ASSESSING THE PERFORMANCE OF HSPF WHEN USING THE HIGH WATER TABLE SUBROUTINE TO SIMULATE HYDROLOGY IN A LOW-GRADIENT WATERSHED

by

Michael Scott Forrester

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

Master of Science
in
Biological Systems Engineering

Brian L. Benham, Chair
Karen S. Kline
Kevin J. McGuire

April 17, 2012
Blacksburg, VA

KEY TERMS: modeling, HSPF, high water table, uncertainty

ASSESSING THE PERFORMANCE OF HSPF WHEN USING THE HIGH WATER TABLE SUBROUTINE TO SIMULATE HYDROLOGY IN A LOW-GRADIENT WATERSHED

by

Michael Scott Forrester

ABSTRACT

Modeling groundwater hydrology is critical in low-gradient, high water table watersheds where groundwater is the dominant contribution to streamflow. The Hydrological Simulation Program-FORTRAN (HSPF) model has two different subroutines available to simulate groundwater, the traditional groundwater (TGW) subroutine and the high water table (HWT) subroutine. The HWT subroutine has more parameters and requires more data but was created to enhance model performance in low-gradient, high water table watershed applications. The objective of this study was to compare the performance and uncertainty of the TGW and HWT subroutines when applying HSPF to a low-gradient watershed in the Coastal Plain of northeast North Carolina. One hundred thousand Monte Carlo simulations were performed to generate data needed for model performance comparison. The HWT model generated considerably higher Nash-Sutcliffe efficiency (NSE) values while performing slightly worse when simulating the 50% lowest and 10% highest flows. Model uncertainty was assessed using the Average Relative Interval Length (ARIL) metric. The HWT model operated with more average uncertainty throughout all flow regimes. Based on the results, the HWT subroutine is preferable when applying HSPF to a low-gradient watershed and the accuracy of simulated stream discharge is important. In situations where a balance between performance and uncertainty is called for, the choice of which subroutine to employ is less clear cut.

ACKNOWLEDGMENTS

There were many people who directly and indirectly helped me during this research and deserve my sincere gratitude. I owe many thanks to my advisor Brian Benham for upholding my work and progress to the highest of standards. Without his guidance this research and paper would not have fulfilled their full potential. I would like to thank my committee member Karen Kline for her real-time help with HSPF and for lessening my TMDL workload during busy research weeks. I would like to thank my third committee member, Kevin McGuire, for bringing his wealth of hydrology knowledge to our committee.

I would like to extend my gratitude to Becky Zeckoski for her seemingly endless patience and knowledge of Visual Basic. The VB.NET program written for this research was a direct result of hundreds of emails back and forth with Becky and her desire to help me get the program working in a timely manner.

Many of my fellow BSE students also helped me along the way. Many thanks to Bobby Jocz for wading through Ahoskie Creek with dysfunctional waders to help me survey stream cross-sections. I would also like to thank Rich Allevi and Daniel Chuquin for allowing me to bother them at their desks whenever I needed a break from formatting figures or dealing with troublesome VB.NET errors.

Finally, I would like to thank Theresa Sweeney for her love and support while I was conducting this research. Though she was sometimes halfway around the world, she was always enthused to hear about my successes (and failures) even when they did not make much sense.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The century old science of hydrological modeling is focused on developing tools that allow users to simulate hydrologic responses to driving forces at various scales. Early models were empirically-based and sought to simulate relatively simple processes like infiltration (Green and Ampt, 1911; Horton, 1933) and runoff (Sherman L, 1932; Horton, 1945). Models have evolved over the years and are now used to provide simulations of nearly every hydrologic component in a watershed, often at once in an all-inclusive model (Singh, 1995). These more recent, complex models synthesize multiple equations and theories to provide a detailed hydrologic simulation. Watershed-scale models are often physically based, allowing users to input the land-use and soil characteristics of the watershed and then use that information to generate model output(Migliaccio and Srivastava, 2007). Aside from physical attributes (e.g., slope, land use), models use parameters that characterize the various hydrologic processes such as evapotranspiration, infiltration, and groundwater flow(Beven, 2000). Models are frequently used as decision support tools to aid in management decisions such as the application of best management practices to address pollutant loading in a watershed (Borah and Bera, 2003; Wagener and Gupta, 2005).

The Hydrological Simulation Program-FORTRAN (HSPF) is an example of a watershed-scale, physically based model (Bicknell *et al.*, 2001). HSPF combines physical data (e.g. land use, slope) and a substantial set of parameters to simulate hydrology and pollutant fate and transport for various constituents such as bacteria, nutrients, and sediment(Migliaccio and Srivastava, 2007).HSPF has two subroutines that can be chosen to simulate groundwater, the traditional groundwater subroutine (TGW) and the high water table groundwater subroutine (HWT). The HWT model was developed for use in low-gradient, high water table watersheds (Bicknell *et al.*, 2001). While the HWT model was created to provide better simulations in these watersheds, a comparison of the use and performance of the HWT model versus the TGW model including an uncertainty analysis is needed.

An estimate of uncertainty in hydrological models is beneficial because models are simplifications of heterogeneous systems and model structure, parameters, and data input all present an inherent uncertainty (Sivapalan, 2003; Wagener and Gupta, 2005; Wagener and Montanari, 2011). Despite the unavoidable uncertainty, simulation models are frequently used

and needed to explore the various assumptions made about the real world hydrologic cycle and to predict behavior within this cycle for decision making purposes (Beven, 1989). How to deal with and assess the present uncertainty is a challenge for modelers. In fact, as Pappenberger and Beven (2006) argue, much of the scientific/modeling community is still resistant to the idea of estimations of uncertainty in hydrological modeling. Despite this, an estimation of model uncertainty, especially when assessing the performance of models, allows for a comparison of the models' input and structural error.

CHAPTER 2: LITERATURE REVIEW

Hydrological Simulation Program – FORTRAN (HSPF)

　　HSPF is a continuous, semi-lumped watershed-scale model developed in collaboration with the United States Environmental Protection Agency (USEPA) and is based off of the Stanford Watershed Model IV (Crawford and Linsley, 1966; Donigian *et al.*, 1995; Bicknell *et al.*, 2001). HSPF combines multiple models and empirical equations to simulate hydrology and water quality constituents such bacteria, sediment, and nutrients, among others (Donigian *et al.*, 1995). HSPF is part of USEPA's Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) package frequently used in total maximum daily load (TMDL) studies (Borah, 2011). HSPF is considered a "semi-lumped" model because it allows stream reaches to be routed into others but lumps together land uses and parameter values by subwatershed (Cho *et al.*, 2009). The model generally uses an hourly time step to translate meteorological and physical input data into streamflow and other hydrological components such as infiltration, interflow, evapotranspiration, and runoff (Migliaccio and Srivastava, 2007). Different water quality modules can then be called upon to simulate a variety of pollutants.

　　From a water balance perspective, HSPF breaks up the watershed into a series of zones that vary by subwatershed and by land-use (Bicknell *et al.*, 2001). Water in each zone is subject to a different set of governing equations and their respective parameters. Figure 2-1 shows the top to bottom breakdown of the water budget zones used in HSPF. The option to simulate the snow zone is generally only used in areas where snow and snowmelt have large effects on the water budget. In HSPF, the Upper, Lower, and Groundwater zones all lie in the subsurface and water movement between the zones is simulated using one of two groundwater subroutines (TGW, HWT). While the two models handle the movement of water differently, they both use the concept of zones (Figure 2-1) to aid in model parameterization. The difference lies in how the models simulate interactions between the zones in calculating the water budget.
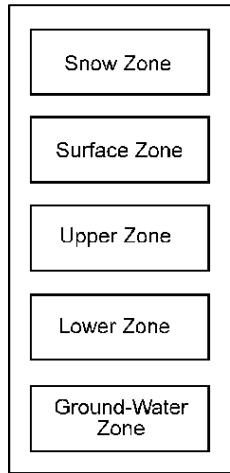
FIGURE 2-1. Zones or Storage Volumes Used by HSPF to Simulate the Water Balance. (Bicknell *et al*., 2001) (Used under fair use, 2012)

Because of extensive parameterization, HSPF is both extremely flexible and complex. While the HSPF users guide provides typical ranges for parameter values (Bicknell *et al.*, 2001), calibration is necessary to improve simulation accuracy (Refsgaard, 1997; Kim *et al.*, 2007). Calibration is often performed with the assistance of the digitally-based HSPF expert system for calibration, HSPEXP, used to analyze model output. Based on this analysis, HSPEXP suggests changes in parameter values designed to improve selected model performance statistics. The default flow-related model performance statistic criteria used in HSPEXP (Lumb *et al.*, 1994) are shown in Table 2-1. More restrictive criteria can be used if desired (Kim *et al.*, 2007).

TABLE 2-1. HSPEXP Default and Optional Model Performance Criteria.

Flow Statistic	Default Criteria Percent Error	Optional Criteria Percent Error
Error in total volume	± 10	± 10
Error in low flow recession	± 0.03	± 0.01
Error in 50% lowest flows	± 10	± 10
Error in 10% highest flows	± 15	± 15
Error in storm peaks	± 20	± 15
Seasonal volume error	± 30	± 10
Summer storm volume error	± 50	± 15

HSPF is typically calibrated using one set or time period of observed data and validated for a different set of observations (Kim *et al.*, 2007), sometimes referred to as a "differential split-sample" test (Klemes, 1986). Validation seeks to corroborate the parameter values

determined during calibration by obtaining adequate model output during a separate time period. Using the same model performance criteria for both periods, model parameters are adjusted through calibration until a satisfactory model is developed for both the calibration and validation periods.

Modeling Low-Gradient Watersheds with High Water Tables

Most watershed-scale models incorporate groundwater into the water balance to some extent. Models such as HSPF and the Soil and Water Assessment Tool (SWAT) keep track of the volume of water in active groundwater storage zones and calculate the release of this water into streams using various approaches/relationships (Migliaccio and Srivastava, 2007). HSPF, SWAT and the Watershed Assessment Model (WAM) also let users define deep percolation parameter values where water routed to these deep aquifers never returns to the stream or river (Migliaccio and Srivastava, 2007). Accurate modeling of groundwater contributions to streamflow is critical, especially in areas where the groundwater contribution from shallow/high water tables is a dominant component of streamflow (Yeh and Eltahir, 2005).

HSPF has two methods for simulating groundwater. The TGW subroutine is more basic and is typically used in HSPF applications. When using the TGW subroutine, the groundwater contribution to streams is some fraction of the active groundwater storage (AGWS) (Zhang *et al.*, 2010). The AGWS parameter simulates a shallow groundwater aquifer that readily contributes to both evapotranspiration and baseflow (Migliaccio and Srivastava, 2007). Calibrated parameters determine how much of the groundwater in this zone is lost to deep aquifers (DEEPFR), evapotranspiration (AGWETP), or contributes to baseflow (AGWRC) (Bicknell *et al.*, 2001). The water that percolates into the deep aquifer (DEEPFR) is lost to a separate, inactive storage (IGWI) and never contributes to baseflow or evapotranspiration (Bicknell *et al.*, 2001). Figure 2-2 is a flow diagram showing the various storages, parameters, and linkages present in the TGW subroutine (Bicknell *et al.*, 2001).
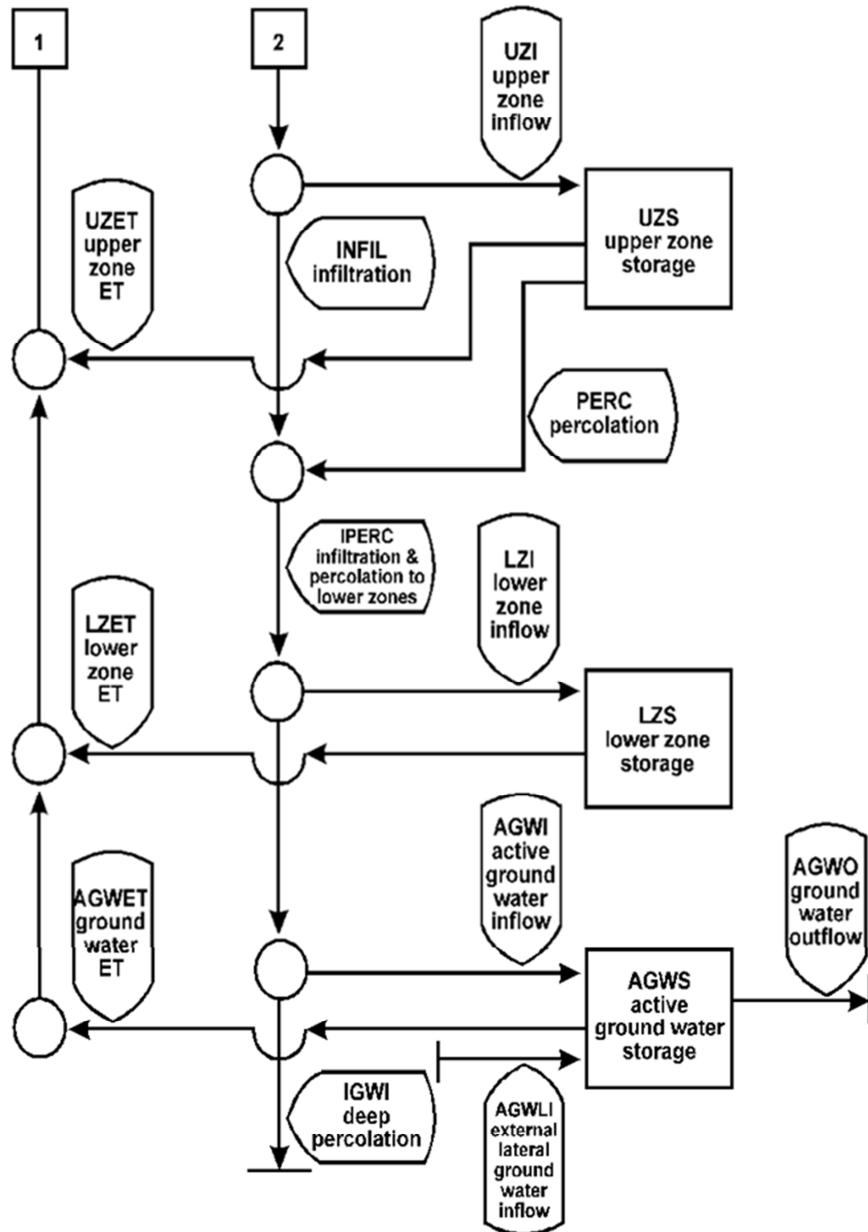
FIGURE 2-2. Flow Chart of the TGW Subroutine. (Bicknell *et al*., 2001) (Used under fair use, 2012)

