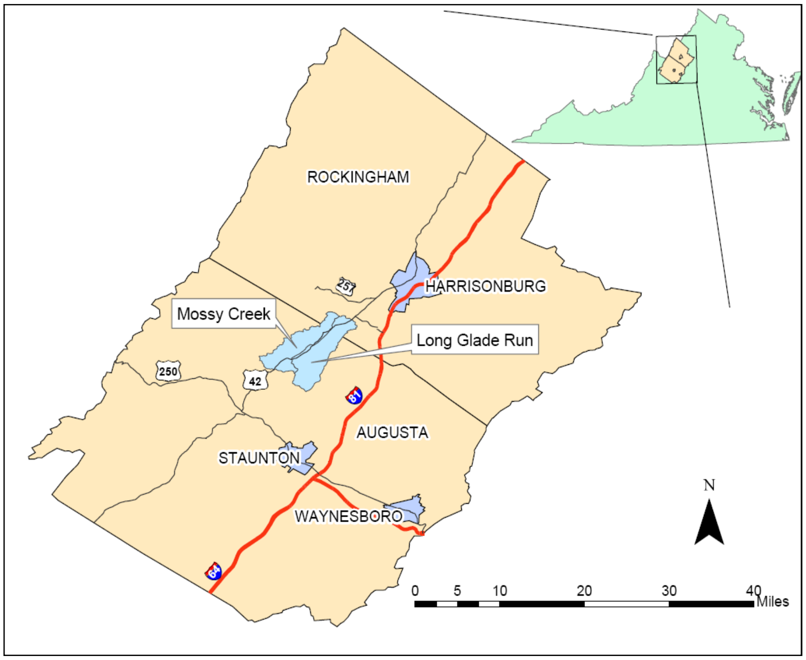


Figure 4.1 Mossy Creek Watershed (Benham et al., 2004)
Used under fair use guidelines, 2011

Mossy Creek was monitored monthly by the Virginia Department of Environmental Quality (DEQ) between July 1992 and March 2003 for FC concentration and other selected water quality constituents at the station ID 1BMSS001.35 (QMA in figure 4.2) located near the outlet of the Mossy Creek watershed. BSE monitored Mossy Creek semi-monthly between February 1998 and December 2001 for selected water quality constituents including FC concentration near the DEQ site. Daily flow data were also collected from May 1998 to December 2002 at the same site.

4.1.2 Mossy Creek Watershed Model

HSPF was used in developing the Mossy Creek bacterial impairment TMDL (Benham et. al., 2004). Mossy Creek was divided into eight subwatersheds for modeling and land use identification purposes (figure 4.2, table 4-1). Other data required by the model included rainfall, FC loading from cattle and wildlife, inflows from springs, solar radiation, and temperature as time series. TMDL modeling data development and acquisitions are described in the Mossy Creek TMDL (Benham et. al., 2004).

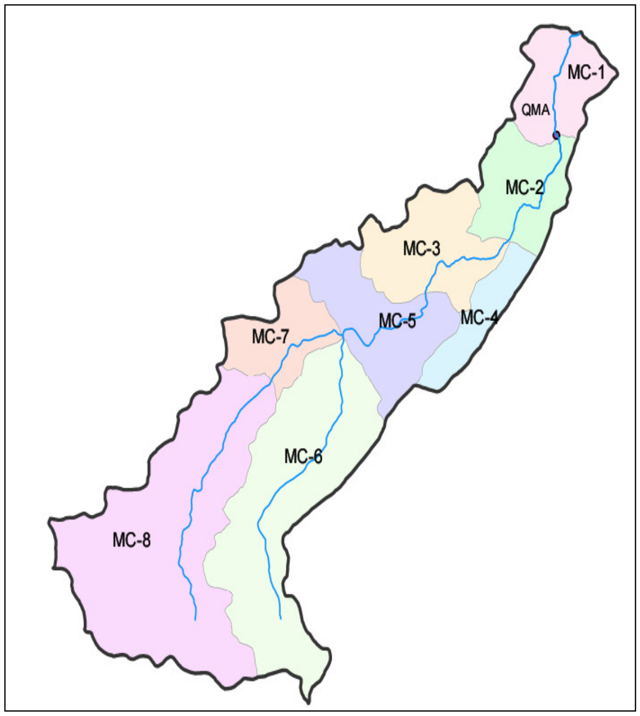


Figure 4.2 Mossy Creek watershed and its subwatersheds (Benham et al., 2004)
Used under fair use guidelines, 2011

Table 4-1 Land use distribution of Mossy Creek watershed (Benham et al., 2004)

Land use	Area (ha)	Percent of total area
Forest	1025.1	25.15
Cropland	556.0	13.64
Pasture	2347.6	57.59
Farmstead	55.0	1.35
Low Density Residential	87.0	2.13
High Density Residential	3.6	0.09
Loafing Lot	1.6	0.04

Used under fair use guidelines, 2011

Using HSPF to simulate FC requires information about several hydrologic and water quality parameters. The GLUE procedure requires that the input distribution of these parameters or "prior" distributions be provided by the modeler. In many previous studies, modelers have often assigned model input parameters uniform distributions (Beven and Freer, 2001). This is typically done to avoid modeler's bias. For the research reported here, we assigned both uniform and triangular input parameter prior distributions based on values found in the literature, expert opinion and GIS data available for Mossy Creek. Details of how the HSPF input parameter distributions were developed are described in Chapter 3 of this dissertation. The input parameter distributions are shown in table 4-2 through table 4-4.

Table 4-2 Distribution of hydrology parameters that apply to all land uses

Parameter	Parameter Description	Type of Distribution
LZSN (inches)	Lower zone nominal soil moisture storage	Uniform (3,8)†
AGWRC	Groundwater recession rate	Uniform (0.92, 0.99)
DEEPFR	The fraction of infiltrating water lost to deep aquifers	Uniform (0.0, 0.2)
BASETP	Evapotranspiration by riparian vegetation as active groundwater enters streambed	Uniform (0.0, 0.05)
AGWETP	Fraction of model segment that is subject to direct evaporation from groundwater storage	Uniform (0.0, 0.05)
IRC	Interflow Recession Coefficient	Triangular (0.5, 0.7, 0.8)‡
INTFW	Coefficient that determines the amount of water which enters the ground from surface detention and becomes interflow	Uniform (1.0, 3.0)

† Numbers in parentheses show lower and upper limit of the uniform distribution, respectively. ‡Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively.

Table 4-3 Distribution of hydrologic parameters that vary according to land use and time of year (for the month of January).

Parameter	Land use	Distribution
INFILT (in/hr) Index to mean infiltration rate.	Forest	Triangular (0.05, 0.1, 1)†
	Cropland	Triangular (0.03, 0.17, 0.24)
	Pasture	Triangular (0.04, 0.09, 0.9)
	Low Density Residential	Triangular (0.03, 0.17, 0.26)
	High Density Residential	Triangular (0.01, 0.01, 0.1)
	Farmstead	Triangular (0.03, 0.15, 0.23)
	Loafing Lot	Uniform (0.15, 0.23)‡
UZSN (inches) Nominal upper zone soil moisture storage	Forest	Uniform (0.2, 0.3)†
	Cropland	Uniform (0.06, 0.1)
	Pasture	Uniform (0.06, 0.1)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.06, 0.1)
CEPSC (inches) Interception Storage Capacity	Forest	Uniform (0.05, 0.075)
	Cropland	Uniform (0.05, 0.075)
	Pasture	Uniform (0.05, 0.075)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.05, 0.075)
LZETP Index to lower zone evapotranspiration	Forest	Uniform (0.1, 0.2)
	Cropland	Uniform (0.1, 0.2)
	Pasture	Uniform (0.1, 0.2)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.1, 0.2)

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

Table 4-4 Summary of water quality parameters which have been reported as sensitive and are typically calibrated when using HSPF.

Parameter	Land use	Type of Distribution
ACQOP-PERLND (cfu day^{-1}) (Accumulation of fecal coliform on pervious land per day)	Pasture Loafing Lot Cropland (January)	Log-triangle (1×10^9, 1×10^{10}, 1×10^{11})† ‡ Log-triangle (1.12×10^{11}, 1.12×10^{12}, 1.12×10^{13}) Log-triangle (2×10^6, 2×10^7, 2×10^8)
SQOLIM adjustment Factor (Factor which is multiplied to ACCUM values to obtain SQOLIM	All	Uniform (2.5, 11.5)*
SQOLIM-PERLND (Maximum accumulation of FC on pervious land)	All	ACQOP-PERLND (for each land use) x SQOLIM adjustment Factor
WSQOP-PERLND (Rate of surface runoff that will remove 90% of stored bacteria from pervious land surface)	All	Uniform (0.5, 2.4)
FSTDEC (day^{-1}) (First order decay rate of bacteria)	All	Triangular (0.12, 1.1, 2.52)

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡Log-triangle distribution implies that the logarithm of lower limit, mode and higher limit follows a triangular distribution. *Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

4.1.3 Generalized Likelihood Uncertainty Estimation (GLUE)

Likelihood function formulation is an important step in conducting GLUE. For this application, the likelihood function was calculated using the variance of the residuals (equation 4-1). For the hydrologic calibration, the value was calculated using observed and simulated daily flow, and for the water quality calibration, the likelihood value was calculated using both non-transformed and log-transformed simulated daily average FC concentration and observed instantaneous FC concentrations. For water quality, only the days for which observed data were available were used to calculate the likelihood function.

$$L_e = (\sigma_e^2)^{-N} \tag{4-1}$$

where,

$$\sigma^2{}_e = 1/n\left(\sum_{i=1}^{n}\left(Y_i - Q_i\right)^2\right),$$

L_e = likelihood value,

σ_e^2 = variance of the residuals or mean square error,

n = number of data points,

Y_i = observed data point,

Q_i = simulated data point, and

N = response surface shaping parameter, chosen by the user.

Equation 4-1 has been used frequently with other GLUE applications (e.g., Beven and Binley, 1992). As the value of N increases, the magnitude of difference between the likelihood values of parameter sets with similar variance increases. When using equation 4.1, N must be started with a small value (starting from 1) to make sure the model simulations bracket the observed data (Keith Beven, personal communication). In this research, N value of 2 was used and the resulting model did bracket the observed data.

Once the model runs were completed, the input parameter sets that were not an acceptable simulator of the system were rejected. Previous GLUE applications have reported a wide variety of parameter set rejection criteria. Beven (1992) considered the parameter sets with very low likelihood values as non-behavioral parameter sets that can be rejected. Balin (2004) reported GLUE application using the topology model (TOPMODEL), in which he did not reject any parameter sets. The parameter rejection criteria may be subjectively decided by the modeler, depending upon the modeling objective. In the research reported here, to determine a suitable simulation rejection criteria a cumulative distribution function (CDF) of likelihood values for hydrologic calibration was plotted (figure 4.3). The visual assessment of the CDF curve suggests an inflection point around a cumulative probability of 0.9 (or 90%), and therefore the simulations that produced likelihood values less than the 90th percent value were rejected. The likelihood values of accepted model simulations were normalized to unity. The normalized likelihood values for each simulation were plotted against the parameter values resulting in dotty plots (illustrated later in Results). The dotty plots are used to obtain the posterior distribution of input parameters using Bayesian equation (Beven and Binley, 1992).

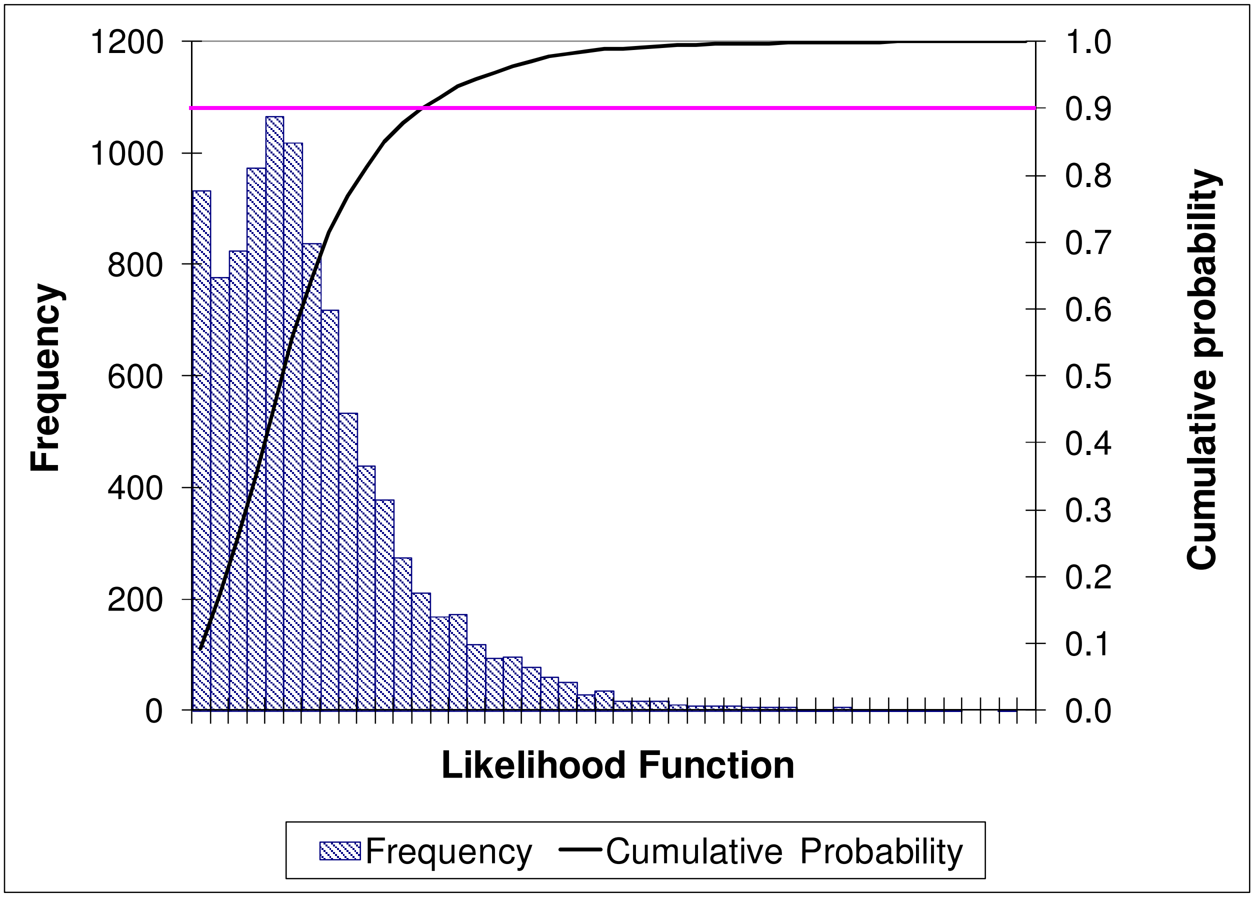


Figure 4.3 Histogram and cumulative distribution function of likelihood functions for hydrologic calibration.

4.1.4 Hydrologic and Water Quality Calibration of Mossy Creek Watershed Model

To conduct hydrologic calibration, the GLUE technique was used for the period – 1 September 1998 to 31 December 1999. The posterior parameter distributions obtained following the GLUE application can also be termed as calibrated parameter distributions. The posterior parameter distributions were used to conduct Monte Carlo simulation of the Mossy Creek model for the validation period – 1 January 2000 to 30 September 2002. To conduct water quality calibration, the GLUE technique was used to estimate the posterior distributions of water quality parameters for the calibration period – 1 October 1998 to 31 December 2001. Insufficient water quality data prevented a water quality validation.

4.1.5 TMDL Pollutant Allocation Scenarios

A TMDL allocation scenario allocates the pollutant load among different sources and hence suggests the amount of reduction in pollutant loading from each source needed to meet the applicable water quality standard. The Mossy Creek TMDL included several allocation scenarios with the two preferred allocation scenarios (table 4-5). To simulate the TMDL allocation scenarios in HSPF, a reduction factor was applied to the pollutant load from each source. For example, to simulate 94% reduction in cattle direct deposit of FC load, the pollutant load from cattle direct deposit was multiplied by 0.06.

Table 4-5 TMDL pollutant allocation scenarios for Mossy Creek TMDL resulting in no violations (Benham et al., 2004)

	Required source-specific fecal coliform load reductions (%)						
TMDL Allocation Scenario	Cattle Direct Deposit	Cropland	Pasture	Loafing Lot	Wildlife Direct Deposit	Straight Pipes	All residential pervious land segments
S1	99	90	98	100	30	100	95
S2	94	95	98	100	0	100	95

Used under fair use guidelines, 2011

Both TMDL allocation scenarios require a 100% reduction in FC loading from illegal straight pipes that discharge waste directly from homes. The major differences between the scenarios are the reduction in cattle direct deposit, wildlife direct deposit and loadings from cropland. Due to low production of FC by wildlife, uncertainty in wildlife direct deposit was not considered in this study. A period of three and a half years (1218 days) that represents a range of hydrological events in Mossy Creek was selected to simulate the in-stream FC concentration under the two allocation scenarios. The hydrology and water quality parameters posterior distributions obtained using the GLUE technique were used to conduct the Monte-Carlo simulations for the scenarios.

4.2 Results and Discussions

The hydrologic parameter posterior distributions for the Mossy Creek watershed model were developed using GLUE. Figure 4.4 illustrates examples of the dotty plots and the posterior distributions generated using GLUE for two hydrology parameters: LZSN-pasture (lower zone nominal soil moisture storage in the pasture land use), and DEEPFER (fraction of infiltrating water lost to deep aquifers). The shape of the LZSN-pasture posterior distribution (figure 4.4 a, the solid line) is very different from its prior distribution (figure 4.4 a, the dashed line), whereas the shape of DEEPFR posterior distribution is similar to the prior distribution (figure 4.4 b). The difference in prior and posterior distributions for the two parameters implies that the observed data provided greater information about the parameter LZSN-pasture than DEEPFR. In other words, the model was more sensitive to LZSN-pasture parameter than DEEPFR in the Mossy Creek watershed model. Plots similar to those illustrated in figure 4.4 were generated for all

HSPF hydrologic parameters, and the posterior distributions of all the hydrologic parameters were calculated (table 4-6).

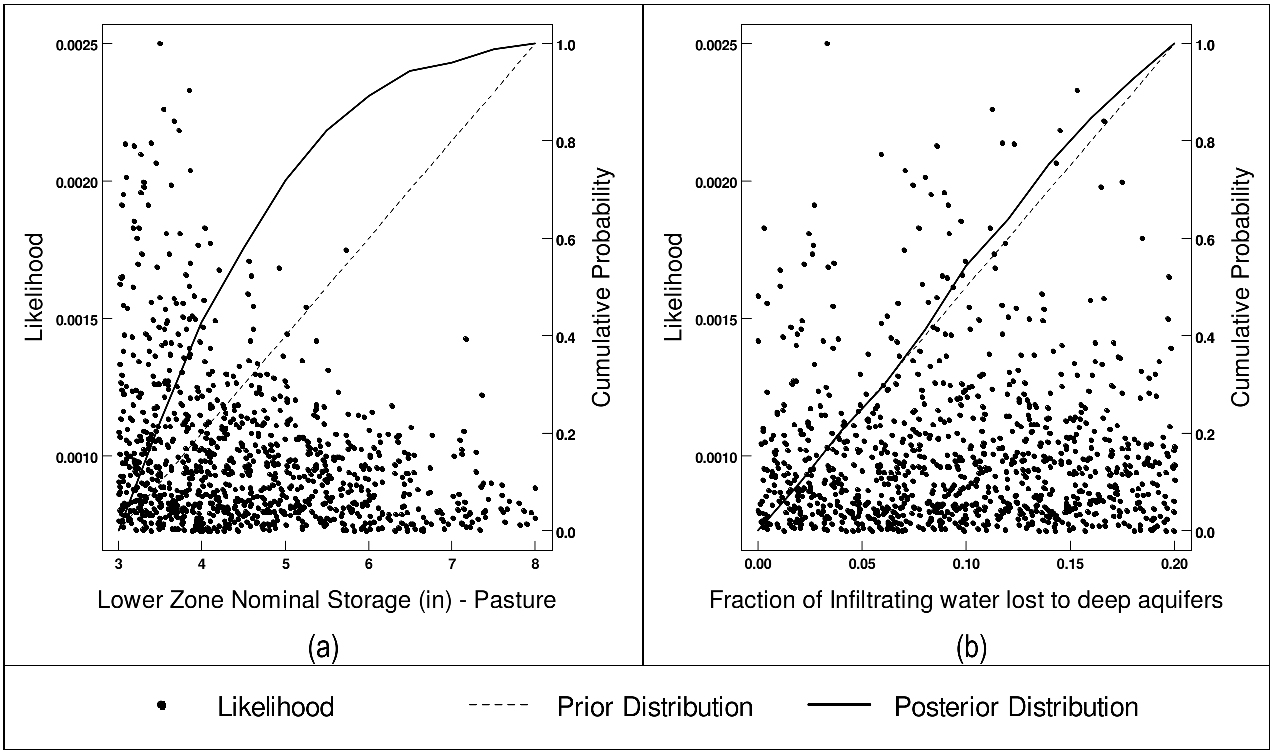


Figure 4.4 Posterior distribution of two hydrologic parameters, (a) LZSN – Pasture, and (b) DEEPFR obtained using GLUE technique

Table 4-6 Posterior distribution of all the hydrology parameters in Mossy Creek watershed model.

