2012-05-04

New Methods for Improving the Watershed-Scale Modeling of In-Stream Pathogen Indicator Bacteria

Jeffrey J. Iudicello
University of Miami, jeffiudicello@gmail.com

UNIVERSITY OF MIAMI

NEW METHODS FOR IMPROVING THE WATERSHED-SCALE MODELING OF IN-STREAM PATHOGEN INDICATOR BACTERIA

By

Jeffrey J. Iudicello

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

May 2012

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

NEW METHODS FOR IMPROVING THE WATERSHED-SCALE MODELING OF IN-STREAM PATHOGEN INDICATOR BACTERIA

Jeffrey J. Iudicello

Approved:

David Chin, Ph.D. Professor of Civil & Environmental Engineering	Terri A. Scandura, Ph.D. Dean of the Graduate School
James Englehardt, Ph.D. Professor of Civil & Environmental Engineering	Victor Pestien, Ph.D. Associate Professor of Mathematics
Fernando Miralles-Wilhelm, Ph.D. Professor, Department of Earth and Environment Florida International University	

IUDICELLO, JEFFREY (Ph.D., Civil Engineering)
New Methods for Improving the Watershed-Scale Modeling of In-Stream Pathogen Indicator Bacteria (May 2012)

Abstract of a dissertation at the University of Miami.

Disssertation supervised by Professor David Chin.
No. of pages in text. (144)

Excessive bacteria levels are the leading cause of impairment in U.S. water bodies. This dissertation looked at the use of watershed-scale computer models to predict in-stream bacteria concentrations. The study site was the Little River Experimental Watershed (LREW) in Tifton, GA, and fecal coliform fate and transport models were built for four of the LREW catchments over the period Jan 1996 - Dec 2002. A multi-model approach was used in the study to examine the current capacity of industry-standard models and to avoid conclusions unique to a particular catchment. Three models were examined: HSPF, SWAT, and a new model based on the principles of hydrograph separation called the Characteristic Concentration (CC) model. Sensitive hydrology and water-quality parameters were identified in HSPF and SWAT and a response-surface iterative scheme was used to calibrate the sensitive parameters of the models, while a simpler calibration method was used for the 2-parameter CC model. Model performance, both hydrology and water-quality, was evaluated by the Nash-Sutcliffe statistic. The research was conducted in three substudies: an examination of the three models' performance in modeling bacteria concentrations in all four catchments, an examination in HSPF and SWAT to determine the relationship between model performance and the hydrologic state of the watershed, and an examination of model combination possibilities to further analyze model performance and provide methods for combining model output to maximize model results beyond what the individual models could achieve.

The research revealed that while the hydrology components of HSPF and SWAT could be considered strong the water-quality components were not as strong. The model parameters describing in-stream bacteria processes were consistently more sensitive than parameters describing terrestrial processes of bacteria, a result that was reinforced when considering the hydrologic state of the watersheds. A Latin Hypercube analysis revealed that parameter uncertainty is significant in the models, but that structural uncertainty resulting from the model process equations is the dominant source of uncertainty in model predictions. The model combination methods were able to provide an improved set of model predictions and showed that in some cases the use of a single calibrated model may still not be the best representation of a watershed. In all cases the CC model performed comparably or better than HSPF and SWAT, and provided a new model framework for analyzing the environmental fate and transport of bacteria. The CC model is worth using in future modeling studies and may be particularly useful in model combination applications since it is comparatively much simpler to use and less data-intensive than either HSPF or SWAT.

DEDICATION

To Micah, Sam, Gabriel, and #4.

This is for you.

And to all the places in the world like Wehr's Dam, Lehigh County, PA, where people can go to interact with and be inspired by land and water.

But blessed is the man who trusts in the LORD,

whose confidence is in him.

He will be like a tree planted by the water

that sends out its roots by the stream.

It does not fear when heat comes;

its leaves are always green.

It has no worries in a year of drought

and never fails to bear fruit.

Jeremiah 17:7-8 (NIV)

By wisdom the LORD laid the earth's foundations,

by understanding he set the heavens in place;

by his knowledge the deeps were divided,

and the clouds let drop the dew.

Proverbs 3:19-20 (NIV)

ACKNOWLEDGEMENTS

In the spirit of *God, Family, and everything else*, I first and foremost would like to thank my heavenly Father for providing me with the ability and opportunity to study His creation in such detail. The knowledge and instincts I have gained in the past five years through studying the environment have revealed to me more of His character, wisdom, and power, and I consider myself blessed to spend the rest of my life in a career field where I work to understand the environment more.

I would also like to thank my immediate and extended family, because there is no way I could have survived this process without their help, encouragement, and prayers. In particular, my lovely wife Elsie has been with me through the thick and the thin of the last 5 years. I could not have made it without her.

On the technical side of things, I must acknowledge: Paul Duda of AquaTerra in Atlanta, Georgia; Tom Jobes of the St. Johns River Water Management District in Paletka, Florida; Nancy Sammons and Mike White of the USDA-ARS in Temple, Texas; Chris Hanson in the Department of Geography at the University of Miami; and Shihab Asfour and Khaled Zakaria at the University of Miami who graciously provided extra computer resources for the modeling efforts.

I must also thank the CAE department for funding me as a TA for 5 years, and to my Ph.D. committee members for their time and expertise both in and out of the classroom. And, of course, I am greatly indebted to my advisor, Dr. David Chin, for being my guide through my entire time at UM. His wealth of knowledge and experience in all things engineering is endless, and I could not have asked for a better mentor through my graduate studies. I know that the skills I developed under his tutelage will be with me for the rest of my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 BACKGROUND

Excessive concentrations of bacteria are under investigation worldwide as a significant problem in water bodies. In the United States, for example, elevated levels of pathogen-indicator bacteria are the leading cause of impaired waters (USEPA 2008), while the topic has drawn attention in other countries as well (e.g., Jamieson et al. 2005; Tyrrel and Quinton 2003; Obiri-Danso and Jones 1999; Bougeard et al. 2011). High indicator bacteria levels may imply fecal contamination and an increased risk of exposure to pathogens like *Cryptosporidium*, *Salmonella*, *Giardia* and *Escherichia Coli*. Fecal contamination is of notable concern as evidenced by an outbreak in Milwaukee in 1993, where *Cryptosporidium* of fecal origin contaminated drinking water and led to roughly 403,000 illnesses, 50 fatalities, and nearly $100 M in total medical costs and productivity losses (Corso et al. 2003; Hoxie et al. 1997). Similarly, in May of 2000, contamination of the Walkerton, Canada municipal water system with *Escherichia coli* 0157:H7 and *Campylobacter jejuni* led to seven fatalities and 2300 people requiring medical attention. The most likely cause of contamination of the municipal well water was determined

to be the infiltration of manure-associated bacteria to the aquifer used for drinking water (Unc and Goss 2004).

Water-quality standards for bacteria are written in terms of *indicator organisms*, and the ideal criteria for describing an indicator organism are (Bitton 2005):

1. It should be one of the intestinal microflora of warm-blooded animals.
2. It should be present when pathogens are present, and absent in uncontaminated samples.
3. It should be present in greater numbers than the pathogen.
4. It should be at least equally resistant as the pathogen to environmental insults and to disinfection in water and wastewater treatment plants.
5. It should not multiply in the environment.
6. It should be detectable by means of easy, rapid, and inexpensive methods.
7. The indicator organism should be nonpathogenic.

While no organism perfectly meets these criteria, the most common indicators are members of the Total Coliform group, defined as the aerobic and facultative anaerobic, gram-negative, nonspore-forming, rod-shaped bacteria that ferment lactose with gas production within 48 hours at 35°C (Bitton 2005). Most current freshwater water-quality standards use Fecal Coliform bacteria (FC) as the indicator, where FC are all coliforms that can ferment lactose at 44.5 °C. FC are indicative of fecal material from warm-blooded animals (Bitton 2005) and can account for upwards of 98% of the total coliforms in feces from warm-blooded animals (Guber et al. 2006). *Escherichia coli* (EC) is a member of the FC group, and some states are now in transition to EC standards. In Texas, for example, the standard indicator organism for bacteria was changed from FC to EC for freshwater and *Enterococci* for saltwater (LaWare and Rifai 2006), although FC is still used when adequate data of the updated indicator organisms is not available. In some cases regression equations are available to convert between FC and EC concentrations (e.g., VADEQ 2004). The most common unit of expressing quantities of bacteria cells is *Colony Forming Units* (CFU).

Sources of in-stream pollutants are generally grouped into two categories: point sources (PS) and non-point sources (NPS). Point sources are discrete sources such as effluent discharges from wastewater treatment plants (WWTP). Because of the discrete, localized nature of PS, legislative efforts such as the National Pollutant Discharge Elimination System (USEPA 2001) have succeeded in quantifying PS loads, reducing them to acceptable levels, and maintaining them at those levels. Non-point sources, on the other hand, generate loads from distributed sources and are generally considered to be the more problematic source (Jamieson et al. 2004). For example, runoff from rainfall is known to transport materials from land surfaces to water bodies, but the exact pollutant loads depend on a variety of factors such as rainfall amount and intensity, constituent solubility, land use, and soil properties such as infiltration capacity and saturation level. In agricultural areas, land-use practices such as the application of manure as fertilizer can contribute significantly to NPS bacteria loads. Feces of livestock such as sheep, chickens, and cattle typically contain loads of 16, 1.3, and 0.23×10^6 FC per gram, respectively (Metcalf & Eddy 2003). Similarly, fresh, pasture-collected feces of sheep, cattle, and horses have been shown to contain loads of 11.2, 7.59, and 0.617×10^5 EC per gram, respectively (Weaver et al. 2005). Such quantities are sometimes collected from feed lots and barns and stored by land managers for later application, but when not collected the quantities may be deposited by the animals on-land or directly into surface waters depending on grazing patterns and water access (e.g. Shirmohammadi et al. 1997). Once released from the source and depending on soil conditions, bacteria may be transported to surface waters via overland flow (Tyrrel and Quinton 2003; Muirhead et al. 2006b); be transported via infiltration and percolation to subsurface soils after land-surface deposition (Muirhead et al. 2006a); travel horizontally in subsurface soils (Gagliardi and Karns 2000; Hunter et al. 1992; McMurry et al. 1998); survive in the soil matrix for extended periods of time, even through extreme cold and freeze-thaw cycles (Ishii et al. 2006; Byappanahalli et al. 2003); survive in streambed sediment (Jamieson et al. 2005; Crowther et al. 2002); and/or demonstrate significantly

varying levels of persistence and survivability depending on environmental factors (Bonjoch et al. 2011). The Total Maximum Daily Load (TMDL) program was developed as an additional policy-based measure to allocate acceptable loadings throughout a watershed (e.g. USEPA 2001). Legislative efforts such as NPDES and TMDL have paved the way for addressing pathogenic contamination, but much work is needed to truly mitigate the problem; as of 2009 there were over 10,000 TMDLs in the U.S. approved or being developed for pathogens alone (USEPA 2009).

A variety of mathematical models have been developed to analyze the fate and transport of environmental pollutants at the watershed scale. The models are powerful tools for use in environmental monitoring as they provide users the flexibility of inputting site-specific landscape and meteorological data and the ability to model pollutant loadings under various land-use scenarios. For example, models allow users to specify livestock and poultry defecation rates and fecal loadings, manure applications, point sources in rivers, and types of land use, to examine the overall effect that different environmental processes have on nearby water quality. The models typically require significant amounts of input data such as meteorological, hydrological, topographical, geographical, geological, and contaminant data. Although these models can be powerful tools, their deficiencies and limitations are well documented (e.g., Oliver et al. 2009; Shirmohammadi et al. 2006; Hantush and Kalin 2008; Haan 1989). The most significant hindrance is the inability to program or build a model able to perfectly replicate complex environmental phenomena, resulting in model predictions laden with a degree of uncertainty. One way of reducing the limitations of inherent uncertainty in a model is by using more than one model; the implementation of several differing models at the same site will provide a variety of results, and a proper synthesis of the model outputs may reveal insight into actual processes that might not be evident by the use of a single model. Additionally, combining models may produce new model predictions that fit data better than the predictions of an individual model and provide additional insight.

While certain model input quantities are relatively easy to measure and use with confidence, such as average daily temperature or humidity at a meteorological station, average daily streamflow, or land use (via GIS) in a catchment, other quantities have much less certainty. This is particularly true with a water-quality constituent like bacteria, where data collection requires a site visit, a grab sample, and several hours of laboratory analysis before concentrations are enumerated. Moreover, whereas meteorological and hydrological data are frequently recorded on a sub-hourly, hourly, or daily time-scale, the cost- and time-intensive nature of collecting water samples for bacterial analysis means it is typically done on a weekly or monthly time-scale, and the grab samples are sometimes assumed to be daily-average values (e.g. Jia and Culver 2008). For practical reasons, however, these grab samples are rarely taken during or directly after a significant rainfall event, when it is suspected that in-stream bacterial fate and transport dynamics are highly active (e.g. McDonald et al. 1982). Inevitably, some bacteria samples are taken during baseflow conditions and may not appropriately represent the contributions from land surfaces via runoff. (In an attempt to address the lack of data during rainfall events, Davies-Colley et al. (2008) used an automated bacteria sampling device installed near a hydrological weir that could be triggered remotely by cell phone, and took samples every 2 hours for 48 hours.) The difficulties associated with attaining and using bacteria data remain significant issues in the development of bacteria models, as reflected by the limited data available in most bacteria modeling studies and the relatively low levels of accuracy achieved by the models compared to other water quality constituents (Baffaut and Sadeghi 2010). Nonetheless, only through further modeling efforts can model uncertainty be understood and remedied to better address the link between PS, NPS, and in-stream bacteria concentrations.

1.2 PURPOSE AND SIGNIFICANCE

The purpose of the research presented in this dissertation is twofold: first, to examine the current capacity of watershed-scale modeling of bacteria, and second to present several new

methods of modeling bacteria such as a new model, a new seasonal analysis for current models, and new model-combination possibilities to maximize model predictions relative to the data. Two complex, state-of-the-art, watershed-scale, industry-standard models with U.S. government approval were used: HSPF (EPA- Bicknell et al. 2001) and SWAT (USDA-Neitsch et al. 2005). The new model presented is the Characteristic Concentration (CC) model (Chin, 2011a), and is based on the principles of hydrograph separation. The study site is a USDA experimental watershed in south-central Georgia, USA, and HSPF (v. 2.3), SWAT (v.2009), and CC models were built for four subcatchments within the watershed to minimize localized effects unique to a single watershed. The research was divided into three studies: 1) all three models were tested to evaluate their performance in modeling bacteria; 2) an investigation was conducted in HSPF and SWAT to determine the relationship between model performance and the hydrologic state of the watershed; and 3) model combination possibilities were considered in an attempt to both provide new modeling options and to maximize model results beyond what the individual models could achieve.

The work presented here is significant because it directly addresses the leading environmental water pollutant in the nation. The research methods and results are directly applicable to current work, provide tools to improve current modeling capacity, and highlight important fields for future research. The net result is an enhanced capacity to address a major concern of both human and environmental health, to ensure that the risk of tragedies like what happened in Milwaukee and Walkerton is drastically reduced, both domestically and internationally.

1.3 ORGANIZATION AND CONVENTIONS

The remainder of the dissertation is structured as follows: Chapter 2 provides a literature review on the complexity of environmental bacteria processes; Chapter 3 introduces the site location and provides a description of the models; Chapter 4 addresses the theory of model

calibration and explains how the models used in the research were calibrated; Chapter 5 presents the first study; Chapter 6 presents the second study; Chapter 7 presents the third study; and Chapter 8 provides an executive summary.

Throughout the course of this dissertation, HSPF and SWAT will frequently be discussed together. As a convention, HSPF will always be addressed first, based simply on alphabetical order and not as a reflection of model capacity or preference. In cases when all three models are being discussed, the CC model will be discussed last.

Chapter 2

Literature Review

2.1 BACKGROUND

In the natural environment, bacteria exhibit a wide variety of complex behavioral patterns that make the analysis of environmental fate and transport difficult. A summary of the principal phenomena is provided here.

2.2 GENETIC VARIABILITY

Within the realm of Fecal Coliforms (FC), a large degree of variation in bacteria types and behavior exists. The variation is due not only to the genetic makeup of different species but also to adaptability to various environmental stimuli. For example, Bolster et al. (2009) conducted a laboratory study to look at the diversity of several strains of EC taken from six sources: beef cattle, dairy cattle, horse, human, poultry, and wildlife. Two isolates were taken from each source, and each isolate was used in a column transport test. Despite being subjected to the same controlled storage and growth conditions the results revealed a wide range of cell properties and transport behavior among the different isolates. Moreover, in many cases the two isolates obtained from the same source showed statistically significant

differences in cell properties. The study used a 2^{nd}-order (time dependent) deposition model to describe the EC transport through the column, yet the researchers found that "a large diversity exists in the transport behavior among the different EC isolates and that model parameters obtained for one EC strain will likely be inadequate for describing the transport behavior of other strains of EC." The authors conclude by cautioning against making generalized comments about the transport of EC based on the performance of a single EC strain and suggest using a distribution of bacterial parameters when modeling environmental EC, even when the samples originate from a single fecal source.

Notable bacteria diversity has also been observed outside the laboratory. A study by Ishii et al. (2006) examined the presence of EC in three soils near Duluth, MN. The three sites were found to contain 32, 84, and 49 unique EC strains, and the strains were found to be highly genetically related. They found that the strains from each site clustered together, indicating a tendency for different populations to become independently naturalized at each site. Moreover, the authors found that the EC strains in the soil were easily differentiated from EC strains obtained from nearby water, could exist in the soils over seasonal freeze-thaw cycles, and were different than those expected of local water and wildlife, indications that the bacteria were likely naturalized to the soil and not directly deposited or replaced by EC from animal feces or water.

Bacteria diversity in environmental samples is further complicated by the number of possible sources contributing to a given sample. Any given riverine sample could consist of bacteria from livestock, wildlife, streambed sediment, or upstream sources like WWTP outfalls. Bacteria source tracking is a useful technique that can reveal varied percentages of human- and/or animal-generated bacteria in a single sample (e.g. Scott et al. 2003). Other studies have shown significant variability in bacteria sources between samples in a given day and between sampling days and sites (Meays et al. 2006). One examination of swine

manure found 14 different species of bacteria that can cause disease in humans (Entry et al. 2000a).

2.3 BACTERIA FATE AND TRANSPORT IN THE ENVIRONMENT

There is a consensus in the literature that increased in-stream bacteria concentrations are observed after rainfall events. A description of the major known processes contributing to this phenomena are presented here.

2.3.1 Overland Flow

Overland flow can be defined as the portion of surface runoff that reaches a stream channel by traveling over the surface of the ground, and it occurs when precipitation intensity exceeds infiltration capacity (Gupta 2001; Black 1991). Bacteria in overland flow tend to occur either in flocs of cells, as single cells, or attached to soil particles (Tyrrel and Quinton 2003). Cell motility is limited to small scales, and the main transport mechanism at larger scales is via passive transport in water fluxes of surface runoff or infiltration (Unc and Goss 2004). Removal mechanisms that reduce bacteria quantities in overland flow include settling to the ground surface, filtration due to passage through the soil matrix, and filtration due to contact with vegetation (Muirhead et al. 2006a). Single bacteria are small and neutrally buoyant in water, reducing the likelihood of removal by these mechanisms for individual cells (Fiener and Auerswald 2003; Muirhead et al. 2006a). For sediment-attached bacteria and flocs of cells, however, these removal mechanisms are significant. It has been estimated that bacteria attached to soil particles $> 63\mu$m in diameter would settle out of overland flow and that flocs of cells would need to be $> 500\mu$m in diameter before being filtered out by grasses (Fiener and Auerswald 2003). Muirhead et al. (2006b) demonstrated that bacteria attached to large soil particles are transported less than unattached cells, since large soil particles tend to settle to the surface and are more easily filtered. However, Muirhead et al. (2006b) also showed that bacteria tend to attach to smaller soil particles

(fines) in the water flow. The significance of this behavior was apparent under saturation-excess runoff conditions, where single EC cells were rapidly transported via advection in the flow and EC attached to smaller soil particles did not settle. Previous work showed that EC did not erode in flocs and only 8% of EC cells from cowpats were attached to particles, indicating that in some cases the vast majority of bacteria cells may be transported in overland flow as single cells. Bacteria adsorption to sediment may be further complicated by the soil type of the sediment (Ling et al. 2002) and whether the bacteria are hydrophobic or hydrophilic (Unc and Goss 2004).

2.3.2 Infiltration and Subsurface

As rainfall and surface waters infiltrate into the ground, bacteria cells are inevitably carried into the soil matrix. The soil matrix can serve as both a source and a sink of bacteria in terrestrial processes (e.g. Ishii et al. 2006; Muirhead et al. 2006a), although a uniform matrix is thought to function as a filter that prevents bacteria from passing into subsurface layers. Certain practices, however, facilitate the transport of bacteria to and through groundwater and can create significant health concerns if the bacteria encounter subsurface sources used for drinking water (e.g. Macler and Merkle 2000). One such practice is the disposal of human waste to subsurface layers, normally done through legal septic systems (which are prone to leak over time), illegal point-source dumping, or in some cases, especially in undeveloped areas and countries, flushing wastewater directly into a covered hole in the ground (LaWare and Rifai 2006). These actions not only increase the risk of groundwater pollution, but do so with human waste. Exposure to human waste is significant since quantities of human-specific pathogens are much higher in human waste than in waste from other warm blooded animals (LaWare and Rifai 2006; Macler and Merkle 2000).

Another such practice is the use of manure as a fertilizer, which may be applied as a solid or liquid to the land surface or injected directly into the ground (Pappas et al. 2008). Manure typically holds large quantities of fecal bacteria and is one of the most significant non-point

sources addressed by land-use regulation efforts (e.g. Goss and Richards 2008; Hooda et al. 2000; Shirmohammadi et al. 1997). For a perspective on the quantities of manure produced in agricultural areas, Aitken (2003) estimates that a 706 km^2 (273 mi^2) catchment in Scotland with above-national-average livestock densities produces a daily EC loading equivalent to the EC loading of untreated sewage from 1.4 million people. While an effort is made in agricultural areas like this to collect and store manure for later use, the sheer daily volume may be prohibitive and ultimately facilitates the movement of surface bacteria to subsurface layers (Shirmohammadi et al. 1997). Studies have revealed elevated fecal bacteria levels in subsurface drain water within minutes to hours after manure application, and in one case a 30- to 900-fold increase in fecal bacteria was observed in subsurface water after land application of manure (Pappas et al. 2008).

The transport of bacteria in subsurface layers is also facilitated by the existence of subsurface irregularities that disrupt the uniformity of the soil and seriously compromise the soil's filtration capacity. Such irregularities are considered "soil macropores" and can result from capillary-sized pores, isolated vertical zones of loose or porous soil, soil cracks, areas surrounding plant roots, and non-biological voids (Hunter et al. 1992; McMurry et al. 1998; Unc and Goss 2004). These macropore networks create regions of preferential flow that provide a pathway for bacteria to reach new (and sometimes unexpected) locations and in shorter time intervals than would be expected, posing a threat to groundwater sources of drinking water. Results have shown that bypass flow through macropores in structured soils is more the rule than the exception (Unc and Goss 2004). Subsurface transport is magnified in areas with Karst geology, where a network of large and connected subsurface voids increases flow and decreases pathogen filtering, making Karst systems especially vulnerable to pathogen transport (Baffaut and Benson 2009).

2.3.3 Die-Off and Persistence

Since bacteria are living organisms the process of die-off is fundamental to the analysis of environmental survival and fate. First-order kinetics are typically used to describe bacteria die-off, most commonly expressed as Chick's Law (e.g. Wilkinson et al. 1995):

$$N_t = N_0 \exp(-k_d t) \tag{2.1}$$

where N_t is the bacteria population at time t (CFU), N_0 is the initial population (CFU), and k_d [T^{-1}] is a decay constant. It is widely accepted that temperature is the primary factor that affects fecal bacteria die-off, where higher temperatures increase die-off. Moreover, k_d varies with temperature and has been shown to fluctuate over orders of magnitude depending on environmental conditions. For example, an EPA study analyzed 30 separate in-situ studies and found that the decay rate for coliform in fresh water ranged from 0.12 to 26 d^{-1} with a median of 1.0 d^{-1}. Wilkinson et al. (1995) summarized die-off rates for FC from five different studies and showed that for a temperature range of 0-30 °C, the die-off rate varies from less than 0.03 d^{-1} to greater than 1.0 d^{-1}. Due to the wide range of reported die-off rate values and the highly site-specific nature of the results, they commented "Direct comparison of die-off rates taken from different studies should be treated with caution." Chick's law is often rearranged to express T_{90}, the time needed for 90% of the initial bacteria population to die (e.g. Kashefipour et al. 2006; Wilkinson et al. 1995).

Although the die-off rate is the dominant factor in describing bacteria survival, other important factors are sunlight exposure, turbidity, salinity, soil type, soil pH, moisture content, nutrient availability, vegetation, and predation, and the relative effects of these factors may be strain and species specific (Pappas et al. 2008; Tyrrel and Quinton 2003; Ishii et al. 2006). Wilkinson et al. (1995) report that sunlight, for example, has both direct

and indirect effects, where sunlight directly causes cell damage and indirectly affects temperature and moisture. They report that T_{90} values between direct sunlight and shade or darkness can vary between 1-2 hours and several days. Similarly, turbidity directly affects exposure to sunlight so that clean rivers or lakes will have higher die-off due to sunlight than dirtied waters like those that receive sewage or effluent. One modeling study addressed the issue by using a total decay rate comprised of two components, one as a function of sunlight decay rate and the other a function of dark decay rate (Kashefipour et al. 2006). The net result of all these factors is a broad range of k_d values reported in the literature for FC and EC and in fresh and salt waters.

In addition to die-off, bacteria can persist and adapt to environments to create sustained populations. In the Ishii et al. (2006) study the authors were able to conclude that "soilborne, naturalized strains were adapted to this soil and were not continuously added from external sources." Similarly, Hunter and McDonald (1991) found high concentrations of fecal bacteria in upland streams and suggest that in some areas a semi-permanent land store of bacteria may exist in surface soils, and the store may act as a potential source of bacteria to streams during hydrological events. Moreover, pathogen survival in the upper soil layer has been shown to vary from 4 to 160 days (Entry et al. 2000b) while Gagliardi and Karns (2000) add that "it has been shown that the levels of members of several genera of pathogenic bacteria decrease only slightly during 100 days in groundwater alone, and several studies have shown that sediments serve as reservoirs for fecal pathogens."

2.3.4 Streambed Sediment

While various terrestrial processes are acknowledged to be a store and/or source of bacteria, river channels can also be a significant source of bacteria (e.g. Jamieson et al. 2004; Obiri-Danso and Jones 1999). Other than point sources, the most commonly suspected source of bacteria in river channels is in the river bed sediment, although direct deposition of grazing animals and wildlife may also be significant (Meays et al. 2006; Hooda et al.

2000). River bed sediment is an ideal environment for bacteria survival due to minimal light exposure, increased protection from predators, and high nutrient concentrations (Russo et al. 2011; Bai and Lung 2005).

When streambed sediment is agitated or disrupted, the entrained bacteria become dislodged and are suspended in the water column. This phenomena is induced by increased streamflow, usually in the wake of a rainfall event, although something as simple as an animal crossing a stream and disrupting sediment with their hooves can have the same effect (Goss and Richards 2008). This concept has been illustrated by experiments using artificial flows in rivers in otherwise dry periods to examine what occurs in the absence of rainfall (e.g. McDonald et al. 1982; Davies-Colley et al. 2008). The studies showed that significant peaks in bacteria concentrations occurred, but that the bacteria peaks arrived *before* the flow peak, indicating that the peak bacteria were more likely from flow-agitated streambed sediment than terrestrial sources. This observation does not negate the contribution from the land store as a result of rainfall and runoff, but does reveal that in-stream sources contribute significantly to the observed heightened concentrations after rainfall events. Moreover, McDonald et al. (1982) showed that in a 1-hour interval between artificial stream pulses, bacteria initially suspended by the first pulse attempted to settle but were re-suspended by the second pulse. An additional complication occurs when increased river flow contacts wetted stream-banks and suspends or dislodges bacteria otherwise attached to bank sediment and soil (Byappanahalli et al. 2003). A similar phenomena may occur in tidal areas, where EC concentrations were observed to increase with high tide as an urban river covered more of its streambank and when it was determined low tide (predominantly baseflow) conditions were not a source of EC (Solo-Gabriele et al. 2000).

2.4 CONCLUSION

This chapter has shown that environmental bacteria processes are complicated and depend on a variety of localized environmental factors such as meteorology, soil type, land use, and livestock species. Moreover, enough behavioral variety has been observed, even among single samples or identical strains of bacteria, that making universal conclusions about these processes beyond general relationships is a difficult and likely impossible task. Using these conclusions to develop process equations for a bacteria model is therefore a challenging and uncertain process. For example, how can equations governing overland transport and sediment adsorption of bacteria be accurately programmed into a model when the process is so dependent on varying micro-scale factors? How can parameters be used to accurately describe soil types and infiltration rates when these properties are rarely homogeneous beyond the plot scale, let alone in application of watershed-scale models? How can data collected at time scales of weeks or months be used to evaluate bacteria processes that might only last minutes or hours? It is apparent that models can not accurately account for all of these complex processes, and therefore model performance will always carry a degree of uncertainty. And yet, these uncertain models are the most useful tools available for addressing the nation's premier water-quality problem. Therefore, it is only through further research and model development that these tools can be sharpened, resulting in the reduction of model uncertainty and an improved ability to replicate and predict bacteria concentrations.

Chapter 3

Model Setup

3.1 SITE DESCRIPTION

The site chosen for this study is the Little River Experimental Watershed (LREW) near Tifton, GA (Fig. 3.1), due to the availability of appropriate data needed for modeling. The LREW has an area of 334 km^2 centered approximately at N31.61, W83.66, has broad floodplains, river terraces, and gently sloping uplands, and is representative of the agricultural areas in the Southeast Coastal Plain (Bosch et al. 2007). The LREW is located in Turner, Worth, and Tift counties and is part of the headwaters of the Suwannee River Basin (Sullivan et al. 2007; Feyereisen et al. 2007). Stream-channel slopes range from 0.1%-0.4% (Bosch and Sheridan 2007), while land slopes are mostly less than 5%, although some valley slopes range from 5 to 15% (Bosch et al. 2007). Land use types in the LREW are predominantly agriculture and forest with small percentages of water and residential/commercial (Sheridan 1997). Non-pasture agricultural lands are primarily row crops of cotton, peanuts, corn, and fruit and vegetable crops (Feyereisen et al. 2007; Sullivan et al. 2007). Riparian buffers are common in the region and are dominant hydrological features of the landscape, while the subsurface properties of the region

facilitate significant subsurface flow both in the vadose zone and shallow groundwater (Bosch et al. 2010). The USDA began hydrological and meteorological monitoring in the LREW in the late 1960s and an extensive database has been developed for the seven sub-catchments within the LREW (F, I, J, K, M, N, O), which range in size from 3-115 km^2 (Sullivan et al., 2007; Bosch et al., 2007a; Bosch et al., 2007b, Bosch and Sheridan, 2007; Feyereisen et al., 2007). Several studies have characterized the hydrological properties of the LREW (e.g. Sheridan, 2007; Bosch et al., 2005; Feyereisen et al., 2008; Feyereisen et al., 2007a; Bosch et al., 2010; Chin et al., 2009; Sheridan et al., 1997; Zhang et al., 2011; White et al., 2009; Cho et al., 2010). The LREW database can be accessed at `ftp://www.tiftonars.org/`.

Catchments I, J, K, and O were selected for this study, where Catchments J and K are nested within I, and Catchment O contains more intensive dairy practices. General

Table 3.1: LREW catchment summary data

Catchment	ID	Stream Order	Area (km^2)	Land-Use %†			Max Elev. (m)	Min Elev. (m)
				Ag.	For.	Urb.		
I	4950	4	49.9	52	48	0	146	102
J	4850	3	22.1	52	48	0	146	105
K	4860	3	16.7	46	54	0	146	106
O	6860	4	15.9	83	12	5	122	86

* -from Feyereisen et al. 2007

† -area percentages from HSPF and SWAT model build

information regarding each catchment is shown in Table 3.1. With the exception of a small urban portion in Catchment O, agricultural and forest lands are the only land types in the catchments. Precipitation measurements are collected in five-minute intervals by the LREW instrumentation, TE525 tipping-bucket gages (Texas Instruments, Inc., Dallas, Texas) with a measurement precision of 0.254 mm and accuracy $\pm$0.5 mm/hr. Rain gages 43 (UTM 3513276 m N, 242618 m E) and 64 (UTM 3488321 m N, 259444 m E) were used for catchments K and O, respectively, while rain gage 38 (UTM 3511453 m N, 241558 m E) was used for both catchments I and J. Streamflow measurements were made by weir

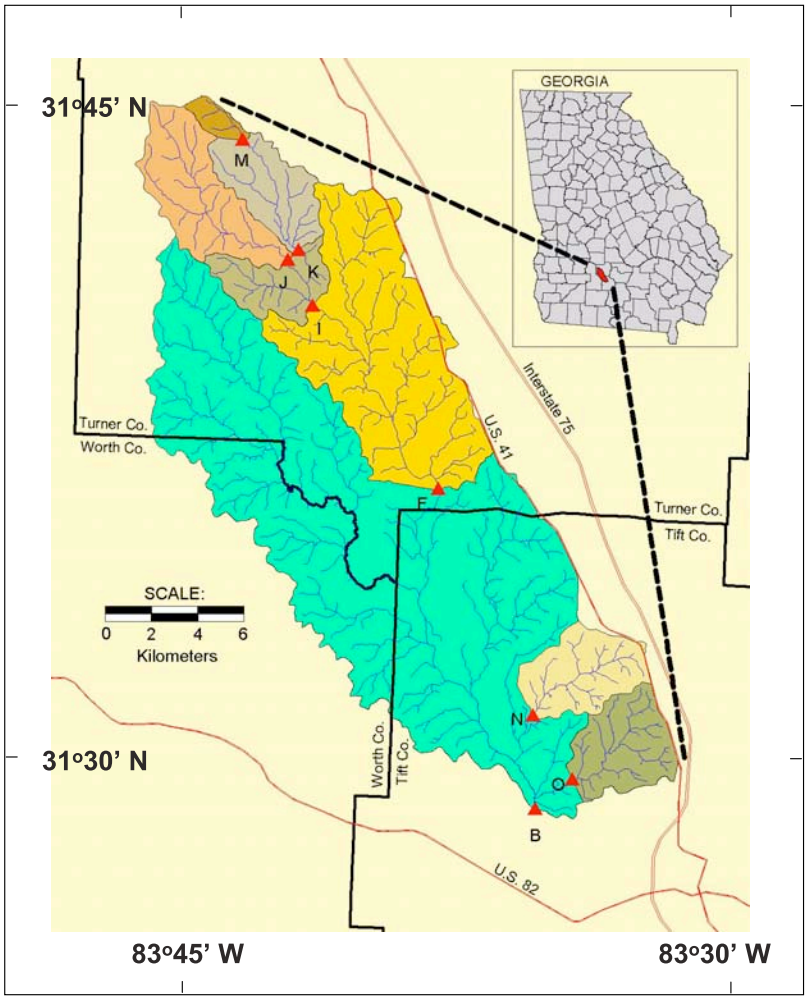


Figure 3.1: The Little River Experimental Watershed in Tifton, GA (Bosch et al. 2007)

structures with design specifications as described in Bosch and Sheridan 2007, and daily average streamflow values are reported in the LREW database. Hourly measurements of cloud cover, wind velocity, dewpoint, and air temperature were taken from the Moody Air Force Base (station 747810) in Valdosta, GA, approximately 70 km from the LREW. To guarantee that the HSPF and SWAT models were built as similarly as possible, identical shape files for the elevation, land use and soil types were used to build the models. The shape files were downloaded from the EPA BASINS (v4.0) system, and the models were built using the USGS National Elevation Dataset, USGS GIRAS, and STATSGO data sets for elevation, land use, and soil type, respectively. BASINS was used to delineate the

catchments in HSPF, while the ArcSWAT/ArcGIS interface was used to build and delineate the catchments in SWAT. Because SWAT handles land-use types differently than HSPF, the thresholds per land use and soil type were adjusted in the building of the SWAT models to match the setup in HSPF. Since HSPF and SWAT operate at different time steps, the meteorological data was aggregated appropriately for input to the models.

All models were developed for the period Jan 1996- Dec 2002. FC datasets with approximately monthly measurements during the period Jan 1997-Apr 2002 were used (Vellidis et al. 2010), and some summary statistics of the FC datasets are shown in Table 3.2. Meteorological activity during the study period yielded wet years from 1996-1998

Table 3.2: Summary of fecal coliform datasets

	LREW Catchment			
	I	J	K	O
# FC Data Points	47	55	53	48
Min Value (CFU/100mL)	6	7	17	15
Max Value (CFU/100mL)	12000	7000	6500	4000
Geometric Mean (CFU/100mL)	208	153	215	157

and dry years from 1999-2002, based on whether the annual precipitation was greater or less than the long-term average annual precipitation (Feyereisen et al. 2007). Although the average annual precipitation facilitates division into wet and dry years, it is not completely representative of the meteorological patterns in the LREW, since rainfall in the region is poorly distributed throughout the year and frequently leads to drought conditions in the summer growing season (Bosch et al. 2005). The monthly rainfall totals were below average for January, February, April, May, and December for the dry years (1999-2002) (Bosch et al. 2005). The LREW was affected by Hurricane Earl in Aug/Sep 1998, and Hurricane Gordon in September of 2000 caused the monthly precipitation to be 3.5 times the long-term average (Bosch et al. 2005).

