



UNCERTAINTY IN TMDL MODELS

A. Shirmohammadi, I. Chaubey, R. D. Harmel, D. D. Bosch, R. Muñoz-Carpena,
C. Dharmasri, A. Sexton, M. Arabi, M. L. Wolfe, J. Frankenberger, C. Graff, T. M. Sohrabi

ABSTRACT. *Although the U.S. Congress established the Total Maximum Daily Load (TMDL) program in the original Clean Water Act of 1972, Section 303(d), it did not receive attention until the 1990s. Currently, two methods are available for tracking pollution in the environment and assessing the effectiveness of the TMDL process on improving the quality of impaired water bodies: field monitoring and mathematical/computer modeling. Field monitoring may be the most appropriate method, but its use is limited due to high costs and extreme spatial and temporal ecosystem variability. Mathematical models provide an alternative to field monitoring that can potentially save time, reduce cost, and minimize the need for testing management alternatives. However, the uncertainty of the model results is a major concern. Uncertainty is defined as the estimated amount by which an observed or calculated value may depart from the true value, and it has important policy, regulatory, and management implications. The source and magnitude of uncertainty and its impact on TMDL assessment has not been studied in depth. This article describes the collective experience of scientists and engineers in the assessment of uncertainty associated with TMDL models. It reviews sources of uncertainty (e.g., input variability, model algorithms, model calibration data, and scale), methods of uncertainty evaluation (e.g., first-order approximation, mean value first-order reliability method, Monte Carlo, Latin hypercube sampling with constrained Monte Carlo, and generalized likelihood uncertainty estimation), and strategies for communicating uncertainty in TMDL models to users. Four case studies are presented to highlight uncertainty quantification in TMDL models. Results indicate that uncertainty in TMDL models is a real issue and should be taken into consideration not only during the TMDL assessment phase, but also in the design of BMPs during the TMDL implementation phase. First-order error (FOE) analysis and Monte Carlo simulation (MCS) or any modified versions of these two basic methods may be used to assess uncertainty. This collective study concludes that a more scientific method to account for uncertainty would be to develop uncertainty probability distribution functions and transfer such uncertainties to TMDL load allocation through the margin of safety component, which is selected arbitrarily at the present time. It is proposed that explicit quantification of uncertainty be made an integral part of the TMDL process. This will benefit private industry, the scientific community, regulatory agencies, and action agencies involved with TMDL development and implementation.*

Keywords. *Latin hypercube sampling, Margin of safety, Monte Carlo simulation, TMDL, Uncertainty.*

Uncertainty is defined as the estimated amount by which an observed or calculated value may depart from the true value (Lapedes, 1978). There is a degree of uncertainty associated with almost all predictive models and measured data. Of greater significance, however, is the uncertainty that arises when only the average values of data that do not characterize the system (e.g., infrequently collected grab samples, short-duration intensive

storm sampling, samples collected at limited locations, etc.) are used to describe highly stochastic and heterogeneous problems. Several researchers have associated sources of error in model applications with measurement errors in input parameter values, model algorithms, scale, spatial heterogeneity, and initial and boundary conditions (Haan, 1989; Beven, 1989; Luis and McLaughlin, 1992; Sohrabi et al., 2003). Beven (1989) reported that defining a consistent effective pa-

Submitted for review in February 2006 as manuscript number SW 6354; approved for publication by the Soil & Water Division of ASABE in July 2006.

The authors are **Adel Shirmohammadi, ASABE Fellow**, Professor, Fischel Department of Bioengineering, University of Maryland, College Park, Maryland; **Indrajeet Chaubey, ASABE Member**, Associate Professor, Department of Biological and Agricultural Engineering, University of Arkansas, Fayetteville, Arkansas; **R. Daren Harmel, ASABE Member**, Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; **David D. Bosch, ASABE Member**, Hydraulic Engineer, USDA-ARS Southeast Watershed Research Laboratory, Tifton, Georgia; **Rafael Muñoz-Carpena, ASABE Member**, Associate Professor, Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida; **Cecil Dharmasri**, Project Engineer, Syngenta Crop Protection, Inc., Greensboro, North Carolina; **Aisha M. Sexton, ASABE Student Member**, PhD Candidate, Fischel Department of Bioengineering, University of Maryland, College Park, Maryland; **Mazdak Arabi, ASABE Member**, Post-Doctoral Research Assistant, Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana; **Mary Leigh Wolfe, ASABE Member**, Associate Professor, Department of Biological Systems Engineering, Virginia Tech, Blacksburg, Virginia; **Jane Frankenberger, ASABE Member**, Associate Professor, Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana; **Carrie Graff**, Post-Doctoral Research Associate, USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland; and **Teymour M. Sohrabi, ASABE Member**, Professor, Department of Irrigation Engineering, University of Tehran, Karaj, Iran. **Corresponding author:** Adel Shirmohammadi, Fischel Department of Bioengineering, University of Maryland, College Park, MD 20742; phone: 301-405-1185; fax: 301-314-9023; e-mail: ashirmo@umd.edu.

parameter value to reproduce the response of a spatially variable pattern of parameter values was not possible. Sohrabi et al. (2003) stated, "Although the introduction of stochastic methods of analysis into the process calculations cannot reduce the inherent uncertainty of a given problem, it can lead to better decision making." Several researchers have also emphasized such analysis as essential for water quality models (Sohrabi et al., 2003; Reckhow, 2003; Chapra, 2003; NRC, 2001; Haan et al., 1995; Beck, 1987).

Congress established the TMDL program in the original Clean Water Act of 1972, Section 303(d) (U.S. Congress, 1972). Most of the effort by EPA and states between 1972 to the 1990s focused on implementing the point-source pollution provision of the Act through the National Pollutant Discharge Elimination System (NPDES), (EPA, 2005; Novotny, 2003). The EPA did not publish any guidelines for state implementation of Section 303(d) addressing nonpoint-source pollution (NPS) until 1991 despite the fact that Section 208 of the 1972 Act had acknowledged the need. The U.S. District Court for Northern District of California reaffirmed the TMDL program in its handling of NPS pollution and stated that states have the responsibility to implement TMDLs on impaired water bodies (EPA, 2005). In 2001, the validity of the TMDL process was reaffirmed after Congress requested a committee to assess the scientific basis to reduce water pollution (Novotny, 2003).

Currently, two methods are available for tracking pollution in the environment and assessing the effectiveness of the TMDL process in improving water quality of impaired water bodies: field monitoring and mathematical/computer modeling. Field monitoring may be the most appropriate and valuable method to support TMDL development, but its use is limited due to high costs and extreme spatial and temporal ecosystem variability. Therefore, mathematical models provide an alternative to monitoring and can save time, reduce cost, and minimize the need for testing management alternatives. Models can be used to assess and develop TMDL plans on large watersheds where broad-scale monitoring is not practical and can provide future forecasts on the impact of TMDL implementation. However, the uncertainty of mathematical model simulation results is a concern. The issue of uncertainty has important policy, regulatory, and management implications. Sexton et al. (2005) and NRC (2001) indicate that uncertainty in the TMDL process is considered under the MOS (margin of safety) component, which is an arbitrary load added to the allocated point-source and nonpoint-source loads to represent the TMDL for a water body:

$$\text{TMDL} = \text{WLA} + \text{LA} + \text{MOS} \quad (1)$$

where TMDL is the total maximum daily load, WLA is the waste load allocation for point sources, LA is the load allocation for nonpoint sources, and MOS is the margin of safety. However, quantifying the magnitude and impact of uncertainty and accounting for that uncertainty in the TMDL process has not been well studied or implemented in practice.

The objectives of this article are to review and synthesize possible sources of uncertainty in simulated outputs of TMDL models, to quantify such uncertainties, and to propose approaches for better incorporation of uncertainty into the TMDL process. In addition, four case studies are presented to illustrate different uncertainty analysis methods and applications in the TMDL process.

SOURCES OF UNCERTAINTY

The watershed models used for TMDL development (whether lumped or distributed parameter models) can be mathematically represented as (Haan, 1989):

$$\underline{Q} = \underline{f}(\underline{I}, \underline{P}, t) + \underline{e} \quad (2)$$

where, \underline{Q} is an $n \times k$ matrix of watershed response to be modeled, \underline{f} is a collection of functional relationships, \underline{I} is an $n \times m$ matrix of inputs, \underline{P} is a vector of p parameters, t is the temporal scale of prediction/calculation, \underline{e} is an $n \times k$ matrix of errors, n is the number of data points, k is the number of outputs, and m is the number of inputs. Even though the distinction between \underline{I} and \underline{P} is not always clear, \underline{I} generally represents a known or a measured model input value, and \underline{P} represents a parameter that is estimated based on some functional relationship with the known inputs (Haan, 1989). Equation 2 indicates that sources of uncertainty may be due to input variability (i.e., parameterization), algorithm selection in defining the simulated processes, accuracy and level of available observed functional data for model calibration and validation, specification of initial and boundary conditions, and scale of application (Beven, 1989).

Although watershed models are developed to mimic natural processes, actual watershed processes are more complex and variable than what can be represented in most sophisticated models (Haan et al., 1995). To understand an event in our natural environment, we may need to provide a scientific explanation of it by inferring a set of general laws or theoretical principles (L_1, L_2, \dots, L_n) and a set of empirical circumstances (C_1, C_2, \dots, C_n) (Woolhiser and Brakensiek, 1982). Each of these laws or empirical circumstances relies on several parameters both to describe the processes and to accommodate spatial and temporal variability. However, reproducing natural complexity challenges modelers and their models in terms of prediction certainty or lack of it. Several researchers have reported sources of the error (e) in model applications (e.g., Haan, 1989; Luis and McLaughlin, 1992). In general, these include measurement error, model error, and spatial heterogeneity. In addition, Beven (1989) concluded that defining a consistent effective parameter value to reproduce the response of a spatially variable pattern of parameter values was not possible.

The number of parameters needed to run the models supporting TMDLs is usually large. Most of these models require spatially distributed information on soils, climate, cropping practices, crop growth parameters, chemical inputs (fertilizer and pesticides), as well as a variety of estimated physical properties such as Manning's roughness coefficient. Because many of these parameters may be sensitive, uncertain, and difficult to measure or estimate, mean values are often used, which are not always representative of the mean behavior of the system (Sohrabi et al., 2003). Additionally, these models need to be parameterized in space, both horizontally and vertically, and in time, so populating the input databases can require substantial measurement and processing resources. Uncertainty can arise from using default values that may or may not be appropriate for the area being modeled and also from transforming values between units.

AnnAGNPS (Annualized Agricultural Nonpoint-Source Pollution) (Yu et al., 2001) and SWAT (Soil Water Assessment Tool) (Arnold et al., 1998) are both examples of

spatially based models frequently used for TMDL development that rely on accurate input on the spatial variability of soils data both laterally and vertically in the environment. Many of the spatial soils databases such as SSURGO are incomplete when it comes to identifying all the parameters within the different layers, so estimates need to be made to fill in the data gaps. Furthermore, estimating parameters that are sensitive to the model output can add to the uncertainty, where small changes in the sensitive input parameter lead to large changes in output. Chu et al. (2004) identified biological and uptake parameters as sensitive parameters that contributed to the sensitivity of nutrient and sediment output in the SWAT model. In the WEPP (Water Erosion Prediction Program) model, slope, hydraulic roughness, Manning's n, and soil erodibility and infiltration parameters were also found to contribute significantly to nutrient and sediment outputs (Baffaut et al., 1997). A sensitivity analysis of AGNPS found that many of the most sensitive parameters were those contained in the Universal Soil Loss Equation and also slope and curve number (Yu et al., 2001). Many of the above sensitive parameters are not directly measurable, nor is it possible to collect a large, random sample to describe the uncertainties (Haan et al., 1995). With the many input parameters required to run these models, it is even more important to accurately quantify uncertainty.