The HWT subroutine is more sophisticated and specialized. The HWT subroutine tracks the groundwater elevation (GWEL) and simulates interactions between the Upper, Lower, and Groundwater zones as GWEL rises and falls. Figure 2-3 illustrates the soil profile as defined in the HWT subroutine, showing the location of the two influence elevation levels that separate the sub-surface water storage zones (AGWS, LZS, UZS, IFWS) (Bicknell *et al.*, 2001). These influence elevations separate the soil column into three regions, as they are referred to in the HSPF documentation. In the HWT subroutine, the PCW parameter represents the porosity

present in micropores and is constant for all three regions. The PGW parameter represents the porosity present in macropores in Regions 1 and 2 and the UPGW parameter represents porosity of the macropores in Region 3. UPGW is generally greater than PGW based on the assumption that soil in the upper layer is more proliferated by roots and contains more structure (Bicknell et al., 2001). The porosities for the three regions (PGW, PCW, UPGW) are typically quantified using available soil data and averaged over the different land-uses and subwatersheds.
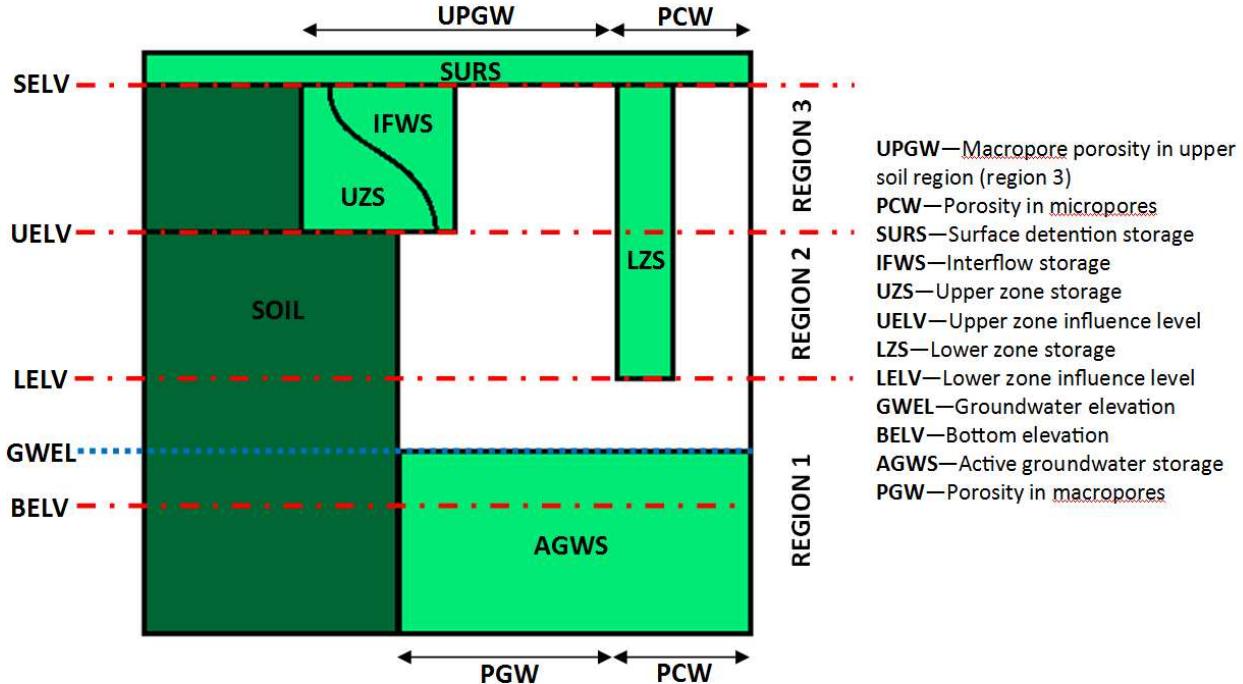
FIGURE 2-3. Soil Profile Representation Used in the HWT Subroutine. (Bicknell *et al*., 2001) (Used under fair use, 2012)

As can be seen in Figure 2-3, Region 1 is the deepest, lying below the lower zone influence elevation (LELV). Groundwater in this region is quantified by AGWS. Region 1 is assumed to be saturated; the groundwater in Region 1 does not interact with the unsaturated zone (above GWEL). Region 2 lies above the LELV with the top of Region 2 being defined by the upper zone influence elevation (UELV). If the GWEL rises into Region 2, groundwater storage interacts with the lower zone storage (LZS). Groundwater that occupies micropores in Region 2 is quantified with the parameter LZS. Increases in soil moisture within LZS are assumed to be uniform throughout the entire zone. Water that enters LZS is held by cohesive forces, meaning it cannot percolate back into the groundwater and can only exit through evapotranspiration (LZETP). The remaining groundwater within Region 2 occupies macropores, is subject to the groundwater parameters DEEPFR and AGWETP, and will continue to contribute to baseflow

(AGWRC). When the GWEL reaches Region 3, above the UELV, the groundwater shares storages with interflow (IFWS) and upper zone storage (UZS). Changes in groundwater that are not attributed to LZS are assigned to IFWS and UZS based on their relative levels of saturation. If the water table rises to the surface any additional water is added to surface storage and can potentially contribute to runoff (Bicknell *et al.*, 2001).

Use of either groundwater subroutine in HSPF requires parameter calibration. Calibration of HSPF parameters relies heavily on guidance provided in the HSPF documentation and on trial and error. Said et al. (2007) proposed using wells placed in strategic locations in a watershed to help calibrate HSPF groundwater parameters. The researchers used observed water table records to provide a more accurate set of parameters. Using the TGW model discussed above and two wells within a watershed, the AGWS curve was fitted to water table elevation data to provide empirical estimates for groundwater parameters. A similar technique may be appropriate for estimating parameters using the HWT model.

Uncertainty in Hydrological Modeling

Uncertainty in hydrological modeling is unavoidable (Beven, 2000). This uncertainty stems from multiple sources during the formation/execution of a model. First, the observed data used to develop and calibrate the model is not without error (Beven, 2000). For example, the resolution of weather stations near a watershed of interest may provide adequate, but not entirely accurate, estimates of precipitation, temperature, etc. that are needed to input into the model (Tetzlaff and Uhlenbrook, 2005). Second, and more importantly, any hydrological model is a simplification of a much more complicated natural system. Vast heterogeneities exist in all scales of hydrological systems (Sivapalan, 2003; Beven, 2006b). This forces hydrologic modelers to make certain assumptions in order to simplify the system. One example of this is the Darcian flow partial differential equations for flow through the soil matrix used in most distributed-parameter models (Beven, 1993). In its simplest form, Darcy's law states that the velocity of a fluid through a porous media is proportional to the pressure gradient and a constant characterizing the permeability of the soil (Muskat, 1937). This equation, however, was derived from small-scale lab experiments. Once this equation attempts to estimate subsurface flow in the hillslope or watershed-scale it will face large, three-dimensional heterogeneities that force the modeler to begin making assumptions (Beven, 1993).

Semi-lumped parameter models like HSPF, for example, can assign different values for the permeability of soils only to a certain extent. In an HSPF application, subwatersheds are broken down into different pervious land segments (e.g. forest, pasture, cropland) and impervious land segments, with parameter values assigned to each segment. Within the individual land segments, however, all parameter values are assumed to be constant. This means that the model assumes all of the soil properties within a given subwatershed's separate land uses are constant, an assumption that is not true but necessary given the limitations of semi-lumped models.

The heterogeneity and complexity of hydrological systems almost always requires some calibration of parameters (Refsgaard, 1997; Gupta *et al.*, 1998). Parameters can be subjectively adjusted, often with some guidance, until model outputs are within an acceptable range of observed values. Calibration with the goal of bringing simulated values closer to observed values brings about two issues. First, the model parameters can begin to stray away from an actual physical representation of a particular component of the system in favor of a value that yields more accurate results (Wagener and Gupta, 2005). Attempts to choose effective parameter values based on field measurements is most often futile because the scale at which the parameters are defined in the model is usually much larger than the scale that measurements can be made in the field. In addition, there are often conceptual model components that do not have a direct physical meaning (Beven, 1993; Sivapalan, 2003).

The second issue that arises during calibration is a concept known as equifinality. Equifinality is the phenomenon where many different parameter sets output similarly acceptable results (Spear *et al.*, 1994; Beven and Freer, 2001; Beven, 2006a). In a model such as HSPF with numerous parameters, the interactions between both the physical characteristics in the watershed and the parameters that attempt to represent these characteristics can result in numerous acceptable parameter sets. With so many viable parameter sets, and possibly no single optimum set, there is considerable uncertainty in parameter values and the simulated results (Beven and Freer, 2001). A main goal of estimating uncertainty, therefore, is to quantify the probability that a particular parameter set's output lies within a given range of the observed value (Beven, 2000). An uncertainty analysis takes into account both the uncertainty inherent in model input data and the uncertainties that arise by simplifying the watershed's processes into model form.

Monte Carlo (MC) simulations provide a basis for generating multiple parameter sets for models where one optimum parameter set does not exist, addressing equifinality. MC

simulations generate parameter sets by sampling from distributions of uncertain parameters (Whitehead and Young, 1979; Wagener and Kollat, 2007). Each parameter set is used to model the system and response variables (e.g. discharge, pollutant loading) and can then be used to assess performance and parameter uncertainty. Multiple MC runs generate distributions of response variables at each time step. If the 2.5% and 97.5% quantiles of these distributions are calculated, the 95% confidence interval for prediction can be determined. The confidence interval represents the band or range of uncertainty for the response variable at each time step (Beven, 1993).

Jin et al. (2010) developed a simple method to quantify model uncertainty called the Average Relative Interval Length (ARIL). The ARIL metric averages the width of the response variable confidence interval at each time step relative to the observed data at that time step over the entire simulation period. ARIL has been used in other studies to assess and compare and quantify the uncertainty present in hydrological models (Jin *et al.*, 2010; Li *et al.*, 2010; Kizza *et al.*, 2011).

By performing multiple MC runs or by using an MC optimization method such as the Shuffled Complex Evolution Metropolis algorithm (Vrugt *et al.*, 2003), behavioral parameter sets can be derived. The concept of behavioral versus non-behavioral parameter sets was first introduced by Spear and Hornberger (1980) and is used in the Generalized Likelihood Uncertainty Estimation (GLUE) method of estimation uncertainty (Beven and Binley, 1992). In the GLUE methodology, behavioral parameter sets are determined using an objective function and a threshold performance value. Parameter sets resulting in model runs that fall above the threshold value are considered to be behavioral and provide an acceptable characterization of the watershed (Beven and Binley, 1992). By filtering MC results into behavioral and non-behavioral parameter sets, only MC model runs identified as behavioral parameter sets are included in the model performance and uncertainty analyses. Selecting behavioral parameter sets also allows for a regional sensitivity analysis. Posterior behavioral parameter distributions that are not uniform and show a trend towards a particular value or range indicate a parameter that is more sensitive in the model (Spear and Hornberger, 1980).

Conclusion

The assumptions and simplifications present in hydrological models are at once necessary and problematic for modelers. The vast and not fully understood heterogeneities in hydrological systems make an exact model impossible. Performing Monte Carlo simulations, and thus accepting the concept of equifinality, allows for an assessment of model performance that implicitly recognizes the uncertainty present in both the model input data and the structure of the model itself. Uncertainty analyses should be incorporated in any comparison of different models or different forms of the same model. HSPF has a separate groundwater subroutine for low-gradient areas with a high water table. The HWT subroutine estimates the water table elevation and uses that as a basis to estimate how groundwater affects discharge. This research compares the performance and uncertainty between the TGW and HWT subroutines when using HSPF to simulate hydrology in a low-gradient, high water table watershed.

CHAPTER 3: ASSESSING THE PERFORMANCE OF HSPF WHEN USING THE HIGH WATER TABLE SUBROUTINE TO SIMULATE HYDROLOGY IN A LOW-GRADIENT WATERSHED

Modeling groundwater hydrology is critical in low-gradient, high water table watersheds where groundwater is the dominant contribution to streamflow. The Hydrological Simulation Program-FORTRAN (HSPF) model has two different subroutines available to simulate groundwater, the traditional groundwater (TGW) subroutine and the high water table (HWT) subroutine. The HWT subroutine has more parameters and requires more data but was created to enhance model performance in low-gradient, high water table watershed applications. The objective of this study was to compare the performance and uncertainty of the TGW and HWT subroutines when applying HSPF to a low-gradient watershed in the Coastal Plain of northeast North Carolina. One hundred thousand Monte Carlo simulations were performed to generate data needed for model performance comparison. The HWT model generated considerably higher Nash-Sutcliffe efficiency (NSE) values while performing slightly worse when simulating the 50% lowest and 10% highest flows. Model uncertainty was assessed using the Average Relative Interval Length (ARIL) metric. The HWT model operated with more average uncertainty throughout all flow regimes. Based on the results, the HWT subroutine is preferable when applying HSPF to a low-gradient watershed and the accuracy of simulated stream discharge is important. In situations where a balance between performance and uncertainty is called for, the choice of which subroutine to employ is less clear cut.

(KEY TERMS: modeling, HSPF, high water table, uncertainty)

3.1: INTRODUCTION

Hydrologic models have evolved over time and are now used to provide simulations of nearly every hydrologic component in a watershed, often at once in an all-inclusive model (Singh and Woolhiser, 2002). These complex models synthesize multiple equations and theories to provide a detailed hydrologic simulation. Watershed-scale models are often physically based, allowing users to input the land-use and soil characteristics of the watershed, using that

information to formulate results (Beven, 1989; Grayson *et al.*, 1992). Aside from physical attributes (e.g., slope, land use), models use parameters that characterize the various hydrologic processes such as evapotranspiration, infiltration, and groundwater flow (Refsgaard, 1997; Migliaccio and Srivastava, 2007).

The Hydrological Simulation Program-FORTRAN (HSPF) is an example of a complex, physically-based watershed-scale model (Donigian *et al.*, 1995). HSPF combines physical data and a substantial set of parameters to simulate hydrology and pollutant fate and transport for various water quality constituents such as bacteria, nutrients, and sediment (Migliaccio and Srivastava, 2007). HSPF is part of the USEPA's Better Assessment Science Integrating Point and Non-point Sources (BASINS) tool and is frequently used in the development and analysis of Total Maximum Daily Loads (TMDL) (Borah and Bera, 2003).

