WATER QUALITY CALIBRATION CRITERIA FOR BACTERIA TMDL DEVELOPMENT

S. M. Kim, B. L. Benham, K. M. Brannan, R. W. Zeckoski, G. R. Yagow

ABSTRACT. *The objective of this study was to report the development of statistics and associated criteria for assessing calibration endpoints of a watershed-scale model used to simulate in-stream bacteria concentrations. Development of these statistics grew out of the authors' experience modeling in-stream bacteria for TMDLs addressing bacterial impairments. The calibration statistics include: geometric mean, daily violation rate of bacteria water quality criterion, and the minimum-maximum range associated with a "temporal window" that spans a period of several days. Because in-stream bacteria concentrations are typically sampled infrequently (on a monthly basis, at best) and represent only an instant in time, it is not reasonable to expect any model to simulate a daily average concentration equal to an observed value on a particular day. For this reason, the authors developed the temporal-window statistic. The temporal-window statistic uses simulated hourly-concentrations over a period of several days (five days were used by the authors) to calculate the minimum-maximum range that is compared to observed in-stream bacteria concentrations. Each temporal window is centered on the day the observed data was collected. Thus, this measure of model calibration determines how frequently the observed data falls within the range of simulated data during a time period that extends several days before and after the observation. Criteria for assessing the sufficiency of a model calibration for simulating in-stream bacteria concentration were developed using the relative difference between the simulated and observed statistics. The reported calibration statistics and criteria can be used to guide the water quality model calibration process for simulation of in-stream bacteria concentration. This article illustrates the application of the calibration statistics to HSPF simulations for the Beaver Creek watershed in Virginia.*

Keywords. *TMDL, Bacteria, Fecal coliform, Water quality, Model calibration, Modeling.*

Under the current demands of the nation's Total Maximum Daily Load (TMDL) program, a large amount of watershed-scale modeling is being conducted in the attempt to identify and quantify pollutant sources so that the pollution from those sources may be reduced to improve water quality. Much of the modeling is directed toward TMDL development for waters impaired due to excessive levels of bacteria. In Virginia alone, over 8,000 km of streams and rivers (constituting nearly 600 segments) are impaired due to violations of indicator bacteria water quality criteria (VADEQ, 2006a); nationwide, over 7,800 segments are reported to have pathogen (mainly indicator bacteria) impairments (USEPA, 2005). Within the TMDL context, the load refers to the maximum amount of pollutant that a given water body can assimilate without negatively affecting its designated use. This 'allowable' load is the target load for water quality remediation. The objective of a TMDL is to allocate loads among different pollutant sources so that the appropriate corrective actions can be taken to meet water quality standards (USEPA, 1991).

The Virginia Department of Environmental Quality (VADEQ), the agency responsible for TMDL development in Virginia, requires the use of a watershed-scale water quality model to develop TMDLs for bacteria impairments except in rare circumstances (VADEQ, 2006a). The Hydrological Simulation Program-FORTRAN (HSPF) is the model most often used to characterize the bacteria sources in the impaired watershed and to establish the relationships between loads from those sources (both point and nonpoint) and in-stream bacteria concentrations. HSPF is a watershed-scale, process-based, lumped-parameter model designed for continuous simulation of hydrologic and water quality (both land loading and in-stream) processes (Bicknell et al., 2001; Duda et al., 2001). HSPF represents spatial variability by dividing the watershed into hydrologically homogenous land segments. In HSPF, the watershed processes for pervious areas, impervious areas, and reaches or reservoirs are modeled separately. The hydrologic simulation in HSPF is based on a mass balance approach. The water quality simulation in HSPF assumes that constituents are associated with one or more of the following: sediment, overland flow, interflow, and groundwater flow. Most frequently, fecal coliform is assumed to be flow associated. Flow-associated constituents are assumed to be accumulated on the land surface until the occurrence of a rainfall event (Paul et al., 2004). In Virginia, in-stream bacteria have typically been simulated as completely dissolved pollutants (VADEQ, 2006a).