Parameter Name	Distribution Limits	Cumulative Distribution[¥]
LZSN-Forest	(3,8)	{0.14, 0.30, 0.45, 0.55, 0.63, 0.71, 0.78, 0.85, 0.93, 1.00}
LZSN-Cropland	(3,8)	{0.13, 0.26, 0.38, 0.48, 0.56, 0.66, 0.75, 0.83, 0.92, 1.00}
LZSN-Pasture	(3,8)	{0.24, 0.43, 0.61, 0.76, 0.86, 0.93, 0.97, 0.99, 1.00, 1.00}
LZSN-Farmstead	(3,8)	{0.09, 0.18, 0.27, 0.37, 0.46, 0.56, 0.68, 0.79, 0.90, 1.00}
LZSN-LDR	(3,8)	{0.09, 0.21, 0.32, 0.43, 0.54, 0.66, 0.73, 0.83, 0.92, 1.00}
LZSN-HDR	(3,8)	{0.11, 0.21, 0.31, 0.41, 0.49, 0.59, 0.69, 0.78, 0.91, 1.00}
LZSN-Loafing Lot	(3,8)	{0.13, 0.24, 0.35, 0.45, 0.56, 0.65, 0.72, 0.78, 0.88, 1.00}
INFILT-Forest	(0.05, 1.0)	{0.06, 0.38, 0.62, 0.78, 0.88, 0.94, 0.98, 0.99, 1.00, 1.00}
INFILT-Cropland	(0.03, 0.8)	{0.01, 0.15, 0.41, 0.62, 0.77, 0.89, 0.95, 0.99, 1.00, 1.00}
INFILT-Pasture	(0.04, 0.9)	{0.00, 0.03, 0.21, 0.42, 0.67, 0.83, 0.93, 0.98, 1.00, 1.00}
INFILT-Farmstead	(0.03, 0.5)	{0.01, 0.11, 0.38, 0.59, 0.77, 0.89, 0.95, 0.98, 1.00, 1.00}
INFILT-LDR	(0.03, 0.26)	{0.00, 0.02, 0.07, 0.18, 0.37, 0.60, 0.82, 0.93, 0.99, 1.00}
INFILT-HDR	(0.01, 0.1)	{0.27, 0.52, 0.70, 0.83, 0.90, 0.95, 0.98, 0.99, 1.00, 1.00}
INFILT-Loafing Lot	(0.15, 0.23)	{0.11, 0.21, 0.33, 0.43, 0.53, 0.61, 0.70, 0.78, 0.90, 1.00}
DEEPFR (all land uses)	(0.0, 0.2)	{0.11, 0.21, 0.33, 0.43, 0.52, 0.63, 0.73, 0.82, 0.91, 1.00}
BASETP (all land uses)	(0.00, 0.05)	{0.32, 0.57, 0.74, 0.86, 0.91, 0.95, 0.98, 0.99, 0.99, 1.00}
AGWETP (all land uses)	(0.00, 0.05)	{0.17, 0.34, 0.48, 0.63, 0.73, 0.82, 0.90, 0.95, 0.98, 1.00}
INTFW (all land uses)	(1.0, 3.0)	{0.11, 0.22, 0.31, 0.40, 0.51, 0.63, 0.74, 0.81, 0.90, 1.00}
IRC (all land uses)	(0.5, 0.8)	{0.00, 0.01, 0.06, 0.14, 0.30, 0.51, 0.77, 0.93, 0.99, 1.00}
MON INTERCEP-Forest-Jan	(0.03, 0.075)	{0.13, 0.31, 0.46, 0.62, 0.75, 0.82, 0.88, 0.92, 0.96, 1.00}
UZSN-Forest-January	(0.1, 0.3)	{0.03, 0.08, 0.13, 0.23, 0.34, 0.45, 0.56, 0.70, 0.83, 1.00}
LZETP-Forest-January	(0.07, 0.2)	{0.03, 0.10, 0.20, 0.29, 0.39, 0.53, 0.66, 0.79, 0.92, 1.00}

[¥]The cumulative distributions are the values at ten equal intervals between the distribution limits.

The posterior distributions of hydrologic parameters were used to conduct Monte Carlo simulations for the validation period – 1 January 1999 to 31 December 2002. The total daily flow volume from each HSPF Monte Carlo iteration was used to calculate the HSPEXP (HSPF Expert System) statistics (Lumb et al., 1994). Several quantiles were calculated using the HSPEXP statistics to validate the posterior distributions (table 4-7).

Table 4-7 Quantiles of the HSPEXP (Expert system for HSPF) statistics for the validation period when Monte Carlo simulations were conducted with "prior" and "posterior" distributions

Calibration Sufficiency Statistics	Default criteria (percent error)	Quantiles for validation period when "prior" distributions were used to conduct Monte Carlo simulations		Quantiles for validation period when "posterior" distributions were used to conduct Monte Carlo simulations	
		2.5	97.5	2.5	97.5
Total Volume	±10	-13.1	13.8	-10.3	8.2
50% Lowest Flows	±10	-8.8	23.8	*-4.7*	*16.8*
10% Highest Flows	±15	-16.5	19.0	*-16.9*	*1.0*
Storm Peaks	±20	-15.4	32.0	-16.1	1.7
Seasonal Volume Error	±30	0.9	16.3	0.2	11.1
Summer storm volume error	±50	-19.1	15.3	-15.7	7.0

Table 4-7 illustrates that overall, the hydrologic parameter posterior distributions or calibrated parameter distributions were acceptable, based on the results for the validation period.

Using the hydrology parameter posterior distributions produced calibration sufficiency statistics within HSPEXP criteria bounds for all except two statistics, 50% lowest flows and 10% highest flows. These results imply that including calibration sufficiency statistics in addition to the difference between observed and simulated daily runoff volume in the likelihood evaluation could perhaps improve the calibration.

Following hydrologic calibration and validation, the GLUE technique was applied to determine water quality parameter posterior distributions. For the water quality parameters, the likelihood function was evaluated using equation 4-1, using both non-transformed and log-transformed simulated daily average and observed instantaneous FC concentrations for the days for which observed data were available. Figure 4.5 illustrates an example of the dotty plots and posterior distributions generated using GLUE for two water quality parameters, FSTDEC (first order decay rate of bacteria) and ACQOP-Pasture (rate of FC accumulation on pasture land use) using both log-transformed and non-transformed data.

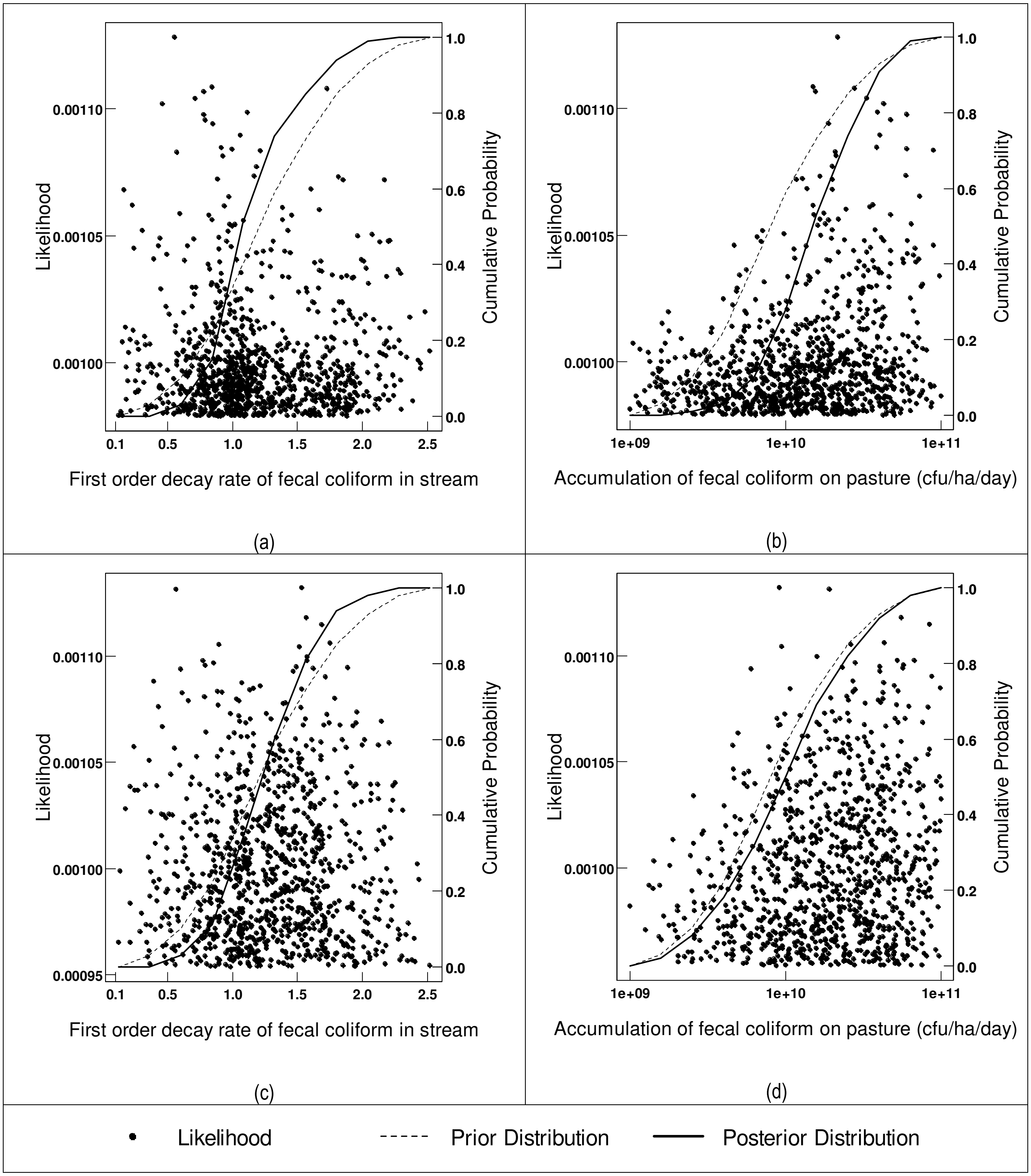


Figure 4.5 Posterior distribution of two water quality parameters obtained using GLUE technique with non-transformed fecal coliform concentrations (a and b), and log-transformed fecal coliform concentrations (c and d).

The difference in skewness of the likelihood values (illustrated by dots in figure 4.5) when the FC concentration was log-transformed versus non-transformed for the two parameters implies that the transformation of FC concentration can affect the distribution of likelihood values. The likelihood values were more skewed when FC concentration was not transformed (figure 4.5 a and b), compared to when the FC concentration was log-transformed (figure 4.5 c and d). As the FC concentration data can easily vary by orders of magnitude, a few high FC concentration values can affect the likelihood values and consequently the posterior probability distribution. The figures also illustrate that the posterior parameter distribution was different from the prior distribution for the two parameters, illustrating the effect of observed data. The posterior distributions of all the water quality parameters were similarly calculated with non-transformed FC concentration data (table 4-8) and log-transformed FC concentration data (table 4-9).

Table 4-8 Posterior water quality parameters obtained using GLUE without non log-transformed fecal coliform concentrations

Parameter Name	Distribution Limits	Cumulative Distribution¥
SQOLIM-FACTOR	(2.5, 11.5)	{0.08, 0.16, 0.23, 0.33, 0.40, 0.55, 0.65, 0.74, 0.85, 1.00}
WSQOP-Forest	(0.5, 2.4)	{0.11, 0.21, 0.32, 0.43, 0.55, 0.64, 0.72, 0.81, 0.91, 1.00}
WSQOP-Cropland	(0.5, 2.4)	{0.12, 0.22, 0.33, 0.42, 0.48, 0.55, 0.64, 0.78, 0.94, 1.00}
ACCUM-Pasture	(1E9, 1E11)	{0.00, 0.02, 0.07, 0.18, 0.39, 0.62, 0.83, 0.94, 1.00, 1.00}
WSQOP-Pasture	(0.5, 2.4)	{0.16, 0.30, 0.42, 0.51, 0.61, 0.71, 0.78, 0.87, 0.94, 1.00}
WSQOP-Farmstead, LDR	(0.5, 2.4)	{0.11, 0.19, 0.30, 0.40, 0.53, 0.61, 0.70, 0.80, 0.90, 1.00}
ACCUM_Loafing Lot	(1.2E11, 1.2E13)	{0.00, 0.03, 0.17, 0.23, 0.45, 0.69, 0.87, 0.95, 0.99, 1.00}
WSQOP-Loafing Lot	(0.05, 2.4)	{0.13, 0.24, 0.34, 0.45, 0.54, 0.63, 0.72, 0.82, 0.91, 1.00}
ACCUM-Cropland-Jan	(2E6, 2E8)	{0.00, 0.02, 0.07, 0.23, 0.54, 0.74, 0.89, 0.97, 1.00, 1.00}
FSTDEC	(0.12, 2.52)	{0.00, 0.03, 0.15, 0.51, 0.74, 0.85, 0.94, 0.99, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES1	(0.03, 0.5)	{0.00, 0.03, 0.10, 0.23, 0.50, 0.74, 0.89, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES2	(0.1, 10)	{0.00, 0.03, 0.10, 0.23, 0.49, 0.73, 0.86, 0.96, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES3	(0.1, 10)	{0.00, 0.03, 0.09, 0.24, 0.48, 0.75, 0.88, 0.95, 0.99, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES4	(0.1, 10)	{0.00, 0.03, 0.09, 0.27, 0.50, 0.77, 0.91, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES5	(0.1, 10)	{0.00, 0.04, 0.10, 0.24, 0.49, 0.73, 0.91, 0.97, 0.99, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES6	(0.1, 10)	{0.00, 0.01, 0.05, 0.14, 0.33, 0.57, 0.74, 0.88, 0.98, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES7	(0.1, 10)	{0.00, 0.04, 0.11, 0.24, 0.49, 0.73, 0.92, 0.97, 0.99, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES8	(0.1, 10)	{0.00, 0.01, 0.04, 0.10, 0.24, 0.35, 0.53, 0.77, 0.96, 1.00}

¥The cumulative distributions are values at 10 intervals between the distribution limits.

Table 4-9 Posterior water quality parameters obtained using log-transformed fecal coliform concentrations

Parameter Name	Distribution Limits	Cumulative Distribution¥
SQOLIM-FACTOR	(2.5, 11.5)	{0.04, 0.10, 0.20, 0.29, 0.40, 0.51, 0.61, 0.72, 0.84, 1.00}
WSQOP-Forest	(0.5, 2.4)	{0.11, 0.19, 0.29, 0.40, 0.50, 0.61, 0.71, 0.82, 0.90, 1.00}
WSQOP-Cropland	(0.5, 2.4)	{0.10, 0.22, 0.34, 0.43, 0.56, 0.66, 0.75, 0.82, 0.91, 1.00}
ACCUM-Pasture	(1E9, 1E11)	{0.00, 0.01, 0.03, 0.10, 0.28, 0.53, 0.74, 0.91, 0.99, 1.00}
WSQOP-Pasture	(0.5, 2.4)	{0.16, 0.30, 0.42, 0.51, 0.61, 0.71, 0.78, 0.87, 0.94, 1.00}
WSQOP-Farmstead, LDR	(0.5, 2.4)	{0.13, 0.24, 0.34, 0.45, 0.55, 0.66, 0.73, 0.83, 0.91, 1.00}
ACCUM-Loafing Lot	(1.2E11, 1.2E13)	{0.00, 0.02, 0.07, 0.18, 0.37, 0.67, 0.84, 0.95, 0.99, 1.00}
WSQOP-Loafing Lot	(0.05, 2.4)	{0.16, 0.28, 0.39, 0.50, 0.61, 0.68, 0.77, 0.85, 0.91, 1.00}
ACCUM-Cropland-Jan	(2E6, 2E8)	{0.00, 0.03, 0.07, 0.28, 0.58, 0.76, 0.91, 0.96, 1.00, 1.00}
FSTDEC	(0.12, 2.52)	{0.00, 0.03, 0.11, 0.34, 0.60, 0.81, 0.94, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES1	(0.03, 0.5)	{0.00, 0.04, 0.16, 0.39, 0.62, 0.81, 0.93, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES2	(0.1, 10)	{0.00, 0.04, 0.16, 0.34, 0.62, 0.78, 0.92, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES3	(0.1, 10)	{0.00, 0.04, 0.16, 0.39, 0.66, 0.85, 0.95, 0.99, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES4	(0.1, 10)	{0.00, 0.04, 0.14, 0.30, 0.56, 0.78, 0.91, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES5	(0.1, 10)	{0.00, 0.05, 0.17, 0.40, 0.67, 0.85, 0.94, 0.99, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES6	(0.1, 10)	{0.00, 0.04, 0.19, 0.45, 0.81, 0.99, 1.00, 1.00, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES7	(0.1, 10)	{0.00, 0.04, 0.12, 0.33, 0.56, 0.75, 0.89, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES8	(0.1, 10)	{0.00, 0.03, 0.12, 0.34, 0.66, 0.91, 1.00, 1.00, 1.0, 1.00}

¥The cumulative distributions are values at 10 intervals between the distribution limits.

The water quality parameter posterior distributions were not validated due to insufficient observed data. The Mossy Creek watershed model input parameter posterior distributions that were obtained both with and without log-transformations of the FC concentration were used to conduct Monte Carlo simulations for two of the TMDL pollutant allocation scenarios suggested in the Mossy Creek TMDL (table 4-5). The daily average simulated in-stream FC concentration time series from each HSPF iteration was used to compute an average, 2.5%, 10%, 90% and 97.5% quantiles for each day (figure 4.6 and figure 4.7). Each occurrence of FC concentration greater than instantaneous FC criteria of 400 cfu/100 ml was considered a violation in each time series. The percent of water quality criterion violations for each time series was calculated by dividing the number of violations on a daily basis by the number of days in the simulation period (table 4-10).