The HSPF subroutine of interest for this research is the high water table (HWT) subroutine developed for use in wetlands and low-gradient areas (Bicknell *et al.*, 2001). The HWT subroutine differs from the traditional HSPF groundwater subroutine (TGW) in that it simulates the groundwater elevation relative to an arbitrary datum (e.g. sea level) rather than just assuming a perched groundwater aquifer (Donigian *et al.*, 1995; Bicknell *et al.*, 2001). Though the HWT model was created to provide a better representation of low-gradient watersheds with a high water table, there is currently no peer-reviewed literature discussing the use of this subroutine. The objective of this research was to use the TGW and HWT subroutines to simulate discharge in a low-gradient watershed and compare the performance and uncertainty of each.

An estimation of the uncertainty associated with each model is necessary for a full comparison because hydrologic models are simplifications of heterogeneous systems and model structure, parameters, and data input all present an inherent uncertainty (Beven and Binley, 1992; Sivapalan, 2003; Wagener and Gupta, 2005; Wagener and Montanari, 2011). Because both models use the same input data such as weather and land use, both are subject to the same uncertainty that this data presents (see Tetzlaff and Uhlenbrook, 2005). Therefore, the calculations of uncertainty within this study reflect the uncertainty due to parameter variability and how this variability, along with differences in groundwater subroutine model structure, affects model output.

Groundwater Simulation in HSPF

The TGW subroutine is typically used in HSPF applications. When using the TGW subroutine, the groundwater contribution to streams is some fraction of the active groundwater storage (AGWS) (Zhang *et al.*, 2010). The AGWS parameter simulates a shallow groundwater aquifer that readily contributes to both evapotranspiration and baseflow (Said *et al.*, 2007). Calibrated parameters determine how much of the groundwater in this zone is lost to deep aquifers (DEEPFR), evapotranspiration (AGWETP), or contributes to baseflow (AGWRC) (Bicknell *et al.*, 2001). The water that percolates into the deep aquifer (DEEPFR) is lost to a separate, inactive storage (IGWI) and never contributes to baseflow or evapotranspiration (Bicknell *et al.*, 2001).

The HWT subroutine is more sophisticated and specialized. The HWT subroutine tracks the groundwater elevation (GWEL) and simulates interactions between the Upper, Lower, and Groundwater zones in HSPF as GWEL rises and falls. Figure 3-1 illustrates the soil profile as defined in the HWT subroutine, showing the location of the two influence elevation levels that separate the sub-surface water storage zones (AGWS, LZS, UZS, IFWS) (Bicknell *et al.* 2001). These influence elevations separate the soil column into three regions, as they are referred to in the HSPF documentation. The PCW parameter represents the porosity present in micropores and is constant for all three regions. The PGW parameter represents the porosity present in macropores in Regions 1 and 2 and the UPGW parameter represents porosity of the macropores in Region 3. UPGW is generally greater than PGW based on the assumption that soil in the upper layer is more disturbed (Bicknell *et al.*, 2001). In practice, the porosities for the three regions (PGW, PCW, UPGW) are typically quantified using available soil data and averaged over the different land-uses and subwatersheds.

FIGURE 3-1. Soil Profile Representation Used in the HWT Subroutine. (Bicknell *et al.*, 2001) (Used under fair use, 2012)
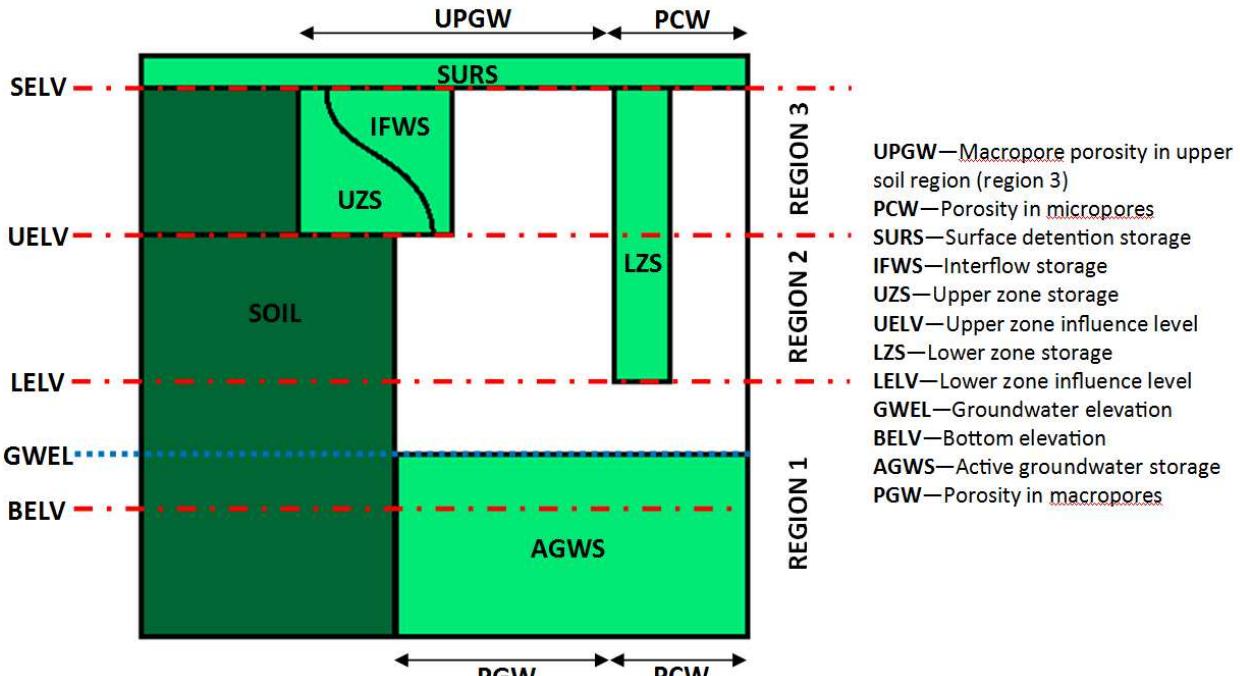
As can be seen in Figure 3-1, Region 1 is the deepest, lying below the lower zone influence elevation (LELV). Groundwater in this region is quantified by AGWS. Region 1 is assumed to be saturated and does not interact with the unsaturated zone (above GWEL). When GWEL is in Region 1 the HWT subroutine operates in the same way as the TGW subroutine. Region 2 lies above the LELV with the top of Region 2 being defined by the upper zone influence elevation (UELV). If the GWEL rises into Region 2, groundwater storage interacts with the lower zone storage (LZS). Groundwater that occupies micropores in Region 2 is quantified with the parameter LZS. Increases in soil moisture within LZS are assumed to be uniform throughout the entire zone. Water that enters LZS is held by cohesive forces, meaning it cannot percolate back into the groundwater and can only exit through evapotranspiration (LZETP). The remaining groundwater within Region 2 occupies macropores and is subject to the groundwater parameters DEEPFR and AGWETP and will continue to contribute to baseflow (AGWRC). When the GWEL reaches Region 3, above the UELV, the groundwater shares storages with interflow (IFWS) and upper zone storage (UZS). Changes in groundwater that are not attributed to LZS are assigned to IFWS and UZS based on their relative levels of saturation. If the water table rises to the surface any additional water is added to surface storage and can potentially contribute to runoff (Bicknell *et al.*, 2001).

15

The implications of the interactions between groundwater and the storages in the upper and lower zones reach beyond hydrology alone. When considering the fate of nutrients in a watershed, the modeling of groundwater and streamflow generation processes can affect delivery of nutrients to the stream (Haith and Shoenaker, 1987). Water reaching the stream from groundwater sources will face different nutrient transport processes than surface runoff (Beaulac and Reckhow, 1982; Valiela *et al.*, 1999). These differences make accuracy and uncertainty associated with modeling the balance between streamflow generation processes important.

Parameter Uncertainty and Identifiability

Uncertainty in hydrological modeling is present in model input data, model structure, and the variability of individual parameters (Beven, 2000). The vast heterogeneities in all scales of hydrological systems and limitations in data collection techniques make this uncertainty unavoidable (Sivapalan, 2003; Tetzlaff and Uhlenbrook, 2005; Beven, 2006b). A full comparison of model performance should attempt to quantify and compare relevant sources of uncertainty. In conjunctions with estimates of model uncertainty due to parameter variability, a regional sensitivity analysis of model parameters can be useful in determining parameters to which the model is more sensitive (Spear and Hornberger, 1980). These parameters will have an identifiable value or range that tends to yield better model outputs respective to a particular performance metric. Models with more identifiable parameters can be easier to calibrate to observed data (Brun *et al.*, 2001).

3.2: METHODS

Study Site

The Ahoskie Creek watershed (14,700 ha) in northeastern North Carolina was selected for this study (Figure 3-2). Ahoskie Creek lies in the Lower Coastal Plain physiographic province, a region characterized by flat surface slopes (<2%) and channel slopes (<0.1%) (Sheridan, 2002). Ahoskie Creek is in the Chowan River basin which drains to the Albemarle Sound. The dominant land uses are forest (50.7%) and cropland (22.9%) (Homer *et al.*, 2012). The stream reaches in the watershed were channelized between 1969 and 1972 to improve drainage. Prior to channelization, the watershed did not have clearly discernible channels

resulting in poor drainage and frequent flooding events (USDA-ARS, 1977). The low-gradient, documented flooding events, and presence of wetlands are indications of a high, active water table making the Ahoskie Creek watershed a good candidate for this study.
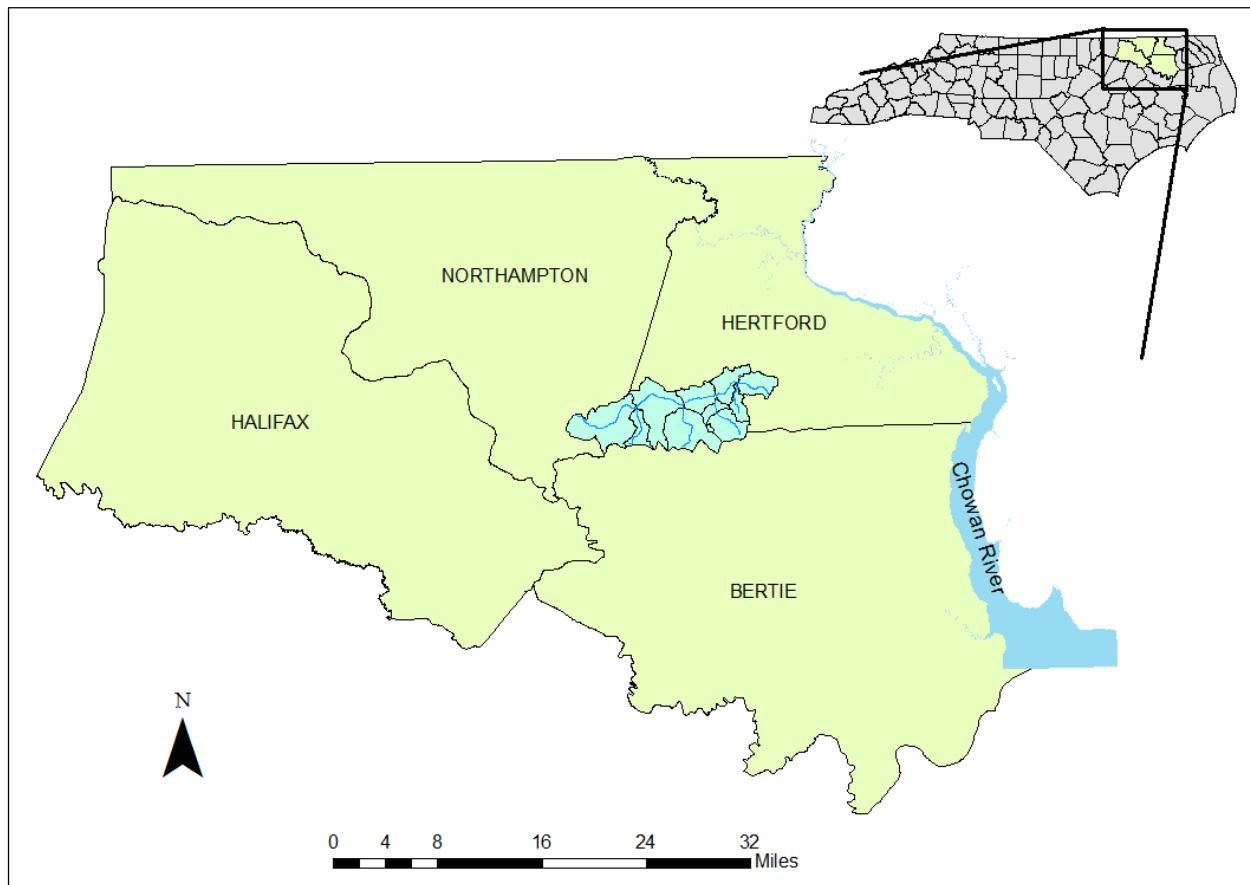
FIGURE 3-2. Location of the Ahoskie Creek Watershed, North Carolina, USA.