UNCERTAINTY DUE TO INPUT VARIABILITY

Much of the uncertainty associated with modeling results can be traced back to uncertainty in the estimates of the parameters used as input for the analysis. Variability in these parameters within and between various land uses and soil types is an important consideration if realistic description of the hydrologic behavior is to be made (Peck et al., 1977). Sohrabi et al. (2003) provided appropriate probability distributions for each of the most sensitive hydrologic and chemical input parameters of SWAT2000 and showed the combined effect of such variability on model output in the form of cumulative probability distribution curves. Variability in model output caused by inaccurate input estimates can be reduced by decreasing the uncertainty in the inputs through two methods: (1) increasing the number of measurements of the parameter and thus increasing confidence in the

mean estimate, and (2) improving the methods used to measure the parameter (Haan and Skaggs, 2003). A brief discussion of the level of uncertainty in some of the common model inputs is provided below.

Climatic Data

Climatic data are required inputs for virtually all hydrologic and water quality models. At a minimum, these data include precipitation, air temperatures, and solar radiation, which can all exhibit considerable variability based on measurement methods. Precipitation data are frequently collected with tipping bucket type devices, which typically have an associated error of 1% to 5%, or weighing lysimeters, which can have up to 2% error. Rainfall measured from NEXRAD radar can have very large errors (Hardegee et al., 2003). Because precipitation is one of the most critical input characteristics to any hydrologic simulation, these errors can have significant impact on the accuracy of the results. Furthermore, the natural distribution of rainfall is only captured with a dense network of sensors, which is not typically available. The impact of variability in other climatic data, such as temperature and solar radiation, on simulation results is similarly important and should be considered in uncertainty analysis.

Soils Data

Most water quality models rely heavily on the accurate input of soil characteristics and often include estimates of soil hydraulic conductivity, texture, bulk density, and water holding capacity. While well-accepted methods have been established for these measurements, considerable uncertainty can be introduced based on personnel expertise and equipment. In addition to measurement error, soil characteristics tend to be highly spatially variable, and inaccuracies related to where soils are located on the earth's surface lead to another form of error. Examples of the typical spatial variability of several parameters are presented in table 1. Parameters associated with the flow of solutes through the soils (e.g., apparent diffusion coefficient and pore water velocity) typically have much higher CVs than do parameters associated only with the flow of water through the soil (table 1).

Table 1. Degree of variability of various soil physical and chemical characteristics (after Warrick and Nielsen, 1980).

Degree of Variability	Parameter	Mean	SD	CV (%)
Low	Bulk density (g cm ⁻³)	1.3	0.09	6.9
	Bulk density (g cm ⁻³)	1.4	0.10	6.8
	Bulk density (g cm ⁻³)	1.5	0.11	7.3
	Water content at zero tension (cm ³ cm ⁻³)	40-45	4.5-4.8	11
	Water content at zero tension (cm ³ cm ⁻³)	47	4.8	10
Medium	Sand/silt/clay fractions (%)	53/28/19	15/9.1/6.8	28/32/36
	Sand/silt/clay fractions (%)	26/27/47	11/6/8	42/22/17
	Water content at 0.1/15 bar tension (g g ⁻¹)	27/9.5	5.4/3.1	20/33
	Water content at 0.1/15 bar tension (g g ⁻¹)	23/7.5	9.2/3.8	40/51
	Water content at 0.2 bar tension (cm ³ cm ⁻³)	32	5.4	17
	Water content at 2.2 bar tension (cm ³ cm ⁻³)	34	4.1	12
High	Water content at 15 bar tension (g ³ g ⁻³)	4.5	1.4	31
	Saturated hydraulic conductivity (cm h ⁻¹)	14	26	190
	Saturated hydraulic conductivity (cm day ⁻¹)	20	22	110
	Unsaturated hydraulic conductivity at 90%/60% saturation (cm day ⁻¹)	0.63/0.0026	1.75/0.01	280/420
	Apparent diffusion coefficient (cm ² day ⁻¹)	370	2.4×10 ⁶	6.5×10 ⁶
Pore water velocity from water/solute (cm day ⁻¹)	44/40	7300/4400	1.7×10 ⁴ /1.1×10 ⁴	

The level of variability of other hydrologic and water quality input parameters, for example, those associated with watershed models such as SWAT2000, is described in case study II.

UNCERTAINTY DUE TO MODEL ALGORITHMS

Shirmohammadi et al. (2001) provide a detailed discussion of modeling concepts, modeling philosophy, model classification, and types of water quality models. Most of the models proposed or used for TMDL assessment (e.g., SWAT, AnnAGNPS, HSPF, etc.) fall under the “Formal (Mathematical Models)” category and often include both empirical and theoretical (or physically based) algorithms. Empirical algorithms are regression equations that are developed based on a set of observed data for defined climatic, soils, and land use scenarios. Thus, their application to conditions other than those for which these algorithms were developed may produce outputs with significant uncertainty. On the other hand, theoretical models are developed based on certain physical laws, and they can be applied to diverse climatic and physiographic regions given proper input values for the parameters of interest. In addition, theoretical models require well-replicated, long-term observed data for calibration and validation. Uncertainty due to the type of algorithms used in a given model can significantly impact the accuracy of model outputs; therefore, it is essential to identify possible uncertainties associated with the model algorithms/structure used in TMDL assessment.

UNCERTAINTY DUE TO MODEL CALIBRATION AND VALIDATION DATA

The most common approach to minimize model output uncertainty is through model calibration (training), in which model parameters are adjusted so that model predictions match measured data within a predefined accuracy level. Once a model is calibrated, we assume that it accurately represents the watershed processes under investigation. However, the uncertainty inherent in calibration and validation data should also be included in the overall assessment of model uncertainty and can be broken down into four

procedural categories: discharge measurement, sample collection, sample preservation and storage, and laboratory analysis (Harmel et al., 2006).

In their study, Harmel et al. (2006) created various “data quality” scenarios representative of measured data sets used in model calibration and validation. The scenarios ranged from a best case with outstanding quality assurance/quality control (QA/QC) and ideal hydrologic conditions to a worst case with little QA/QC and difficult hydrologic conditions. A range of typical scenarios with moderate QA/QC and typical hydrologic conditions was also examined. For these scenarios, the authors compared errors introduced by each procedural category and estimated the uncertainty in measured streamflow and load data. The root mean square error propagation method of Topping (1972) was used to estimate the cumulative probable uncertainty:

$$E_p = \sqrt{\sum_{i=1}^n (E_1^2 + E_2^2 + E_3^2 + \dots + E_n^2)} \quad (3)$$

where E_p is the probable error range ($\pm \%$), n is the total number of sources of potential error, and E_1, E_2, \dots, E_n are the potential sources of error ($\pm \%$).

Results indicated that substantial error can be introduced by each of the procedural categories (fig. 1). In a worst-case scenario, sample collection, sample preservation and storage, and laboratory analysis can contribute more uncertainty than streamflow measurement. In the range of typical scenarios, sample collection had the greatest potential to contribute large amounts of uncertainty, but streamflow measurement, sample preservation/storage, and laboratory analysis also contributed considerable error. In the best-case scenario, the uncertainty was less than 5% for each procedural category. Based on these results, a change is needed in QA/QC methodology for storm water data so that streamflow measurement and sample collection receive the same attention as typically given to sample preservation/storage and laboratory analysis. In fact, sample collection may deserve the most attention, as it can introduce the most uncertainty in typical scenarios (fig. 1).

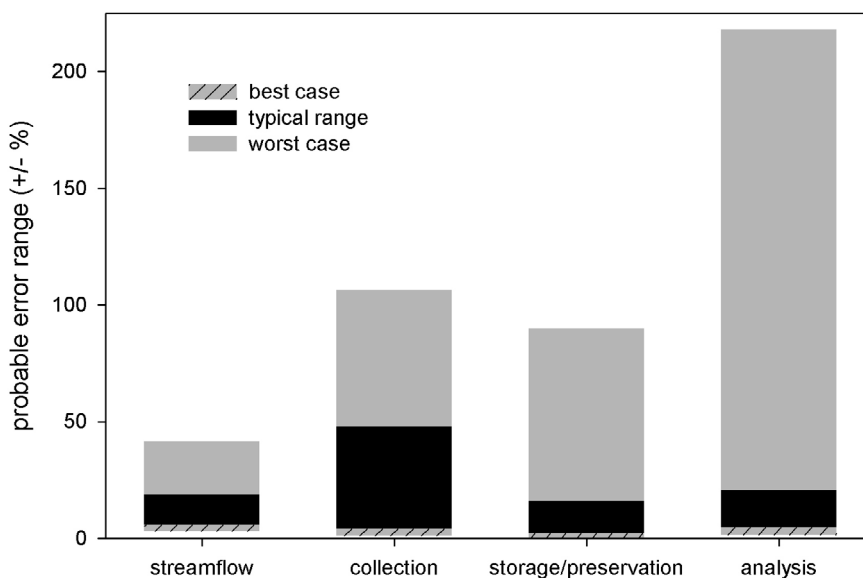


Figure 1. Comparison of probable uncertainty contributed by each procedural category for best case, typical, and worst case “data quality” scenarios; data presented were averaged across constituent type (Harmel et al., 2006).

Table 2. Cumulative probable uncertainty, represented by the probable error range (E_P), for streamflow and nutrient and sediment storm loads for worst case, best case, and typical scenarios (Harmel et al., 2006).

	Streamflow (%)	NO ₃ -N (%)	NH ₄ -N (%)	Total N (%)	Dissolved P (%)	Total P (%)	TSS (%)
Worst-case scenario	42	421	246	168	417	249	117
Typical scenario maximum	19	69	100	70	104	110	53
Typical scenario average	10	17	31	29	23	30	18
Typical scenario minimum	6	8	11	11	11	8	7
Best-case scenario	3	4	3	6	4	3	3

When the uncertainty introduced by each procedural category was propagated to the resulting streamflow and water quality data, results indicated that streamflow is much less uncertain than nutrient and sediment loads (Harmel et al., 2006). Estimated E_P values for streamflow were relatively low, ranging from 6% to 19% under typical conditions and only reaching 42% in the worst-case scenario (table 2). In contrast, estimated E_P values for dissolved nutrients ranged from 8% to 104%, E_P for total nutrients ranged from 8% to 110%, and E_P for TSS ranged from 7% to 53% under typical conditions and from 117% to 421% in the worst-case scenario.

When storm water quality concentrations were examined instead of constituent loads, the uncertainty was reduced only slightly. The reduction was relatively small because streamflow measurement, which is unnecessary for concentration measurement, typically contributes less uncertainty than the other procedures. In contrast, measured baseflow constituent concentrations, which do not require streamflow measurement or storm water sample collection, experienced substantial E_P reductions, ranging from 15% to 83% in the worst-case scenario, from 3% to 35% in typical scenarios, and from 1% to 3% in the best-case scenario.

Most TMDL models have several parameters that can be adjusted to fit measured watershed response data. The model parameters resulting from the calibration process generally apply only to that watershed and corresponding measured data. Even when parameter estimates are available for a watershed of interest, the estimates should be treated as random variables since their values depend on observed data, which themselves are random variables (Haan, 1989). Since any function of a random variable is also a random variable, model outputs should be viewed as random variables each having a probabilistic structure. Thus, model outputs should be described with a probability density function and confidence intervals (Haan, 1989). Such a description helps quantify the range of uncertainty in model outputs (Haan et al., 1995). With uncertainty estimates for both measured data and model outputs, more rigorous, scientifically defensible, and potentially cost-effective TMDL assessments will result.