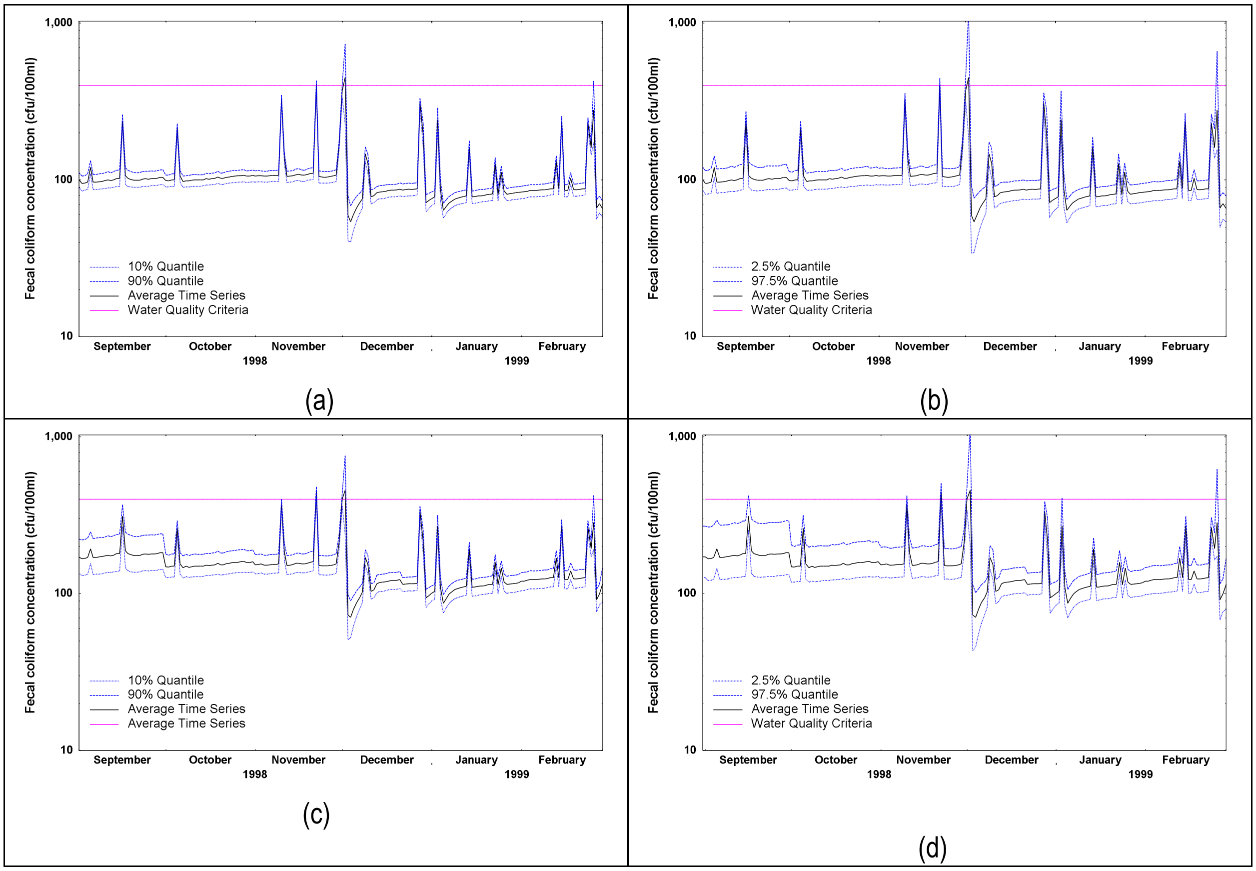


Figure 4.6 TMDL allocation scenario S1, 80% probability interval (a) and 95% probability interval (b); TMDL allocation scenario S2, 80% probability interval (c) and 95% probability interval (d); using non-transformed FC concentration. Representative plots showing first six months of simulation period.

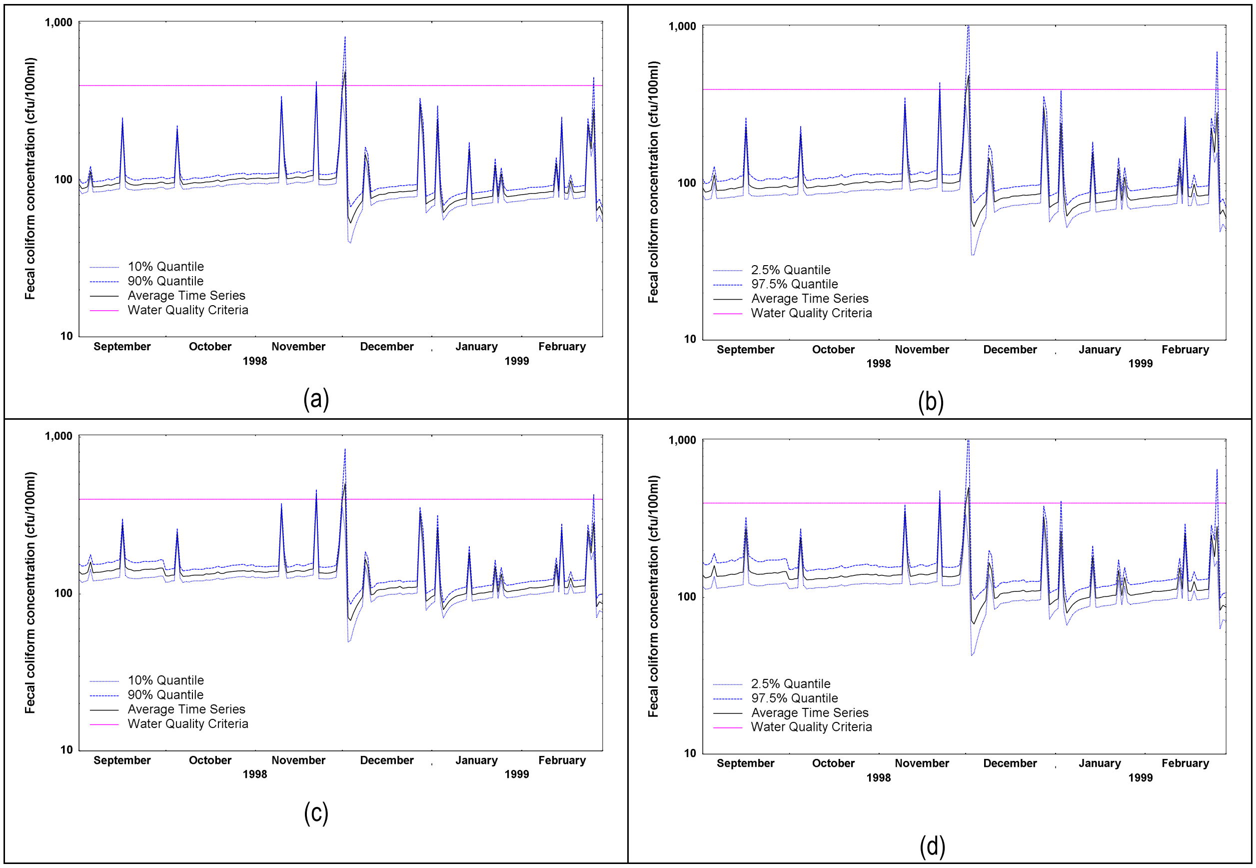


Figure 4.7 TMDL allocation scenario S1, 80% probability interval (a) and 95% probability interval(b); TMDL allocation scenario S2, 80% probability interval (c) and 95% probability interval (d); using log-transformation of FC concentration. Representative plots show results for first six months of simulation period.

Table 4-10 Percent of water quality criterion violations by the average time series, and the probability intervals for two TMDL allocation scenarios when GLUE was performed with and without log-transformation of FC concentration.

Fecal coliform concentration transformation	TMDL Allocation Scenario	Average time series violations (%)	80% probability interval violations (%)	95% probability interval violations
Non-transformed	S1	1.2	(0.2, 1.6)†	(0.0, 2.6)
Non-transformed	S2	1.9	(0.6, 14.6)	(0.2, 22.7)
Log-transformed	S1	1.2	(0.2, 1.6)	(0.0, 2.6)
Log-transformed	S2	1.3	(0.4, 1.8)	(0.1, 3.1)

† Numbers in parentheses show the percent of violation incidences over a period of 1218 days by the respective time series for the upper and lower bounds of the probability interval.

In figure 4.6 and figure 4.7, the 95% probability intervals–frames (b) and (d)–are wider than the 80% probability interval–frames (a) and (c)–for both allocation scenarios. These graphs illustrate that uncertainty increases as we seek greater confidence in predicted results. The percent of FC concentration criterion violations by the average time series was not significantly different across the allocation scenarios when the GLUE technique was applied with or without log-transformation of FC concentration (table 4-10). The S2 scenario exhibited greater uncertainty in FC concentration criterion violations than S1.

For the S2 allocation scenario, the FC log-transformation reduced the percent violations of the upper bound by a factor of 8 for the 80% probability interval and a factor of 7 for the 95% probability. Observed and simulated FC concentration can vary by several orders of magnitude, and the log transformation reduced the effect of some observed data that were orders of magnitude greater than others. In Mossy Creek, for example, the observed FC concentration varied from 25 cfu/100ml to as high as 50,000 cfu/100 ml. Application of the GLUE technique using log-transformed FC concentration resulted in parameter sets that were not affected by the large fluctuations of FC concentration and hence helped reduce the uncertainty in predicted FC concentration.

4.3 Summary and Conclusions

Generalized Likelihood Uncertainty Estimation (GLUE) was used to estimate the predictive uncertainty in FC criterion violations when using a HSPF model. The study used the data and the HSPF model developed to generate the bacterial impairment TMDL for Mossy Creek, located in Virginia's Shenandoah Valley. This study illustrates a unique application of the GLUE technique with an HSPF based watershed model, and presented a framework to estimate predictive uncertainty in water quality modeling. Two Mossy Creek TMDL pollutant allocation scenarios were compared. Although the two recommended allocation scenarios were represented as the scenarios that would result in zero percent of water quality criterion violations, the analysis performed here illustrates that when uncertainty was taken into account, the FC water quality criterion would be violated 1-2% of the time for both allocation scenarios assuming full implementation. The amount of uncertainty, however, differed for the two allocation scenarios, and the allocation scenario allowing greater FC input from cattle direct deposit (S2) exhibited the greatest uncertainty. Although the input uncertainty in FC direct deposit and FC cropland loading were similar (log-transformed distribution spread over two orders of magnitude), the differences in uncertainty between the two scenarios illustrate that the cattle direct deposit is a greater source of uncertainty in FC criterion violations than cropland runoff in Mossy Creek watershed.

From a TMDL perspective, these results give stakeholders and decision makers more insight into moving forward with TMDL implementation. As discussed in Chapter 3, the allocation scenario that meets the water quality standard with greater confidence is generally more expensive to implement, and therefore the decision about the level of confidence required to achieve a specific watershed management objective should be made by local interests (i.e., stakeholders, elected officials). Concerned parties might like to have greater confidence in predicting future water quality under different watershed management scenarios if the ecosystem is fragile and there are potentially grave consequences of water quality violations. In an adaptive TMDL implementation program where the water quality monitoring, modeling, and research continues during the implementation process, the stakeholders might like to tackle the pollutant source that contributes greater uncertainty in model prediction ahead of other sources.

Formulation of a likelihood function is an important step in GLUE application. The choice of likelihood formulation can affect the posterior distributions of parameters and hence the uncertainty estimates of model predictions. In this research, likelihood was evaluated using both log-transformed and non-transformed simulated and observed FC concentrations. The posterior distributions of the water quality parameters obtained for the two cases were used to estimate the predictive uncertainty in FC concentration criterion violation. When the log-transformed data were used, the number of water quality violations reported by 90 and 97.5% quantiles decreased by 8 and 7 times, respectively, for one of the scenarios. In other words, reported uncertainty reduced when the FC concentration data were log-transformed. Observed and simulated FC concentration can vary by orders of magnitude and few high concentrations can affect the posterior distributions and eventually uncertainty in model output, when we use GLUE without normalizing the data. These results underscore the importance of normalizing the observed data in GLUE application for uncertainty estimation.

The research presented here is one of the first applications of GLUE with water quality model developed using HSPF, and the results underscore its applicability in estimating predictive uncertainty for complex watershed models. The uncertainty estimation using the GLUE technique, however, may depend upon the factors like the choice of likelihood function and parameter set acceptance/rejection criteria, which warrants further research in similar applications.

References:

Balin, D. 2004. *Hydrological Behaviour through Experimental and Modelling Approaches. Application to the Haute-Mentue Catchment.* School of Acrhitecture, Civil and Environmental Engineering. Lausanne, Switzerland. Ecole Polytechnique Fédérale de Lausanne.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Benaman, J., and C.A. Shoemaker. 2002. Sensitivity and uncertainty analysis of a distributed watershed model for the TMDL process. National TMDL Science and Policy 2002 Speciality Conference, Water Environment Federation.

Benham, B., K. Branna, K. Christophel, T. Dillaha, L. Henry, S. Mostaghimi, R. Wagner, J. Wynn, G. Yagow, and R. Zeckoski. 2004. *Total maximum daily load development for Mossy Creek and Long Glade Run: Bacteria and general standard (Benthic) impairments.* Richmond, Va.: Virginia Department of Environmental Quality. http://www.deq.state.va.us/tmdl/homepage.html. Assessed Jan. 25, 2005.

Benham, B.L., K.M. Brannan, T.A. Dillaha, S.M. Mostaghimi, J. Wynn, G. Yagow, and R. Zeckoski. 2003. *Total maximum daily load development for Linville Creek: Bacteria and General standard (Benthic) impairments.* Richmond, Va.: Virginia Department of environmental quality. http://www.deq.state.va.us/tmdl/apptmdls/shenrvr/linville.pdf. Assess Jan 30, 2006.

Benham, B.L., K. Brannan, T. Dillaha, S. Mostaghimi, G. Yagow. 2002. TMDLs (Total Maximum Daily Loads) – Terms and Definitions. *Virginia Cooperative Extension.* Pub No. 442-550.

Beven, K., and J. Freer. 2001. Equifinality, Data Assimilation, and Uncertainty Estimation in mechanistic modelling of Complex Environmental Systems using GLUE Methodology. *Journal of Hydology* 249(1-4): 11-29.

Beven, K.. and A. Binley (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr. T.H. Jobes, and A.S. Donigian, Jr. 2005. *HSPF Version 12.2 User's Manual.* AQUA TERRA Consultants. Mountain View, CA.

Freer, J. K. Beven, and B. Ambroise. 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research.* 32(7): 2161-2173.

Lumb, A. M., R. B. McCammon, and John L. Kittle, Jr. 1994. *Users Manual for an Expert Systems (HSPEXP) for Calibration of the Hydrological Simulation Program - FORTRAN.* Reston, Virginia, U.S. Geological Survey, Water-Resources Investigations Report.

Setegn, S.G., R. Shrinivasan, A.M. Melesse, B. Dargahi. 2009. SWAT model application and uncertainty analysis in the Lake Tana Basin, Ethiopia. *Hydological processes.* 24(3): 357-367

Stedinger J.R., R.M. Vogel, S.U. Lee, and R. Batcheldar. 2008. Appraisal of the Generalized Likelihood Uncertainty Estimation (GLUE) method. *Water Resources Research.* 44, W00B06, doi:10.1029/2008WR006822.

Stow, C.A., K.H. Reckhow, S.S. Qian, E.C. Lamon, G.B. Arhonditsis, M.E. Borsuk, and D. Seo. 2007. Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. *J. Amer. Water Resources Ass.* 43(6): 1499-1507

Suter, G.W. II, L.W. Barthouse and R.V. O'Neill. 1987. Treatment of risk in environmental impact assessment. *Environmental Mgmt.* 11(3): 295-303.

USEPA. 2006. National Section 303(d) List Fact Sheet. United States Environmental Protection Agency. Available at http://oaspub.epa.gov/waters/national_rept.control. Accessed Mar. 20, 2006.

USEPA. 2001. *The National Cost of the Total Maximum Daily Load Program (Draft Report).* Office of Water, United States Environmental Protection Agency. Washington D.C.

USEPA. 2000. *BASINS Technical Note 6.* United States Environmental Protection Agency. Washington D.C. Office of Water.

VADCR. 2003. Bacteria TMDLs for Abrams Creek and Upper and Lower Opequon Creek Located in Frederick and Clarke County, Virgnia. Virginia Department of Environmenatal Quality, and Virginia Department of Conservation and Recreation. Richmond, VA, USA.

Yagow, G. 2001. Fecal *Coliform TMDL Mountain Run watershed, Culpeper County, Virginia.* Biological Systems Engineering, Virginia Tech, Blacksburg, Va. http://www.deq.virginia.gov/tmdl. Accessed September 10, 2006.

Zheng, Y., and A.A. Keller. 2007. Uncertainty Assessment in Watershed-Scale Water Quality Modeling and Management:1. Framework and Application of Generalized Likelihood Uncertainty Estimation (GLUE) Approach. *Water Res. Res.* 43. doi: 10.1029/2006WR005345

Chapter 5. Evaluation of the applicability of Generalized Likelihood Uncertainty Estimation and Markov Chain Monte Carlo to estimate uncertainty in HSPF based water quality modeling.

Abstract: *Generalized Likelihood Uncertainty Estimation (GLUE) and Markov Chain Monte Carlo (MCMC) were used to estimate uncertainty in percent violations of instantaneous FC concentration criteria predicted by a watershed model developed using Hydrological Simulation Program – FORTRAN (HSPF), for Mossy Creek watershed in Virginia. GLUE and MCMC are based on similar concepts and can be used to obtain posterior (or calibrated) distributions of model parameters based on observed data and prior parameter distribution using the Bayesian equation. The posterior distributions were used to simulate in-stream fecal coliform (FC) concentration for two pollutant allocation scenarios presented in the Mossy Creek TMDL. The TMDL allocation scenarios differed in the reductions specified from cattle directly depositing FC in the stream, and FC loadings from cropland. Results showed that with either technique, the instantaneous FC criterion was violated approximately 1% of the time (on a daily basis) for the prediction period for both allocation scenarios. The scenario allowing greater input of FC direct deposit in streams produced greater uncertainty, illustrating that FC direct deposit in streams is a greater source of uncertainty in in-stream FC criterion violations than FC loadings from cropland. Decision makers can use the results of an uncertainty assessment like this to select among alternative TMDL allocation scenarios, to set realistic targets for water quality achievements, and to prioritize implementation efforts. These results also show that uncertainty reported by GLUE and MCMC were similar, however, MCMC is more computationally intensive than GLUE.*

Keywords: *GLUE, Generalized Likelihood Uncertainty Estimation, Markov Chain Monte Carlo, MCMC, HSPF, fecal coliform, uncertainty analysis.*

Introduction

The Clean Water Act classifies the water bodies that do not meet water quality standards as ''impaired,'' and requires total maximum daily loads (TMDLs) to be developed for those water bodies. A Total Maximum Daily Load (TMDL) specifies the maximum amount of a particular pollutant a waterbody can receive and still meet applicable water quality standards (Benham et al., 2002). In 2001, the U.S. Environmental Protection Agency (USEPA) estimated the annual average cost of developing TMDLs to be $63-69 million per year for the next fifteen years and the cost of implementing TMDLs to be between $1 and 3.4 billion per year for the next decade

(USEPA, 2001). Pathogens, typically represented by the surrogate indicator bacteria (IB), being the second most widespread cause of water quality impairments (USEPA, 2006) will be responsible for a significant share of this expense.

TMDL development often includes the application of water quality modeling software. Most water quality modeling software currently used to aid in developing TMDL includes modules that are process-based, empirical or a mixture. These software do not include detailed uncertainty analysis capabilities. Uncertainty can be a result of limited knowledge of the system or stochastic variability (Beck 1987; Suter et al, 1987). Typically, a margin of safety is included to account for the inherent uncertainty present in determining the TMDL, but there is a limited science-based guidance available to estimate the margin of safety. Without some formal measure of the uncertainty, one cannot accurately assess the probability of achieving a given water quality criterion, or the risk of violating it.