Conceptual models like HSPF must be calibrated to insure that the model adequately represents watershed characteristics and processes (van Griensven, 2002; Van Liew et al.,

Submitted for review in July 2006 as manuscript number SW 6556; approved for publication by the Soil & Water Division of ASABE in December 2006.

The authors are **Sang Min Kim,** Post-doctoral Research Associate, **Brian L. Benham, ASABE Member Engineer,** Extension Specialist, Assistant Professor, **Kevin M. Brannan, ASABE Member Engineer,** Research Associate, **Rebecca W. Zeckoski, ASABE Member Engineer,** Research Associate, and **Gene R. Yagow, ASABE Member Engineer,** Research Scientist, Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. **Corresponding author:** Brian L. Benham, 209 Seitz Hall, Biological Systems Engineering Dept., Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; phone: 540-231-5705; fax: 540-231-3199; e-mail: benham@vt.edu.

  ISSN 0883−8542

2005). Calibration is the process of adjusting appropriate model parameter values in an attempt to improve the agreement between simulated output and observed data. Even under ideal conditions, calibration is as much art as science and requires the use of the best professional judgment to know which parameters to adjust, by how much, and when agreement between simulated and observed values is "reasonably good" (Moore and Doherty, 2005).

In Virginia, the VADEQ provides some limited calibration guidance when using HSPF to develop bacteria impairment TMDLs in a guidance memo to TMDL developers (Lawson, 2003). The memo directs the modeler to use the 'Expert System for the Calibration of HSPF' (HSPEXP) (Lumb et al., 1994) when performing hydrology calibrations, provides a list of model parameters that can be adjusted, discusses selected criteria to assess the quality of the hydrology calibration, and specifies how the hydrology calibration results are presented. The memo also provides some minimal guidance for calibrating HSPF to simulate in-stream bacteria concentration. The memo again suggests list of model parameters that can be adjusted, but discusses only very general guidelines to assess the quality of the water quality calibration, and no specific calibration criteria are provided. A search of the peer reviewed literature revealed no citations specifically related to calibration approaches or criteria for simulating in-stream bacteria concentration.

The greatest challenge when calibrating a model to simulate water quality constituents is the paucity of observed data (Paul et al., 2004). Monitoring programs are generally not designed to capture watershed responses that are central to water quality calibration, and even when resources are available for collecting suitable data, it is not straightforward to judge which data will be relevant prior to modeling the system in detail (Beck, 1987). There are always errors, such as biases and random effects in sampling and analytical techniques, associated with water quality observations (Keith, 1990; Jarvie et al., 2002).

In Virginia, in-stream bacteria concentrations are enumerated from grab samples that are collected at fixed frequency (typically monthly) at a given location. Simulation of in-stream bacteria concentrations in Virginia is performed at an hourly time-step in HSPF; the hourly concentrations are then averaged to produce a simulated daily average in-stream bacteria concentration (Lawson, 2003). Because models are simplifications of reality, it is not reasonable to capture all sources of variability that occur in nature. Therefore, it is unrealistic to expect simulated daily average in-stream bacteria concentrations to match observed data precisely. We suggest that a more reasonable expectation is that an observed concentration will fall within a range of simulated concentrations over some short period of time, a "temporal window" that includes the sample date. Our approach, then, is to compare observed data with a range of simulated data around the time an in-stream bacteria concentration observation is made. This approach considers the inherent variability in the observed data and presents what we consider a more realistic and achievable calibration target.

Water quality calibration guidance is needed to ensure a consistent approach when simulating in-stream bacteria concentrations and to determine appropriate calibration endpoints. This article presents a set of statistics and calibration end-point criteria that can be used when simulating in-stream bacteria concentration with HSPF. These statistics are based on the temporal-window concept and evolved from the authors' experience developing 26 bacteria impairment TMDLs. The objective of this article is to illustrate the use of these statistics using a Virginia watershed where the authors developed a bacterial impairment TMDL.

