

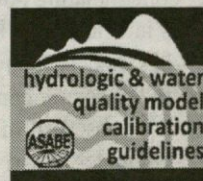
1-1-2015

# Uncertainty Considerations in Calibration and Validation of Hydrologic and Water Quality Models

J. Hernandez  
*Boise State University*



# UNCERTAINTY CONSIDERATIONS IN CALIBRATION AND VALIDATION OF HYDROLOGIC AND WATER QUALITY MODELS



J. A. Guzman, A. Shirmohammadi, A. M. Sadeghi, X. Wang, M. L. Chu, M. K. Jha, P. B. Parajuli, R. D. Harmel, Y. Khare, J. Hernandez

**ABSTRACT.** *Hydrologic and water quality models (HWQMs) are increasingly used to support decisions on various environmental issues and policy directions for present and future scenarios, at scales varying from watershed to continental levels. Uncertainty associated with such models may affect the ability of the models to accurately evaluate the response of complex systems, leading to misguided assessments and risk management decisions. Current well-known HWQMs contain numerous input parameters, many of which are not known with certainty, and in other cases model users can hardly recognize the genesis of uncertainty. Uncertainty in data, model structure, and model parameters can propagate throughout model runs, causing the model output to substantially deviate from the expected response of the natural system. Various uncertainty assessment methods have been used with different HWQMs, creating concerns about an adequate approach for handling uncertainty in these models and how such an approach can be implemented across various discretization complexities and scales. In this article, our primary intention is to review uncertainty in the currently used HWQMs and to provide guidance and useful information for researchers and investigators. In this regard, we explore the genesis of uncertainty in hydrologic and water quality modeling (i.e., spatiotemporal scales, model representation, model discretization, model parameterization) and provide strategies for assessing uncertainty in hydrologic and water quality modeling on local and global scales when interpreting the model output.*

**Keywords.** *Hydrologic modeling, Model calibration, Model validation, Spatiotemporal, Uncertainty.*

**U**ncertainty is a term used across many disciplines without identical definitions, but in essence it is related to the variable associated with the difference between the true state of a system

Submitted for review in April 2014 as manuscript number NRES 10710; approved for publication by the Natural Resources & Environmental Systems Community of ASABE in December 2014.

Mention of company or trade names is for description only and does not imply endorsement by the USDA. The USDA is an equal opportunity provider and employer.

The authors are **Jorge A. Guzman, ASABE Member**, Research Hydrologist, USDA-ARS Grazinglands Research Laboratory, El Reno, Oklahoma; **Adel M. Shirmohammadi, ASABE Fellow**, Professor and Associate Dean, College of Agriculture and Natural Resources, University of Maryland, College Park, Maryland; **Ali Sadeghi**, Soil Physicist, USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland; **Xiuying Wang**, Research Scientist, Blackland Research and Extension Center, Temple, Texas; **Maria L. Chu**, Assistant Professor, Department of Agricultural and Biological Engineering, University of Illinois, Urbana-Champaign, Illinois; **Manoj K. Jha**, Assistant Professor, Department of Civil, Architectural, and Environmental Engineering, North Carolina University A&T State University, Greensboro, North Carolina; **Prem B. Parajuli, ASABE Member**, Associate Professor, Department of Agricultural and Biological Engineering, Mississippi State University, Mississippi State, Mississippi; **R. Daren Harmel, ASABE Member**, Research Leader, USDA-ARS Grassland, Soil and Water Research Laboratory, Temple, Texas; **Yogesh Khare, ASABE Member**, Post-Doctoral Research Associate, Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida; **Jairo Hernandez, ASABE Member**, Assistant Professor, Department of Civil Engineering, Boise State University, Boise, Idaho. **Corresponding author:** Jorge Guzman, USDA-ARS Grazinglands Research Laboratory, 7207 West Cheyenne St., El Reno, OK 73036-2144; phone: 405-262-5291; e-mail: jorge.guzman@okstate.edu.

and its observational or theoretical assessment at a specific space and time. For example, Lapedes (1978) defined uncertainty as the estimated amount by which an observed or calculated value departs from the true value. Note that this definition implies that the true values are known. However, true values in natural systems may be unknown but measurable under specific set of conditions in time and space (i.e., environmental observations, dynamic system parameters, etc.). In hydrologic and water quality models (HWQMs) observations are commonly accepted to represent true values, as they are the best known assessment of the system. However, it has been well documented that a degree of uncertainty is always associated with all observations (Shirmohammadi et al., 2006; Harmel et al., 2006, 2009). Walker et al. (2003) described uncertainty as any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system. The "ideal" concept was introduced in a manner similar to the true value commonly found in observational uncertainty, in which measurements can be repeatedly conducted on the same object to derive the statistical properties of the observations. Because HWQMs are made to be systemic representations of the real world with unique observations, uncertainty analysis provides modelers and model users the capacity to quantify and assess the sources of uncertainty in models' outputs to improve model robustness.