The Hydrological Simulation Program–FORTRAN (HSPF) is a continuous simulation model that simulates various hydrological and water quality processes (Bicknell et al., 2005), and has been widely used to develop IB impairment TMDLs (e.g., Benham et al., 2005; Benham et al., 2003; VADCR, 2003; Yagow, 2001). HSPF is a lumped parameter, watershed scale model, and produces a deterministic time-series of hydrology and water quality. Despite its widespread usage, there have been few applications of HSPF that included a detailed uncertainty analysis. In this dissertation, the application of two-phase Monte Carlo approach to estimate uncertainty is illustrated in Chapter 3, and application of the Generalized Likelihood Uncertainty Estimation (GLUE) approach is illustrated in Chapter 4. GLUE uses the Bayesian equation (Beven and Binley, 1992) to estimate inferences about the input parameters of the model. These estimated inferences or probability distributions are termed as 'posterior' distributions as opposed to 'prior' distributions that reflect the pre-existing knowledge about the parameter. The posterior distributions are then used to conduct Monte Carlo simulations for the prediction period and estimate uncertainty in the model output (Beven and Binley, 1992).

Another technique that can be utilized to estimate inferences about the model parameters is Markov Chain Monte Carlo (MCMC) (Kuczera and Parent, 1998). The MCMC method generates samples of parameter values from the posterior distribution by constructing a Markov Chain that has the posterior distribution as its equilibrium distribution (Robert and Casella, 1999). The posterior distribution of input parameters can be used to estimate uncertainty in model output. The MCMC approach has been used by many researchers to estimate posterior distributions and uncertainty with various hydrologic software (Balin, 2004; Makowski et al., 2002;

Marshall et al, 2005) and it has been suggested as a viable approach for estimating uncertainty in water quality modeling (Stow et al., 2007). Application of Bayesian techniques like GLUE and MCMC to estimate uncertainty related to water quality modeling is limited.

The objective of this research is to compare the applicability of GLUE and MCMC techniques in estimating the uncertainty in in-stream FC concentrations as it relates to TMDL development. GLUE and MCMC techniques were used to estimate the posterior distributions of input parameters in the Mossy Creek HSPF watershed model. These posterior distribution parameters were used to simulate in-stream FC concentration in the Mossy Creek. Simulation results were used to estimate the uncertainty in FC concentration criterion violations for the prediction period. The research also establishes a methodology to use MCMC for uncertainty analysis with HSPF in similar applications.

5.1 Materials and Methods

The materials and methods in Chapter 5 include some details that were discussed in Chapters 3 and 4. The common sections are discussed only briefly here.

5.1.1 Study Area

Mossy Creek, located in Rockingham and Augusta counties in Virginia (figure 5.1) was selected for this research. Mossy Creek was listed as impaired in 1996 due to violations of Virginia's Primary Contact Recreational standard's FC criterion. The Department of Biological Systems Engineering (BSE) at Virginia Tech developed a bacterial TMDL for Mossy Creek (Benham et al., 2004). The Mossy Creek watershed (4076 ha) is characterized as a rolling valley with Blue Ridge Mountains to the east and the Appalachian Mountains to the west. The predominant land uses in Mossy Creek watershed are pasture and agriculture. The primary sources of FC identified in the Mossy creek TMDL were direct deposition of feces in the stream by cattle (cattle loitering and defecating in the stream), and runoff from pastures where grazing animals defecate.

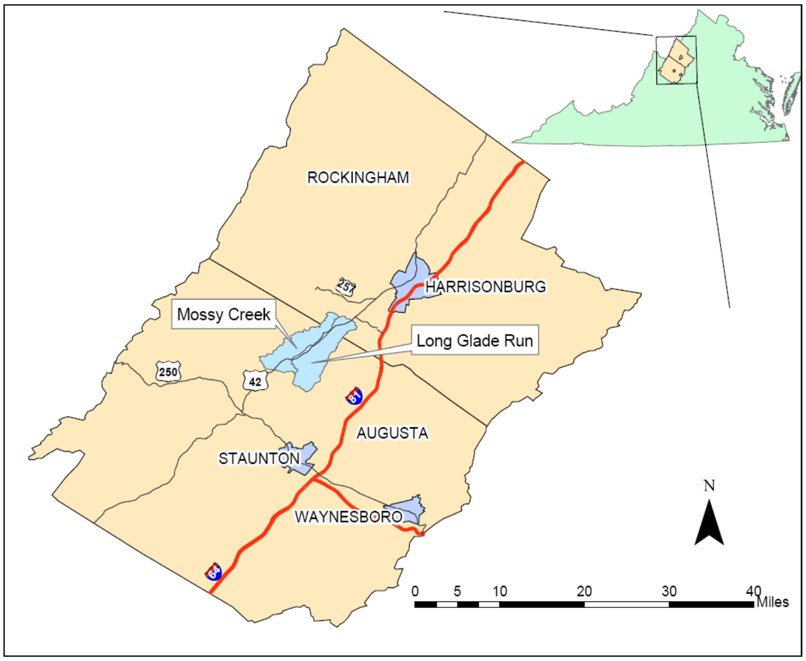


Figure 5.1 Mossy Creek Watershed (Benham et al., 2004)
Used under fair use guidelines, 2011

Mossy Creek was monitored monthly by the Virginia Department of Environmental Quality (DEQ) between July 1992 and March 2003 for FC concentration and other selected water quality constituents at the station ID 1BMSS001.35 located near the outlet of the Mossy Creek watershed. BSE monitored Mossy Creek semi-monthly between February 1998 and December 2001 for selected water quality constituents including FC concentration near the DEQ site. Daily flow data were also collected from May 1998 to December 2002 at the same site.

5.1.2 Mossy Creek Watershed Model

HSPF was used in developing the Mossy Creek bacterial impairment TMDL (Benham et. al., 2004). Mossy Creek was divided into eight subwatersheds for modeling and land use identification purposes (figure 5.2, table 5-1). Other data required by the model included rainfall, FC loading from cattle and wildlife, inflows from springs, solar radiation, and temperature as time series. The model development and data acquisitions are described in the Mossy Creek TMDL (Benham et. al., 2004).

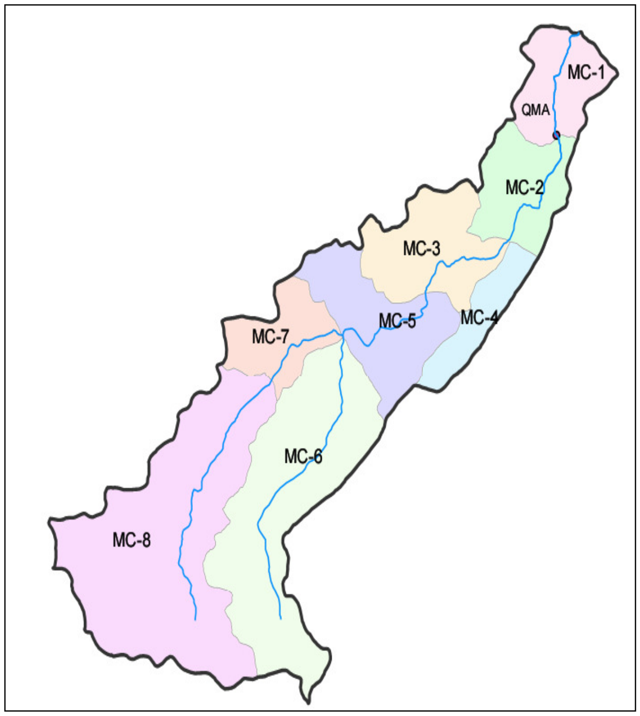


Figure 5.2 Mossy Creek watershed and its subwatersheds (Benham et al., 2004)
Used under fair use guidelines, 2011

Table 5-1 Land use distribution of Mossy Creek watershed (Benham et al., 2004)

Land use	Area (ha)	Percent of total area (%)
Forest	1025.1	25.15
Cropland	556.0	13.64
Pasture	2347.6	57.59
Farmstead	55.0	1.35
Low Density Residential	87.0	2.13
High Density Residential	3.6	0.09
Loafing Lot	1.6	0.04

Used under fair use guidelines, 2011

Using HSPF to simulate FC requires information about several hydrologic and water quality parameters. GLUE and MCMC procedures require the probability distribution of these input parameters. These input distributions are also termed as “prior distribution” as they reflect the knowledge of the modeler about the model parameters prior to the assessment of model with the observed values. In many previous studies, modelers assigned a uniform distribution to most of the input parameters (Beven and Freer, 2001) to avoid a modeler’s bias. For this research, we assigned prior distributions to model parameters based on literature review, expert opinion and GIS data available for Mossy Creek watershed. The detailed process of assigning the distributions to the model parameters is described in the Chapter 3 of this dissertation. The input parameter distributions are listed in table 5-2 through table 5-4.

Table 5-2 Distribution of hydrology parameters that apply to all and land uses

Parameter	Parameter Description	Type of Distribution
LZSN (inches)	Lower zone nominal soil moisture storage	Uniform (3,8)†
AGWRC	Groundwater recession rate	Uniform (0.92, 0.99)
DEEPFR	The fraction of infiltrating water lost to deep aquifers	Uniform (0.0, 0.2)
BASETP	Evapotranspiration by riparian vegetation as active groundwater enters streambed	Uniform (0.0, 0.05)
AGWETP	Fraction of model segment that is subject to direct evaporation from groundwater storage	Uniform (0.0, 0.05)
IRC	Interflow Recession Coefficient	Triangular (0.5, 0.7, 0.8)‡
INTFW	Coefficient that determines the amount of water which enters the ground from surface detention and becomes interflow	Uniform (1.0, 3.0)

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

Table 5-3 Distribution of hydrologic parameters which vary according to the land use and time of year, for the month of January.

Parameter	Land use	Distribution
INFILT (in/hr) Index to mean infiltration rate.	Forest	Triangular (0.05, 0.1, 1)†
	Cropland	Triangular (0.03, 0.17, 0.24)
	Pasture	Triangular (0.04, 0.09, 0.9)
	Low Density Residential	Triangular (0.03, 0.17, 0.26)
	High Density Residential	Triangular (0.01, 0.01, 0.1)
	Farmstead	Triangular (0.03, 0.15, 0.23)
	Loafing Lot	Uniform (0.15, 0.23)‡
UZSN (inches) Nominal upper zone soil moisture storage	Forest	Uniform (0.2, 0.3)
	Cropland	Uniform (0.06, 0.1)
	Pasture	Uniform (0.06, 0.1)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.06, 0.1)
CEPSC (inches) Interception Storage Capacity	Forest	Uniform (0.05, 0.075)
	Cropland	Uniform (0.05, 0.075)
	Pasture	Uniform (0.05, 0.075)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.05, 0.075)
LZETP Index to lower zone evapotranspiration	Forest	Uniform (0.1, 0.2)
	Cropland	Uniform (0.1, 0.2)
	Pasture	Uniform (0.1, 0.2)
	Farmstead, low and high density residential areas and loafing lots	Uniform (0.1, 0.2)

†Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

Table 5-4 Summary of water quality parameters which have been reported as sensitive and are typically calibrated when using HSPF

Parameter	Land use	Type of Distribution
ACQOP-PERLND (cfu day^{-1}) (Accumulation of fecal coliform on pervious land per day)	Pasture Loafing Lot Cropland (January)	Log-triangle (1×10^9, 1×10^{10}, 1×10^{11})† ‡ Log-triangle (1.12×10^{11}, 1.12×10^{12}, 1.12×10^{13}) Log-triangle (2×10^6, 2×10^7, 2×10^8)
SQOLIM adjustment Factor (Factor which is multiplied to ACCUM values to obtain SQOLIM	All	Uniform (2.5, 11.5)*
SQOLIM-PERLND (Maximum accumulation of FC on pervious land)	All	ACQOP-PERLND (for each land use) x SQOLIM adjustment Factor
WSQOP-PERLND (Rate of surface runoff that will remove 90% of stored bacteria from pervious land surface)	All	Uniform (0.5, 2.4)
FSTDEC (day^{-1}) (First order decay rate of bacteria)	All	Triangular (0.12, 1.1, 2.52)

† Numbers in parentheses show lower limit, mode, and upper limit of the triangular distribution, respectively. ‡ Log-triangle distribution implies that the logarithm of lower limit, mode and higher limit follows a triangular distribution. * Numbers in parentheses show lower and upper limit of the uniform distribution, respectively.

5.1.3 Generalized Likelihood Uncertainty Estimation

Likelihood function formulation is an important step in conducting GLUE. For this application, the likelihood function was calculated using the variance of the residuals (equation 5-1). For the hydrologic calibration, the value was calculated using observed and simulated daily flow, and for the water quality calibration, the likelihood value was calculated using log-transformed simulated daily average FC concentration and observed instantaneous FC concentrations. The FC concentrations can generally vary by orders of magnitude and, as illustrated in Chapter 4, log-transformation can help reduce the effect of a few high FC concentrations on uncertainty estimates.

$$L_e = (\sigma_e^2)^{-N} \qquad \text{5-1}$$

Where,

$$\sigma^2{}_e = 1/n\left(\sum_{i=1}^{n}\left(Y_i - Q_i\right)^2\right),$$

L_e = likelihood value,

σ_e^2 = variance of the residuals or mean square error,

n = number of data points,

Y_i = observed data point,

Q_i = simulated data point, and

N = shaping parameter, chosen by the user.

Equation 5-1 has been used frequently with other GLUE applications (e.g., Beven and Binley, 1992). As the value of *N* increases, the magnitude of difference between the likelihood values of parameter sets with similar variance increases. When using equation 5.1, *N* must be started with a small value (starting from 1) to make sure the model simulations bracket the observed data (Keith Beven, personal communication). In this research, *N* value of 2 was used and the resulting model did bracket the observed data.

Once the model runs were completed, the input parameter sets that were not an acceptable simulator of the system were rejected. The parameter rejection criteria may be subjectively decided by the modeler, depending upon the modeling objective. To decide rejection of parameter sets, a cumulative distribution function of likelihood values for hydrologic calibration (figure 5.3) was plotted. Visual inspection of CDF curves illustrates an inflection point at about 0.9 (or 90%), and therefore we rejected the 90% iterations resulting in lower likelihood values than others. The likelihood values of remaining iterations were normalized to unity. The normalized likelihood values for each simulation were plotted against the parameter values resulting in dotty plots. The dotty plots are used to obtain the posterior distribution of input parameters using Bayesian equation (Beven and Binley, 1992).

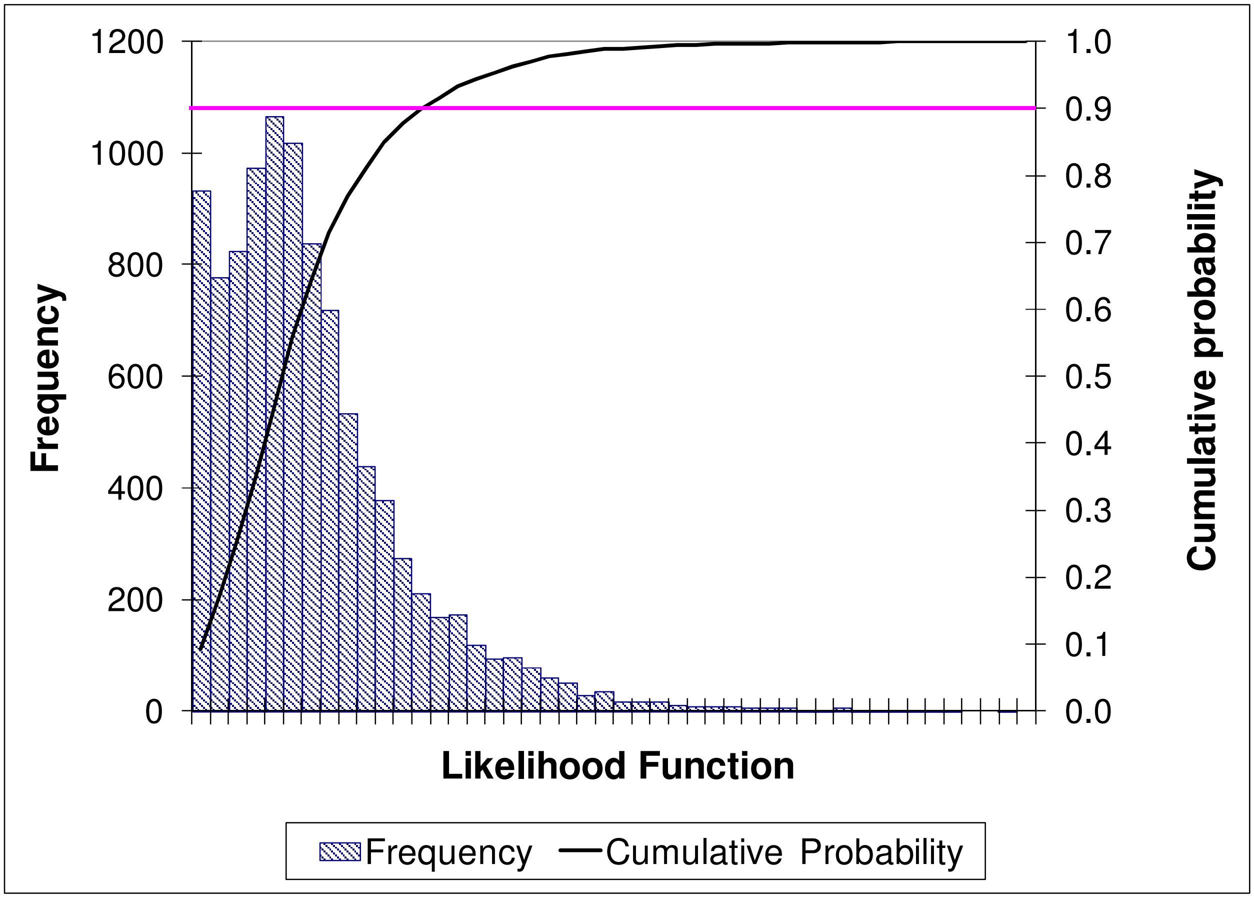


Figure 5.3 Histogram and cumulative distribution function of likelihood functions for hydrologic calibration.

5.1.4 Markov Chain Monte Carlo (MCMC)

Markov Chain Monte Carlo can be used as a Bayesian technique to estimate input parameter posterior distributions. If there are *p* unknown model parameters, θ: (θ_1, θ_2, … θ_p), and there is some inherent information about these parameters, this inherent information can be represented in the form of a probability distribution $\pi(\theta)$. This probability distribution is termed as the 'prior' distribution. The dependence of observed data *Y* on the *p* parameters θ is termed as a likelihood function, $L(Y/\theta)$. Likelihood is a function of the simulated and observed data of the modeled system. This likelihood function is used to update each parameter's probability distribution using the following Bayesian equation.

$$\pi(\theta \mid Y) = \frac{\pi(\theta) L(Y \mid \theta)}{\int_{\theta} \pi(\theta) L(Y \mid \theta) d\theta} \qquad \text{5-2}$$

Where $\pi(\theta/Y)$ is known as the parameter posterior distribution that expresses the probability of the parameter after using the observed data, Y, of the system being modeled. The denominator in equation 5-2 is a normalizing constant and hence it can be written as

$$\pi(\theta \mid Y) \propto \pi(\theta) L(Y \mid \theta) \qquad 5\text{-}3$$

To obtain the posterior distribution as described above, MCMC approach creates a random walk or a Markov process that has $\pi(\theta/Y)$ as its stationary distribution. The process is run sufficiently long so that the resulting sample closely approximates a sample from $\pi(\theta/Y)$ (Robert and Casella, 1999). Metropolis et al. (1953) proposed the Metropolis algorithm to build a Markov chain. An important step in implementing the MCMC approach is the choice of a statistical likelihood[1] function. For n observations, as is the case with time-series output, the statistical likelihood function is given by

$$L(\theta \mid Y) = \frac{1}{(\sqrt{2\pi})^n \sigma^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - Q_i)^2 \right) \qquad 5\text{-}4$$

Where, σ = variance of residuals, Y_i = i th observed data point, and Q_i = i th simulated data point.