Methods

Calibration Statistics

In Virginia, when developing load allocations for a bacterial impairment TMDL, simulated in-stream fecal coliform bacteria concentrations must not exceed two criteria that are part of the state's water quality standard: the calendar month geometric mean and the single-sample criterion (VADEQ, 2006a). The geometric mean criterion applies when multiple samples are collected in any given calendar month (at least two samples must be collected during the month), while the single-sample (or instantaneous) criterion applies to the maximum allowable instantaneous concentration in any individual sample (VADEQ, 2006a). Since multiple samples are rarely collected in any given calendar month, the geometric means used to assess model calibration are the geometric mean of all daily-average simulated concentrations for the entire calibration period and the geometric mean of all the observed data for the same period (VADEQ, 2006a).

A period of 5 days was selected for the temporal window for the case study presented herein. The 5-day temporal window was selected based on the authors' experience with modeling in-stream bacteria concentrations while developing TMDLs. The selection of the 5-day temporal window considered: the agreement in storm timing between observed rainfall and flow data, agreement between observed and simulated storm peaks, and uncertainty associated with bacteria input data (especially information about wildlife and livestock access to, and time spent in, streams). A 5-day temporal window may not be appropriate for all watersheds. More research is needed to suggest what period might be more appropriate for specific watershed conditions and modeling objectives.

For the calibration process reported here, a series of 5-day temporal windows (5-day windows) were created, one for each observed data point. Each 5-day window was centered on the day the observation was collected and includes 2 days before and 2 days after. The range (minimum and maximum) of simulated hourly concentrations were determined for each 5-day window and compared with the observed data to calculate the following temporal-window calibration statistics:

- PCT_Within: Percent of observed values within 5-day window minimum-maximum range
- PCT_Above: Percent of observed values > 5-day window maximum
- PCT_Below: Percent of observed values < 5-day window minimum

The PCT_Within statistic represents the measure of agreement between the observed data and the range of simulated concentrations for each observation's 5-day window; the more observations that fall within the simulated range of values for their 5-day windows, the higher the percentage. PCT_Above and PCT_Below are used to quantify the degree to which the model is over- or under-predicting in-stream concentration. The authors recommend that PCT_Above be roughly equal to PCT_Below to avoid over-prediction or under-prediction.

Each of the temporal-window statistics were developed and tested through the authors' TMDL development experience. The temporal-window statistics are considered to be an informative measure of the amount of dispersion of simulated data centered on the date when a sample is collected. Other calibration statistics used by the authors include the average, median, geometric mean, and the instantaneous violation rate of Virginia's bacteria water quality standard's single-sample criterion. The calibration criteria for each of these statistics and the temporal-window statistics are listed in table 1. In addition to these statistics, the authors compare the observed and simulated data graphically. Visually comparing the observed and simulated data permit an overall assessment of the model predictions and often assist in diagnosing difficult calibration issues. The calibration of HSPF for the Beaver Creek watershed is presented to demonstrate the use of the calibration statistics. The results of three runs will be presented: Initial Run – the first model run using initial estimates of model input; Intermediate Run – a run selected after several significant changes to model input were made; Final Run – the simulation results for the model input accepted as the calibration end-point.