UNCERTAINTY DUE TO SCALE

With developments in computing technology and information management systems and increased accessibility of spatial data, GIS-integrated modeling frameworks have been widely used in TMDL assessment. The role of GIS varies from developing spatial data for modeling inputs to running models in part or in full on a GIS platform (Shirmohammadi et al., 2001). These developments greatly benefit pre-processing of input data, batch processing a large number of model runs, and post-processing of model outputs and visualization. However, with the excitement of the technology and its convenience, users may overlook factors contribut-

ing to the uncertainty of model predictions that are introduced by these techniques.

Issues related to scale in TMDL assessment can be divided into two major categories. One involves model-related issues, where the simulated processes that were developed to run at a certain scale or a range of scales are not consistent with the spatial units of the model being applied (i.e., plot scale, landscape level, watershed level, etc.), (Shirmohammadi et al., 2001). The second category involves the spatial data used in generating model inputs, including digital elevation models (DEMs), land use and land cover (LULC) data, soils data, stream networks, and weather data.

Hartkamp et al. (1999) also discussed issues related to integrating GIS and modeling by defining two classes of uncertainty related to spatial data: positional and thematic. Spatial errors can arise from several sources: the spatial unit of the input data layers, or the aggregation and generalization processes used for the data development. Scaling can also be a form of positional error where the lack of appropriate spatial resolution for the overlaying input data layers that are integrated within the model framework. Additional sources of uncertainty arise when soils input data are generated from digital soil surveys within GIS (i.e., DiLuzio et al., 2004). This is mainly caused by a one-to-many relationship between mapping units and soil components in the digital soil surveys. As a result, the user may generate a manipulated soil attribute data file that may not be representative of the original data.

DEM data are available at various scales, and the drainage networks derived from them within GIS can vary widely. To quantify the effect of GIS data resolution, Cotter et al. (2003) evaluated uncertainty in flow, sediment, NO₃-N, and TP loads from an agricultural watershed as predicted by the SWAT model. The authors evaluated seven different spatial resolutions of DEM, land use, and soils data (30 × 30 m, 100 × 100 m, 150 × 150 m, 200 × 200 m, 300 × 300 m, 500 × 500 m, and 1000 × 1000 m) and found that the SWAT model output was most affected by input DEM data resolution (fig. 2), where a decrease in resolution resulted in smaller watershed area and slope and larger slope length. Similarly, land use data resolution was found to affect distribution of pasture, forest, and urban areas within the watershed. However, soils data resolution was found to affect only sediment and TP (total phosphorus) predictions for the range of resolutions evaluated. Queija et al. (2005) pointed out that DEM data will be available at a wide range of resolutions and users will be able to access the appropriate data sets for the region of interest. Data sources such as LIDAR (light detection and ranging) will provide fine-resolution DEMs and will be helpful in reducing the uncertainty associated with coarser-resolution DEMs.

The accuracy of the DEM greatly depends on terrain morphology, sampling density, and the interpolation method (Aguilar et al., 2005) used to derive elevation between

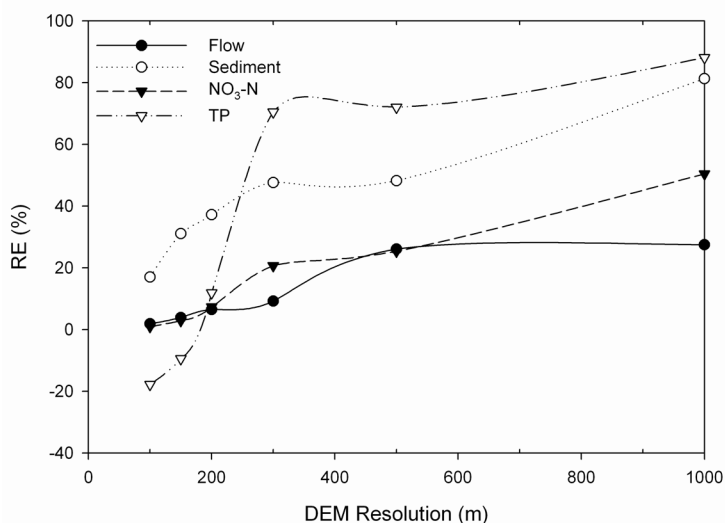


Figure 2. Effect of DEM data resolution on uncertainty in flow, sediment, NO₃-N, and TP load predicted by the SWAT model, represented by percent relative error (RE %) (Cotter et al., 2003).

sampling locations. During the DEM data development process, depressions can be formed within the landscape, where the number of depressions is inversely related to scale (Lindsay and Creed, 2005). Some of the depressions may be artifacts of the process, but the actual depressions should be accurately captured within the data. Uncertainties in digital terrain modeling have been studied widely (i.e., Schneider, 2001; Hugentobler, 2001). According to Wise (2000, 2001), the DEM quality may influence uncertainty more at small scales than at large scales. Desmet (1997) cautioned on the use of different interpolation methods in generating DEMs and demonstrated that same source data could result in erosion estimates varying by orders of magnitude.

UNCERTAINTY EVALUATION METHODS

The most common strategies for quantifying uncertainty in mathematical models used in watershed-scale water quality analysis are Monte Carlo (MC) simulation and first-order error (FOE) or first-order approximation (FOA) analysis. The majority of other types of approaches have been derived from these initial methods. Examples of these derivations include the Latin hypercube sampling (LHS) method (McKay et al., 1979), the mean value first-order reliability analysis method (MFORM) (Madsen et al., 1986), the mean value first-order second-moment (MFOSM) method, the advanced mean value first-order second-moment (AFOSM) method (Hasofer and Lind, 1974), and the mean value second-order (MSO) method (Mailhot and Villeneuve, 2003). Newly introduced methods, whether they have resurfaced after many years or are newly developed, include Bayesian analysis (e.g., Borsuk and Stow, 2000; Franks and Beven, 1997), a probabilistic approach by Borsuk et al. (2002), and the deterministic equivalent modeling method (DEMM) (Cryer and Applequist, 2003a, 2003b). Selected methods appropriate for uncertainty analysis in TMDL models are discussed below.

FIRST-ORDER APPROXIMATION (FOA)

The FOA method can be used to estimate the amount of uncertainty in the output (O) as a function of uncertainty in parameter P using the method described by Haan (2002). In general, the expected or mean value of O can be estimated as:

$$E[O] \approx f(\bar{P}) \quad (4)$$

and the variance in the model output can be estimated by:

$$\text{Var}(O) = \sigma_o^2 \approx \sum_{i=1}^n \left[\left(\frac{\partial f}{\partial P_i} \right)_{\bar{P}_i} \sigma_{P_i} \right]^2 \quad (5)$$

where O is the output of interest modeled as a function of a vector of n parameters (\underline{P}), σ_o^2 is the variance in output, and σ_p is the standard deviation of P . Summers et al. (1993) discussed the strengths and limitations of FOA analyses in water quality modeling. Generally, FOA analysis under small parameter uncertainty is considered to give reasonable estimates of model output uncertainty (Zhang et al., 1993).

MEAN VALUE FIRST-ORDER RELIABILITY METHOD (MFORM)

MFORM is an uncertainty analysis approach that allows the user to express uncertainty in terms of variance and is a developed form of FOA (Melching and Yoon, 1996). Variance is an indication of the closeness of the values of a sample or population to the mean. MFORM allows the user to determine the variance of the dependent variable as well as the variance contributed by each input parameter (basic variable). Each basic variable should be standardized to receive equal consideration. By determining the parameters contributing to the most uncertainty, one can go back and re-evaluate those parameters to determine their values with greater certainty. This provides the user with updated model calibration and less output uncertainty. The model can then be run to determine the uncertainty of results based on validation data. Results are usually expressed in terms of

variance, probability (of failure), confidence intervals, or other descriptive statistics (e.g., coefficient of variation).

MFORM is derived by performing a Taylor series expansion of the model output function as:

$$Y = g(X_e) + \sum_{i=1}^n (x_i - x_{ie}) \left(\frac{\partial g}{\partial x_i} \right)_{X_e} \quad (6)$$

where Y is the dependent variable or model output of interest, $g(\cdot)$ is the function representing the simulation process (algorithms, set of equations) to obtain Y , X_e is the vector of basic variables at the expansion point, n is the number of basic variables x_i , x_{ie} is the vector value of basic variable, and $\partial g/\partial x_i$ represents the rate of change of the model output with respect to a unit change in each basic variable, usually referred to as the sensitivity coefficient.

In MFORM, the expansion point is at the mean value of the basic variables. Therefore, the mean and variance of the dependent variable can be approximated as:

$$E(Y) \approx g(X_m) \quad (7)$$

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 \approx & \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i} \right)_{X_m}^2 \sigma_i^2 \\ & + 2 \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial g}{\partial x_i} \right)_{X_m} \left(\frac{\partial g}{\partial x_j} \right)_{X_m} \\ & \times C_v(x_i, x_j) \end{aligned} \quad (8)$$

where $E(Y)$ is the expected value (mean) of random variable Y , X_m is the vector of basic variables at the mean values, σ_i^2 is the variance of basic variable i , $C_v(x_i, x_j)$ is the covariance of basic variables i and j , and all other variables are previously defined. The first term represents the variance of statistically independent parameters, while the second term is used to tabulate the variance of correlated parameters. $C_v(x_i, x_j)$ can be tabulated by using the identity:

$$C_v(x_i, x_j) = E[(x_i - x_{mi})(x_j - x_{mj})] \quad (9)$$

where x_{mi} is the mean value of all x_i values, and x_{mj} is the mean value of all x_j values.

If basic variables are not correlated, then $C_v(x_i, x_j)$ is equal to zero. In this case, the variance of output can be written as:

$$\text{Var}(Y) = \sigma_Y^2 \approx \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i} \right)_{X_m}^2 \sigma_i^2 \quad (10)$$

This term represents the fraction of model output variance (FOV) contributed by each basic variable. When using complex models, the best way to solve for $\partial g/\partial x_i$ (eqs. 8 and 10) is by using numerical methods. Melching and Bauwens (2001) tabulated $\partial g/\partial x_i$ using forward difference with change in x_i equal to 0.01. The unit change of x_i depends on the sensitivity of the model to change in parameters.

Melching and Bauwens (2001) evaluated uncertainty in coupled nonpoint source and stream water quality models applied to a suburban watershed. They used LHS and mean value first-order reliability methods (MFORM) to determine

prediction uncertainty of dissolved oxygen (DO) concentrations. LHS was used to identify the basic variables that significantly contribute to output uncertainty, while MFORM was used to provide estimates of the percentage contribution of the variables to output uncertainty. In LHS, input parameters were ranked in terms of their correlation with output values to determine their importance. The coupled models included a nonpoint-source loading model, a constant treatment efficiency model, and a river water quality model. General conclusions could not be made about the overall uncertainty of the system because of the limitations inherent in each individual model. However, the study was able to help identify key sources of uncertainty, i.e., the main parameters that significantly affect the uncertainty in simulated DO concentrations at the location most prone to low concentrations.