3.2 MODEL DESCRIPTIONS

Previous studies that have examined bacteria modeling in both HSPF and SWAT include Singh et al. (2005) and Chin et al. (2009). Benham et al. (2006), in particular provide a detailed study of the process equations relating to bacteria fate and transport in HSPF and SWAT.

3.2.1 HSPF

The Hydrologic Simulation Program Fortran (HSPF) (Bicknell et al. 2001) is a watershed-scale, process-oriented model that operates on an hourly time step. The model uses three main modules to represent a watershed: pervious land segments (PERLND), impervious land segments (IMPLND), and reaches and reservoirs (RCHRES), each of which are considered to be hydrologically homogeneous. Storage during runoff and infiltration are determined using the Manning and Philip equations, respectively, a storage routing technique is used to route water in reaches, and a mass-balance approach is used to determine the distribution of water in the system. The soil structure of pervious areas is divided into four layers or zones. The surface layer receives rainfall and irrigation, and water from the surface layer may become overland flow. The upper zone represents water in depressional storage on the surface and in the top layers of the soil. Water in the upper zone that does not percolate to lower layers may become direct evapotranspiration. The lower zone represents the root zone of plants, and water in the lower zone may become evapotranspiration depending on factors like vegetation type, rooting depth, and soil properties. The upper and lower zones both have water storage components and can both receive water from and contribute water to interflow. Water that infiltrates and percolates through the upper zone and lower zone arrives at the groundwater layer, at which point it is further distributed between active groundwater and inactive groundwater.

A summary of the process equations in HSPF describing infiltration and runoff, interflow, groundwater, and water quality, as described in Bicknell et al. 2001, are provided below.

Description of Infiltration and Runoff Processes

In HSPF, the moisture available to land surfaces, M_{ls}, is distributed between infiltration, interflow, and surface runoff according to the relationships shown in Figure 3.2. The abscissa of Figure 3.2 is the % of watershed area, the ordinate is the amount of water considered per interval (in), and Lines 1 and 2 distribute the total moisture available (M_{ls}) into regions of infiltration capacity (below Line 1), potential interflow and inflow (between Lines 1 and 2), and potential surface detention and runoff (above Line 2). The shape of the figure is based on the calculation of the mean infiltration capacity over the land segment, $\overline{I_1}$ (in/interval), found by

$$\overline{I_1} = \left(\frac{I_f}{L^{I_e}}\right) I_g \tag{3.1}$$

where I_f is a parameter describing infiltration capacity of the soil (in/interval), L is a ratio of lower zone storage (in) to lower zone nominal capacity (in), I_e is an exponent parameter that must be greater than 1.0 [-], and I_g is a factor to account for frozen ground, if applicable [-]. Upon determining the value for $\overline{I_1}$, the remainder of Line 1 is described by the equations

$$I_{\max,1} = I_d \overline{I_1} \tag{3.2}$$

and

$$I_{\min,1} = \overline{I_1} - \left(I_{\max,1} - \overline{I_1}\right) = 2\overline{I_1} - I_{\max,1} \tag{3.3}$$

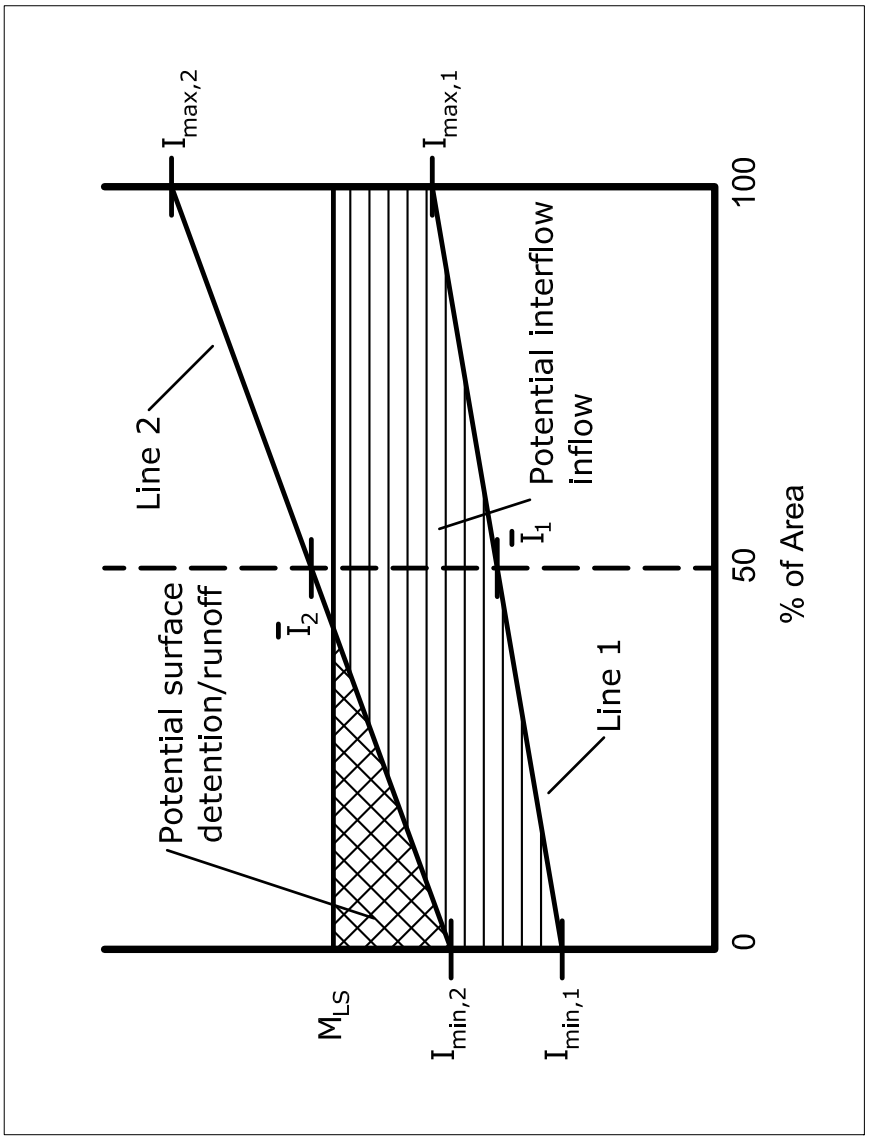


Figure 3.2: Distribution of infiltration and interflow inflow (Modified from Bicknell et al. 2001)

where $I_{max,1}$ is the ordinate of Line 1 at 100% of the area; I_d is a parameter describing the ratio of maximum to mean infiltration capacity over the land segment [-] and has a range of

[1.0-2.0]; and $I_{\min,1}$ is the ordinate of Line 1 at 0% of the area. Upon establishing Line 1, Line 2 is determined using the relationship

$$R = I_n\left(2.0^L\right) \tag{3.4}$$

where R is the ratio of the ordinates of Line 2 to Line 1 and I_n is a parameter describing interflow inflow [-]. It is apparent that the distribution of M_{ls} shown in Figure 3.2, as described by Equations 3.1-3.4, is primarily dependent on $\overline{I_1}$ as determined in Equation 3.1. Moreover, since I_g is simply a factor accounting for frozen ground, the most important parameters affecting $\overline{I_1}$ are those describing infiltration (I_f), the ratio between lower zone storage and lower zone nominal capacity (L), and the exponential parameter I_e. The ratio L can be greater than, less than, or equal to 1.0. The slope of Line 1 is determined by the parameter I_d. Since I_d is limited to the range [1.0-2.0], HSPF requires that the maximum infiltration ($I_{\max,1}$) can not be more than twice the mean infiltration $\overline{I_1}$. It is also apparent in Figure 3.2 that as the % of area increases, the amount of M_{ls} assigned to potentail surface detention and runoff decreases, and may even be zero, depending on the slopes of Lines 1 and 2. The interflow parameter (I_n) in Equation 3.4 has the most direct affect on the slopes of the lines, and as I_n increases, the potential interflow and inflow increase while the potential surface detention and runoff decrease.

Description of Interflow Processes

In HSPF, the interflow outflow is assumed to have a linear relationship to storage and is therefore a function of inflow, storage, and a recession parameter. Contributions to the interflow component may originate at the surface or come from upslope external lateral flows, and are either retained in storage or routed as discharge from the land segment. The interflow outflow is calculated by

$$I_{fw} = I_{k1} I_{fl} + I_{k2} I_{fs} \tag{3.5}$$

where I_{fw} is the interflow outflow (in/interval); I_{fl} is the inflow into the interflow storage, including lateral inflow (in/interval); I_{fs} is the interflow storage at the start of the interval (in); and I_{k1} and I_{k2} are variables determined by

$$I_{k1} = 1.0 + \left[\frac{I_{k2}}{\ln(I_{rc})\delta/24.0} \right] \tag{3.6}$$

$$I_{k2} = 1.0 - \exp\left[\ln(I_{rc})\delta/24.0 \right] \tag{3.7}$$

where I_{rc} is an interflow recession parameter (d^{-1}) and δ is the number of hours per time interval considered. I_{rc} is the ratio of the present rate of interflow outflow to the value 24 hours earlier, given no inflow. I_{rc} is inversely proportional to I_{fw} such that as I_{rc} increases, the interflow outflow decreases and vice versa. Values of I_{rc} are in the range (0 1) and can be input on a monthly basis to allow for variation in soil properties throughout the year.

Description of Groundwater Processes

In HSPF, water that infiltrates or percolates from the upper zone but does not go to lower zone storage is assigned to either active groundwater or inactive groundwater. Inactive groundwater is considered lost from the system, and a user-defined parameter exists to describe the fraction of the groundwater inflow that goes to inactive groundwater. The infiltrating and percolating water not distributed to the inactive groundwater combines with any lateral inflow and/or irrigation water to form the total inflow to the active groundwater storage. The outflow from active groundwater is calculated by

$$A_{go} = K_{gw} A_{gs} \left(1.0 + K_v G_s\right) \tag{3.8}$$

where A_{go} is the active groundwater outflow (in/interval), K_{gw} is the groundwater outflow recession parameter (interval^{-1}), A_{gs} is the active groundwater storage at the start of the interval (in), K_v is a parameter which can make the active groundwater storage relation to outflow nonlinear (in^{-1}), and G_s is the index to groundwater slope (in). The parameter K_{gw} is calculated as

$$K_{gw} = 1.0 - \left(A_{gr}^{(\delta/24.0)}\right) \tag{3.9}$$

where A_{gr} is the daily recession constant of groundwater flow if K_v or G_s = 0.0; A_{gr} can be thought of as the ratio of current groundwater discharge to groundwater discharge 24-hr earlier.

Description of Water-Quality Processes

HSPF does not distinguish between water-quality constituents but rather uses the same process equations for all contaminants, while accounting for conservative and non-conservative properties. In pervious areas (PERLND), water-quality constituents can be simulated in surface and subsurface outflow. The concentrations in surface outflow are considered as either associated with sediment transport or part of a system based on deposition (mass per area per time), where accumulation and depletion of the constituent dictate land-surface storage and the concentration in surface outflow is a function of the surface-water flow and the land-storage concentrations. In addition to overland flow, HSPF can model water-quality constituents in interflow, in groundwater, and sorbed to sediment; users can specify which features to include or omit. Of the four methods of representing

water quality constituents, only the overland flow method can receive deposition, since it is the only one to maintain a storage feature. Deposition can be represented as a wet flux associated with a concentration in rainfall, or as a dry flux applied directly to land surfaces. In addition to overland flow, water-quality constituents can be accumulated and removed by processes independent of storm events. In impervious areas (IMPLND), water-quality constituents are only considered in overland flow and sorbed to sediment, and the algorithms for these processes are the same as those for the pervious areas. In reaches and reservoirs (RCHRES), water-quality constituents are considered in either a dissolved or sediment-associated state. For non-sediment-associated cases, advection and decay processes are modeled, while for sediment-associated cases the additional processes considered are advection of suspended material, deposition and scour of sediment, decay of suspended and bed material, and adsorption/desorption between dissolved and sediment-associated phases.

The fundamental relationship describing surface storage of water-quality constituents is

$$S = A_c + S_0(1.0 - R_r) \tag{3.10}$$

where S is the constituent storage on the surface (#/acre), A_c is the constituent accumulation rate (#/acre $\cdot$ d), S_0 is the constituent storage at the start of the interval (#/acre), and R_r is a unit removal rate of the stored constituent (d^{-1}). When surface runoff occurs while the constituent is in surface storage, the constituent washoff is calculated by

$$C_w = S\left[1.0 - \exp\left(-S_w \frac{2.30}{S_{90}}\right)\right] \tag{3.11}$$

where C_w is the washoff of the constituent from the land (#/acre $\cdot$ d), S_w is the surface outflow of water (in/hr), and S_{90} is the surface runoff rate that induces 90% constituent washoff in 1 hour (in/hr).

In-stream die-off is calculated by a first-order relationship

$$c_t = c_0 \exp(-kt) \tag{3.12}$$

where c_t (#/ft^3) is the concentration at time t, c_0 is the initial concentration, and k is a first-order decay rate parameter (hr^{-1}). Constituent concentration in interflow outflow is user-specified with a dedicated concentration parameter.

3.2.2 SWAT

The Soil and Water and Assessment Tool (SWAT) (Neitsch et al. 2005) is a watershed-scale, process-oriented model that operates on a daily time step. The model divides a watershed into subbasins to account for varying soil types and land uses. Hydrologic response units (HRUs) are lumped land areas within the subbasin comprised of similar soil, land cover, and management combinations. A shallow aquifer and a deep aquifer are simulated for each subbasin. The shallow aquifer is unconfined and contributes to flow in the main reach of the subbasin while the deep aquifer is confined and any water entering the deep aquifer is assumed to contribute to streamflow outside the watershed, i.e. it is lost from the system. For given daily rainfall amounts, runoff is calculated per HRU using the NRCS Curve Number method and routed to obtain total watershed runoff.

A summary of the process equations in SWAT describing surface runoff, groundwater, and water quality, as described in Neitsch et al. (2005), are provided below.

Description of Surface Runoff Processes

SWAT uses the NRCS Curve Number (CN) method to estimate runoff volumes from varying land use and soil types. The CN method is described by

$$Q_{CN} = \frac{\left(R - I_a\right)^2}{\left(R - I_a + S\right)} \tag{3.13}$$

where Q_{CN} is the accumulated runoff or rainfall excess, R is the daily rainfall depth, I_a is the initial abstraction, and S is the retention parameter, and all of these parameters are in terms of mm of H_2O. Runoff occurs when R is greater than I_a. The initial abstraction term (I_a) accounts for surface storage, interception, and infiltration prior to runoff, and the retention parameter (S) is calculated by

$$S = 25.4\left(\frac{1000}{CN} - 10\right) \tag{3.14}$$

where CN is the curve number. I_a is commonly assumed to equal $0.2S$ and Equation 3.13 becomes

$$Q_{CN} = \frac{\left(R - 0.2S\right)^2}{\left(R + 0.8S\right)} \tag{3.15}$$

The CN method, although empirical, is commonly used and relatively straightforward, requiring only estimates of rainfall (R) and CN. Higher CN values lead to higher Q_{CN} and estimates of CN for various land uses and soil types can be found in Neitsch et al. (2005) and Chin (2006).

Description of Groundwater Processes

SWAT uses the following water balance to describe the shallow aquifer:

$$A_{s,i} = A_{s,i-1} + W_{r,i} - Q_{gw,i} - W_{rp,i} - W_{pump} \tag{3.16}$$

where $A_{s,i}$ is the amount of water stored in the shallow aquifer on day i, $A_{s,i-1}$ is the amount of water stored in the shallow aquifer on day $i-1$, $W_{r,i}$ is the amount of recharge entering the shallow aquifer on day i, $Q_{gw,i}$ is the groundwater flow, or base flow, into the main channel on day i, $W_{rp,i}$ is the amount of water moving into the soil zone in response to water deficiencies on day i, and W_{pump} is the amount of water removed from the shallow aquifer by pumping on day i. All terms have dimension mm of H_2O. Aquifer recharge occurs when water percolates beyond the lowest depth of the soil profile or when bypass flow passes through the vadose zone, and the recharge to both aquifers on a given day is determined using the exponential relationship

$$w_{rt,i} = (1 - exp[-1/\delta_{gw}]) \cdot w_{sp} + exp[-1/\delta_{gw}] \cdot w_{rt,i-1} \tag{3.17}$$

where $w_{rt,i}$ is the total amount of recharge entering the aquifers on day i (mm H_2O), δ_{gw} is the delay time or drainage time of the overlying geologic formations (days), w_{sp} is the total amount of water exiting the bottom of the soil profile on day i (mm H_2O), and $w_{rt,i-1}$ is the total amount of recharge entering the aquifers on day $i-1$ (mm H_2O). The term δ_{gw} cannot be directly measured, but can be estimated by using different values of δ_{gw} to simulate aquifer recharge and comparing the resulting variations in water table level with the observed values. The portion of the total daily recharge routed to the deep aquifer is described by the relation

$$w_d = \beta_d \cdot w_{rt} \tag{3.18}$$

where w_d is the amount of water moving into the deep aquifer on day i (mm H_2O), β_d is the aquifer percolation coefficient [-], and w_{rt} is the total amount of recharge entering both aquifers on day i (mm H_2O).

The water in the shallow aquifer may contribute to baseflow in the reach depending on a threshold parameter that stipulates the minimum amount of water stored in the shallow aquifer for baseflow contribution to occur. The groundwater flow to the reach on day i, $Q_{gw,i}$, is calculated using

$$Q_{gw,i} = Q_{gw,i-1} \cdot exp[-\alpha_{gw} \cdot \Delta t] + w_{r,sh} \cdot (1 - exp[-\alpha_{gw} \cdot \Delta t]) \tag{3.19}$$

if $aq_{sh} > aq_{sht}$ and

$$Q_{gw,i} = 0 \tag{3.20}$$

if $aq_{sh} \leq aq_{sht}$, where $Q_{gw,i}$ is the groundwater flow into the main channel on day i (mm H_2O), $Q_{gw,i-1}$ is the groundwater flow into the main channel on day $i-1$ (mm H_2O), α_{gw} is the baseflow recession constant, Δt is the time step (1 day), $w_{r,sh}$ is the amount of recharge entering the shallow aquifer on day i (mm H_2O), aq_{sh} is the amount of water stored in the shallow aquifer at the beginning of day i (mm H_2O) and $aq_{sh,t}$ is the threshold water level in the shallow aquifer for groundwater contribution to the main channel to occur (mm H_2O).

SWAT makes a distinction between water removed from the unsaturated zone for evapotranspiration (ET) and the water displaced from the saturated zone to the unsaturated

zone to replace the water lost to ET. This movement of water from the shallow aquifer to the overlying unsaturated zone is referred to in SWAT as *revap*, and is modeled as a function of water demand for ET:

$$w_{rev} = \beta_{rev} \cdot E_0 \tag{3.21}$$

where w_{rev} is the maximum amount of water displaced to the unsaturated zone in response to water deficiencies (mm H_2O), β_{rev} is the revap coefficient, and E_0 is the potential ET for the day (mm H_2O).

Description of Water-Quality Processes

In contrast to HSPF, SWAT uses a suite of process equations designed specifically to address bacteria fate and transport. The main source of terrestrial bacteria in the watershed is the application of manure to land surfaces. Bacteria concentrations within the manure are specified by a parameter governing land application (kg/ha) and a parameter describing bacteria content in the manure (CFU/g). When manure is applied, the bacteria in the manure is either attached to plant foliage or applied directly to the surface layer, defined by SWAT as the top 10 mm of soil, and the partitioning of bacteria to these two surfaces is calculated as a function of the ground cover. Washoff of the bacteria from the foliage to the surface layer only occurs when daily precipitation exceeds 2.54 mm and is governed by a wash-off fraction parameter that represents the amount of bacteria on the foliage considered dislodgeable. Bacteria in the surface soil layer is considered to be either in solution or adsorbed to surface soil particles, and bacteria in soil solution may be transported to deeper soil layers via tillage or percolation, where it is assumed to die. Bacteria that wash off foliage remain in solution in the surface layer. Die-off and regrowth, described by Chick's law, are considered for the addition and subtraction of bacteria to the system, and are

calculated on foliage, in the surface soil solution, and on the soil particles. Bacteria in soil solution may interact directly with surface runoff, and depending on soil properties, bacteria attached to soil particles may be transported to the main channel via surface runoff. For large subbasins with a time of concentration greater than one day, SWAT accounts for the bacteria in the surface runoff which does not arrive at the main channel on the day it is generated but rather goes to storage. For bacteria in-stream or in reservoirs, the only process considered is die-off. Finally, SWAT allows for two types of bacteria to be modeled by distinguishing between *persistent* and *less-persistent* bacteria. The fate and transport of the two types are governed by distinct die-off and regrowth rates. The rationale for this approach is to account for long-term impacts in soils, where persistent bacteria may become significant in time. While this feature provides modelers with more flexibility in addressing bacterial processes in their watersheds, Baffaut and Sadeghi 2010 comment that in practice most SWAT models only use one of the types of bacteria due to a lack of information.

The die-off equation used for the soil surface solution is

$$b_{sl,i} = b_{sl,i-1} \cdot exp(-\mu_{sl}) - b_{min} \tag{3.22}$$

where $b_{sl,i}$ is the amount of bacteria present in soil solution on day i (CFU/m^2), $b_{sl,i-1}$ is the amount of bacteria present in soil solution on day $i-1$ (CFU/m^2), μ_{sl} is the overall rate constant for die-off/re-growth of bacteria in soil solution (1/day), and b_{min} is the minimum daily loss of bacteria, which was maintained at the default value of zero. The leaching of bacteria in soil solution in the top 10 mm into the first soil layer is

$$b_{perc} = \frac{b_{sl} \cdot w_{ps}}{10 \cdot \rho_b \cdot d \cdot k_{bp}} \tag{3.23}$$

where b_{perc} is the amount of bacteria transported from the top 10 mm into the first soil layer (CFU/m^2), b_{sl} is the amount of bacteria present in soil solution (CFU/m^2), w_{ps} is the amount of water percolating to the first soil layer from the top 10 mm on a given day (mm H_2O), ρ_b is the bulk density of the top 10 mm (Mg/m^3) (assumed to be equivalent to bulk density of first soil layer), d is the depth of the "surface" layer (10 mm), and k_{bp} is the bacteria percolation coefficient (10 m^3/Mg). The amount of bacteria in the soil solution transported in surface runoff is

$$b_{sro} = \frac{b_{sl} \cdot Q_{CN}}{\rho_b \cdot d \cdot k_{bs}} \tag{3.24}$$

where b_{sro} is the amount of bacteria lost in surface runoff (CFU/m^2), b_{sl} is the amount of bacteria present in soil solution (CFU/m^2), Q_{CN} is the amount of surface runoff calculated by the CN method on a given day (mm H_2O), and k_{bs} is the bacteria soil partitioning coefficient (m^3/Mg). The amount of bacteria transported to the stream via attachment to soil particles is calculated by

$$b_{sed} = 0.0001 \cdot c_{sed} \cdot \frac{y_{sed}}{A} \cdot \varepsilon \tag{3.25}$$

where b_{sed} is the amount of bacteria transported with sediment in surface runoff (CFU/m^2), c_{sed} is the concentration of bacteria attached to sediment in the top 10 mm (CFU/metric ton soil), y_{sed} is the sediment yield on a given day (metric tons), A is the HRU area (ha), and ε is the bacteria enrichment ratio. The concentration of bacteria in the sediment is determined by

$$c_{sed} = 1000 \cdot \frac{b_{sb}}{\rho_b \cdot d} \tag{3.26}$$

where b_{sb} is the amount of persistent bacteria sorbed to the soil (CFU/m^2). When the time of concentration in a watershed is greater than one day, the bacteria load in the surface runoff is used to estimate the bacteria released to the main channel by:

$$b_{sro} = (b'_{sro} + b_{sros,i-1}) \cdot \left(1 - exp\left[\frac{-C_L}{t_c}\right]\right) \tag{3.27}$$

$$b_{sed} = (b'_{sed} + b_{seds,i-1}) \cdot \left(1 - exp\left[\frac{-C_L}{t_c}\right]\right) \tag{3.28}$$

where b_{sro} is the amount of persistent bacteria discharged to the main channel in surface runoff on a given day (CFU/m^2), b'_{sro} is the amount of surface runoff persistent bacteria generated in the HRU on a given day (CFU/m^2), $b_{sros,i-1}$ is the surface runoff persistent bacteria stored or lagged from the previous day (CFU/m^2), b_{sed} is the amount of sediment-attached persistent bacteria discharged to the main channel in surface runoff on a given day (CFU/m^2), b'_{sed} is the amount of sediment-attached persistent bacteria generated in the HRU on a given day (CFU/m^2), $b_{seds,i-1}$ is the sediment-attached persistent bacteria stored or lagged from the previous day (CFU/m^2), C_L is the surface runoff lag coefficient, and t_c is the time of concentration for the HRU (hrs.). The die-off relationship for bacteria in stream reaches and reservoirs is calculated using the first-order decay function

$$b_{w,i} = b_{w,i-1} \cdot exp(-\mu_w) \tag{3.29}$$

where $b_{w,i}$ is the amount of bacteria present in the water body (reach or reservoir) on day i (CFU/100mL), $b_{w,i-1}$ is the amount of bacteria present in the water body on day $i-1$ (CFU/100mL), and μ_w is the rate constant for die-off of bacteria in the water body (1/day).

3.2.3 The CC Model

The Characteristic Concentration (CC) model (Chin, 2011a) combines the principles of hydrograph separation with the concept of *event mean concentration* (EMC) which assumes that the runoff contaminant loading is linearly proportional to the runoff volume. The CC model therefore models bacteria by assigning a (constant) characteristic bacteria concentration to the runoff and a separate characteristic concentration to the baseflow. Whereas the EMC approach has been used primarily in urban watersheds, the application of the concept to bacteria modeling in agricultural areas is new, and preliminary studies have shown promising results (Chin, 2011a; Chin, 2011b). The model makes no assumptions about terrestrial loadings or micro-scale processes like washoff or adsorption, but rather quantifies the contributions of runoff and baseflow processes to in-stream concentrations.

In the CC model, the model-predicted in-stream bacteria concentration, $c(t)$, is calculated using the relation (Chin, 2011b):

$$\ln c(t) = q_r(t) \ln c_r + [1 - q_r(t)] \ln c_b + \varepsilon(0, \sigma) \tag{3.30}$$

where q_r is the surface-runoff fraction of the streamflow, c_r and c_b are the (constant) characteristic concentrations associated with runoff and baseflow, respectively, and ε is a Gaussian random variable with a mean of zero and a standard deviation of σ, where σ is the standard error of fitting a straight line to a plot of $\ln c$ versus q_r derived from concentration and streamflow measurements. The geometric mean, C_G, of N concentration measurements is then

$$C_G = c_r^{\langle q_r \rangle_N} c_b^{\langle 1 - q_r \rangle_N} 10^{\varepsilon(0, \sigma/\sqrt{N})} \tag{3.31}$$

where $\langle q_r \rangle_N$ is the average of $q_r(t)$ over N measurements, and it is the calculated C_G values that are used as model predictions to compare with the observed log-concentrations.

The surface-runoff and baseflow portions of the streamflow are estimated using the three hydrograph separation methods of the USGS HYSEP computer program (Sloto and Crouse 1996). The three techniques are named the fixed interval (FI), the sliding interval (SI), and the local minimum (LM) methods, and are based on the duration of surface runoff calculated as

$$D = A^{0.2} \tag{3.32}$$

where D is the number of days after which surface runoff stops, and A is the drainage area in mi^2. An interval "$2d^*$" is defined as the odd integer between 3 and 11 days nearest to $2D$. The fixed-interval method assigns the lowest discharge in each $2D^*$ interval to all days in that interval. The sliding-interval method finds the lowest discharge in one half the $2D^*$-interval-minus-one-day before and after the day being considered and assigns it to that day. The local minimum method examines each day to see if it is the lowest discharge in one half the $2D^*$-interval-minus-one-day before and after the current day. If it is, then it is a local (discharge) minimum and it is connected to adjacent local minimums by straight lines, from which the baseflow values on each day between local minimums are interpolated. In the context of pollutant transport via surface runoff, it is worth noting that of the three methods, the LM method produces the lowest baseflow ratios and therefore highest runoff quantities.

Chapter 4

Model Calibration

4.1 BACKGROUND

Calibration is the process of adjusting model parameters to minimize the error between measurements and model predictions. Model errors occur because of model predictive uncertainty, which is a combination of model structural uncertainty, parameter uncertainty, and data uncertainty. Structural uncertainty is a result of an imperfect understanding of natural processes as well as the extrapolation of relationships developed in a controlled laboratory environment or at a plot-scale to watershed-scale applications. These shortcomings result in a model structural code built on imperfect process equations. Parameter uncertainty results from the model's use of constant values to represent properties over large areas that, realistically, may be scale dependent and change in space and time. Parameter uncertainty also results from determining the values of parameters in certain process equations that mathematically require variables with no physical meaning. Data uncertainty results from errors in measuring and recording the data used in calibrating the model and as an input to the model. All things considered, structural uncertainty tends to be the dominant source of uncertainty. Probabilistic methods are used to quantify

uncertainty and model results are typically presented with confidence bounds such as a 95% confidence interval (e.g., Vrugt et al. 2003, Abbaspour et al. 2007).

The topic of calibration is common in the environmental modeling literature, and numerous techniques, programs, and methodologies have been developed. Some examples are SCE (Duan et al. 1992), GLUE (Beven and Binley 1992; Franks and Beven 1997; Blasone et al. 2008; Beven 2006), SCEM-UA (Vrugt et al. 2003; Blasone et al. 2008), DDS (Tolson and Shoemaker 2007), and GML (Skahill and Doherty 2006). However, despite the copious amount of literature addressing calibration, the topic is still highly contested (e.g. Mantovan and Todini 2006; Beven et al. 2007; Mantovan et al. 2007). Although a universally accepted methodology for calibration does not exist, a perusal of the relevant literature reveals four key concepts common to most calibration approaches:

Identification of sensitive parameters. Since most models have many parameters, sometimes many dozen parameters, a group of sensitive parameters are typically selected for calibration. Changes in the sensitive parameters more predominantly affect the model output than do changes in the other parameters, and the less-sensitive parameters are considered negligible in the calibration process and are typically held constant at default values. Included with the selection of sensitive parameters is usually a stipulation of the parameter space to be explored in the calibration of each parameter (e.g. Jia and Culver 2008), where the parameter space can be the entire range allowed for that parameter within the model framework or a refined sub-region of the space selected due to prior knowledge of the parameter. Examples of modeling studies which have identified sensitive parameters and corresponding parameter space search ranges are Doherty and Johnston (2003) and Skahill et al. (2009) in HSPF, Cho et al. (2010) and Glavan et al. (2011) in SWAT, and Singh et al. (2005) in both HSPF and SWAT.

Selection of Objective Function. Objective functions (OFs) are used to determine how well the output of a given parameterization of a model fits the data (e.g. Moriasi et al. 2007;

Legates and McCabe 1999). While the use of multiple OFs in the calibration of model parameters generally uses more of the information inherent in a data set and prevents the parameter estimates from focusing disproportionally on one component of the model (Haan 1989), the choice of OF(s) is critical as different OFs will likely generate different calibrated parameter sets (e.g. Al-Abed and Whiteley 2002; Yilmaz et al. 2010), and sometimes the use of a single OF is appropriate. Most OFs employ a least-squares technique between the observations and model predictions (e.g. Chin 2009; Legates and McCabe 1999; Troutman 1985; Brion et al. 2002; Brion and Lingireddy 2003). A least-squares approach is advantageous, since the error that results from the difference between observations and model predictions inherently incorporates the components of model structural uncertainty, parameter uncertainty, and data uncertainty (Hantush and Kalin 2008). OFs in hydrological studies commonly match average streamflows or volumes at various time steps, such as daily or monthly average streamflows or monthly or yearly flow volumes (e.g. Singh et al. 2005; Feyereisen et al. 2007). Log-transformation of data and model output may be used to minimize the effect of peaks in the data. A simple example of a multiple-OF weighting scheme is described by Haan (1989) as

$$W_p = w_1 \sum e_1^2 + w_2 \sum e_2^2 + \ldots \tag{4.1}$$

where W_p is the weighted prediction, the subscripts refer to different objectives, w is a weight and e is the error.

Establishing a Default System State. Since several calibration trials are often required to calibrate or analyze a system (e.g. Duan et al. 1992; Doherty and Johnston 2003), a default starting state is required to objectively evaluate the success of each trial. A default starting state guarantees that all trials begin on an equal playing field and allows the use of different starting points for different trials. This is advantageous because different starting points

often yield different end points, since the trials likely followed various paths or trajectories through the parameter space en route to the final calibrated parameter set (Beven 2006; Skahill et al. 2009). The selection of starting point is particularly important, as a starting point near optimal values of the system may drastically reduce calibration computational time and overall provide a more efficient calibration (Skahill et al. 2009; Yang et al. 2007; Al-Abed and Whiteley 2002). The establishment of a default starting state is important for both manual and automated calibration methods.

Interpretation and Implementation of Calibrated Parameters. Two discussions of calibration philosophy commonly arise in the interpretation and implementation of calibrated parameter sets. The first discussion revolves around whether a single, optimal parameter set really exists or whether there are more than one (possibly many) parameter sets that achieve the same "optimal" model calibration. A conceptual analogy of the discussion is whether the parameter space has a corresponding likelihood/probability surface with a single (maximum) peak or one with multiple equal peaks. (Conversely, if the calibration goal is to minimize an objective function instead of maximize a likelihood function, the analogy is of a bowl with a single low point or of one with several basins with low points of equal depth (e.g. Skahill and Doherty 2006)). On one hand, researchers and modelers insist, perhaps idealistically, that a single optimal set can be found (Beven 2006; Journel 1997). On the other hand, most modeling studies acknowledge that as a given model increases in complexity and parameterization, that parameter correlation and interaction also increases, where "correlation is the term used to describe the phenomenon whereby two or more parameters can be varied in harmony in such a way as to have virtually no effect on the calibration objective function" (Doherty and Johnston 2003). The net effect of correlation is multiple parameter sets that achieve the same level of objectivity (e.g. Duan et al. 1992; Troutman 1985; Beven 2006; Al-Abed and Whiteley 2002; Skahill et al. 2009; Jia and Culver 2008). Conceptually, the parameter space would have a region of higher likelihood than other regions, but the high region of the likelihood surface would be

flat and broad instead of a well-defined peak. Vrugt et al. (2003) explain that "Most likely, a search conducted on the feasible parameter space close to the global optimum will reveal many behavioral parameter sets with quite similar performance in reproducing the observed data." Doherty and Johnston (2003) add that "Calibration of a model such as HSPF can rarely be achieved without at least some degree of nonuniqueness in the estimated parameters". This latter take on the discussion is championed by the Generalized Likelihood Uncertainty Estimation (GLUE) methodology (Beven and Binley 1992), which supports the concept of *equifinality*, where it is acknowledged that model, parameter, and data uncertainties prevent the identification of a single optimal parameter set. Blasone et al. (2008) explain that "the GLUE approach calls for rejecting the concept of a unique global optimum parameter set within some particular model structure, instead recognizing the acceptability, within a model structure, of different parameter sets that are similarly good in producing fit model predictions." One way of reducing the effect of correlation is by removing one or more of the correlated parameters from the calibration process and instead holding them at constant values (Troutman 1985).

The second discussion revolves around whether parameter estimates should be considered as physically realistic or not. Some modeling studies place noteworthy emphasis on deriving and using parameters with realistic ranges and values (e.g. Al-Abed and Whiteley 2002; Feyereisen et al. 2008; Glavan et al. 2011). However, an inconsistency with this viewpoint arises when equifinality in a model yields more than one parameter set that achieve the same level of model optimization while using parameterizations that reflect different physical processes (e.g. Jia and Culver 2008). In such a case it is clear that model uncertainty is too prohibitive to derive any physically realistic meaning out of the parameters that optimize the model. The modeler is then presented with a problem of 'decidability' (Beven 2006), where a decision must be made about which parameter set to use for the calibrated model, even if it means selecting a parameter set that is not realistic but provides the best model fit to the data. Troutman (1985) expresses the sentiment that if a

physically-realistic parameterization of a physically-based model yields predictions that match the data, then the model is acceptably deterministic, but "if a calibrationlike definition of the parameters gives values that are totally at odds with what one would expect physically, then the model probably is less deterministic than we would like to believe, and it makes no sense to expect to be able to use physically realistic parameter values in it and obtain decent runoff predictions in the first place."

In summary, Al-Abed and Whiteley (2002) appropriately comment that "when a model has a large number of parameters that need calibration, when the connection of these parameter values to physical processes is not straightforward, and when parameters are strongly interrelated, successful completion of a calibration exercise is a real challenge."

4.2 IMPLEMENTATION

Based on previous studies (Van Liew et al. 2003; Chin et al. 2009), the sensitive parameters governing the hydrology and bacteria processes in HSPF and SWAT were selected for calibration. These parameters are listed with a brief description in Tables 4.1 and 4.2 along with the total parameter ranges as defined by the upper and lower limits in HSPF and SWAT. Select regions of the total ranges were examined in this study, and these refined ranges are also shown in Tables 4.1 and 4.2. The refined ranges are a combination of values recommended in the literature and values determined heuristically.