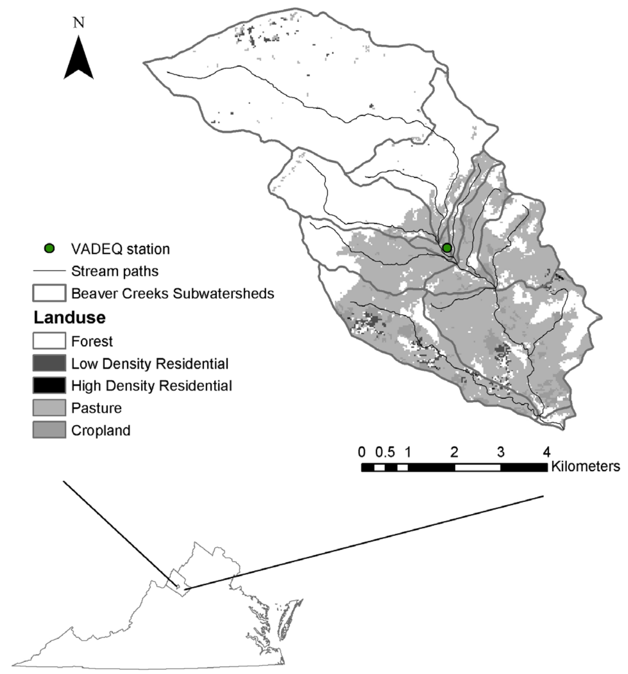

Figure 1. Beaver Creek watershed location, land use, and water quality sampling stations.

Beaver Creek Bacteria Calibration

Study Watershed

The use of the proposed calibration statistics and criteria are illustrated for the Beaver Creek watershed where the authors developed a bacterial impairment TMDL (Al-Smadi et al., 2005). The Beaver Creek watershed, located west of the City of Harrisonburg, Virginia, is a tributary of the North River (USGS hydrologic unit code 02070005), which is a tributary of the Shenandoah River (fig. 1). The watershed is 41.3 km^2 in size. Beaver Creek is a mainly forested watershed (about 61%) with the remaining 39% in agricultural land and some low density residential (fig. 1).

Observed Water Quality Data

The VADEQ has monitored water quality in Beaver Creek (1BBVR003.60) periodically from September 1994 to the present collecting a total of 33 samples; 52% of those samples violated the instantaneous bacteria water quality criterion (Al-Smadi et al., 2005). VADEQ collected one sample per year from 1994 to 1997, after which sampling stopped for almost two years. From August 1999 to June 2003, VADEQ collected samples monthly. Figure 2 shows the timeline and the observed fecal coliform (FC) concentrations (VADEQ, 2006b). During the calibration period (1 January 1999 to 30 June 2003), there were a total of 25 samples; 12 of which violated the bacteria standard's instantaneous sample criterion of 400 CFU/100 mL that was in effect at the time, an instantaneous violation rate (IVR) of 48%.

Results

Water Quality Model Calibration

Table 2 lists the simulated and observed calibration statistics values for the Initial Run, a representative Intermediate Run, and the Final (calibrated) Run, with the bold values in the "% Diff" column indicating which calibration statistics meet the criteria presented in table 1. For the Initial Run, the simulated Geomean and IVR were much less than the observed data (-44% and -40%, respectively), while the simulated median was much greater than the observed median (+89%). Also, PCT_Above (24%) was larger than PCT_Below (12%). The graphical comparison of the simulated data 5-day window and the observed data is shown in figure 3a. The low Geomean and IVR calibration statistics for the Initial Run indicated that the model was generally under-predicting in-stream bacteria concentrations. To increase simulated concentrations, the amount of time wildlife spends in streams (a parameter with a great deal of uncertainty) was increased (Intermediate Run). As expected,

Table 1. Suggested calibration statistics and criteria for use when simulating in-stream bacteria concentration.

Statistic	Abbreviation	Criteria
Average	Average	±100%
Median	Median	±100%
Geometric mean	Geomean	±100%
Instantaneous violation rate	IVR	±10%
Percent of observed values within 5-day minimum-maximum range	PCT_Within	>70%
Percent of observed values > 5-day maximum	PCT_Above	Roughly equal to percent of observed values < 5-day minimum
Percent of observed values < 5-day minimum	PCT_Below	Roughly equal to percent of observed values > 5-day maximum

