

WATERSHED-SCALE FATE AND TRANSPORT OF BACTERIA

D. A. Chin, D. Sakura-Lemessy, D. D. Bosch, P. A. Gay

ABSTRACT. *The added dimensionality provided by using multiple models to predict the fate and transport of bacteria at the watershed scale were investigated. Both HSPF and SWAT were applied to the 15.6 km² catchment K of the Little River Experimental Watershed (LREW) in Georgia. Over the seven-year period from 1996 to 2002, SWAT provided a more accurate description of fecal coliform concentrations, with an NSE of 0.73 compared to 0.33 for HSPF. For this particular watershed, the SWAT process equations are more representative of the watershed-scale fate and transport of bacteria than the HSPF process equations. Based on this comparative analysis, it can be inferred that elevated levels of fecal coliform in the receiving stream are primarily due to in-stream sources. This source characterization could not be achieved by using only the HSPF model, which indicates a much greater contribution from groundwater and terrestrial nonpoint sources. A model-averaging approach in which a weighted average of the HSPF and SWAT predictions are used to predict bacteria concentrations in the receiving stream demonstrates that model weights can be determined such that the NSE of the combined models will be greater than either of the models taken individually. However, in the present case, the marginal improvements in NSE obtained through this integration were small.*

Keywords. *Bacteria, Fecal coliform, HSPF, Hydrology, Modeling, SWAT, Watershed.*

The leading cause of impairment of rivers and streams in the U.S. is excessive levels of pathogen indicator bacteria (USEPA, 2008), and the most common indicator bacteria in freshwater bodies are fecal coliform and *Escherichia coli* (*E. coli*). Federal regulations require that all states monitor and assess public water bodies within their borders, identify those that are impaired, and determine the limiting contaminant loadings that would be required for the impaired waters to meet their applicable water-quality standards; these limiting loadings are called total maximum daily loads (TMDLs). Implementation of loading reductions to comply with adopted TMDLs generally requires identification of the likely sources of contamination and estimation of the quantitative relationship between contaminant source loadings and contaminant concentrations in receiving waters. Since contaminant sources causing impairments are generally a combination of nonpoint and point sources, watershed-scale fate and transport models are particularly useful in identifying likely contaminant sources and quantifying the relationship between source loadings and contaminant concentrations in receiving waters. The most commonly used watershed-scale fate and transport models for TMDL applications are HSPF (Bicknell et al., 2001) and

SWAT (Neitsch et al., 2005), with HSPF more commonly used in urbanized watersheds and SWAT more commonly used in agricultural watersheds, although both models can be used in most watersheds.

An essential component in utilizing watershed-scale fate and transport models is accounting for predictive uncertainty, which can be attributed to a combination of parameter uncertainty, structural uncertainty, and data uncertainty. In most models, structural uncertainty is dominant and limits the minimum predictive uncertainty that can be achieved by any given model. Since the genesis of structural uncertainty in any model is the inadequacy of the process equations incorporated in the model, the only way to reduce structural uncertainty and lower the limit of predictive uncertainty is to use a different model with more accurate process equations. A model-averaging approach (Claeskens and Hjort, 2008; Ajami et al., 2006) can be particularly useful in reducing structural uncertainty relative to the single-model approach, where the weights allocated to different models are adjusted depending on the predictive uncertainty of the individual models. The weighted-average model will likely have less predictive uncertainty than if only one model was universally used. A second benefit of the multi-model approach is that the process equations incorporated in the better-performing model are presumably a more accurate representation of the fate and transport processes in the watershed, and provide valuable insight into and quantification of the sources and movement of contaminants within the watershed.

This article demonstrates the use of the multi-model approach in which both HSPF and SWAT are applied to predict the fate and transport of indicator bacteria in a particular watershed, and the prediction results are used to identify the dominant fate and transport processes within the watershed. The reduction in predictive uncertainty using the two-model approach is also quantified, and conclusions are drawn regarding the efficacy of using the two-model approach to iden-

Submitted for review in October 2008 as manuscript number SW 7749; approved for publication by the Soil & Water Division of ASABE in January 2009.

The authors are **David A. Chin**, Professor, Department of Civil Engineering, University of Miami, Coral Gables, Florida; **Donna Sakura-Lemessy**, Assistant Professor, Department of Natural Sciences, Albany State University, Albany, Georgia; **David D. Bosch**, ASABE Member, Research Hydraulic Engineer, USDA-ARS Southeast Watershed Research Laboratory, Tifton, Georgia; and **Paige A. Gay**, Assistant Research Scientist, Department of Biological and Agricultural Engineering, University of Georgia, Tifton, Georgia. **Corresponding author:** David A. Chin, Department of Civil Engineering, 1251 Memorial Dr., University of Miami, Coral Gables, FL 33124; phone: 305-284-3391; fax: 305-284-3798; E-Mail: dchin@miami.edu.

tify source-loading reductions that would be effective in implementing TMDLs.

PROCESS EQUATIONS

Both HSPF and SWAT codes contain hydrology and water-quality components. In developing water-quality models using these codes, it is standard practice to first calibrate the hydrology component of the model by adjusting the hydrologic process parameters until model streamflow predictions agree with observed streamflows to an acceptable degree. The second step is to calibrate the water-quality component by adjusting the water-quality process parameters of the model until the model tracer concentration predictions at specified locations in streams agree with observed concentrations to an acceptable degree. The hydrologic process equations in HSPF and SWAT are fundamentally different, with the most notable external difference being that HSPF simulates the rainfall/runoff process at hourly time steps, while SWAT simulates the rainfall/runoff process at daily time steps. Many previous studies have documented the comparative hydrologic process equations and performances of HSPF and SWAT (e.g., Van Liew et al., 2003). Bacterial fate and transport processes incorporated in HSPF and SWAT codes are also fundamentally different, thereby increasing the likelihood that reduction in structural uncertainty can be achieved using a multi-model approach. A comparative analysis of the HSPF and SWAT process equations for modeling the fate and transport of bacteria can be found in Benham et al. (2006), and a brief review of these processes with some supplementary material is provided here.

In both HSPF and SWAT codes, bacteria are added to hydrologically homogeneous land segments within the watershed at a specified rate that can vary with time. In HSPF, the bacteria loading rate is specified directly ($\text{cfu ha}^{-1} \text{h}^{-1}$) and can be either constant or vary monthly, while in SWAT the loading is specified as the product of the bacteria content of manure (cfu g^{-1}) and the manure loading rate ($\text{kg ha}^{-1} \text{d}^{-1}$), which can be either constant or vary daily. In HSPF, deposited bacteria can be transported by either direct entrainment in overland flow or by association with sediment contained in overland flow, in which case a potency factor (Chin, 2006) must be specified. In SWAT, bacteria are partitioned into both soluble and sorbed phases using a linear isotherm of the form:

$$S = k_p c \quad (1)$$

where S is the sorbed bacteria density (cfu g^{-1}), k_p is the partitioning coefficient (mL g^{-1}), and c is the bacteria concentration in solution (cfu mL^{-1}) within the top 10 mm of soil.