The key HSPF hydrology parameters used in model calibration were: INFILT, the parameter describing infiltration capacity of the soil (Eq. 3.1); LZSN, the nominal storage in the lower zone (Eq. 3.1); INFILD, the ratio of maximum to mean infiltration capacity over the land segment (Eq. 3.2); INFEXP, an exponent in the infiltration equation (Eq. 3.1); UZSN, the nominal storage in the upper zone; DEEPFR, the fraction of groundwater inflow which enters the deep (inactive) groundwater; AGWRC, the basic groundwater recession rate if KVARY is zero and there is no inflow to groundwater (Eq. 3.9); IRC, the ratio of

present rate of interflow outflow to the value 24 hours earlier, if there was no inflow (Eqs. 3.6-3.7); KVARY, a parameter which affects the behavior of groundwater recession flow, enabling it to be non-exponential in its decay with time (Eq. 3.8); INTFW, the interflow inflow parameter (Eq. 3.4); and KS, a weighting factor for hydraulic routing. The key HSPF water-quality parameters used in model calibration were: FSTDEC, the in-stream first-order decay rate (Eq. 3.12); ACQOP, the land accumulation of bacteria (Eq. 3.10); WSQOP, the parameter relating land accumulation to concentrations in washoff, defined as the rate of surface runoff that removes 90% of the accumulated bacteria in one hour (Eq. 3.11); IOQC, the concentration of bacteria in interflow; and PSRC, a custom input to HSPF as a mass flux from a direct source into the receiving stream.

The key SWAT hydrology parameters used in model calibration were: CN2, the NRCS runoff curve number for moisture condition II (Eq. 3.14); GW_REVAP- the revap coefficient governing the maximum amount of water that will be removed from the aquifer via 'revap' on a given day (Eq. 3.21); REVAPMN, a threshold water level in the shallow aquifer for revap to occur, such that revap only occurs when the water level is above the threshold; GWQMN, a threshold water level in the shallow aquifer for base flow to occur, such that groundwater contributes to the main channel only when the water level is above the threshold (Eq. 3.19); RCHRG_DP, a percolation coefficient governing the amount of water moving from the shallow aquifer to the deep aquifer on a given day (Eq. 3.18); ALPHA_BF, the baseflow recession constant (Eq. 3.19). Values vary from 0.1-0.3 for land with slow response to recharge to 0.9-1.0 for land with a rapid response; GW_DELAY, the delay time or drainage time for aquifer recharge, to account for the lag that may occur between the time that water exits the soil profile and enters the shallow aquifer as a result of geologic properties and the depth of the water table (Eq. 3.17); CH_K(2), effective hydraulic conductivity of the channel; CH_N(2), Manning's "n" value for the main channel.

The key SWAT water-quality parameters used in model calibration were: CFRT_KG, the land application rates of manure; BACTMIX, the bacteria percolation coefficient (Eq. 3.23); BACTKDQ, the bacteria soil partitioning coefficient (Eq. 3.24); WDPRCH and WDPQ, the die-off coefficients for in-stream and soil solution bacteria, respectively (Eq. 3.29); and BCNST, a custom input to SWAT as a direct NPS flux concentration. In SWAT, BACTMIX affects the leaching of bacteria in the soil solution to lower soil layers, such that as BACTMIX increases, the leaching to lower soil levels decreases, and vice-versa. Similarly, BACTKDQ affects the transport via surface runoff of bacteria in soil solution, such that as BACTKDQ increases, the amount of soil solution bacteria transported in runoff decreases, and vice-versa. BACTKDQ can be thought of as the ratio of concentrations of bacteria in the soil solution of the surface 10 mm to the surface runoff.

The HSPF models were built to model the fate and transport of bacteria in overland flow and interflow (concentrations in groundwater and sediment were not considered), where bacteria deposition was applied directly to the land surfaces. The SWAT models only considered persistent bacteria and used a washoff factor of 1.0 so that all bacteria deposited on foliage are washed off. Bacteria die-off was only considered in the soil solution phase. For the purposes of facilitating the manure application in the model, the manure was assigned as *beef* manure with a bacteria density of 10^5 CFU/g. Impervious areas were absent or minimal in the study areas.

4.3 THEORY

Calibration of the HSPF and SWAT models was done using the sensitive parameters as listed in Tables 4.1 and 4.2. The models were first calibrated for hydrology, then for water quality. The following single objective function (OF) was used for the hydrology calibration

Table 4.1: HSPF sensitive parameter ranges and descriptions

HSPF Hydrology				
Parameter	Units	Total Range	Refined Range	Description
INFILT	mm/hr	2.54E-3 – 2540	0.0254 – 7.62	Infiltration capacity index
LZSN	cm	0.0254 – 254	0.127 – 20.3	Lower zone nominal storage
INFILD	–	1 – 2	1 – 2	Ratio between maximum and mean infiltration
INFEXP	–	0 – 10	0 – 5	Exponent in the infiltration equation
UZSN	cm	0.0254 – 25.4	0.254 – 12.7	Upper zone nominal storage
DEEPFR	–	0 – 1	0 – 1	Fraction of groundwater lost from the system
AGWRC	1/d	0.001 – 0.999	0.1 – 0.999	Basic groundwater recession rate
IRC	1/d	1E-30 – 0.999	0.1 – 0.999	Interflow recession parameter
KVARY	1/cm	0 – ∞	0 – 1.97	Groundwater recession flow parameter
INTFW	–	0 – ∞	0 – 15	Interflow inflow parameter
KS	–	0 – 0.99	0 – 0.99	Weighting factor for hydraulic routing
HSPF Water Quality				
ACQOP_AG	FC/ha-d	0 – ∞	2.47 – 2.47E7	FC accumulation rate, Agricultural land
ACQOP_FOR	FC/ha-d	0 – ∞	2.47 – 2.47E7	FC accumulation rate, Forest land
WSQOP	cm/hr	0.0254 – ∞	0.0254 – 0.508	Rate of surface runoff that removes 90% of FC in 1 hour
IOQC	FC/m^3	0 – ∞	1.77E5 – 1.77E9	FC concentration in interflow
PSRC	FC/d	0 – ∞	5.0E7 – 5.0E9	Mass flux from direct source
FSTDEC	1/d	1E-5 – ∞	0.1 – 8	First-order decay rate of FC in stream

Table 4.2: SWAT sensitive parameter ranges and descriptions

SWAT Hydrology				
Parameter	Units	Total Range	Refined Range	Description
CN2_AG	–	0 – 100	30 – 98	Runoff curve number for moisture condition II, Agricultural land
CN2_FOR	–	0 – 100	30 – 98	Runoff curve number for moisture condition II, Forest land
GW_REVAP	–	0 – 1.0	0.02 – 1.0	Groundwater "revap" coefficient
REVAPMN	mm	0 – ∞	0 – 5	Depth in shallow aquifer for percolation to deep aquifer
GWQMN	mm	0 – ∞	0 – 1	Threshold depth in shallow aquifer for return flow
RCHRG_DP	–	0 – 1	0 – 1	Deep aquifer percolation fraction
ALPHA_BF	d	0 – 1	0.01 – 1	Baseflow recession constant
GW_DELAY	d	0 – ∞	0 – 50	Groundwater delay time
CH_K(2)	mm/hr	0 – ∞	0 – 300	Effective hydraulic conductivity
CH_N(2)	–	0.01 – ∞	0.01 – 0.3	Manning's *n* in main channel
SWAT Water Quality				
CFRT_KG_Ag	kg/ha-d	0 – ∞	1 – 4000	Manure application rate, Agricultural land
CFRT_KG_For	kg/ha-d	0 – ∞	1 – 8000	Manure application rate, Forest land
BACTKDQ	m^3/Mg	0 – ∞	0 – 50	Bacteria soil partitioning coefficient
BACTMIX	10 m^3/Mg	1.0E-6 – ∞	1 – 20	Bacteria percolation coefficient
WDPRCH	1/d	0 – ∞	0 – 5	In-stream bacteria die-off coefficient
WDPQ	1/d	0 – ∞	0 - 0.5	Soil solution bacteria die-off coefficient
BCNST	CFU/100mL	0 – ∞	1.0E6 – 1.0E10	Direct-source concentration, flow = 0.01 m^3/d

$$L_H\left(\mathbf{y}|\theta_i,\hat{\theta}_{j\neq i}\right)=\frac{1}{S^N}\prod_{i=1}^{N}\exp\left[\frac{-1}{2S^2}(y_i-m_i)^2\right] \tag{4.2}$$

where L_H is the conditional likelihood for hydrology parameter θ_i, $\boldsymbol{y}$ is the set of N measured data (i.e. average daily flows), θ_i is the value of the parameter being varied, $\hat{\theta}_{j\neq i}$ is the set of values of the other parameters being held constant, and S is the standard deviation of the measurements (y_i) relative to the model predictions (m_i) and is given by

$$S=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i-m_i)^2} \tag{4.3}$$

While some hydrological studies calibrate with multiple OFs that examine weekly, monthly, and/or yearly flow and volumes (e.g., Jia and Culver 2008; Doherty and Johnston 2003; Al-Abed and Whiteley 2002; Haan 1989), the use of a single, daily-based OF in this case is justified because of its use in a bacteria model, where a daily time step is more consistent with bacterial dynamics than larger time steps. The Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) was the metric used to measure model accuracy, and is described by

$$NSE=1-\frac{\sum_{i=1}^{N}[y_i-m_i]^2}{\sum_{i=1}^{N}[y_i-\bar{y}]^2} \tag{4.4}$$

where N, y_i, and m_i are previously defined and $\bar{y}$ is the mean of the measurements. While the NSE is one of the most common statistics used in hydrological modeling (e.g. Feyereisen et al. 2007a; Arnold et al. 2010; Eckhardt et al. 2003; Paul et al. 2004; Laroche et al. 1996; Chin et al. 2009; Moriasi et al. 2007; Gupta and Kling 2011), it may be sensitive to extremely large values (Legates and McCabe 1999; Chin et al. 2009). One way of reducing the effect of large values is by log-transforming the data and the model predictions.

Since bacteria concentrations are commonly observed to be lognormal (e.g., USEPA 2001; Wilkinson et al. 1995), the water quality calibration was similarly performed using the following objective function:

$$L_{\mathrm{WQ}}\left(\mathbf{y}|\theta_i,\hat{\theta}_{j\neq i}\right)=\frac{1}{S_{\ln}^{N}}\prod_{i=1}^{N}\exp\left[\frac{-1}{2S_{\ln}^{2}}(\ln y_i-\ln m_i)^2\right] \tag{4.5}$$

where L_{WQ} is the conditional likelihood for water quality parameter θ_i, $\boldsymbol{y}$ is the set of N measured data (FC data), θ_i is the parameter being varied, $\hat{\theta}_{j\neq i}$ is the set of values of the other parameters being held constant, and $S_{\ln}$ is the standard deviation of the log of the measurements ($\ln y_i$) relative to the log of the model predictions ($\ln m_i$) and is given by

$$S_{\ln}=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\ln y_i-\ln m_i)^2} \tag{4.6}$$

where all terms are previously defined. Log-scaled NSE values ($\mathrm{NSE}_{\ln}$) were calculated according to

$$NSE_{\ln}=1-\frac{\sum_{i=1}^{N}[\ln y_i-\ln m_i]^2}{\sum_{i=1}^{N}[\ln y_i-\ln \bar{y}]^2} \tag{4.7}$$

To assess the assumption of normal residuals required by the preceding likelihood equations, the Shapiro-Wilk statistic (Shapiro and Wilk 1965) was used, and normality was assessed at the 95-percent confidence limit.

4.4 PROCEDURE

The calibration procedure used was a response-surface iterative scheme that searches the parameter space looking for regions of highest likelihood. The algorithm is detailed in Chin

(2009) and is repeated here verbatim with permission. For a water-quality model with p parameters, θ_1, ..., θ_p and evaluation of the model at n discrete values of each parameter in the defined parameter space, the procedure is:

STEP 1: Initialize the parameter vector as $\theta^{(1,1)}=[\theta_1, \theta_2^{(0)}, ..., \theta_p^{(0)}]^T$ where the "(1,1)" superscript for θ indicates the first iteration on the first parameter (θ_1), and the "(0)" superscript for θ_i indicates that θ_i is held constant at its initial (zero-iteration) value. The conditional likelihood function for θ_i is calculated using the relation

$$L(\theta_1|\theta_2^{(0)}, ..., \theta_p^{(0)}, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(1,1)})]^2 y_i^{\lambda-1} \right] \tag{4.8}$$

The maximum-likelihood value of θ_1 is determined from the conditional likelihood function and is denoted by $\theta_1^{(1,*)}$.

STEP 2: Update the parameter vector to $\theta^{(1,2)}=[\theta_1^{(1,*)}, \theta_2, \theta_3^{(0)}, ..., \theta_p^{(0)}]^T$ and calculate the conditional likelihood function for θ_2 using the relation

$$L(\theta_2|\theta_1^{(1,*)}, \theta_3^{(0)}, ..., \theta_p^{(0)}, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(1,2)})]^2 y_i^{\lambda-1} \right] \tag{4.9}$$

The maximum-likelihood value of θ_2 is determined from the conditional likelihood function and is denoted by $\theta_2^{(1,*)}$.

STEP 3: Repeat Step 2 for all parameters to complete the first iteration which ends with taking the parameter vector as $\theta^{(1,p)}=[\theta_1^{(1,*)}, ..., \theta_{p-1}^{(1,*)}, \theta_p]^T$ and calculating the conditional likelihood function for θ_p using

$$L(\theta_p|\theta_1^{(1,*)}, ..., \theta_{p-1}^{(1,*)}, \theta_p, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(1,p)})]^2 y_i^{\lambda-1} \right] \tag{4.10}$$

STEP 4: Start the second iteration by taking the parameter vector as $\theta^{(1,2)}=[\theta_1, \theta_2^{(1,*)}, ..., \theta_p^{(1,*)}]^T$ where the "(1,*)" superscript for θ_i indicates that θ_i is held constant at its first-iteration maximum-likelihood value. The conditional likelihood function for θ_i is calculated using the relation

$$L(\theta_1|\theta_2^{(1,*)}, ..., \theta_p^{(1,*)}, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(1,2)})]^2 y_i^{\lambda-1} \right] \tag{4.11}$$

STEP 5: Repeat Step 4 for all parameters to complete the second iteration, which ends with taking the parameter vector as $\theta^{(2,p)}=[\theta_1^{(2,*)}, ..., \theta_{p-1}^{(2,*)}, \theta_p]^T$ and calculating the conditional likelihood function for θ_p using

$$L(\theta_p|\theta_1^{(2,*)}, ..., \theta_{p-1}^{(2,*)}, \theta_p, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(2,p)})]^2 y_i^{\lambda-1} \right] \tag{4.12}$$

STEP 6: Repeat the iterations described by Steps 4 and 5 until the conditional likelihood function converges to its asymptotic value. The j-th iteration ends with the parameter vector given as $\theta^{(2,p)}=[\theta_1^{(2,*)}, ..., \theta_{p-1}^{(j,*)}, \theta_p]^T$ and the conditional likelihood function for θ_p is given by

$$L(\theta_p|\theta_1^{(j,*)}, ..., \theta_{p-1}^{(j,*)}, \theta_p, \mathbf{y}, \lambda) = \frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} [z_i - g_i^{(\lambda)}(\theta^{(j,p)})]^2 y_i^{\lambda-1} \right] \tag{4.13}$$

The final iteration cycle provides the maximal conditional likelihood function of all the parameters, where the maximal conditional likelihood function for each parameter is defined as the likelihood function for that parameter conditioned on the other parameters being at their maximum-likelihood values. For example, the maximal conditional likelihood function of θ_p is given by Equation 4.13.

4.5 COMPUTATIONAL DETAILS

The implementation of the calibration algorithm utilizes a system of inner and outer loops, where the outer loop represents the P sensitive parameters and the inner loop searches the designated parameter space of an individual parameter. Starting with default parameter values and the parameter ranges listed in Tables 4.1 and 4.2, the inner loop divides the initial parameter range of the first parameter in the outer loop into N equally spaced divisions by evaluating $N+1$ model runs, while all other parameters are held at default values. At each model run, the corresponding conditional likelihood is determined, and after $N+1$ model runs, the parameter is set to the value that generated the highest (conditional) likelihood. The parameter range of the evaluated parameter is then refined to focus sequential searches in the region of highest likelihood. The outer loop then increments to the next parameter in sequence, and the process is repeated: the inner loop divides the initial parameter space of the second parameter into N equally spaced divisions ($N+1$ model runs) and the corresponding conditional likelihoods are evaluated, the parameter is set to the value of the $N+1$ model runs that generated the highest likelihood, the parameter range is refined to focus sequential iterations of that parameter in the region of highest likelihood, and the outer loop is incremented to the next parameter. The process of searching the parameter space, identifying parameter values of highest likelihood, and refining the parameter space for future runs is repeated for the P parameters until a parameter convergence threshold is reached or a maximum of M iterations through the outer loop has occurred.

A Matlab code was developed to automate the calibration of the P sensitive parameters in HSPF (11 hydrology, 6 WQ) and SWAT (10 hydrology, 7 WQ). In all cases, a burn-in period of length $2P$ outer-loop increments was established, where at the end of the inner loop of these iterations the parameter space was not refined but was re-established as the initial parameter range. This was shown to improve overall calibration results and efficiency, guaranteeing that the first two passes of each parameter were not too affected by

default values. N was set to 20 divisions (21 model runs) for the first seven complete outer-loop increments (7P) in both HSPF models and the SWAT hydrology model and the first ten outer-loop increments (10P) in the SWAT WQ model, and (N) was reduced to six (seven model runs) for the remainder of the calibration, if necessary. In most cases it was found that seven passes through each parameter sufficiently refined the parameter search space for each of the P parameters, and the reduction of N aided in final convergence and greatly reduced computational time. If, during an intermediary inner loop, the parameter value generating the highest likelihood was found to be at the limit of the refined parameter range (but not at an upper or lower limit of the considered parameter space as listed in Tables 4.1-4.2), the parameter range of that parameter was broadened for the next outer loop iteration of that parameter.

On an Intel Core 2, 2.4 GHz PC, a single model run of HSPF and SWAT took approximately 2.5 and 0.75 seconds, respectively, and a single outer loop of 21 model evaluations (including likelihood calculations) took approximately 140 and 25 seconds in HSPF and SWAT, respectively. For the hydrology calibrations, the outer-loop exit limit M was set to 201, so that in cases where the model parameters did not converge, a maximum of [(21 model runs)×(77 outer loop increments)+(7 model runs)×(201-77 outer loop increments)] = 2,485 total model runs in HSPF and [(21 model runs)×(70 outer loop increments)+(7 model runs)×(201-70 outer loop increments)] = 2,387 model runs in SWAT. In cases when the models did not converge, reaching M = 201 took roughly 6 hours for HSPF and 40 minutes for SWAT. For the water quality calibrations, M was set to 101 in HSPF and 151 in SWAT, due to both models having fewer parameters for water quality than for hydrology, and because the simulation in SWAT was less time-intensive. Additionally the threshold for switching from N = 21 to N = 7 in SWAT was done after 10P increments of the outer loop. In cases when the model parameters did not converge, a maximum of [(21 model runs)*(42 outer loop increments)+(7 model runs)*(101- 42 outer loop increments)] = 1,295 total model runs occurred in HSPF and [(21 model runs)*(70 outer loop increments)+(7 model

runs)*(151-70 outer loop increments)] = 2,037 model runs occurred in SWAT. While the calibration method described here is temporally and computationally intensive, it is not uncommon for the calibration of a model to require thousands of model evaluations. For example, Franks and Beven (1997) and Eckhardt et al. (2003) used at least 10,000 model evaluations to sample the parameter space for the calibration of their models, while Skahill and Doherty (2006) used a calibration scheme that required about 52,000 model evaluations and took almost 2.5 days to complete. The study by Razavi et al. (2010) specifically examines ways of improving the computational efficiency of automatic calibration methods.

It was found in many cases in both HSPF and SWAT that starting the automated calibration with different parameters often yielded different calibration results. In other words, from the established default starting state of the system at the beginning of the procedure, selecting different parameters as the starting point would yield different calibrated parameter sets. This behavior was found true in both the hydrology and water-quality calibrations. For example, it was found that starting the hydrology calibration of HSPF in Catchment J with the parameter UZSN, then proceeding in turn to DEEPFR, AGWRC, etc., as listed in Table 4.1 yielded a different calibrated parameter set with higher NSED than starting the calibration with parameter KS and proceeding to INFILT, LZSN, etc. Similarly in SWAT, it was found that starting the hydrology calibration in Catchment I with the parameter CH_K(2), then proceeding in turn to CH_N(2), CN2_Ag, etc., as listed in Table 4.2, yielded a different calibrated parameter set with higher NSED than starting the calibration with parameter CN2_For and proceeding to GW_REVAP, GWQMN, etc. Implementing the burn-in loop helped reduce this dependence on starting point and overall provided consistently better results. But the burn-in loop did not eliminate the relationship between starting parameter and final calibration result. To ensure that the most optimal parameter set (and therefore highest NSED) was determined, an automated calibration was conducted from a default state for each model in each catchment, using each parameter as a starting point and proceeding in the order shown in Tables 4.1 and 4.2. So 11 automated

hydrology calibrations in HSPF and 10 automated hydrology calibrations in SWAT were conducted in each catchment. It was generally found that most (not all) of the calibration attempts, aided by the burn-in loop, converged to a similar region of maximum likelihood and parameter values. As explained earlier (under the subheading “Interpretation and Implementation of Calibrated Parameters”) this behavior is indicative of both a single-peaked yet broad and flat likelihood surface, as well as parameter correlation that leads to non-unique parameter sets in the region near the peak. The hydrology parameter set in each model in each catchment that yielded the highest NSED was used as the hydrology component in the corresponding water-quality model. Then, using the optimal hydrology parameter sets, 6 automated water-quality calibrations in HSPF and 7 automated water-quality calibrations in SWAT were conducted in each catchment.

Chapter 5

A Multi-Model, Multiple Watershed Examination of In-Stream Bacteria Modeling

5.1 BACKGROUND

HSPF and SWAT are considered state-of-the-art watershed models (Van Liew et al. 2003). Perhaps more importantly, the models have a stamp of approval from the EPA (HSPF) and the USDA (SWAT), signifying that despite their limitations the models are and will be considered as standards in environmental modeling. The influence and importance of these models reaches beyond domestic borders as studies using the models have been carried out in countries such as China (Yang et al. 2007), Canada (Al-Abed and Whiteley 2002; Laroche et al. 1996), South Korea (Lee et al. 2010), Mexico (LaWare and Rifai 2006), Ireland (Coffey et al. 2010), France (Bougeard et al. 2011; Baffaut and Sadeghi 2010), Switzerland (Abbaspour et al. 2007), and the UK (Glavan et al. 2011).

Although the models are widely used and receive government backing, they represent two very different formulations of the physical world. In the context of bacteria modeling the most notable differences are: the hourly time resolution in HSPF vs. the daily time resolution in SWAT; the division of watershed area into pervious and impervious segments

in HSPF vs. the division of area into HRUs based on soil type and land use in SWAT; the generic water-quality component in HSPF vs. the bacteria-specific water-quality component in SWAT; the existence of bacteria concentrations in subsurface waters in HSPF vs. the assumption that bacteria concentrations in subsurface waters are zero in SWAT; and the surface layer, upper zone, lower zone, and groundwater layer watershed structure in HSPF vs. the shallow and deep aquifer subbasin structure in SWAT. In light of these differences it is fair to assume that the models will perform differently given the same location and data. This chapter presents the primary HSPF and SWAT models in the LREW that are the basis of the research in this dissertation and provides an analysis of the model performance.

The CC model is included for comparison with the other two models. Although the CC model has a much simpler formulation than HSPF and SWAT, it has two major advantages over the other models. First, it uses measured streamflow data instead of a modeled hydrological component. Using the data eliminates the significant model uncertainty that results from imperfect hydrologic process equations and the use of numerous hydrological parameters. In using the streamflow data, the daily hydrology NSE of the CC model is guaranteed to be 1.00. Second, the 2-parameter water-quality component of the CC model is conceptually simple, induces much less parameter uncertainty or risk of parameter correlation than HSPF and SWAT, and is easily calibrated.

The HSPF, SWAT, and CC models were developed for the LREW as described in Section 3.1. The HSPF and SWAT models were calibrated as described in Chapter 4. Since the CC model only has two parameters, a simple least-squares calibration routine was implemented to find values of c_r and c_b that maximize the NSE_{ln} (Eq. 4.7). The results presented are for the calibrated models; validation and forecasting of the models were not considered. While validation is normally an important aspect of a modeling effort, two factors led to the consideration of only the calibrated models. First, the FC datasets were sufficiently limited to further divide into a set for calibration and a set for validation, so the whole dataset was

used for calibration. Secondly, when calibration and validation phases are conducted for a model, the calibrated models typically achieve the best performance (e.g. NSE_{ln}) and the results presented are from the calibrated models. This fact is true in the previous work in the LREW, where the majority of the results were based on the calibrated versions of the models (e.g. Feyereisen et al. 2007; White et al. 2009; Chin et al. 2009; Sakura-Lemessy 2009). Since this study was an examination of the model performance, and since the previous work in the LREW was presented for calibrated models, the calibrated models were only considered here.

5.2 RESULTS AND DISCUSSION

5.2.1 HSPF and SWAT Hydrology Calibration Results

The calibrated maximum likelihood (ML) parameters of the HSPF and SWAT hydrology models are shown in Tables 5.1 and 5.2. It is apparent that for each model there is little

Table 5.1: HSPF hydrology calibration results

Parameter	Units	Default*	Catchment			
			I	J	K	O
INFILT	mm/hr	2.54	7.16	5.62	1.78	2.17
LZSN	cm	15.2	1.12	0.864	3.99	6.86
INFILD	–	1.91	1.80	1.70	2.00	1.03
INFEXP	–	2.00	3.49	3.05	1.02	2.48
UZSN	cm	2.82	8.10	8.36	7.82	4.75
DEEPFR	–	0.100	0.269	0.219	0	0.260
AGWRC	1/d	0.980	0.974	0.987	0.972	0.958
IRC	1/d	0.500	0.415	0.400	0.373	0.235
KVARY	1/cm	0	0	0.258	0.180	0.223
INTFW	–	0.750	15.0	15.0	5.72	5.78
KS	–	0.50	0	0	0	0

* Note: slight variations of some default values occurred across catchments

consistency among calibrated parameter values between catchments. It is not uncommon for models of nearby watersheds to have different parameterizations (e.g. Laroche et al. 1996) and this behavior has been well documented in the LREW (Feyereisen et al. 2007). It

Table 5.2: SWAT hydrology calibration results

Parameter	Units	Default*	Catchment			
			I	J	K	O
CN2_AG	–	77	52	35	35	64
CN2_FOR	–	55	77	69	70	78
GW_REVAP	–	0.020	0.362	0.933	0.437	0.230
REVAPMN	mm	1.00	0.670	0	0	0.134
GWQMN	mm	0.000	20.0	20.0	20.0	17.5
RCHRG_DP	–	0.050	0	0.135	0	0.288
ALPHA_BF	d	0.048	0.516	0.642	0.658	0.572
GW_DELAY	d	31.0	29.5	0.750	21.5	18.8
CH_K(2)	mm/hr	0	300	300	218	185
CH_N(2)	–	0.014	0.300	0.129	0.037	0.031

*Note: slight variations of some default values occurred across catchments

is important to remember that these parameterizations maximize the model output relative to average daily streamflow, and that it might not be possible to derive consistent or physically realistic conclusions from the parameter relationships shown.

Nonetheless, there are several noteworthy items in the calibrated hydrology parameters. The HSPF parameter INFILD (Eq. 3.2) calibrated to 2.00 in Catchment K. INFILD is an important parameter in defining the relationship of Lines 1 and 2 in Figure 3.2 but has an HSPF-defined upper limit of 2.00, signifying that the maximum infiltration can never be twice as large as the mean infiltration. Yet in Catchment K, if allowed, the parameter would have gone above 2.00 in an attempt to better fit the daily streamflow. In running several calibration attempts on different pervious land segments Lowe (2004) found the median value of INFILD was 2.00, indicating that at least half of the trials resulted in an INFILD value of 2.00. This parameterization is an example of a potential source of structural uncertainty since there appears to be no physical rationale for limiting INFILD to be less than or equal to 2.00. Moreover, this is an example where INFILD, if it was known to correlate with another parameter, could be left at the default value of 2.00 and taken out of the calibration process. Similarly, all four catchments in HSPF have calibrated AGWRC values between 0.900 and 0.999, which is a common result in the literature (e.g. Lowe

2004; Chin et al. 2009; Van Liew et al. 2003). If AGWRC was found to be correlated with another parameter, it could be held at the default value of 0.980 and taken out of the calibration. If, in cases like those highlighted here, a particular parameter is known to be highly correlated with several other parameters, the technique of holding the parameter at a constant value and removing it from consideration in the calibration process will ultimately reduce parameter uncertainty in the model results, reduce fluctuation in the calibrated values of the other parameters and reduce the overall time needed for an automated calibration algorithm (like the one employed here) to calibrate the model.

In SWAT the CN parameters calibrated to lower values in the agricultural areas than the forest areas in all four catchments. Since higher CN values indicate more runoff, the models are representing the agricultural area as being more pervious than the forest areas. While these values may not be realistic in the LREW (Feyereisen et al. 2008), similar values have been successfully used in other LREW models (Chin et al. 2009; Sakura-Lemessy 2009).

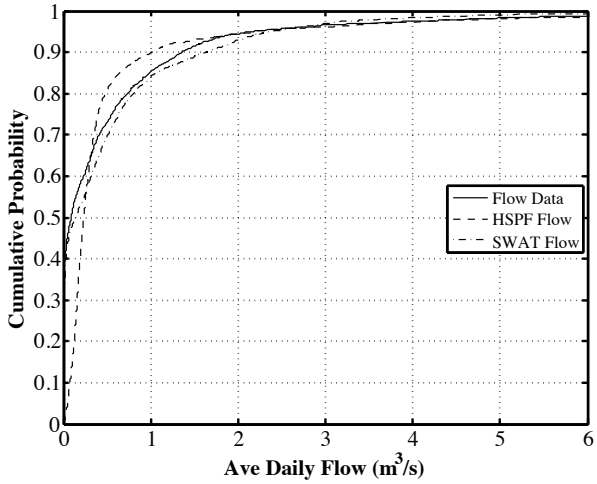
The parameter correlation was determined from the calibration process by tracking the optimal values of each outer-loop cycle of each parameter. For the 10 hydrology parameters in SWAT, for example, if the maximum amount of outer-loop cycles was set to 201 and the calibration did not converge but was stopped upon reaching the maximum limit, then each of the 10 parameters would have at least 20 outer-loop cycles from which an optimal parameter value was selected for that iteration of the calibration. The parameter correlation was then determined between the 20 values of each parameter and the 20 values of each other parameter. It was observed that statistically significant parameter correlation occurred frequently and at times between several parameters, although no consistent relationship between correlated parameters existed and instead the parameter correlation depended on the starting point of the calibration in each model in each catchment.

The NSE values of the calibrated HSPF and SWAT hydrology models are shown in Table 5.3, where the NSE values using a daily time step (NSED) and a monthly time step (NSEM)

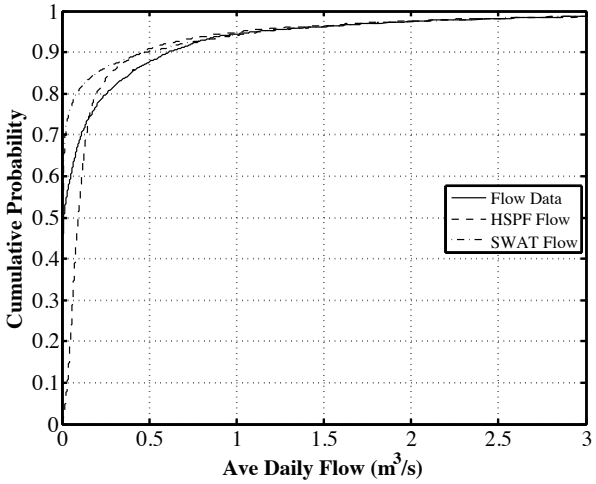
Table 5.3: Summary of HSPF and SWAT model results

Catchment	Model Type	Metric	HSPF	SWAT
I	Hydrology	NSED	**0.884**	0.595
		NSEM	**0.903**	0.897
	Water Quality	NSE_{ln}	0.243	**0.275**
		p value	0.074	0.86
J	Hydrology	NSED	**0.880**	0.626
		NSEM	**0.881**	0.762
	Water Quality	NSE_{ln}	**0.263**	-1.671
		p value	0.57	**3.1E-4**
K	Hydrology	NSED	**0.897**	0.670
		NSEM	**0.894**	0.886
	Water Quality	NSE_{ln}	**0.387**	0.334
		p value	0.82	0.48
O	Hydrology	NSED	**0.941**	0.693
		NSEM	**0.955**	0.932
	Water Quality	NSE_{ln}	**0.342**	-0.028
		p value	0.21	0.10

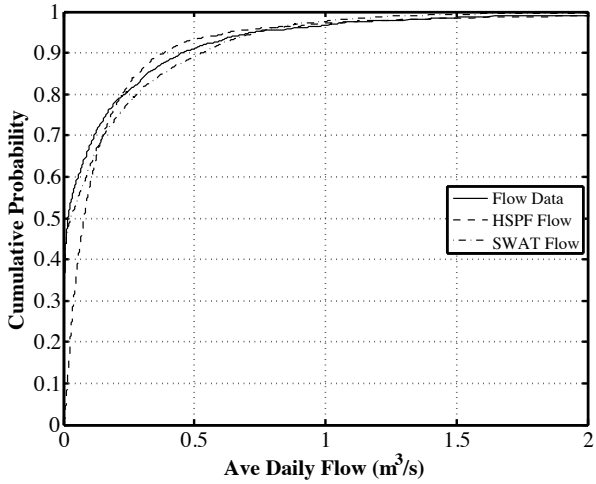
are given. Note that the NSEM values were calculated based on the daily OF described in Equation 4.2; the OF was not first readjusted to a monthly time step before calculating NSEM. The hydrology calibrations achieved in HSPF and SWAT were *satisfactory* ($0.50 < \text{NSE} \leq 0.65$) to *very good* ($\text{NSE} > 0.65$) based on NSE criteria suggested in Moriasi et al. (2007), and are at least comparable if not better than other modeling efforts in the LREW using similar or identical catchments and/or time periods (Feyereisen et al. 2007a; Chin et al. 2009; Sakura-Lemessy 2009). Having strong NSE agreement is important since the water-quality predictions are directly affected by the hydrology components of the models. The hydrology calibrations reveal that HSPF produces a better NSED value than SWAT, a result previously reported in the literature (Van Liew et al. 2003; Chin et al. 2009) and attributed to the fact that HSPF performs calculations on an hourly time step while SWAT uses a daily time step. The frequency distributions of the modeled HSPF and SWAT hydrology components are shown with the measured streamflow for the catchments in Figure 5.1. The models show good agreement in matching the streamflow records at higher flows, roughly the 90%-tile and above, and varying levels of agreement at lower flows. Note



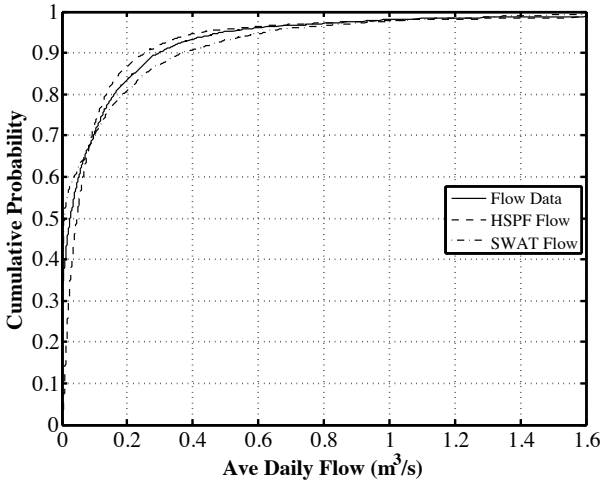
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 5.1: Frequency distributions of streamflow

that the magnitude of the flows correlate with the size of the catchment areas (Table 3.1), which from largest to smallest are I, J, K, and O.