HSPF Model Development

HSPF relies on a combination of remotely sensed data and field-based observations to characterize the watershed and parameterize the model. Using digital raster graphics (DRG) and digital elevation model (DEM) data (USGS, 2006a; USGS, 2010) the Ahoskie Creek watershed was manually delineated into ten (10) subwatersheds, with an average subwatershed area of 1559 ha (Figure 3-3).
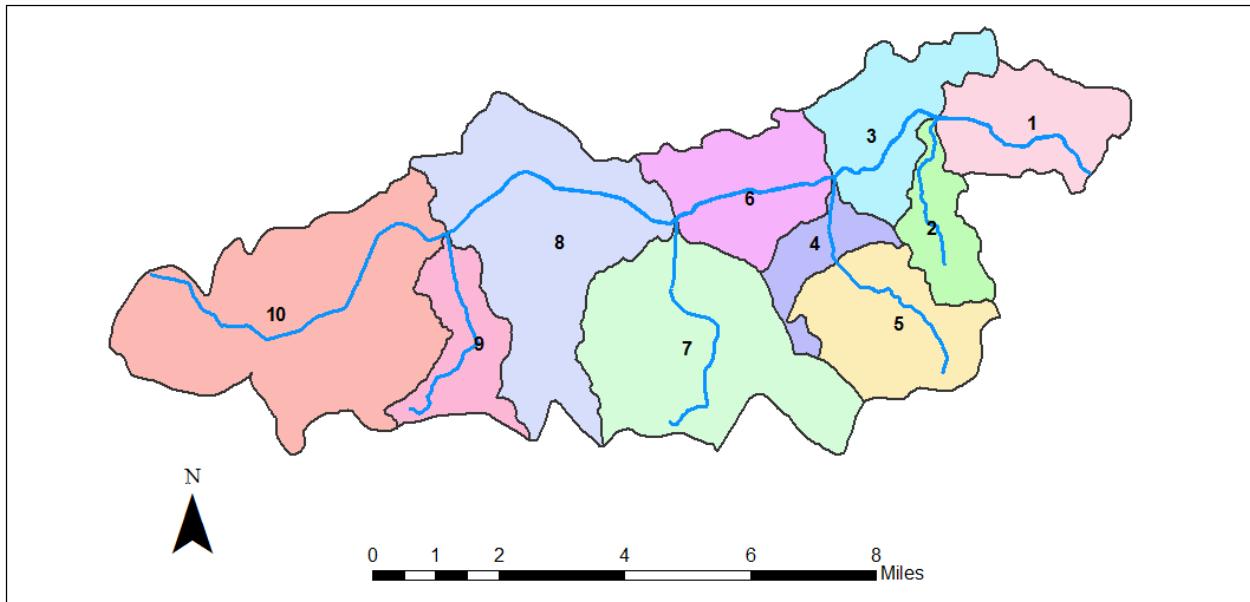
FIGURE 3-3. Ahoskie Creek Subwatersheds.

The land uses for Ahoskie Creek were derived from the National Land Cover Database (NLCD) (USGS, 2006b). The NLCD land uses present in the watershed were aggregated into five broad land use categories: forest, cropland, pasture, residential, and wetlands (Figure 3-4, Table 3-1). When using HSPF, each subwatershed can include pervious land segments (PLS) and/or impervious land segments (IPS). Watershed characterization involves parameterizing the various PLSs and IPSs for each subwatershed.
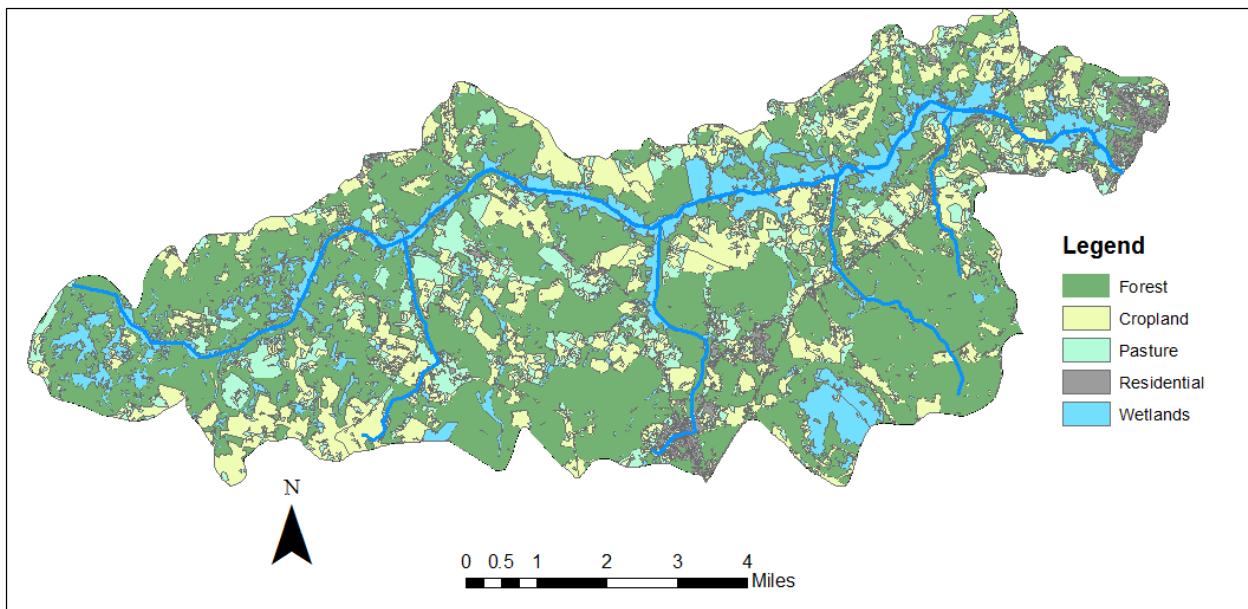
FIGURE 3-4. Ahoskie Creek Watershed Land Use Distribution.

TABLE 3-1. Ahoskie Creek Watershed Land Use Distribution.

Land Use Type	Percentage of Total Area (%)
Forest	50.7
Cropland	22.9
Pasture	11.3
Residential	2.2
Wetlands	12.9

The Soil Survey Geographic database (SSURGO) (USDA-NRCS) was used to generate the HSPF infiltration parameters (INFILT) used in both models and soil porosity parameters (PCW, PGW, UPGW) used only in the HWT subroutine. Unique values for INFILT, PCW, PGW, and UPGW were assigned to each PLS by taking an area weighted average of the soils intersecting the five different land uses.

HSPF uses stage-volume-discharge relationships or Hydraulic Function Tables (FTABLES) at the outlet of each subwatershed to quantify discharge. For this study, FTABLES were developed using surveyed stream cross-section data. The cross-sectional profiles were assumed to be uniform within a subwatershed. The length of each reach and the cross-sectional profiles formed stage-volume relationships. Manning's equation was then used for each reach, taking into account the surface roughness, slope of the subwatershed, and cross-sectional profile, to calculate discharge and complete the stage-volume-discharge relationship. The Manning's

roughness coefficient (n) was determined for each reach by examining the streambed and banks and comparing them to reference stream channel photos (Chow, 1959).

Observed daily flow data (1950 to 2011) were obtained from the USGS gaging station located on Ahoskie Creek in Ahoskie, NC (02053500). Precipitation and air temperature data were obtained primarily from the National Weather Service Cooperative Observer Program (COOP) station in Murfreesboro, NC (315996) that has an average annual rainfall of 116.05 cm. The Murfreesboro station is located 16 km northwest of the Ahoskie Creek watershed and is the closest active COOP station to the watershed. Gaps or inconsistencies in the Murfreesboro climate record were filled using data from Roanoke Rapids, NC (317319) and Wakefield, VA (448800) located 64 km northwest and 72 km north, respectively, from the Ahoskie Creek watershed. Wind speed, percent sun, and dew point temperature were all obtained from the Norfolk International Airport COOP station (446139).

Monte Carlo Simulation

Monte Carlo (MC) simulations were conducted to assess the performance and quantify the uncertainty of the TGW and HWT models. One hundred thousand (100,000) runs were performed for each model. The MC simulations were executed using Visual Basic.NET program (VBP) script file created specifically for this research. The VBP randomly sampled parameter values from *a priori* distributions to populate the fixed-format HSPF User Control Input file (UCI file) for each MC model run. The template UCI file used in the simulations contained deterministic parameter values for parameters not directly associated with either groundwater subroutine. Only parameters directly associated with each groundwater subroutine were treated as stochastic. The deterministic parameters were calibrated using the TGW model. The VBP compiled the UCI file for each run and executed HSPF. For the comparisons made in this research, the model was run on an hourly time step and the output was average daily flow in cubic meters per second (m^3/s). Model performance was evaluated using multiple objective functions to compare the simulated and observed daily average discharge values. A three-year simulation period (1 January 2006 to 31 December 2009) was used to compare model performance with a one-year startup period run prior to the simulation period to allow state variables to equilibrate. It is important to note that while the baseline TGW model used to calibrate deterministic parameters had lower zone evapotranspiration (LZETP) values that varied

monthly, the two models used in the study had LZETP values that were constant throughout the entire year.

Deterministic Parameter Selection

Physically based parameters and all parameters not involved the groundwater subroutines were treated as deterministic. The HSPF guidance document BASINS Technical Note 6 (USEPA, 2000) was used to assign initial values for these parameters and were refined through calibration. Calibration of the deterministic parameters was performed by running HSPF with the TGW subroutine. Calibration sufficiency was based on comparing simulated and observed discharge at an hourly time step for the period of 1990 through 2009. Model calibration was deemed satisfactory when model output met the optional, more restrictive calibration criteria used in lieu of the default calibration criteria specified by HSPEXP shown in Table 3-2 (Kim *et al.*, 2007).

TABLE 3-2. HSPF Model Performance Calibration Criteria and Calibration Crtieria Values for the Ahoskie Creek Model.

Flow Statistic	Calibration Criteria Percent Error (%)	Calibrated Value (%)
Error in total volume	± 10	9.5
Error in low flow recession	± 0.01	0.01
Error in 50% lowest flows	± 10	9.2
Error in 10% highest flows	± 15	7.2
Error in storm peaks	± 15	-12.1
Seasonal volume error	± 10	10
Summer storm volume error	± 15	13.7

The HSPF infiltration parameter (INFILT) has been shown to be a sensitive parameter in previous studies (Al-Abed and Whiteley, 2002; Donigian, 2002) and is frequently adjusted when calibrating HSPF. For this study, after the initial calibration, INFILT was held constant across all land uses. Holding INFILT constant further isolated the groundwater subroutines from the rest of the HSPF model. Because each MC run used the same weather data and had the same INFILT values, the groundwater subroutines acted with more consistent inputs. In the same vein, the soil porosity parameters (PCW, PGW, UPGW) used in the HWT model were held constant

21

throughout the MC simulation. Again, this allows for more consistency between runs and isolation of the groundwater subroutine and parameters.

TGW Subroutine Parameterization

For the MC simulations, TGW subroutine parameters were treated as stochastic, sampled from uniform distributions over predefined ranges. The parameter ranges were obtained from the "typical" ranges listed in BASINS Technical Note 6 (USEPA, 2000) (Table 3-3).

TABLE 3-3. Stochastic Parameters and Ranges for the Traditional Groundwater (TGW) Subroutine.

Parameter	Parameter Description, Units	Minimum	Maximum
INTFW[†]	Interflow inflow	1.000	3.000
IRC[†]	Interflow recession coefficient, 1/day	0.100	1.000
UZSN[†]	Upper zone nominal storage, in.	3.000	8.000
LZSN	Lower zone nominal storage, in.	0.500	0.700
AGWRC	Groundwater recession rate, 1/day	0.920	0.990
DEEPFR[‡]	Fraction of groundwater lost to deep, inactive aquifers	0.001	0.050
BASETP[‡]	Fraction of remaining ET potential satisfied by baseflow	0.001	0.050
AGWETP	Fraction of remaining ET potential satisfied by active groundwater storage	0.001	0.200
LZETP - Forest	Lower zone ET potential	0.600	0.800
LZETP - Cropland	Lower zone ET potential	0.500	0.700
LZETP - Pasture	Lower zone ET potential	0.400	0.600
LZETP - Residential	Lower zone ET potential	0.200	0.500
LZETP - Wetlands	Lower zone ET potential	0.600	0.900

[†] Each land use was represented in HSPF with a separate parameter but all were given the same initial distribution.
[‡] The forest, cropland, pasture, and residential land uses were all represented in HSPF with one parameter. Wetlands were represented with a separate parameter but given same initial distribution.

HWT Subroutine Parameterization

The HWT subroutine uses the same parameters used in the TGW model plus three additional parameters (SRRC, SREXP, and IFWSC). The surface runoff recession coefficient (SRRC) regulates surface runoff as a function of surface storage. The surface runoff exponent (SREXP) is a multiplicative factor used in the calculation of surface runoff. And the interflow storage capacity (IFWSC) is the maximum interflow storage capacity when the groundwater elevation (GWEL) rises above the upper influence elevation (UELV). These HWT-specific

22

parameters were assigned uniform distributions with initial distribution ranges based on HSPF documentation (Bicknell *et al.*, 2001). The initial distribution ranges were modified for IFWSC and SREXP to insure model stability. The maximum value for the interflow storage capacity parameter (IFWSC) was set at of 6.8 to ensure that the lower zone influence elevation (LELV) would not rise above the upper zone influence elevation (UELV) and generate an error. The upper bound of the surface runoff exponent parameter (SREXP) was set to 2.0 to prevent excess water from being attributed to surface runoff. Table 3-4 shows the stochastic HWT subroutine parameters and distribution ranges.

TABLE 3-4. Stochastic Parameters and Ranges for the High Water Table Groundwater (HWT) Subroutine.

Parameter	Parameter Description	Minimum	Maximum
INTFW	Interflow inflow	1.000	3.000
IRC	Interflow recession coefficient	0.100	1.000
UZSN	Upper zone nominal storage	3.000	8.000
LZSN	Lower zone nominal storage	0.500	0.700
AGWRC	Groundwater recession rate	0.920	0.990
DEEPFR	Fraction of groundwater lost to deep, inactive aquifers	0.001	0.050
BASETP	Fraction of remaining ET potential satisfied by baseflow	0.001	0.050
AGWETP	Fraction of remaining ET potential satisfied by active groundwater storage	0.001	0.200
LZETP - Forest	Lower zone ET potential	0.600	0.800
LZETP - Cropland	Lower zone ET potential	0.500	0.700
LZETP - Pasture	Lower zone ET potential	0.400	0.600
LZETP - Residential	Lower zone ET potential	0.200	0.500
LZETP - Wetlands	Lower zone ET potential	0.600	0.900
SRRC	Surface runoff recession coefficient	0.010	0.990
SREXP	Surface runoff exponent	0.100	2.000
IFWSC	Max Interflow storage capacity	0.100	6.800

[†] Each land use was represented in HSPF with a separate parameter but all were given the same initial distribution.
[‡] The forest, cropland, pasture, and residential land uses were all represented in HSPF with one parameter. Wetlands were represented with a separate parameter but given same initial distribution.