In reality, bacterial transport is associated with both dissolution in overland flow and sorption onto sediments. However, significant uncertainty in modeling both bacterial partitioning and sediment transport has caused most models to treat bacteria transport as being entirely associated with dissolution in surface runoff (e.g., Paul et al., 2004; Jamieson et al., 2004). This approach is adopted in this article by specifying the potency factor (in HSPF) and partition coefficient (in SWAT) both equal to zero. This formulation does not neglect bacteria transport by sediment attachment, but simply assumes that such transport can be described by an effective dissolution rate. HSPF indirectly simulates the dieoff of bacteria on the land surface by limiting the amount of accu-

mulated bacteria, while SWAT explicitly simulates the dieoff of sorbed and solution bacteria on the land surface using Chick's law, which can be expressed in the form:

$$N_t = N_0 \exp(-k_d t) \quad (2)$$

where N_t is the number of bacteria at time t (cfu), N_0 is the initial number of bacteria (cfu), and k_d is a net decay constant that varies with temperature.

HSPF tracks a single type of bacteria deposited on the ground, while SWAT divides bacteria into persistent and less-persistent categories, and can also simulate the deposition and washoff of bacteria on foliage. During each time step (Δt), the amount of solution bacteria released (ΔM_r) from the soil solution into the surface runoff is given by:

$$\Delta M_r = \begin{cases} M_s [1 - \exp(-k_2 \Delta Q)] & \text{(HSPF)} \\ M_s k_1 \Delta Q & \text{(SWAT)} \end{cases}$$

where M_s is the amount of bacteria in solution in the soil-storage layer at the beginning of the time interval, ΔQ is the runoff within the time interval (cm), and k_1 and k_2 are release-rate constants (cm^{-1}).

Rainfall that infiltrates into the soil can contribute to streamflow via interflow and outflow from the shallow groundwater. However, a major difference between codes is that HSPF allows specification of constant or monthly variations in bacteria concentration in the interflow and shallow groundwater inflow to streams, while SWAT requires that subsurface bacteria concentrations be zero. Experiments on bacteria transport in soils and groundwater usually conclude that bacteria move in the subsurface, and that temporal or permanent saturation allows indicator bacteria to move quite far along preferential pathways (Benham et al., 2006). In addition to bacteria transported to streams by surface runoff, both HSPF and SWAT allow the direct input of bacteria into stream segments as either point sources or direct nonpoint sources, with HSPF allowing hourly variation of loading (cfu h^{-1}) and SWAT determining loading as the product of a specified daily source flow ($\text{m}^3 \text{d}^{-1}$) and source concentration (cfu m^{-3}). Loadings from terrestrial and direct sources are added together to give the total flux of bacteria into the receiving stream segment, and each segment is simulated as a completely mixed reactor with first-order decay.

The relative validity of the above watershed-scale fate and transport process equations can only be inferred by comparing the performance of models based on these fundamentally different process equations, and associating the more accurate process equations (for a particular watershed) with the better-performing models. Such an intercomparison within the Little River Experimental Watershed is described in this article, and the implications of these results on reducing predictive uncertainty and in developing implementation plans for TMDLs are subsequently discussed.

LITTLE RIVER EXPERIMENTAL WATERSHED

Hydrologic monitoring in the Little River Experimental Watershed (LREW) in south-central Georgia is conducted by the USDA-ARS Southeast Watershed Research Laboratory (SEWRL) in Tifton, Georgia, and the location of the LREW within the state of Georgia is shown in figure 1. The LREW

is centered at approximately 31.61° N and 83.66° W and covers an area of approximately 334 km². The LREW is instrumented to measure rainfall and streamflow within the primary watershed and the seven nested subwatersheds, as shown in figure 1, where the subwatersheds are identified by their pour points (M, K, J, I, F, N, O, B) and the subwatersheds range in size from 3 to 115 km². Overall, the LREW lies within an area of broad floodplains, river terraces, and gently sloping uplands. Most land-surface slopes are less than 5%, although some valley slopes range from 5% to 15%. Streams in the LREW have channel slopes ranging from 0.1% to 0.4% (Bosch and Sheridan, 2007). The LREW contains sandy soils underlain by limestone, and a seasonally dependent shallow phreatic aquifer exists throughout the watershed. Physiographically, the LREW is located in the Tifton Upland, and stratigraphically the subsurface consists of Quaternary sand overlaying the parent Miocene Hawthorne Formation (Assmussen, 1971). The depth of the surficial alluvium ranges from 2 m in the headwater streams to about 6 m at the lower end of the watershed, and the hydrology is typical of that found in the southern Coastal Plain, where surface materials have high infiltration rates, low surface runoff, and high groundwater inflow to streams.

There are 46 continuously recording rain gauges covering the LREW and the surrounding area, and these gauges are spaced approximately 2.4 km apart in the upper watershed and 4.8 km apart in the lower watershed (Bosch and Sheridan, 2007). Most rain gauges consist of TE525 tipping-bucket gauges (Texas Electronics, Inc., Dallas, Texas) with minimum measurement precision of 0.254 mm, accuracy of ±0.5 mm h⁻¹, and a recording interval of 5 min. The present study focuses on catchment K, which is shown in figure 1 and covers an area of approximately 16.7 km². Hourly rainfall in catchment K was calculated from 5 min measurements at rain gauge RG43 (UTM 3513276 m N, 242618 m E) located approximately in the center of catchment K. The correlation

of the upscaled daily rainfall at RG43 with the catchment-averaged daily rainfall provided by SEWRL was 0.98, which provided confidence that RG43 can be used to characterize the hourly catchment-averaged rainfall.

Other meteorological data available for catchment K include hourly solar radiation, cloud cover, air temperature, dew point, and potential evapotranspiration. Data collected at Met station 747810 located in Valdosta, Georgia, approximately 70 km from the LREW were used to describe these climatic variables. Although this distance is not ideal, it was deemed tolerable since these latter climatic variables are not expected to vary significantly over this length scale and are of secondary importance to rainfall in simulating the rainfall/runoff process in the catchment.