5.2.2 HSPF and SWAT Water-Quality Parameter Sensitivity

The calibrated maximum likelihood (ML) parameters of the HSPF and SWAT water-quality models are shown in Tables 5.4-5.5. As in the hydrology models, there is a variety of parameter values in the models across catchments. Sensitivity plots of the water-quality

parameters in HSPF and SWAT are shown in Figures 5.2 and 5.3, respectively. In these figures the relative likelihood distribution of each parameter over a range of 0.5-2.0 times

Table 5.4: HSPF water-quality calibration results

Parameter	Units	Default*	Catchment			
			I	J	K	O
ACQOP_AG	FC/ha-d	247	7.56E6	9.93	4.62	2.94
ACQOP_FOR	FC/ha-d	247	7.88E6	5.93	3.24	4.77
WSQOP	cm/hr	4.17	0.0254	0.217	0.180	0.508
IOQC	FC/m^3	3530	3.23E6	5.68E6	1.36E7	1.40E7
PSRC	FC/d	1.00E8	2.15E9	2.25E8	1.25E9	1.99E8
FSTDEC	1/d	1.00	0.100	0.100	5.75	0.100

* Note: slight variations of some default values occurred across catchments

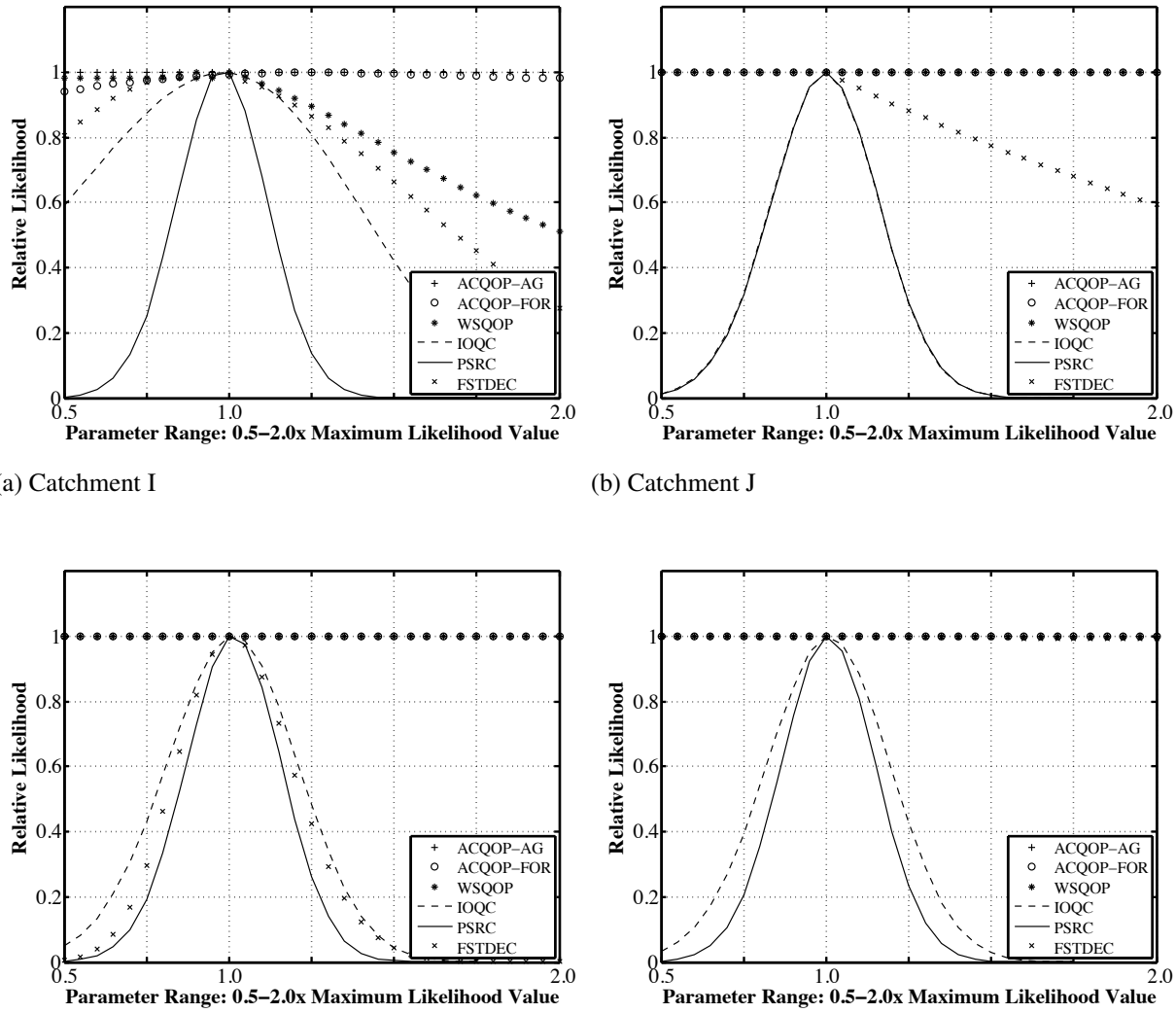
Table 5.5: SWAT water-quality calibration results

Parameter	Units	Default*	Catchment			
			I	J	K	O
CFRT_KG_Ag	kg/ha-d	30.0	0.210	2640	670	0.230
CFRT_KG_For	kg/ha-d	30.0	3330	6330	1180	7830
BACTKDQ	m^3/Mg	175	1.07	1.38	1.45	0.0540
BACTMIX	10 m^3/Mg	10.0	14.1	20.0	18.5	7.81
WDPRCH	1/d	0	2.30	1.85	2.88	3.99
WDPQ	1/d	0	0	0	0	0.128
BCNST	CFU/100mL	1.00E8	2.53E9	8.30E7	2.64E8	2.21E8

* Note: slight variations of some default values occurred across catchments

the ML parameter value is shown, where the rest of the parameters are held constant at their ML values. Since the SWAT parameter WDPQ calibrated to zero in catchments I, J, and K it was not included on the relative scale shown in Figure 5.3, but rather is shown for the actual parameter values in Figure 5.4.

The most notable observation that can be made from the HSPF and SWAT parameter sensitivity plots (Figs. 5.2 and 5.3) is that the parameters directly affecting in-stream processes are clearly the most sensitive (i.e. IOQC, PSRC, and FSTDEC in HSPF and WDPRCH and BCNST in SWAT). This result is a reflection of how the in-stream parameters have a constant effect, even during baseflow conditions, and are less dependent



(a) Catchment I

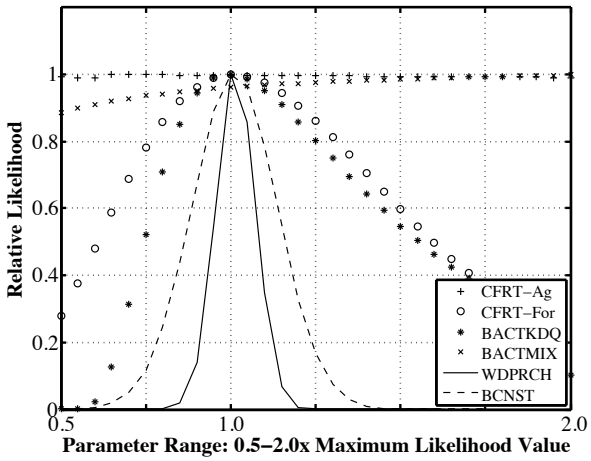
(b) Catchment J

(c) Catchment K

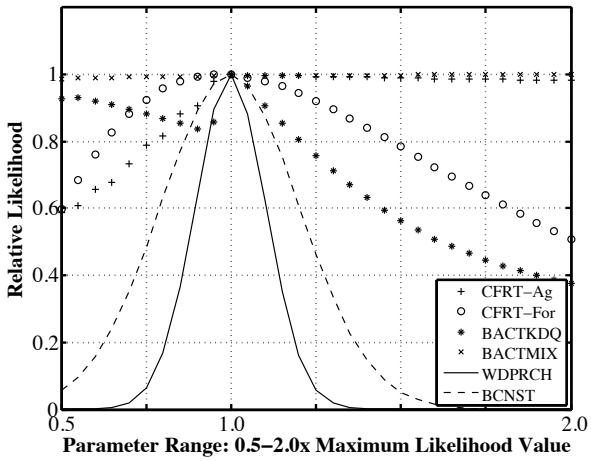
(d) Catchment O

Figure 5.2: HSPF water quality parameter sensitivity

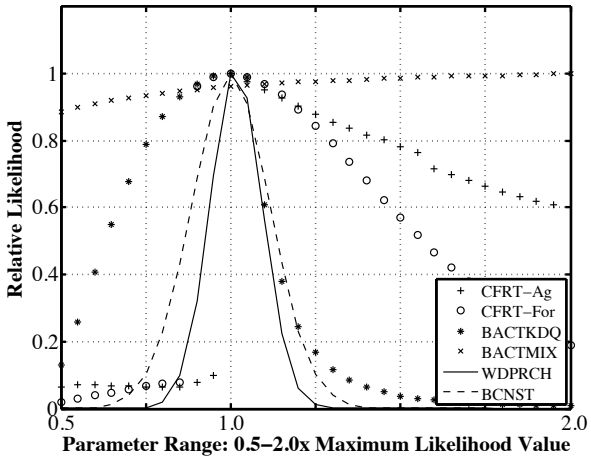
on rainfall/runoff phenomena. The HSPF parameters relating to the terrestrial accumulation (ACQOP) and washoff (WSQOP) of bacteria are notably less sensitive than in-stream parameters and even insensitive in some cases. This is likely due to the decrease in runoff for larger areas as calculated by HSPF and shown in Figure 3.2. The SWAT parameters relating to surface deposition of bacteria (CFRT_KG) show differing levels of sensitivity, from insensitive to very sensitive, which can be attributed to differences in land use and runoff. Crowther et al. (2002) discuss the varying contribution of bacteria from land sources



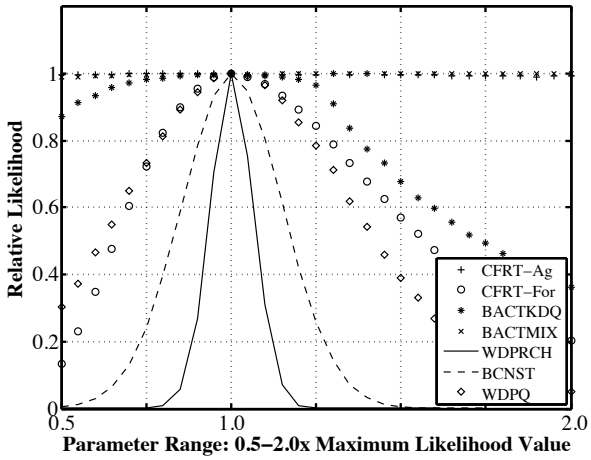
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 5.3: SWAT water quality parameter sensitivity

and suggest that terrestrial sources are more significant in smaller catchments while in larger catchments the land sources are relatively less important than in-stream sources. This phenomena is likely a result of larger catchments having both increased contributions of bacteria from stream bed sediments and longer and more complex land-to-stream bacteria pathways. Similarly, Byappanahalli et al. (2003) found that EC bacteria populations were ubiquitous in riparian sediment, likely independent of fecal contamination, and decreased with distance from the stream. The SWAT parameters BACTKDQ and BACTMIX are key

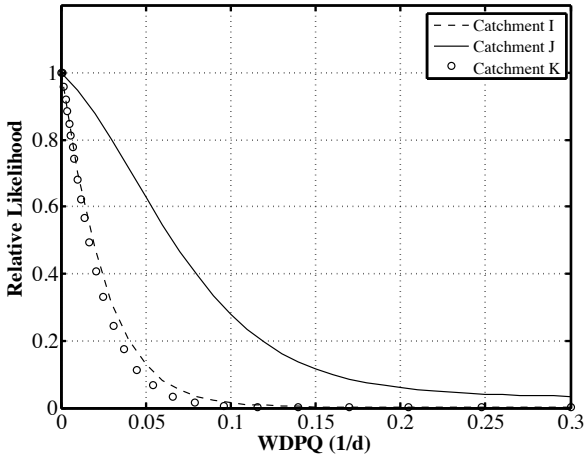


Figure 5.4: Sensitivity of WDPQ in Catchments I, J, and K

parameters in SWAT's unique suite of codes describing bacteria processes. Yet BACTMIX is consistently insensitive while BACTKDQ shows a wide range of sensitivities. These two parameters are an example of how increased parameterization in an attempt to better describe bacteria processes may or may not improve the model. In the case of the insensitive BACTMIX, it appears the extra parameters and process equations are more likely to increase uncertainty instead of add meaning to the model. It is worth noting that the most sensitive hydrological parameter in SWAT is the curve number (in this case CN2_AG and CN2_FOR), a well-known fact which has been subsequently verified within the LREW (Feyereisen et al. 2008, 2007; Cho et al. 2010; Jeong et al. 2010; White et al. 2009; Chin et al. 2009). The curve number of each land use, therefore, will have a direct but varying emphasis on the runoff contribution of bacteria from terrestrial sources. This phenomena is reflected in Figure 5.3, where the parameter describing land-application of manure in forested lands (CFRT_For) is generally more sensitive than the equivalent parameter in agricultural lands (CFRT_Ag). The parameter sensitivity reflected here is likely due to the fact that the forest areas calibrated to higher CN values than the agricultural areas (Table 5.2), indicating that more runoff is generated from forest areas. Although SWAT represents the source of terrestrial bacteria in forest lands as manure

application, in reality no manure would be applied to forested lands as an agricultural practice, and it is more realistic to assume that the source of terrestrial bacteria applied to the forested lands is some sort of wildlife defecation.

To assess terrestrial deposition, most bacteria models use data or make assumptions about animal quantities (i.e. livestock, poultry, etc.), defecation rates, and manure bacteria quantities (e.g. Baffaut and Sadeghi 2010; LaWare and Rifai 2006), although much care must be taken in doing such analyses (Haydon and Deletic 2009). The approach used in this study was to include the equivalent parameters (ACQOP, CFRT_KG) as unknowns in the calibration because most of the agricultural activity in the four catchments is crop-oriented (not livestock-oriented) and because no other data was available. This approach avoids the contribution of uncertainty that results from estimated loading rates and allows the model parameters to freely calibrate to their optimal values without any initial artificial emphasis on particular parameter values. In this manner, in the absence of a loading-bias, the true parameter sensitivities are revealed as shown in Figures 5.2 and 5.3.

The sensitivity and magnitude of the IOQC parameter in HSPF highlights an important and fundamental difference in how HSPF and SWAT address subsurface bacteria concentrations. The calibrated IOQC values shown in Table 5.4 indicate that HSPF is assigning a bacteria concentration of 10^6-10^7 CFU/m^3 to the interflow, which contrasts with SWAT, which has a core assumption that interflow concentration is zero. (Since the CC model is based on hydrograph separation, which divides total flow into runoff and baseflow, the CC model does not directly address interflow concentrations.) In reality, intact and uniform soils act as a filter and bacteria transport through such soils is highly unlikely (McMurry et al. 1998). However, as discussed in Chapter 2, bacteria can move in subsurface flow (e.g. Oliver et al. 2009; Haydon and Deletic 2006), and sometimes to significant degrees, when the movement is facilitated by subsurface irregularities like macropore networks and areas of preferential flow (Hunter et al. 1992; McMurry et al.

1998), or in regions with karst hydrology (Baffaut and Benson 2009). Bacteria fate and transport in the subsurface is nearly impossible to predict with any degree of certainty, and the inability to appropriately address these areas is a source of structural uncertainty in both models. The issue is further complicated in an area like the LREW, with soils of high infiltration rates and where lateral subsurface flow has been shown in some cases to be more than 20% of the annual rainfall (Bosch et al. 2005).

5.2.3 HSPF and SWAT Water-Quality Calibration Results

The NSE values of the calibrated HSPF and SWAT water-quality models are shown in Table 5.3, where the NSE_{ln} is given for water quality. With the exception of SWAT in Catchment J, the HSPF and SWAT water-quality model residuals were normal at the 95% confidence limit, as evidenced by the *p* values shown in the table. For the water-quality models, both models attained better NSE_{ln} values in Catchments K and I, although HSPF performed better in Catchment K while SWAT performed better in Catchment I. In Catchments J and O, HSPF performed notably better than SWAT. Although the guidelines suggested by Moriasi et al. (2007) do not include bacteria model results, the NSE_{ln} values attained by the models in this study would be considered *unsatisfactory* for other water-quality constituents. Baffaut and Sadeghi 2010 provide NSEs of five different SWAT bacteria models but none of those models used log-scaled NSEs, making direct comparison with this study impossible. When log-scaling is not used, maximizing the NSE through calibration tends to focus on hitting peak concentrations while not placing as much emphasis on non-peak values.

While negative NSE values are not desirable, they are not uncommon in bacteria modeling. Some of the NSE values reported in the Baffaut and Sadeghi 2010 study were negative, for both calibration and validation, as are the NSE_{ln} values for SWAT in Catchments J and O. Due to the definition of the NSE (Eq. 4.4), using the average of the data to predict the observations produces an NSE value of zero (Legates and McCabe 1999), and therefore a

model that produces a negative NSE is undesirable. The negative NSE_{ln} in Catchment O can be attributed to SWAT's performance in the immediate wake of a rainfall event. For example, the three largest residuals from the model predictions (Fig. 5.5d) are of (log-scaled) magnitude 1.7, 1.6, and 1.2 orders of magnitude, and all occur within 1-3 calendar days after a rainfall event. The SWAT FC predictions in the same 1-3 day interval are all higher than those on the measurement days, and re-adjusting the model predictions to the higher values results in an NSE_{ln} of 0.323. Or, eliminating these three points from the record completely results in an NSE_{ln} of 0.141. These results show that in Catchment O, where the NSE_{ln} was only slightly negative, a few small changes can improve the model NSE_{ln} from an undesirable negative value to a more favorable positive value. It is likely that these large model residuals are due to SWAT's inability to model subdaily rainfall. Hantush and Kalin (2008) address this issue in depth and recommend using a 3-day average for daily flows to help compensate for the errors that may result from using a daily time step. (Similarly, Russo et al. (2011) considered a three-day window for comparing HSPF FC predictions with data, and found that the best match of the three days considered was within an order of magnitude of the corresponding data point at least 84% of the time.) The performance of SWAT in Catchment J is similar but more extreme to that in Catchment O- there are nine model-predicted residuals of (log-scaled) magnitude greater than 1.0: five underpredictions (two of which are nearly 3.0 orders of magnitude) and four overpredictions. The five underpredictions all occur within two calendar days of a significant rainfall, and in all cases a higher FC prediction can be found within that two-day interval. The four overpredictions all occur in periods of low flow and no or little rainfall, and result from the model's inability to reproduce the low concentrations recorded in the data. SWAT's inability to account for these points results in the large, negative NSE_{ln} value in Catchment J. Interestingly, HSPF and the CC model do not encounter the same difficulty in Catchment J.

The HSPF and SWAT model predictions are compared with corresponding measurements in Figure 5.5, where the abscissa is the log-scaled FC data and the ordinate is the log-scaled model predictions. The solid line in the figures has a slope of 1 and represents a perfect model prediction whereas the dashed line indicates order-of-magnitude agreement. In all cases the models tended to over-predict at lower measurement values and under-predict at higher measurement values. Moreover, the model predictions deviate the most at extreme highs and lows, which can result from random events or from wetter-than or drier-than usual weather conditions. These highs and lows are clearly the most difficult to replicate and reflect the limitations in modeling bacteria processes using an idealized numerical model. Regardless of the difficulty found in modeling extreme events and the overall scatter of model predictions in Figure 5.5, almost all the models showed a linear correlation between the model predictions and the FC data in the log domain, as shown in Table 5.6.

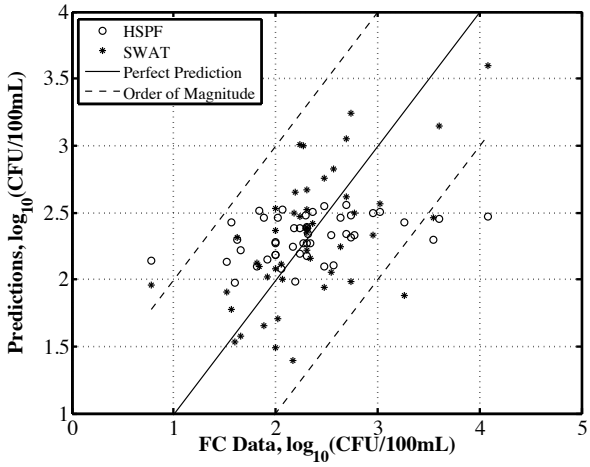
Table 5.6: Correlation of FC data and model predictions, log-scaled

Catchment	Data-HSPF	Data-SWAT
I	0.525	0.581
J	0.537	-0.131*
K	0.645	0.608
O	0.606	0.423

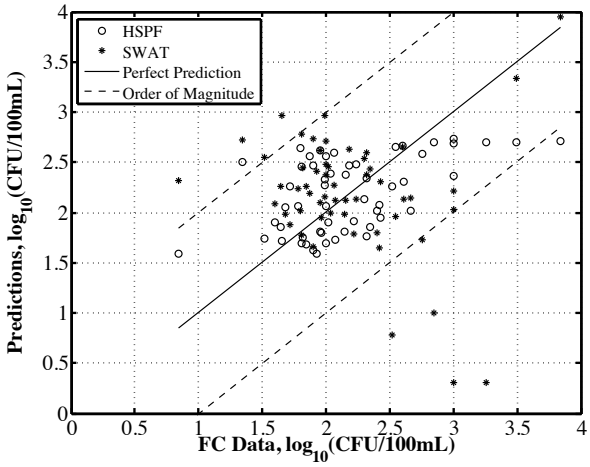
*- Correlation not statistically significant

The instance where the correlation is not statistically significant is marked with an asterisk (*)- in all other cases the correlation is statistically significant ($p < 0.05$), supporting a linear correlation between the FC data and the model predictions.

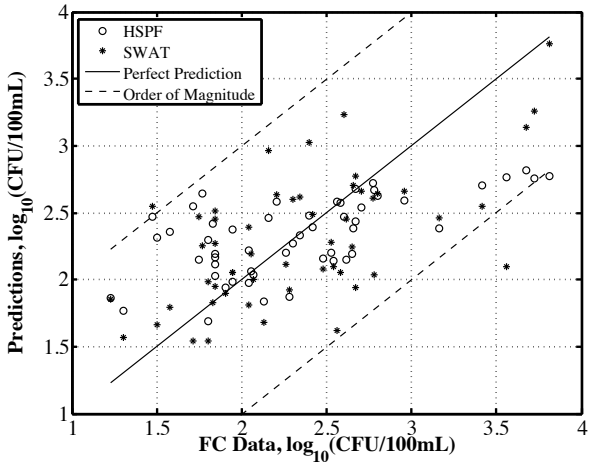
The model predictions tended to be closest to the line of perfect prediction in the data range of roughly 32-320 CFU/100 mL (1.5-2.5 in the log domain), although some predictions near the perfect-correlation line can be seen at higher magnitudes (e.g. SWAT in J, K, and O). Moreover, in Catchments I, J, and K, the HSPF model predictions tended to stay within a smaller range of magnitude (Y-axis) while the SWAT predictions covered a wider range. This is particularly obvious in Catchment I (Fig. 5.5a) where almost all the HSPF



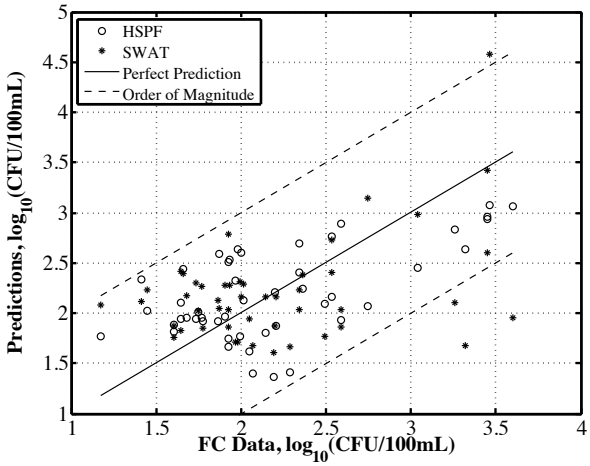
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 5.5: HSPF and SWAT model results- log-scaled predictions

predictions are restrained within a magnitude of 100-320 FC/100 mL (2.0-2.5 in the log domain), while the data and the SWAT predictions both fluctuate significantly above and below that range. Despite the tendency to fluctuate, SWAT achieved the higher NSE_{ln} in Catchment I.

The Chin et al. 2009 study examined very similar HSPF and SWAT models in Catchment K, with the same time period and FC data set. However, the Chin et al. 2009 study did not use a log-scaled objective function for the water quality portion, and the reported HSPF and

SWAT NSE (water quality) values are 0.33 and 0.73, respectively. Comparing these values with the log-scaled NSE_{ln} values in Catchment K for HSPF (0.387) and SWAT (0.334) as shown in Table 5.3, it appears the log-scaling had a less-influential effect on model predictions in HSPF than in SWAT. The different effect of log-scaling is likely due to the fundamental difference in time resolution between the models as previously addressed, where HSPF uses an hourly scale and SWAT uses a daily scale (e.g. Hantush and Kalin 2008; Russo et al. 2011). Therefore, it is likely that SWAT is more successful than HSPF at hitting peak concentrations, as seen in the results here as well as in Chin et al. 2009, a feature that is de-emphasized by log-transformation.

5.2.4 HSPF and SWAT Water-Quality Parameter Uncertainty

To quantify the role of parameter uncertainty on model error, a Latin Hypercube (LHC) analysis (McKay et al. 1979) was conducted in both HSPF and SWAT in all four catchments. LHCs use a stratified sampling technique to sample the model parameter space and are considerably more efficient and produce lower variances than simple random sampling methods such as Monte Carlo techniques (Flores et al. 2010; Christiaens and Feyen 2002; Stein 1987). The first step is to determine the marginal probability density function (PDF) of the water-quality parameters (Flores et al. 2010; Urban and Fricker 2010). This step is often problematic, however, since parameter PDFs are usually unknown. Some studies address the issue by assuming parameter PDF distributions (e.g. Kros et al. 1999; Sexton et al. 2011; Stein 1987; Paul et al. 2004), others estimate the PDFs from data (e.g. Flores et al. 2010; Christiaens and Feyen 2002), while others provide no information about how the parameter PDFs used in the LHC were determined (e.g. Jia and Culver 2008; Sándor and András 2004; Shen et al. 2008; Shirmohammadi et al. 2006; Griensven et al. 2006; Yu et al. 2001). The approach used here is to assume independent and normally distributed parameter PDFs. This is a common assumption since the multivariate normal distribution is known and easily applied to higher dimensions (Pebesma and Heuvelink

1999; Kros et al. 1999; Haan 1989). The PDF of the k-variate normal distribution is described by (Wilks 1962):

$$f(x_1,\ldots,x_k) = \frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{1/2k}} \exp\left[-1/2 Q(x_1,\ldots,x_k)\right] \tag{5.1}$$

where $(x_1,\ldots,x_k)$ is any point in R_k, and

$$Q(x_1,\ldots,x_k) = \sum_{i,j=1}^{k} \sigma^{ij}(x_i-\mu_i)(x_j-\mu_j) \tag{5.2}$$

with $\| \sigma^{ij} \|$ being the inverse of the covariance matrix $\| \sigma_{ij} \|$ and μ_i the means of the x_i.

Although the mean and standard deviation of the marginal PDF of each sensitive parameter are required yet unknown, they can be estimated conditionally from either the starting point (default) or the end point (optimal) of the calibration process. In this study the mean and standard deviation were conditionally determined from the calibrated parameter values, since the calibrated state yields smaller variances in the parameter distributions than the default state. The statistics are found by sampling the parameter space of a single parameter while holding all other parameters at the constant values associated with the given calibration state. In other words, with all parameters held at their calibrated values, one water-quality parameter was incrementally varied through its possible parameter range (Tables 4.1 and 4.2) and the model was evaluated at each incremented parameter value. The mean and standard deviation of the parameter were then estimated from the resulting values of the likelihood function. Conceptually, this process is equivalent to cutting slices through the multi-dimensional likelihood surface, where each parameter has a unique, normally-distributed slice that represents the parameter distribution.

With the assumption of independence and normality, the multivariate distribution is then described by Equation 5.1 using the mean and standard deviation of the marginal distributions of the parameters conditionally derived from the calibrated state of the models. Since the parameter distributions used here were derived from actual parameter values no negative values occurred, although implementing the parameter distributions in the LHC from theoretical distributions must be done carefully to assure that no negative values are derived from the normal PDF. Incidentally, HSPF and SWAT have built-in warnings to notify users when improper (e.g. negative) parameter values are used. While it is acknowledged that parameter correlation may have an effect on LHC analyses (e.g. Flores et al. 2010; Melching and Bauwens 2001; Pebesma and Heuvelink 1999; Iman and Conover 1982), the inclusion of correlation may ultimately be negligible (Pebesma and Heuvelink 1999) or at worst include model runs with statistically implausible parameter combinations (Flores et al. 2010). The assumption of parameter independence is justified here to facilitate the use of the multivariate normal distribution and since parameter correlation is more likely due to uncertain process equations than physically-realistic processes.

Once the marginal parameter PDFs are determined, each PDF is divided into Q regions of equal probability and a parameter value is sampled from within each region. Q is recommended to be greater than $4/3*P$ or at least between 2-5 times P (Christiaens and Feyen 2002), where P is the number of parameters. In some cases the analysis is conducted with several values of Q to examine LHC efficiency (Pebesma and Heuvelink 1999). The model is then evaluated Q times, where the parameter set used in each model realization is a random combination of parameter values taken from the divisions of the marginal PDFs. This guarantees that each region of the PDF has equal probability of being sampled.

Because $P = 6$ for HSPF and $P = 7$ for SWAT, Q was selected as 50, so that the marginal PDF of each parameter was divided into 50 regions, or 2% probability per region. The parameter value corresponding to the centroid of each region was used, although some

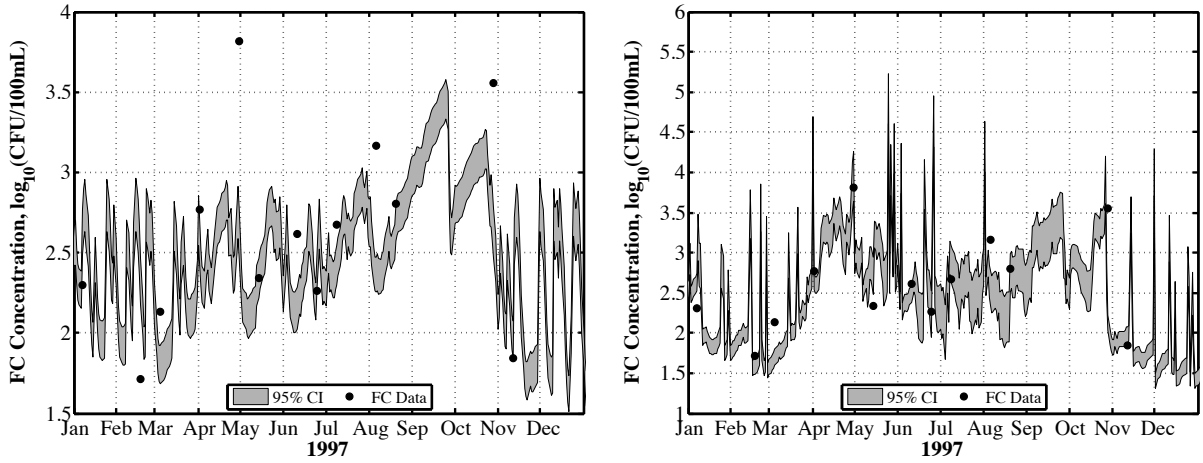
studies sample randomly or use the midpoint within each region (e.g. Pebesma and Heuvelink 1999; Melching and Bauwens 2001). For each of the 50 model evaluations, the parameter set used was a random combination of parameter values selected from the sub-divided PDFs. Insensitive parameters (Figs. 5.2 and 5.3) were held at ML values.

The results of the LHC analysis are summarized in Table 5.7, where for each catchment the distribution of FC data points within and outside the 95% confidence interval of the 50 model runs is quantified. A visual example of the process is shown in Figure 5.6, where the

Table 5.7: Number of FC data points relative to LHC 95% confidence interval (CI)

HSPF				
Catchment	Below CI (%)	Within CI (%)	Above CI (%)	Total
I	15 (32)	17 (36)	15 (32)	47
J	23 (42)	5 (9)	27 (49)	55
K	19 (36)	9 (17)	25 (47)	53
O	21 (44)	6 (12)	21 (44)	48
SWAT				
Catchment	Below CI (%)	Within CI (%)	Above CI (%)	Total
I	19 (40)	12 (26)	16 (34)	47
J	25 (45)	17 (31)	13 (24)	55
K	23 (43)	10 (19)	20 (38)	53
O	18 (38)	16 (33)	14 (29)	48

FC data and LHC analysis are plotted for the year 1997 in Catchment K. The results of the LHC analysis (Table 5.7) show that 9-36% and 19-33% of the data falls within the 95% confidence interval for HSPF and SWAT, respectively. All of the models produced at least half of the data points outside the confidence interval and seven of the eight models produced at least two-thirds of the data points outside of the modeled confidence interval, showing that even when parameter uncertainty is maximally reduced through calibration, the structural component of the model uncertainty is still dominant and prohibitive (Abbaspour et al. 2007). These results also show that the natural processes involved include more variability than can be accounted for simply by parameter variation, or, as stated by Yang et al. (2007), “prediction uncertainty in hydrological modeling can hardly be



(a) HSPF (b) SWAT

Figure 5.6: Latin Hypercube Analysis- Catchment K, 1997

described by parameter uncertainty only". The only way to achieve better model predictions, then, is to revisit and improve the process equations describing bacteria fate and transport in each model. A notable simplification of the complex process equations used by HSPF and SWAT may be needed to improve model performance.

5.2.5 CC Model

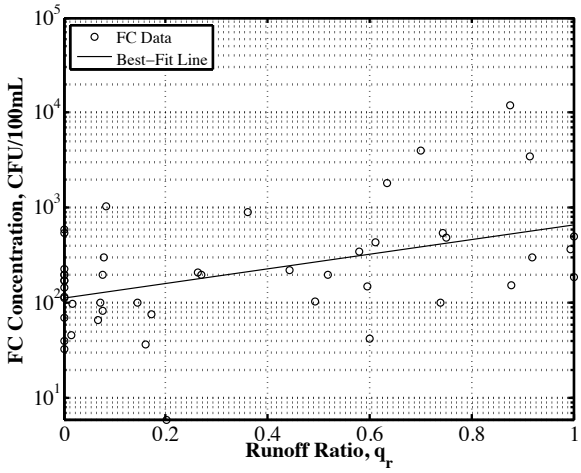
The CC model results are shown in Table 5.8, where the parameters c_b and c_r were determined from the linear best-fit relationship between $\ln c$ and q_r as shown in Figure 5.7. It is apparent from Table 5.8 that the hydrograph separation technique that attained the highest NSE_{ln}, while maintaining the assumption of normal errors, was the LM method in Catchments I, K, and O, and the FI method in J. These versions of the models will be used for the remainder of the analysis. It is interesting to note that while the FI model did not have normal residuals at the 95% confidence limit in all Catchments, it attained a significantly higher NSE_{ln} than either HSPF or SWAT in Catchments I, J, and O, and a NSE_{ln} in Catchment K between those attained by HSPF and SWAT. The CC model did not produce any negative NSE_{ln} values under any circumstances. Moreover, in most cases c_r

Table 5.8: Summary of CC model results

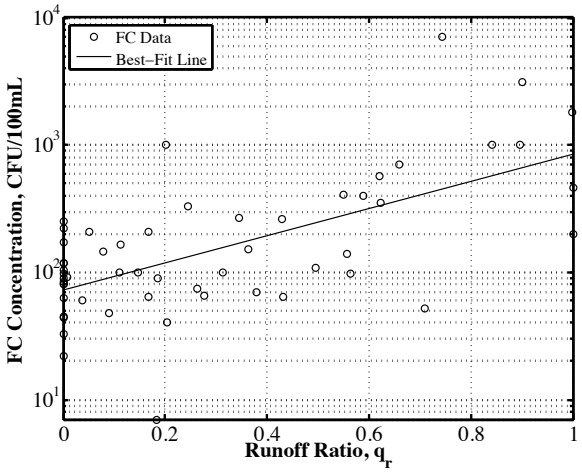
Catchment	Model	NSE_{ln}	p Value	Characteristic Concentration (CFU/100 mL)	
				c_b	c_r
I	FI	0.439	2E-3	99	2668
	SI	0.329	0.03	88	1777
	LM	0.217	0.36	114	662
J	FI	0.402	0.20	73	845
	SI	0.325	0.01	65	910
	LM	0.281	0.23	81	477
K	FI	0.370	4E-3	92	1243
	SI	0.277	0.11	95	1110
	LM	0.302	0.19	86	894
O	FI	0.297	0.92	77	1087
	SI	0.370	0.10	63	2257
	LM	0.395	0.68	66	913

was at least an order of magnitude greater than c_b, and the LM method consistently produced the lowest c_r value in all catchments while the FI method produced the highest c_r parameter in I and K and the SI method produced the highest c_r parameter in J and O.

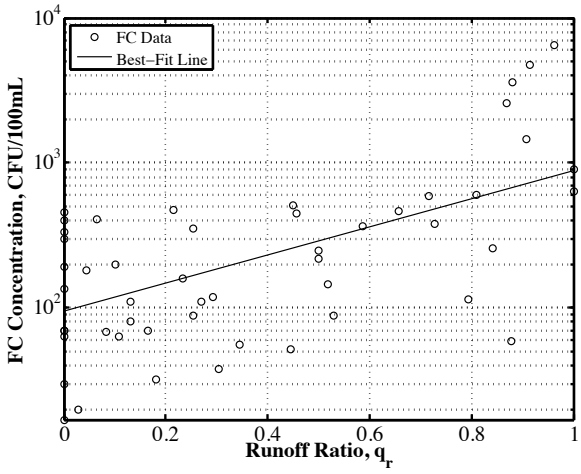
The CC model predictions are plotted with measurements in Figure 5.8, where the abscissa is the log-scaled FC data and the ordinate is the log-scaled model predictions. The solid line in the figures has a slope of 1 and represents a perfect model prediction whereas the dashed line indicates order-of-magnitude agreement. As in the HSPF and SWAT models, the CC model tended to over-predict at low data values and under-predict at high data values, and the model predictions tended to be closest to the line of perfect prediction in the data range of roughly 32-320 CFU/100 mL (1.5-2.5 in the log domain). Similar to the HSPF model, the predictions tended to stay within a tightly defined range, albeit a slightly larger range of roughly 100-1,000 CFU/100 mL (2.0-3.0 in the log domain). Determining the best-fit parameters (and NSE_{ln} values) of the CC model was done only on the days corresponding to the FC data and did not incorporate the random term shown in Equation 3.30.