Beck (1987) provided a review of the history and the analysis of uncertainty with regard to water quality model-



ing. The review covered the role of uncertainty both in the development of water quality models for interpretation and explanation of past behavior and in the applications of such models for predicting future behavior. According to Beck (2009), “contemplating now climate change, and what to do about it, is inconceivable in the absence of very high-order models—big models. They, and their reliability and uncertainties, are the central concern. We are surrounded by ‘big data,’ both as observed of the real world and as generated from the computer world with these big models. And yet, there remains this deep uncertainty to be handled at the interfaces among science, policy, and society.”

Studies have reported that uncertainties in predicted output could be associated with many model components, including model structure (e.g., unaccounted processes, empirical equations, and numerical errors), model parameters (e.g., incomplete knowledge of parameter values, ranges, and their physical meaning), and spatial and temporal variability in required input values (Vicens et al., 1975; Loague and Corwin, 1996; Loucks et al., 2005; Harmel and Smith, 2007; Yang et al., 2008; Zhang et al., 2009a, 2009b; Wang and Yen, 2014; Yen et al., 2014). Haan (1989) represented the sources of uncertainty with an equation as:

$$O = f(I, P, t) + e \quad (1)$$

where  $O$  is an  $n \times k$  matrix of the watershed response to be modeled,  $f$  is a collection of functional relationships,  $I$  is an  $n \times m$  matrix of inputs,  $P$  is a vector of  $p$  model parameters,  $t$  is the temporal scale of prediction and calculation, and  $e$  is an  $n \times k$  matrix of errors (note that this term should be interpreted as uncertainty in predicted output), with  $n$  data points,  $k$  outputs, and  $m$  inputs. Even though the distinction between  $I$  and  $P$  is not always clear,  $I$  generally represents a known or measured model input value, and  $P$  represents a parameter that is estimated based on some functional relationship with the known inputs. Equation 1 indicates that uncertainty may be due to variability in the input data (e.g., spatiotemporal scales, parameterization), in the algorithms selected for simulation of the processes included in the model, in the accuracy and level of available measured functional data (e.g., runoff, subsurface flow, nutrient discharge, etc.) for model calibration and validation, in setting proper initial and boundary conditions, and in the scale of application (Beven, 1989). For details on the sources of uncertainty, refer to Shirmohammadi et al. (2006).

HWQMs are frequently used as a cost-effective alternative to monitoring data for watershed management, assessment of nonpoint-source pollution, and to set water resource policy (Graham and Butts, 2005; Yang et al., 2008; Zhang et al., 2009a; Sexton et al., 2011b). However, major concerns include the level of reliability in model outputs and the capacity to quantify the degree of certainty associated with the predictions and estimations, especially when modeling complex ecosystems. Achieving reliable outputs from HWQMs is not straightforward, as uncertainty in modeling complex systems can arise at any stage of the model assessment, propagate non-linearly to the model output, and be masked or amplified due to parameterization (Beven, 1989, 1993; Kirchner, 2006). In addition, model

type (e.g., physically based, conceptual, or data driven), spatial discretization (e.g., distributed, semi-distributed, or lumped), the mathematical representation of the natural system, and the model objectives (e.g., forensic or forecasting) are intrinsically related (Wagener and Gupta, 2005) and dependent on each other; thus, uncertainty in one will affect uncertainty in the ultimate outcome of the model.

In general, model parameters that are intended to represent processes in space and time (patterns) are bound to the spatiotemporal properties of observations that cannot be separated from each other (Hiebler and Michaud, 2012). Defining a consistent effective parameter value to reproduce the response of a spatially variable pattern of parameter value is not possible (Beven, 1989). Conversely, parameters in HWQMs are seldom evaluated for their representativeness and are commonly reverse-estimated or “polished” by calibration techniques based on model performance objective functions (e.g., metrics). In addition, parameters are treated as deterministic rather than stochastic (described by probabilistic distributions) when associated with physical observations (Wagener and Gupta, 2005) and seldom allowed to change (dynamics) along model runs.