Flow exits catchment K via a third-order stream, and flow-rates from the catchment are measured by a compound rectangular weir with a centered V notch (Bosch and Sheridan, 2007). The rectangular weir has a crest length of 17.8 m, a V-notch depth of 44.2 cm with 10:1 side slopes, and is designed for a 25-year maximum flowrate of 16.5 m³ s⁻¹. The catchment K outlet coincides with a four-barrel box culvert that provides roadway cross-drainage, and the flow-measurement weir is located between the upstream wing walls, approximately 3 m upstream from the culvert. Water-surface elevations upstream and downstream of the flow-measurement weir were measured to the nearest 2 mm every 5 min with a pressure transducer and recorded by a digital data logger. The rating curve for the flow-measurement structure was developed based on laboratory model testing and field measurements (Bosch and Sheridan, 2007).

HSPF MODEL

The USEPA BASINS system was used to delineate catchment K based on the National Elevation Dataset (NED) and the specified catchment outlet. The catchment area delineated using this digital elevation model was 15.6 km², compared to 16.7 km² commonly associated with catchment K (Bosch and Sheridan, 2007). For consistency in using the NED for stream delineation, a catchment area of 15.6 km² was used in the model simulations. The regional land-use GIRAS layer indicates that the catchment consists of 7.1 km² of agricultural land (45%) and 8.5 km² of forest land (55%). This proportion of agricultural land (45%) in the GIRAS database is higher than has been reported by Bosch et al. (2006) at 35% to 40% and Feyereisen et al. (2007a) at 37%. The relationship between stage and runoff in the stream segment (FTABLE) was generated using the implicit channel geometry in BASINS (Technical Note 1, 2007). The seven-year calibration period for the model was from 2 January 1996 to 28 December 2002, during which time hourly flow measurements and a total of 53 instantaneous measurements of fecal coliform (FC) concentration at the catchment outlet were available. Analysis for fecal coliform was conducted using the EC Medium test, which utilizes a membrane filtration technique and mFC media to identify fecal coliform colonies (APHA, 1995). Precision and accuracy of the FC analyses were documented using sample duplicates, laboratory blanks, and National Institute of Standards and Technology (NIST) traceable reference standards. All of the duplicated samples deviated by less than 25% from the calculated mean, with over 75% deviating by 10% or less from the mean.

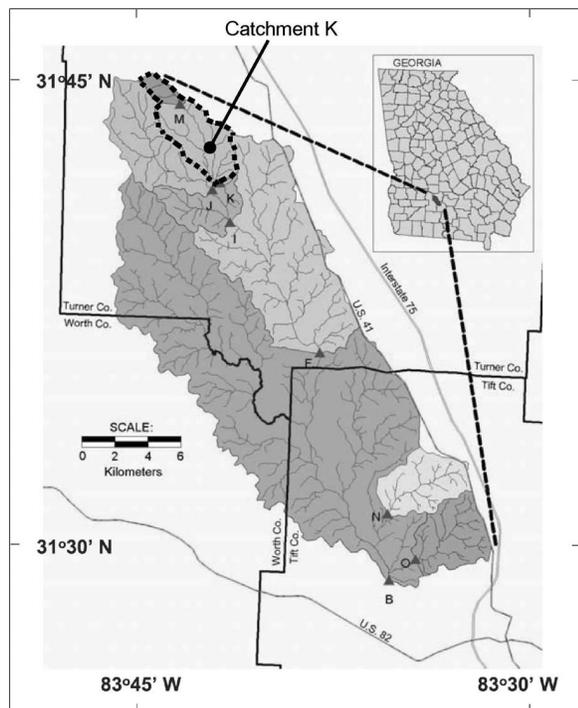


Figure 1. Little River Experimental Watershed, Georgia.

A three-step approach was used to calibrate the hydrology and water-quality components of the model. The steps followed were: (1) identify the model parameters that potentially influence the output variable of interest; (2) sequentially maximize the conditional likelihood of each parameter; and (3) repeat step 2 until there are no changes to maximum-likelihood parameter values.

Parameters having a significant influence on surface runoff in HSPF models are commonly taken as (Van Liew et al., 2003): infiltration rate (INFILT), lower-zone nominal storage (LZSN), ratio of maximum to mean soil infiltration capacities (INFEXP), exponent that determines how much a deviation from nominal lower-zone storage affects the infiltration rate (INFILD), fraction of groundwater that is lost to the deep aquifer (DEEPPFR), active groundwater recession coefficient (AGWRC), and the interflow recession coefficient (IRC). Other potentially influential variables include the upper-zone nominal storage (UZSN), groundwater recession flow parameter (KVARY), interflow inflow parameter (INTFW), and the weighting factor for hydraulic routing (KS). Following the calibration procedure described in detail by Chin (2009), all of the aforementioned parameters were varied sequentially to determine the maximum conditional likelihood, where the conditional likelihood (L) for parameter θ_i is given by:

$$L(y | \theta_i, \hat{\theta}_{j \neq i}) = \frac{1}{S^N} \prod_{i=1}^N \exp \left\{ -\frac{1}{2S^2} (y_i - m_i)^2 \right\} \quad (4)$$

where y is the set of N measured data (i.e., daily flows), θ_i are values 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 corresponding model predictions (m_i) and is given by:

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - m_i)^2} \quad (5)$$

At each step in the calibration sequence, all parameters values except one (θ_i) were fixed, and the maximum condi-

tional likelihood value of that variable (θ_i) was found iteratively. This parameter value was then fixed at its maximum conditional likelihood value, and the maximum conditional likelihood value of the next parameter was found. This process was repeated for subsequent parameters, cycling through all parameters until the parameter set converged to the maximum-likelihood parameter set shown in table 1.

This iterative approach converged to the maximum-likelihood parameter set within three cycles and was automated by developing control software to run the iterations without stopping. It is recognized that the maximum-likelihood parameter set might not be unique due to correlations between model parameters. In such cases where significant parameter correlations exist, alternative maximum-likelihood parameter sets could be identified by using different initial parameter estimates (Chin, 2009). Using the maximum-likelihood parameter set, the predicted and measured flows are compared in figure 2a for a three-year time window within the seven-year calibration period (1996-2002). The three-year time window in figure 2 and subsequent figures was selected primarily to facilitate illustration, since it is long enough to show a significant segment of modeled results versus measured data, and yet still discriminate between daily measurements. Quantitative comparison between the measured and predicted flows is most easily done using the Nash-Sutcliffe efficiency (NSE), defined as:

$$NSE = 1 - \frac{\sum_{i=1}^N (y_i - m_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (6)$$

where \bar{y} is the mean of the measurements.