Equation 5-4 assumes that the residuals between observed and simulated values are normally distributed. To build a Markov chain, where a new parameter value is sampled starting from the previous one, a jump specification is needed. For the random walk, the new parameter value is sampled around the previous one using a symmetric probability distribution or $\pi(\theta_{old}|\theta_{new}) = \pi(\theta_{new}|\theta_{old})$, where θ_{old} is the previous parameter vector and θ_{new} is the new parameter vector. This symmetric distribution is centered on the last accepted parameter value by the relationship $\theta_{new}|\theta_{old} = N(\theta_{old}, s.I)$, where s is the variance scaling factor and $\boldsymbol{I}$ is the identity matrix. The variance scaling factor is very important as it affects how the Markov chain moves towards equilibrium. A high variance scaling factor may lead to slow chain movement and a small variance scaling factor may result in a haphazard parameter chain across possible parameter space. Although there is guidance to estimate the scaling factor, it is generally obtained by trial and error (Gelman et al., 2000).

Once a new parameter set is obtained, it is accepted or rejected. This step is the central point to the Metropolis algorithm. Acceptance or rejection of the new parameter set is determined by the ratio of the posterior probability density function between the previous parameter set and the sampled parameter set (equation 5-5).

[1] the use of statistical adjective is to differentiate between the likelihood term used in GLUE and MCMC approach.

$$r = \frac{\pi(\theta_{new} \mid Y)}{\pi(\theta_{old} \mid Y)} \qquad 5\text{-}5$$

The Metropolis rule is used to accept or reject a new parameter:

If $r > 1$, accept the new parameter set

If $r < 1$, generate a random number $\boldsymbol{u}$ from a uniform distribution [0,1]

If $r > u$, accept the new parameter set

If $r < u$, reject the new parameter set.

There are several special cases of Metropolis algorithm, including Metropolis-Hastings Algorithm (Hastings 1970), Gibbs algorithm (Geman and Geman, 1984), and Metropolis within Gibbs Algorithm (Gelfand and Smith, 1990). For this research, we used Metropolis within Gibbs algorithm to sample and accept or reject the new parameter sets. Metropolis within Gibbs algorithm allows us to move a parameter in its state space (or sample around its previous value), calculate *r*, accept or reject the new parameter, and then move to the next parameter in the set.

To implement the MCMC approach for use with the Mossy Creek HSPF model, a software utility was developed using Microsoft® Visual Basic. The utility supplied HSPF with the new parameter set, and analyzed the model output for the next iteration in the Markov Chain. The software utility also ensured that the new parameter values were always inside the upper and lower bounds of parameter distributions. Each new sample of a parameter was checked against the bounds of its prior distribution, and if the new sample was beyond the bounds, it was sampled again, before going through the acceptance/rejection process. Each parameter vector and the simulation results were stored in a Microsoft® SQL Server database.

The length of a Markov chain is important as it influences the convergence to the posterior parameter distribution. The period before convergence occurs is referred to as the 'burn in period.' Several methods have been proposed to monitor convergence of Markov chain. These include the Geweke test (Geweke, 1992) and the convergence diagnostic proposed by Gelman and Rubin (1992). The Gelman-Rubin diagnostic was used here to monitor convergence for all parameters. The diagnostic is based on generating multiple Markov chains and calculating the mixture of chain variance, within chain variance and Gelman-Rubin statistic. The Gelman-Rubin statistic is the ratio of the mixture of chain variance and within chain variance multiplied by a correction factor. The Markov chain is considered to be converged when these variances stabilize and the Gelman-Rubin statistic approaches unity. Chain convergence and the assessment procedures are illustrated graphically later when model calibration is discussed.

5.1.5 Mossy Creek Model Calibration and Validation

The Mossy Creek HSPF model was calibrated and validated, using GLUE and MCMC techniques, independently. For the GLUE technique, the hydrologic calibration was conducted for the period of 1 September 1998 to 31 December 1999. The posterior parameter distributions obtained using GLUE were considered to be the calibrated parameter distributions. The posterior parameter distributions were used to validate the hydrologic model for the period – 1 January 2000 to 30 September 2002. To conduct the water quality calibration, GLUE was used to estimate the posterior distributions of water quality parameter for the calibration period – 1 October 1998 to 31 December 2001. Insufficient water quality data prevented a water quality validation. The posterior parameter distributions obtained using GLUE are described in the results section.

For the MCMC technique, each sampled parameter set depends upon the previous parameter set, therefore, it is impractical to perform hydrologic and water quality parameter calibration separately. The Mossy Creek HSPF model was calibrated using four years of concurrent hydrologic and water quality data (1 September 1998 to 30 September 2002). For each iteration, two likelihood functions were calculated, one for hydrology which was based on daily flow volume at the watershed outlet (watershed inches), and one for water quality which was based on observed instantaneous in-stream FC concentration (cfu / 100 ml) and simulated daily average FC concentration. A log transformation was performed on the observed and simulated flow volume and in-stream FC before calculating the likelihood function to normalize the scale of residuals. The posterior distributions obtained following the calibration are described in the result section.

5.1.6 TMDL Pollutant Allocation Scenarios

A TMDL allocation scenario allocates the pollutant loads among different sources and hence suggests the amount of reduction in pollutant loading from each source to meet the applicable water quality standard. The Mossy Creek bacterial TMDL included several allocation scenarios, with two preferred scenarios (table 5-5). To simulate the TMDL allocation scenarios in HSPF, a reduction factor was applied to the pollutant load from each source.

Table 5-5 TMDL pollutant allocation scenarios resulting in no violations of instantaneous criteria for indicator bacteria (Benham et al., 2004).

	Required source-specific fecal coliform load reductions (%)						
TMDL Allocation Scenario	Cattle Direct Deposit	Cropland	Pasture	Loafing Lot	Wildlife Direct Deposit	Straight Pipes	All residential pervious land segments
S1	99	90	98	100	30	100	95
S2	94	95	98	100	0	100	95

Both TMDL allocation scenarios require 100% reduction in FC loadings from illegal straight pipes that discharge waste directly from homes. The major difference between the scenarios is the reduction in cattle direct deposit, wildlife direct deposit and loadings from cropland. Due to low production of FC by wildlife, uncertainty in wildlife direct deposit was not considered in this study. A period of three and a half years (1218 days) that represents a range of hydrological events in Mossy Creek was selected to simulate the in-stream FC concentration under the two allocation scenarios. The hydrology and water quality parameters posterior distributions obtained using GLUE and MCMC techniques were used to conduct the Monte-Carlo simulations for the two allocation scenarios.

5.2 Results and Discussions

5.2.1 GLUE

The hydrologic parameter posterior distributions for the Mossy Creek watershed model were developed using GLUE. Figure 5.4 illustrates the dotty plots and posterior distributions of two hydrology parameters LZSN-pasture (Lower zone nominal soil moisture storage in the pasture land use) and DEEPFER (fraction of infiltrating water lost to deep aquifers). The shape of the posterior distribution of LZSN-Pasture (figure 5.4 a, the solid line) is very different from the shape of its prior probability distribution (figure 5.4 a, the dashed line), whereas the shape of posterior distribution of DEEPFR (figure 5.4 b) is similar to the prior distribution. The difference in prior and posterior distributions for the two parameters implies that the observed data provided greater information about the parameter LZSN-pasture than DEEPFR. In other words, the model is more sensitive to LZSN-pasture than DEEPFR in the Mossy Creek watershed model. Plots similar to those illustrated in figure 5.4 were generated for all the hydrologic parameters and the posterior distributions were calculated (table 5-6).

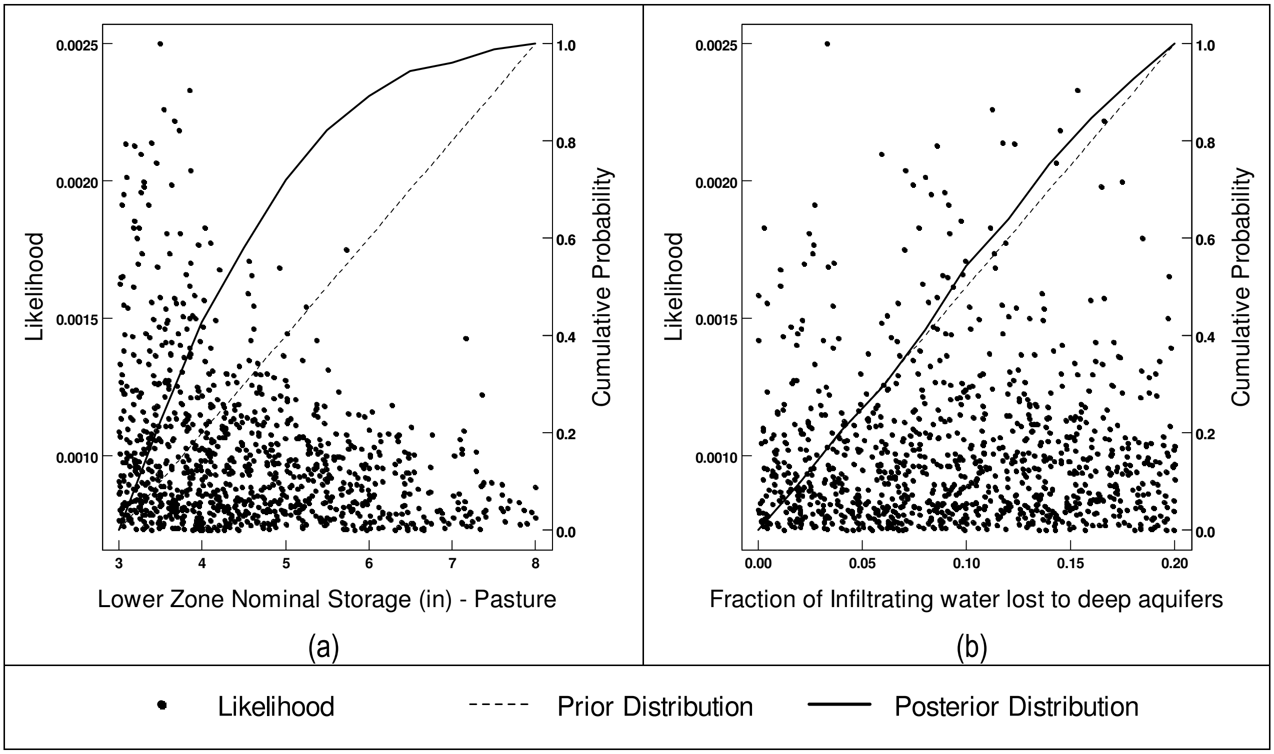


Figure 5.4 Posterior distribution of two hydrologic parameters, (a) LZSN – Pasture, and (b) DEEPFR obtained using GLUE technique

Table 5-6 Posterior distribution of all the hydrology parameters in Mossy Creek watershed model.

Parameter Name	Distribution Limits	Cumulative Distribution¥
LZSN-Forest	(3,8)	{0.14, 0.30, 0.45, 0.55, 0.63, 0.71, 0.78, 0.85, 0.93, 1.00}
LZSN-Cropland	(3,8)	{0.13, 0.26, 0.38, 0.48, 0.56, 0.66, 0.75, 0.83, 0.92, 1.00}
LZSN-Pasture	(3,8)	{0.24, 0.43, 0.61, 0.76, 0.86, 0.93, 0.97, 0.99, 1.00, 1.00}
LZSN-Farmstead	(3,8)	{0.09, 0.18, 0.27, 0.37, 0.46, 0.56, 0.68, 0.79, 0.90, 1.00}
LZSN-LDR	(3,8)	{0.09, 0.21, 0.32, 0.43, 0.54, 0.66, 0.73, 0.83, 0.92, 1.00}
LZSN-HDR	(3,8)	{0.11, 0.21, 0.31, 0.41, 0.49, 0.59, 0.69, 0.78, 0.91, 1.00}
LZSN-Loafing Lot	(3,8)	{0.13, 0.24, 0.35, 0.45, 0.56, 0.65, 0.72, 0.78, 0.88, 1.00}
INFILT-Forest	(0.05, 1.0)	{0.06, 0.38, 0.62, 0.78, 0.88, 0.94, 0.98, 0.99, 1.00, 1.00}
INFILT-Cropland	(0.03, 0.8)	{0.01, 0.15, 0.41, 0.62, 0.77, 0.89, 0.95, 0.99, 1.00, 1.00}
INFILT-Pasture	(0.04, 0.9)	{0.00, 0.03, 0.21, 0.42, 0.67, 0.83, 0.93, 0.98, 1.00, 1.00}
INFILT-Farmstead	(0.03, 0.5)	{0.01, 0.11, 0.38, 0.59, 0.77, 0.89, 0.95, 0.98, 1.00, 1.00}
INFILT-LDR	(0.03, 0.26)	{0.00, 0.02, 0.07, 0.18, 0.37, 0.60, 0.82, 0.93, 0.99, 1.00}
INFILT-HDR	(0.01, 0.1)	{0.27, 0.52, 0.70, 0.83, 0.90, 0.95, 0.98, 0.99, 1.00, 1.00}
INFILT-Loafing Lot	(0.15, 0.23)	{0.11, 0.21, 0.33, 0.43, 0.53, 0.61, 0.70, 0.78, 0.90, 1.00}
DEEPFR (all land uses)	(0.0, 0.2)	{0.11, 0.21, 0.33, 0.43, 0.52, 0.63, 0.73, 0.82, 0.91, 1.00}
BASETP (all land uses)	(0.00, 0.05)	{0.32, 0.57, 0.74, 0.86, 0.91, 0.95, 0.98, 0.99, 0.99, 1.00}
AGWETP (all land uses)	(0.00, 0.05)	{0.17, 0.34, 0.48, 0.63, 0.73, 0.82, 0.90, 0.95, 0.98, 1.00}
INTFW (all land uses)	(1.0, 3.0)	{0.11, 0.22, 0.31, 0.40, 0.51, 0.63, 0.74, 0.81, 0.90, 1.00}
IRC (all land uses)	(0.5, 0.8)	{0.00, 0.01, 0.06, 0.14, 0.30, 0.51, 0.77, 0.93, 0.99, 1.00}
MON INTERCEP-Forest-Jan	(0.03, 0.075)	{0.13, 0.31, 0.46, 0.62, 0.75, 0.82, 0.88, 0.92, 0.96, 1.00}
UZSN-Forest-January	(0.1, 0.3)	{0.03, 0.08, 0.13, 0.23, 0.34, 0.45, 0.56, 0.70, 0.83, 1.00}
LZETP-Forest-January	(0.07, 0.2)	{0.03, 0.10, 0.20, 0.29, 0.39, 0.53, 0.66, 0.79, 0.92, 1.00}

¥The cumulative distributions are the values at 10 equal intervals between the distribution limits.

The posterior distributions of hydrologic parameters were used to conduct Monte Carlo simulations for the validation period – 1 January 1999 to 31 December 2002. The total daily flow volume resulting from each HSPF Monte Carlo iteration was used to calculate the HSPEXP (HSPF Expert System) statistics (Lumb et al., 1994). Several quantiles were calculated using the HSPEXP statistics to validate the posterior distributions (table 5-7).

Table 5-7 Quantiles of the HSPEXP (expert system for HSPF) statistics for the validation period when Monte Carlo simulations were conducted with "posterior" and "prior" distributions

Calibration Sufficiency Statistics	Default criteria (percent error)	Quantiles for validation period when "prior" distributions were used to conduct Monte Carlo simulations		Quantiles for validation period when "posterior" distributions were used to conduct Monte Carlo simulations	
		2.5	97.5	2.5	97.5
Total Volume	±10	-13.1	13.8	-10.3	8.2
50% Lowest Flows	±10	-8.8	23.8	*-4.7*	*16.8*
10% Highest Flows	±15	-16.5	19.0	*-16.9*	*1.0*
Storm Peaks	±20	-15.4	32.0	-16.1	1.7
Seasonal Volume Error	±30	0.9	16.3	0.2	11.1
Summer storm volume error	±50	-19.1	15.3	-15.7	7.0

Table 5-7 illustrates that overall, the hydrologic parameter posterior distributions were acceptable. Using the hydrology parameter posterior distributions produced calibration sufficiency statistics within HSPEXP criteria bounds 95% of the time for all except two statistics, 50% lowest flows and 10% highest flows. These results imply that including calibration sufficiency statistics in addition to the residuals between observed and simulated daily flow volume in the likelihood evaluation could perhaps improve the calibration.

Following the hydrologic calibration and validation, GLUE technique was used to estimate the posterior distributions of water quality parameters. For water quality parameters, the likelihood function was calculated using equation 5-1, after log-transformation of observed instantaneous FC concentration and simulated daily average FC concentration for the days for which observed data were available. Figure 5.5 shows the example of the dotty plots and posterior distributions generated using GLUE for two water quality parameters, FSTDEC (first order decay rate of bacteria) and ACQOP-Pasture (rate of FC accumulation on pasture land use) using log-transformed data. The shape of posterior distributions of FSTDEC and ACCUM-pasture are different from the shape of their prior distributions indicating the effect of observed water quality data on the posterior distributions of these parameters. In other words, these two parameters are sensitive in Mossy Creek watershed model. The posterior distribution of each water quality parameter was developed similarly (table 5-8).

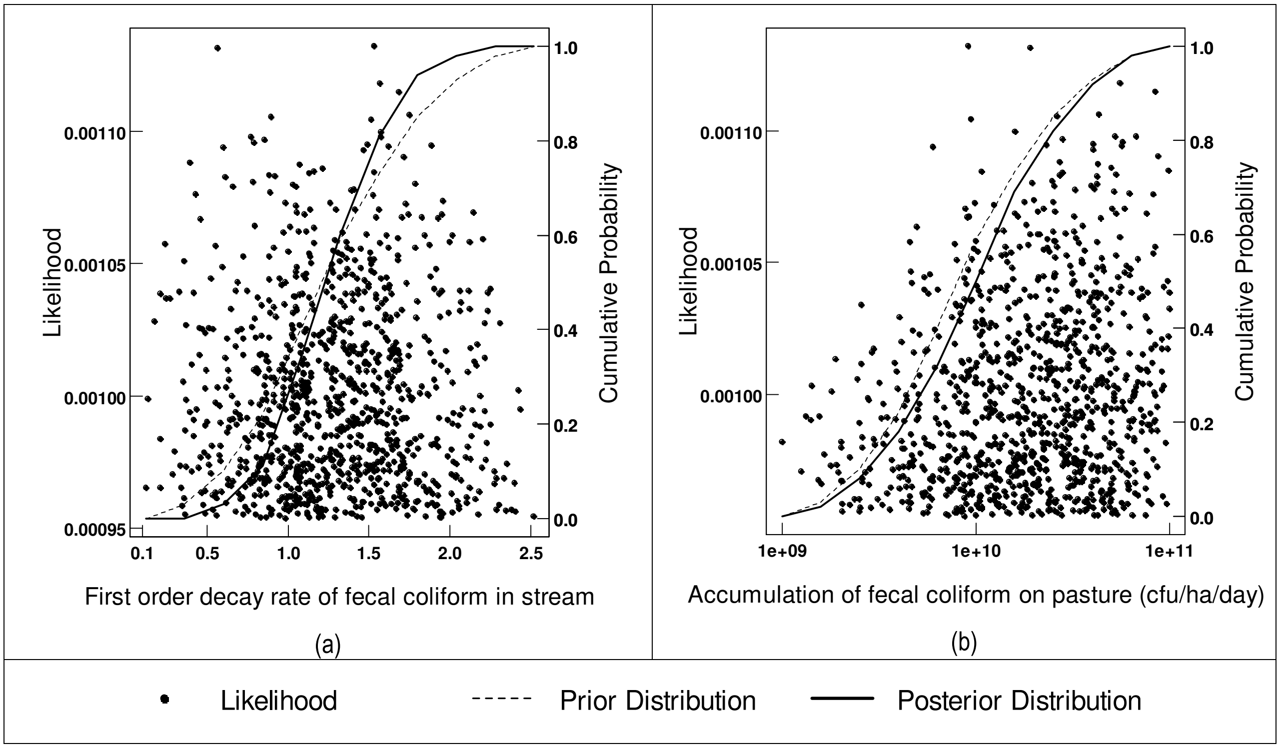


Figure 5.5 Posterior distributions of two water quality parameters (a) FSTDEC, and (b) Accumulation of fecal coliform in pasture; obtained using GLUE technique.