MONTE CARLO SIMULATION (MCS)

In MCS analysis, the effect of uncertainty in model parameter P on output O is estimated by repeated simulations using randomly selected parameter values. Effects of uncertain knowledge of one or more parameter values can be reliably estimated using MCS analysis. MCS analysis has been reported to be the most robust method for estimating uncertainty in water quality models (Hession et al., 1996) and is commonly selected as a standard for comparing against other methods (Yu et al., 2001). Haan et al. (1995) proposed methods to validate water quality models using MCS analysis. The accuracy of output uncertainty estimates depends on the number of model simulations performed and on the adequacy of the assumed parameter distribution (Haan, 2002). The number of model simulations should be sufficiently large to reliably estimate the probability distribution of the output variables (Gardner and O'Neill, 1983). However, this requires considerable computing resources. In addition, the distribution of many of the parameters used in water quality models is not known. If the model parameters under study are correlated, a multivariate simulation of the parameters must be used.

LATIN HYPERCUBE SIMULATION (LHS) WITH CONSTRAINED MCS

When the model complexity under study is such that computational requirements for MCS analysis become prohibitive, the LHS method can be used to efficiently estimate uncertainty in the model output (Iman and Shorten-carier, 1985; Iman et al., 1980). LHS is a stratified sampling approach in which the probability distribution of each parameter is subdivided into n non-overlapping ranges. Each range of parameter value is considered to have equal probability of occurrence. One random value of the parameter under evaluation is selected from each range for each parameter. Then, the model output is obtained for each set of randomly selected parameter values without repeated simulations as is performed with MCS analysis. In LHS analysis, the order of selection of parameter ranges is also randomized and the model is executed n times, corresponding to n number of parameter ranges. The output distribution and statistics can be obtained from the sample of n output values. The LHS technique used to obtain n different values from each of k variables operates as follows:

1. The range of each variable is divided into n non-overlapping intervals on the basis of equal probability.
2. Values are then randomly selected from each interval.
3. The n subintervals of each input variables (X_i) are then randomly permuted relative to other input variables (X_j), so that every combination of subinterval is equally likely.

The choice of sample size (n) in LHS depends on a number of considerations but is dominated by the cost of making a single computer run and by the number of input variables (k). Iman and Helton (1985) found that good results could be obtained with $n > (4/3)k$. This is a significant improvement relative to standard MCS, which requires at least an order of magnitude more runs per input variable.

GENERALIZED LIKELIHOOD UNCERTAINTY ESTIMATION (GLUE)

GLUE (Beven and Binley, 1992) is a sampling method that uses the behavior/non-behavior classification concept (Spear and Hornberger, 1980) to distinguish between the fraction of Monte Carlo simulations that matches the system behavior and the complement fraction that does not. The main argument that led to development of GLUE is that for a selected model structure, several combinations of model parameters (in contrast to a unique optimum) may exist that are equally good for reproducing the system behavior (Freer and Beven, 1996). Beven and Binley (1992) suggested that uncertainties due to errors in model structure and effects of correlation structure in the parameter space are implicitly handled within the GLUE procedure.

In the GLUE method, the relative importance of model parameters is not evaluated individually and independently but as set of parameters. Each set of model parameters within a given model structure is evaluated as a single unit. In other words, it is the parameter set that is important to produce system behavior, not the individual parameters. First, retained behavior-giving Monte Carlo simulations are given weights based on a likelihood function. A likelihood measure based on the Nash-Sutcliffe efficiency criterion with shaping factor N is defined as (Freer and Beven, 1996):

$$L(\alpha|y) = \left(1 - \frac{\sigma_\epsilon^2}{\sigma_o^2}\right)^N ; \sigma_\epsilon^2 < \sigma_o^2 \quad (11)$$

where $L(\alpha|y)$ is the likelihood of parameter set (α) given the observed data (y), and σ_ϵ^2 and σ_o^2 refer to the error variance between model simulations and observed data, and the variance of the observed data, respectively. For $N = 1$, equation 11 is the well-known Nash-Sutcliffe efficiency coefficient (E_{N-S}) that is often used for calibration of hydrologic and water quality models. A negative E_{N-S} value indicates that the corresponding model output is dissimilar to the behavior of the system under study, and the likelihood of such simulation in mimicking the system behavior is zero. Then, the likelihood estimates associated with model simulations are rescaled such that they sum up to 1.0 and are used as likelihood weights to form a cumulative weighted distribution for the output of interest. It should be noted that GLUE has been extensively used in a wide variety of watershed and water quality studies (Zak et al., 1997; Hankin and Beven, 1998; Romanowicz and Beven, 1998; Schulz et al., 1999; Zak and Beven, 1999).

CASE STUDIES

Four case studies are presented to illustrate the use of uncertainty evaluation methods. Three of the case studies focus on uncertainty related to TMDL analysis with the SWAT model. The first two cases show the impact of input parameter uncertainty on uncertainty in model output, with case I using the MCS method and case II using LHS with constrained MCS. Case III uses GLUE to test the SWAT model's output uncertainty due to both input variability and model structure. Case IV was selected to illustrate the evaluation of uncertainty in the design of BMPs during the TMDL implementation phase. This case uses a combination of ranking sensitive input parameters, developing PDFs (probability distribution functions) for those parameters, and finally uses the MCS method to develop a PDF for sediment trapping efficiency under different BMPs.

CASE STUDY I - UNCERTAINTY EVALUATION: SWAT AND MCS ASSESSMENT

Use of MCS analysis in quantifying uncertainty in watershed response prediction is illustrated using the Soil and Water Assessment Tool (SWAT). This model was selected only for illustration, and the framework can be applied to any type of conceptual (black box) models and for any outputs by replacing the parameters of the model.

Description of the SWAT Model

The SWAT model (Arnold et al., 1998) is a physically based, distributed parameter, watershed-scale model. It divides the study watershed into sub-basins and identifies smaller homogeneous areas within each sub-basin called hydrologic response units (HRU) (Arnold et al., 1998; Neitsch et al., 2001). All model calculations are performed at the HRU level. The EPA currently supports this model for developing TMDLs in agricultural watersheds. SWAT consists of three major components: (1) sub-basin, (2) reservoir routing, and (3) channel routing. GIS interfaces have been developed to facilitate the aggregation of input data. This interface requires a land cover map, soils map, and DEM. A detailed description of the model can be found at the SWAT website (www.brc.tamus.edu/swat/index.html).

Description of the MCS Methodology

MCS analysis involves sensitivity analysis to identify the model parameters that have the most influence on predicted outputs, generating a probability distribution of those parameters, running the model using each realization of the model parameter, and generating a probability distribution of model outputs to quantify output uncertainty. A sensitivity analyses may not be needed when the model contains only a few parameters. However, when the number of parameters becomes large, and as more than one output is affected by the same parameter, a sensitivity analysis is needed to identify the model parameters that have the greatest impact on model outputs. A relative sensitivity (S_r) coefficient can be calculated as follows and used to rank model parameters:

$$S_r = \frac{\partial O}{\partial P} \frac{P}{O} \quad (12)$$

where O is output of interest, and P is the parameter under investigation. The S_r calculated using this method is dimensionless and can be used to compare all parameters. One of the limitations of using S_r to identify the most sensitive model

parameters is that it does not account for covariance among parameters. Haan (1989) has argued that the model parameters should be treated as random variables and their uncertainty can be described as PDFs. The PDF of a parameter provides a range and distribution of uncertainty, whereas S_r quantifies the impact of this uncertainty on model outputs.

The outputs of interest in this study were mean annual flow from 2000 to 2002. Runoff from the watershed was found to be most affected by curve number (CN), soil evaporation compensation factor (ESCO), and evaporation coefficient for the stream reach (EVRCH), based on detailed sensitivity analyses (White and Chaubey, 2005). Values of CN are based on physical variables, such as land cover and soil types, and do not follow a specific distribution. Therefore, the required input PDF could not be directly generated from CN. Research has shown that the retention parameter (S) is log-normally distributed (Haan and Schulze, 1987). S is related to CN by:

$$S = \frac{1000}{CN} - 10 \quad (13)$$

The mean and standard deviation (SD) for S were required to describe the lognormal distribution. The mean values of S were assumed to be the default S (calculated from the original CN for each HRU). The SD of S was assumed to be 0.5 of the mean, as suggested by Haan and Schulze (1987).

To simulate these parameters as a distribution, Monte Carlo analysis was employed. Output was generated for 500 model simulations and evaluated as a distribution for each sub-basin/HRU combination (Haan, 2002). The parameter PDF was created by randomly generating 500 values of S for each HRU in the watershed using a lognormal distribution with mean and SD as described above, and then calculating the corresponding CN using equation 13. The

ESCO and EVRCH were assumed to be uniformly distributed, and 500 random values of these parameters were obtained based on range of each parameter distribution. The model was run 500 times under two conditions: (1) assuming that only one parameter was uncertain, and (2) assuming that all three parameters were uncertain. Output PDFs were calculated for flow values and were used to establish a 95% confidence interval (CI) for the output.

Effects of CN uncertainty on predicted mean annual flow for the Illinois River watershed for 2000-2002 is shown as PDFs in figure 3. These PDFs represent a conditional uncertainty, i.e., it is assumed that all other model parameters are known values and are held constant at the beginning of the simulation, and CN is the only uncertain parameter. In figures 3a through 3c, the solid vertical line represents the measured flow data. The PDF of CN in figure 3d shows an area-weighted CN for the entire watershed. Figure 3d also shows the uncertainty in the CN itself as a PDF and with a 95% confidence interval on CN ranged from 62.2 to 65.8. It should be noted that each HRU in the model has a unique curve number. Even though the area-weighted CN does not show much variability, a greater uncertainty in CN at the HRU level can be expected.

Figure 3 also shows that the uncertainty in mean annual flow due to uncertain knowledge of CN was not very high, as indicated by a narrow 95% CI for all three years. Depending on the structure of the model, a small uncertainty in the parameter can result in relatively small or large uncertainty in model output. For applications where a small uncertainty in a parameter results in a relatively large uncertainty in the output, the model structure should be critically examined. If the selected model has a particular parameter that is very uncertain and results in even greater uncertainty in modeled

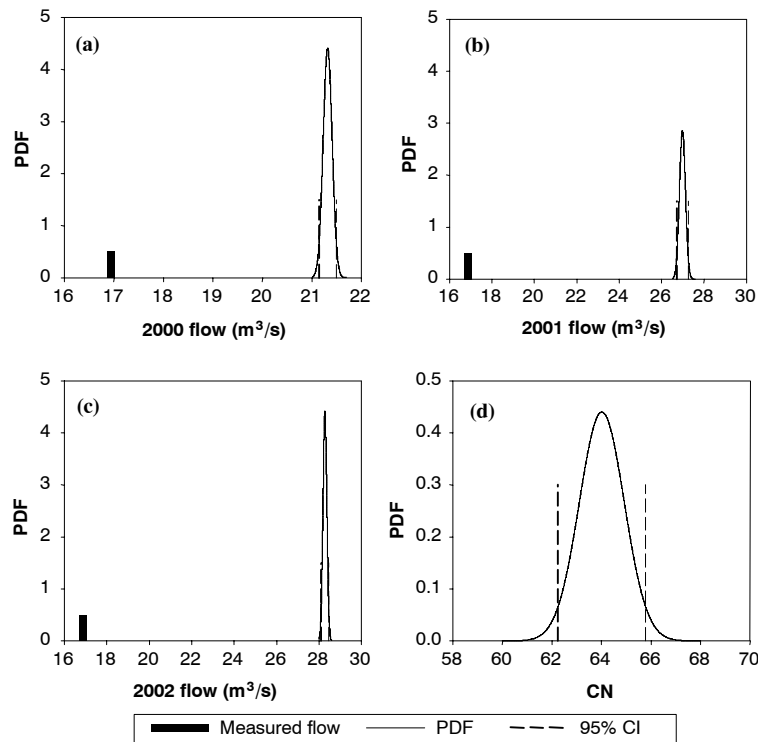


Figure 3. Effects of uncertainty in CN on predicted mean annual flow for the Illinois River watershed. Dashed vertical lines represent 95% confidence interval and solid vertical line represents the measured flow data.