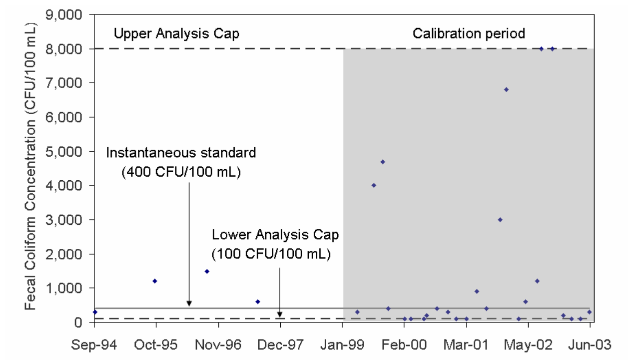

Figure 2. Time series of in-stream bacteria (fecal coliform) concentrations sampled in the Beaver Creek watershed.

this change increased the simulated in-stream bacteria concentration. This change also increased all calibration statistic %Diff values (table 2) except PCT_Above, which decreased from 24% to 8%, as now only 2 of 25 observed values were above their respective 5-day window maximum values (fig. 3b). For the Intermediate Run, the simulated Geomean and IVR were greater than the observed data (+25% and +10%, respectively) and the simulated median increased to 209% above the observed median. The PCT_Above decreased (8%) while PCT_Below increased (20%). These statistics and the graphical comparison (fig. 3b) indicate the model was over predicting in-stream bacteria concentration. To address the Intermediate Run over-prediction, several adjustments were made. The population per household was reduced to decrease the PCT_Below statistic. This adjustment decreased the contribution of bacteria from straight pipes (households where excrement is not treated in any way). To address the Geomean and IVR over-prediction, livestock access to streams was decreased (i.e. the number of livestock defecating directly into the stream was decreased). The calibration's Final Run resulted in small %Diff for the Geomean (-2%) and IVR (0%) calibration statistics (table 2). The PCT_Within showed that 72% of the observed values fell within the simulated 5-day window ranges and the agreement between the PCT_Above and PCT_Below statistics was acceptable. The graphical comparison for the Final Run is shown in figure 3c. The average daily fecal coliform concentration output from the calibrated model is shown with the observed data is in figure 4.

DISCUSSION

The temporal-window calibration statistics presented here provide modelers with feedback on the distribution of simulated in-stream bacteria concentration with respect to observed data. These temporal-window statistics, when combined with more traditional measures of average, median, Geomean, and IVR can be used by the modeler to adjust and refine the model's calibration to ensure that the model is adequately simulating in-stream bacteria concentration. In conjunction with these statistics, graphical (visual) comparisons of simulated and observed data are recommended as good modeling practice. A graphical comparison is indispensable because it can show particular trouble spots (e.g. outliers, questionable data) and suggest a reason for not meeting one or more of the calibration criteria. Visually comparing simulated and observed data also provides insight into the overall quality of the simulation to the modeler and to the people using the modeling results for decision making. The bacteria calibration procedure and criteria discussed here were applied in the development of other TMDL plans in Virginia and proved useful in achieving a acceptable calibration (Benham et al., 2005, 2006; Mostaghimi et al., 2004).

Additionally, other factors such as total number of samples, frequency and timing of sample collection (e.g., season of the year), and quality of hydrology calibration should be taken into consideration when calibrating a model to simulate in-stream bacteria concentration. If the total number of observations is small, it may be easier to calibrate the model, but much more difficult to arrive at a meaningful calibration in terms of most of the criteria listed in table 1. It may also be necessary to put more weight on the graphical comparison. If the samples are routinely collected in one season of the year and not others, one would expect the geometric mean and single-sample violation rates of the observed data to be biased to conditions in the over-sampled season, however, limited data should still mostly fall within the minimum-maximum ranges of the temporal window. Finally, since simulated water quality processes build on hydrologic processes, the quality of the hydrologic calibration has a major impact on the quality of the in-stream bacteria concentration calibration. Therefore, a poor hydrologic calibration (or poor simulation of one or more specific seasons) would greatly affect the ability to calibrate in-stream bacteria concentrations.

In the Beaver Creek case-study presented here, a 5-day temporal window was used. Selection of the 5-day window was based on the authors' experience. For the Beaver Creek calibration, the agreement between the simulated and observed Geomean and IVR calibration statistics were exceptionally good (table 2). Based on these statistics, the relatively good agreement between the average and median statistics, and the graphical comparison (fig. 3c), the

Table 2. Comparison of model calibration statistics for Beaver Creek calibration runs; initial, representative intermediate, and final run.