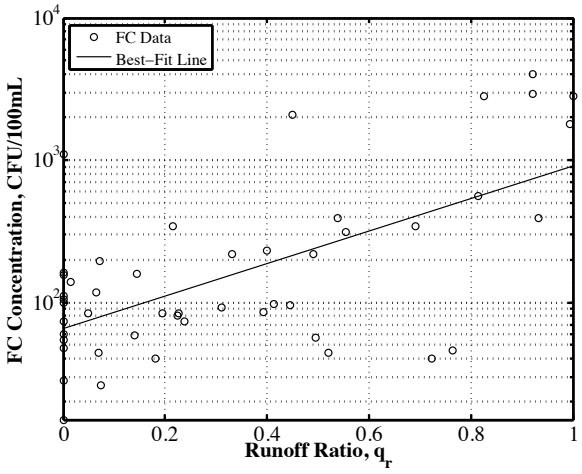
(a) Catchment I- Local Minimum



(b) Catchment J- Fixed Interval



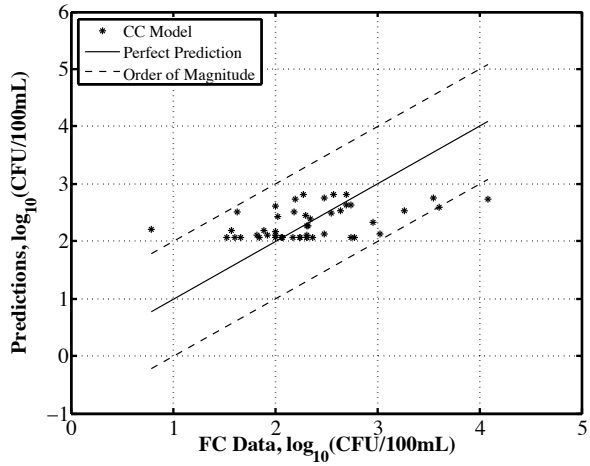
(c) Catchment K- Local Minimum



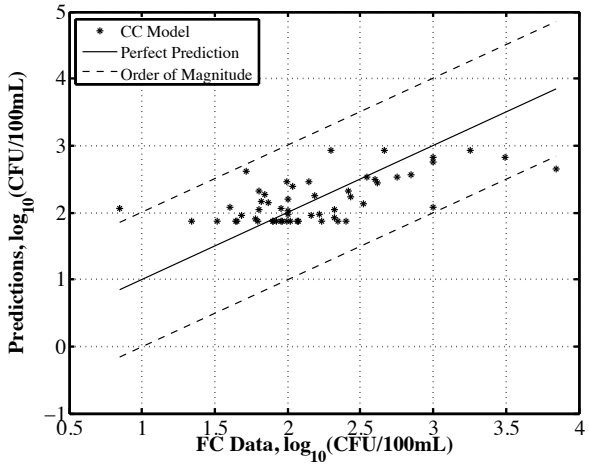
(d) Catchment O- Local Minimum

Figure 5.7: CC model results- best-fit line

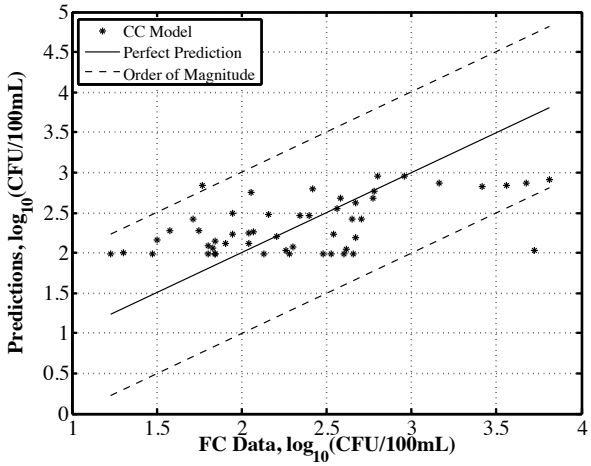
A further examination of the CC model in contrast to HSPF and SWAT reveals several distinct advantages that merit its inclusion in bacteria modeling studies. First and foremost, the data requirements (streamflow and bacteria concentrations) are much less than for HSPF and SWAT, which makes model setup and use significantly easier and can be advantageous in areas with little available data. Secondly, while HSPF and SWAT have several hydrology and water quality parameters that must be calibrated, and complex process equations for modeling environmental phenomena, the 2-parameter CC model is structurally much



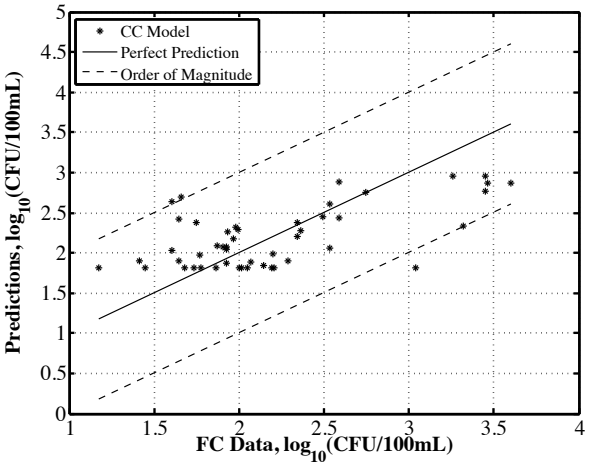
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 5.8: CC model results- log-scaled predictions

simpler and easier to calibrate and has been shown to achieve comparable results (Tables 5.3 and 5.8). Another advantage is that the CC model uses the actual streamflow measurements to model the bacteria concentrations, while HSPF and SWAT both use a modeled hydrology component to model the bacteria concentrations instead of the streamflow measurements. While the CC model does not provide any land-use to in-stream concentration relationships, the c_b and c_r parameters directly quantify the origins of the in-stream bacteria loads, can be adjusted based on land-use changes, and can be used to

quantify the load reductions necessary to achieve compliance with given water-quality standards. Perhaps most importantly, the CC model predicts bacteria concentrations while directly accounting for the runoff ratio of the streamflow on the day of data collection. This important relationship can be seen in Figure 5.7, where $q_r = 0$ corresponds to baseflow conditions, $q_r = 1$ corresponds to runoff equaling streamflow, and the best-fit line at $q_b = 0$ and $q_r = 1$ equals the calibrated parameters c_b and c_r, respectively. The distributions of FC data shown in Figure 5.7 show that while some points were indeed collected during baseflow conditions ($q_r = 0$), the majority of the data was collected under the effect of baseflow plus runoff conditions. Moreover, the peak values in each catchment occurred at $q_r > 0.7$, although not at $q_r = 1$. The FC versus q_r relationships shown in Figure 5.7 capture the randomness and the difficulty encountered in bacteria modeling. While expected relationships generally hold (such as higher FC concentrations with increasing q_r), there are contradictory exceptions that also occur. Extreme exceptions can be seen in Catchments I, J, and K, (Fig. 5.7a-5.7c), where two data points differ by almost two orders of magnitude around $q_r = 0.9$, 0.75, and 0.85, respectively. The CC model, however, accounts for the fluctuations at baseflow conditions as seen in Figure 5.7, such that when $q_r = 0$ in Equation 3.30, the in-stream concentration varies depending on the baseflow and random components. The inclusion of the random term is an adequate way to represent the randomness of the bacteria resuspension and sediment-associated processes that occur in-stream. While HSPF and SWAT have much more complicated algorithms for detailing in-stream processes, a large amount of uncertainty is incorporated due to not being able to perfectly describe this phenomena.

Finally, it is acknowledged that a linear relationship exists between model complexity and model uncertainty, where more complicated models require more parameters but the increase in parameters decreases their identifiability and makes calibration and uncertainty analysis more difficult (Shirmohammadi et al. 2006; Yang et al. 2007). Conversely, simpler model structures with less parameters will result in a more robust parameter calibration but

may contribute uncertainty due to reduced process equations that replicate simpler physical processes (Franks and Beven 1997). This study revealed that the 2-parameter CC model attained comparable or better NSE_{ln} values than the much more complicated HSPF and SWAT models (Tables 5.4, 5.5, 5.8), indicating that a simpler model does not necessarily underperform more complex models.

5.3 CONCLUSIONS

This chapter presented the HSPF and SWAT models and preliminary results that are the foundation of all the research efforts included in this dissertation. The hydrology components of the models were strong but it was found that many of the parameters in both models were significantly correlated with several other parameters. The parameter correlation is a result of structural uncertainty and induces an element of equifinality into the calibrated parameter sets. The water-quality components of the models were not as strong due to minimal data and the difficulty in modeling the highly random and uncertain nature of bacteria dynamics. The water-quality components are also inherently complicated by the parameter correlation in the calibrated hydrology components, since the water-quality is dependent on the hydrology. The most sensitive water-quality parameters were those dealing with in-stream bacteria processes, while parameters relating to terrestrial processes varied from sensitive to completely insensitive. The process equations involving the insensitive parameters contribute uncertainty to the model and should be refined or removed. The LHC analysis confirmed that significant parameter uncertainty occurs in HSPF and SWAT, but that the structural uncertainty is still the dominant source of model uncertainty. The only way to reduce structural uncertainty then is to change the way the models are built. All three models over-predicted low bacteria levels and under-predicted peak bacteria levels. The CC model is much simpler than both HSPF and SWAT, yet attained comparable results, making it an attractive option for those not experienced with HSPF and SWAT, or in areas where minimal data is available. The CC model is

fundamentally different than HSPF and SWAT and therefore can provide another perspective from which to analyze the problem of bacteria modeling and possibly contribute understanding into bacteria fate and transport.

Chapter 6

In-Stream Bacteria Modeling as a Function of the Hydrologic State of a Watershed

6.1 BACKGROUND

While there is consensus in the literature that increased in-stream bacteria concentrations occur after rainfall, an exact synthesis of all related phenomena to best predict in-stream concentrations has been elusive. As described in Chapter 2, many of the processes involved in bacteria fate and transport are known, such as overland flow, infiltration, subsurface flow, absorption to streambed sediment, point-source contributions, and soil matrix habitation and persistence (e.g. Wilkinson et al. 1995; Harmel et al. 2010; Jamieson et al. 2005; Auer and Niehaus 1993). However, the exact roles played by these complex phenomena are not constant and can vary significantly from site to site. For example, the transport of bacteria in overland flow can be significant and has been shown to depend on the moisture content and/or wet or dry state of the soil (e.g. Tyrrel and Quinton 2003; Muirhead et al. 2006a,b), and yet transport of bacteria in overland flow may be insignificant in large/upper catchments (Crowther et al. 2002) or at increasing distances from riparian and streambank locations (Byappanahalli et al. 2003). Other studies, however, have revealed that the

increased bacteria quantities in the wake of a rainfall event are likely due to in-stream processes instead of terrestrial processes. For example, bacteria concentrations have been observed to peak on the rising limb of a hydrograph under both natural (Davies-Colley et al. 2008) and man-made (McDonald et al. 1982) stormflow conditions. These results contradict what might be expected, that bacteria concentrations would climax after the hydrograph peak as terrestrial fecal matter is transported to the stream, and instead show that the increase in channel flow stirs up and releases bacteria stored in channel sediment. Solo-Gabriele et al. (2000) showed that increased bacteria concentrations in a coastal subtropical environment were due to the heightened river stage interacting with channel bank sediment and soil instead of runoff-induced terrestrial transport phenomena. Similarly, in studying a midwestern stream that discharges into Lake Michigan, Byappanahalli et al. (2003) observed that EC was common in the stream basin, especially in "submerged, margin, and wetted bank sediments". Their results led to the conclusion that EC was ubiquitous in riparian areas and likely independent of fecal contamination. Davies-Colley et al. (2008) suggest that the main role of fecal matter washed into streams from terrestrial sources is to replenish the in-channel stores of bacteria that are flushed on the rising limb of the hydrograph.

Bacterial water-quality analyses are often conducted on a seasonal basis based on wet and dry times of the year. The state of Georgia, for example, implements a seasonal water-quality standard for FC where the allowable monthly geometric mean from May-October is five times more stringent than that for November-April (GA DNREP 2005). A seasonal breakdown is convenient because it generally distinguishes wet and dry rainfall periods, higher and lower temperature (which can be significant in bacteria die-off processes), and times when recreational use of water bodies such as bathing, boating, and fishing are more likely to occur (e.g. Liu et al. 2006). A seasonal analysis is also convenient in agricultural areas where the transport of dairy effluent and animal manure in overland flow has been shown to be highest in the hydrologically wet time of the year

(Shirmohammadi et al. 1997), and where there is a distinct difference in land use during the growing season than during other times of the year (e.g. Feyereisen et al. 2008). Modeling studies have acknowledged that: better results can be achieved when using data from wet conditions instead of dry conditions (Jeong et al. 2010); in some cases, the specific adjustment of calibrated parameter values is required in cold and wet spring periods and very wet summer periods (Al-Abed and Whiteley 2002); parameters may be adjusted based on growing season (White et al. 2009); and that using "dryer than average" streamflow records may be useful in specifically examining low flow conditions, although they are likely to substantially underestimate hydrologic processes over the long-term (Van Liew et al. 2003).

Although a seasonal breakdown based on the calendar year is simple and convenient, it may not accurately describe the hydrologic state of the watershed, since occasional dry days will inevitably occur in the "wet" season, and vice versa. An additional complication occurs from the FC data collection, where in response to time and financial constraints, long-term studies will usually collect data at weekly, monthly, or multi-month intervals, while short-term (seasonal) studies often utilize an intensified sampling routine over a smaller time interval that coincides with recreational use. A more accurate method of defining hydrologic state is to look at the actual river flow on any given day, regardless of the calendar month in which it occurs. Such an approach is particularly useful in long-term water-quality studies, where the overall state of the river over the course of the study can not be assumed to conform to what is expected of a particular season, and therefore the specification of the actual state of the river on the day of the weekly, monthly, or other time interval when the data is collected is a major improvement.

The research presented in this chapter implements such an approach, where "wet" and "dry" conditions were defined by a probabilistic assessment of the daily river flow as determined by the flow duration curve. The FC data sets were divided into corresponding

wet and dry groups and the bacteria modeling was performed for the wet and dry conditions. The analysis provides unique insight into model performance, parameter sensitivity, and the FC data on a wet or dry basis. This insight can be used for guiding the design of future bacteria studies or for refining the current model process equations.

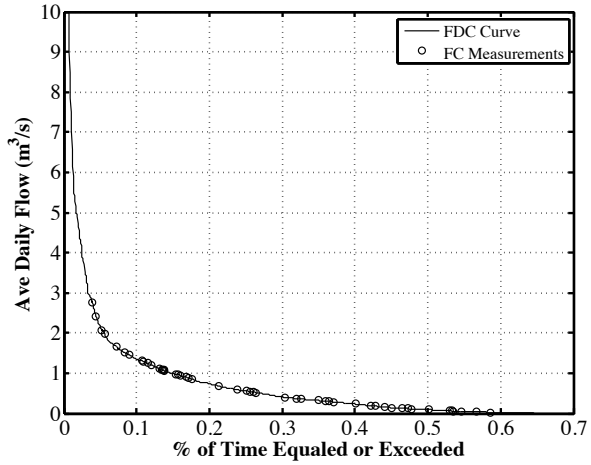
6.2 THEORY

A Flow Duration Curve (FDC) is a probability distribution that expresses the flow values in terms of an exceedance probability relative to the entire flow record. The FDC was determined using the Weibull formula (Weibull 1939; Chin 2006). First, the daily streamflow record of length N for each catchment was ranked such that a rank of 1 corresponded to the observation of lowest magnitude, and a rank of N corresponded to the observation of largest magnitude. The exceedance probability of the m-ranked observation, x_m, was then estimated as a cumulative distribution function by:

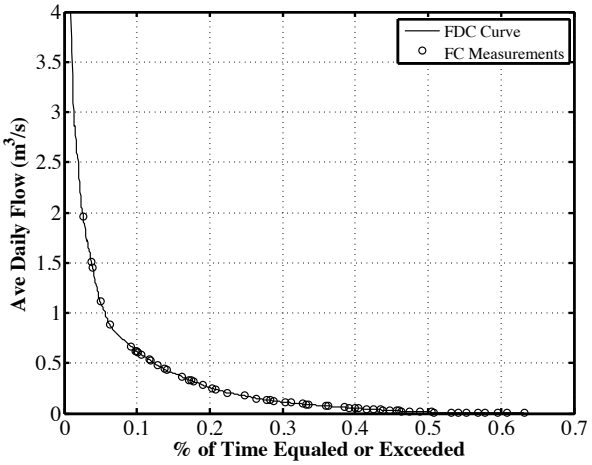
$$P_x(X < x_m) = 1 - \frac{m}{N+1}, \qquad m = 1, ..., N \tag{6.1}$$

FDCs are typically broken into five regions based on exceedance probabilities, where 0-10% is considered *High Flows*, 10-40% is considered *Moist Conditions*, 40-60% is considered *Mid-Range Flows*, 60-90% is considered *Dry Conditions*, and 90-100% is considered *Low Flows* (USEPA 2007). Flow duration curves (FDC) were developed from the daily streamflow records of each catchment. The FDCs for all four catchments are shown in Figure 6.1. The FC data sets were included in the figure and the data points are plotted based upon the exceedance probability of the flow on the day when the FC data measurement was made.

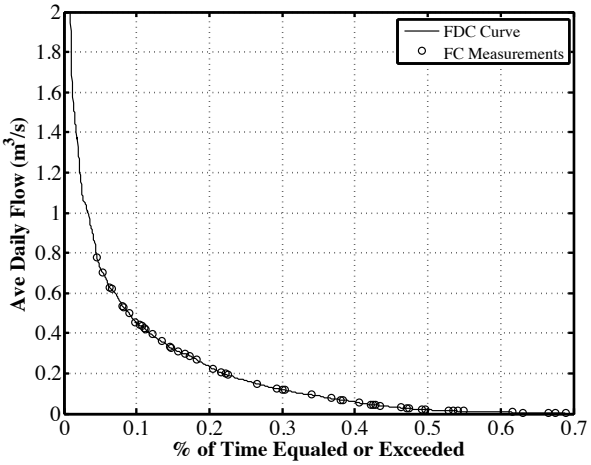
In Table 6.1 a breakdown is shown of the FC data points in each catchment based on exceedance probabilities of 20, 25, 30, and 35%. In each case, the data points taken on days



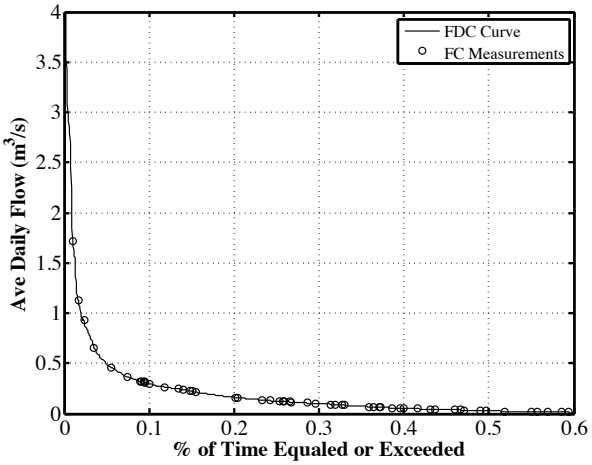
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 6.1: Flow Duration Curves

of higher flows (lower exceedance probability) were considered to be “wet” and the data points taken on days of lower flow (higher excceedance probability) were considered “dry”. Based on the distribution of data as shown in Table 6.1, the 30% exceedance probability was selected as the demarcator between wet and dry because the data is more evenly divided between wet and dry at 30% than at the other exceedance probabilities. Therefore, for the remainder of the analysis, wet will refer to conditions with flow exceedance values less than or equal to 30% and dry will refer to conditions with flow exceedances greater

Table 6.1: Distribution of FC data points into wet and dry states for several exceedance probabilities

Catchment	I		J		K		O	
Total FC Points	47		55		53		48	
Flow Condition	wet	dry	wet	dry	wet	dry	wet	dry
20% Exc. Prob.	20	27	19	36	21	32	15	33
25% Exc. Prob.	22	25	24	31	26	27	18	30
30% Exc. Prob.	27	20	28	27	29	24	24	24
35% Exc. Prob.	29	18	32	23	31	22	28	20

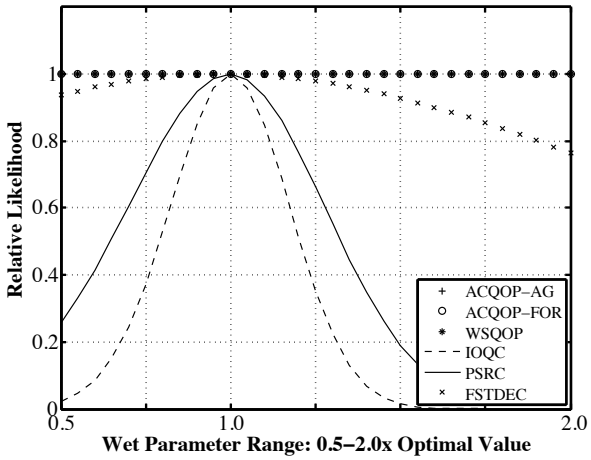
than 30%, even though some flows greater than 30% are still considered *moist conditions* or *mid-range* according to the traditional definition.

Upon establishing the definition of the hydrologic states, separate calibrations were performed for the wet and dry water-quality models in HSPF and SWAT. The hydrology components of the models were kept the same as the calibrated versions in Chapter 5, but the water-quality components were calibrated separately to the wet and dry FC data sets. The calibration procedure and default starting state of the models were the same as described earlier.

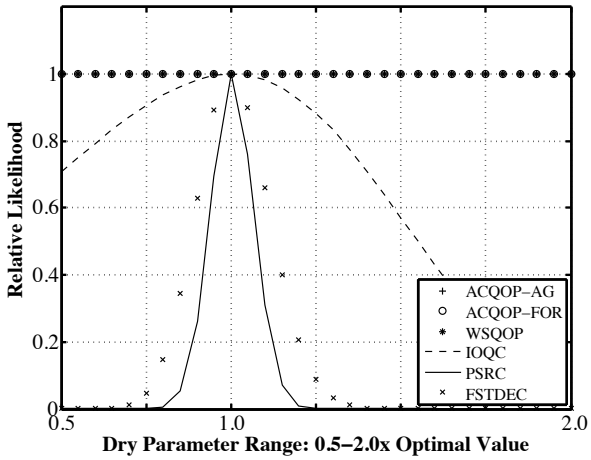
6.3 RESULTS AND DISCUSSION

6.3.1 HSPF and SWAT Wet/Dry Parameter Sensitivity

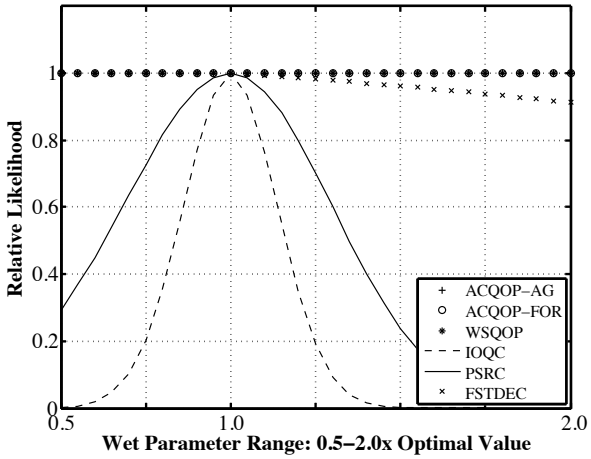
The HSPF and SWAT water-quality parameters considered were the same as discussed earlier, and the calibrated maximum likelihood (ML) parameters of the wet and dry models are shown in Tables 6.2-6.5. Sensitivity plots of the water-quality parameters in HSPF and SWAT in the wet and dry states are shown in Figures 6.2-6.5. In these figures the relative likelihood distribution of each parameter over a range of 0.5-2.0 times the ML parameter value is shown, where the other parameters are held constant at their ML values. The most notable observation in HSPF (Figs. 6.2 and 6.3) is that, similar to what was found in Chapter 5, the in-stream parameters PSRC and IOQC are consistently the most sensitive



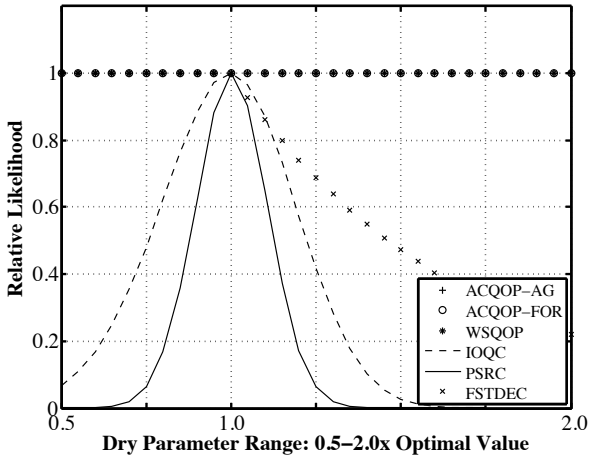
(a) Catchment I-Wet



(b) Catchment I-Dry

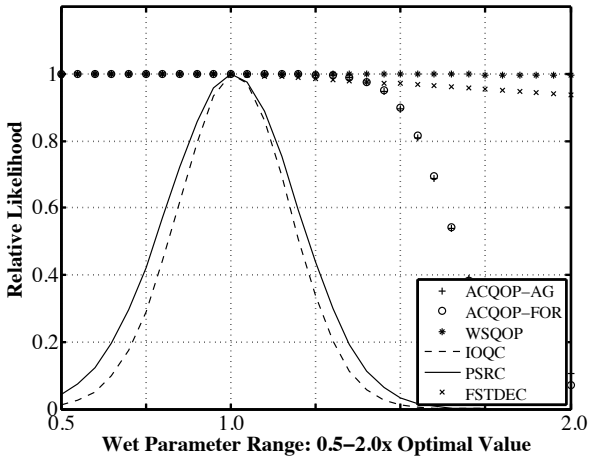


(c) Catchment J-Wet

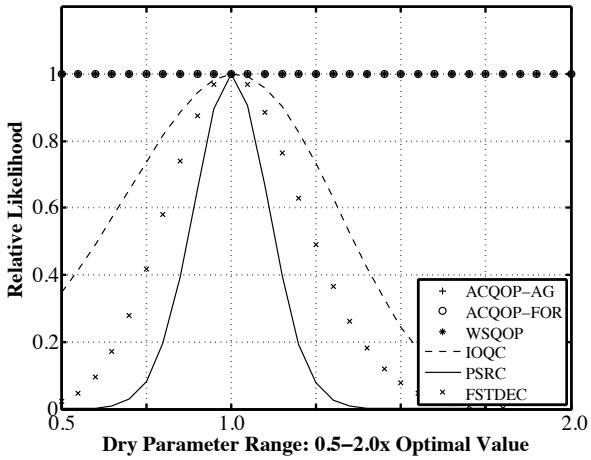


(d) Catchment J-Dry

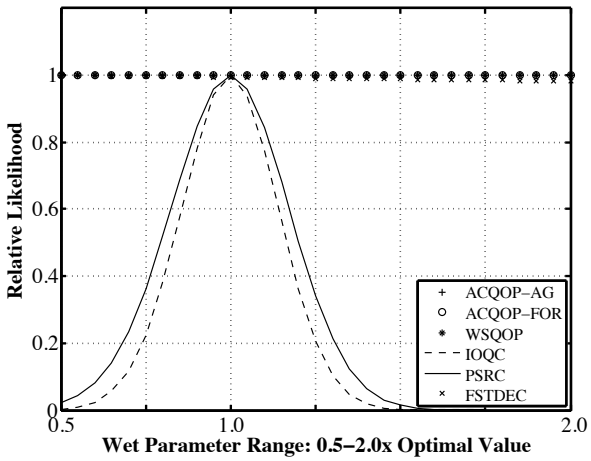
Figure 6.2: HSPF water-quality parameter sensitivity, Catchments I and J



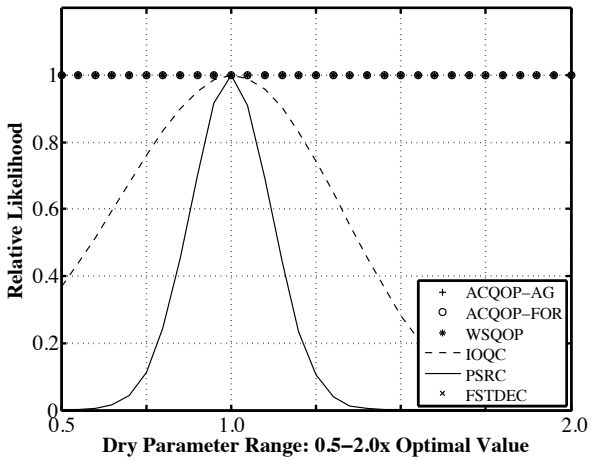
(a) Catchment K-Wet



(b) Catchment K-Dry

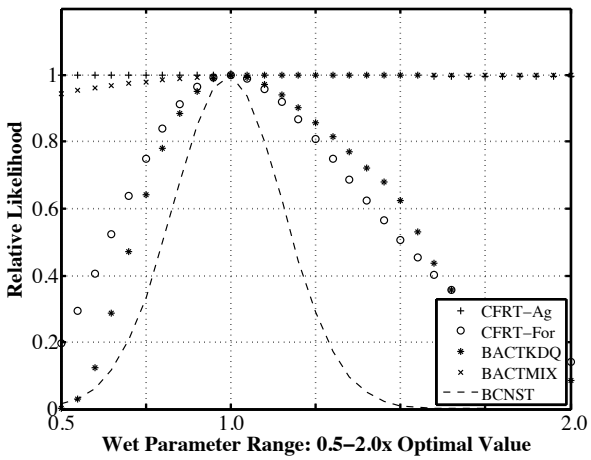


(c) Catchment O-Wet

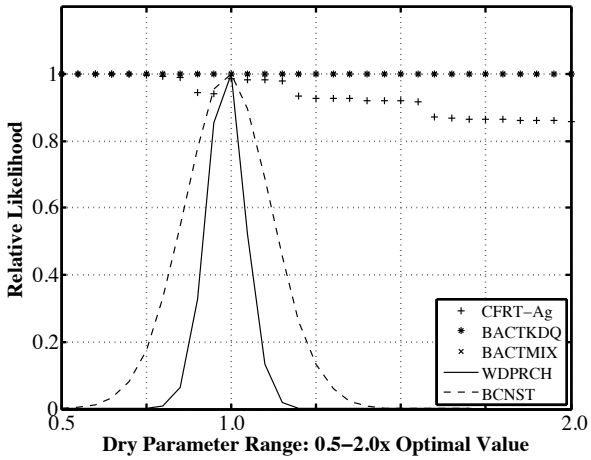


(d) Catchment O-Dry

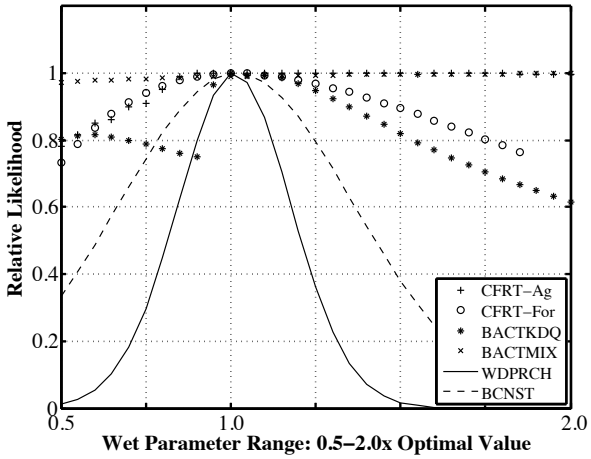
Figure 6.3: HSPF water-quality parameter sensitivity, Catchments K and O



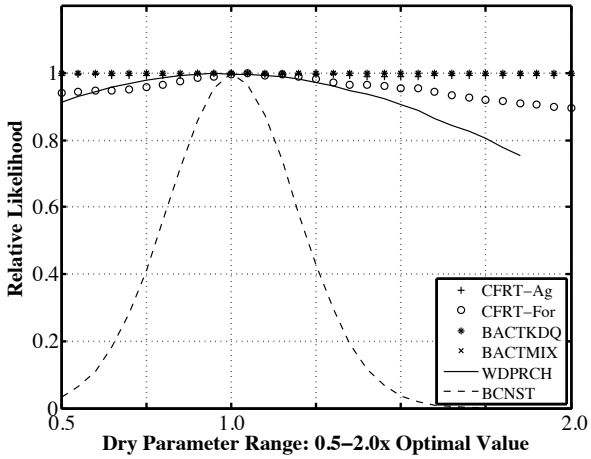
(a) Catchment I-Wet



(b) Catchment I-Dry

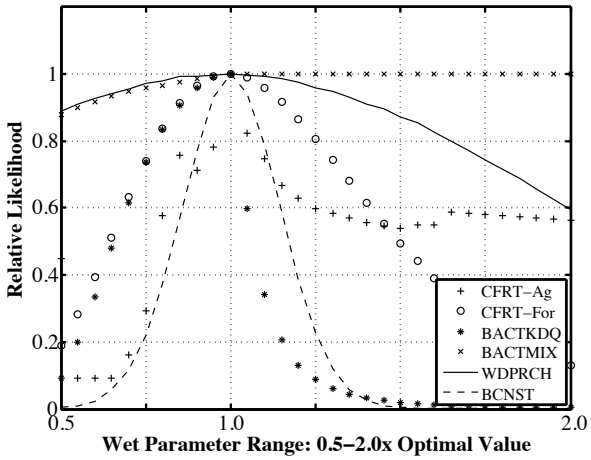


(c) Catchment J-Wet

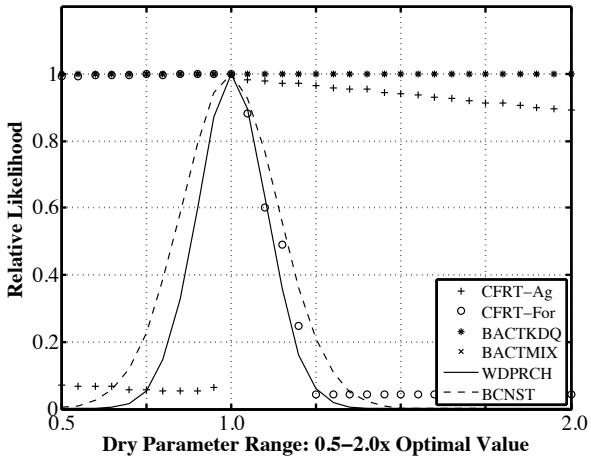


(d) Catchment J-Dry

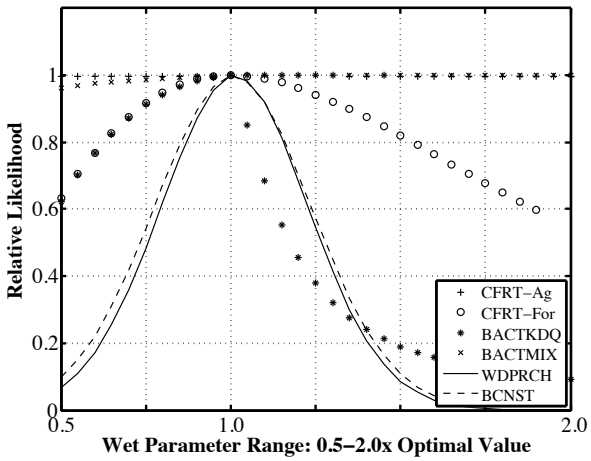
Figure 6.4: SWAT water-quality parameter sensitivity, Catchments I and J



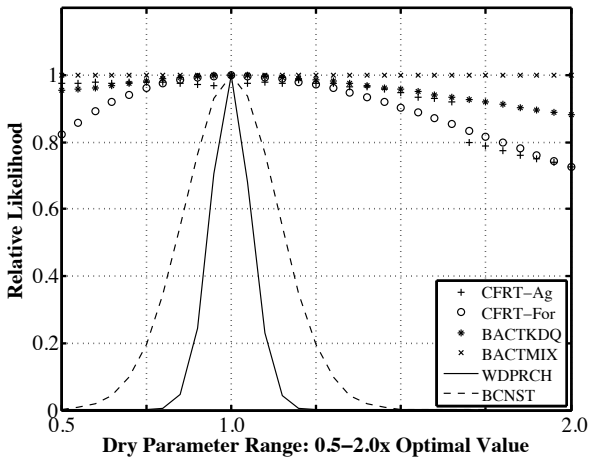
(a) Catchment K-Wet



(b) Catchment K-Dry



(c) Catchment O-Wet



(d) Catchment O-Dry

Figure 6.5: SWAT water-quality parameter sensitivity, Catchments K and O

Table 6.2: HSPF water-quality calibration results- wet state

Parameter	Units	Default*	Catchment			
			I	J	K	O
ACQOP_AG	FC/ha-d	247	9.41	7.96	1.63E4	4.50
ACQOP_FOR	FC/ha-d	247	12.0	10.3	1.48E4	13.5
WSQOP	cm/hr	4.17	0.510	0.210	0.090	0.510
IOQC	FC/m^3	3530	6.35E6	5.38E6	7.03E6	1.35E7
PSRC	FC/d	1.00E8	7.37E8	1.50E8	4.93E8	3.22E8
FSTDEC	1/d	1.00	0.340	0.100	0.100	0.100

* Note: slight variations of some default values occurred across catchments

Table 6.3: HSPF water-quality calibration results- dry state

Parameter	Units	Default*	Catchment			
			I	J	K	O
ACQOP_AG	FC/ha-d	247	24.7	2.76E5	5.34	24.7
ACQOP_FOR	FC/ha-d	247	24.7	2.47E5	4.27	24.7
WSQOP	cm/hr	4.17	0.508	0.0437	0.508	0.508
IOQC	FC/m^3	3530	2.80E6	9.11E6	1.44E7	1.57E7
PSRC	FC/d	1.00E8	5.00E9	2.30E8	1.12E9	1.65E8
FSTDEC	1/d	1.00	2.66	0.100	4.00	4.00

* Note: slight variations of some default values occurred across catchments

parameters while the parameters relating to terrestrial processes (ACQOP and WSQOP) are consistently insensitive. These observations are generally consistent for HSPF in all catchments and in both wet and dry states. There are several differences that can be seen between the wet and dry states, however. First, the interflow concentration (IOQC) is notably less sensitive in the dry state than in the wet state in all catchments. This observation reflects the fact that in the dry state the contribution from baseflow to the total river flow is likely dominant over the contribution from interflow. In fact, in very dry conditions when total stream flow consists solely of baseflow, the interflow and therefore interflow contribution to in-stream bacteria will be zero, which further explains the lessened sensitivity of IOQC in the dry state. Secondly, the order of sensitivity between IOQC and PSRC reverses between wet and dry states, where IOQC is more sensitive than PSRC in the wet state in all four catchments and PSRC is more sensitive than IOQC in the dry state in all four catchments. This too reflects the effect that sources of flow have on the in-stream

Table 6.4: SWAT water-quality calibration results- wet state

Parameter	Units	Default*	Catchment			
			I	J	K	O
CFRT_KG_Ag	kg/ha-d	30.0	0.010	2920	26.5	0.210
CFRT_KG_For	kg/ha-d	30.0	1004	4850	979	6740
BACTKDQ	m^3/Mg	175	1.22	1.45	1.50	1.75
BACTMIX	10 m^3/Mg	10.0	20.0	16.5	15.6	8.31
WDPRCH	1/d	0	0	2.34	0.790	4.25
WDPQ	1/d	0	0	0	0	0
BCNST	CFU/100mL	1.00E8	9.72E8	2.22E8	3.26E8	4.82E8

* Note: slight variations of some default values occurred across catchments

Table 6.5: SWAT water-quality calibration results- dry state

Parameter	Units	Default*	Catchment			
			I	J	K	O
CFRT_KG_Ag	kg/ha-d	30.0	9420	95.2	656	4.75
CFRT_KG_For	kg/ha-d	30.0	0	1.12	0.290	8000
BACTKDQ	m^3/Mg	175	1.07	1.38	1.45	0.846
BACTMIX	10 m^3/Mg	10.0	14.1	20.0	18.5	20.0
WDPRCH	1/d	0	2.05	0.140	1.91	3.04
WDPQ	1/d	0	0	0	0	0
BCNST	CFU/100mL	1.00E8	1.12E9	9.03E6	1.03E8	7.76E7

* Note: slight variations of some default values occurred across catchments

bacteria concentration, where the background source (PSRC) is more sensitive in the dry state when it is the only constant and direct bacteria source, while the interflow concentration is more sensitive in the wet state when interflow is likely to contribute more significantly to total streamflow. The calibrated water-quality parameters shown in Tables 6.2 and 6.3 show that PSRC is approximately the same order of magnitude in wet and dry states, but in Catchments I, J, and K, the value of PSRC in the dry state is greater than in the wet state. This further reflects the additional bacteria load contribution made by PSRC in the dry state. Similarly, the order of magnitude of IOQC is approximately the same between wet and dry states of the catchments, and the value of IOQC in Catchments J, K, and O is greater in the dry state than in the wet state. Finally, while the in-stream bacteria decay parameter (FSTDEC) shows little to no sensitivity in the wet state of any catchment, it is

sensitive in the dry states of Catchments I, J, and K. Bacteria die-off, as represented by FSTDEC, is likely a more significant process in the dry state than in the wet state when the stream has more flow and components like the interflow are contributing a significant quantity of bacteria to the stream.