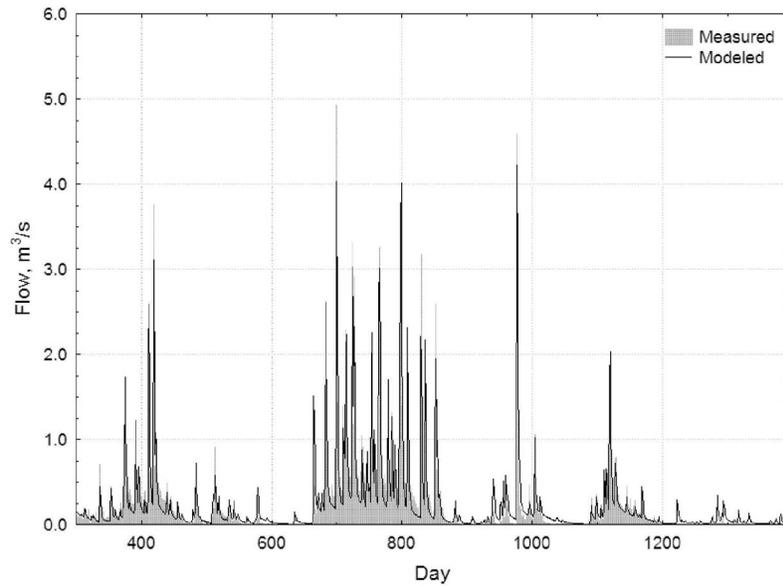
For the seven-year period of calibration, the NSE is 0.87 for daily averaged flows and 0.89 for monthly averaged flows, and the discrepancy in the cumulative outflow volume is 7%. Collectively, these metrics indicate that the model provides excellent agreement between predictions and measurements at both daily and monthly time scales.

Table 1. HSPF maximum likelihood parameters.

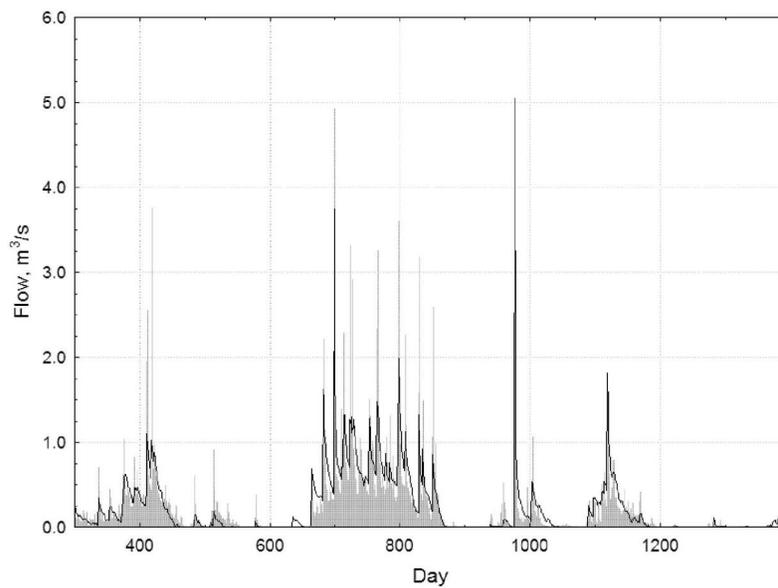
	Parameter	Value	Units	Description
Hydrology	AGWRC	0.95	d ⁻¹	Basic groundwater recession rate
	DEEPPFR	0.05	--	Fraction of groundwater lost from the system
	INFILT	0.86	mm h ⁻¹	Infiltration capacity index
	INFEXP	2.0	--	Exponent in the infiltration equation
	INFILD	2.0	--	Ratio between maximum and mean infiltration
	INTFW	27	--	Interflow inflow parameter
	IRC	0.51	d ⁻¹	Interflow recession parameter
	KVARY	0.	cm ⁻¹	Groundwater recession flow parameter
	LZSN	6.1	cm	Lower zone nominal storage
	UZSN	64.5	mm	Upper zone nominal storage
	KS	0.58		Weighting factor for hydraulic routing
Water quality	ACQOP	0.25 ^[a] , 1.45 ^[b]	× 10 ⁸ FC ha ⁻¹ d ⁻¹	Rate of accumulation of FC
	FSTDEC	2.09	d ⁻¹	First-order decay rate of FC in stream
	IOQC	1423	cfu/100 mL	Interflow FC concentration
	PSRC	17.4	× 10 ⁹ FC d ⁻¹	Mass flux from direct source
	SOQLIM	8.25	× 10 ¹¹ FC ha ⁻¹	Maximum surface storage of FC
	WSQOP	0.94	mm h ⁻¹	Rate of surface runoff that removes 90% of FC in 1 h

^[a] Agricultural land.

^[b] Forest land.



(a) HSPF



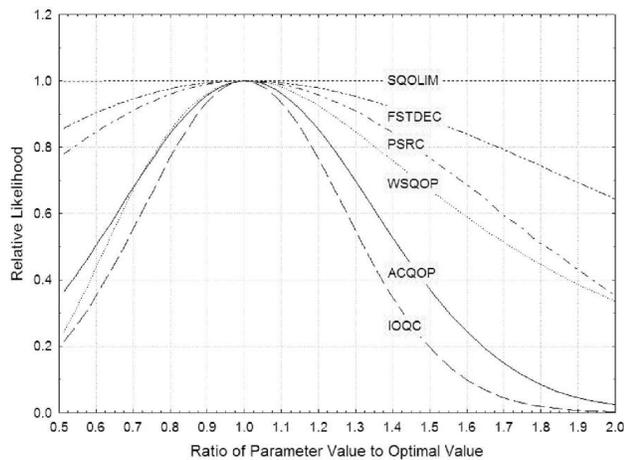
(b) SWAT

Figure 2. Comparison of measured and simulated flows.

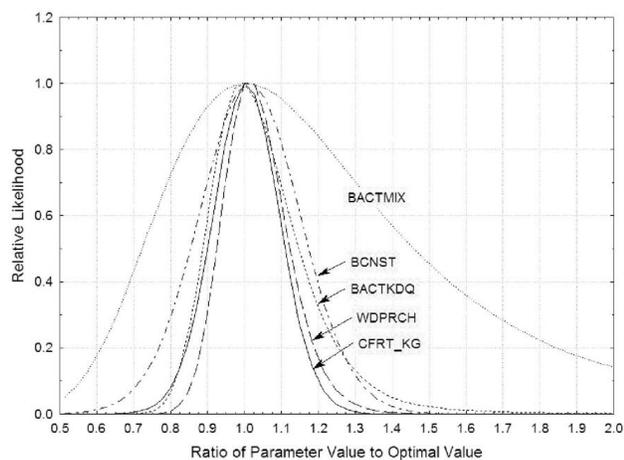
The mechanism of fecal coliform (FC) transport from the land surface to streams is assumed to be by both surface runoff and subsurface transport by interflow, with bacteria concentrations in the groundwater assumed to be zero. In addition, contributions by distributed in-stream sources are accounted for by direct nonpoint source inputs. The parameters that contribute to FC fate and transport were identified as: the rate of terrestrial accumulation of FC (ACQOP), the storage limit of FC on the land surface (SQOLIM), the rate of surface runoff that removes 90% of stored FC per hour (WSQOP), interflow FC concentration (IOQC), FC mass flux from a direct nonpoint source (PSRC), and the first-order decay rate of FC in the receiving stream (FSTDEC). These parameters were varied sequentially using the same approach as for the hydrology calibration, with the likelihood measure

given by equations 4 and 5, and y is taken as the set of FC measurements. The maximum-likelihood parameter set is listed in table 1.