To estimate potential uncertainties in outputs due to model structure, model input, and model parameter values, several uncertainty analysis methods have been developed. Some of these methods include: generalized likelihood uncertainty estimation (GLUE; Beven and Binley, 1992), the parameter solution (ParaSol) method (van Griensven et al., 2006), the sequential uncertainty fitting (SUFI-2) method (Abbaspour et al., 2004), Latin hypercube sampling (LHS) with constrained Monte Carlo simulation (MCS; Sohrabi et al., 2002; Sohrabi et al., 2003), the Bayesian Markov chain Monte Carlo (MCMC) method (Yang et al., 2007), the Bayesian model averaging (BMA) method (Hoeting et al., 1999), and dynamically dimensioned search - approximation of uncertainty (DDS-AU; Tolson and Shoemaker, 2008). Ajami et al. (2007) proposed the integrated Bayesian uncertainty estimator (IBUNE) to account for the major uncertainties associated with model input, parameters, and structure in hydrologic prediction. Yen et al. (2014) extend IBUNE to further account for measurement uncertainty in their integrated parameter estimation and uncertainty analysis tool (IPEAT) for SWAT. Melching and Bauwens (2001) used a combination of LHS and MFORM (mean value first-order reliability method) to evaluate uncertainty in HWQMs combined with stream water quality models. In their study, LHS was used to identify the basic parameters that significantly contribute to the model output uncertainty, while MFORM was used to quantify the percentage contribution of the variables to the output uncertainty. It is important to note that these methods differ in their approach and do not necessarily use the same uncertainty analysis philosophy (Yang et al., 2008). For instance, GLUE and SUFI-2 account for uncertainties from all sources, including model input, model structure, model parameters, and response through parameter uncertainty. On the other hand, ParaSol accounts for prediction uncertainty, which is constructed by equally weighting all uncertain parameters (Van Griensven and Meixner, 2006).



Uncertainty analysis is a rather complex process that has merit for implementation because it allows "intellectual honesty" to prevail in model results. Uncertainty analysis also allows the interpretation of risks associated with model output uncertainty. Furthermore, assessing model simulation results with a relatively known level of certainty will allow policymakers and conservation program managers to improve the implementation of pollution abatement programs to more effectively and efficiently achieve the nation's goals for cleaner water and healthier watersheds (Sexton et al., 2011b). For example, one of the major challenges is the use of HWQMs for assessing total maximum daily loads (TMDLs; NAS, 2001; NRC, 2001; EPA, 2002). However, most HWQMs were not developed with a comprehensive uncertainty assessment in mind, and future model development should integrate a hierarchical framework in which system identification, parameter estimation, state estimation, and output prediction are systematically evaluated (Liu and Gupta, 2007).

The primary goal of this article is to review the methods used to evaluate uncertainty in current HWQMs and to provide guidance for developers, model users, and researchers. Specifically, our intention is to explore the genesis of uncertainty in HWQMs while considering the spatio-temporal scales, model representation, model discretization, and model parameterization. We also provide strategies for assessing the associated uncertainties on local and global scales when interpreting model output and introduce best possible approaches to quantify such uncertainties, especially during HWQM calibration and validation.

## GENESIS OF UNCERTAINTY IN HYDROLOGIC MODELING

### OBSERVATIONAL ERRORS IN MODEL INPUT DATA

Measurements in monitoring natural processes for hydrologic investigation are normally associated with some degree of inaccuracy (e.g., random and systematic errors), which can change during the observational period (Fiebrich et al., 2005; Harmel et al., 2006). These inaccuracies are a function of the nature of the sensor, sensor degradation, mathematical translation, environmental conditions, data collection operation, and other factors. Thus, measurement uncertainties are inherent in the observations used to infer quantities or scientific judgment. The degree of variability captured during the observational period (sampled on a certain spatiotemporal scale) compared to the system's expected variability as well as the deviation of the mean from the theoretical mean of the observations define the observation uncertainty (Sohrabi et al., 2002). However, in hydrological systems, the theoretical mean is unknown in many cases and subject to dynamic processes in which the stationary condition of the dataset is not preserved. For instance, areal precipitation, which is commonly used in hydrologic models, is frequently underestimated or overestimated, which is attributed to the characteristics of the environment; changes in instrumentation, measurements, and data processing protocols; recording practices; and meteorological dynamics. Groisman and Legates (1994) estimated

that bias in precipitation measurements accounts on average for 5% to 25% of the precipitation in the continental U.S., with larger bias at higher elevations and latitudes of as much as 400% in some locations.