Model Performance Comparison

Model performance was compared using three (3) metrics. The first, Nash-Sutcliffe efficiency (NSE), was used as a global or overall evaluation of performance over the three year

simulation period and was calculated using Equation (1). NSE values can range from -∞ to 1 (Nash and Sutcliffe, 1970). The NSE index was chosen because it is commonly used in hydrologic literature and incorporates both variation about the mean and total variance not described by the model (Jain and Sudheer, 2008). A NSE value was calculated for each run of the two models.

$$E = 1 - \frac{\sum_{t=1}^{T}(Q_o(t) - Q_m(t))^2}{\sum_{t=1}^{T}(Q_o(t) - \overline{Q_o})^2} \tag{1}$$

where

Q_o = observed discharge

Q_m = simulated discharge

t = time, in days

The second metric used to compare the performance of the models provided a measure of how effectively the models simulated the lowest 50% of observed flows. The third metric used was a measure of how effectively the models simulated the highest 10% of observed flows. These metrics are often used by HSPF modelers who use the HSPEXP calibration assessment tool (Lumb *et al.*, 1994) to assess calibration adequacy. When computing the 50% lowest metric, HSPEXP sums the volume of the lowest 50% of observed flows and then sums the simulated flow volumes for those time intervals. The volume of the highest 10% flows are summed similarly. The percent error for both metrics was computed using Equation (2).

$$\% \; Error = \frac{Simulated \; Sum - Observed \; Sum}{Observed \; Sum} \tag{2}$$

Based on the calibration criteria, an error of +/- 10% is considered acceptable for the low flow metric and +/- 15% is acceptable for the high flow metric. While the response variable considered in this study was daily average flow rate instead of total flow volume, the HSPEXP guidance for both metrics was still used to assess model performance. The performance metric results from the two models were compared using the non-parametric Wilcoxen Rank-Sum test.

Behavioral Parameter Set Selection

The concept of behavioral versus non-behavioral parameter sets was first introduced by Spear and Hornberger (1980) and is used in the Generalized Likelihood Uncertainty Estimation (GLUE) method of estimation uncertainty (Beven and Binley, 1992). Parameter sets resulting in model runs that fell above a threshold criteria were considered to be behavioral and provide an acceptable characterization of the watershed (Beven and Binley, 1992).

The criteria for selecting the behavioral parameter sets for each model was determined by examining the cumulative distribution function (CDF) curves of each of the three aforementioned metrics. An upper inflection point in the curve indicated a threshold value below which the CDF value dropped sharply. Parameter sets for the model runs that yielded a metric value above the threshold value were considered behavioral and used in the performance and uncertainty analyses. All other parameter sets were considered non-behavioral and were not considered in any subsequent analyses. This process was conducted for each of the three metrics resulting in three different groups of behavioral parameter sets for both models.

Quantifying Model Uncertainty

The Average Relative Interval Length (ARIL) metric, Equation (3), introduced by Jin et al. (2010), was used to quantify the uncertainty present in the two models resulting from groundwater subroutine structural and parameter uncertainty. The ARIL metric provides a single measure for comparing the relative width of the confidence intervals throughout a specified time period. The 95% confidence intervals were formed using the 2.5% and 97.5% quantiles of the distribution of simulated flow values at each time step. The ARIL metric provides a single number that can be used to compare model performance. Smaller ARIL values indicate narrower confidence bands and less model uncertainty.

$$ARIL = \frac{1}{n} \sum \frac{Limit_{upper,t} - Limit_{lower,t}}{Q_{obs,t}} \tag{3}$$

where

n = Number of time steps

$Limit_{upper,t}$ = 97.5% quantile for time step t

$Limit_{lower,t}$ = 2.5% quantile for time step t

$Q_{obs,t}$ = Observed discharge value for time step t

Assessing Parameter Identifiability

Parameter identifiability, especially in heavily parameterized models such as HSPF, can be challenging during calibration (Brun *et al.*, 2001). Parameters that exhibited a trend towards a particular value or range were more sensitive within the model structure were considered to be "identifiable" (Spear and Hornberger, 1980; Beck, 1987). Parameter identifiability in this study was determined by comparing the *a priori*, uniform distributions of each parameter to their respective posterior distributions. The posterior parameter distributions were determined using the behavioral parameter sets discussed earlier.

A straightforward metric was used to assess parameter identifiability. First, the *a priori* and posterior parameter ranges were rescaled from 0 to 1. The CDF curve of a uniform distribution with a range from 0 to 1 exhibits a 1 to 1 slope. The CDF curve of the rescaled posterior distribution was then compared to the CDF of the *a priori*, uniform distribution. The vertical distance between the two distributions was calculated at 1000 points along the rescaled parameter range. Any parameter that showed a maximum vertical difference between the *a priori* and posterior CDF greater than 0.1, or 10% of the cumulative probability was considered to be identifiable (i.e. non-uniform). While the selected parameter identifiability threshold was arbitrary, it allowed for relative comparison of the number of identifiable parameters between the two models. Figure 3-5 shows an example of a non-identifiable and an identifiable parameter.
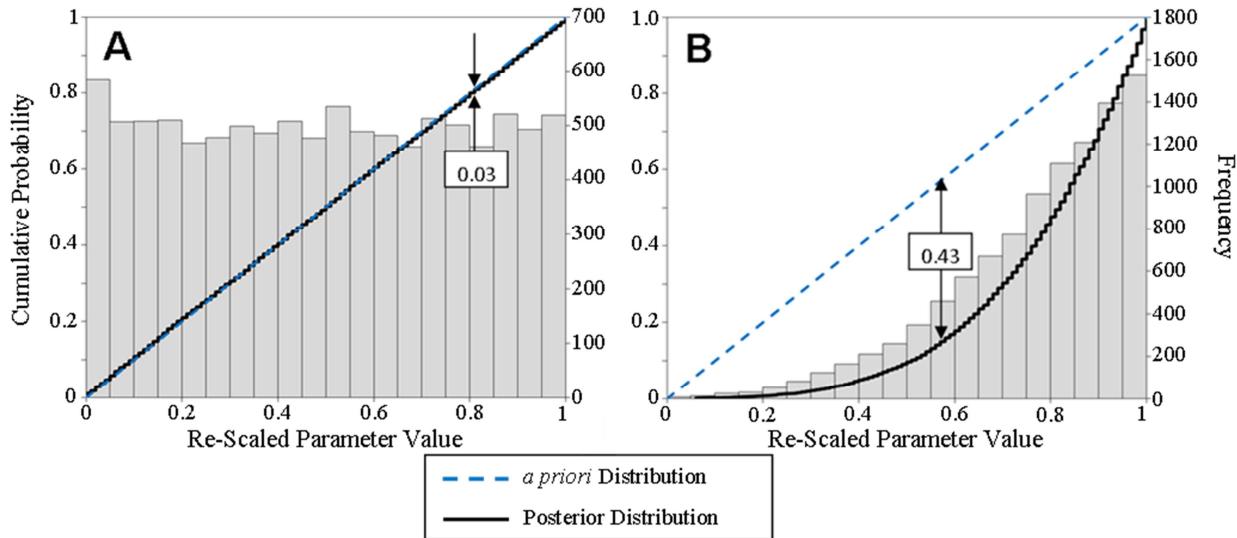
FIGURE 3-5. Examples of a Non-Identifiable (A) and an Identifiable Parameter (B).

This metric was used in lieu of the Kolmogorov-Smirnov (KS) statistic due to the large sample sizes of parameter values in this study. The KS statistic has been used in previous studies to detect differences in *a priori* and posterior parameter distributions (Brun *et al.*, 2001; McIntyre *et al.*, 2005). Because of large posterior parameter distribution sample sizes (10,000 values), even distributions that appeared visually uniform generated KS statistic p-values less than 0.05, indicating a non-uniform distribution.

3.3: RESULTS AND DISCUSSION

Behavioral Parameter Sets

The model performance histograms and resulting CDF curves used to determine behavioral parameter sets based on each of the three performance metrics are shown in Figure 3-6. Four of the six CDF curves exhibited an upper inflection point where the sharp trend towards better model performance decreased at or near the 90th percentile. For consistency, the 90th percentile was chosen as the threshold value used to determine behavioral parameter sets for all three of the performance metrics. Only model runs performing at or above the 90th percentile value for each of the metrics were used in the model performance and uncertainty analyses. Given 100,000 model runs, the 90th percentile threshold generated 10,000 behavioral parameter sets for each of the three performance metrics.
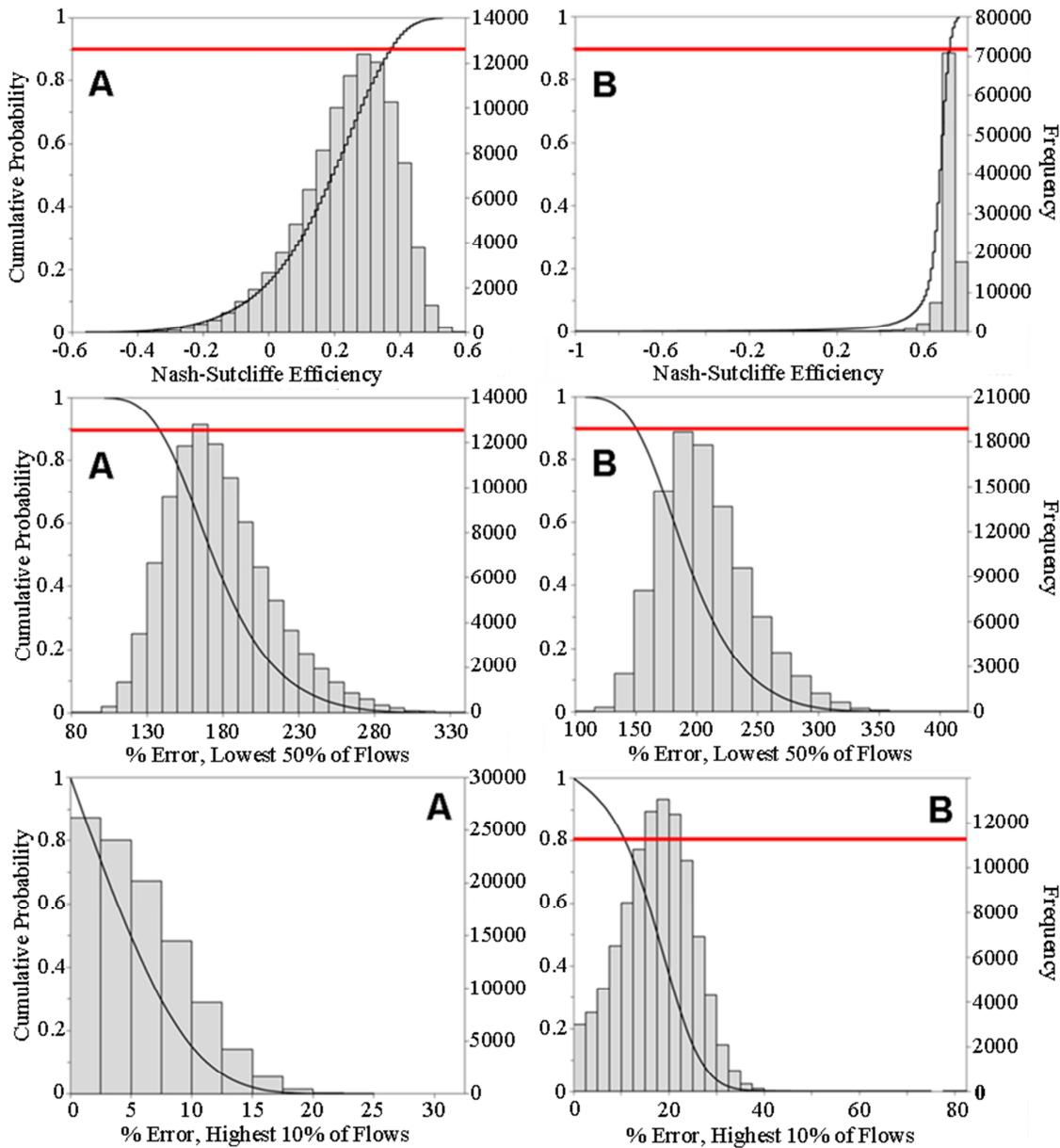
FIGURE 3-6. CDF Curves and Histograms Used to Determine Behavioral Parameter Sets for the TGW (A) and HWT (B) Models. The solid, horizontal line indicates the inflection point.

Performance Comparison

Using only the behavioral parameter sets, the performance of the two models was compared using the NSE and the 50% lowest and 10% highest daily average flow metrics. Because the sample sizes for these comparisons were so large (10,000 values each), any difference in the means of the performance metric distributions was statistically significant when compared using the Wilcoxen Rank-Sum test, i.e. all p-values were much less than 0.05.

As shown in Figure 3-7, the HWT model significantly outperformed the TGW model with regards to NSE. In fact, the maximum NSE achieved by the TGW model (0.60) did not even reach the worst performing behavioral HWT model run (0.72). With respect to the 50% lowest flow metric, the TGW model outperformed the HWT model though both models did a poor job of simulating low flows; the percent errors were greater than 100% for both models. It is important to note, however, that many of the observed flow rates used to assess this metric were quite low (<0.15 m^3/s), so even a small error in the simulated flow rate will result in a large percent error. For example, an observed flow rate of 0.05 m^3/s that is modeled as 0.1 m^3/s results in a percent error of 100%. With respect to the 10% highest flow metric, the TGW again outperformed the HWT model. However, means of the 10% highest flow distributions for both models, fell well within the HSPEXP calibration criteria of +/-15%.
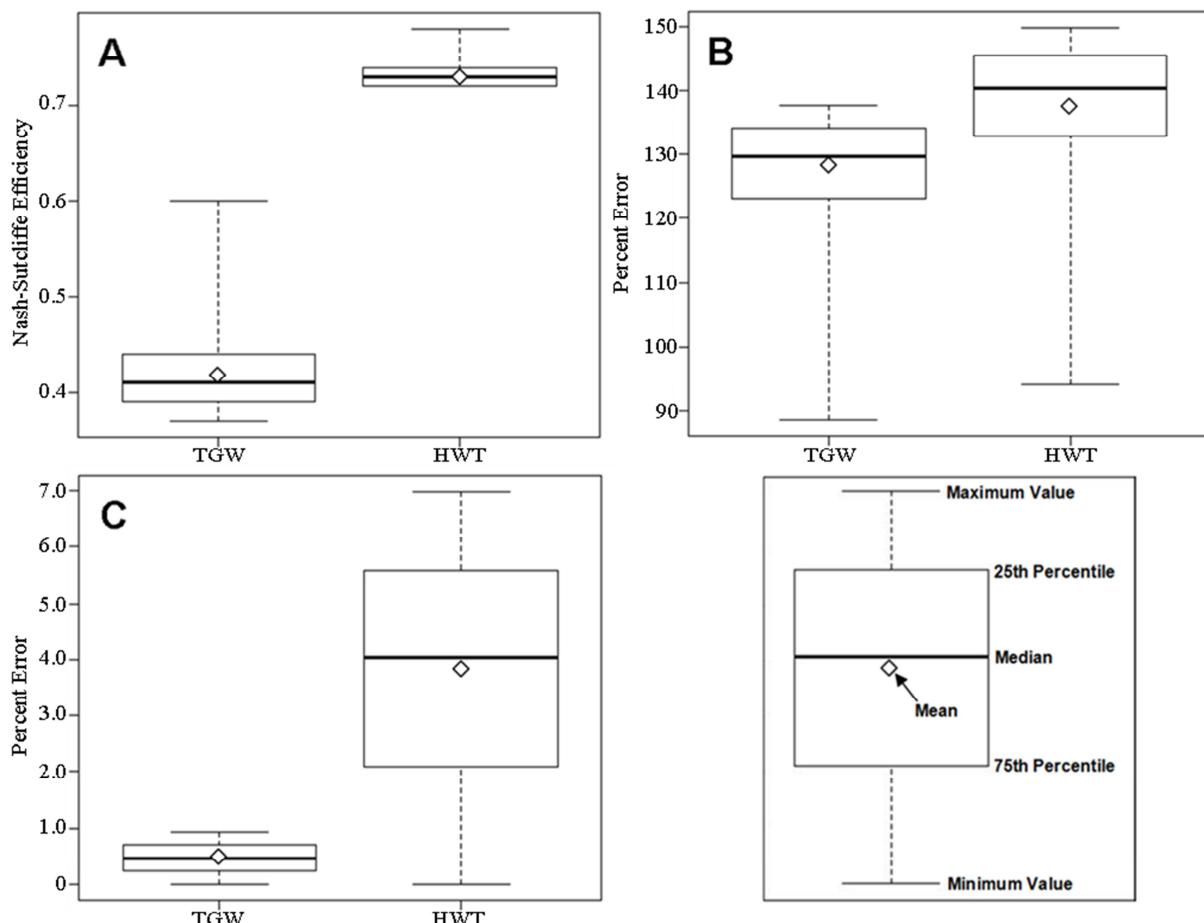

FIGURE 3-7. Boxplots Showing Mean, Median, and Distribution of the TGW and HWT Models When Compared Using Three Metrics: NSE (A), Percent Error When Simulating the Lowest 50% of Daily Average Flow Rates (B), and the Percent Error When Simulating the Highest 10% of Average Flow Rates.