In SWAT (Figs. 6.4 and 6.5), the most consistently sensitive parameter in wet and dry states in all four catchments is the in-stream background concentration (BCNST), which is more sensitive in the dry state than the wet state of each catchment with exception to Catchment K, where the sensitivity is approximately equal between states. The other sensitive in-stream parameter in SWAT from Chapter 5 is the bacteria die-off rate (WDPRCH) which does not show a consistent pattern of sensitivity in the wet and dry states, although it is generally more sensitive in the dry state than the wet state. Since WDPQ in Catchment I calibrated to zero it was not included on the relative scale shown in Figure 6.4a, but rather is shown for the actual parameter values in Figure 6.6. Also, while the land-application

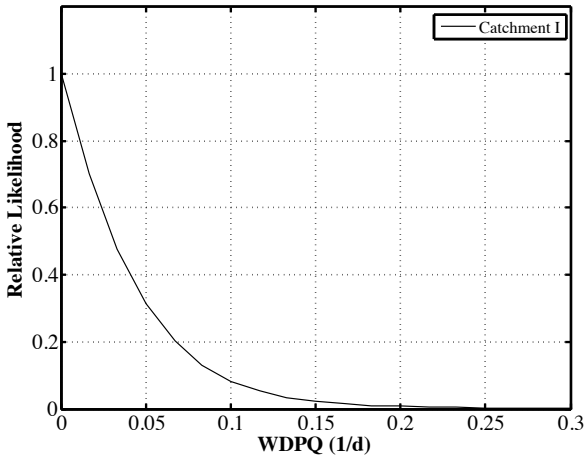


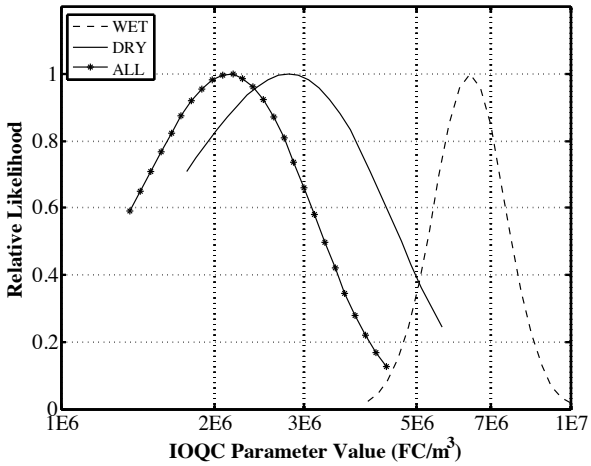
Figure 6.6: SWAT parameter WDPRCH in wet model of Catchment I

parameters (CFRT) reveal a variety of sensitivities, as in Chapter 5, they are relatively insensitive in the dry states of the catchments, which is a reflection of the reduced bacteria contributions made by terrestrial sources in dry conditions. Tables 6.4 and 6.5 reveal that there is little consistency in the CFRT-Ag and CFRT-For parameters between catchments or

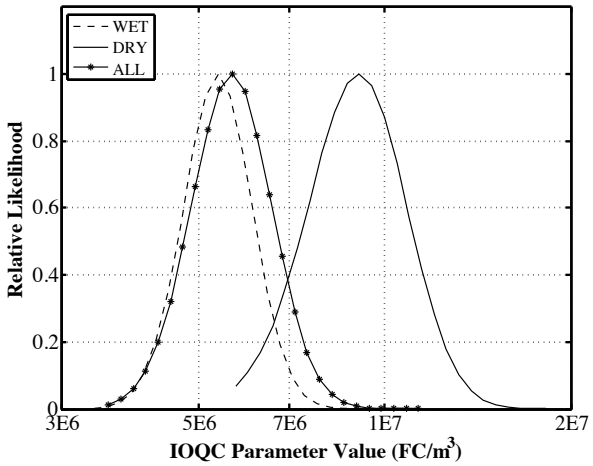
in wet or dry states. This lack of a predictable pattern is likely a result of the calibration procedure attempting to find optimal values of these parameters of varying sensitivity in order to maximize the NSE_{ln} of the model output. This result is noteworthy in SWAT since the model does not permit subsurface bacteria and therefore must account for all in-stream concentrations via direct sources like BCNST or via terrestrial rainfall-runoff processes affected by parameters like CFRT. Since CFRT-For in the dry model of Catchment I calibrated to zero it was not included on the relative scale shown in Figure 6.4b, but was relatively insensitive at non-zero values similar to CFRT-For in Catchment J (Fig. 6.4d). The SWAT parameters that affect bacteria concentrations in soil solution in runoff (BACTKDQ) and the leaching of bacteria to subsurface layers (BACTMIX) are both insensitive in the dry state, and BACTMIX is also insensitive in the wet state. The die-off coefficient for soil-solution bacteria (WDPQ) calibrated to zero for the wet and dry models in all catchments and was otherwise insensitive at nonzero values, and was therefore left out of the sensitivity plots shown in Figures 6.4 and 6.5.

The effects of the wet/dry analysis were further explored for the sensitive parameters IOQC and PSRC in HSPF and BCNST in SWAT. The sensitivity curves of the selected parameters from Figures 5.2, 5.3, and 6.2-6.5 were combined with the actual calibrated parameter values in Tables 5.4, 5.5, and 6.2-6.5 to compare the relative values and sensitivities of IOQC, PSRC, and BCNST, as shown in Figures 6.7, 6.8, and 6.9, respectively. The sensitive parameters FSTDEC (HSPF) and WDPRCH (SWAT) were not considered since they both play identical roles as first-order decay rates in the models and therefore do not highlight differences between HSPF and SWAT, although the parameters do show a variety of calibrated values between catchments and hydrologic states in Tables 5.4, 5.5, and 6.2-6.5.

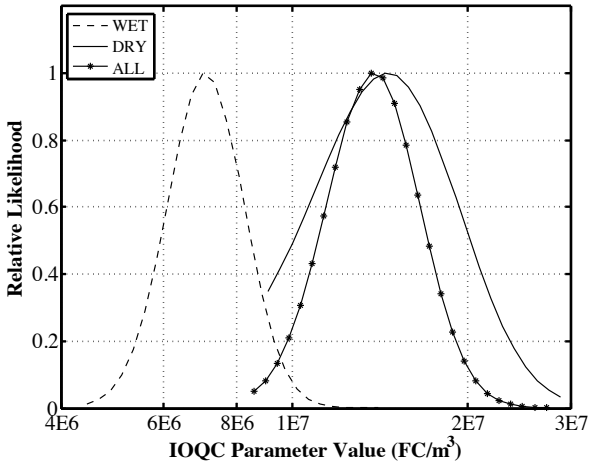
In Figure 6.7 it can be seen that IOQC in the wet state is generally more sensitive than the dry state in each catchment. The value of the calibrated IOQC parameter is least in the wet state in all cases except for Catchment I. Similarly, the value of IOQC in the “all” state is



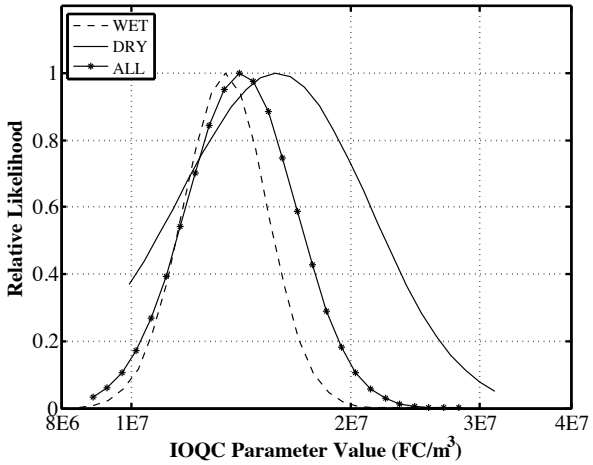
(a) Catchment I



(b) Catchment J

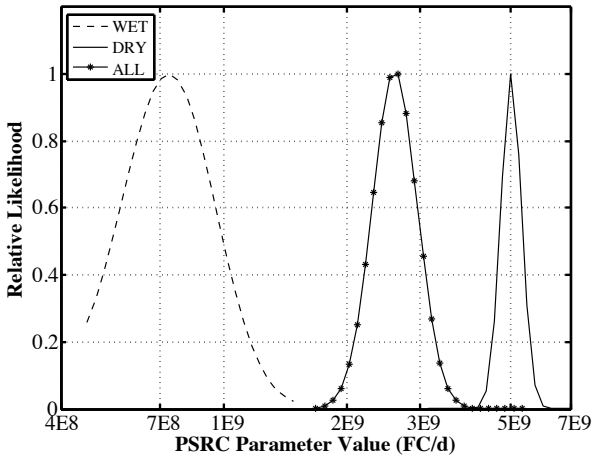


(c) Catchment K

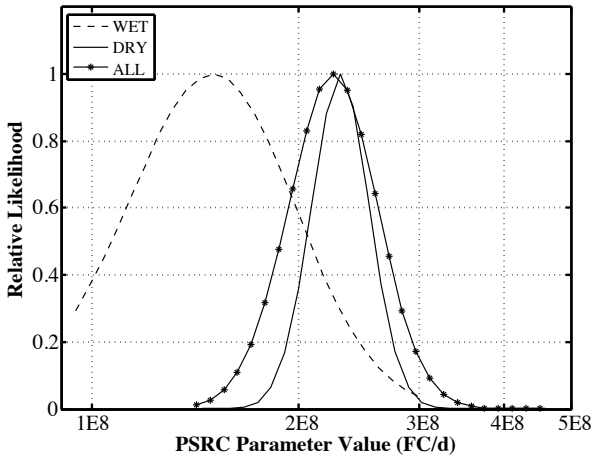


(d) Catchment O

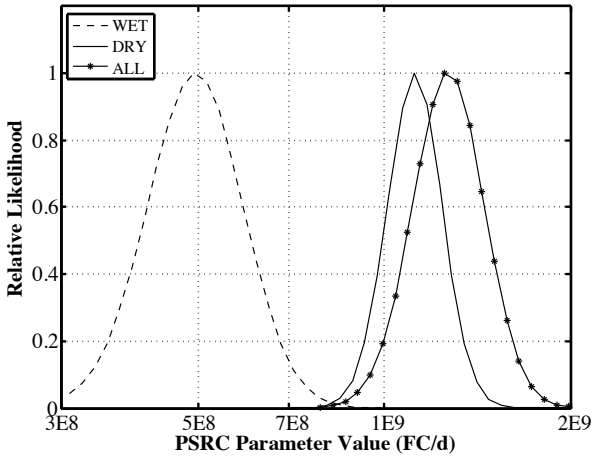
Figure 6.7: IOQC (HSPF) sensitivity in three hydrologic states



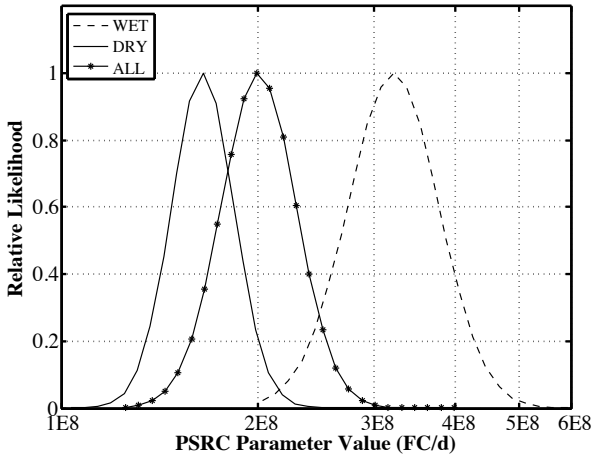
(a) Catchment I



(b) Catchment J

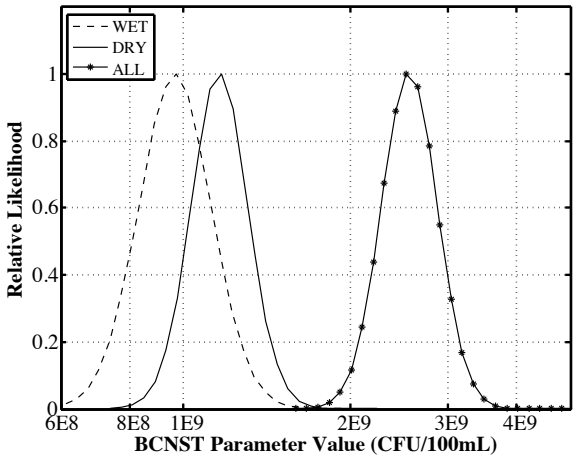


(c) Catchment K

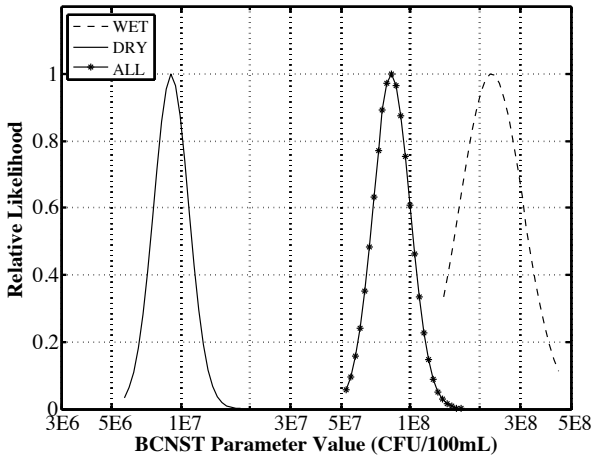


(d) Catchment O

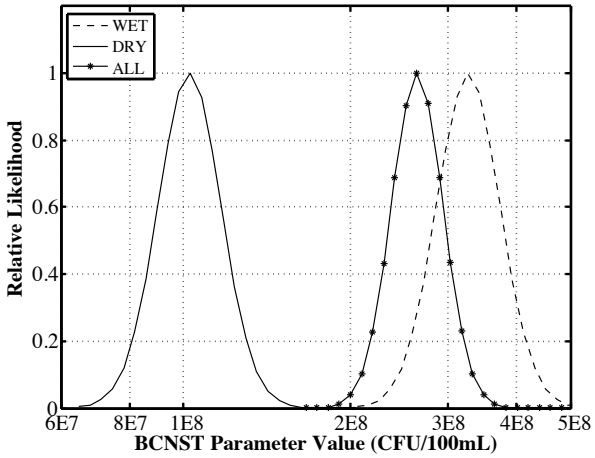
Figure 6.8: PSRC (HSPF) sensitivity in three hydrologic states



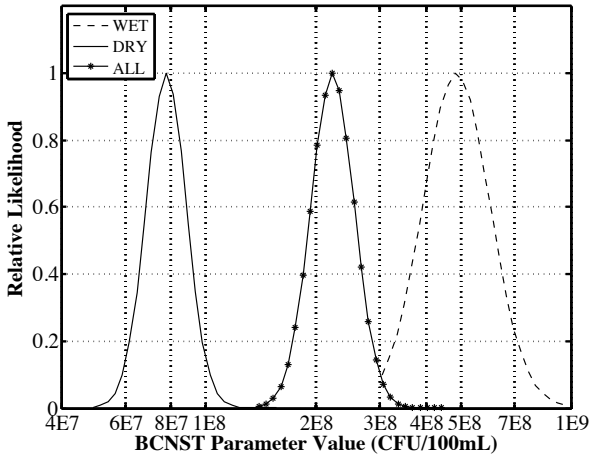
(a) Catchment I



(b) Catchment J



(c) Catchment K



(d) Catchment O

Figure 6.9: BCNST (SWAT) sensitivity in three hydrologic states

between that of the wet and dry states in all cases except for Catchment I. The value of IOQC is of the same magnitude on a per-catchment basis. In Figure 6.8 it can be seen that the value of PSRC is greater in the dry state than the wet state in all cases but Catchment O. The shape of the curves show that PSRC is more sensitive in the dry state than the wet state, and that all the states are within an order of magnitude. Additionally, as was shown in Figures 6.2-6.3, the magnitude of the PSRC sensitivity was comparable to IOQC in most cases. This sensitivity analysis of PSRC is important since it is not a default parameter in HSPF but was added as a custom input. It is fair to assume, then, that the HSPF models would perform significantly different and not as well without the inclusion of the sensitive background parameter PSRC. In Figure 6.9 it can be seen that the parameter BCNST has greater magnitude in the wet state than the dry state in all but Catchment I, but the dry state is more sensitive. Similarly, the magnitude of the parameter in the "all" state is between the wet and dry states in all but Catchment I. BCNST tends to be more sensitive in the dry state than the other two states, and the parameter is within an order of magnitude in the three states on a per catchment basis. As with PSRC in HSPF, BCNST is not a default parameter in SWAT and was added as a custom parameter to the models in this study. It is fair to assume that not including BCNST would result in models that perform significantly different and not as well as the models considered here.

6.3.2 HSPF and SWAT Calibration Results

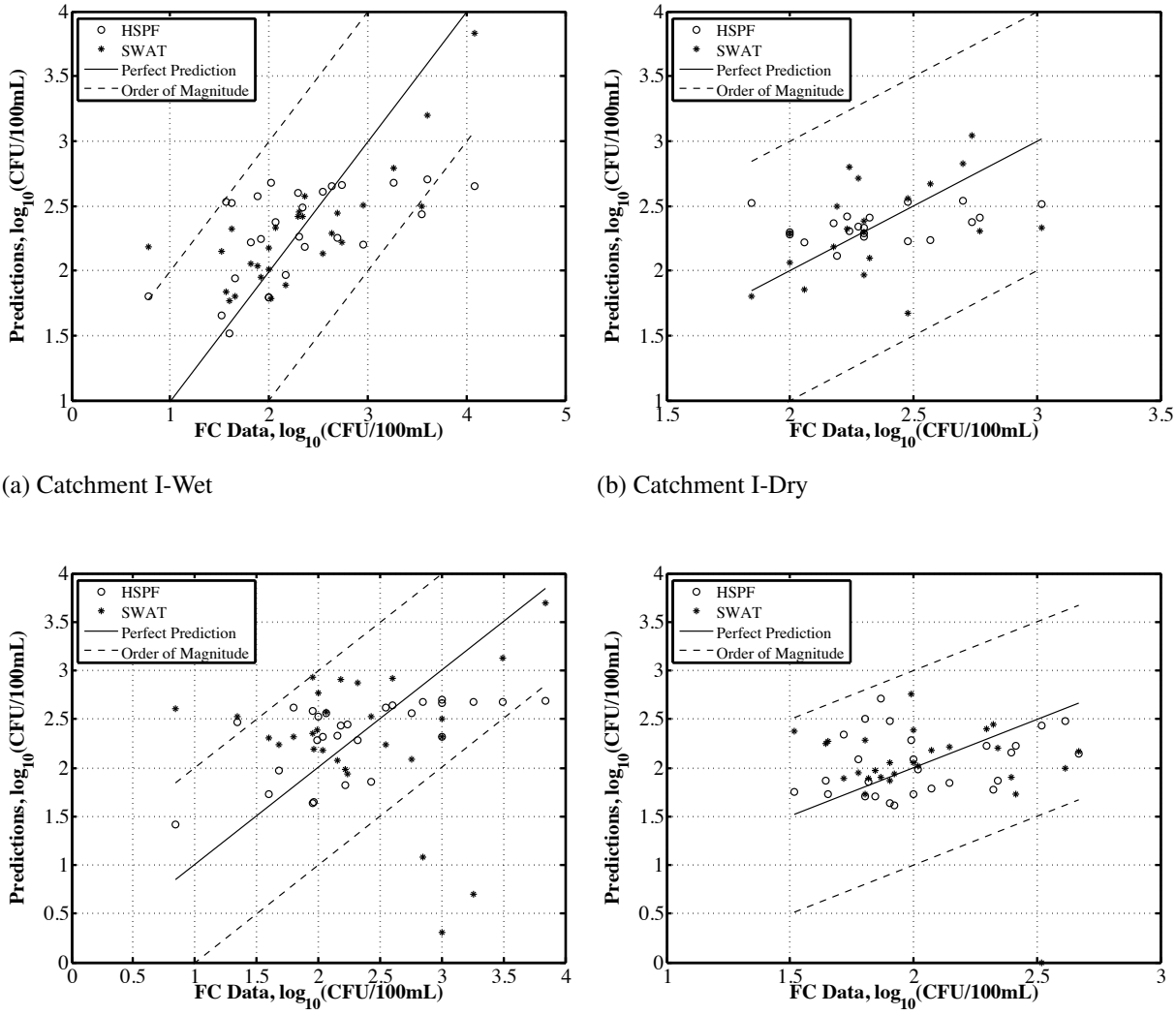
The NSE_{ln} and p values of the calibrated wet and dry water-quality models are shown in Table 6.6. Also shown in Table 6.6, as a reference and labelled "all", are the calibrated water-quality models using the whole data set as reported in Chapter 5. The results show that in all cases except for Catchment O in SWAT that the wet models achieved the highest NSE_{ln} followed by the models using the whole data set ("all") and finally the dry models. The pattern is reversed in SWAT in Catchment O, where the dry achieved the highest NSE_{ln} followed by the wet and the "all". Even though the wet versions of both models make

Table 6.6: Summary of water-quality model calibrations

Catchment	Hydrologic State	Model NSE_{ln}		*p* Value	
		HSPF	SWAT	HSPF	SWAT
I	all	0.243	**0.275**	0.07	0.86
	wet	0.313	**0.588**	0.78	0.06
	dry	**0.120**	-0.432	0.85	0.59
J	all	**0.263**	-1.67	0.57	**3.1E-4**
	wet	**0.370**	-1.31	**1.9E-3**	**0.01**
	dry	**-0.436**	-2.84	0.12	**1.7E-4**
K	all	**0.383**	0.334	0.82	0.48
	wet	0.438	**0.592**	0.79	0.26
	dry	**0.295**	0.260	0.79	0.79
O	all	**0.342**	-0.028	0.21	0.10
	wet	**0.484**	0.066	0.05	0.06
	dry	-0.168	**0.152**	0.43	0.17

improvements over the "all" models in Catchments I, J, and K, it appears the magnitude of improvement in each case is greater in SWAT than in HSPF. This is particularly obvious in Catchments I and K, where HSPF and SWAT have nearly equal NSE_{ln} values in the "all" state, but the wet state of SWAT in these catchments have a notably higher NSE_{ln} than the corresponding HSPF models. Conversely, the dry states of SWAT in Catchments I and K have NSE_{ln} values lower than the dry states in HSPF. This result is likely due to the daily vs. hourly difference in time resolution between the models, which tends to make SWAT fluctuate to greater extremes, as previously discussed.

The HSPF and SWAT wet and dry calibrated model predictions are compared with corresponding measurements in Figures 6.10-6.11, where the abscissa is the log-scaled wet/dry FC data and the ordinate is the log-scaled model predictions. The solid line in the figures has a slope of 1 and represents a perfect model prediction whereas the dashed line indicates order-of-magnitude agreement. In all cases the models tended to over-predict at lower measurement values and under-predict at higher measurement values, as was found in Chapter 5. In some, but not all cases, the model predictions deviated the most at extreme highs and lows. All of the dry models, with exception to one HSPF point in Catchment K,



(a) Catchment I-Wet

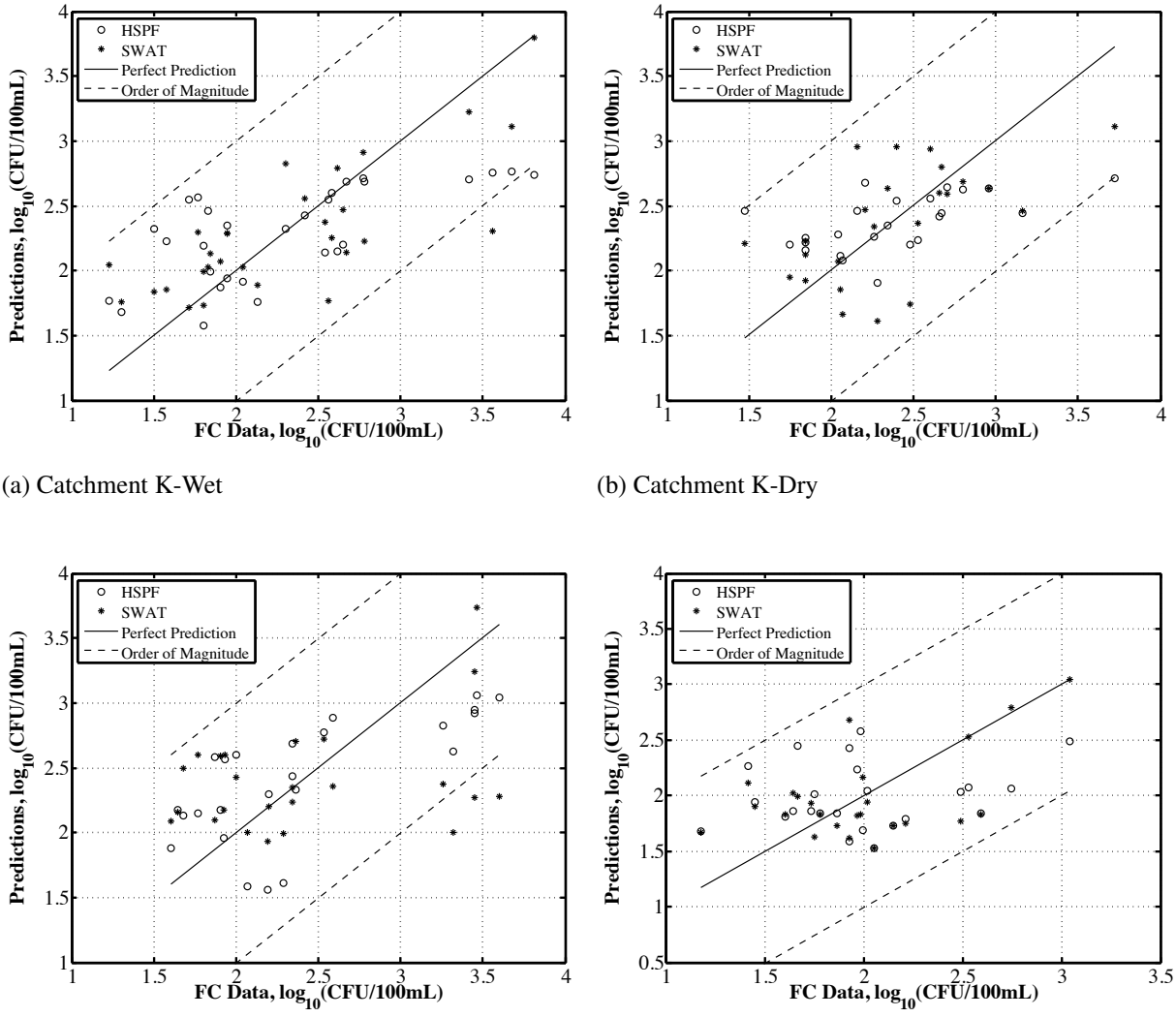
(b) Catchment I-Dry

(c) Catchment J-Wet

(d) Catchment J-Dry

Figure 6.10: HSPF and SWAT model results- log-scaled seasonal predictions

had model predictions within one order of magnitude of the data. Conversely, all the wet models had at least two points with residuals greater than one order of magnitude. It is interesting to note that the division of FC data points into wet and dry states, as described earlier, results in a wider range of data values considered in the wet state of each catchment than in the dry state, the only exception being the lowest values considered in Catchment O. In other words, the largest data point considered in the wet state is of greater magnitude than the largest data point in the dry state, and the smallest data point considered in the wet



(a) Catchment K-Wet

(b) Catchment K-Dry

(c) Catchment O-Wet

(d) Catchment O-Dry

Figure 6.11: HSPF and SWAT model results- log-scaled seasonal predictions

state is of lesser magnitude than the smallest data point in the dry state (with exception to the low points in Catchment O). This implies that even though the wet state includes more streamflow, the extremes of the data sets, both high and low, occur under wet conditions. Conversely, the data sets collected in dry conditions incorporated a smaller overall range of bacteria concentrations.

6.4 CONCLUSION

The wet and dry models presented in this chapter reflect the complex relationship of in-stream bacteria concentrations and the hydrologic state of a watershed. Several important conclusions can be drawn in analyzing the differences between how HSPF and SWAT perform in the wet and dry states. For example, the fundamental difference in how the models handle subsurface bacteria concentrations has a more pronounced effect in a wet/dry analysis, where HSPF becomes more sensitive to interflow concentrations (IOQC) and still has insensitive land-application parameters (ACQOP) in the wet state, while SWAT does not model interflow concentrations but instead accounts for the difference with land-application parameters (CFRT) that are notably more sensitive in the wet state than in the dry state. Both models reveal that the parameters describing background sources (PSRC, BCNST) and first-order decay (FSTDEC, WDPRCH) are more sensitive in the dry state, when the system has less water and these parameters have a greater effect on the model output. The results also show that the observed increase in concentrations in wet conditions are modeled in HSPF as contributions from interflow and in SWAT as a function of a rainfall-runoff terrestrial process. In reality the process is a complex combination of both phenomena and therefore neither model accounts for both processes appropriately. Finally, in light of model uncertainty, the wet and dry analysis further revealed that certain parameters and/or process equations can be reduced or completely removed from consideration, simultaneously reducing parameter and structural uncertainty while not affecting overall model performance. A reduction in uncertainty may be particularly useful in dry conditions when only in-stream processes are actually active.

Future research can extend the wet/dry analysis to reveal more of the natural processes occurring and then to refine the models to reflect those processes. Extensive data collection relating specifically to bacteria concentrations in wet and dry states would clearly be a major asset and should be one of the first steps taken.

Chapter 7

Model Combinations

7.1 BACKGROUND

This study examines several methods of combining the models presented thus far to attain an overall better fit to the data. The first section provides a framework for integrating the water-quality component of the CC model with the hydrological components of HSPF and SWAT. The second section provides two methods for combining the outputs of HSPF and SWAT to achieve a higher NSE_{ln}.

7.2 HSPF-CC AND SWAT-CC MODEL INTEGRATION

7.2.1 Theory

An additional application of the CC approach is to use it as a water-quality model in combination with the hydrological components of models like HSPF and SWAT. The CC-combination model can be formulated as (Chin 2011):

$$C_j = \left(\frac{Q_{r,j}}{Q_j}\right)C_r + \left(\frac{Q_{i,j}}{Q_j}\right)C_i + \left(\frac{Q_{b,j}}{Q_j}\right)C_b \tag{7.1}$$

where the subscript j is the time index, C_j is the predicted concentration, Q_j is the total streamflow, $Q_{r,j}$ is the surface-runoff component, $Q_{i,j}$ is the interflow component, $Q_{b,j}$ is the baseflow component, and c_r, c_i, and c_b are the surface-runoff, interflow, and base-flow characteristic concentrations, respectively. The flow quantities Q_j, $Q_{r,j}$, $Q_{i,j}$, and $Q_{b,j}$ are derived from the hydrology output of HSPF and SWAT. Using these HSPF-CC and SWAT-CC model formulations, a calibration was performed to find values of c_r, c_i, and c_b that yield the best match between the log-scaled model predictions (C_j) and the log-scaled data. The calibration results of these models, along with the NSE_{ln} and p value are included in Table 7.1.