The quality of observed data should not be considered a trivial problem. Observational errors can induce or mask variability, result in undesired trends, and impact the model calibration and validation results. Therefore, uncertainty in the observations used for model calibration and validation should be taken into consideration (Montgomery and Sanders, 1986; Harmel et al., 2006, 2009). Moreover, datasets may not contain sufficient statistically viable measurements (e.g., number of samples or measurements, frequency of sampling or measurements, etc.) to properly characterize the system for accurate inferences or extrapolations. Finally, it is not common to find data assessment, statistical analysis, and data adjustment performed on observed datasets prior to their use for model calibration and validation (e.g., data screening; Chu et al., 2013; Guzman et al., 2014).

Data inconsistency results from changes in the observational error (e.g., random or systematic error) that become problematic due to the inherent assumption in HWQMs that the statistics describing input data remain unchanged over time (Kahya and Kalayci, 2004; Xu et al., 2003; Chu et al., 2013). In addition, natural or human processes may result in significant changes in statistical descriptors (e.g., mean, variance, and trends) of the observations (i.e., non-homogeneity; Yevjevich and Jeng, 1969; Guzman et al., 2014). Moreover, conditions of stationarity are often implicitly invoked when calibrating model parameters (e.g., based on current input datasets) and then used to predict the watershed hydrologic response under changing scenarios within non-stationary conditions. Thus, these scenarios use new datasets (e.g., precipitation, land use, soil physical properties) that may possess statistically significant changes in statistical descriptors, thus propagating uncertainty in the overall model predictions.

### SPATIOTEMPORAL SCALE AND FUNDAMENTAL DATA PROPERTIES

Observation of hydrological phenomena, as stated earlier, is bound to spatiotemporal properties of the observations that are mutually dependent (Hiebele and Michaud, 2012), and this introduces uncertainty (Beven, 1989). Perception of certain hydrological phenomena is related to the intrinsic data properties (e.g., sampling frequency and density, and quality of the observations). At watershed scales, point measurements are often extrapolated to represent a larger area or volume (i.e., spatial scale), thus missing the representation of spatial heterogeneity. When this heterogeneity is spatially represented in the model (geometric spatial discretization), it is intrinsically related to the computational methods and mathematical translations. As an example, hydrologic models are commonly classified by their geometrical spatial discretization as distributed, semi-distributed, or lumped. Heterogeneity is represented by unique parameters related to the specific space (e.g., lumped or semi-distributed) or by a set of parameters (e.g., distributed) derived from raster or vector data structures.



Thus, on one hand, observations are constrained by the area or volume that properly represents all the physical phenomena (e.g., representative elementary volume; Constanza-Robinson et al., 2011). Conversely, the spatial representation (e.g., structured, unstructured grids, units, etc.) is linked to mathematical and numerical approximations (e.g., finite differences, finite volume, finite element, spectral, pseudo-spectral, etc.) occurring in the virtual world. These three levels of data disaggregation (elementary volume, spatial discretization, and mathematical computation) may result in different spatial scales to accurately represent the variability of hydrologic phenomena, which consequently affects uncertainty in the model output.

Similarly, temporal data aggregation or disaggregation (e.g., hourly, daily, monthly, yearly) introduce another data transformation that potentially affects the model calibration and validation (Harmel et al., 2009). As an example, changes in the sampling frequency of rainfall (subdaily, daily, monthly, etc.) result in changes in the representation of the magnitude and spatial patterns of rainfall events, and consequently runoff estimations (St-Hilaire et al., 2003) and other related transport phenomena. Temporal downscaling of rainfall is commonly achieved by using numerical solutions that may not be linearly valid. Mathematical operations applied on the spatiotemporal scale of the observations may impact the variance of time series data that, when placed within the context of other observations, may result in unintended spatial and temporal patterns (Blöschl and Sivapalan, 1995; Grayson et al., 2002; Poveda, 2011). Thus, modelers must carefully evaluate the objectives of the model with the uniformity of the spatiotemporal scales of the observations and the model runs to properly assess the phenomena of interest (e.g., runoff, chemical and microbial fate and transport) at a well-defined space and time, thus minimizing model uncertainty.