29

Uncertainty Analysis

Model performance variability due to structural and parameter uncertainty was quantified using the ARIL metric and the behavioral model runs for each of the three model performance metrics (Table 3-5). The ARIL values computed for NSE were calculated over the entire three year simulation period (1461 days). The ARIL values for the lowest 50% and highest 10% of daily average flow rates were calculated only during those low and high flow events, 731 and 146 days, respectively.

TABLE 3-5. ARIL Values Calculated Based on the Three Performance Metrics.

Metric	TGW	HWT
NSE	1.898	2.105
50% Lowest Flows	2.826	3.105
10% Highest Flows	0.366	0.447

The ARIL values for the TGW model were smaller than the HWT model for each of the three performance metrics. This is an indication that the TGW model has less overall uncertainty than the HWT throughout the entire time period (NSE) and during the periods of low and high flows. One cause of the greater uncertainty associated with the HWT model is the additional HWT subroutine parameters not used in the TGW subroutine (PCW, PGW, UPGW, IFWSC, SRRC, SREXP). These additional parameters and the complexity of the HWT subroutine created more model input and structural uncertainty which increased the overall model uncertainty.

Parameter Identifiability

Parameter identifiability for the TGW and HWT subroutines was compared by examining the difference between the posterior parameter distributions and the *a priori*, uniform distributions that each parameter was given. Some parameters showed posterior distributions that were uniform while others showed a significant trend towards a particular parameter value or range. Figure 3-8 contains illustrations of selected parameters with posterior distributions where there was a clearly identifiable preference for a particular value or range. Table 3-6 contains the maximum vertical distance between the *a priori* and posterior parameter distributions. Vertical distances equal to or greater than 0.1 are bolded.
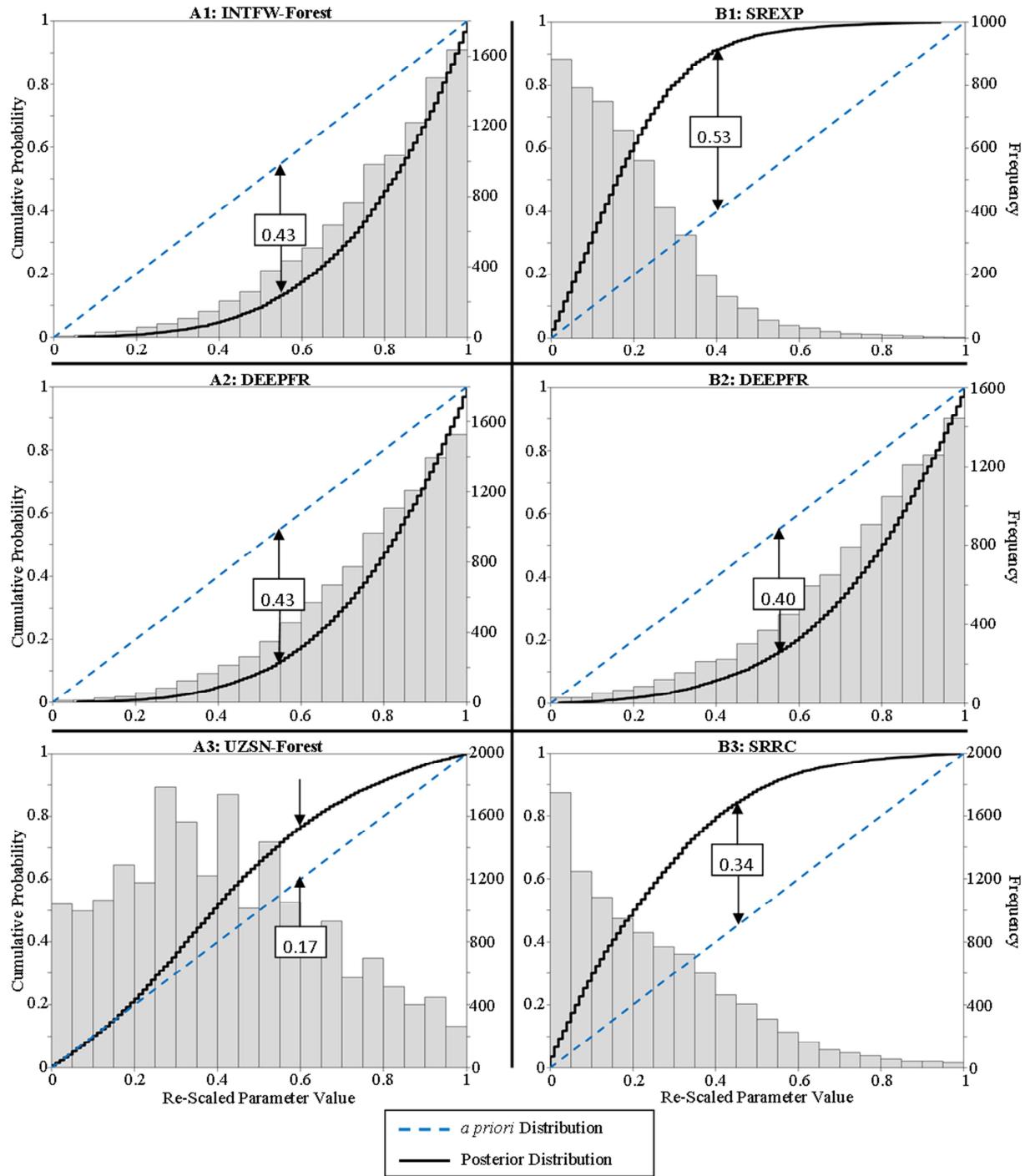
FIGURE 3-8. Illustration of Selected Identifiable Posterior Parameter Distributions for the TGW (A) and HWT (B) Models and the Three Performance Metrics: NSE (1), 50% Lowest Flows (2), and 10% Highest Flows (3).

TABLE 3-6. Maximum Vertical Distance Between Parameter *A Priori* Distributions and Posterior Distributions.

	TGW			HWT		
Parameter	NSE	Lowest 50% of Flows	Highest 10% of Flows	NSE	Lowest 50% of Flows	Highest 10% of Flows
AGWETP	0.05	0.19	0.01	0.01	**0.24**	0.02
AGWETP-Wetlands	0.01	0.03	0.02	0.01	0.03	0.01
AGWRC	0.02	0.09	0.03	0.07	**0.35**	**0.12**
BASETP	**0.12**	**0.33**	0.03	0.01	**0.27**	0.04
BASETP-Wetlands	0.03	0.05	0.02	0.02	0.05	0.02
DEEPFR	0.04	**0.43**	0.03	0.08	**0.40**	0.07
DEEPFR-Wetlands	0.02	0.06	0.01	0.02	0.06	0.02
INTFW-Forest	**0.43**	0.07	0.03	0.02	0.04	0.05
INTFW-Cropland	**0.26**	0.02	0.03	0.02	0.03	0.03
INTFW-Pasture	**0.14**	0.02	0.02	0.01	0.02	0.02
INTFW-Residential	0.04	0.02	0.02	0.01	0.01	0.01
INTFW-Wetlands	**0.14**	0.03	0.02	0.01	0.03	0.03
IRC-Forest	0.06	**0.13**	0.05	**0.13**	0.12	**0.13**
IRC-Cropland	0.05	0.05	0.03	0.06	0.06	0.06
IRC-Pasture	0.02	0.03	0.02	0.04	0.03	0.03
IRC-Residential	0.01	0.01	0.02	0.01	0.02	0.02
IRC-Wetlands	0.03	0.03	0.02	0.05	0.04	0.03
LZETP-Forest	**0.11**	**0.14**	0.02	0.05	**0.14**	0.04
LZETP-Cropland	0.06	0.08	0.02	0.03	0.08	0.04
LZETP-Pasture	0.07	0.09	0.02	0.03	0.08	0.03
LZETP-Residential	0.03	0.04	0.01	0.02	0.04	0.02
LZETP-Wetlands	0.06	0.06	0.02	0.03	0.05	0.02
LZSN	**0.33**	0.09	0.08	0.06	0.07	**0.16**
UZSN-Forest	**0.33**	**0.14**	**0.17**	**0.27**	**0.18**	**0.29**
UZSN-Cropland	**0.18**	0.08	0.06	0.12	0.08	**0.13**
UZSN-Pasture	**0.12**	0.06	0.04	0.06	0.07	0.07
UZSN-Residential	0.03	0.02	0.02	0.02	0.03	0.02
UZSN-Wetlands	0.08	0.04	0.03	0.06	0.05	0.08
IFWSC	–	–	–	0.01	0.01	0.02
SREXP	–	–	–	**0.53**	0.06	**0.34**
SRRC	–	–	–	**0.22**	**0.10**	**0.40**
Identifiable Parameters	10	6	1	5	8	7

*Numbers shown in bold indicate an identifiable parameter

The HWT subroutine yielded 20 identifiable parameters across all three model performance metrics, while the TGW subroutine yielded 17. Considering the fact that the HWT subroutine used three more stochastic parameters than the TGW model, 21.5% of the HWT parameters across all three metrics were found to be identifiable compared to 20.2% of the TGW parameters.

Discussion and Conclusion

The high water table (HWT) groundwater subroutine was developed for HSPF when modeling low-gradient, high water table watersheds (Bicknell *et al.*, 2001). This specialized subroutine requires more time and effort on the part of the modeler to collect the necessary data to parameterize the model and it uses more land use-specific parameters than HSPF's traditional groundwater subroutine (TGW). This study compared the performance and uncertainty of the two models when being applied to Ahoskie Creek, a low-gradient, high water table watershed in northeast North Carolina.

Because of the nature of hydrological models and the heterogeneity of watershed characteristics, the various forms of model uncertainty are absolutely unavoidable (Beven, 2000; Sivapalan, 2003). The balance between model performance and uncertainty should be considered when selecting an appropriate model. Both the TGW and HWT models showed variability due to parameter uncertainty. This uncertainty means that two different modelers can obtain similar model performance with different parameter sets, a phenomenon often referred to as equifinality (Spear *et al.*, 1994; Beven and Freer, 2001; Beven, 2006a). If the TGW and HWT models were used to simulate pollutant loading, differences in the balance between baseflow and surface runoff due to differences in the two groundwater subroutines could lead a modeler or decision maker to attribute pollutant loading to alternate, perhaps incorrect sources. This potential misattribution could then affect decisions such as the type, number, and location of pollution control measures prescribed mitigate the pollution problem.

While this study did not directly address the level of effort required to calibrate either groundwater subroutine, if one subroutine had shown a significantly higher number of identifiable parameters, it follows that calibration of that subroutine would likely be more straightforward. The parameter identifiability analysis revealed that the number of identifiable

groundwater parameters was relatively similar between the two groundwater subroutines, indicating essentially no advantage in calibration effort for either subroutine.

When choosing between the two groundwater subroutines, the context within which HSPF is being applied must be considered. In this study, both models performed very well when simulating the 10% highest daily average flow rates, falling well within the HSPEXP criteria of +/- 15%. Furthermore, both models had similarly high percent error value ranges when simulating the 50% lowest daily average flow rates, greatly exceeding the HSPEXP criteria of +/-10%. The HWT model significantly outperformed the TGW model with respect to NSE values, an indicator of model performance over the entire range of simulated flows. In an application where the accuracy of hydrology simulation is an important part of a larger study involving pollutant loading, such as a TMDL, the increase in uncertainty present in the HWT model may be tolerable because of the improved model performance. If a balance between model performance and uncertainty is sought, the choice of which subroutine to employ is less clear cut.

CHAPTER 4: CONCLUSIONS

The results of this study showed that when applying HSPF in a low-gradient watershed where accuracy of simulated discharges is most important, the HWT subroutine is preferable. If a balance between model performance and uncertainty is sought, it is unclear which groundwater subroutine is preferable.

One of the limitations of the study was the aggregation of the hourly discharge data output by HSPF into average daily time steps. This was done to reduce the computational resource needed to analyze the data. Performing this study using an hourly time step may have reduced the percent errors during low and high flows by decreasing the effect that the aggregating of flows had on the data. In addition, while the model was calibrated using LZETP values that varied monthly, the two models used in the study had LZETP values that were constant throughout the entire year. This lack of seasonal variation could have limited the accuracy of the models.

To fully compare the TGW and HWT models, this study should be carried out in multiple low-gradient, high water table watersheds. This would show if the results found in this study using the Ahoskie Creek watershed would be similar if the two models were applied elsewhere.