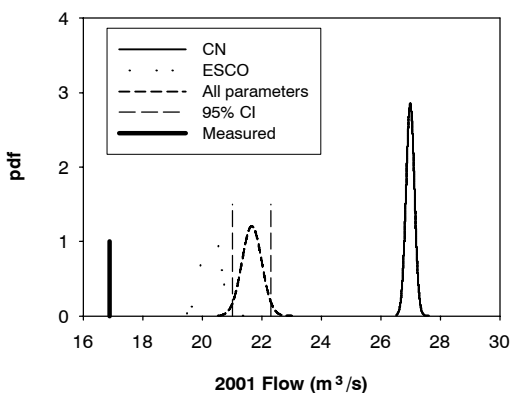


Figure 4. Effect of interactions among parameters on predicted flow for the Illinois River watershed for 2001.

output, such a model should not be used in making critical environmental decisions. In this case, even though CN was found to be one of the most sensitive parameters, the uncertainty in output was not magnified due to uncertainty in CN, as indicated by a narrow CI (fig. 3).

If CN was the only uncertain model parameter, one could examine the measured watershed response at this point to stochastically validate the model. Figure 3 shows that mean annual flow was overestimated for all three years. Since measured flow data fall outside the 95% CI for all years, it can be concluded that the model was not stochastically validated. This indicates that either the model itself or the method of estimating the CN is inadequate, and the model cannot capture the true watershed response at 95% CI. Here it is assumed that the measured data are known with certainty. In reality, there is always some uncertainty present in the measured data (Harmel et al., 2006). Under such conditions, a PDF of the measured data can also be developed and used to calculate the probability that the modeled output will be within a certain tolerance of the measured data.

The uncertainty in the model output can be expected to change when more than one model parameter is uncertain. Figure 4 shows the uncertainty in predicted mean annual flow for 2001 as PDF and cumulative density function (CDF) when CN, ESCO, and EVRCH are all uncertain. In addition, mean, % coefficient of variation (%CV), and 95% CI for flow are shown in table 3 when all three parameters are uncertain, along with the measured watershed response data. It should be noted that both mean and range of uncertainty, as indicated by 95% CI, are different for flow predictions under multiple uncertain parameters.

The range of the 95% CI gives an idea of the practical applicability of the model for TMDL development. Ideally, this range should be as small as possible. A large range in 95% CI suggests that the model outputs are very uncertain and that watershed response predictions are also uncertain,

Table 3. Effects of uncertainty in CN, ESCO, and EVRCH parameters on predicted flow for the Illinois River watershed.

Year	Flow (m ³ s ⁻¹)			Measured
	Mean	%CV	95% CI	
2000	15.87	2.00	15.24-16.51	16.95
2001	21.66	1.50	21.01-22.31	16.88
2002	23.89	0.99	23.43-24.36	16.89

Table 4. Fraction of total variance in predicted flow attributable to each uncertain parameter.

Year	CN	ESCO	EVRCH
2000	0.10	0.89	0.01
2001	0.28	0.71	0.01
2002	0.07	0.92	0.01

even though the model could be stochastically validated. Results from this study indicate that the SWAT model does not suffer from this limitation. It should be noted that the range of uncertainty, as indicated by %CV and 95% CI, was relatively small for the SWAT model in this study, which is desirable for watershed models. For Monte Carlo uncertainty analysis, the fraction of total variance in the model output (F_i) attributable to the i th parameter can be estimated as (Haan, 2002):

$$F_i = \frac{r_i^2}{\sum_{i=1}^n r_i^2} \quad (14)$$

where r_i is the correlation between i th parameter and the output. Table 4 shows the fraction of the total variance in the predicted flow and sediment due to uncertainty in CN, ESCO, and EVRCH. Most of the variance in predicted flow resulted from uncertainty in ESCO.

CASE STUDY II - UNCERTAINTY EVALUATION: SWAT AND LHS WITH CONSTRAINED MCS ASSESSMENT

The MCS analysis combined with LHS was used to evaluate the SWAT2000 model's output variability due to input variability for a small (350 ha) watershed in Frederick County, Maryland. The ARRAMEIS (Advanced Risk and Reliability Assessment Modeling Information System) software package equipped with LHS and MCS was used to perform the uncertainty analysis (Iman et al., 1980).

A total of 34 input parameters (Sohrabi et al., 2003) representing hydrologic and water quality characteristics of the watershed for which SWAT2000 was found to be most sensitive were selected for assessing the impact of their variability on model uncertainty. The first step in probabilistic risk assessment is therefore the determination of appropriate PDFs for the input variables considered in the assessment. Determining the uncertainty PDF to assign to model input parameters is one of the major hurdles in overall evaluation of uncertainty associated with hydrologic and water quality modeling (Haan et al., 1995). This difficulty is compounded by parameter variations with depth due, for example, to the presence of horizons in the soil.

To characterize uncertainty, each input variable was assigned a probability density and a range in values based on professional judgment, SWAT model's default values, and/or literature information (table 5). If the range of values for a parameter was within a factor of 10 or greater, it was assigned a log-uniform distribution, while all other parameters were assigned a uniform distribution.

The field-scale variability in many inputs has been reported to be adequately modeled by normal or log-normal distribution (Jury et al., 1991). In addition, Jensen and Refsgaard (1991) showed soil hydraulic conductivity to be approximately log-normally distributed. Therefore, saturated hydraulic conductivity of soil in different soil layers and soil series and hydraulic conductivity both in the main

Table 5. Description of the distribution and range of values for SWAT2000 input parameters selected for evaluation in the uncertainty analysis (Sohrabi et al., 2003).

Parameter	Distribution	Assigned Range	Units
CMN	Uniform	0.00015-0.00045	--
NPERCO	Log-uniform	0.001-1.00	--
PHOSKD	Uniform	87.5-262.5	m ³ mg ⁻¹
PPERCO	Uniform	10.0-17.0	--
UBN	Uniform	10.0-30.0	--
UBP	Uniform	10.0-30.0	--
USLE_C	Log-uniform	0.03-0.50	--
GW_DELAY	Uniform	5.0-15.0	day
RCHRG_DP	Uniform	0.0-1.0	--
GW_REVAP	Uniform	0.02-0.2	--
GWQMN	Log-uniform	5.00-200.00	mm H ₂ O
OV_N	Log-uniform	0.008-0.480	--
CN2	β	50.0-90.0	--
USLE_P	Uniform	0.25-0.75	--
CH_COV	Uniform	0.0-1.0	--
CH_EROD	Uniform	0.0-1.0	--
CHK_2-C	β	0.05-0.5	mm h ⁻¹
CHK_2-M	β	16.1-50.1	mm h ⁻¹
CHK_2-P	β	5.1-16.0	mm h ⁻¹
CH_N(2)	Uniform	0.025-0.15	--
SOL_AWC(1M)	Uniform	0.10-0.15	mm H ₂ O mm soil ⁻¹
SOL_AWC(1P)	Uniform	0.14-0.20	mm H ₂ O mm soil ⁻¹
SOL_AWC(2M)	Uniform	0.05-0.10	mm H ₂ O mm soil ⁻¹
SOL_AWC(2P)	Uniform	0.07-0.12	mm H ₂ O mm soil ⁻¹
SOL_K(1M)	β	50.8-160.0	mm h ⁻¹
SOL_K(1P)	β	16.0-160.0	mm h ⁻¹
SOL_K(2M)	β	16.0-50.8	mm h ⁻¹
SOL_K(2P)	β	5.1-16.0	mm h ⁻¹
USLE_K	Uniform	0.15-0.45	(see note) ^[a]
CH_K(1C)	β	0.05-0.5	mm h ⁻¹
CH_K(1M)	β	16.1-50.1	mm h ⁻¹
CH_K(1P)	β	5.1-16.0	mm h ⁻¹
CH_N(1)	Uniform	0.025-0.15	--

^[a]USLE_K = (0.013 metric ton m² h)/(m³ metric ton cm).

channel and in tributaries are assumed to have log-normal distributions in this work. Due to the lack of data to estimate

mean and standard deviation for PDFs thought to be Gaussian, the β distribution was used to assign proper values for shape factors to fall into approximate log-normal distribution. In fact, the β distribution is generally used as a rough model in the absence of sufficient data (Wyss and Jorgensen, 1998).

The range for each parameter was selected to cover all physically realistic values. Minimums, maximums, ranges, means, medians, variances, and coefficients of variation for the 50 generated samples for each of the selected model inputs were computed (Sohrabi et al., 2003). Table 6 shows description of the SWAT2000 state variables used in uncertainty analysis in this study.

Table 7 presents statistics of the SWAT2000 model's simulation results using LHS/MCS with variable input parameter values and a comparison of the means with those obtained using average input parameter values, and measured flow and water quality constituents. Data indicate that predicted streamflow values range from 215.0 to 530.0 mm, with a standard deviation of 102.4 mm. The average streamflow predicted using mean input parameter values is above the upper end of this range and certainly higher than the measured value, indicating that inclusion of parameter uncertainty would significantly improve the mean predictions. The fact that the mean of LHS predictions is less than the prediction from mean inputs essentially indicates that negative deviations of input parameters from their mean causes a larger decrease in predicted streamflow than the increase caused by positive deviations of inputs of equal magnitude. Relatively large coefficients of variations in input parameters (10% to 76%) are transformed to a coefficient of variation of only 28.6% in predicted outputs. This underscores the importance of the present MCS approach. Assuming a linear transfer of uncertainty through the model, for example, would have led to a grossly underestimated uncertainty in flow and other results that cover a range of possible outcomes that in reality are highly unlikely.

Knowledge of risk level, along with the corresponding range of expected streamflows, is a clear advantage over knowing only means (which are possibly incorrect, as discussed above). The risk level can, for example, be chosen ei-

Table 6. Description of SWAT2000 state variables used in uncertainty analysis (Sohrabi et al., 2003).

Variable	Description
Flow-related variable	FLOW_OUT
	Average daily streamflow at the outlet of watershed, includes channel losses (m ³ s ⁻¹)
Sediment-related variable	SED_OUT
	Sediment transported with water at the outlet of watershed, includes channel losses (metric ton)
Nutrient-related variables	ORGN_OUT
	Organic nitrogen transported with water at the outlet of watershed (kg N)
	ORGP_OUT
	Organic phosphorus transported with water at the outlet of watershed (kg P)
	NO3_OUT
	Nitrate transported with water at the outlet of watershed (kg N)
	NH4_OUT
	Ammonium transported with water at the outlet of watershed (kg N)
	MINP_OUT
	Mineral phosphorus transported with water at the outlet of watershed (kg P)

Table 7. Values of ranges, means, standard deviations (SD), and coefficient of variations for each model output for one-year (1995) simulations (Sohrabi et al., 2003).