Table 5-8 Posterior water quality parameters distributions obtained using GLUE.

Parameter Name	Distribution Limits	Cumulative Distribution¥
SQOLIM-FACTOR	(2.5, 11.5)	{0.04, 0.10, 0.20, 0.29, 0.40, 0.51, 0.61, 0.72, 0.84, 1.00}
WSQOP-Forest	(0.5, 2.4)	{0.11, 0.19, 0.29, 0.40, 0.50, 0.61, 0.71, 0.82, 0.90, 1.00}
WSQOP-Cropland	(0.5, 2.4)	{0.10, 0.22, 0.34, 0.43, 0.56, 0.66, 0.75, 0.82, 0.91, 1.00}
ACCUM-Pasture	(1E9, 1E11)	{0.00, 0.01, 0.03, 0.10, 0.28, 0.53, 0.74, 0.91, 0.99, 1.00}
WSQOP-Pasture	(0.5, 2.4)	{0.16, 0.30, 0.42, 0.51, 0.61, 0.71, 0.78, 0.87, 0.94, 1.00}
WSQOP-Farmstead, LDR	(0.5, 2.4)	{0.13, 0.24, 0.34, 0.45, 0.55, 0.66, 0.73, 0.83, 0.91, 1.00}
ACCUM-Loafing Lot	(1.2E11, 1.2E13)	{0.00, 0.02, 0.07, 0.18, 0.37, 0.67, 0.84, 0.95, 0.99, 1.00}
WSQOP-Loafing Lot	(0.05, 2.4)	{0.16, 0.28, 0.39, 0.50, 0.61, 0.68, 0.77, 0.85, 0.91, 1.00}
ACCUM-Cropland-Jan	(2E6, 2E8)	{0.00, 0.03, 0.07, 0.28, 0.58, 0.76, 0.91, 0.96, 1.00, 1.00}
FSTDEC	(0.12, 2.52)	{0.00, 0.03, 0.11, 0.34, 0.60, 0.81, 0.94, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES1	(0.03, 0.5)	{0.00, 0.04, 0.16, 0.39, 0.62, 0.81, 0.93, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES2	(0.1, 10)	{0.00, 0.04, 0.16, 0.34, 0.62, 0.78, 0.92, 0.98, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES3	(0.1, 10)	{0.00, 0.04, 0.16, 0.39, 0.66, 0.85, 0.95, 0.99, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES4	(0.1, 10)	{0.00, 0.04, 0.14, 0.30, 0.56, 0.78, 0.91, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES5	(0.1, 10)	{0.00, 0.05, 0.17, 0.40, 0.67, 0.85, 0.94, 0.99, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES6	(0.1, 10)	{0.00, 0.04, 0.19, 0.45, 0.81, 0.99, 1.00, 1.00, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES7	(0.1, 10)	{0.00, 0.04, 0.12, 0.33, 0.56, 0.75, 0.89, 0.97, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES8	(0.1, 10)	{0.00, 0.03, 0.12, 0.34, 0.66, 0.91, 1.00, 1.00, 1.0, 1.00}

¥The cumulative distributions are values at 10 intervals between the distribution limits.

The water quality posterior parameter distributions were not validated due to insufficient observed data. The Mossy Creek watershed model input parameter posterior distributions that were obtained using the GLUE techniques were used to conduct Monte Carlo simulations for two of the TMDL pollutant allocation scenarios suggested in the Mossy Creek TMDL (table 5-5). The simulations were conducted for a period of approximately three and a half years (1218 days) chosen to represent a range of hydrological events.

The daily average simulated in-stream FC concentration time series from each Monte Carlo simulation were used to compute an average, 2.5%, 10%, 90% and 97.5% quantiles for each day. The quantiles and the average time series were plotted for the two allocation scenarios (figure 5.6). The percent of water quality criterion violations for each time series was calculated by dividing the number of daily violations by the number of days in the simulation period (table 5-9).

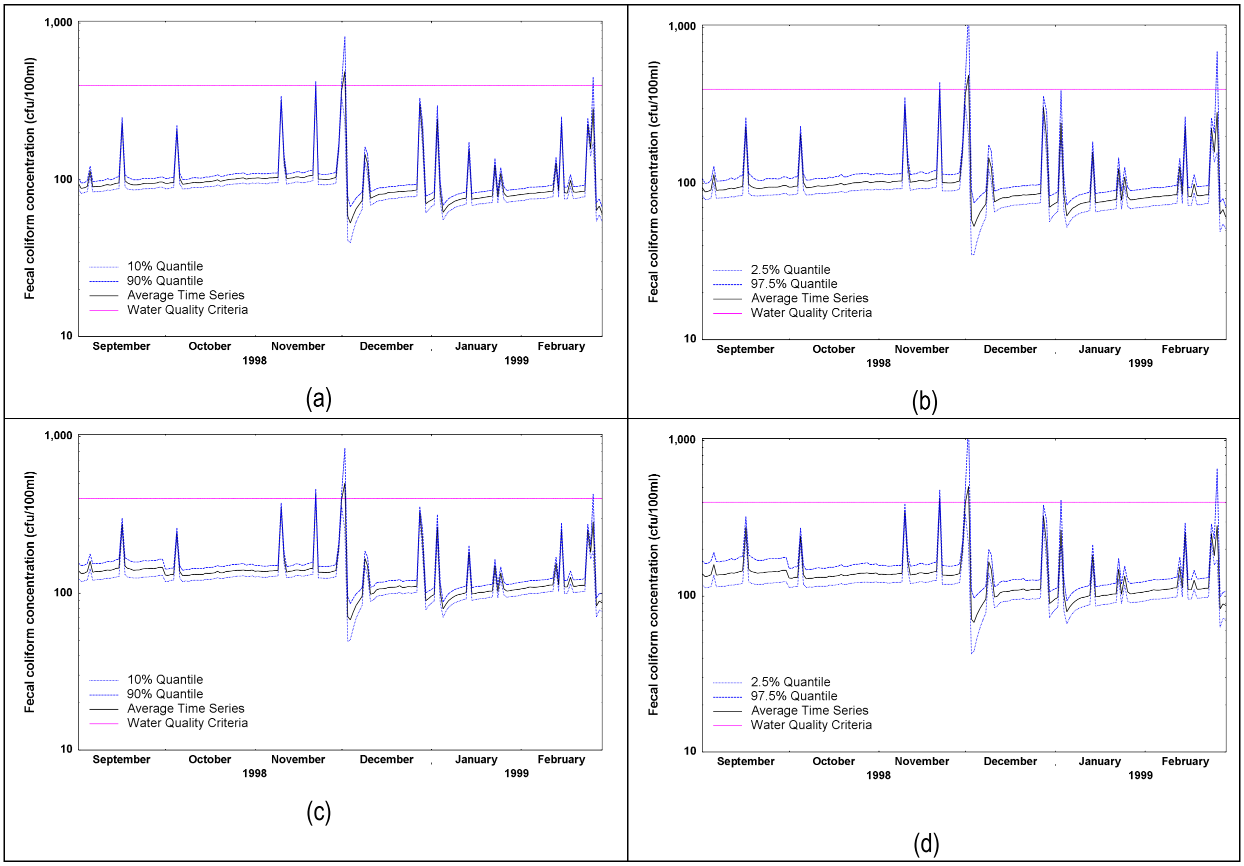


Figure 5.6 For TMDL allocation scenario S1, 80% probability interval (a) and 95% probability interval (b); TMDL allocation scenario S2, 80% probability interval (c) and 95% probability interval (d). Representative plots showing first six months of simulation period.

Table 5-9 Percent of water quality criterion violations by the average time series and the probability intervals for two TMDL allocation scenarios, when GLUE was used for estimating posterior parameter distributions.

TMDL Allocation Scenario	Average time series violations (%)	80% probability interval violations (%)	95% probability interval violations
S1	1.2	(0.2, 1.6)†	(0.0, 2.6)
S2	1.3	(0.4, 1.8)	(0.1, 3.1)

† Numbers in parentheses show the percent of violation incidences over a period of 1218 days by the respective time series for the upper and lower bounds of the probability interval.

In figure 5.6, the 95% probability interval – frames (b) and (d) are wider than the 80% probability intervals –frames (a) and (c) for both the allocation scenarios. These graphs illustrate that uncertainty increases as we seek greater confidence in predicted results. The S2 scenario exhibited greater uncertainty in FC concentration criterion violations than S1. The S2 scenario had greater input of FC from direct deposit by cattle when compared to S1, but less input of FC through cropland, which illustrates that the direct deposit of FC is a greater source of uncertainty than cropland in Mossy Creek watershed as both the pollutant sources had similar input uncertainty. S2 also had greater input of FC through wildlife, but wildlife was not considered an uncertain source, and a deterministic time series was used to model the wildlife FC input in Mossy Creek.

5.2.2 MCMC

The MCMC technique was used to obtain posterior distributions of hydrology and water quality parameters for the Mossy Creek HSPF model. The posterior parameter distributions were obtained to estimate predictive uncertainty in in-stream FC concentration. One-hundred thousand iterations of model simulations were performed to obtain the Markov chains for hydrology and water quality parameters. Examples of Markov Chain of two hydrology parameters LZSN-Cropland (lower zone nominal soil moisture storage in the cropland land use) and DEEPFR (fraction of infiltrating water lost to deep aquifers); and two water quality parameters ACQOP-Pasture (rate of FC accumulation on pasture land use), and FSTDEC (first order decay rate of bacteria) are illustrated in figure 5.7 and figure 5.8, respectively.

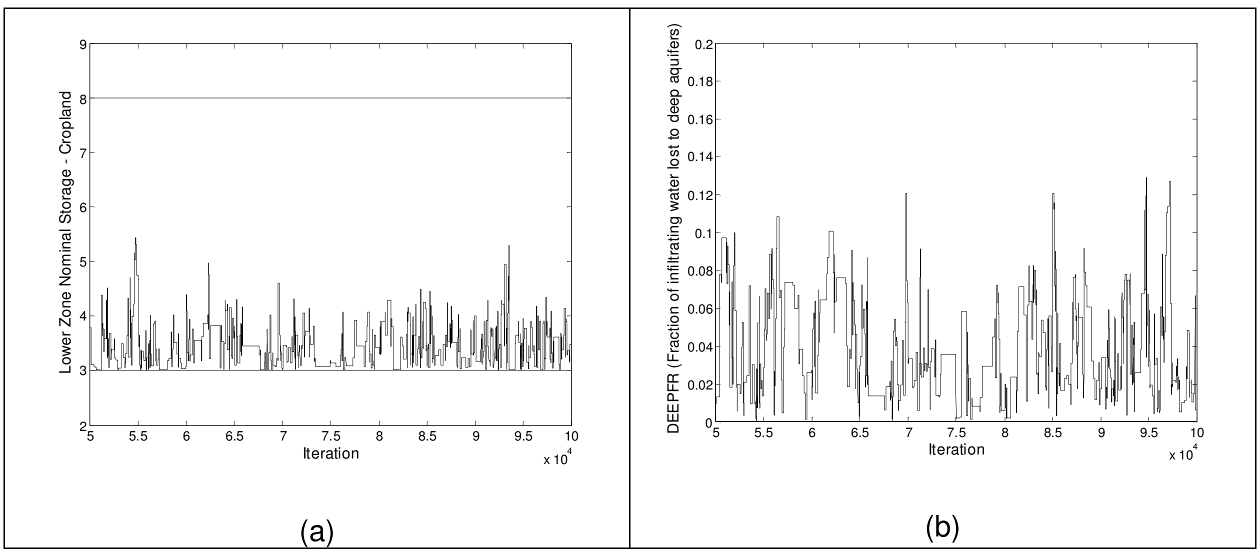
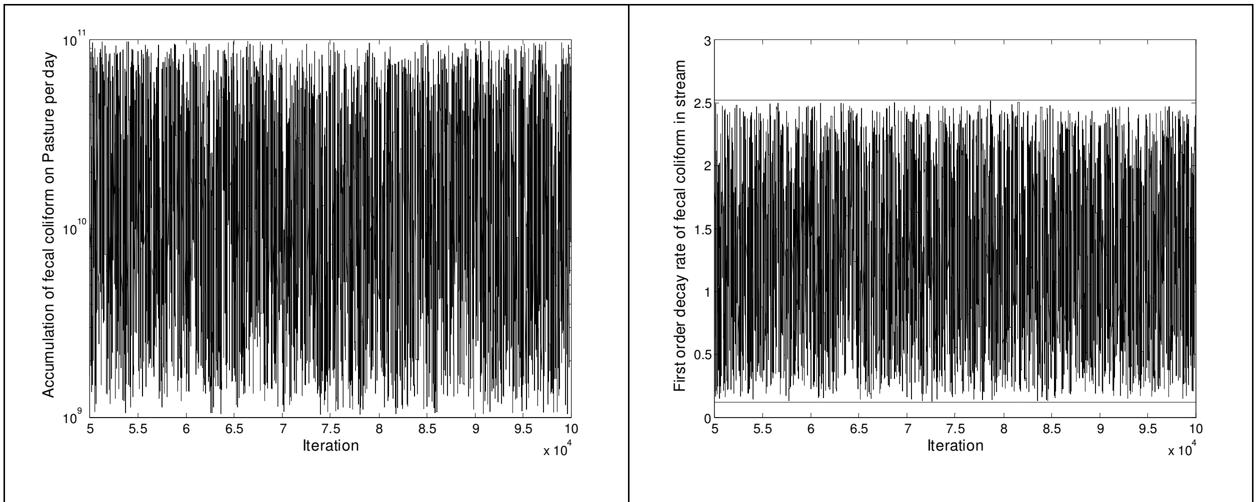


Figure 5.7 Markov Chains of two hydrology parameters. The chains illustrated here are for last 50,000 iterations out of 100,000 iterations.



(a) **(b)**

Figure 5.8 Markov chains of two water quality parameters. The chains illustrated here are for last 50,000 iterations out of 100,000 iterations.

As evident from figure 5.7and figure 5.8, the values of hydrology parameters changed slower (figure 5.7) when compared to the water quality parameters (figure 5.8). In other words, Markov chains for hydrology parameters moved slower than the Markov chains for water quality parameters. In addition, the Markov chains for hydrology parameters did not traverse the whole parameter space, while the water quality parameters did. Figure 5.7 and figure 5.8 illustrate that the observed data had greater effect on LZSN-Cropland and DEEPFR than ACQOP-Pasture and FSTDEC, which was expected, as the observed data for flow was available for the complete simulation period, whereas observed FC concentration was available for only 90 days out of the entire simulation period of over than three years.

Three or more Markov chains are required for each parameter to estimate the convergence of Markov chain (Gelman and Rubin, 1992) using the Gelman-Rubin statistic. To obtain three Markov chains, three different instances of MCMC were conducted simultaneously on three different computers as parallel processes. The starting points of the three chains for each parameter were selected randomly. To verify the convergence of the parameters we conducted Gelman-Rubin tests (figure 5.9 and figure 5.10).

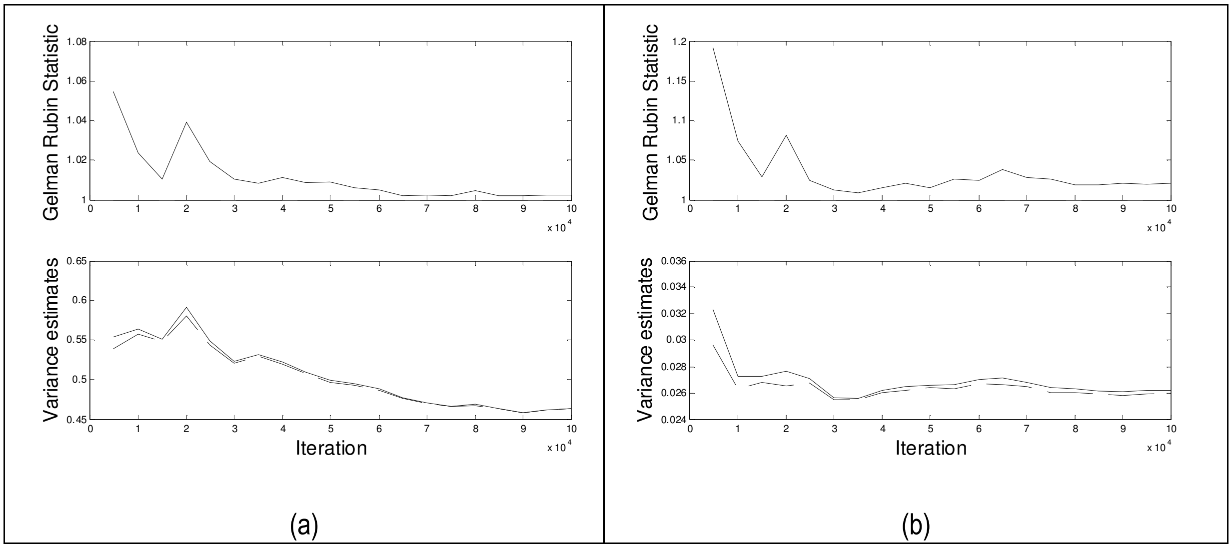


Figure 5.9 Gelman-Rubin statistic and variance estimates of (a) Lower zone nominal soil moisture (LZSN-Cropland), and (b) fraction of water lost to deep aquifers (DEEPFR).

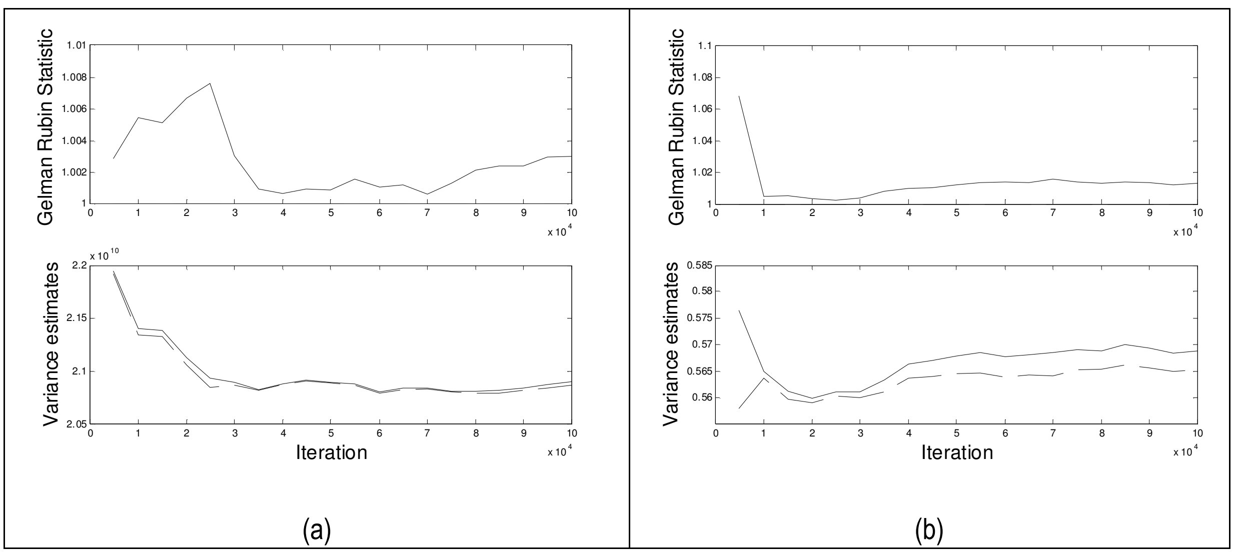


Figure 5.10 Gelman-Rubin statistic and variance estimates of (a) fecal coliform on pasture per day (ACQOP), and (b) first order decay rate of fecal coliform (FSTDEC).