LITERATURE CITED

Al-Abed, N. A. and H. R. Whiteley, 2002. Calibration of the Hydrological Simulation Program Fortran (HSPF) Model Using Automatic Calibration and Geographical Information Systems. Hydrological Processes 16:3169-3188.

Beaulac, M. N. and K. H. Reckhow, 1982. An Examination of Land Use - Nutrient Export Relationships. Journal of the American Water Resources Association 18:1013-1024.

Beck, M. B., 1987. Water Quality Modeling: A Review of the Analysis of Uncertainty. Water Resour. Res. 23:1393-1442.

Beven, K., 1989. Changing Ideas in Hydrology -- The Case of Physically-Based Models. Journal of Hydrology 105:157-172.

Beven, K., 1993. Prophecy, Reality and Uncertainty in Distributed Hydrological Modelling. Advances in Water Resources 16:41-51.

Beven, K., 2006a. A Manifesto for the Equifinality Thesis. Journal of Hydrology 320:18-36.

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

Beven, K. and J. Freer, 2001. Equifinality, Data Assimilation, and Uncertainty Estimation in Mechanistic Modelling of Complex Environmental Systems Using the GLUE Methodology. Journal of Hydrology 249:11-29.

Beven, K. J., 2000. *Rainfall-Runoff Modelling - The Primer*. Chichester, UK, John Wiley & Sons Ltd.

Beven, K. J. ed. 2006b. Introduction. *Streamflow Generation Processes.* Wallingford, UK, IAHS Press.

Bicknell, B. R., J. C. Imhoff, J. L. Kittle, T. H. Jobes and A. S. Donigian, 2001. Hydrological Simulation Program - Fortran: HSPF Version 12 User's Manual. U.S. Environmental Protection Agency, Athens, GA.

Borah, D. K., 2011. Hydrologic Procedures of Storm Event Watershed Models: A Comprehensive Review and Comparison. Hydrological Processes 25:3472-3489

Borah, D. K. and M. Bera, 2003. Watershed-Scale Hydrologic and Nonpoint-Source Pollution Models: Review of Mathematical Bases. Transactions of the American Society of Agricultural and Biological Engineers 46:1553-1566.

Brun, R., P. Reichert and H. R. Künsch, 2001. Practical Identifiability Analysis of Large Environmental Simulation Models. Water Resoures Research 37:1015-1030.

Cho, J., V. A. Barone and S. Mostaghimi, 2009. Simulation of Land Use Impacts on Groundwater Levels and Streamflow in a Virginia Watershed. Agricultural Water Management 96:1-11.

Chow, V. T., 1959. *Open-Channel Hydraulics*. New York, McGraw-Hill.

Crawford, N. H. and R. K. Linsley, 1966. Digital Simulation in Hydrology, Stanford Watershed Model IV. Department of Civil Engineering Report Technical Report 39, Stanford University.

Donigian, A. S., 2002. Watershed Model Calibration and Validation: The HSPF Experience. Proceedings of the Water Enviroment Federation 2002:44-73.

Donigian, A. S., Jr., B. R. Bicknell and J. C. Imhoff, 1995. Hydrological Simulation Program - Fortran (HSPF). *In*: *Computer models of watershed hydrology,* V. P. Singh, ed. Water Resources Publications, Colorado, pp. 395-442.

Grayson, R. B., I. D. Moore and T. A. McMahon, 1992. Physically Based Hydrologic Modeling: 1. A Terrain-Based Model for Investigative Purposes. Water Resour. Res. 28:2639-2658.

Green, W. H. and G. A. Ampt, 1911. Studies on Soil Physics. Journal of Agricultural Science 4:1-24

Gupta, H. V., S. Sorooshian and P. O. Yapo, 1998. Toward Improved Calibration of Hydrologic Models: Multiple and Noncommensurable Measures of Information. Water Resources Research 34:751-763.

Haith, D. A. and L. L. Shoenaker, 1987. Generalized Watershed Loading Functions for Stream Flow Nutrients. Journal of the American Water Resources Association 23:471-478.

Homer, C. H., J. A. Fry and C. A. Barnes, 2012. National Land Cover Database.

Horton, R. E., 1933. The Role of Infiltration in the Hydrologic Cycle. Transanctions of the American Geophysical Union 14:446-460.

Horton, R. E., 1945. Erosional Development of Streams and Their Drainage Basins; Hydrophysical Approach to Quantitative Morphology. GSA Bulletin 56:275-370.

Jain, S. K. and K. P. Sudheer, 2008. Fitting of Hydrologic Models: A Close Look at the Nash--Sutcliffe Index. Journal of Hydrologic Engineering 13:981-986.

Jin, X., C.-Y. Xu, Q. Zhang and V. P. Singh, 2010. Parameter and Modeling Uncertainty Simulated by GLUE and a Formal Bayesian Method for a Conceptual Hydrological Model Journal of Hydrology 383:147-155.

Kim, S. M., B. L. Benham, K. M. Brannan, R. W. Zeckoski and J. Doherty, 2007. Comparison of Hydrologic Calibration of HSPF Using Automatic and Manual Methods. Water Resoures Research 43:W01402, doi:01410.01029/02006WR004883

Kizza, M., A. Rodhe, C.-Y. Xu and H. K. Ntale, 2011. Modelling Catchment Inflows into Lake Victoria: Uncertainties in Rainfall–Runoff Modelling for the Nzoia River. Hydrological Sciences Journal 56:1210-1226.

Klemes, V., 1986. Dilettantism in Hydrology: Transition or Destiny? Water Resources Research 22:177-188.

Li, L., J. Xia, C.-Y. Xu and V. P. Singh, 2010. Evaluation of the Subjective Factors of the GLUE Method and Comparison with the Formal Bayesian Method in Uncertainty Assessment of Hydrological Models. Journal of Hydrology 390:210-221.

Lumb, A. M., R. B. McCammon and J. John L. Kittle, 1994. Users Manual for an Expert System (HSPEXP) for Calibration of the Hydrological Simulation Program - Fortran. U. S. Geological Survey Report 94-4168, Reston, VA.

McIntyre, N., B. Jackson, A. J. Wade, D. Butterfield and H. S. Wheater, 2005. Sensitivity Analysis of a Catchment-Scale Nitrogen Model. Journal of Hydrology 315:71-92.

Migliaccio, K. W. and P. Srivastava, 2007. Hydrological Components of Watershed-Scale Models. Transactions of the American Society of Agricultural and Biological Engineers 50:1695-1703.

Muskat, M., 1937. The Flow of Fluids Through Porous Media. Journal of Applied Physics 8:274-283.

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

Refsgaard, J. C., 1997. Parameterisation, Calibration and Validation of Distributed Hydrological Models. Journal of Hydrology 198:69-97.

Said, A., M. Ross and K. Trout, 2007. Calibration of HSPF Using Active Ground Water Storage. In: *World Environmental and Water Resources Congress 2007: Restoring Our Natural Habitat*, K. C. Kabbes (Editor). ASCE, Tampa, Florida, USA, pp. 342-342.

Sheridan, J. M., 2002. Peak Flow Estimates for Coastal Plain Watersheds. Transactions of the American Society of Agricultural and Biological Engineers 45:1319-1326.

Sherman L, K., 1932. Streamflow From Rainfall By the Unit-Graph Method. Engineering News Record 108:501-505.

Singh, V. P., 1995. *Computer Models of Watershed Hydrology*. Colorado, Water Resources Publications.

Singh, V. P. and D. A. Woolhiser, 2002. Mathematical Modelling of Watershed Hydrology. Journal of Hydrology 7:270-292.

Sivapalan, M., 2003. Prediction in Ungauged Basins: A Grand Challenge for Theoretical Hydrology. Hydrological Processes 17:3163-3170.

Spear, R. C., T. M. Grieb and N. Shang, 1994. Parameter Uncertainty and Interaction in Complex Environmental Models. Water Resources Research 30:3159-3169.

Spear, R. C. and G. M. Hornberger, 1980. Eutrophication in Peel Inlet—II. Identification of Critical Uncertainties Via Generalized Sensitivity Analysis. Water Research 14:43-49.

Tetzlaff, D. and U. Uhlenbrook, 2005. Effects of Spatial Variability of Precipitation for Process-Oriented Hydrological Modelling: Results From Two Nested Catchments. Hydrology and Earth System Sciences Discussions 2:119-154.

USDA-ARS, 1977. The Hydrology and Hydrogeology of Ahoskie Creek Watershed, North Carolina: Data and Analysis. U. S. Department of Agriculture - Agricultural Research Service, Washington, D.C.

USDA-NRCS, U. S. Department of Agriculture - Natural Resources Conservation Service Soil Survey Geographic (SSURGO) Database. www.soils.usda.gov/survey/geography/ssurgo, *accessed* January 2011.

USEPA, 2000. BASINS Technical Note 6. United States Environmental Protection Agency, Washington D.C.

USGS, 2006a. U. S. Geological Survey National Elevation Dataset (NED). www.ned.usgs.gov, *accessed* January 2011.

USGS, 2006b. U. S. Geological Survey National Land Cover Database (NLCD). www.landcover.usgs.gov, *accessed* January 2011.

USGS, 2010. U. S. Geological Survey Digital Raster Graphics (DRG). www.topomaps.usgs.gov/drg, *accessed* January 2011.

Valiela, I., J. Costa, K. Foreman, J. Teal, B. Howes and D. Aubrey, 1999. Transport of Groundwater-Borne Nutrients From Watersheds and Their Effects on Coastal Waters. Biodegradation 10:177-197.

Vrugt, J. A., H. V. Gupta, W. Bouten and S. Sorooshian, 2003. A Shuffled Complex Evolution Metropolis Algorithm for Optimization and Uncertainty Assessment of Hydrologic Model Parameters. Water Resources Research 39:1201-1219.

Wagener, T. and H. Gupta, 2005. Model Identification for Hydrological Forecasting Under Uncertainty. Stochastic Environmental Research and Risk Assessment 19:378-387.

Wagener, T. and J. Kollat, 2007. Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. Environmental Modelling & Software 22:1021-1033.

Wagener, T. and A. Montanari, 2011. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. Water Resources Research 47:W06301.

Whitehead, P. and P. Young, 1979. Water Quality in River Systems: Monte-Carlo Analysis. Water Resources Research 15:451-459.

Yeh, P. J. F. and E. A. B. Eltahir, 2005. Representation of Water Table Dynamics in a Land Surface Scheme. Part I: Model Development. Journal of Climate 18:1861-1880.

Zhang, J., A. Said and M. Ross, 2010. Approach Using Active Groundwater Storage for Hydrologic Model Calibration in West-Central Florida. Journal of Irrigation and Drainage Engineering 136:58-62.

APPENDIX A: VB.NET SOURCE CODE

A.1: MAIN MODULE

```vb
Module Main
    Public MCruns As Integer 'number of Monte Carlo runs
    Private intnumparams As Integer 'number of parameters being varied in Monte Carlo
simulation
    Private finalparams(,) As Double 'array containing the final parameter values for
each parameter for each MC run
    Public parmarray As New List(Of parameter) 'List containing name, line #, column #,
length, min, and max for each parameter
    Private timesteps As Integer 'timestep counter
    Private simulatedarray(,) As Double
    Private strParamName As String
    Public Sub Central() 'runs MC
        Dim randomarray(,) As Double 'array of random numbers from 0 to 1 to determine
value of each parameter
        Dim pltgen As Object
        Dim strFplt As String
        Dim RunCommand As String

        getparamarray(randomarray) 'pulls in parmarray, assigns random 0 to 1 value for
each parameter for each MC run
        selection(randomarray) 'uses randomarray to populate finalparams
        For i = 1 To MCruns 'begins MC
            uciwriter(i)
            ChDir("c:\Basins\models\HSPF\bin")
            RunCommand = "winhspflt.exe C:\Users\forreste\Desktop\MonteCarlo\Run" &
Format(i, "#") & "\ahoskie.uci"
            RunShell(RunCommand)
            GetPLTGEN(i)
            objfunctions(simulatedarray, timesteps, i)
            My.Computer.FileSystem.DeleteFile("C:\Users\forreste\Desktop\MonteCarlo\Run"
& Format(i, "#") & "\MET_TIDE.wdm") 'deletes copy of weather data file
            My.Computer.FileSystem.DeleteFile("C:\Users\forreste\Desktop\MonteCarlo\Run"
& Format(i, "#") & "\MC.ech") 'deletes useless .ECH file
            My.Computer.FileSystem.DeleteFile("C:\Users\forreste\Desktop\MonteCarlo\Run"
& Format(i, "#") & "\MC.out") 'deletes useless .OUT file
        Next

        WriteOutputs() 'calls WriteOutputs sub

        MsgBox("I'm Done!")
        End

    End Sub
    Public Sub WriteOutputs() 'writes parameter values for each run
        Dim strFparam As String

My.Computer.FileSystem.CreateDirectory("C:\Users\forreste\Desktop\MonteCarlo\ParameterVal
ues")
        For i As Integer = 1 To intnumparams
            strFparam = "C:\Users\forreste\Desktop\MonteCarlo\ParameterValues\" &
parmarray(i).Name & ".txt"
            Dim paramwrite As System.IO.StreamWriter
            paramwrite = My.Computer.FileSystem.OpenTextFileWriter(strFparam, True,
System.Text.Encoding.ASCII)
```