Selection of the appropriate spatiotemporal scale of analysis should be directly related to the modeling purpose. At the same time, methods of uncertainty estimation should reflect the variation in model performance at the selected spatiotemporal scales (Kirkby, 1976; Heuvelink, 1998). For example, if the purpose is to determine the size of a culvert, then calibration and validation should focus on peak runoff rates driven by subdaily rainfall data (e.g., breakpoint rainfall). As discussed by Goodrich et al. (2012), at relatively small scales, carefully calibrated process-based models are superior to simpler models. At larger scales, however, inadequate knowledge of boundary conditions, critical inputs, representative properties, and the emergence of additional processes makes process-based modeling more challenging. Regardless of model complexity, model performance measures should focus on the scale of variability (i.e., modeling purpose), whether short-term dynamic response (e.g., design of culverts) or long-term impact (e.g., determining the relative impacts of best management practices).

Increasing the spatial scale from field to farm to hillslope to watershed to basin scale introduces new processes, simplifies assumptions with additional modeling parameters, and disproportionately incorporates new observations (Hiebler and Michaud, 2012; Loe et al., 2012; Kelly et al., 2011). For example, infiltration and erosion

parameters established at a point scale may not apply at a larger scale in aggregate or lumped model applications. At small scales (areas less than a few square meters), soil is detached by the energy of impacting raindrops and is transported by rainfall-driven shallow flows. As scale increases to the size of typical erosion plots (tens of square meters), runoff may become concentrated in rills, where detachment and transport are increased by rill erosion. At larger scales (hundreds to thousands of square meters), gully erosion may become a major sediment source. At watershed scales, depending on the local and upslope conditions, stream channels and associated floodplains may be important sources or sinks of sediment. All critical processes must be adequately reflected in the model structure when models are parameterized and validated at different spatiotemporal scales.

Data resolution is another major issue when changing the spatial scale of modeling. Of the many types of data (topography, land use, soil, land management, etc.) involved in modeling, precipitation data are probably the biggest driver of uncertainty (Beeson et al., 2014), which is aggravated by non-stationary conditions. HWQMs are precipitation-driven models, so correct representation of spatial precipitation is essential. One way to achieve this is to collect data at a finer spatial scale, but this approach may be impractical and cost-ineffective in most cases. Weather generators may not be reliable for generating an entire dataset. NEXRAD (next generation radar) may offer a promising alternative for regions where a limited rain gauge network exists for spatial representation of precipitation (Sexton et al., 2010; Beeson et al., 2011).

Impact assessment of climate variability and change has increasingly relied on forcing hydrologic models with climate model projections (Jha et al., 2004, 2013, 2014; Takle et al., 2005) that are based on global circulation models. As opposed to watershed models, global circulation models produce results in structured grids (i.e., uniform spatial discretization), while HWQMs are calibrated against monitoring data (unstructured networks) with limited accuracies. Forcing these models with global circulation models introduces additional uncertainties, primarily in prediction at differing spatial and temporal scales.

#### MODEL REPRESENTATION AND PROPAGATION OF STRUCTURAL UNCERTAINTY

Model structural uncertainty results from unaccounted simulations of processes, lack of understanding of actual physical phenomena, or improper mathematical representation. Model structural uncertainty can arise from unaccounted feedback processes, such as the effects of wetlands and reservoirs on hydrology, and chemical transport associated with the occurrence of landslides in the watershed. Model inaccuracy due to oversimplification of processes and process interactions also leads to model structural uncertainty (Yang et al., 2008). Moreover, approximations or errors derived from numerical algorithms or binary numerical accuracy contributes to uncertainty as a function of the model's physical representation and discretization approach. For example, physically based models that solve non-symmetric equation systems derived from use of the



finite difference (e.g., MODFLOW, MIKE SHE), finite element (e.g., HYDRUS), or finite volume methods may prompt numerical issues (such as numerical stability, "good" approximation, and local conservancy), satisfy the discrete maximum principle, and work for general domains and arbitrary grids (Michev, 1996).

Building a conceptual model linked to the objectives of the problem is recommended before model application or before developing or updating the mathematical model. The latter is merely a mathematical representation of the former. Errors in the conceptual model due to incomplete understanding or representation of the system are bound to propagate through the modeling process regardless of the model's mathematical sophistication, accuracy of the input parameter values, and observed data. Walker et al. (2003) categorized this type of uncertainty as epistemic uncertainty. Thus, improving the knowledge of various fundamental processes and linkages is vital to reducing the uncertainty emanating from oversimplification, overrepresentation, or the lack of fundamental processes in the model. Refsgaard et al. (2007) and Walker et al. (2003) suggested ways to deal with epistemic uncertainty, some of which include: (1) additional research through extensive analysis of available data to explore less understood phenomena, (2) empirical efforts focused on improving knowledge of system variability through additional data collection (e.g., longer time series, collection frequency), and (3) expert opinion and peer review, all of which lead to better system understanding.