Model Output	Units	Range	Mean	SD	CV (%)	Measured Value
Streamflow	mm	215.0 - 530.0	357.1	102.4	28.6	540.9 ^[a] , 381.3 ^[b]
Sediment	kg ha ⁻¹	4827.0 - 9285.0	9673.0	3469.6	35.8	4595.7
NO ₃	kg ha ⁻¹	10.9 - 226.6	72.5	73.6	101.2	24.9
NH ₄	kg ha ⁻¹	6.6 - 8.4	7.6	0.48	6.3	8.1
ORG-N	kg ha ⁻¹	12.7 - 15.9	14.7	1.01	6.8	Not available
ORG-P	kg ha ⁻¹	2.8 - 12.7	3.7	2.13	57.5	Not available
MIN-P	kg ha ⁻¹	43.1 - 66.5	54.2	6.6	12.2	Not available

^[a] Unadjusted value, includes baseflow contribution from outside the watershed.

^[b] Adjusted value, excludes baseflow contribution from outside the watershed (Chu and Shirmohammadi, 2004).

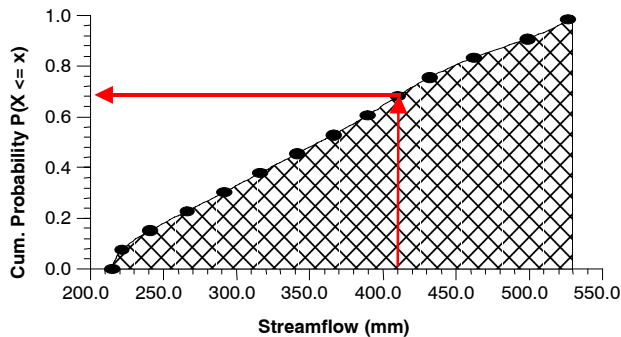


Figure 5. Model output distribution of streamflow at the watershed outlet based on variable input parameter values using LHS/MCS analysis. The arrow shows 9-year average streamflow predicted by SWAT based on average input parameter values (Sohrabi et al., 2003).

ther lower or higher based on the criticality of the intended application of the results and on the significance of any resource it is aimed at protecting or enhancing.

Figure 5 represents the SWAT output cumulative probability distribution of streamflow (Sohrabi et al., 2003). The 9-year average streamflow simulated by the SWAT model based on average input parameter values under BMP4 (strip cropping under no-till) was 411 mm and is depicted by the arrow in figure 5. This 9-year average streamflow corresponds to cumulative probability of 0.7, which means there is 70% confidence that streamflow is equal to or less than 411 mm. In other words, the probability of streamflow being greater than 411 mm is 30%. PDFs similar to figure 5 can provide the reliability of certain quantitative values and their associated uncertainty for each model output, thus helping managers and decision makers assess the results of modeling scenarios.

CASE STUDY III - UNCERTAINTY ANALYSIS OF SWAT'S SEDIMENT YIELD USING GLUE

Figure 6 provides an example of GLUE (Generalized Likelihood Uncertainty Estimation) analysis for sediment yield estimation in the Dreisbach watershed (approx. 6.32 km²) in Indiana. SWAT was utilized to simulate fate and transport of sediment in the watershed. Five thousand SWAT simulations were performed for the 1974-1978 time period when hydrologic and water quality data were collected at the outlet of the watershed. Monte Carlo analysis was used to determine the number of SWAT model simulations in this study, as was performed in case study I. Multiple numbers of model simulations are necessary to cover the range of variability in input parameter values. In addition, the accuracy of uncertainty highly depends on the number of model simulations (Haan, 2002; Gardner and O'Neil, 1983). In figure 6, the 50th percentile of the cumulative GLUE likelihood function was used as the expected value for estimated sediment yield, while 20th and 80th percentiles were used as lower and upper bounds, respectively. It is evident that the observed data (represented by the dashed line) fell well within the uncertainty bounds. To establish a margin of safety for future predictions, the parameter sets that represent the 20th, 50th, and 80th percentiles could be used for SWAT simulations. For the Dreisbach watershed, average monthly sediment yields (t ha⁻¹ month⁻¹) fell in the range of [0.009, 0.0205] with an expected value of 0.014.

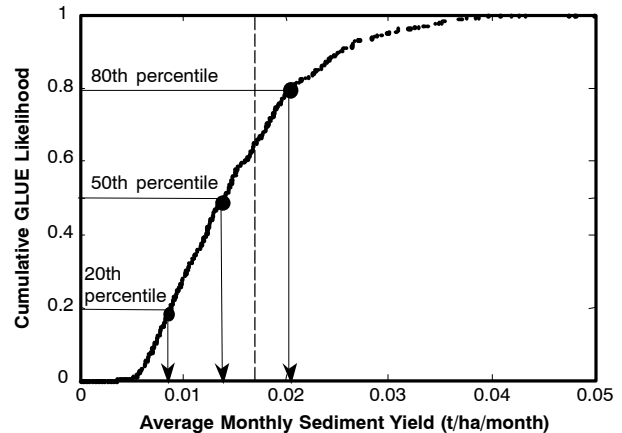


Figure 6. Cumulative GLUE likelihoods for 5000 Latin hypercube sampling simulations for average monthly sediment yield at the outlet of Dreisbach watershed, Indiana. Simulations were performed for 1974-1978 period. The vertical dashed line represents measured data.

CASE STUDY IV - UNCERTAINTY ANALYSIS IN MODEL-BASED BMP DESIGN DURING THE IMPLEMENTATION PHASE OF TMDLS

The work presented so far relates to the assessment of TMDLs using watershed-scale models. This case study demonstrates how uncertainty analysis is also a critical factor to consider in the design of BMPs in the TMDL context. Parsons and Muñoz-Carpena (2001) proposed integrating sensitivity and uncertainty analyses in the modeling and design process of a common BMP, vegetative filter strips (VFS), used in the implementation phase of TMDLs to control runoff and sediment outflow from upslope disturbed areas. The vegetative filter strip modeling design system, VFSSMOD-W, which was developed and tested originally at North Carolina State University (Muñoz-Carpena et al., 1999) and later extended as a design tool (Muñoz-Carpena and Parsons, 2004; Parsons and Muñoz-Carpena, 2002) was modified to enable built-in sensitivity and uncertainty analyses. VFSSMOD-W is a field-scale, mechanistic, storm-based model developed to route the incoming hydrograph and sedigraph from an adjacent field through a VFS and to calculate the resulting outflow, infiltration, and sediment trapping efficiency. A front-end model, UH, generates the necessary source area design inputs for each design storm, including a rainfall hyetograph, a runoff hydrograph, and sediment loss from the source area using a combination of the NRCS curve number method, the unit hydrograph, and the modified Universal Soil Loss Equation. With these inputs, a set of response curves, i.e., sediment and runoff reduction vs. filter construction characteristics, can be developed from VFSSMOD-W outputs for a given design scenario (Muñoz-Carpena and Parsons, 2004).

The response curves can be evaluated with respect to the TMDL plan design goal, i.e., a required sediment reduction expressed in terms of sediment delivery ratio (SDR = sediment out from filter/sediment into the filter), or runoff reduction expressed in terms of runoff delivery ratio (RDR = runoff out from filter/runoff into the filter). The procedure to evaluate the uncertainty for a given design case is based on the following steps (Parsons and Muñoz-Carpena, 2001): (1) identify and rank the input parameters of UH and VFSSMOD relative to their sensitivity on sediment trapping; (2) develop probability density functions for the most

Table 8. Input distributions for MCS used in the uncertainty analysis of a BMP designed to meet a TMDL runoff sediment reduction.

Parameter	Base Value	Distribution	Statistics		
Curve number, CN	85	Triangular	Peak = 85	Min. = 79	Max. = 90
Soil erodibility, K	0.33	Normal	Mean = 0.33	SD = 0.05	
Sat. hydraulic conductivity Green-Ampt, K_{sat} (cm h ⁻¹)	11.99	Lognormal	Mean = 12.0	SD = 3.0	
Initial soil moisture Green-Ampt (cm ³ cm ⁻³)	0.239	Uniform	Min. = 0.05	Max. = 0.25	
Sediment particle class diameter, d_p (μm)	66	Normal	Mean = 66	SD = 10	

sensitive input parameters; and (3) use Monte Carlo simulation to sample the input parameters and develop a probability density function for sediment trapping. This procedure allows the estimation of confidence intervals for the BMP effectiveness to meet a TMDL requirement. In this way, the user can use *a priori* knowledge of local variability and obtain better (or more certain) predictions.

The analysis is illustrated with a typical application in the Piedmont region of North Carolina. An agricultural field is upslope from the planned VFS. The agricultural production system consists of a row crop (with a curve number of 85) growing in sandy clay. The slope of the source area is 2%. Six-hour storms with 1-year (54 mm) and 5-year (85 mm) return periods were selected for evaluation. The VFS parameters were selected to represent a good stand of grass such as fescue. The VFS length was fixed at 5 m for the 1-year storm and 10 m for the 5-year storm and is designed to meet a required TMDL sediment reduction of 80% (SDR = 20%). A sensitivity analysis facilitates the selection of the parameters of the model that should be evaluated with the uncertainty analysis. Table 8 shows the base values for the sensitive parameters along with the PDFs and statistics selected to be representative of the area (Parsons and Muñoz-Carpena, 2001). The objective for selecting the distributions and statistics was to represent possible selections based on the design problem. For example, the initial soil water content was assigned a uniform distribution and allowed to vary randomly between typical field values of 0.05 to 0.25 cm³ cm⁻³ (Muñoz-Carpena et al., 1999).

A total of 3300 simulations (1800 and 1500 for the 1-year and 5-year storms, respectively) were run in the MCS procedure. As an example, figure 7 shows the sampled distributions for curve number and soil erodibility for the 1-year storm. It is apparent that a relatively large number of simulations are needed to capture the desired input PDFs, especially for the normally distributed erodibility input.

The resulting cumulative probability density functions for SDR and RDR are presented in figure 8. The PDFs for RDR do not show drastic probability differences for the tested

scenarios, but the SDR probability density function is shifted left for the 5-year/10 m combination, as compared to the 1-year/5 m combination. Results show that the probabilities of achieving the required SDR of 0.2 or less are 70% and 20% for the 5-year and 1-year scenarios, respectively. Conversely, figure 8 shows that if the same probability of 70% were desired for the 1-year design, the TMDL objective would not be met, since only a sediment reduction of 60% (SDR = 0.4), a 25% loss of efficiency for this BMP, would be achieved.

INCORPORATION AND COMMUNICATION OF UNCERTAINTY TO USERS

The previous sections of this article describe sources of uncertainty and several approaches for quantifying and expressing uncertainty related to the use of models in TMDL development. Incorporation of uncertainty into the TMDL development process is not well-defined or well-practiced. A variety of stakeholders in the TMDL process, including model users developing TMDLs and TMDL implementation plans, decision makers (agency personnel), and watershed residents, could benefit from consideration of uncertainty in the TMDL process.