A Markov chain is assumed to be converged when the Gelman Rubin statistic reaches unity, and within chain variance (solid line in figure 5.9) and among chain variance (dashed) line in figure 5.9) estimates match. The parameter LZSN-Cropland illustrated in figure 5.9(a) converged at about 70,000 iterations. Although, the Gelman-Rubin statistics started reaching the value of 1 at about 30,000 iterations, the variance estimates started stabilizing at about 70,000 iterations. The parameter DEEPFR (figure 5.9(b)) converged at about 30,000 iterations. The Gelman-Rubin statistics for ACQOP-Pasture (figure 5.10(a)) started increasing after 70,000 iterations, variances started stabilizing long before that. Moreover, this small increase of Gelman-

Rubin statistics can be neglected as the order of magnitude of variances of ACQOP-Pasture was about 10. Similarly, all the individual parameters were analyzed and all the parameters appeared to converge at about 70,000 iterations or earlier.

The values of the parameters in Markov chain for the first 70,000 iterations were rejected as the ‘burn-in’ period and remaining parameter values were used to approximate the hydrology and water quality parameters posterior distributions. Figure 5.11 illustrates examples of the prior and posterior distributions for LZSN-Cropland, DEEPFR, ACQOP-Pasture, and FSTDEC.

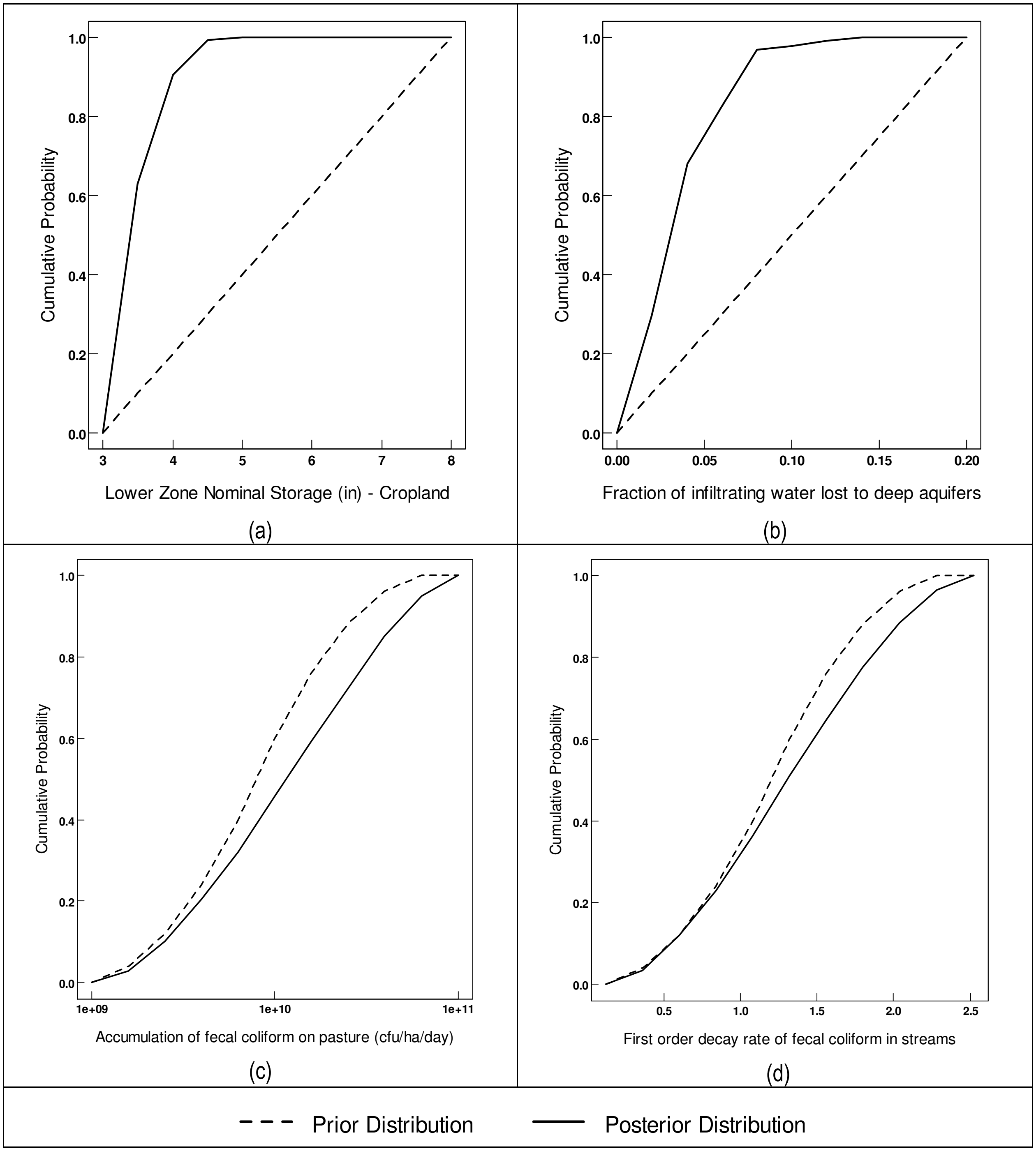


Figure 5.11 Posterior distribution of two hydrology and two water quality parameters obtained using MCMC

Instead of fitting any predefined distributions to the posterior distributions, the empirical posterior distributions were used for uncertainty analysis. In general, there was a smaller effect of observed data on the water quality parameters compared to the hydrological parameters, as indicated by the differences in shapes of prior and posterior distributions. A similar trend was noted in GLUE analysis. This trend was expected, as there was a greater amount of observed hydrology data than water quality data. The posterior distribution of all hydrology and water quality parameters is listed in table 5-10, and table 5-11, respectively.

Table 5-10 Posterior distributions of hydrology parameters obtained after the application of MCMC technique.

Parameter Name	Distribution Limits	Cumulative Distribution¥
LZSN-Forest	(3,8)	{0.74, 0.98, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00}
LZSN-Cropland	(3,8)	{0.63, 0.90, 0.99, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00}
LZSN-Pasture	(3,8)	{0.97, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00}
LZSN-Farmstead	(3,8)	{0.18, 0.37, 0.60, 0.67, 0.72, 0.80, 0.87, 0.95, 0.99, 1.00}
LZSN-LDR	(3,8)	{0.18, 0.47, 0.65, 0.73, 0.79, 0.88, 0.94, 0.98, 0.99, 1.00}
LZSN-HDR	(3,8)	{0.01, 0.07, 0.17, 0.27, 0.38, 0.55, 0.69, 0.77, 0.91, 1.00}
LZSN-Loafing Lot	(3,8)	{0.15, 0.31, 0.37, 0.56, 0.64, 0.72, 0.81, 0.89, 0.92, 1.00}
INFILT-Forest	(0.05, 1.0)	{0.20, 0.42, 0.55, 0.64, 0.72, 0.80, 0.91, 0.97, 0.99, 1.00}
INFILT-Cropland	(0.03, 0.8)	{0.00, 0.08, 0.20, 0.32, 0.44, 0.56, 0.71, 0.82, 0.95, 1.00}
INFILT-Pasture	(0.04, 0.9)	{0.00, 0.05, 0.26, 0.62, 0.81, 0.88, 0.90, 0.98, 1.00, 1.00}
INFILT-Farmstead	(0.03, 0.5)	{0.05, 0.09, 0.16, 0.24, 0.41, 0.54, 0.69, 0.82, 0.93, 1.00}
INFILT-LDR	(0.03, 0.26)	{0.05, 0.18, 0.35, 0.52, 0.63, 0.75, 0.83, 0.92, 0.97, 1.00}
INFILT-HDRs	(0.01, 0.1)	{0.14, 0.25, 0.44, 0.56, 0.64, 0.72, 0.81, 0.89, 0.92, 1.00}
INFILT-Loafing Lot	(0.15, 0.23)	{0.05, 0.18, 0.30, 0.45, 0.54, 0.64, 0.69, 0.79, 0.90, 1.00}
DEEPFR (all land uses)	(0.0, 0.2)	{0.30, 0.68, 0.83, 0.97, 0.98, 0.99, 1.00, 1.00, 1.00, 1.00}
BASETP (all land uses)	(0.00, 0.05)	{0.90, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00}
AGWETP (all land uses)	(0.00, 0.05)	{0.00, 0.00, 0.00, 0.02, 0.07, 0.18, 0.40, 0.73, 0.92, 1.00}
INTFW (all land uses)	(1.0, 3.0)	{0.12, 0.25, 0.36, 0.53, 0.67, 0.72, 0.80, 0.89, 0.95, 1.00}
IRC (all land uses)	(0.5, 0.8)	{0.02, 0.07, 0.16, 0.27, 0.42, 0.55, 0.66, 0.81, 0.96, 1.00}
MON INTERCEP-Forest-January	(0.03, 0.075)	{0.06, 0.18, 0.39, 0.58, 0.71, 0.85, 0.94, 0.97, 0.99,1.00}
UZSN-Forest-January	(0.1, 0.3)	{0.00, 0.00, 0.00, 0.01, 0.03, 0.09, 0.20. 0.36, 0.62, 1.00}
LZETP-Forest-January	(0.07, 0.2)	{0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00. 0.02, 1.00}

¥The cumulative distributions are values at 10 intervals between the distribution limits.

Table 5-11 Posterior distributions of water quality parameters obtained after the application of MCMC technique.

Parameter Name	Distribution Limits	Cumulative Distribution¥
SQOLIM-FACTOR	(2.5, 11.5)	{0.04, 0.10, 0.18, 0.28, 0.39, 0.51, 0.63, 0.77, 0.90, 1.00}
WSQOP-Forest	(0.5, 2.4)	{0.08, 0.19, 0.30, 0.42, 0.53, 0.64, 0.75, 0.85, 0.94, 1.00}
WSQOP-Cropland	(0.5, 2.4)	{0.07, 0.18, 0.28, 0.39, 0.50, 0.61, 0.72, 0.83, 0.93, 1.00}
ACCUM-Pasture	(1E9, 1E11)	{0.03, 0.10, 0.20, 0.32, 0.46, 0.59, 0.72, 0.85, 0.95, 1.00}
WSQOP-Pasture	(0.5, 2.4)	{0.09, 0.20, 0.31, 0.42, 0.53, 0.64, 0.75, 0.85, 0.94, 1.00}
WSQOP-Farmstead, LDR	(0.5, 2.4)	{0.09, 0.20, 0.31, 0.43, 0.54, 0.65, 0.76, 0.86, 0.94, 1.00}
ACCUM_Loafing Lot	(1.2E11, 1.2E13)	{0.04, 0.11, 0.21, 0.32, 0.44, 0.56, 0.68, 0.81, 0.94, 1.00}
WSQOP-Loafing Lot	(0.05, 2.4)	{0.08, 0.19, 0.31, 0.42, 0.53, 0.64, 0.75, 0.84, 0.93, 1.00}
ACCUM-Cropland-Jan	(2E6, 2E8)	{0.10, 0.22, 0.31, 0.49, 0.61, 0.69, 0.80, 0.87, 0.98, 1.00}
FSTDEC	(0.12, 2.52)	{0.04, 0.12, 0.23, 0.36, 0.51, 0.65, 0.77, 0.88, 0.96, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES1	(0.03, 0.5)	{0.03, 0.11, 0.22, 0.35, 0.49, 0.62, 0.75, 0.88, 0.96, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES2	(0.1, 10)	{0.05, 0.13, 0.25, 0.38, 0.53, 0.67, 0.80, 0.91, 0.98, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES3	(0.1, 10)	{0.04, 0.15, 0.29, 0.43, 0.59, 0.73, 0.86, 0.95, 0.99, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES4	(0.1, 10)	{0.03, 0.13, 0.25, 0.39, 0.53, 0.67, 0.79, 0.89, 0.97, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES5	(0.1, 10)	{0.04, 0.14, 0.27, 0.41, 0.57, 0.71, 0.84, 0.94, 0.99, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES6	(0.1, 10)	{0.06, 0.20, 0.39, 0.60, 0.80, 0.94, 0.99, 1.00, 1.00, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES7	(0.1, 10)	{0.04, 0.13, 0.25, 0.39, 0.54, 0.68, 0.81, 0.91, 0.98, 1.00}
Direct Deposit Cattle Time Series, Multiplication Factor for RCHRES8	(0.1, 10)	{0.02, 0.09, 0.20, 0.37, 0.58, 0.81, 0.95, 1.00, 1.00, 1.00}

¥The cumulative distributions are values at 10 intervals between the distribution limits.

The input parameter posterior distributions were used to conduct Monte Carlo simulations of the model for the two different TMDL pollutant allocation scenarios suggested in Mossy Creek TMDL. The daily average in-stream FC concentration resulting from all the simulations were used to used to compute an average, 2.5%, 10%, 90% and 97.5% quantiles for each day. These quantiles and the average time series were plotted for the two allocation scenarios (figure 5.12). The percent of days the water quality criterion was violated by each time series was also calculated (table 5-12).

Table 5-12 Percent of water quality criterion violations by the average time series and the probability intervals for two TMDL allocation scenarios, when MCMC was used for estimating posterior parameter distributions.

TMDL Allocation Scenario	Average time series violations (%)	80% probability interval violations (%)	95% probability interval violations (%)
S1	0.9	(0.0, 1.5)	(0.0, 2.8)
S2	1.1	(0.0, 1.6)	(0.0, 3.5)

† Numbers in parentheses show the percent of violation incidences over a period of 1218 days by the respective time series for the upper and lower bounds of the probability interval.

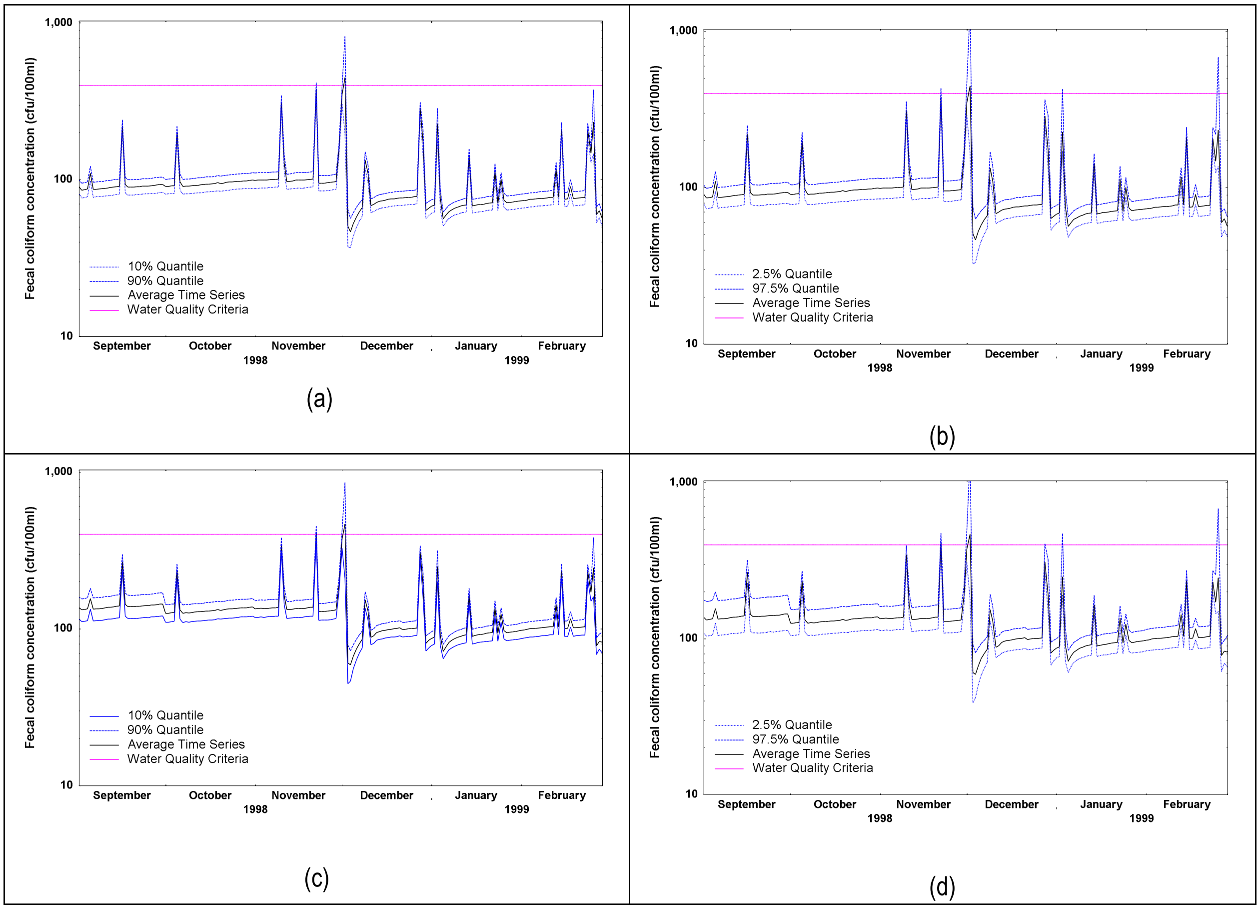


Figure 5.12 For TMDL allocation scenario S1, (a) 80% probability interval (b) 95% probability interval; for TMDL allocation scenario S2 (c) 80% probability interval (d) 95% probability interval. Representative plots show results for first six months of simulation period. The posterior distributions were obtained using MCMC.

The width of 95% probability interval is greater than 80% probability interval for the two allocation scenarios (figure 5-10), indicating greater uncertainty, similar to the results from the GLUE analysis. MCMC also predicted greater uncertainty in scenario S2 compared to S1. The number of violations of instantaneous FC criteria that occurred for each allocation scenario by the daily average time series was obtained for a period of 1218 days.

The percent of violation incidences of the instantaneous FC concentration criteria by the daily average FC concentration time series was very similar for the two TMDL allocation scenarios. The percentage of violation incidences for the 80 and 95% probability intervals illustrate that although the uncertainty increases for allocation scenario S2 compared to S1, this increase is minimal.

5.3 Summary and Conclusions

The objective of this research was to compare the predictive uncertainty in in-stream FC criterion violations using Generalized Likelihood Uncertainty Estimation (GLUE) and Markov Chain Monte Carlo (MCMC) with an HSPF based water quality model. The study used the data collected and modeling files developed for the Mossy Creek, VA bacterial impairment TMDL (Benham et al., 2004). Both techniques use Bayesian equation to develop posterior parameter distributions based on prior parameter distributions, observed data and model response. The application of GLUE does not require a modeler to make any implicit assumptions about the response variable(s). The GLUE likelihood value is calculated using a model goodness of fit measure, while MCMC expects the errors to be normally distributed and calculates a statistical likelihood function for the application of Bayesian equation.

For both TMDL allocation scenarios, the average FC criterion violation was about 1% for both techniques. Both the 80 and 95% probability intervals were similar for both the techniques, with the 95% probability interval reporting less than 4% violations for either scenario. The S2 scenario produced more instantaneous FC criterion violations and greater uncertainty than scenario S1 that had lower input of FC direct deposit in streams, but greater input of FC through cropland runoff. The input uncertainty in cropland direct runoff was similar to FC direct deposit in streams, therefore, the results indicated that the FC direct deposit was a greater source of uncertainty compared to FC runoff from croplands. Stakeholders and decision makes can use this information about uncertainty in selecting among the pollutant allocation scenarios or addressing the pollutant source that is a greater source of uncertainty in model prediction than others in the case of adaptive TMDL implementation.