40

```vbnet
            For j As Integer = 1 To MCruns
                paramwrite.WriteLine(finalparams(i, j))
            Next
            paramwrite.Close()
        Next


    End Sub

    Public Sub uciwriter(ByVal runnumber As Integer) 'populates .uci file
        Dim uci As System.IO.FileStream '.uci file generated for each MC run
        Dim strFin As String 'pathway to .uci template file
        Dim strFout As String 'pathway to populated .uci file
        Dim ctr As Integer 'counter of line number in .uci file
        Dim ctr1 As Integer 'counter of line number in parameter list
        Dim length As Integer 'length of parameter space in .uci file
        Dim spaces As Integer 'empty spaces within parameter space
        Dim ctrX As Integer 'counter within lines of .uci file
        Dim ctrY As Integer
        Dim copy As String
        Dim paste As String
        Dim linetext As String 'text in .uci file template
        copy = "C:\Users\forreste\Desktop\MonteCarlo\MET_TIDE.wdm" 'brings in .wdm file
for new run
        paste = "C:\Users\forreste\Desktop\MonteCarlo\Run" & Format(runnumber, "#") &
"\MET_TIDE.wdm"
        My.Computer.FileSystem.CreateDirectory("C:\Users\forreste\Desktop\MonteCarlo\Run"
& Format(runnumber, "#"))
        System.IO.File.Copy(copy, paste)
        strFin = "C:\Users\forreste\Desktop\MonteCarlo\ahoskie.uci" 'path to .uci file
        strFout = "C:\Users\forreste\Desktop\MonteCarlo\Run" & Format(runnumber, "#") &
"\ahoskie.uci" 'path to individual runs
        Dim objOut As System.IO.StreamWriter
        objOut =
My.Computer.FileSystem.OpenTextFileWriter("C:\Users\forreste\Desktop\MonteCarlo\Run" &
Format(runnumber, "#") & "\ahoskie.uci", True, System.Text.Encoding.ASCII)
        ctr = 1
        ctr1 = 1
        Dim reader As System.IO.StreamReader
        reader = My.Computer.FileSystem.OpenTextFileReader(strFin) 'reads in text from
template .UCI file
        linetext = reader.ReadLine()
        Do Until linetext Is Nothing
            If ctr1 <= intnumparams Then

                If ctr = parmarray(ctr1).LineNumber Then 'if the parameter line number in
the paramarray equals the current line number being written
                    ctrX = parmarray(ctr1).LineNumber 'counter X = the current linenumber
being written
                    ctrY = ctr1 + 1 'counter Y = the linenumber for the next parameter in
paramarray
                    length = Len(parmarray(ctr1).ParamFormat) 'length of parameter space
within .uci
                    spaces = parmarray(ctr1).Length - length  'empty spaces within
parameter space
                    objOut.Write(Left(linetext, (parmarray(ctr1).ColumnNumber - 1) +
spaces)) 'write from template until first varied parameter
```

```vb
                        objOut.Write(Format(finalparams(ctr1, runnumber),
parmarray(ctr1).ParamFormat)) 'write parameter value
                    If ctrY <= intnumparams Then
                        Do While ctrX = parmarray(ctrY).LineNumber 'loop that runs within
the line in .uci file
                            length = Len(Format(finalparams(ctrY, runnumber),
parmarray(ctrY).ParamFormat)) 'length of parameter
                            spaces = parmarray(ctrY).Length - length 'spaces in parameter
                            objOut.Write(Mid(linetext, parmarray(ctrY - 1).ColumnNumber +
parmarray(ctrY - 1).Length, _
                                    parmarray(ctrY).ColumnNumber - 1 -
(parmarray(ctrY - 1).ColumnNumber - 1 + parmarray(ctrY - 1).Length))) 'write in parameter
                            For i As Integer = 1 To spaces
                                objOut.Write(" ")
                            Next
                            objOut.Write(Format(finalparams(ctrY, runnumber),
parmarray(ctrY).ParamFormat))

                            ctrY += 1
                            If ctrY > intnumparams Then
                                Exit Do
                            End If
                        Loop
                    End If

                    objOut.WriteLine(Right(linetext, Len(linetext) - parmarray(ctrY -
1).ColumnNumber + 1 - parmarray(ctrY - 1).Length)) 'writes rest of line

                    ctr1 = ctrY

                Else
                    objOut.WriteLine(linetext) 'if parameter line # doesn't match .uci
line, write from template
                End If
            Else
                objOut.WriteLine(linetext)
            End If
            ctr += 1
            linetext = reader.ReadLine() 'linetext = full line of text from .uci template
        Loop

        reader.Close()
        objOut.Close()

    End Sub
    Public Sub selection(ByVal randomarray(,) As Double) 'Routine to select random
parameter value
        ReDim finalparams(intnumparams + 1, MCruns + 1)
        For i = 1 To intnumparams
            For j = 1 To MCruns
                finalparams(i, j) = (randomarray(parmarray(i).RandIndex, j) *
(parmarray(i).Max - parmarray(i).Min)) + parmarray(i).Min 'selects random value from
range, assuming uniform dist
            Next
        Next

        'Monte Carlo parameter selection
    End Sub
```

```vb
    Public Sub getparamarray(ByRef randomarray(,) As Double) 'populates parameter array
based on random number (0-1) array
        Dim objFileSys As Object
        Dim objFile As Object
        Dim strFname As String
        Dim linetext As String
        strFname = "C:\Users\forreste\Desktop\MonteCarlo\inputparams.txt" 'path to
parameter input file
        objFileSys = CreateObject("Scripting.FileSystemObject")
        Dim reader As System.IO.StreamReader
        reader = My.Computer.FileSystem.OpenTextFileReader(strFname)
        reader.ReadLine()
        intnumparams = reader.ReadLine() 'number of parameters being varied in MC
simulation
        reader.ReadLine()
        Dim current As String
        Dim temp As Array
        Dim newParameter As New parameter
        Dim maxIndex As Integer
        maxIndex = 0
        parmarray.Add(newParameter) 'adds each new parameter to array
        For i = 1 To intnumparams
            current = reader.ReadLine()
            temp = Split(current, ",")
            newParameter = New parameter
            parmarray.Add(newParameter)
            With parmarray(i) 'reads in information for each parameter
                .Name = temp(0)
                .LineNumber = temp(1)
                .ColumnNumber = temp(2)
                .Length = temp(3)
                .Min = temp(4)
                .Max = temp(5)
                .ParamFormat = temp(6)
            End With

            Dim intExistIndex As Integer
            strParamName = parmarray(i).Name
            intExistIndex = parmarray.FindIndex(AddressOf FindByName)
            If intExistIndex >= maxIndex + 1 Then
                'code to execute if the parameter is not already in use
                parmarray(i).RandIndex = maxIndex + 1
                maxIndex = maxIndex + 1
            Else
                'code to execute if the parameter is already in use
                'this would be something that assigns the new RandIndex
                parmarray(i).RandIndex = parmarray(intExistIndex).RandIndex

            End If


        Next

        Dim rnd As New Random() 'creates random number array
        ReDim randomarray(intnumparams + 1, MCruns + 1)
        For ctr1 As Integer = 1 To intnumparams
            For ctr As Integer = 1 To MCruns
                randomarray(ctr1, ctr) = (rnd.Next(0, 101)) / 100
```

```vb
            Next ctr
        Next ctr1

        reader.Close()

    End Sub
    Function FindByName(ByVal parm As parameter) As Boolean  'FindByName function used as
part of predicate function to assign index values
        If parm.Name = strParamName Then
            Return True
        Else
            Return False
        End If
    End Function

    Public Sub objfunctions(ByVal simulatedarray(,) As Double, ByVal tsteps As Integer,
ByVal runnumber As Integer) 'computes objective functions
        Dim strFname As String
        Dim observed() As Double
        Dim nsenum As Integer
        Dim nsenumx As Integer
        Dim nsedenom As Integer
        Dim nsedenomx As Integer
        Dim nse As String
        Dim strFnse As String
        Dim ssex As Integer
        Dim sse As Integer
        Dim strFsse As String
        Dim sstotx As Integer
        Dim sstot As Integer
        Dim rmse As Integer
        Dim strFrmse As String
        Dim Le As String
        Dim strFLe As String

My.Computer.FileSystem.CreateDirectory("C:\Users\forreste\Desktop\MonteCarlo\ObjectiveFun
ctions")
        Dim Lewriter As System.IO.StreamWriter
        strFLe = "C:\Users\forreste\Desktop\MonteCarlo\Objectivefunctions\Le.txt"
        Lewriter = My.Computer.FileSystem.OpenTextFileWriter(strFLe, True,
System.Text.Encoding.ASCII)
        Dim nsewriter As System.IO.StreamWriter
        strFnse = "C:\Users\forreste\Desktop\MonteCarlo\Objectivefunctions\nse.txt"
        nsewriter = My.Computer.FileSystem.OpenTextFileWriter(strFnse, True,
System.Text.Encoding.ASCII)
        Dim ssewriter As System.IO.StreamWriter
        strFsse = "C:\Users\forreste\Desktop\MonteCarlo\Objectivefunctions\sse.txt"
        ssewriter = My.Computer.FileSystem.OpenTextFileWriter(strFsse, True,
System.Text.Encoding.ASCII)
        Dim rmsewriter As System.IO.StreamWriter
        strFrmse = "C:\Users\forreste\Desktop\MonteCarlo\Objectivefunctions\rmse.txt"
        rmsewriter = My.Computer.FileSystem.OpenTextFileWriter(strFrmse, True,
System.Text.Encoding.ASCII)
        Dim reader As System.IO.StreamReader
        strFname = "C:\Users\forreste\Desktop\MonteCarlo\observedtimeseries.txt"
        reader = My.Computer.FileSystem.OpenTextFileReader(strFname)
        ReDim observed(tsteps + 1)
```

```vbnet
        nsenum = 0
        nsedenom = 0
        sse = 0
        sstot = 0
        For i As Integer = 1 To tsteps
            observed(i) = CDbl(reader.ReadLine())
            nsenumx = (observed(i) - simulatedarray(i, runnumber)) ^ 2  'Nash Sutcliffe
efficiency
            nsedenomx = (observed(i) - 52.529) ^ 2
            nsenum = nsenum + nsenumx
            nsedenom = nsedenom + nsedenomx

            ssex = (observed(i) - simulatedarray(i, runnumber)) ^ 2 'Sum of squared
errors
            sse = ssex + sse


            sstotx = (observed(i) - 52.529) ^ 2 'Total SS
            sstot = sstotx + sstot
            'compute objective functions
        Next
        nse = Format(1 - (nsenum / nsedenom), "#.00")
        rmse = ((sse / tsteps)) ^ 0.5
        Le = (sse / tsteps) ^ -2
        nsewriter.WriteLine(nse)
        ssewriter.WriteLine(sse)
        rmsewriter.WriteLine(rmse)
        Lewriter.WriteLine(Le)
        'write objective function files

        nsewriter.Close()
        ssewriter.Close()
        rmsewriter.Close()
        Lewriter.Close()
        reader.Close()
    End Sub

    Private Sub RunShell(ByVal cmdline As String)          'Function to define RunShell
        'such as VB can wait until shell
        'completes processing
        Shell(cmdline, AppWinStyle.Hide, True)


    End Sub

    Public Sub GetPLTGEN(ByVal runnumber As Integer) 'reads in HSPF output data from
PLTGEN file
        'reads to a specific line based on how many years of data you want
        Dim strFplt As String
        Dim text As String
        Dim reader As System.IO.StreamReader
        Dim reader2 As System.IO.StreamReader
        Dim steps As Integer
        strFplt = "C:\Users\forreste\Desktop\MonteCarlo\Run" & Format(runnumber, "#") &
"\pltgen.plt"
        reader = My.Computer.FileSystem.OpenTextFileReader(strFplt)
        reader2 = My.Computer.FileSystem.OpenTextFileReader(strFplt)
```

```vb
        Dim strFtimes As String
        strFtimes = "C:\Users\forreste\Desktop\MonteCarlo\timeseries.txt"
        Dim times As System.IO.StreamWriter
        times = My.Computer.FileSystem.OpenTextFileWriter(strFtimes, True,
System.Text.Encoding.ASCII)


        Dim i As Integer
        i = 1
        For i = 1 To 391
            reader.ReadLine()
        Next
        timesteps = 0
        text = reader.ReadLine()
        Do Until text Is Nothing
            timesteps = timesteps + 1
            times.Write(Mid(text, 25))
            times.Write(",")
            text = reader.ReadLine()
        Loop
        times.Write(Environment.NewLine)

        times.Close()
        reader.Close()

        ReDim simulatedarray(timesteps + 1, MCruns + 1)
        Dim k As Integer
        k = 1
        For k = 1 To 391
            reader2.ReadLine()
        Next
        steps = 0
        text = reader2.ReadLine()
        Do Until text Is Nothing
            steps = steps + 1
            simulatedarray(steps, runnumber) = Mid(text, 25)
            text = reader2.ReadLine()
        Loop
        reader2.Close()


    End Sub
End Module
```

A.2: START FORM

```
Public Class StartForm

    Private Sub cmdStart_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles cmdStart.Click
        Me.Hide()
        Main.MCruns = Me.NumberofRuns.Text
        Main.Central()
    End Sub


    Private Sub StartForm_Load(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles MyBase.Load

    End Sub
End Class
```

A.3: PARAMETER CLASS

```
Public Class parameter
    Public Name As String   'Parameter name
    Public LineNumber As Integer  'Parameter line number in .uci file
    Public ColumnNumber As Integer   'Parameter column number in .uci file
    Public Length As Integer  'spaces occupied by parameter in .uci file
    Public Min As Double  'minimum parameter value in uniform distribution
    Public Max As Double  'maximum parameter value in uniform distribution
    Public ParamFormat As String
    Public RandIndex As Integer

End Class
```

APPENDIX B: DETERMINISTIC PARAMETER VALUES

Parameter	Definition	Units	Calibrated Value	Function
PERLND				
PWAT-PARM2				
FOREST	Fraction forest cover	none	1.0 forest, 0.0 other	Forest cover
INFILT	Index to infiltration capacity	in/hr	0.023-0.109[a]	Soil and cover conditions
LSUR	Length of overland flow	feet	452-513[a]	Topography
SLSUR	Slope of overland flowplane	none	0.0039-0.0322[a]	Topography
KVARY	Groundwater recession variable	1/in	0	Calibrate
PWAT-PARM3				
PETMAX	Temp below which ET is reduced	deg. F	40	Climate, vegetation
PETMIN	Temp below which ET is set to zero	deg. F	35	Climate, vegetation
INFEXP	Exponent in infiltration equation	none	2	Soil properties
INFILD	Ratio of max/mean infiltration capacities	none	2	Soil properties
PWAT-PARM4				
CEPSC	Interception storage capacity	inches	0.06-0.27[b]	Vegetation
NSUR	Mannings' n (roughness)	none	0.45 forest; 0.30 pasture and residential; 0.27 crop; 0.40 wetlands	Land use, surface condition
IMPLND				
IWAT-PARM2				
LSUR	Length of overland flow	feet	497	Topography
SLSUR	Slope of overland flowplane	none	0.0116	Topography
NSUR	Mannings' n (roughness)	none	0.1	Land use, surface condition
RETSC	Retention/interception storage capacity	inches	0.1	Land use, surface condition
IWAT-PARM3				
PETMAX	Temp below which ET is reduced	deg. F	40	Climate, vegetation
PETMIN	Temp below which ET is set to zero	deg. F	35	Climate, vegetation
RCHRES				
HYDR-PARM2				
KS	Weighting factor for hydraulic routing		0.5	

[a]Varies with land use, [b]Varies by month and with land use, [c]Varies monthly