Awareness of uncertainty and how to determine its effects on modeling results empowers stakeholders to quantify the uncertainty in the parameter that affects them the most. For example, if a large portion of the uncertainty in the model output of interest were due to the estimate of a particular model parameter, e.g., the soil evaporation compensation factor (ESCO) in case study I, then efforts could be focused on reducing the uncertainty of that parameter estimate and thus reducing the uncertainty of the model output. If more of the uncertainty were due to the measured data used for calibration, then decision makers could judge the benefit of collecting more measured data for calibration to reduce the uncertainty attributed to that source. In practice, additional data are collected in some TMDL studies; however, this is typically done without quantitative information on how

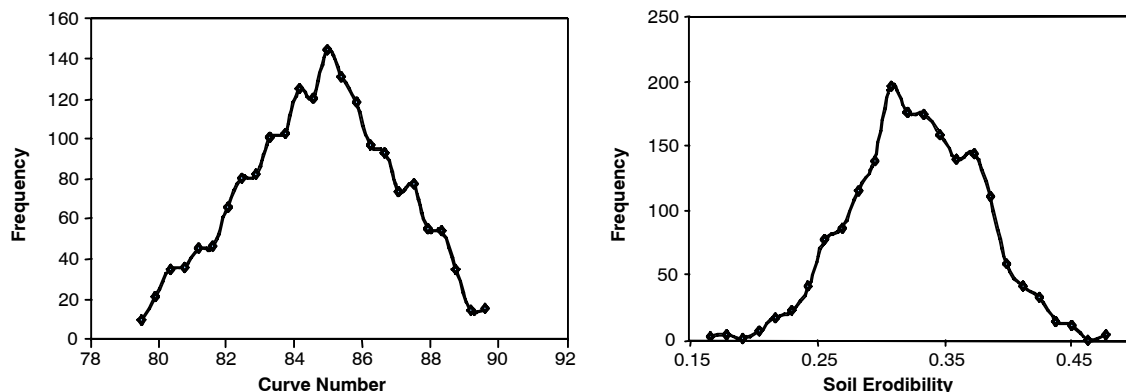


Figure 7. Sampled distributions for curve number and soil erodibility for the 1-year storm (1800 simulations).

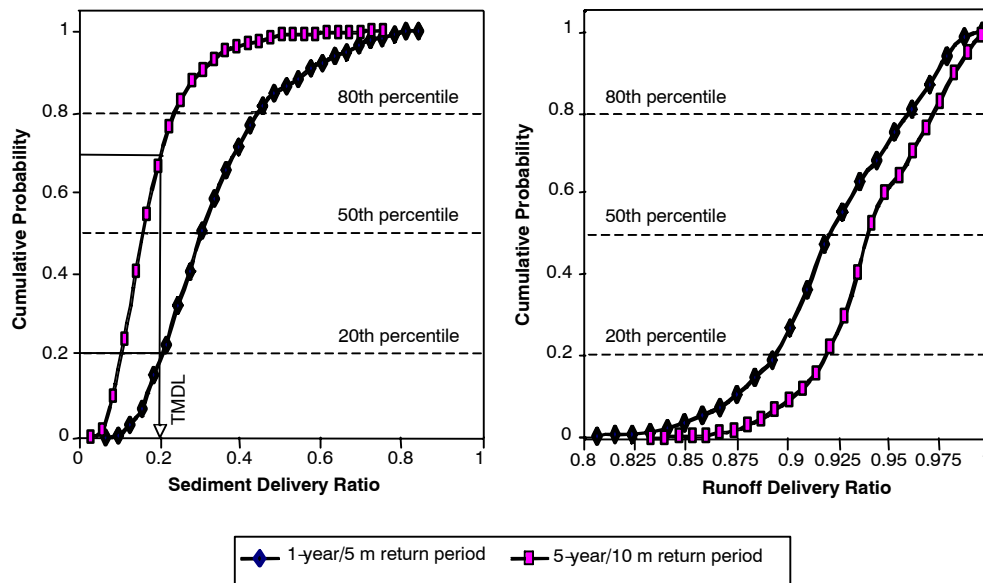


Figure 8. Comparison of simulated probability density functions for sediment and runoff delivery ratios obtained in the uncertainty analysis.

much the uncertainty can be reduced. Including quantification of uncertainty in the process would provide the basis for a quantitative cost-benefit analysis related to collecting additional field data.

The margin of safety (MOS) term in the TMDL equation is intended to account for uncertainty in the development process. However, the MOS is often taken to be an assumed value or percentage of the TMDL (Sexton et al., 2005) rather than explicitly including estimates of uncertainty. If considerations of uncertainty are included directly in the estimates of the waste load allocation (WLA) and the load allocation (LA), then the MOS is not necessary. Typically, the WLA is assumed to be the sum of the permitted loadings in the watershed, with no uncertainty assumed. However, if some level of uncertainty is assumed in the WLA estimate, such as $\pm 10\%$, then the MOS can be expressed as $0.10 \times \text{WLA}$. As described earlier, the uncertainty analysis of models used to estimate the load allocation (LA) generally yields a distribution of output values (e.g., fig. 4). One approach for translating a distribution into an estimate of the LA plus an MOS would be to take the LA as the value corresponding to 50% probability of occurrence and then estimate the MOS as a percentage of that value.

A better approach would be to more fully utilize the output distribution and, therefore, the estimate of uncertainty in developing a TMDL. A TMDL is defined as the load that a water body can assimilate without violating water quality standards. Thus, in developing a TMDL, the model results for

the constituent of interest are compared with the applicable water quality standard. By considering the uncertainty in the model output, one could determine the probability of meeting (or violating) the water quality standard. For example, if the maximum average monthly sediment yield allowed by the standard was 0.02 t ha^{-1} , there would be an 80% probability that the sediment yield will be less than that value, or an 80% probability of meeting the water quality standard, for the scenario represented by figure 6. If a higher level of confidence is desired by the decision makers or other stakeholders, additional management (or reduction) scenarios would need to be simulated with the model to find scenarios with a higher probability of meeting the standard. The TMDL could then be presented as shown in table 9, thus allowing decision makers and other stakeholders to make the decision as to the trade-off between required reductions and the probability of meeting the standard. As indicated in the example in table 9, increasing the probability of meeting the standard by a small percentage could require substantial reductions in loading, and thus substantial financial and personnel resources.

Information of the type shown in table 9 would also benefit the development of TMDL implementation plans (IPs), particularly in the public participation phase and in the allocation of resources for implementation. Public participation is a very important part of IP development. Presenting management scenarios to the public along with the costs and probability of meeting specific water quality standards for

Table 9. Reductions in bacteria loadings from a sample watershed required to achieve particular probabilities of meeting the water quality standard.

Probability of Meeting Standard (%)	Required Reduction in Loading							
	Cattle Direct Deposit	Cropland	Pasture	Loafing Lot	Wildlife Direct Deposit	Impervious Land Surfaces	Forest	Residential Pervious Areas
49	0	0	0	0	0	0	0	0
79	0	0	0	0	0	100	0	0
89	0	0	0	0	50	100	0	0
98	95	60	60	60	20	70	0	70
99	88	30	30	30	0	98	30	30
100	95	60	60	60	25	97	0	97

each scenario would give the stakeholders a basis for deciding which scenarios they prefer. Such information would also be important in establishing implementation phases. Resource agencies could also use such information to determine how to distribute resources most cost-effectively to achieve water quality goals.

SUMMARY AND CONCLUSIONS

Spatial and temporal variability in parameter values used in simulation models results in uncertain model outputs when describing or predicting a natural event. The types of model structure and algorithms used introduce additional uncertainty to predictions. Finally, measured data used for model calibration and validation introduce other potential sources of uncertainty in predictions. Models used for TMDLs are not immune from such uncertainties. This multi-state study identified possible sources of uncertainty in models used for TMDL assessment and discussed several methods of uncertainty analysis. Three methods, Monte Carlo simulation (MCS), Latin hypercube sampling with constrained MCS (LHS-MCS), and generalized likelihood uncertainty estimation (GLUE), were used to estimate uncertainty in the SWAT model's output due to input variability in the TMDL assessment phase of different watersheds located in three different states. In addition, a case study using VFSSMOD-W, a vegetative filter strip modeling design system, demonstrated the applicability of uncertainty analysis in the design of BMPs, a key component of TMDL implementation plans. Results emphasize the importance of uncertainty and the need to evaluate and report uncertainty in model simulations in PDF or equivalent forms. This may be performed using any method derived from first-order error (FOE) analysis or Monte Carlo simulation (MCS).

Determination and presentation of model outputs with associated probabilities for each simulation output can improve management decisions related to TMDL allocation and implementation. Including explicit quantification of uncertainty due to different sources in the TMDL process would provide more complete information for decision makers and other stakeholders. Quantitative cost-benefit analyses could be conducted with respect to establishing TMDLs and in developing TMDL implementation plans. A higher level of buy-in by stakeholders, particularly watershed residents and landowners, could be achieved by providing quantitative information on uncertainty as well as on the steps taken by TMDL developers and decision makers to reduce that uncertainty.

ACKNOWLEDGEMENTS

This article is the one of a series of articles developed by members of the USDA-CSREES Regional Project S-1004 "Development and Evaluation of TMDL Planning and Assessment Tools and Processes" and the ASABE SW-21 Hydrology Committee. The authors would like to acknowledge the editorial leadership and coordination provided by Rafael Muñoz-Carpena, Adel Shirmohammadi, and George Vellidis in this group effort. The authors also wish to thank Drs. Bosch, Srivastava, and Stoeckler for their time and effort in reviewing the manuscript before it was submitted.

REFERENCES

- Aguilar, F. J., F. Aguera, M. A. Aguilar, and F. Carvajal. 2005. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogram. Eng. and Remote Sensing* 71(1): 805-816.
- Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large-area hydrologic modeling and assessment: Part I. Model development. *J. American Water Resources Assoc.* 34(1): 73-89.
- Baffaut, C., M. A. Nearing, J. C. Ascough II, and B. Liu. 1997. The WEPP watershed model: II. Sensitivity analysis and discretization on small watersheds. *Trans. ASAE* 40(4): 953-943.
- Beck, M. B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Res.* 23(8): 1393-1442.
- Beven, K. 1989. Changing ideas in hydrology: The case of physically based models. *J. Hydrology* 105(1-2): 157-172.
- Beven, K. J., and A. M. Binley. 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes* 6(3): 279-298.
- Borsuk, M. E., and C. A. Stow. 2000. Bayesian parameter estimation in mixed-order model of BOD decay. *Water Resources* 34(6): 1830-1836.
- Borsuk, M. E., C. A. Stow, and K. H. Reckhow. 2002. Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development. *Environ. Sci. and Tech.* 36(10): 2109-2115.
- Chapra, S. C. 2003. Engineering water quality models and TMDLs. *J. Water Resources Planning and Mgmt.* 129(4): 247-256.
- Chu, T. W., and A. Shirmohammadi. 2004. Evaluation of SWAT model's hydrology component in the Piedmont physiographic region of Maryland. *Trans. ASAE* 47(4): 1057-1073.
- Chu, T. W., A. Shirmohammadi, H. Montas, and A. Sadeghi. 2004. Evaluation of the SWAT model's sediment and nutrient components in the Piedmont physiographic region of Maryland. *Trans. ASAE* 47(5): 1523-1538.
- Cotter, A. S., I. Chaubey, T. A. Costello, T. S. Soerens, and M. A. Nelson. 2003. Water quality model output uncertainty as affected by spatial resolution of input data. *J. American Water Resources Assoc.* 39(4): 977-986.
- Cryer, S. A., and G. E. Applequist. 2003a. Direct treatment of uncertainty: I. Applications in aquatic invertebrate risk assessment and soil metabolism of chlorpyrifos. *Environ. Eng. Sci.* 20(3): 155-167.
- Cryer, S. A., and G. E. Applequist. 2003b. Direct treatment of uncertainty: II. Applications in pesticide runoff, leaching, and spray drift exposure modeling. *Environ. Eng. Sci.* 20(3): 169-181.
- Desmet, P. J. J. 1997. Effects of interpolation errors on the analysis of DEMs. *Earth Surface Processes and Landforms* 22(6): 563-580.
- DiLuzio, M., J. G. Arnold, and R. Srinivasan. 2004. Integration of SSURGO maps and soil parameters within a geographic information system and nonpoint-source pollution model system. *J. Soil Water Cons.* 59(4): 123-133.
- EPA. 2005. National management measures to control nonpoint-sources pollution from forestry. Report No. EPA 841-B-05-001. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water.
- Franks, S. W., and K. J. Beven. 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. *J. Geophysical Res.* 102(D20): 23911-13999.
- Freer, J., and K. J. Beven. 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Res.* 32(7): 2161-2173.
- Gardner, R. H., and R. V. O'Neill. 1983. Parameter uncertainty and model predictions: A review of Monte Carlo results. In *Uncertainty and Forecasting of Water Quality*, 245. M. B. Beck and G. van Straten, eds. Berlin, Germany: Springer-Verlag.