The research demonstrated that GLUE and MCMC techniques are effective tools for estimating uncertainty in FC water quality concentration criterion violations. An important feature of the two techniques was to predict the posterior distribution of input parameters using prior distribution and observed data. These posterior distributions of model parameters can be further improved as more observed data becomes available. For Mossy Creek watershed, the two techniques provided similar results in uncertainty estimates.

The likelihood formulation is an important step in the application of GLUE, and it provides the modeler flexibility in selecting the model responses that are important for the current application. The GLUE technique, however, does not require a statistical likelihood function, and it is often criticized for that. GLUE also requires the modeler to define the model parameter acceptance/rejection criteria, which may affect the posterior distributions. On the other hand, MCMC has a robust formulation that utilizes a statistical likelihood function and does not require modeler input in rejecting the parameter sets that do not perform well. MCMC, however, requires user input in selecting a variance scaling factor that ensures proper chain movement, and in selecting the burn-in period of Markov chain. MCMC also is more resource intensive than GLUE. A single MCMC run that consisted of 100,000 HSPF iterations took about one week to run on a Dell Precision 670 workstation with Intel® XEON™ 3.39 GHz processor and 2.00 GB RAM. Three similar runs were conducted on three computers, as at least three Markov chains are required to determine convergence. A single GLUE run that consisted of 10,000 HSPF iterations, however, took just less than 1 day. It is important to note here that Mossy Creek was a relatively simple model with a single meteorological station and eight reaches with about four years of simulation period. The computational time for each HSPF run would increase for a more complex model and/or longer simulation period.

The research illustrates that, as per computational requirements and the predictive uncertainty, GLUE is a more practical technique for uncertainty estimation for a water quality model developed with HSPF, compared to MCMC. However, with the increase in computational power and reduced run time for HSPF software, these advantages of GLUE over MCMC may blur in future. Further research in sampling algorithm in MCMC application with HSPF may be able to decrease the computing cost. Further research is also required in analyzing the effects of using different likelihood functions and different parameter acceptance/rejection criteria when using GLUE for an application like the one reported here.

References:

Balin, D. 2004. *Hydrological Behaviour through Experimental and Modelling Approaches. Application to the Haute-Mentue Catchment.* School of Acrhitecture, Civil and Environmental Engineering. Lausanne, Switzerland. Ecole Polytechnique Fédérale de Lausanne.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442

Benham, B.L., K. Branna, K. Christophel, T. Dillaha, L. Henry, S. Mostaghimi, R. Wagner, J. Wynn, G. Yagow, and R. Zeckoski. 2004. *Total maximum daily load development for Mossy Creek and Long Glade Run: Bacteria and general standard (Benthic) impairments.* Richmond, Va.: Virginia Department of Environmental Quality. http://www.deq.state.va.us/tmdl/homepage.html. Assessed Jan. 25, 2005.

Benham, B.L., K.M. Brannan, T.A. Dillaha, S.M. Mostaghimi, J. Wynn, G. Yagow, R.W. Zeckoski. 2003. *Total Maximum Daily Load Development for Linville Creek: Bacteria and General Standard (Benthic) Impairments.* Richmond, VA.: Virginia Department of Environmental Quality. http://www.deq.virginia.gov/tmdl/apptmdls/shenrvr/linville.pdf. Assessed Jan 30, 2006.

Benham, B.L., K. Brannan, T. Dillaha, S. Mostaghimi, G. Yagow. 2002. TMDLs (Total Maximum Daily Loads) – Terms and Definitions. *Virginia Cooperative Extension.* Pub No. 442-550.

Beven, K., and J. Freer. 2001. Equifinality, Data Assimilation, and Uncertainty Estimation in mechanistic modelling of Complex Environmental Systems using GLUE Methodology. *Journal of Hydology* 249(1-4): 11-29.

Beven, K.. and A. Binley. 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr. T.H. Jobes, and A.S. Donigian, Jr. 2005. *HSPF Version 12.2 User's Manual.* AQUA TERRA Consultants. Mountain View, CA.

Gelfand, A.E., and A.F.M. Smith. 1990. Sampling based approaches to calculating marginal densities. *Journal of American Statistical Association.* 85(410): 398-409.

Gelman, A., J.B. Carlin, H.S. Stern, and R.D.B. 2000. *Bayesian Data Analysis.* Boca Raton, London, New York, Washington D.C., Chapman&Hall/CRC.

Gelman A., and D. Rubin. 1992. Inference from Iterative Simulation using Multiple Sequences. *Statistical Science.*7: 457-511.

Geman, S. and D. Geman. 1984. Stochastic Relaxation, Gibbs' distribution and Bayesian restoration of images. *IEEE trans.* PAMI 6: 721-741.

Geweke, J. 1992. Evaluating the accuracy of sampling – based approaches to calculating posterior moments (with discussion). *Bayesian Statistics 4.* A.F.M. Smith. Oxford Clarendon Press: 169-193.

Hastings, W.K. 1970. Monte Carlo methods using Markov chains and their applications. *Biometrika.*57: 97-109.

Kuczera, G., and E. Parent. 1998. Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: the Metropolis Algorithm. *Journal of Hydrology. 211:* 69-85.

Lumb, A. M., R. B. McCammon, and John L. Kittle, Jr. 1994. *Users Manual for an Expert Systems (HSPEXP) for Calibration of the Hydrological Simulation Program - FORTRAN.* Reston, Virginia, U.S. Geological Survey, Water-Resources Investigations Report.

Marshall, L. D. Nott, and A. Sharma. 2005. Hydrological Model Selection: A Bayesian alternative. *Water Resources Research.* 41. doi:10.1029/2004WR003719, 2005.

Makowski, D., D. Wallach, and M. Tremblay. 2002. Using a Bayesian approach to parameter estimation; Comparison of GLUE and MCMC methods. *Agronomie*. 22: 191-203.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chemical Phys.* 21(6): 1087-1092.

Robert, C.P., and G. Casella. 1999. *Monte Carlo Statistical Methods.* Springer-Verlag, New York, Inc. New York, NY.

Stow, C.A., K.H. Reckhow, S.S. Qian, E.C. Lamon, G.B. Arhonditsis, M.E. Borsuk, and D. Seo. 2007. Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. *J. Amer. Water Resources Ass.* 43(6): 1499-1507

Suter, G.W. II, L.W. Barthouse and R.V. O'Neill. 1987. Treatment of risk in environmental impact assessment. *Environmental Mgmt.* 11(3): 295-303.

USEPA. 2001. *The National Cost of the Total Maximum Daily Load Program (Draft Report).* Office of Water, United States Environmental Protection Agency. Washington D.C.

USEPA. 2006. *National Section 303(d) List Fact Sheet.* United States Environmental Protection Agency. Available at http://oaspub.epa.gov/waters/national_rept.control. Accessed Mar. 20, 2006.

Chapter 6. Estimating Uncertainty in Indicator Bacteria TMDL Developed Using HSPF: Reflection on the applications of Monte Carlo based techniques.

Water quality modeling is often conducted to develop Total Maximum Daily Loads (TMDLs). Uncertainty is always present in water quality modeling (Beck, 1987), and without a scientific estimation of uncertainty, it is difficult to estimate the probability of achieving a given water quality standard or the risk of violating it. The research reported here compared and contrasted four different Monte Carlo-based techniques for estimating uncertainty in water quality modeling related to bacterial TMDL development: single-phase Monte Carlo, two-phase Monte Carlo (Helton, 1994; McIntosh et al., 1994), Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992), and Markov Chain Monte Carlo (MCMC) (Kuczera and Parent, 1998). In this chapter, the results from the application of the four techniques are summarized and a recommendation is made regarding the most appropriate technique for this type of application.

The techniques were applied to the Mossy Creek watershed in Virginia. Mossy Creek was listed as impaired in 1996 due to violations of instantaneous fecal coliform (FC) criteria, and a TMDL was developed in 2004 (Benham et al., 2004). TMDL development included the application of HSPF to model the hydrology and water quality of the watershed. The modeling files and data developed for the TMDL were used in this study.

To apply the four uncertainty estimation techniques, distributions were assigned to each of the typically calibrated input hydrology and water quality parameters. For single- and two-phase Monte Carlo simulation, the input parameter distributions were manually calibrated and distributions were adjusted as needed to best match the model output and observed data. The parameters were assumed to be independent and no covariance information was provided for this application. The application of GLUE and MCMC generated posterior or calibrated distributions of typically calibrated parameters using a Bayesian equation. The posterior distributions obtained from GLUE and MCMC reflect the covariance implicitly.

The calibrated or posterior parameter distributions obtained using each technique were used to simulate in-stream FC concentrations in Mossy Creek under two proposed pollutant allocation scenarios presented in the TMDL (table 6-1) for a future prediction period of 1218 days. The major differences between the two scenarios were the level of reduction in FC loading from cattle direct deposit and FC loading in cropland runoff. The first or S1 scenario required a greater reduction of cattle direct deposits to the stream compared to S2 (99% for S1 vs. 94% for S2), but required a lower reduction in FC loads from cropland compared to S2 (90% for S1 vs. 95% for

S2). The scenarios also differed in the reduction called for from wildlife loading, but that FC source category provided a minor contribution to overall FC production in the watershed and was excluded from the analysis conducted here.

Table 6-1 TMDL pollutant allocation scenarios resulting in no violations of instantaneous criteria for indicator bacteria (Benham et al., 2004)

	Required source-specific fecal coliform load reductions (%)						
TMDL Allocation Scenario	Cattle Direct Deposit	Cropland	Pasture	Loafing Lot	Wildlife Direct Deposit	Straight Pipes	All residential pervious land segments
S1	99	90	98	100	30	100	95
S2	94	95	98	100	0	100	95

The predicted average daily FC concentrations from the Monte Carlo simulation were analyzed and an average, 2.5%, 10%, 90%, and 97.5% quantile time series were calculated. For each time series, each occurrence of daily FC concentration greater than instantaneous FC criteria of 400 cfu/100 ml was considered a violation incidence. The number of violations were added for each time series and divided by prediction period (1218 days) to calculate percent violations.

The percent violations reported by the average time series for all the uncertainty estimation techniques for the two scenarios were less than 2% (figure 6.1). The single- and two-phase MC simulation techniques illustrated similar percent violations from average time series. This result was expected, since the distribution of input parameters in both techniques for the prediction period was the same. The range of percent violations for the S1 scenario was similar for simple MC and two-phase, but the range was smaller for S2 for the two-phase MC technique. In other words, the estimated uncertainty reported by two-phase MC was lower compared to single-phase MC. The primary reason for the two-phase MC to report lower uncertainty was that there were likely insufficient iterations to simulate the complete parameter space. In a MC simulation, the modeler’s aim is to conduct sufficient iterations to sample as much parameter space as possible. In the two-phase MC analysis, the parameters were varied at most 300 times, as compared to 12,000 times in single-phase MC. The two-phase MC might yield similar results to single-phase MC if the number of knowledge uncertain and stochastic variability iterations were increased. However, any increase in iterations will increase in computing cost, which must be balanced against additional information.

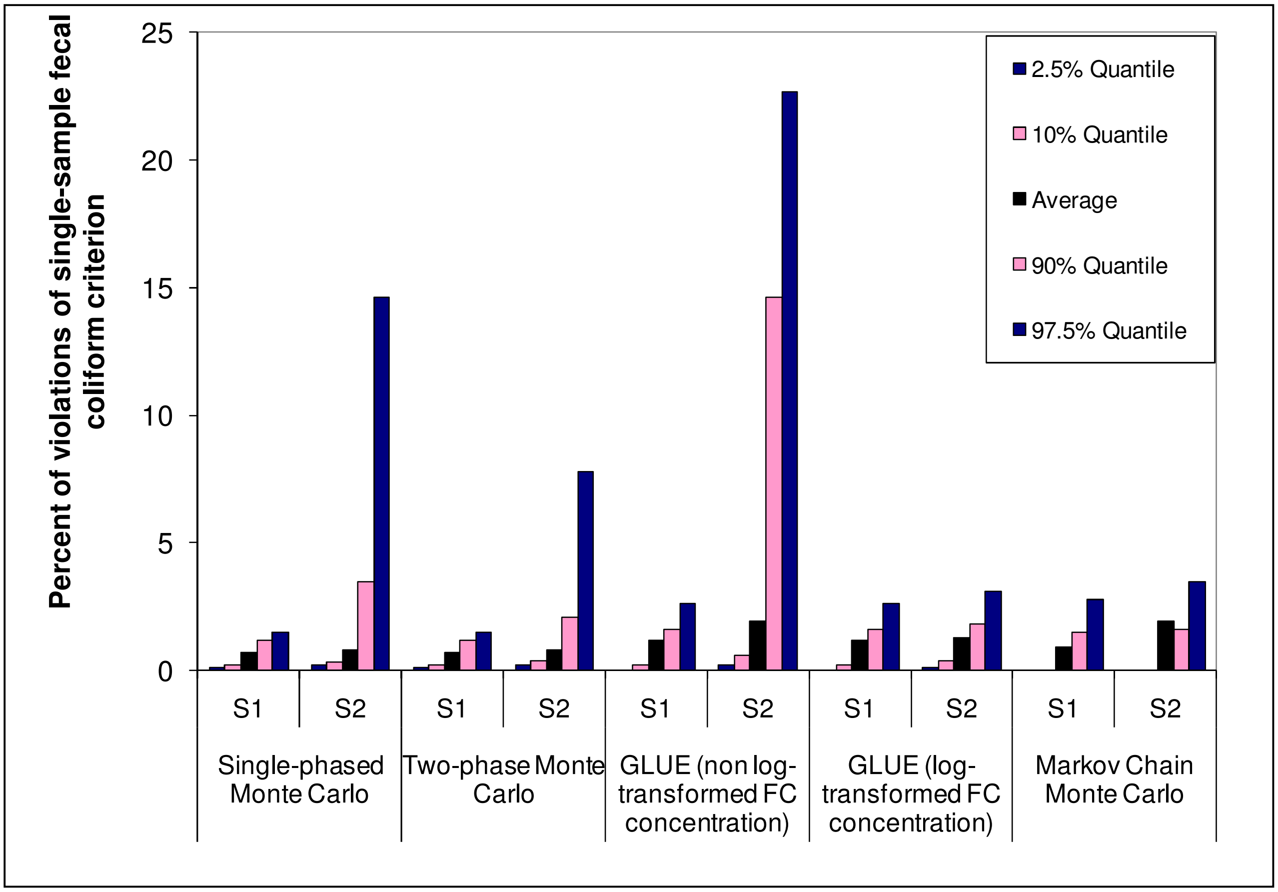


Figure 6.1 The percent violations and uncertainty reported by four uncertainty estimation techniques for two of the allocation scenarios in the Mossy Creek watershed TMDL.

Two-phase MC simulations provided additional information about the effect of knowledge uncertainty and stochastic variability in the model output. In this application, the effect of knowledge uncertainty was significantly greater than the effect of stochastic variability on uncertainty estimates. A modeler can use this information to focus future efforts in collecting more information about the knowledge uncertain parameters. In this application, this result was probably an artifact of assigning only one parameter as stochastically variable. This categorization of a parameter as knowledge uncertain or stochastically variable can be subjective. In the case of watershed modeling, where the parameters are applied to significantly large areas for a simulation period spanning multiple years, it is difficult to classify a parameter as either knowledge uncertain or stochastically variable. Further, to sample complete parameter space with two-phase MC and estimate uncertainty, the number of simulations need to be greater compared to single-phase MC. Therefore, between single phase and two-phase Monte Carlo analysis, single phase is the preferred approach for uncertainty estimation in this type of application.

The uncertainty estimates from the two applications of GLUE underscore the importance of log-transforming the observed and simulated FC concentrations. In the GLUE application with non log-transformed FC concentration data, a few very high observed FC concentrations can affect the likelihood estimation and, consequently, posterior distributions and uncertainty estimates. The percent violations reported by 90 and 97.5% quantiles for the scenario S2 reduced by 8 and 7 times respectively following the log-transformation of FC concentration (figure 6.1). Given these results, it is recommend that the modeler evaluate model response and normalize the observed data and model response, if required, to obtain uncertainty estimates that are not skewed by few data points.

When compared to single-phase MC simulation, GLUE (log-transformed) reported similar uncertainty estimates for scenario S1 and lower uncertainty for scenario S2 (table 6-2). With the GLUE (log-transformed) application, the violations by 90 and 97.5% quantiles reduced by 2 to 4 times, respectively, illustrating that the calibrated (posterior distribution) parameter sets obtained using GLUE reduced uncertainty in the model output.

The application of GLUE was more straightforward compared to the single-phase MC application, as GLUE provides a framework to estimate posterior distributions that are essentially the calibrated parameter distributions, whereas, the calibration process in single-phase and two-phase MC is manual. In the simple MC application, covariance among the parameters must also be provided, whereas the posterior distributions obtained from GLUE reflect the covariance implicitly. The GLUE technique is the preferred approach when selecting between single-phase MC and GLUE to estimate uncertainty in water quality modeling. The application of GLUE, however, is dependent upon the choice and formulation of likelihood function, and model acceptance/rejection criteria. Further research is needed in these areas for an application like this.

The uncertainty reported by GLUE (log-transformed) and MCMC application was very similar. Both techniques use the Bayesian equation to develop posterior distributions and therefore provide a framework for model calibration. However, the application and basic assumptions of the two techniques are different. GLUE uses any goodness of fit criteria to estimate likelihood functions, whereas MCMC has a robust formulation that utilizes statistical likelihood and requires model residuals to be normally distributed. The ability to use any goodness of fit criteria provides flexibility in the application of GLUE, but it is also a controversial aspect of GLUE. As mentioned earlier, GLUE application requires the input from modeler in selecting and formulating a likelihood function and selection of model acceptance/rejection criteria, and more research is needed to identify suitable options for modeling FC concentration using HSPF. MCMC requires input from the modeler in selecting the variance scaling factor (that affects the movement of Markov Chain) and identifying the burn-in periods for the Markov Chain.

MCMC is also more computationally expensive than GLUE for watershed model applications. The MCMC application required three independent Markov chains with 100000 iterations each to identify the posterior parameter distributions. For the Mossy Creek watershed model, the difference in computing time between 10000 model simulations for the GLUE technique and the 100000 model simulations for the MCMC technique was of few hours to about a week on an average desktop computer. The Mossy Creek watershed model was relatively simple with only eight subwatersheds and six reaches. As the number of watersheds, land uses and constituents increase, the HSPF simulation time will also increase. However, as the computing power of computers is increasing exponentially, the difference in computing cost among GLUE and MCMC might reduce in the near future. When comparing GLUE and MCMC, based on the uncertainty estimates by the techniques, computing cost, and the flexibility in selection of likelihood functions, GLUE is the recommended alternative for estimating uncertainty in water quality modeling.

Overall, among single-phase MC, two-phase MC, GLUE and MCMC, GLUE is the preferred approach to estimate uncertainty in water quality modeling in the types of applications similar to the one presented here. However, we also recognize that with respect to water quality modeling with HSPF, several aspects of GLUE applications do need further research.

References:

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Benham, B.L., K. Branna, K. Christophel, T. Dillaha, L. Henry, S. Mostaghimi, R. Wagner, J. Wynn, G. Yagow, and R. Zeckoski. 2004. Total maximum daily load development for Mossy Creek and Long Glade Run: Bacteria and general standard (Benthic) impairments. Richmond, Va.: Virginia Department of Environmental Quality. Available at http://www.deq.state.va.us/tmdl/homepage.html. Assessed Jan 25, 2005.

Beven, K.. and A. Binley (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes* 6(3): 279-298.

Helton, J.C. 1994. Treatment of uncertainty in performance assessment for complex systems. *Risk Analysis*. 14(4): pp 483-511.

Kuczera, G., and E. Parent. 1998. Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: the Metropolis Algorithm. *Journal of Hydrology. 211:* 69-85.

MacIntosh, D.L., G.W. Suter., and F.O. Hoffman. 1994. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. *Risk Analysis*. 14(4): 405-419.