A particularly appealing feature of the calibration scheme used in this study is that the conditional likelihood distribution of each parameter can be determined for the maximum-likelihood parameter set, as shown in figure 3a. These curves show how the likelihood of each parameter changes as it varies between one-half and twice its optimal value. The results in figure 3a provide a direct illustration of the sensitivity of the model output to that parameter, showing that the model predictions are relatively insensitive to values of SQOLIM (= maximum accumulation of bacteria) and most sensitive to IOQC (= interflow bacteria concentration).



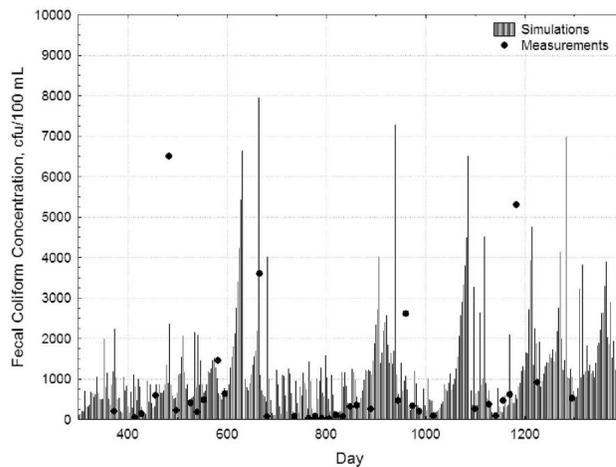
(a) HSPF



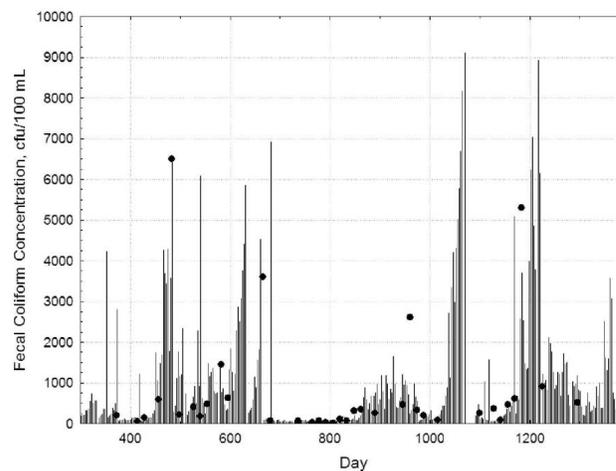
(b) SWAT

Figure 3. Marginal likelihood functions of model parameters.

Using the maximum-likelihood parameter set, the predicted and measured FC concentrations in the receiving stream are compared in figure 4a for the same three-year time window used to compare flows in figure 2a. For the seven-year calibration period, the Nash-Sutcliffe efficiency (NSE) is 0.33, which is comparable to the performance of HSPF models reported in other studies. For example, in using HSPF to develop a TMDL implementation plan in a similar watershed in Georgia (Georgia DNR, 2002), the NSE was found to be 0.24. It is debatable whether it is appropriate to use the NSE without normalizing the FC values, since the FC concentration can vary by orders of magnitude. However, the conventional NSE used in this study is desirable since it gives more weight to matching high FC values, which normally determine impairment designations, and is the benchmark for comparison with other studies. The relatively low agreement between measurements and observations could be attributed to a variety of factors. For example, the comparison is between instantaneous FC measurements and predictions that are both stream averaged and daily averaged, while, in reality, there are likely to be significant within-stream and intra-day variations in FC concentrations.



(a) HSPF



(b) SWAT

Figure 4. Comparison of measured and simulated bacteria concentrations.

SWAT MODEL

The ArcSWAT system (version 2.1) was used to delineate catchment K using the same elevation, land use, and soil database as used in the HSPF model. A minor modification was that the 1 arc-second National Elevation Dataset (NED) used in SWAT had to be downloaded from the USGS National Map Server since the NED built into BASINS was not readily transferable to ArcGIS. For consistency with HSPF, the SWAT model of the watershed contained two hydrologic response units (HRUs): a 7.1 km² agricultural-use HRU and an 8.5 km² forest-use HRU.

The parameter that has the most influence on surface runoff in SWAT is the runoff curve number for moisture condition 2 (CN2) (Van Liew et al., 2003). For subsurface response, the most influential parameters that are commonly cited (e.g., Van Liew et al., 2003; Feyereisen et al., 2007b) are the parameter that controls the amount of water that moves from the shallow aquifer to the root zone (GW_REVAP), threshold depth of water in the shallow aquifer for movement to the root zone (REVAPMN), threshold depth of water in the shallow aquifer required for return flow to occur to the stream

Table 2. SWAT maximum-likelihood parameters.

	Parameter	Value	Units	Description
Hydrology	ALPHA_BF	0.75	d ⁻¹	Base flow recession constant
	CH_K2	132.5	mm h ⁻¹	Effective hydraulic conductivity
	CH_N2	0.044	--	Manning's n in main channel
	CN2	35 ^[a] , 61 ^[b]	--	Curve number for moisture condition II
	GW_DELAY	18.5	d	Groundwater delay time
	GWQMN	0.11	mm	Threshold depth in shallow aquifer for return flow
	GW_REVAP	0.51	--	Groundwater "revap" coefficient
	RCHRG_DP	0.0	--	Deep aquifer percolation fraction
	REVAMMN	0.11	mm	Depth in shallow aquifer for percolation to deep aquifer
Water quality	BACTKDQ	0.53	m ³ Mg ⁻¹	Bacteria soil partitioning coefficient
	BACTMIX	5.6	10 m ³ Mg ⁻¹	Bacteria percolation coefficient
	BCNST	3.55	× 10 ⁸ cfu/100 mL	Direct-source concentration, flow = 0.01 m ³ d ⁻¹
	CFRT_KG	2.0 ^[a] , 22.1 ^[b]	kg ha ⁻¹ d ⁻¹	Application rate at 10 ⁵ cfu g ⁻¹
	WDPRCH	2.33	d ⁻¹	Bacteria dieoff coefficient in streams
	WDPO	0.0	d ⁻¹	Bacteria dieoff coefficient in soils

^[a] Agricultural land.

^[b] Forest land.