7.2.2 Results and Discussion

A breakdown of the flow volumes produced by the different models is shown in Table 7.2, where the HSPF and SWAT models shown for each catchment are the calibrated models discussed previously. It can be seen that the total volumes produced in HSPF and SWAT are of the same order of magnitude and nearly the same value as the flow record for the seven-year period considered. In all cases, HSPF over-predicted the total flow volume and SWAT over-predicted in Catchment K and under-predicted in I, J, and O. The distribution of flow volume between runoff (V_r), interflow (V_i), and baseflow (V_b) reveals some of the differences in the HSPF and SWAT model structure. HSPF tends to generate little runoff volume (0-19%) and assigns a significant proportion (21-67%) of the total volume to interflow. This distribution of flow in HSPF is a direct representation of Figure 3.2 for larger areas. SWAT places a greater emphasis on runoff volume (16-33%) and a smaller emphasis on interflow (5-19%). The baseflow volumes in HSPF and SWAT range from 33-60% and 59-76%, respectively. The baseflow volume percentages, or baseflow index (BFI), can be

Table 7.1: CC Integration Model Results and Bacteria Load Distributions

				Concentrations (CFU/100mL)			Bacteria loads ($\times 10^{13}$ CFU) (%)				
Catchment	Model	NSE_{ln}	p	c_r	c_i	c_b	M_r	M_i	M_b	M_0†	M
	HSPF	0.243	0.74	-	-	-	359 (93)	25.2 (7)	-	0.55 (0)*	384
	SWAT	0.275	0.86	-	-	-	152 (70)	-	-	64.7 (30)	217
	HSPF-CC	0.090	0.71	10000‡	482	150	3.16 (7)	37.6 (78)	7.25 (15)	-	48.0
I	SWAT-CC	0.254	0.09	1607	301	127	51.1 (80)	2.14 (4)	10.4 (16)	-	63.7
	CC-SI	0.329	0.03	1777	-	88	91.6 (93)	-	6.39 (7)	-	98.0
	CC-FI	0.439	0.002	2668	-	99	140 (95)	-	7.08 (5)	-	147
	CC-LM	0.217	0.36	662	-	114	42.6 (86)	-	6.82 (14)	-	49.4
	HSPF	0.263	0.57	-	-	-	0.16 (1)	21.5 (99)	-	0.06 (0)*	21.7
	SWAT	-1.671	3E-4	-	-	-	139 (98)	-	-	2.12 (2)	141
	HSPF-CC	0.237	0.48	51	556	97	6E-3 (0)*	21.0 (92)	1.82 (8)	-	22.9
J	SWAT-CC	0.338	0.24	242	108	912	1.57 (6)	0.87 (3)	24.5 (91)	-	26.9
	CC-SI	0.325	0.01	910	-	65	23.8 (93)	-	1.85 (7)	-	25.7
	CC-FI	0.402	0.20	845	-	73	22.4 (92)	-	2.05 (8)	-	24.5
	CC-LM	0.281	0.23	477	-	81	15.3 (89)	-	1.82 (11)	-	17.1
	HSPF	0.387	0.92	-	-	-	3.59 (13)	23.7 (86)	-	0.32 (1)	27.6
	SWAT	0.334	0.48	-	-	-	40.6 (86)	-	-	6.75 (14)	47.3
	HSPF**	0.35	-	-	-	-	6.9 (44)	4.7 (30)	-	4.0 (26)	15.6
	SWAT**	0.12	-	-	-	-	112 (93)	-	-	7.98 (7)	120
K	HSPF-CC	0.050	0.46	1	474	180	4E-4 (0)*	8.22 (67)	4.00 (33)	-	12.2
	SWAT-CC	0.256	0.63	1673	473	119	11.4 (73)	0.95 (6)	3.37 (21)	-	15.7
	CC-SI	0.277	0.11	1110	-	95	15.8 (88)	-	2.12 (12)	-	17.9
	CC-FI	0.370	0.004	1243	-	92	17.9 (90)	-	2.03 (10)	-	19.9
	CC-LM	0.302	0.19	894	-	86	16.5 (91)	-	1.55 (9)	-	18.1
	HSPF	0.342	0.21	-	-	-	16.2 (64)	8.96 (36)	-	0.05 (0)*	25.2
	SWAT	-0.028	0.10	-	-	-	138 (96)	-	-	5.65 (4)	144
	HSPF-CC	0.264	0.06	1	19143	119	6E-3 (0)*	122 (98)	2.17 (2)	-	124
O	SWAT-CC	0.316	0.60	6570	139	106	62.8 (96)	0.33 (1)	1.84 (3)	-	65.0
	CC-SI	0.370	0.10	2257	-	63	31.1 (97)	-	1.00 (3)	-	32.1
	CC-FI	0.297	0.92	1087	-	77	15.2 (93)	-	1.21 (7)	-	16.4
	CC-LM	0.395	0.68	913	-	66	14.9 (94)	-	0.88 (6)	-	15.8

†- background loads derived from PSRC (HSPF) and BCNST (SWAT) model parameters

*- actual % is nonzero but less than 0.5% and therefore rounds to zero

‡- In this case c_r was relatively insensitive and a ceiling of 10,000 (CFU/100mL) was imposed

**- Catchment K model results from Chin (2011)

Table 7.2: Model Flow Distribution

Catchment	Model	NSED	V_r ($\times 10^6$ m^3) (%)	V_i ($\times 10^6$ m^3) (%)	V_b ($\times 10^6$ m^3) (%)	V ($\times 10^6$ m^3)	BFI
I	HSPF	0.884	0.3 (0)*	78.0 (62)	48.4 (38)	126.7	0.38
	SWAT	0.595	31.8 (26)	7.1 (6)	82.2 (68)	121.1	0.68
	CC-SI	1.00	51.5 (42)	-	72.6 (58)	124.2†	0.55
	CC-FI	1.00	52.6 (42)	-	71.6 (58)	124.2	
	CC-LM	1.00	64.3 (52)	-	59.8 (48)	124.2†	
J	HSPF	0.880	0.1 (0)*	37.8 (67)	18.8 (33)	56.7	0.33
	SWAT	0.626	6.5 (16)	8.0 (19)	26.9 (65)	41.4	0.65
	CC-SI	1.00	26.2 (48)	-	28.4 (52)	54.6	0.48
	CC-FI	1.00	26.5 (49)	-	28.0 (51)	54.6†	
	CC-LM	1.00	32.1 (59)	-	22.5 (41)	54.6	
K	HSPF	0.897	0.4 (1)	17.4 (43)	22.2 (56)	40.0	0.56
	SWAT	0.670	6.8 (18)	2.0 (5)	28.3 (76)	37.1	0.76
	HSPF**	0.90	0.3 (1)	19.1 (48)	20.6 (52)	39.9	0.52
	SWAT**	0.67	8.5 (22)	2.7 (7)	27.9 (71)	39.1	0.71
	CC-SI	1.00	14.2 (39)	-	22.3 (61)	36.5	0.57
	CC-FI	1.00	14.4 (39)	-	22.1 (61)	36.5	
	CC-LM	1.00	18.5 (51)	-	18.0 (49)	36.5	
O	HSPF	0.941	5.6 (19)	6.4 (21)	18.2 (60)	30.2	0.60
	SWAT	0.693	9.6 (33)	2.3 (8)	17.3 (59)	29.2	0.59
	CC-SI	1.00	13.8 (46)	-	15.9 (46)	29.6†	0.51
	CC-FI	1.00	14.0 (47)	-	15.7 (47)	29.6†	
	CC-LM	1.00	16.4 (55)	-	13.3 (45)	29.6†	

†- $V_r/V_i/V_b$ components do not sum to V as shown due to rounding error at 10^6 m^3 magnitude

*- actual % is nonzero but less than 0.5% and therefore rounds to zero

**- Catchment K model results from Chin (2011)

compared using the CC (USGS) hydrograph separation methods as a benchmark and are shown together in Table 7.2, where the BFI shown for the CC models is the average BFI of the three methods. In Catchments I, J, and K, HSPF has the lowest BFI, SWAT has the highest BFI, and the CC models are in between, while in Catchment O HSPF and SWAT are nearly equal and both are greater than the CC models. On one hand, HSPF generates a noticeably higher NSED than SWAT in each catchment, indicating that it more-closely matches the streamflow on a daily basis. On the other hand, Bosch et al. (2005) report that for a five-year study period the surface runoff and lateral subsurface flow for a study plot in the LREW were in the range 5-40% and 2-23% of annual precipitation (Note: not percentage of streamflow as shown in Table 7.2), respectively, depending on agricultural land-tilling practices. This data implies that while interflow in the LREW can occur in notable quantities, the high percentage as modeled by HSPF is probably not realistic. Although SWAT is more successful at replicating the realistic runoff and interflow quantities, it overestimates the BFI and in general is less able to replicate the streamflow record as evidenced by the lower NSED. These hydrology results in HSPF and SWAT show that while the models satisfactorily replicate the streamflow on daily (NSED) and monthly (NSEM) bases, the breakdown of flow components contributing to the total flow may not be accurate or realistic. These results have an additional effect on the water-quality components of the models, since the water-quality predictions are dependent on the hydrology components of the models. The CC models have an NSED of 1.00 since they use the actual streamflow data. Of the three methods, the LM method consistently generates the lowest baseflow and highest runoff volumes, while the SI and FI are nearly identical. The Chin (2011) study only examined Catchment K, and the results are included in Table 7.2 as a reference. Although the models in Chin (2011) used different HSPF and SWAT model parameterizations the flow distribution is nearly identical to the models considered here.

A breakdown of the water-quality model performances is shown in Table 7.1, where the NSE_{ln}, characteristic-concentration parameters, and distribution of bacteria loads are shown

for all the models. The HSPF models show little consistency in the breakdown of runoff (M_r), interflow (M_i), and background (M_0) bacteria loads, other than the background load being a small portion of the total. The runoff load, (M_r), varies from 1-93% of the total load, while the interflow load, (M_i), varies from 7-99% of the total load. The SWAT models show a more consistent pattern of load distribution, where the runoff load, (M_i), is 70-98% of the total bacteria load. The more consistent load distribution shown by SWAT is partially due to the simpler model formulation, where SWAT does not permit bacteria loads in subsurface flow and the total load M must therefore come from either surface runoff (M_r) or background (M_0) sources. Although HSPF provides an additional component for modeling and analysis with the inclusion of the interflow term, the extra term does not appear to induce consistent results with respect to the distribution of loads. The loading results for Catchment K from Chin (2011) are included in Table 7.1 as a reference. The HSPF models show order-of-magnitude agreement, while the SWAT model in Chin (2011) has a larger total load (M) and runoff load (M_r) but a similar percentage distribution between runoff (M_r) and background (M_0) sources as the model in this study. While it is acknowledged that bacteria concentrations in interflow may be nonzero, it is unlikely that the high percentages modeled in HSPF and HSPF-CC and shown in Table 7.1 are realistic. On the contrary, it has been shown throughout this dissertation that terrestrial sources may not have a significant effect on in-stream concentrations, meaning that the high percentages in the runoff loads, particularly in SWAT and the CC model, are more a manifestation of the model structure instead of a realistic estimate. The most physically realistic scenario is one with an interflow term like HSPF, a more consistent runoff contribution like in SWAT yet with a lower percentage of the total load, and a larger percentage of the total load coming from the background and in-stream sources.

The HSPF-CC models produce noticeably lower NSE_{ln} values than HSPF in Catchments I and K, and comparable NSE_{ln} values in J and O. The CC parameters in the HSPF-CC model (Table 7.1) reveal small runoff values (c_r), and interflow values (c_i) greater than

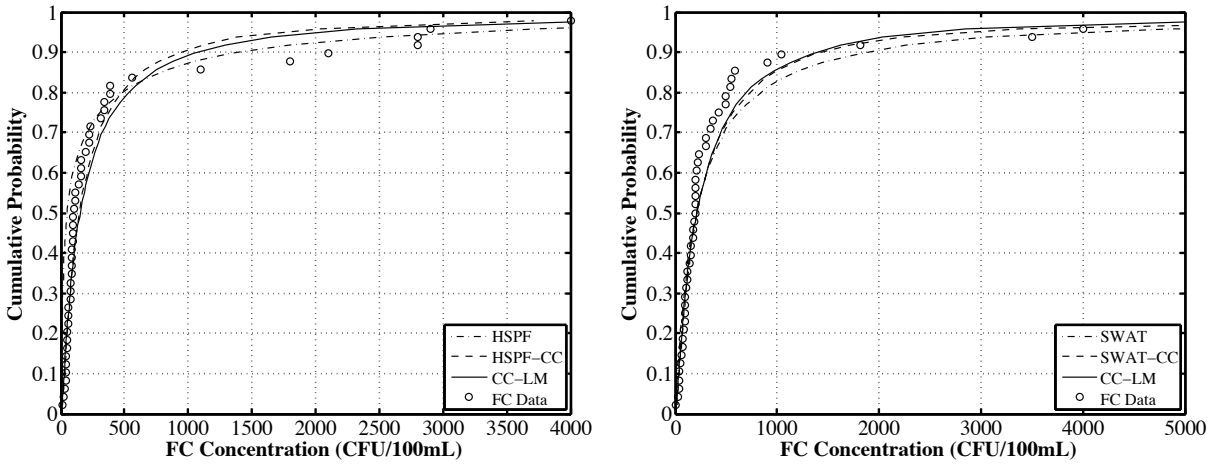
baseflow values (c_b). The lone exception is Catchment I, where the c_r parameter was insensitive and a ceiling of 10,000 (CFU/100mL) was imposed on the parameter (without affecting the performance), to facilitate calibration. The low values of c_r and small contribution of runoff load (M_r) to the total bacteria load (M) as shown in Table 7.1 are a reflection of the small overall contribution of runoff volume (V_r) to total volume (V) as previously discussed and shown in Table 7.2. It is worth noting that the parameter sensitivity analysis from Chapter 5 revealed that the in-stream parameters IOQC and PSRC were the most sensitive water-quality parameters in HSPF. These two parameters are directly related to the interflow (M_i) and background (M_0) loads, respectively. Although PSRC is sensitive it does not contribute more than 1% of the load in any catchment.

The SWAT-CC models achieve similar NSE_{ln} values to the SWAT models in Catchments I and K, and noticeably improve the negative NSE_{ln} values of SWAT in Catchments J and O. In Catchments I, K, and O, the SWAT-CC model yields c_r parameters roughly one order of magnitude greater than c_i and c_b, and c_i is greater than c_b. Likewise, in these three catchments, the runoff loading (M_r) is 73-96% of the total bacteria load (M). The SWAT-CC model in Catchment J does not fit this (expected) trend but rather assigns a high value to the baseflow concentration (c_b), resulting in a large (91%) contribution of baseflow load (M_b) to the total bacteria load (M). While this result may seem unconventional, the SWAT-CC model achieves an NSE_{ln} of 0.338 while the corresponding SWAT model has an NSE_{ln} of -1.671. This result reveals that including subsurface bacteria concentrations in SWAT may noticeably improve the model in certain cases. The sensitivity analysis from Chapter 5 showed that the direct source parameter (BCNST) was one of the sensitive parameters, but the load derived from it (M_0) is 30% of the total load in Catchment I and between 2-7% in Catchments J, K, and O.

The 2-parameter CC models from Chapter 5 are also included in Table 7.1. These CC models perform consistently regarding the load distribution in all cases, where the NSE_{ln}

values range from 0.217-0.439, the c_r parameter is greater than c_b, and the runoff loading, M_r, is 86-97% of the total bacteria load (M).

As a further measure of model performance, the expected frequency distributions of in-stream bacteria concentrations were determined and are shown in Figure 7.1, where Catchment O was selected as an example due to the high NSE_{ln} values in the HSPF, HSPF-CC and CC-LM models, and Catchment I was selected as an example due to the high NSE_{ln} values in the SWAT, SWAT-CC, and CC-LM models. Due to the random component



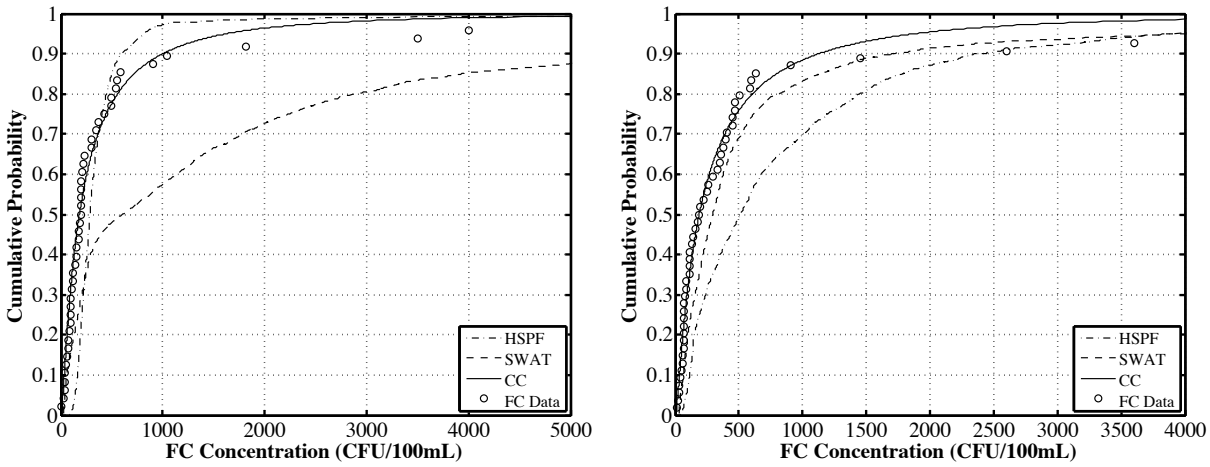
(a) HSPF-Catchment O (b) SWAT-Catchment I

Figure 7.1: Predicted and measured frequency distributions- data measurement days

in the CC model formulation (Eqs. 3.30 and 3.31), the frequency distributions shown in the HSPF-CC, SWAT-CC and CC-LM models are ensemble-averaged distributions from 500 runs. The frequency distributions shown in Figure 7.1a reveal that all three models closely represent the data up to the 80^{th} percentile, after which the HSPF-CC and CC-LM models are similar and more-successfully replicate concentrations of higher magnitude while the HSPF model more closely approximates the data in the range 1000-3000 CFU/100mL. The frequency distributions shown in Figure 7.1b reveal that the SWAT, SWAT-CC, and CC-LM models closely represent the data up to the 50^{th} percentile. All three models under-predict

the data to around the 90^{th} percentile, after which the models again closely approximate the data at higher magnitudes.

While the frequency distributions shown in Figure 7.1 reflect only the days on which the FC data was collected, the frequency distributions shown in Figure 7.2 for Catchments I and K reflect the entire 7-year calibration period. The HSPF and SWAT distributions shown were



(a) Catchment I (b) Catchment K

Figure 7.2: Predicted frequency distributions- entire study period

derived from the model predictions of the calibrated models from Chapter 5, while the CC-LM model was derived from an ensemble average after 500 runs. It can be seen clearly that the CC-LM model in both catchments best replicates the data to the 80^{th}-percentile, and in Catchment I better approximates the data than either HSPF or SWAT, while in Catchment K both HSPF and SWAT better approximate the data at higher magnitudes. In comparing the frequency distributions of the SWAT models in Catchment I, as shown in Figure 7.1b and 7.2a, it is clear that a bias is induced in only considering the days of data measurements, and that the entire record includes more concentrations of higher magnitude than those reflected on only the data days.

7.2.3 Conclusions: CC Model Integration

The HSPF-CC and SWAT-CC model integration technique provides a straightforward way to incorporate the characteristic concentration water-quality model into the established hydrological frameworks of HSPF and SWAT. The HSPF-CC model provided similar NSE_{ln} values to the HSPF model in two of the four catchments considered while the SWAT-CC model provided similar NSE_{ln} values in two catchments and notably improved the NSE_{ln} in two catchments. The HSPF-CC and SWAT-CC approach also quantifies how the models distribute flow volumes and bacteria loadings between runoff, interflow, and baseflow (Tables 7.1 and 7.2). The distribution of flow is a reflection of the calibrated hydrology components of the models and is particularly important since the bacteria prediction capability of the water-quality components is dependent on the hydrology components. In this sense the HSPF models in particular, with an unusually high ratio of interflow (V_i) to runoff (V_r), likely do not adequately model the runoff and groundwater flows and therefore may compromise the water-quality components of the model. The exact distribution of flow as shown in Table 7.2 is likely complicated by parameter correlation in the hydrology components of of the models as mentioned in Chapter 5. Finally, it was shown that the HSPF-CC and SWAT-CC model integrations were simply formulated, the integrated models used fewer parameters than the HSPF and SWAT models, and the integrated models achieved similar or better results than the corresponding HSPF and SWAT models. These results show that the CC model can be an attractive additional option for water-quality models, and may even be included as an external module within the HSPF and SWAT models for future applications.

7.3 SYNTHESIS OF MODEL OUTPUT

One way to attain improved model predictions is by using a multi-model framework, where the outputs of two or more models are combined using a weighting scheme to achieve a

result better than any of the individual input models. The multi-model approach does not directly change the structural components of the input models, but rather finds an optimal combination of the model outputs to arrive at an improved set of predictions that better fits the data. This study examines two weighting schemes for combining HSPF and SWAT to attain improved model predictions: an NSE_{ln} optimization technique and the use of Artificial Neural Networks (ANN).

7.3.1 NSE_{ln} Optimization

Theory

The first approach is to optimize a combination of HSPF and SWAT to achieve a higher NSE_{ln} than the individual models. The procedure is described by (Chin et al. 2009)

$$CP_i = aH_i + (1-a)S_i \tag{7.2}$$

where CP_i is the combined prediction at time step i, H_i and S_i are the corresponding predictions by HSPF and SWAT respectively, and a is a weighting factor between zero and one. Using this multi-model relationship in the definition of the NSE yields:

$$NSE = 1 - \frac{\sum_{i=1}^{N}[M_i - aH_i - (1-a)S_i]^2}{\sum_{i=1}^{N}(M_i - \overline{M})^2} \tag{7.3}$$

where M_i is the measurement at time step i, and $\overline{M_i}$ is the average of the N measurements. Taking the derivative of Equation 7.3 with respect to a gives the optimal value of a for weighting the models:

$$\frac{dNSE}{da} = 0 \Rightarrow a = \frac{\sum_{i=1}^{N}(M_i - S_i)(H_i - S_i)}{\sum_{i=1}^{N}(S_i - H_i)^2} \tag{7.4}$$

It can be further shown that

$$\frac{d^2NSE}{da^2} = -\sum_{i=1}^{N}(S_i - H_i)^2 < 0 \tag{7.5}$$

guaranteeing that using the value of a given by Equation 7.4 will maximize the NSE and will be greater than (or equal to, if a = 1.0 or 0) the NSE value from an individual model. While the above derivation was presented in terms of NSE, the NSE_{ln} can be found by using the log-transformed values of HSPF and SWAT in Equations 7.3 as follows:

$$NSE_{\ln} = 1 - \frac{\sum_{i=1}^{N}[M_i - a\ln H_i - (1-a)\ln S_i]^2}{\sum_{i=1}^{N}(M_i - \overline{M})^2} \tag{7.6}$$

Results and Discussion

The NSE_{ln} optimization approach was used with HSPF and SWAT for the entire FC data set ("all") as well as the wet and dry models. The results are presented in Table 7.3, where the weighting factors a and $1-a$ are shown along with the initial model NSE_{ln} and the

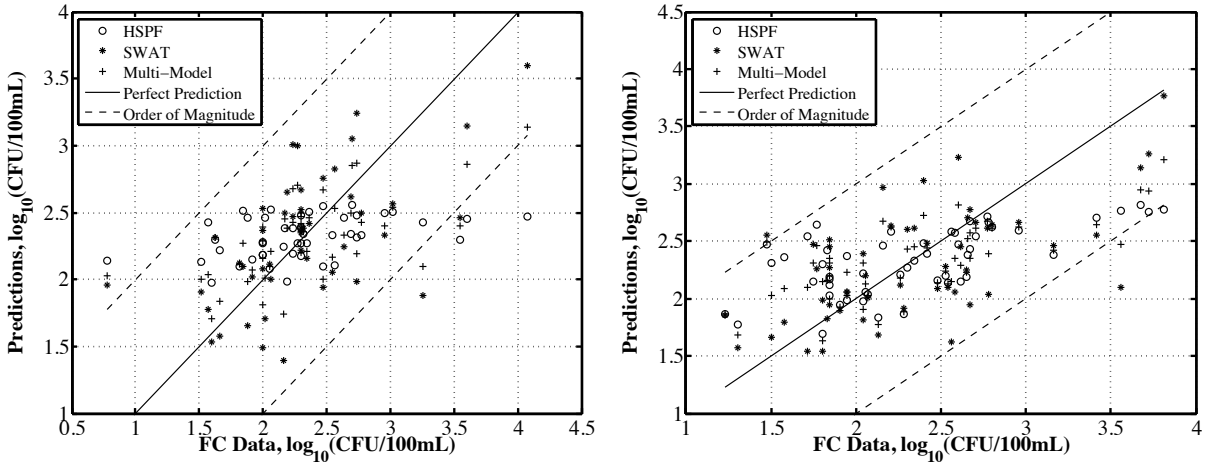
Table 7.3: NSE_{ln} Optimization Multi-Model Summary

Catchment	Previous NSE_{ln}		Multi-Model Results		
	HSPF	SWAT	% HSPF (a)	% SWAT (1-a)	Optimized NSE_{ln}
I all	0.243	0.275	0.47	0.53	0.422
J all	0.263	-1.67	0.93	0.07	0.272
K all	0.383	0.334	0.56	0.44	0.471
O all	0.342	-0.028	0.75	0.25	0.390
I wet	0.313	0.588	0.11	0.89	0.593
J wet	0.370	-1.31	1.00	0.00	0.370
K wet	0.438	0.592	0.31	0.69	0.632
O wet	0.484	0.066	1.00	0.00	0.484
I dry	0.120	-0.432	0.68	0.32	0.278
J dry	-0.436	-2.84	0.80	0.20	-0.258
K dry	0.295	0.260	0.56	0.44	0.357
O dry	-0.168	0.152	0.22	0.78	0.179

optimized multi-model NSE_{ln}. The NSE_{ln} optimization multi-model approach provided an improved NSE_{ln} over either HSPF or SWAT independently in all cases except for the wet models in Catchments J and O, where $a = 1.00$. The improvements were greatest when the initial models were weighted nearly evenly (a near 0.50) such as in Catchments I and K or the dry model of Catchment K. Conversely, the NSE_{ln} optimization approach achieved less significant improvements in NSE_{ln} when one input model was heavily favored (a near 1.00 or 0), such as in the wet models of Catchments J and O, in the dry model of Catchment O, or in Catchment J.

The HSPF weighting factor a can be used as a simple parameter to compare model performance. Of the twelve model scenarios considered in Table 7.3, a was greater than 0.50 eight times, indicating that HSPF outperformed SWAT two-thirds of the time. It should be stated that SWAT did not perform well in Catchment J under any circumstances, achieving a negative NSE_{ln} under "all", wet, and dry conditions. It should also be noted that there is a variety of values of a, both between catchments and throughout hydrological states. The lack of a consistent or predictable pattern of weighting the models is a further reflection on the difficult and unpredictable nature of bacteria modeling. An example of this complexity can be seen in Catchment I, where the models are weighted roughly equal ($a = 0.47$) under normal conditions, but where SWAT performs noticeably better under wet conditions ($a = 0.11$) and HSPF performs better under dry conditions ($a = 0.68$). Similarly, $a = 1.00$ in the wet model of Catchment O while $a = 0.22$ in the dry model of Catchment O.

The HSPF, SWAT, and NSE_{ln} optimized model predictions are compared with corresponding measurements for Catchments I and K in Figure 7.3, where the abscissa is the log-scaled FC data and the ordinate is the log-scaled model predictions. The solid line in the figures has a slope of 1 and represents a perfect model prediction whereas the dashed line indicates order-of-magnitude agreement. It is clear in Figure 7.3 that the NSE_{ln} optimization technique yields model predictions that are between the HSPF and SWAT



(a) Catchment I (b) Catchment K

Figure 7.3: HSPF, SWAT, and NSE_{ln}-optimized model results- log-scaled predictions

predictions at each point and always closer to the line of perfect prediction than the model with the largest residual at that point. Since the magnitude of the HSPF and SWAT model residuals vary at each point, and one model is not consistently better or worse than the other model at each point, the multi-model framework guarantees a higher NSE_{ln} over the course of the whole data set.

7.3.2 Artificial Neural Networks

Background

Artificial Neural Networks (ANN) are a form of artificial intelligence model that is designed to replicate the processes of the human neurological system (Neelakantan et al. 2001). ANNs are advantageous in that they generally have fewer data requirements and have been shown to solve complex, non-linear problems where the modeled relationships between variables are not well understood (Yu et al. 2004; Singh et al. 2009; Brion and Lingireddy 2003). Several studies have used ANNs successfully in modeling the complex environmental fate and transport processes of bacteria and other water-quality constituents

(e.g. Singh et al. 2009; Brion et al. 2002; Brion and Lingireddy 2003; Neelakantan et al. 2001; Ying et al. 2007; Sakura-Lemessy 2009). However, despite the powerful adaptability of an ANN there is concern that the "black-box" nature of an ANN, where little is understood of exactly how the complex inner-workings of the network produce the output from the input, does not reveal or replicate actual processes and therefore the networks should be used carefully before deducing any physical meaning (e.g. Prada-Sarmiento and Obregón-Neira 2009; Brion et al. 2002; Troutman 1985).

This portion of the research was not intended to be a comprehensive overview of ANNs but rather an example of how an ANN designed from standard practices and recommendations in the literature can be used in a multi-model combination context for bacterial water-quality modeling. All networks considered in this study employed a feedforward network structure with back-propagation of errors. The feedforward network is one of the most popular and important ANN structures and back-propagation is a popular training algorithm (Schalkoff 1997; Brion and Lingireddy 1999). Feedforward network structure has an input layer, zero or more hidden layers, and an output layer. Input layers serve to hold the input values and distribute them to nodes in the next layer. Although the number of hidden layers is unlimited, a single hidden layer is typically adequate (Weijters and Hoppenbrouwers 1995). Each layer has at least one node, and the nodes of each layer are linked to each node of the next higher layer, with no connections between non-successive layers (Weijters and Hoppenbrouwers 1995). A weight is applied to each link, and a bias (also considered a weight) is applied to each node as an additional method to adjust network performance. Singh et al. (2009) recommend between ($2N^{1/2}$) to (2N+1) nodes in the hidden layer, where N is the number of input nodes and M is the number of output nodes, while Brion et al. (2002) and Neelakantan et al. (2001) recommend a N:2N:1 network architecture for water-quality models.

The weights applied to each link vary and are the primary parameters affecting network output, and *training* is the process by which these weights are calibrated (Schalkoff 1997; Brion and Lingireddy 2003). ANNs, in fact, are not *programmed* to model something but are *trained* to do so, where training a network is the process of adjusting the weights in response to incorrect results (Weijters and Hoppenbrouwers 1995). Schalkoff (1997) provides the following steps for training a feedforward back-propagation network:

1. Initialize all unit weights in the network.
2. Apply an input (stimulus) vector to the network.
3. Feed forward or propagate the input vector to determine all unit outputs.
4. Compare unit responses in the output layer with the desired or target response.
5. Compute and propagate an error sensitivity measure backward (starting at the output layer) through the network, using this as the basis for weight correction.
6. Minimize the overall error at each stage through unit weight adjustments.

In other words, the network is initialized and run forward to obtain an output. The error of the output relative to the data is computed, and if that error is not acceptable, the linking weights are adjusted by starting with the last layer and working towards the input layer. The network is run again, and the procedure is repeated until an acceptable error is attained or a maximum number of iterations has been reached.

The hidden and output layers have a *transfer function* which governs the input-output relationship for the layer. Common transfer functions for the hidden layers are the hyperbolic tangent sigmoid transfer function (TANSIG)

$$TANSIG(n) = \frac{2}{1+\exp(-2n)} - 1 \tag{7.7}$$

and the log sigmoid transfer derivative function (LOGSIG)

$$LOGSIG(n) = \frac{1}{1+\exp(-n)} \tag{7.8}$$

where n is the input to the hidden layer. The TANSIG function compresses any input into the range [-1 1] while the LOGSIG function compresses any input into the range [0 1]. A commonly used transfer function for the output layer is the linear transfer derivative function (PURELIN), which, unlike TANSIG and LOGSIG does not restrict output to a range but rather allows the outputs to take on any value (Demuth and Beale 2002; Ying et al. 2007; Postma 1995).

The ANNs used in this study were built in Matlab (2011a) using the "nftool" in the Neural Network (NN) Toolbox, and a visual example of the ANN as constructed in Matlab can be seen in Figure 7.4. In accordance with the N:2N:1 recommendation, the networks were built with a 2:4:1 architecture, where the HSPF and SWAT model outputs were the two network

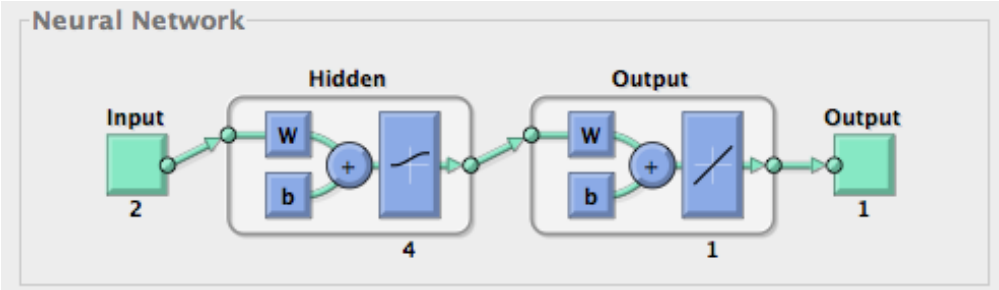


Figure 7.4: Neural Network Portrayed in Matlab

inputs. A 2:4:1 architecture yields a total of 17 weights and biases that must be calibrated in the training process ('w' and 'b' in Figure 7.4): four weights from each input to each node in the hidden layer (8), one bias to each node in the hidden layer (4), a weight from each node in the hidden layer to the output layer (4), and a bias to the output layer (1). The model outputs were log-transformed before being used as inputs to the ANNs, because it was found through trial and error that the non-transformed versions frequently resulted in imaginary numbers. The networks used the LOGSIG transfer function in the input layer and

the PURELIN transfer function in the output layer. Convergence of the training occurred when the default minimum gradient was met or after 1,000 iterations. ANN studies typically divide a data set into portions to use in training,validating, and testing the network, where the assignment of data into these sections is done randomly and usually varies between runs of the same network (e.g. Yu et al. 2004). For example, the default data distribution settings in the Matlab NN tool is 70%, 15%, 15% for training, validating, and testing, respectively. In this study, however, since the data limitations were already significant and therefore difficult to further divide, and since the desired output was the calibrated/trained network and not the predictions of a trained and validated model, all of the FC data was used in the training of the ANN. Although the randomness due to the division of the data was removed from this analysis, the use of the ANNs still included a random element. The randomness comes from the common and efficient Nguyen-Widrow initialization algorithm, used as a default to initialize the weights and biases at the beginning of each network run (Step 1 above) (Demuth and Beale 2002).

Results and Discussion

Since each network was initialized differently, ten runs were made per catchment and hydrologic state, and the results of the three best networks are presented in Table 7.4. The three best networks were selected as the three highest NSE_{ln} values with acceptable *p*-values for normality of residuals. The approach implemented here was similar to that used by Ying et al. (2007).

The best network of all the considered scenarios from Table 7.4 is shown in Table 7.5 with the corresponding HSPF and SWAT models. The results shown in Table 7.5 reveal that in all cases the ANNs greatly improved the input HSPF and SWAT estimates of NSE_{ln}. The ANNs were also able to improve the multi-model estimate in all cases derived from the

Table 7.4: Summary of Neural Networks

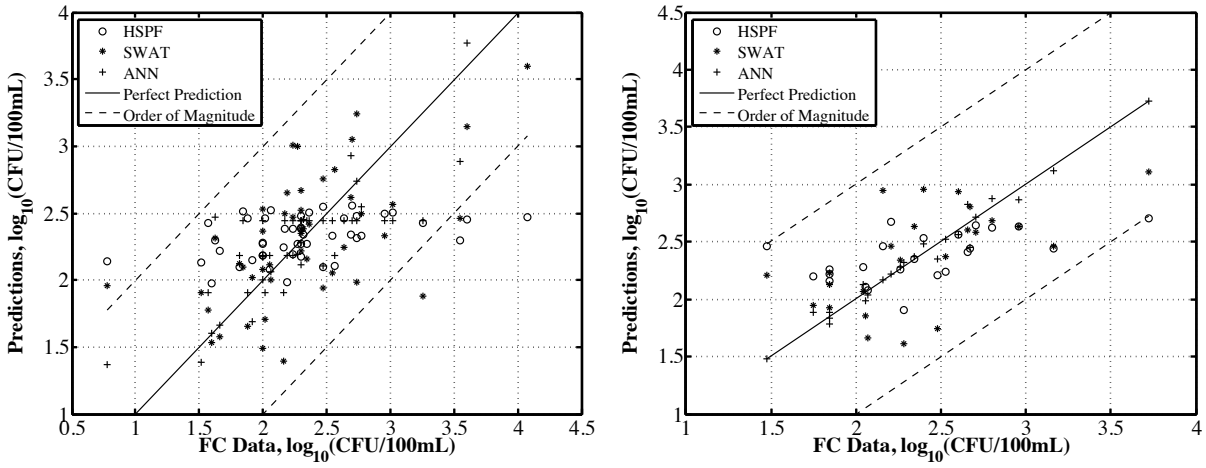
Catchment	Trial	NSE_{ln}
I	1	0.809
	2	0.734
	3	0.615
J	1	0.722
	2	0.741
	3	0.746
K	1	0.718
	2	0.681
	3	0.714
O	1	0.763
	2	0.764
	3	0.762
I Wet	1	0.874
	2	0.900
	3	0.884
J Wet	1	0.879
	2	0.882
	3	0.878
K Wet	1	0.931
	2	0.934
	3	0.935
O Wet	1	0.931
	2	0.934
	3	0.935
I Dry	1	0.867
	2	0.772
	3	0.848
J Dry	1	0.838
	2	0.565
	3	0.695
K Dry	1	0.947
	2	0.973
	3	0.935
O Dry	1	0.945
	2	0.973
	3	0.935

Table 7.5: HSPF, SWAT, and ANN Summary

Catchment	Model	NSE_{ln}
I	HSPF	0.243
	SWAT	0.275
	ANN	0.809
J	HSPF	0.263
	SWAT	-1.67
	ANN	0.746
K	HSPF	0.387
	SWAT	0.334
	ANN	0.718
O	HSPF	0.342
	SWAT	-0.028
	ANN	0.764
I Wet	HSPF	0.313
	SWAT	0.588
	ANN	0.900
J Wet	HSPF	0.370
	SWAT	-1.306
	ANN	0.882
K Wet	HSPF	0.438
	SWAT	0.592
	ANN	0.935
O Wet	HSPF	0.484
	SWAT	0.066
	ANN	0.935
I Dry	HSPF	0.120
	SWAT	-0.432
	ANN	0.867
J Dry	HSPF	-0.436
	SWAT	-2.84
	ANN	0.838
K Dry	HSPF	0.295
	SWAT	0.260
	ANN	0.973
O Dry	HSPF	-0.168
	SWAT	0.152
	ANN	0.973

NSE_{ln} optimization (Table 7.3). The most significant improvement is in both the wet and dry states, where NSE_{ln} values above 0.9 are commonly achieved.

The HSPF, SWAT, and ANN model predictions are compared with corresponding measurements for Catchment I (ANN NSE_{ln} = 0.809) and the dry model in Catchment K (ANN NSE_{ln} = 0.973) in Figure 7.5, where the abscissa is the log-scaled FC data and the ordinate is the log-scaled model predictions. The solid line in the figures has a slope of 1



(a) Catchment I (b) Catchment K-Dry

Figure 7.5: HSPF, SWAT, and ANN model results- log-scaled predictions

and represents a perfect model prediction whereas the dashed line indicates order-of-magnitude agreement. It is obvious in the figure that the ANN predictions are much closer to the line of perfect prediction, which is a reflection of the higher NSE_{ln}. This is particularly true in Figure 7.5b, where the dry model of Catchment K has an NSE_{ln} of 0.973. It is also interesting to note that while the ANN predictions tend to over-predict at low values and under-predict at high values (better seen in Fig. 7.5a), a behavior derived from the input HSPF and SWAT models, there is no consistent or predictable pattern regarding the magnitude of the ANN model residual relative to the magnitudes of the HSPF and SWAT model residuals. This is in sharp contrast with the NSE_{ln} optimization method as

shown in Figure 7.3, where the multi-model predictions are always between the HSPF and SWAT predictions.