		Initial Run			Intermediate Run			Final Run		
Statistic	Criteria	Obs	Sim	% Diff	Obs	Sim	% Diff	Obs	Sim	% Diff
Average	±100	1,620	1,616	**<1**	1,620	2,145	**32**	1,620	917	**-43**
Median	±100	300	568	**89**	300	927	209	300	467	**56**
Geomean	±100	480	287	**-40**	480	599	**25**	480	468	**-2**
IVR (%)	±10	48	27	-44	48	53	**10**	48	48	**0**
PCT_Within (%)	>70		64		**72**			**72**		
PCT_Above (%)	Equals PCT_Below		24		8			20		
PCT_Below (%)	Equals PCT_Above		12		20			8		

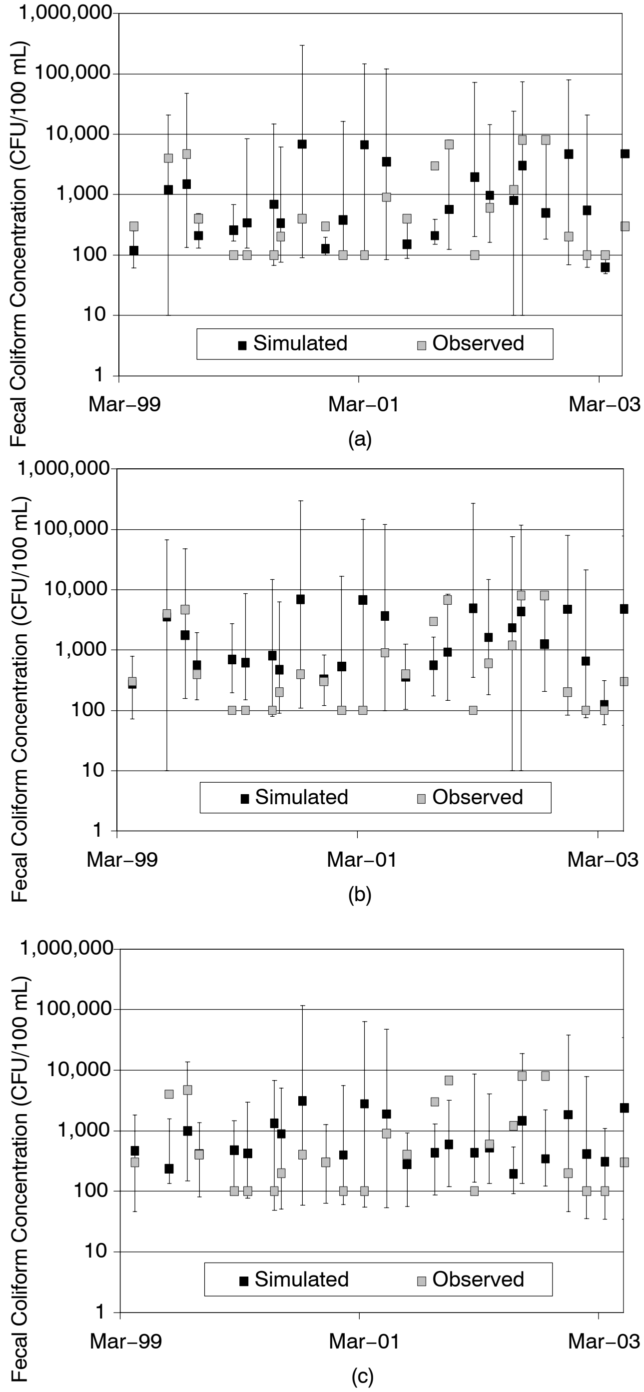

Figure 3. “5-day window” of simulated in-stream bacteria concentrations surrounding each observed value for (a) initial run, (b) intermediate run, and (c) final run.

agreement between the PCT_Above and PCT_Below statistics was deemed acceptable.