(GWQMN), and the fraction of percolation from the root zone that recharges the deep aquifer (RCHRG_DP). Other potentially influential variables include the base flow recession constant (ALPHA_BF), groundwater delay time (GW_DELAY), effective hydraulic conductivity of main channel (CH_K2), and Manning's n of the main channel (CH_N2). Following the same procedure used in calibrating HSPF, the maximum-likelihood parameter set was determined and is given in table 2.

The predicted and measured flows are compared in figure 2b for a three-year time window within the seven-year calibration period (1996-2002). For the period of calibration, the Nash-Sutcliffe efficiency is 0.65 for daily averaged flows, 0.88 for monthly averaged flows, and the discrepancy in the cumulative outflow volume from the watershed is 1%. Collectively, these metrics indicate that the model provides fair agreement with measurements on daily time scales, with excellent agreement on monthly time scales (Moriassi et al., 2007).

The parameters that contribute to FC fate and transport were identified as: bacteria application rate (CFRT_KG), bacteria soil partitioning coefficient (BACTKDQ), bacteria percolation partitioning coefficient (BACTMIX), dieoff factor for bacteria in streams (WDPRCH), dieoff factor for bacteria in soils (WDPO), and direct nonpoint source flux (BCNST). These parameters were varied sequentially using the same approach as for the hydrology calibration, with the likelihood measure taken as the NSE for estimating the 53 FC measurements during the seven-year calibration period. The maximum-likelihood parameter set was determined and is listed in table 2. The conditional likelihood distributions of the parameters influencing FC fate and transport are shown in figure 3b for parameters varying between one-half and twice their optimal value. It is evident that the predictions are sensitive to all parameters, with the greatest sensitivity to the bacteria loading rate (CFRT_KG) and least sensitivity (although still moderately sensitive) to the bacteria percolation partitioning coefficient (BACTMIX). Using the maximum-likelihood parameter set, the predicted and FC concentrations in the receiving stream are compared in figure 4b for the same three-year time window used to compare flows in figure 2b. For the entire calibration period, the NSE for predicting fecal coliform is 0.73.

COMPARISON OF MODELS

The performance of the HSPF and SWAT models as measured by their NSE values are compared in table 3. These results clearly show that both HSPF and SWAT perform well in predicting monthly averaged flows, with HSPF performing much better in predicting daily averaged flows. This is not a surprising or a new observation, since HSPF uses hourly time steps compared with the daily time steps in SWAT, and hence HSPF is better able to resolve the response of the catchment to storms with subdaily durations. These results can be further contrasted with those reported by Feyereisen et al. (2007b), who used SWAT to simulate the hydrology of catchment K for the 1997-2002 subperiod. Feyereisen et al. (2007b) found $NSE_{\text{day}} = 0.56$ and $NSE_{\text{month}} = 0.88$, while the present model applied to this same time period gives $NSE_{\text{day}} = 0.65$ and $NSE_{\text{month}} = 0.89$. The superior performance of the present SWAT model for daily time steps is likely a result of a more comprehensive calibration process involving the sequential identification of conditional maximum-likelihood values of the parameters.

In addition to the consensus result that HSPF produces much better agreement with daily averaged flows than SWAT, the contributions of surface runoff, interflow, and groundwater inflow to the receiving stream indicated by HSPF are also markedly different than SWAT, with direct implications for simulating the fate and transport of contaminants contained in these flow components. For the simulation period (1996-2002), HSPF gives the relative contributions of surface runoff, interflow, and groundwater inflow as 2%, 60%, and 38%, while SWAT gives the relative contributions as 13%, 4%, and 83%.

With respect to predicting bacteria concentrations at the watershed outlet, the results shown in table 3 indicate that SWAT performs better than HSPF, with an NSE of 0.73 versus 0.33. These results suggest that the bacteria fate and transport process equations incorporated in SWAT provide a better representation of the fate and transport of bacteria in this particular watershed. In the context of the present watershed, a significant difference between HSPF and SWAT is that SWAT requires the concentration of bacteria in the interflow and shallow-aquifer inflows to the stream segment be equal

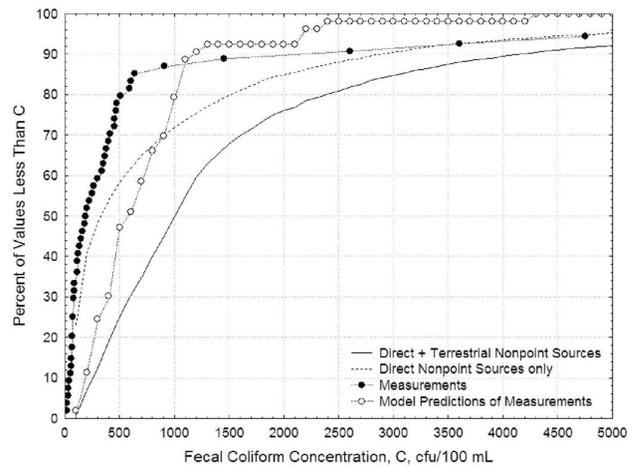
Table 3. Performance comparison of HSPF and SWAT models.

	HSPF	SWAT	a HSPF + (1 - a)SWAT
Hydrology			
NSE _{month}	0.89	0.88	0.90
NSE _{day}	0.87	0.65	0.88
Water quality			
NSE _{all}	0.33	0.726	0.730

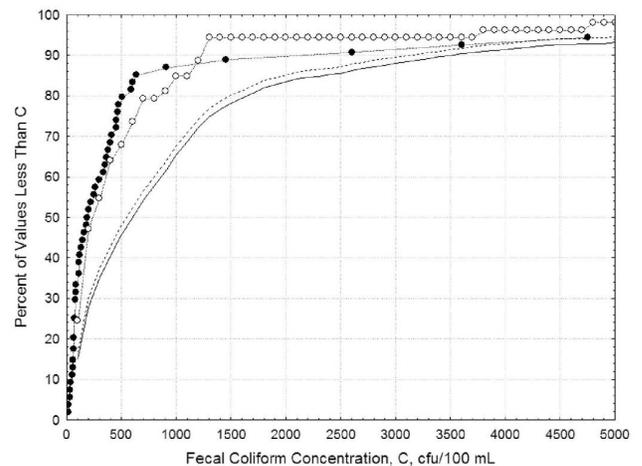
to zero, whereas HSPF allows nonzero concentrations in the subsurface inflow to the stream. In the present case, the hydrology models indicate that interflow and shallow-aquifer inflow contribute between 87% and 98% of the flow to the receiving stream, so this difference in model formulation has significant implications. Secondly, in comparing the conditional likelihood distributions of the HSPF and SWAT fate and transport parameters in figure 3, it is clear that the predicted concentrations are much more sensitive to the SWAT parameters than the HSPF parameters. This gives an indication that the SWAT process equations (and associated parameters) might provide a more precise description of FC fate and transport process in the present case.