A commonly used metric for evaluating the efficiency of models is the Akaike Information Criterion (AIC), described by (Akaike 1974):

$$AIC = -2\log ML + 2P \tag{7.9}$$

where ML is the value of the maximum likelihood attained by a model and P is the number of independently adjusted parameters within a model. Coincidentally, the HSPF (11 hydrology and 6 water quality) and SWAT (10 hydrology and 7 water quality) models considered here both have a total of 17 parameters. Therefore, since P in Equation 7.9 is the same in both models, the AIC then only reflects differences in the maximum likelihood attained by the models. Since the NSE_{ln} serves as a similar reflection on the maximum likelihood attained by the models, the inclusion of the AIC metric is redundant and will not be considered further. Comparing the ANN model performance with HSPF and SWAT, however, is not a direct comparison since an ANN is dependent on HSPF and SWAT as an input and not an independent watershed model in itself. Nonetheless, the AIC of the ANNs also reduces to a reflection of the maximum likelihood since the ANNs considered have a 2:4:1 architecture with 17 parameters. Therefore, the efficiency of the models considered throughout this research is best reflected by the NSE_{ln} value.

7.3.3 Conclusions: Model Synthesis

The NSE_{ln} optimization and ANN model combination methods both showed significant capacity to provide an improved set of model predictions. However, there are two major drawbacks to implementing these methods. First, in order to use a multi-model approach the HSPF and SWAT models must be developed before being used as inputs for the model

combination technique. Building models like HSPF and SWAT is a time-consuming process requiring expertise and significant amounts of data, and it is generally rare that both models are developed for a single project or application; the more common approach is to adopt one of the models and focus solely on it. A model combination approach as presented here may therefore be irrelevant to most modeling studies, simply because only one model is used. The parameter *a*, however, shows that only using one model is not ideal and may be misleading, since a well-developed and optimized model may still underperform a different model in any given catchment. The second drawback is that the model combination methods simply optimize a combination of the model outputs without actually changing or improving the models. This is particularly true with the ANNs, which are black-box models that train themselves to better fit the data instead of actually replicating natural processes. Despite these drawbacks, however, the model-combination methods did achieve higher NSE_{ln} values, which at the very least will provide better predictions to fit the data. Moreover, it is possible that with further analysis of the combined model predictions, particularly in case of the ANNs, which actually vary the predictions at each step without taking a weighted average of the input models like the NSE_{ln} optimization method, that new insight can be gained into the actual processes occurring.

Chapter 8

Executive Summary

8.1 SUMMARY

The objective of this research was to examine the capacity of watershed-scale models in the modeling of pathogen indicator bacteria. Pathogen indicator bacteria such as fecal coliforms are the leading pollutant in U.S. waters and are of serious concern because high levels of the bacteria indicate an increased risk of human exposure to serious, costly, and possibly fatal illnesses. The main problems encountered in addressing FC in the environment are that there are many sources, some of which are of a highly random and unpredictable nature, and that the lifecycle processes of bacteria such as transport, persistance, and die-off depend on a wide array of localized environmental factors and may vary significantly between different bacteria strains. Deriving conclusions from these observations for universal application is difficult to do, since the micro-scale processes that govern bacteria fate and transport are so highly localized. Nonetheless, government authorities are and must continue to be active in drafting and enforcing environmental policies addressing land use and water quality, for the safety of the communities under their leadership. Watershed-scale models are powerful analysis tools to aid in such management decisions.

Watershed models, however, are not perfect and the model's structure, parameters, and input data all contribute to overall model prediction uncertainty. Much of the model uncertainty can be attributed to *scale*, where the time scale of the data or the length scale of the process equations must be extrapolated to other scales to be used in the model. Complex bacteria models have been developed that incorporate the leading research ideas, but as models become more complex the calibration of the models is more complex and the structural, parameter, and data components of the uncertainty increase.

It is in this context- model predictions with uncertainty addressing the nation's leading water-quality problem- that the research in this dissertation was conducted, to examine the current capacity of watershed-scale modeling of bacteria and to present new methods for improving the modeling of bacteria. The research was divided into three substudies, and conclusions of the three substudies are summarized below.

8.2 CONCLUSIONS

1. The research in Chapter 5 showed that the hydrology components of HSPF and SWAT can be considered satisfactory for modeling daily-averaged flow while the water-quality components of both models can be considered unsatisfactory for modeling FC. The most sensitive parameters were those with a direct influence on in-stream processes like the parameters describing bacteria die-off and background sources. Parameters describing terrestrial processes were not as sensitive and were sometimes insensitive. Moreover, process equations involving the insensitive parameters likely contributed uncertainty to the model predictions. A Latin Hypercube analysis showed that significant parameter uncertainty occurs in HSPF and SWAT, but that total model uncertainty is dominated by structural uncertainty. The only way to reduce the structural uncertainty then is to change the structure of the models. While the CC model is structurally more simple than HSPF and

SWAT it performed comparably to the other two models and should be considered a valid option in future bacteria modeling studies.

2. The research in Chapter 6 presented a new method of examining HSPF and SWAT's capacity to model bacteria based on the hydrologic state of the watershed. A wet state for high flow and a dry state for low flow were defined based on the flow duration curve and the water-quality models were re-calibrated based on the corresponding wet and dry FC datasets. The wet water-quality models achieved notably better fits to the data than the dry models. As in Chapter 5, the in-stream parameters were most sensitive, but the degree of sensitivity was heightened in the dry state. The wet/dry modeling analysis confirms the increase in concentrations when streamflow increases. By considering the actual flow on the day when data was collected the wet/dry analysis presented here is a more accurate representation of seasonal affects than considerations based on months of a calendar.

3. The research in Chapter 7 presented a framework for combining a 3-parameter CC model water-quality component with the hydrology components of HSPF and SWAT. The integrated CC models performed comparably in HSPF and made notable improvements in some catchments in SWAT. Analyzing the integrated CC models revealed the distribution of flow volumes in the models and showed how each model distributed the bacteria loading between surface runoff, interflow, baseflow, and background sources.

Also in Chapter 7, the multi-model approaches showed how outputs of models like HSPF and SWAT can be combined to achieve a better fit to the data. The NSE_{ln} optimization method is a simple weighting scheme between the two models that provided an improvement over the individual models, especially in cases where the models performed similarly on an individual basis. The NSE_{ln} optimization method also yielded the model-weighting parameter a, a simple evaluation of which model performed better in any given case. The values of a for all the modeling scenarios in this dissertation (Table 7.3) reveal that HSPF and SWAT outperform each other in different scenarios and caution

against a single-model approach, since the optimization of one model may still underperform another model in any given case. The Artificial Neural Networks provided an advanced method of combining models and overall achieved notable improvements over the individual models. Drawbacks to the multi-model approach are that at least two models like HSPF and SWAT must first be available and that the multi-model methods are simply black-box methods that optimize a combination of the input and do not reveal or improve any physical processes.

8.3 RECOMMENDATIONS

1. **Continue Using the Multi-Model Approach.** The results presented in this dissertation clearly show that a multi-model approach, while more time consuming, is a worthwhile endeavor. An analysis of the results of multiple models reveals the fundamental model prediction capacity relative to other models, may prevent the outright implementation of results from an underperforming model, and facilitates the use of the model combination techniques to maximize the NSE_{ln} as presented in Chapter 7.

2. **Develop Improved Datasets.** The FC data set used in this study had approximately monthly FC measurements and was likely a limiting factor in this research. Ideally, bacteria measurements would be taken on a more frequent weekly or daily basis, and information would be simultaneously gathered describing time of day, baseflow index, and and the weather status at the time of collection. An enhanced data set, accompanied by a thorough characterization of the watershed with regard to bacteria sources, would provide valuable information regarding the bacteria concentrations and the flow. Such data could be used in the HSPF and SWAT models to further identify sensitive parameters and refine process equations involving insensitive parameters as shown in Chapter 5; in a wet/dry analysis to reveal more about the model process equations and sensitive parameters as proposed in Chapter 6; and in the CC model to further quantify bacteria concentrations from runoff and

baseflow sources as proposed in Chapter 5. Detailed data sets are expensive and rare. However, developing a detailed data set with the above considerations in mind would greatly improve modeling capacity and provide better tools for addressing in-stream bacteria pollution.

3. **Review Process Equations.** It was shown in HSPF and SWAT that certain parameters affect model output much more than other parameters (Chapters 5 and 6). The process equations for all parameters in these complex models should be re-examined while asking questions such as: What processes are being described by the insensitive parameters? Can those equations be improved or should they be removed from the model? What are the processes described by the sensitive parameters and how can they be improved? Moreover, the Latin Hypercube Analysis (Chapter 5) showed that the structural uncertainty of the models is the dominant source of uncertainty and that the only way to improve the models is through changing the model structure. One possibility is that HSPF and SWAT structures can be refined via simplified equations and reduced parameters without reducing performance. Or, a complete overhaul of the model structure may be required.

4. **Increase Use of the CC Model.** The CC model is a new and attractive option for watershed-scale bacteria modeling. It is easy to use and does not require the expertise needed for models like HSPF and SWAT. Two particular advantages of the CC model are that it uses streamflow data instead of a modeled hydrology component and that it is conceptually simple and only requires calibration of two parameters. While in some cases simpler process equations and reduced parameterizations limit model performance, the CC model consistently achieves comparable and even better water-quality results than HSPF and SWAT. The CC model may even be an attractive option for use in a multi-model technique so that only one of the HSPF or SWAT models would be required. Regardless of the application, the CC model warrants further examination in the watershed-scale modeling of bacteria.

Bibliography

Abbaspour, K. C., J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, and R. Srinivasan (2007). Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology 333*, 413–430.

Aitken, M. (2003). Impact of agricultural practices and river catchment characteristics on river and bathing water quality. *Water Science and Technology 48*(10), 217–224.

Akaike, H. (1974). A new look at the Statistical Model Identification. *IEEE Trans. Automat. Control 19*, 716–723.

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

Arnold, J., P. Allen, M. Volk, J. Williams, and D. Bosch (2010). Assessment of different representations of spatial variability on SWAT model performance. *Transactions of the American Society of Agricultural and Biological Engineers 53*(5), 1433–1443.

Auer, M. T. and S. L. Niehaus (1993). Modeling fecal coliform bacteria- I. Field and laboratory determination of loss kinetics. *Water Research 27*(4), 693–701.

Baffaut, C. and V. Benson (2009). Modeling flow and pollutant transport in a karst watershed with SWAT. *Transactions of the American Society of Agricultural and Biological Engineers 52*(2), 469–479.

Baffaut, C. and A. Sadeghi (2010). Bacteria modeling with SWAT for assessment and remediation studies: A review. *Transactions of the American Society of Agricultural and Biological Engineers 53*(5), 1585–1594.

Bai, S. and W.-S. Lung (2005). Modeling sediment impact on the transport of fecal bacteria. *Water Research 39*, 5232–5240.

Benham, B., C. Baffaut, R. Zeckoski, K. Mankin, Y. Pachepsky, A. Sadeghi, K. Brannan, M. Soupir, and M. Habersack (2006). Modeling bacteria fate and transport in watersheds to support TMDLs. *Transactions of the American Society of Agricultural and Biological Engineers 49*(4), 987–1002.

Beven, K. (2006). 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., P. Smith, and J. Freer (2007). Comment on 'Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology' by Pietro Mantovan and Ezio Todini. *Journal of Hydrology 338*, 315–318.

Bicknell, B., J. Imhoff, J. Kittle Jr., and A. Donigian Jr. (2001). Hydrological Simulation Program - Fortran (HSPF): User's manual for release 12. *Athens, Georgia : National Exposure Research Laboratory, Office of Research and Development, U.S Environmental Protection Agency*.

Bitton, G. (2005). *Wastewater microbiology, 3rd Ed.* Hoboken, NJ: Wiley-Liss.

Black, P. E. (1991). *Watershed hydrology*. Englewood Cliffs, NJ: Prentice Hall.

Blasone, R.-S., J. A. Vrugt, H. Madsen, D. Rosbjerg, B. A. Robinson, and G. A. Zyvoloski (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources 31*, 630–648.

Bolster, C., B. Haznedaroglu, and S. Walker (2009). Diversity in cell properties and transport behavior among 12 different environmental *Escherichia coli* isolates. *Journal of Environmental Quality 38*, 465–472.

Bonjoch, X., C. García-Aljaro, and A. Blanch (2011). Persistence and diversity of faecal coliform and enterococci populations in faecally polluted waters. *Journal of Applied Microbiology 111*, 209–215.

Bosch, D., J. Arnold, M. Volk, and P. Allen (2010). Simulation of a low-gradient coastal plain watershed using the SWAT landscape model. *Transactions of the American Society of Agricultural and Biological Engineers 53*(5), 1445–1456.

Bosch, D., T. Potter, C. Truman, C. Bednarz, and T. Strickland (2005). Surface runoff and lateral subsurface flow as a response to conservation tillage and soil-water conditions. *Transactions of the American Society of Agricultural Engineers 48*(6), 2137–2144.

Bosch, D. and J. Sheridan (2007, doi:10.1029/2006WR005833). Stream discharge database, Little River Experimental Watershed, Georgia, United States. *Water Resources Research 43*(W09473).

Bosch, D., J. Sheridan, R. Lowrance, R. Hubbard, T. Strickland, G. Feyereisen, and D. Sullivan (2007, doi:10.1029/2006WR005844). Little River Experimental Watershed database. *Water Resources Research 43*(W09470).

Bosch, D., J. Sheridan, and L. Marshall (2007, doi:10.1029/2006WR005834). Precipitation, soil moisture, and climate database, Little River Experimental Watershed, Georgia, United States. *Water Resources Research 43*(W09472).

Bougeard, M., J.-C. Le Saux, N. Pérenne, C. Baffaut, M. Robin, and M. Pommepuy (2011). Modeling of *Eschericia Coli* fluxes on a catchment and the impact on coastal water and shellfish quality. *Journal of the American Water Resources Association 47*(2), 350–366.

Brion, G. and S. Lingireddy (2003). Artificial neural network modelling: A summary of successful applications relative to microbial water quality. *Water Science and Technology 47*(3), 235–240.

Brion, G. M. and S. Lingireddy (1999). A neural network approach to identifying non-point sources of microbial contamination. *Water Research 33*(14), 3099–3106.

Brion, G. M., T. Neelakantan, and S. Lingireddy (2002). A neural-network based classification scheme for sorting sources and ages of fecal contamination in water. *Water Research 36*, 3765–3774.

Byappanahalli, M., M. Fowler, D. Shively, and R. Whitman (2003). Ubiquity and persistence of *Escherichia coli* in a midwestern coastal stream. *Applied and Environmental Microbiology 69*(8), 4549–4555.

Chin, D., D. Sakura-Lemessy, D. Bosch, and P. Gay (2009). Watershed-scale fate and transport of bacteria. *Transactions of the American Society of Agricultural and Biological Engineers 52*(1), 145–154.

Chin, D. A. (2006). *Water-resources engineering, 2nd Ed.* Upper Saddle River, NJ 07458: Pearson Prentice Hall.

Chin, D. A. (2009). Predictive uncertainty in water-quality modeling. *Journal of Environmental Engineering 135*(12), 1315–1325.

Chin, D. A. (2011). Quantifying pathogen sources in streams by hydrograph separation. *Journal of Environmental Engineering 137*(9), 770–781.

Chin, D. A. (20XXb). Identification of sources of bacteria in streams for water-quality control. *Wat. Environ. Res.*. Accepted for publication.

Cho, J., R. R. Lowrance, D. D. Bosch, T. C. Strickland, Y. Her, and G. Vellidis (2010). Effect of watershed subdivision and filter width on SWAT simulation of a coastal plain watershed. *Journal of the American Water Resources Association 46*(3), 586–602.

Christiaens, K. and J. Feyen (2002). Use of sensitivity and uncertainty measures in distributed hydrological modeling with an application to the MIKE SHE model. *Water Resources Research 38*(9), 1169. doi:10.1029/2001WR000478.

Coffey, R., E. Cummins, N. Bhreathnach, V. Flaherty, and M. Cormican (2010). Development of a pathogen transport model for Irish catchments using SWAT. *Agricultural Water Management 97*, 101–111.

Corso, P. S., M. H. Kramer, K. A. Blair, D. G. Addiss, J. P. Davis, and A. C. Haddix (2003, April). Cost of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee, Wisconsin. *Emer. Inf. Dis. 9*(4).

Crowther, J., D. Kay, and M. D. Wyer (2002). Faecal-indicator concentrations in waters draining lowland pastoral catchments in the UK: relationships with land use and farming practices. *Water Research 36*(7), 1725–1734.

Davies-Colley, R., E. Lydiard, and J. Nagels (2008). Stormflow-dominated loads of faecal pollution from an intensively dairy-farmed catchment. *Water Science and Technology 57*(10), 1519–1523.

Demuth, H. and M. Beale (2002). Matlab neural network toolbox user's guide, version 4. Technical report, The MathWorks.

Doherty, J. and J. M. Johnston (2003). Methodologies for calibration and predictive analysis of a watershed model. *Journal of the American Water Resources Association 39*(2), 251–265.

Duan, Q., S. Sorooshian, and V. Gupta (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research 28*(4), 1015–1031.

Eckhardt, K., L. Breuer, and H.-G. Frede (2003). Parameter uncertainty and the significance of simulated land use change effects. *Journal of Hydrology 273*, 164–176.

Entry, J. A., R. K. Hubbard, J. E. Thies, and J. J. Fuhrmann (2000a). The influence of vegetation in riparian filterstrips on coliform bacteria: I. Movement and survival in water. *Journal of Environmental Quality 29*(4), 1206–1214.

Entry, J. A., R. K. Hubbard, J. E. Thies, and J. J. Fuhrmann (2000b). The influence of vegetation in riparian filterstrips on coliform bacteria: II. Movement and survival in water. *Journal of Environmental Quality 29*(4), 1215–1224.

Feyereisen, G., R. Lowrance, T. Strickland, J. Sheridan, R. Hubbard, and D. Bosch (2007b, doi:10.1029/2006WR005835). Long-term water chemistry database, Little River Experimental Watershed, southeast Coastal Plain, United States. *Water Resources Research 43*(W09474).

Feyereisen, G., T. Strickland, D. Bosch, and D. Sullivan (2007). Evaluation of SWAT manual calibration and input parameter sensitivity in the Little River Watershed. *Transactions of the American Society of Agricultural and Biological Engineers 50*(3), 843–855.

Feyereisen, G., T. Strickland, D. Bosch, C. Truman, J. Sheridan, and T. Potter (2008). Curve number estimates for conventional and conservation tillages in the southeastern Coastal Plain. *Journal of Soil and Water Conservation 63*(3), 120–128.

Fiener, P. and K. Auerswald (2003, May/Jun). Effectiveness of grassed waterways in reducing runoff and sediment delivery from agricultural watersheds. *Journal of Environmental Quality 32*(3), 927–936.

Flores, A. N., D. Entekhabi, and R. L. Bras (2010). Reproducibility of soil moisture ensembles when representing soil parameter uncertainty using a Latin Hypercube-based approach with correlation control. *Water Resources Research 46*.

Franks, S. W. and K. J. Beven (1997). Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. *Journal of Geophysical Research 102*(D20), 23991–23999.

GA DNREP (2005, November). Rules and regulations for water quality control. Chapter 391-3-6, Georgia Department of Natural Resources Environmental Protection Division, Atlanta, GA.

Gagliardi, J. V. and J. S. Karns (2000). Leaching of *Escherichia coli* O157:H7 in diverse soils under various agricultural management preferences. *Applied and Environmental Microbiology 66*(3), 877–883.

Glavan, M., S. White, and I. P. Holman (2011). Evaluation of river water quality simulations at a daily time step – Experience with SWAT in the Axe Catchment, UK. *Clean- Soil, Air, Water 39*(1), 43–54.

Goss, M. and C. Richards (2008). Development of a risk-based index for source water protection planning, which supports the reduction of pathogens from agricultural activity entering water resources. *Journal of Environmental Management 87*, 623–632.

Griensven, A. V., T. Meixner, S. Grunwald, T. Bishop, M. Diluzio, and R. Srinivasan (2006). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology 324*, 10–23.

Guber, A. K., D. R. Shelton, Y. A. Pachepsky, A. M. Sadeghi, and L. J. Sikora (2006). Rainfall-induced release of fecal coliforms and other manure constituents: Comparison and modeling. *Applied and Environmental Microbiology 72*(12), 7531–7539.

Gupta, H. V. and H. Kling (2011). On typical range, sensitivity, and normalization of mean squared error and Nash-Sutcliffe efficiency type metrics. *Water Resources Research 47*(W10601).

Gupta, R. S. (2001). *Hydrology and hydraulic systems*. Prospect Heights, IL: Waveland Press.

Haan, C. (1989). Parametric uncertainty in hydrologic modeling. *Transactions of the American Society of Agricultural Engineers 32*(1), 137–146.

Hantush, M. M. and L. Kalin (2008). Stochastic residual-error analysis for estimating hydrologic model predictive uncertainty. *Journal of Hydrologic Engineering 13*(7), 585–596.

Harmel, R., R. Karthikeyan, T. Gentry, and R. Srinivasan (2010). Effects of agricultural management, land use, and watershed scale on *E. coli* concentrations in runoff and streamflow. *Transactions of the American Society of Agricultural and Biological Engineers 53*(6), 1833–1841.

Haydon, S. and A. Deletic (2006). Development of a coupled pathogen-hydrologic catchment model. *Journal of Hydrology 328*, 467–480.

Haydon, S. and A. Deletic (2009). Model output uncertainty of a coupled pathogen indicator-hydrologic catchment model due to input data uncertainty. *Environmental Modeling & Software 24*, 322–328.

Hooda, P., A. Edwards, H. Anderson, and A. Miller (2000). A review of water quality concerns in livestock farming areas. *Science of the Total Environment 250*, 143–167.

Hoxie, N. J., J. P. Davis, J. M. Vergeront, R. D. Nashold, and K. A. Blair (1997). Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am. J. Pub. Hea. 87*(12).

Hunter, C. and A. McDonald (1991, October). The occurrence of coliform bacteria in the surface soils of two catchment areas in the Yorkshire Dales. *J. Inst. Wat. Environ. Manage. 5*, 534–538.

Hunter, C., A. McDonald, and K. Beven (1992). Input of fecal coliform bacteria to an upland stream channel in the Yorkshire Dales. *Water Resources Research 28*(7), 1869–1876.

Iman, R. L. and W. Conover (1982). A distribution-free approach to inducing rank correlation among input variables. *Comm. Stat. 11*(3), 311–334.

Ishii, S., W. B. Ksoll, R. E. Hicks, and M. J. Sadowsky (2006). Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Applied and Environmental Microbiology 72*(1), 612–621.

Jamieson, R., R. Gordon, D. Joy, and H. Lee (2004). Assessing microbial pollution of rural surface waters: A review of current watershed scale modeling approaches. *Agricultural Water Management 70*, 1–17.

Jamieson, R., D. Joy, H. Lee, R. Kostaschuk, and R. Gordon (2004, doi: 10.1139/S04-001). Persistence of enteric bacteria in alluvial streams. *Journal of Environmental Engineering Science 3*, 203–212.

Jamieson, R., D. M. Joy, H. Lee, R. Kostaschuk, and R. Gordon (2005). Transport and deposition of sediment-associated *Escherichia coli* in natural systems. *Water Research 39*, 2665–2675.

Jeong, J., N. Kannan, J. Arnold, R. Glick, L. Gosselink, and R. Srinivasan (2010). Development and integration of sub-hourly rainfall-runoff modeling capability within a watershed model. *Water Resources Management 24*, 4505–4527.

Jia, Y. and T. B. Culver (2008). Uncertainty analysis for watershed modeling using Generalized Likelihood Uncertainty Estimation with multiple calibration measures. *Journal of Water Resources Planning and Management 134*(2), 97–106.

Journel, A. (1997). The abuse of principles in model building and the quest for objectivity. In E. Baafi and N. Schofield (Eds.), *Geostatistics Wollongong '96*. The Netherlands: Kluwer Academic Publishers.

Kashefipour, S., B. Lin, and R. Falconer (2006). Modelling the fate of faecal indicators in a coastal basin. *Water Research 40*, 1413–1425.

Kros, J., E. Pebesma, G. Reinds, and P. Finke (1999). Uncertainty assessment in modelling soil acidification at the European scale: A case study. *Journal of Environmental Quality 28*(2), 366–377.

Laroche, A.-M., J. Gallichand, R. Lagacé, and A. Pesant (1996). Simulating atrazine transport with HSPF in an agricultural watershed. *Journal of Environmental Engineering 122*(7), 622–630.

LaWare, P. and H. S. Rifai (2006). Modeling fecal coliform contamination in the Rio Grande. *Journal of the American Water Resources Association 42*(2), 337–356.

Lee, S.-B., C.-G. Yoon, K. W. Jung, and H. S. Hwang (2010). Comparative evaluation of runoff and water quality using HSPF and SWMM. *Water Science and Technology 62*(6), 1401–1409.

Legates, D. R. and G. J. McCabe (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research 35*(1), 233–241.

Ling, T., E. Achberger, C. Drapcho, and R. Bentson (2002). Quantifying adsorption of an indicator bacteria in a soil-water system. *Transactions of the American Society of Agricultural Engineers 45*(3), 669–674.

Liu, L., M. S. Phanikumar, S. L. Molloy, R. L. Whiteman, D. A. Shively, M. B. Nevers, D. J. Schwab, and J. B. Rose (2006). Modeling the transport and inactivation of *E. coli* and Enterococci in the near-shore region of Lake Michigan. *Environmental Science and Technology 40*(16), 5022–5028.

Lowe, S. (2004). Estimation of input parameters in the HSPF watershed model. *Journal of Engineering Technology 1*(Spr), 16–20.

Macler, B. A. and J. C. Merkle (2000). Current knowledge on groundwater microbial pathogens and their control. *Hydrogeology Journal 8*, 29–40.

Mantovan, P. and E. Todini (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology 330*, 368–381.

Mantovan, P., E. Todini, and M. Martina (2007). Reply to comment by Keith Beven, Paul Smith and Jim Freer on 'Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology'. *Journal of Hydrology 338*, 319–324.

McDonald, A., D. Kay, and A. Jenkins (1982, August). Generation of fecal and total coliform surges by stream flow manipulation in the absence of normal hydrometeorological stimuli. *Applied and Environmental Microbiology 44*(2), 292–300.

McKay, M., W. Conover, and R. Beckman (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics 21*, 239–245.

McMurry, S., M. Coyne, and E. Perfect (1998). Fecal coliform transport through intact soil blocks amended with poultry manure. *Journal of Environmental Quality 27*, 86–92.

Meays, C. L., K. Broersma, R. Nordin, A. Mazumder, and M. Samadpour (2006). Spatial and annual variability in concentrations and sources of *Escherichia coli* in multiple watersheds. *Environmental Science and Technology 40*, 5289–5296.

Melching, C. S. and W. Bauwens (2001). Uncertainty in coupled nonpoint source and stream water-quality models. *Journal of Water Resources Planning and Management 127*(6), 403–413.

Metcalf & Eddy. Revised by: George Tchobanoglous, Franklin L. Burton, and David Stensel (2003). *Wastewater engineering: Treatment and reuse, 4th edition*. New York: Mcgraw Hill.

Moriasi, D., J. Arnold, M. Van Liew, R. Bingner, R. Harmel, and T. Veith (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers 50*(3), 885–900.

Muirhead, R., R. Collins, and P. Bremer (2006a). The association of *E. coli* and soil particles in overland flow. *Water Science and Technology 54*(3), 153–159.

Muirhead, R. W., R. P. Collins, and P. J. Bremer (2006b). Interaction of *Escherichia coli* and soil particles in runoff. *Applied and Environmental Microbiology 72*(5), 3406–3411.

Nash, J. and J. Sutcliffe (1970). River flow forecasting through conceptual models: Part 1. A discussion of principles. *Journal of Hydrology 10*(3), 282–290.

Neelakantan, T., G. Brion, and S. Lingireddy (2001). Neural network modelling of *Cryptosporidium* and *Giardia* concentrations in the Delaware River, USA. *Water Science and Technology 43*(12), 125–132.

Neitsch, S., J. Arnold, J. Kiniry, and J. Williams (2005). Soil and Water Assessment Tool theoretical documentation, version 2005. *Temple, Texas : U.S. Department of Agriculture, Agriculture Research Service, Grassland, Soil, and Water Research Laboratory*.

Obiri-Danso, K. and K. Jones (1999). Distribution and seasonality of microbial indicators and thermophilic campylobacters in two freshwater bathing sites on the River Lune in northwest England. *Journal of Applied Microbiology 87*, 822–832.

Oliver, D. M., A. L. Heathwaite, R. D. Fish, D. R. Chadwick, C. J. Hodgson, M. Winter, and A. J. Butler (2009). Scale appropriate modelling of diffuse microbial pollution from agriculture. *Progress in Physical Geography 33*(3), 358–377.

Pappas, E. A., R. S. Kanwar, J. L. Baker, J. C. Lorimor, and S. Mickelson (2008). Fecal indicator bacteria in subsurface drain water following swine manure application. *Transactions of the American Society of Agricultural and Biological Engineers 51*(5), 1567–1573.

Paul, S., P. Haan, M. Matlock, S. Mukhtar, and S. Pillai (2004). Analysis of the HSPF water quality parameter uncertainty in predicting peak in-stream fecal coliform concentrations. *Transactions of the American Society of Agricultural Engineers 47*(1), 69–78.

Pebesma, E. J. and G. B. Heuvelink (1999). Latin Hypercube sampling of Gaussian random fields. *Technometrics 41*(4), 303–312.

Postma, E. (1995). Optimisation networks. In *Artificial neural networks- An introduction to ANN theory and practice*. Springer.

Prada-Sarmiento, F. and N. Obregón-Neira (2009). Forecasting of monthly streamflows based on artificial neural networks. *Journal of Hydrologic Engineering 14*(12), 1390–1395.

Razavi, S., B. A. Tolson, L. S. Matott, N. R. Thomson, A. MacLean, and F. R. Seglenieks (2010, doi:10.1029/2009WR008957). Reducing the computational cost of automatic calibration through model preemption. *Water Resources Research 46*(W11523).

Russo, S. A., J. Hunn, and G. W. Characklis (2011). Considering bacteria-sediment associations in microbial fate and transport modeling. *Journal of Environmental Engineering 137*(8), 697–706.

Sakura-Lemessy, D.-M. G. (2009). *A multi-model approach to predicting pathogen indicator bacteria loading in TMDL analyses*. Ph. D. thesis, University of Miami, Coral Gables, FL 33124.

Sándor, Z. and P. András (2004). Alternative sampling methods for estimating multivariate normal probabilities. *J. Econ. 120*, 207–234.

Schalkoff, R. J. (1997). *Artificial neural networks*. New York, NY: McGraw-Hill.

Scott, T. M., S. Parveen, K. M. Portier, J. B. Rose, M. L. Tamplin, S. R. Farrah, A. Koo, and J. Lukasik (2003). Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. *Applied and Environmental Microbiology 69*(2), 1089–1092.

Sexton, A., A. Shirmohammadi, A. Sadeghi, and H. Montas (2011). Impact of parameter uncertainty on critical SWAT output simulations. *Transactions of the American Society of Agricultural and Biological Engineers 54*(2), 461–471.

Shapiro, S. and M. Wilk (1965). An analysis of variance test for normality (complete samples). *52*(3 and 4), 591–611.

Shen, Z., Q. Hong, H. Yu, and R. Liu (2008). Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Science of the Total Environment 405*, 195–205.

Sheridan, J. (1997). Rainfall-streamflow relations for coastal plain watersheds. *Applied Engineering in Agriculture 13*(3), 333–344.

Shirmohammadi, A., I. Chaubey, R. Harmel, D. Bosch, R. Muñoz-Carpena, C. Dharmasri, A. Sexton, M. Arabi, M. Wolfe, J. Frankenberger, C. Graff, and T. Sohrabi (2006). Uncertainty in TMDL models. *Transactions of the American Society of Agricultural and Biological Engineers 49*(4), 1033–1049.

Shirmohammadi, A., K. Yoon, and W. Magette (1997). Water quality in mixed land-use watershed– Piedmont region in Maryland. *Transactions of the American Society of Agricultural Engineers 40*(6), 1563–1572.

Singh, J., H. V. Knapp, J. Arnold, and M. Demissie (2005, April). Hydrological modeling of the Iroquois river watershed using HSPF and SWAT. *Journal of the American Water Resources Association 41*(2), 343–360.

Singh, K. P., A. Basant, A. Malik, and G. Jain (2009). Artificial neural network modeling of the river water quality-A case study. *Ecological Modeling 220*, 888–895.

Skahill, B. E., J. S. Baggett, S. Frankenstein, and C. W. Downer (2009). More efficient PEST compatible model independent model calibration. *Environmental Modeling & Software 24*, 517–529.

Skahill, B. E. and J. Doherty (2006). Efficient accommodation of local minima in watershed model calibration. *Journal of Hydrology 329*, 122–139.

Sloto, R. and M. Crouse (1996). HYSEP: A computer program for streamflow hydrograph separation and analysis. *Water-Resources Investigation Report No. 96-4040. Lemoyne, Pennsylvania : United States Geological Survey*.

Solo-Gabriele, H. M., M. A. Wolfert, T. R. Desmarais, and C. J. Palmer (2000). Sources of *Escherichia coli* in a coastal subtropical environment. *Applied and Environmental Microbiology 66*(1), 230–237.

Stein, M. (1987). Large sample properties of simulations using Latin Hypercube sampling. *Technometrics 29*(2), 143–151.

Sullivan, D., H. Batten, D. Bosch, J. Sheridan, and T. Strickland (2007, doi:10.1029/2006WR005836). Little River Experimental Watershed, Tifton, Georgia, United States: A geographic database. *Water Resources Research 43*(W09471).

Tolson, B. A. and C. A. Shoemaker (2007, doi:10.1029/2005WR004723). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research 43*(W01413).

Troutman, B. M. (1985). Errors and parameter estimation in precipitation-runoff modeling, part I: Theory. *Water Resources Research 21*(8), 1195–1213.

Tyrrel, S. and J. Quinton (2003). Overland flow transport of pathogens from agricultural land receiving faecal wastes. *Journal of Applied Microbiology 94*, 87S–93S.

Unc, A. and M. J. Goss (2004). Transport of bacteria from manure and protection of water resources. *App. Soil Ecol. 25*, 1–18.

Urban, N. M. and T. E. Fricker (2010). A comparison of Latin Hypercube and grid ensemble designs for the multivariate emulation of an Earth system model. *Comp. Geosci. 36*, 746–755.

USEPA (2001, January). Protocol for developing pathogen TMDLs. Technical Report EPA-841-R-00-002, Office of Water (4503F), U.S. Environmental Protection Agency.

USEPA (2007, August). An approach for using load duration curves in the development of TMDLs. Technical Report EPA 841-B-07-006, Office of Wetlands, Oceans and Watersheds, U.S. Environmental Protection Agency.

USEPA (2008). Causes of impairment for 303(d) listed waters.

USEPA (2009). National summary of impaired waters and TMDL information. Technical report, US Environmental Protection Agency, Washington, DC. Accessed December 2011: http://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T.

VADEQ (2004, April). Bacteria TMDL for Roses Creek Watershed, Virginia. TMDL Report, Virginia Department of Environmental Quality.

Van Liew, M., J. Arnold, and J. Garbrecht (2003). Hydrologic simulation on agricultural watersheds: Choosing between two models. *Transactions of the American Society of Agricultural and Biological Engineers 46*(6), 1539–1551.

Vellidis, G., R. Lowrance, M. Adams, and A. Milton (2010). Fecal bacteria concentrations in agricultural watersheds of the Georgia coastal plain. Unpublished report.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003, doi:10.1029/2002WR001642). A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research 39*(8).

Weaver, R., J. Entry, and A. Graves (2005). Numbers of fecal streptococci and *Escherichia coli* in fresh and dry cattle, horse, and sheep manure. *Can. J. Microbiol. 51*, 847–851.

Weibull, W. (1939, Stockholm). A statistical theory of the strength of materials. *Ing. Vet. Handl. 151*(15).

Weijters, A. and G. Hoppenbrouwers (1995). Backpropagation networks for grapheme-phoneme conversion: a non-technical introduction. In *Artificial Neural Networks- An Introduction to ANN Theory and Practice*. Springer.

White, E., G. Feyereisen, T. Veith, and D. Bosch (2009). Improving daily water yield estimates in the Little River Watershed: SWAT adjustments. *Transactions of the American Society of Agricultural and Biological Engineers 52*(1), 69–79.

Wilkinson, J., A. Jenkins, M. Wyer, and D. Kay (1995, October). Modelling faecal coliform concentrations in streams. Technical Report 127, Institute of Hydrology, Wallingford, Oxfordshire, U.K.

Wilks, S. S. (1962). *Mathematical statistics*. New York, NY: Wiley.

Yang, J., P. Reichert, K. Abbaspour, and H. Yang (2007). Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference. *Journal of Hydrology 340*, 167–182.

Yilmaz, K. K., J. A. Vrugt, H. V. Gupta, and S. Sorooshian (2010). Model calibration in watershed hydrology. In B. Sivakimar and R. Berndtsson (Eds.), *Advances in data-based approaches for hydrologic modeling and forecasting*. World Scientific.

Ying, Z., N. Jun, C. Fu-yi, and G. Liang (2007). Water quality forecast through application of BP neural network at Yuqiao reservoir. *J. Zhejiang Univ. SCI. A 8*(9), 1482–1487.

Yu, C., W. Northcott, and G. McIsaac (2004). Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds. *Transactions of the American Society of Agricultural Engineers 47*(1), 285–290.

Yu, P.-S., T.-C. Yang, and S.-J. Chen (2001). Comparison of uncertainty analysis methods for a distributed rainfall-runoff model. *Journal of Hydrology 244*, 43–59.

Zhang, X., R. Srinivasan, J. Arnold, R. C. Izaurralde, and D. Bosch (2011). Simultaneous calibration of surface flow and baseflow simulations: a revisit of the SWAT model calibration framework. *Hydrological Processes 25*, 2313–2320.