Summary and Conclusions

In this study, statistics and criteria were presented to aid in the calibration of HSPF used to simulate in-stream bacteria concentration. The calibration statistics included the average, median, and geometric mean and the instantaneous violation rate (IVR) of the observed and daily average simulated concentrations for the calibration period, and

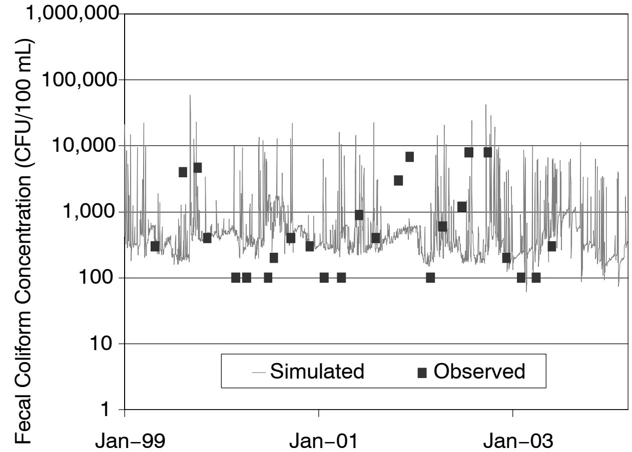

Figure 4. Comparison of observed monthly and simulated average daily in-stream bacteria concentrations in Beaver Creek.

temporal-window statistics. Generally, the number of observed values available for model calibration of in-stream bacteria concentrations is limited. The temporal-window statistics provide a reasonable measure and target for calibration that accounts for greater variability in observed data and provide additional information (minimum, average, and maximum of simulated data several days before, the day of, and several days after the sampling date) to assess the calibration quality, especially where observed data are limited. These statistics and criteria can be used by the water quality modeler as a guide to choosing which parameters to adjust and by how much, and to determine when sufficient agreement is achieved between the observed and simulated data. Graphical comparisons of observed and simulated data are also useful in assessing model calibration.

Application of the calibration statistics and criteria presented here was illustrated for the Beaver Creek watershed in Virginia. A 5-day temporal window was selected for this application based on the authors’ experience in modeling in-stream bacteria concentrations in Virginia. Overall, the calibration statistics provided enough information about the simulation performance to select parameters that needed adjustments. Also, the calibration statistics provided information that was used to impart a rational justification for the needed adjustments. This approach represents a time-saving and logical procedure that can be used to improve the calibration process for simulating in-stream bacteria concentrations and could also be adapted for other water quality parameters with limited data. Other modelers are encouraged to build on this approach in order to increase the confidence in simulations of in-stream bacteria concentrations for the development of TMDLs.

References

Al-Smadi, M., B. L. Benham, K. M. Brannan, G. Yagow, and R. W. Zeckoski. 2005. Bacteria total maximum daily load development for Beaver Creek. Richmond, Va.: Virginia Department of Environmental Quality, Virginia Department of Conservation and Recreation. Available at: http://www.deq.virginia.gov/tmdl/apptmdls/shenrvr/beaver.pdf. Accessed 8 June 2006.

Beck, M. B. 1987. Uncertainty in water quality models. *Water Resources Research* 23(8): 1393-1441.

Benham, B. L., K. M. Brannan, S. M. Morris, G. R. Yagow, and R. W. Zeckoski. 2005. Bacteria total maximum daily load development for North River. Richmond, Va.: Virginia Department of Environmental Quality. Available at: http://www.deq.virginia.gov/tmdl/apptmdls/shenrvr/northbac.pdf. Accessed 27 September 2006.

Benham, B. L., R. W. Zeckoski, and A. Mishra. 2006. Bacteria total maximum daily load development for Pigg River, Snow Creek, Story Creek, and Old Womans Creek. Richmond, Va.: Virginia Department of Environmental Quality, Virginia Department of Conservation and Recreation. Available at: http://www.deq.virginia.gov/tmdl/apptmdls/roankrvr/piggec.pdf. Accessed 28 September 2006.