To further explore this difference, it is instructive to separate the contributions of terrestrial nonpoint sources and direct nonpoint sources to the bacteria concentrations in the receiving stream. The contribution of terrestrial sources can be extracted by putting the direct nonpoint-source mass flux equal to zero and calculating the stream concentrations of bacteria, and the contributions of direct nonpoint sources can be determined by subtracting the concentrations derived from only terrestrial sources from the concentrations resulting from both terrestrial and direct nonpoint sources (i.e., the calibrated model). The cumulative probability distribution of concentrations resulting from only direct nonpoint sources is compared with the cumulative probability distribution of concentration from both direct and terrestrial nonpoint sources in figure 5a for HSPF and figure 5b for SWAT. In contrasting these results, the key consideration is that the direct nonpoint source cumulative distribution represents the distribution of stream concentrations that could be expected if all terrestrial sources were eliminated. With this in mind, figure 5a indicates that significant reductions in both median and 90-percentile stream concentrations could be obtained by reducing the contributions of terrestrial bacteria sources, while figure 5b indicates that only very small reductions in median and 90-percentile stream concentrations could be achieved by reducing (or even eliminating) the contributions of terrestrial sources. Since the SWAT process equations give much better predictions of the observed bacteria concentrations in the receiving stream, it is reasonable to conclude that figure 5b is more representative of field conditions, and that development of strategies to reduce bacteria concentrations in the receiving stream should focus on in-stream sources that cause relatively continuous bacteria input to the stream. This key result could only have been obtained by comparing the relative performance of two fate and transport process models and would have been missed if only the HSPF model were used in the analysis.

A second salient feature of figures 5a and 5b is the comparison of the cumulative distribution of the measurements (using the Weibull formula) represented by the filled circles and the cumulative distribution of the modeled data on the



(a) HSPF



(b) SWAT

Figure 5. Contribution of terrestrial sources to stream bacteria.

same days as the measurements, represented by the open circles. Figures 5a and 5b show fair agreement given by HSPF and good agreement given by SWAT. For the SWAT model, it is particularly informative to compare the distribution of modeled concentrations on measurement days (which is in good agreement with observations) with the distribution of modeled concentrations on all days of the simulation. Based on this comparison, it is clear that median and 90-percentile values of the modeled concentrations for the set of measurement days underestimate the median and 90-percentile values of the modeled concentrations, respectively, for the set of all days. A likely reason for this is that sampling days tend to be times of low rainfall/runoff, thereby missing the high concentrations that result from large surface runoff events and significantly biasing the estimated cumulative probability distribution of concentrations. This phenomenon has also been previously observed by Baffaut (2006) in a watershed in Missouri. The possible significant bias introduced by using measured data to estimate the cumulative probability distribution of concentrations in the stream is particularly troublesome because compliance with percentile-based water-quality standards is usually assessed based on

measurements, and actual percentile concentrations are likely to be higher than those estimated from measurements.

Finally, it is useful to compare the maximum-likelihood process parameters related to terrestrial loading and direct loading derived from the HSPF and SWAT models. Maximum-likelihood HSPF terrestrial loading is in the range of 0.25×10^8 to 1.45×10^8 cfu ha⁻¹ d⁻¹ for agricultural and forest land uses, while maximum-likelihood SWAT terrestrial loading is in the range 2.0×10^8 to 22.1×10^8 cfu ha⁻¹ d⁻¹. Although these maximum-likelihood terrestrial loading estimates differ by an order of magnitude, they provide a reasonable level of confidence in pinning down the existing bacteria loading to be in the range of 10^7 to 10^9 cfu ha⁻¹ d⁻¹. The maximum-likelihood direct nonpoint-source loading given by HSPF is 1.7×10^{10} cfu d⁻¹, and SWAT gives 3.6×10^{10} cfu d⁻¹, providing strong support to the assertion that existing direct loading from in-stream sources is in the fairly narrow range of 2×10^{10} to 4×10^{10} cfu d⁻¹. This provides an instance where using a multi-model approach gives greater confidence in estimated source loadings than could be derived from using a single model. Since the area of the catchment is 15.6 km² (= 1560 ha), terrestrial loading is estimated to be in the range of 10^{10} to 10^{13} cfu d⁻¹, compared to in-stream loadings on the order of 10^{10} cfu d⁻¹. The predominant sources of fecal coliform in the watershed are likely associated with wildlife, since most of the forested lands in catchment K are heavily populated with wildlife, and the catchment is commonly used as a recreational hunting reserve. The most common wildlife species is deer, but the area also has relatively high populations of wild turkeys, geese, raccoons, and migratory ducks. Most the game animals inhabit the riparian areas along the stream banks and likely contribute to in-stream fecal coliform loadings.

MODEL AVERAGING

Model averaging is an approach in which a weighted average of model output is used as the actual predictor (Claeskens and Hjort, 2008). Although this approach has been applied and tested in several hydrologic studies (e.g., Ajami et al., 2006), it has not to date been applied to predicting contaminant levels using watershed-scale models. Considering HSPF and SWAT as competing models, a model-averaged prediction can be expressed in the form:

$$P_i = aH_i + (1-a)S_i \quad (7)$$

where P_i is the prediction at time step i ; H_i and S_i are the corresponding estimates by HSPF and SWAT respectively; and a is a weighting factor between zero and one. If the Nash-Sutcliffe efficiency (NSE) is used to measure the performance of the combined models, then:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N [M_i - aH_i - (1-a)S_i]^2}{\sum_{i=1}^N (M_i - \bar{M})^2} \quad (8)$$

where M_i is the measurement at time step i , and \bar{M} is the average of the N measurements. The value of a that maximizes NSE can be calculated by taking the derivative of equation 8, which yields:

$$\frac{d\text{NSE}}{da} = 0 \Rightarrow a = \frac{\sum_{i=1}^N (M_i - S_i)(H_i - S_i)}{\sum_{i=1}^N (S_i - H_i)^2} \quad (9)$$

and it can be further shown that:

$$\frac{d^2\text{NSE}}{da^2} = -\sum_{i=1}^N (S_i - H_i)^2 < 0 \quad (10)$$

which guarantees that by taking a as specified by equation 9, the value of NSE will be maximized and be greater than the value of NSE derived from either model by itself. Applying equation 9 to HSPF and SWAT simulations give $a = 0.83$ for daily averaged flows, $a = 0.56$ for monthly averaged flows, and $a = 0.09$ for bacteria concentrations. These values of a provide a direct measure of the relative abilities of the competing models to simulate the observed data. The NSE obtained by using model-averaged predictions are shown in table 3, where it is clear that the gain in best model performance is relatively small, although, strictly speaking, the predictions are all improved by model averaging. These results indicate that the real benefit to using multiple models might not be in the fact that improved performance is achieved by model averaging, but that using multiple models has the potential to identify a preferred model that might demonstrate much better performance than a single model, with no downside if the second model proves to be less accurate. For example, if it was decided *a priori* to use the HSPF model, then in the present case an excellent description of the hydrology and a fair description of the water quality would be realized; by using the second model (SWAT), the potential for a much-improved description of the water quality is realized.