- Haan, C. T. 1989. Parametric uncertainty in hydrologic modeling. *Trans. ASAE* 32(1): 137-146.
- Haan, C. T. 2002. *Statistical Methods in Hydrology*. 2nd ed. Ames, Iowa: Iowa State Press.
- Haan, C. T., and R. E. Schulze. 1987. Return period flow prediction with uncertain parameters. *Trans. ASAE* 30(3): 665-669.
- Haan, C. T., B. Allred, D. E. Storm, G. J. Sabbagh, and S. Prabhu. 1995. Statistical procedure for evaluating hydrologic/water quality models. *Trans. ASAE* 38(3): 725-733.
- Haan, P. K., and R. W. Skaggs. 2003. Effect of parameter uncertainty on DRAINMOD predictions: I. Hydrology and yield. *Trans. ASAE* 46(4): 1061-1067.
- Hankin, B. G., and K. J. Beven. 1998. Modeling dispersion in complex open channel flows: Equifinality of model structure (1). *Stochastic Hydrology and Hydraulics* 12(6): 377-396.
- Hardegree, S., S. Van Vactor, K. Healy, C. Alonso, J. Bonta, D. Bosch, D. Fisher, D. Goodrich, D. Harmel, J. Steiner, and M. Van Liew. 2003. Multi-watershed evaluation of WSR-88D (NEXRAD) radar-precipitation products. In *Proc. 1st Interagency Conference on Research in the Watersheds*, 486-491. Washington, D.C.: USDA-ARS.
- Harmel, R. D., R. J. Cooper, R. M. Slade, R. L. Haney, and J. G. Arnold. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE* 49(3): 689-701.
- Hartkamp, A. D., J. W. White, and G. Hoogenboom. 1999. Interfacing geographic information systems with agronomic modeling: A review. *Agron. J.* 91(5): 761-772.
- Hasofer, A. M., and N. C. Lind. 1974. Exact and invariant second-moment code format. *J. Eng. Mechanics Div., ASCE* 100(EM1): 111-121.
- Hession, W. C., D. E. Storm, and C. T. Haan. 1996. Two-phase uncertainty analysis: An example using the universal soil loss equation. *Trans. ASAE* 39(4): 1309-1319.
- Hugentobler, M. 2001. Propagation of uncertainties in digital terrain models with interval methods. In *Proc. GIS Research U.K. 9th Annual Conference (GISRUK 2001)*, 341-344. Pontypridd, Wales: University of Glamorgan, GIS Research Centre.
- Iman, R. L., and J. C., Helton. 1985. A comparison of uncertainty and sensitivity analysis techniques for computer models. NUREG/CR-3904. Albuquerque, N.M.: Sandia National Laboratories.
- Iman, R. L., and M. J. Shortencarier. 1985. A FORTRAN 77 program and user's guide for the generation of Latin hypercube and random samples for use with computer models. NUREG/CR-3624. Albuquerque, N.M.: Sandia National Laboratories.
- Iman, R. L., J. M. Davenport, and D. K. Zeigler. 1980. Latin hypercube sampling (program user's guide). Technical Report SAND79-1473. Albuquerque, N.M.: Sandia National Laboratories.
- Jensen, K. H., and J. C. Refsgaard. 1991. Spatial variability of physical parameters and processes in two field soils Part III: Solute transport at field scale. *Nordic Hydrology* 22: 327-340.
- Jury, W. A., W. R. Gardner, and W. H. Gardner. 1991. *Soil Physics*. New York, N.Y.: John Wiley and Sons.
- Lapedes, D. N. 1978. *McGraw-Hill Dictionary of Scientific and Technical Terms*. 2nd ed. New York, N.Y.: McGraw-Hill.
- Lindsay, J. B., and I. F. Creed. 2005. Sensitivity of digital landscapes to artifact depressions in remotely sensed DEMs. *Photogram. Eng. and Remote Sensing* 71(1): 1029-1036.
- Luis, S. J., and D. McLaughlin. 1992. A stochastic approach to model validation. *Advances in Water Resources* 15(1): 15-32.
- Madsen, H. O., S. Krenk, and N. C. Lind. 1986. *Methods of Structural Safety*. Englewood Cliffs, N.J.: Prentice-Hall.
- Mailhot, A., and J.-P. Villeneuve. 2003. Mean-value second-order uncertainty analysis method: Application to water quality modeling. *Advances in Water Resources* 26(5): 491-499.
- McKay, A. D., W. J. Conover, and R. J. Beckman. 1979. A comparison of three methods for selecting values on input variables in the analysis of output from a computer code. *Technometrics* 21(6): 239-245.
- Melching, C. S., and W. Bauwens. 2001. Uncertainty in coupled nonpoint sources and stream water-quality models. *J. Water Resources Planning and Mgmt.* 127(6): 403-413.
- Melching, C. S., and C. G. Yoon. 1996. Key sources of uncertainty in QUAL2E model of Passaic River. *J. Water Resources Planning and Mgmt.* 122(2): 105-113.
- Muñoz-Carpena, R., and J. E. Parsons. 2004. A design procedure for vegetative filter strips using VFSSMOD-W. *Trans. ASAE* 47(5): 1933-1941.
- Muñoz-Carpena, R., J. E. Parsons, and J. W. Gilliam. 1999. Modeling hydrology and sediment transport in vegetative filter strips. *J. Hydrology* 214: 111-129.
- Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2001. Soil and Water Assessment Tool: Theoretical documentation, version 2000. Temple, Texas: USDA-ARS.
- Novotny, V. 2003. *Water Quality: Diffuse Pollution and Watershed Management*. 2nd ed. New York, N.Y.: John Wiley and Sons.
- NRC. 2001. Assessing the TMDL approach to water quality management. Washington D.C.: National Research Council of the National Academy of Science.
- Parsons, J. E., and R. Muñoz-Carpena. 2001. Impact of uncertainty on the design of vegetative filter strips. ASAE Paper No. 012214. St. Joseph, Mich.: ASAE.
- Parsons, J. E., and R. Muñoz-Carpena. 2002. VFSSMOD-W, a graphical Windows system for the evaluation and design of vegetative filter strips for sediment trapping. In *Proc. Watershed Management to Meet Emerging TMDL Environmental Regulations*, 532-535. A. Saleh and B. Wilson, eds. St. Joseph, Mich.: ASAE.
- Peck, A. J., R. J. Luxmoore, and J. L. Stolzy. 1977. Effect of spatial variability of soil hydraulic properties in water budget modeling. *Water Resources Res.* 13(2): 348-354.
- Queija, V. R., J. M. Stocker, and J. J. Kosovich. 2005. Recent U.S. geological applications of lidar. *Photogram. Eng. and Remote Sensing* 71(1): 5-9.
- Reckhow, K. H. 2003. On the need for uncertainty assessment in TMDL modeling and implementation. *J. Water Resources Planning and Mgmt.* 129(4): 245-246.
- Romanowicz, R. J., and K. J. Beven. 1998. Dynamic real-time prediction of flood inundation probabilities. *Hydrological Processes* 43(2): 181-196.
- Schneider B. 2001. Uncertainty of local form in digital terrain modelling. In *Proc. GIS Research U.K. 9th Annual Conference (GISRUK 2001)*, 336-340. Pontypridd, Wales: University of Glamorgan, GIS Research Centre.
- Schulz, K., K. J. Beven, and B. Huwe. 1999. Equifinality and the problem of robust calibration in nitrogen budget simulation. *SSSA J.* 63(6): 1934-1941.
- Sexton, A. M., A. Shirmohammadi, and H. J. Montas. 2005. A conceptual approach for estimating margin of safety values for TMDL model simulations. In *Proc. Watershed Management to Meet Water Quality Standards and Emerging TMDL III*, 111-117. ASAE Publication No. 701P0105. St. Joseph, Mich.: ASAE.
- Shirmohammadi, A., H. J. Montas, L. Bergstrom, and W. G. Knisel, Jr. 2001. Water quality models. In *Agricultural Nonpoint-Source Pollution: Watershed Management and Hydrology*, 233-256. W. F. Ritter and A. Shirmohammadi, eds. Washington, D. C.: Lewis Publishers.
- Sohrabi, T. M., A. Shirmohammadi, T. W. Chu, H. Montas, and A. P. Nejadhashemi. 2003. Uncertainty analysis of hydrologic and water quality predictions for small watershed using SWAT2000. *Environ. Forensics* 4: 229-238.

- Spear, R. C., and G. M. Hornberger. 1980. Eutrophication in Peel Inlet: II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research* 14(1): 43-49.
- Summers, J. K., H. T. Wilson, and J. Kou. 1993. A method for quantifying the prediction uncertainties associated with water quality models. *Ecological Modelling* 65(3-4): 161-176.
- Topping, J. 1972. *Errors of Observation and their Treatment*. 4th ed. London, U.K.: Chapman and Hall.
- U.S. Congress. 1972. Federal Water Pollution Control Act Amendments. P.L. 92-500.
- Warrick, A. W., and D. R. Nielsen. 1980. Spatial variability of soil physical properties in the field. In *Applications in Soil Physics*, 319-344. D. Hillel, ed. New York, N.Y.: Academic Press.
- White, K. L., and I. Chaubey. 2005. Sensitivity analysis, calibration, and validation for a multisite and multivariable SWAT model. *J. American Water Resources Assoc.* 41(5): 1077-1089.
- Wise, S. 2000. Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological Processes* 14(11-12): 1909-1929.
- Wise, S. 2001. The effect of DEM interpolation errors on hydrological modelling at different scales: Does size really matter? In *Proc. GIS Research U.K. 9th Annual Conference (GISRUK 2001)*, 393-398. Pontypridd, Wales: University of Glamorgan, GIS Research Centre.
- Woolhiser, D. A., and D. L. Brakensiek. 1982. Hydrologic system synthesis. In *Hydrologic Modeling of Small Watersheds*, 3-16. C. T. Haan, H. P. Johnson, and D. L. Brakensiek, eds. ASAE Monograph No. 5. St. Joseph, Mich.: ASAE.
- Wyss, G. D., and K. H. Jorgensen. 1998. A user guide to LHS: Sandia's Latin hypercube sampling software. In *Technical Report SAND98-0210*, 1-140. Albuquerque, N.M.: Sandia National Laboratories.
- Yu, P. S., T. C. Yang, and S. J. Chen. 2001. Comparison of uncertainty analysis methods for a distributed rainfall-runoff model. *J. Hydrology* 244(1-2): 43-59.
- Zak, S. K., and K. J. Beven. 1999. Equifinality, sensitivity, and predictive uncertainty in the estimation of critical loads. *Science of the Total Environ.* 236: 191-214.
- Zak, S. K., K. J. Beven, and B. Reynolds. 1997. Uncertainty in estimation of critical loads: A practical methodology. *Water, Air and Soil Pollution* 98(3-4): 297-316.
- Zhang, H., C. T. Haan, and D. L. Nofziger. 1993. An approach to estimating uncertainties in modeling transport of solutes through soils. *J. Contaminant Hydrology* 12(1-2): 35-50.