Bicknell, B. R., J. C. Imhoff, J. L. Kittle, Jr., T. H. Jobes, A. S. Donigian, Jr., and R. C. Johanson. 2001. Hydrological Simulation Program – FORTRAN: HSPF Version 12 User's Manual. Mountain View, Calif.: AQUA TERRA Consultants. In Cooperation with the U.S. Geological Survey and U.S. Environmental Protection Agency. Available at: http://www.epa.gov/waterscience/basins/b3docs/HSPF12.zip. Accessed 8 June 2006.

Duda, P., J. Kittle, Jr., M. Gray, P. Hummel, and R. Dusenbury. 2001. WinHSPF, Version 2.0, An Interactive Windows Interface to HSPF, User Manual. Contract No. 68-C-98-010. Washington D.C.: U.S. Environmental Protection Agency.

Jarvie, H. P., P. J. A. Withers, and C. Neal. 2002. Review of robust measurement of phosphorus in river water; sampling, storage, fractionation and sensitivity. *Hydrology and Earth System Sciences* 6(1): 113-132.

Keith, L. H. 1990. Environmental sampling: a summary. *Environmental Science and Technology* 24(5): 610-617.

Lawson, L. G. 2003. HSPF Model Calibration and Verification for Bacteria TMDLs. Guidance Memo No. 03-2012. Department of Environmental Quality Water Division, Commonwealth of Virginia. Available at: http://www.deq.state.va.us/waterguidance/pdf/032012.pdf. Accessed 12 November 2006.

Lumb, A. M., R. B. McCammon, and J. L. Kittle. 1994. Users Manual for an EXPert System (HSPEXP) for Calibration on the Hydrological Simulation Program-Fortran. Water-Resources Investigations Report 94-4168. Reston, Va.: U.S. Geological Survey.

Moore, C., and J. Doherty. 2005. Role of the calibration process in reducing model predictive error. *Water Resources Research* 41(W05020): doi:10.1029/2004WR003501.

Mostaghimi, S., B. L. Benham, K. M. Brannan, K. Christophel, T. A. Dillaha, R. Wagner, J. Wynn, G. R. Yagow, and R. W. Zeckoski. 2004. Bacteria TMDLs for Abrams Creek and Upper and Lower Opequon Creek Located in Frederick and Clarke County, Virginia. Richmond, Va.: Virginia Department of Environmental Quality, Available at: http://www.deq.virginia.gov/tmdl/apptmdls/shenrvr/abropefc.pdf. Accessed 27 September 2006.

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

USEPA. 1991. Guidance for water quality-based decisions: The TMDL process. EPA 440/4-91-001. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water.

USEPA. 2005. Total Maximum Daily Loads: National Section 303(d) List Fact Sheet: Top 100 Impairments. Washington, D.C.: U.S. Environmental Protection Agency. Available at: http://oaspub.epa.gov/waters/national_rept.control#TOP_IMP. Accessed 8 June 2006.

VADEQ. 2006a. 9 VAC 25-260 Virginia Water Quality Standards. Richmond, Va.: Virginia Department of Environmental Quality. Available at http://www.deq.virginia.gov/wqs/documents/WQS06_EDIT_001.pdf. Accessed 22 February 2007.

VADEQ. 2006b. Surface Water Monitoring. Richmond: Virginia Department of Environmental Quality. Available at: http://www.deq.virginia.gov/watermonitoring/aqm.html. Accessed 20 September 2006.

van Griensven, A. 2002. Developments towards integrated water quality modeling for river basins. Publication No. 40. Brussels, Belgium: Vrije Universiteit, Department of Hydrology and Hydraulic Engineering.

Van Liew, M. W., J. G. Arnold, and D. D. Bosch. 2005. Problems and potential of autocalibrating a hydrologic model. *Transactions of ASAE* 48(3): 1025-1040.