SUMMARY AND CONCLUSIONS

The development and implementation of TMDLs require the quantitative determination of both terrestrial and direct loadings on impaired waters, the sources of loading, and the relationships between loading reductions and contaminant concentrations in the impaired waters. To these ends, the watershed-scale models HSPF and SWAT are the most widely used in the U.S., and conventional practice is to *a priori* select one of these models in any particular application. The results presented in this article have demonstrated the added dimensionality that can be achieved by using multiple models to predict the fate and transport of bacteria at the watershed scale. Both HSPF and SWAT were applied to the 15.6 km² catchment K of the Little River Experimental Watershed (LREW) in Georgia. It was demonstrated that, over the seven-year period from 1996 to 2002, HSPF provided a much more accurate description of daily averaged flows with a Nash-Sutcliffe efficiency (NSE) of 0.87, both models performed comparably in describing monthly averaged flows with an NSE of around 0.89, and SWAT provided a much more accurate description of fecal coliform concentrations with an NSE of 0.73 compared to 0.33 for HSPF.

The relative performance of the models in predicting daily averaged flows is expected, since the HSPF model uses hourly time steps and is capable of resolving the response of the watershed to storms with subdaily time scales, while SWAT uses daily time steps and is unable to accurately resolve re-

sponses to individual storms. The ability of SWAT to perform much better than HSPF in predicting fecal coliform concentrations in the receiving stream indicates that, in this particular watershed, the SWAT process equations are more representative of the watershed-scale fate and transport than the HSPF process equations. As a consequence, it can reasonably be concluded that elevated levels of fecal coliform in the receiving stream are primarily due to direct in-stream sources of bacteria, and that terrestrial sources distributed throughout the watershed have relatively little effect on bacteria levels in the stream, except under high rainfall/runoff conditions. This source characterization could not be achieved by using only the HSPF model, which indicates a much greater contribution of terrestrial sources, which is discounted due to the much poorer performance of HSPF in predicting stream concentrations of bacteria. A model-averaging approach in which a weighted-average of the HSPF and SWAT predictions was used to predict flows and bacteria concentrations in the receiving stream was investigated. It was shown analytically that model weights can be determined such that the NSE of the combined models will be greater than that of either of the models taken individually. This model-averaging approach was implemented; however, the marginal improvements in best-model NSE were relatively small.

Collectively, the results presented in this article indicate that the primary benefit in using multiple models to simulate watershed-scale fate and transport is the potential to identify a model that might provide a significant improvement in prediction capability over a single model, with no downside if the second model proves to be less accurate. The fate and transport processes incorporated in the more accurate model can then be identified as more representative of the catchment under investigation, a result that is particularly useful in developing implementation plans for TMDLs.

ACKNOWLEDGEMENTS

Nancy Sammons of the USDA-ARS Grassland Soil and Water Research Laboratory in Temple, Texas, provided support and advice in the development of the SWAT model, and Tom Jobes of the St. Johns River Water Management District in Palatka, Florida, provided support and advice in the development of the HSPF model. Chris Hanson of the University of Miami provided critical support on the GIS components of the models.

REFERENCES

Ajami, N., Q. Duan, X. Gao, and S. Sorooshian. 2006. Multi-model combination techniques for hydrological forecasting: Application to distributed model intercomparison project results. *J. Hydrometeorology* 7(4): 755-768.

APHA. 1995. *Standard Methods for the Examination of Water and Wastewater*. 19th ed. Washington, D.C.: American Public Health Association.

Asmussen, L. 1971. Hydrologic effects of Quarternary sediments above the marine terraces in the Georgia Coastal Plain. *Southeastern Geology* 12(3): 189-201.

Baffaut, C. 2006. Little Sac River watershed fecal coliform total maximum daily load. FAPRI-UMC Report No. 11-06. Columbia, Mo.: University of Missouri, Food and Agricultural Policy Research Institute (FAPRI).

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. *Trans. ASABE* 49(4): 987-1002.

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

Bosch, D., and J. Sheridan. 2007. Stream discharge database, Little River experimental watershed, Georgia, United States. *Water Resources Res.* 43(9): W09473, doi:10.1029/2006WR005833.

Bosch, D., D. Sullivan, and J. Sheridan. 2006. Hydrologic impacts of land-use changes in coastal plain watersheds. *Trans. ASABE* 49(2): 423-432.

Chin, D. 2006. *Water-Quality Engineering in Natural Systems*. New York, N.Y.: John Wiley and Sons.

Chin, D. 2009. Predictive uncertainty in water-quality modeling. *J. Environ. Eng.* (in review).

Claeskens, G., and N. Hjort. 2008. *Model Selection and Model Averaging*. Cambridge, U.K.: Cambridge University Press.

Feyereisen, G., R. Lowrance, T. Strickland, J. Sheridan, R. Hubbard, and D. Bosch. 2007a. Long-term water chemistry database, Little River experimental watershed, southeast Coastal Plain, United States. *Water Resources Res.* 43(9): W09474, doi:10.1029/2006WR005835.

Feyereisen, G., T. Strickland, D. Bosch, and D. Sullivan. 2007b. Evaluation of SWAT manual calibration and input parameter sensitivity in the Little River watershed. *Trans. ASABE* 50(3): 843-855.

Georgia DNR. 2002. Total maximum daily loads (TMDLs) for fecal coliform in 303(d) listed streams in the Oconee River Basin. Atlanta, Ga.: Georgia Department of Natural Resources, Environmental Protection Division.

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. *Agric. Water Mgmt.* 70(1): 1-17.

Moriasi, D., J. Arnold, M. V. Liew, R. Bingner, R. Harmel, and T. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50(3): 885-900.

Neitsch, S., J. Arnold, J. Kiniry, and J. Williams. 2005. Soil and Water Assessment Tool theoretical documentation. Version 2005. Temple, Tex.: USDA-ARS Grassland Soil and Water Research Laboratory.

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. *Trans. ASAE* 47(1): 69-78.

USEPA. 2008. Causes of impairment for 303(d) listed waters. Washington, D.C.: U.S. Environmental Protection Agency.

Van Liew, M., J. Arnold, and J. Garbrecht. 2003. Hydrologic simulation on agricultural watersheds: Choosing between two models. *Trans. ASABE* 46(6): 1539-1551.