#### MODEL PARAMETERIZATION (PHYSICAL, EMPIRICAL, AND NESTED PARAMETERS)

The number of parameters (both physical and empirical) is large for most HWQMs. Although not error-free, the best values for these parameters may be obtained by measurements; thus, parameter uncertainty arises from incomplete knowledge of parameter values, ranges, and their physical meaning (Loucks et al., 2005; Yang et al., 2008). However, spatial and temporal heterogeneity in the landscape and overall ecosystem makes it almost impossible to have enough measured values to represent such variables in a statistically sound manner or to cover the entire surface and subsurface watershed to the optimal spatial and temporal scale for a specific hydrologic process. Model users often rely on literature data and sometimes their best "guesstimates" for model parameter values. For example, information on soils, climate, cropping practices, crop growth parameters, chemical inputs, as well as other physically based parameters such as Manning's roughness coefficient, physical dimensions of the landscape in terms of drainage pattern, etc., are difficult to measure in a spatially and temporally representative manner. This leads modelers to use average values for these parameters, despite the fact that the parameters may be sensitive, may significantly affect the model output, and do not reflect the mean behavior of the system (Sohrabi et al., 2003). Briggs et al. (2012) suggested that whether primary data sources are used to estimate input parameters or the information is derived from one or more secondary sources, the parameter estimation methodology should be able to generate a reliable point

estimate and an acceptable measure of precision by defining a defensible range. A point estimate with a sound range for any parameter of interest can be used by others to provide limits or ranges of model simulation uncertainties for prediction of any model outcome of interest. In addition, calibration requires tradeoff judgments by modelers. For example, one portion of a hydrograph can match better than others during a model calibration exercise. Each input parameter has a unique effect on different parts of the hydrograph. Thus, focusing on adjusting parameters without knowing the modeling purpose creates additional uncertainties (Herr and Chen, 2012).

Adjusting parameters unsystematically (e.g., calibrating for nutrients before calibrating for hydrology) can also add uncertainty due to the hierarchy of driven hydrological processes and fluxes controlling other phenomena. A large number of parameters should not be adjusted at once; instead, a hierarchical sequence of parameter calibration is preferred (Arnold et al., 2012). For example, model parameterization for suspended sediment may increase prediction uncertainty if the model is not yet parameterized for surface runoff. An example of a hierarchical sequence for adjusting model parameters could be flow volume, surface runoff, groundwater, temperature, suspended sediments, conservative ions, nutrients, phytoplankton, and dissolved oxygen. In addition, care should be taken when using autocalibration tools, which may adjust the parameters unrealistically if not done correctly, leading to incorrect modeling results or, in the worst-case scenario, to equifinality (Beven, 2006; Kirchner, 2006).

Briggs et al. (2012) discussed situations in which parameter uncertainty estimation is not required (e.g., identification of the quantitative relationships between inputs and outputs, or sensitivity analysis). In model parameterization and uncertainty analysis, care should be taken to distinguish sensitivity analysis from uncertainty analysis, as a highly uncertain parameter with low sensitivity may easily have more impact on the model outputs than a more sensitive parameter estimated or measured with less uncertainty.

Sexton et al. (2011a) conducted a thorough analysis of the impact of parameter uncertainty on critical SWAT outputs. Their results indicated that sensitive parameters should not be the only parameters considered for model calibration and their contribution to model output uncertainty. The study found that parameters to which the model was not highly sensitive contributed to model simulation uncertainty to a larger extent. Therefore, these parameters should be determined more accurately and also considered during model calibration to improve model performance and reduce model uncertainty.

#### UNCERTAINTY IN MONITORING DATA

Analysis of uncertainty in measured model input data, which drive model calibration and validation processes, improves model application and enhances decisions based on the modeling results (Reckhow, 1994; Kavetski et al., 2003; Pappenberger and Beven, 2006; Beven, 2006b; Shirmohammadi et al., 2006; Moriasi et al., 2007; Harmel and Smith, 2007; Harmel et al., 2010). The contribution of uncertainty in measured data (e.g., discharge, concentra-



tions, and loads) that are used for model calibration and validation has traditionally been ignored but has recently been emphasized. Montgomery and Sanders (1986) and Harmel et al. (2006, 2009) produced comprehensive uncertainty analysis methods specifically for measured discharge and water quality data to facilitate uncertainty estimation. These methods were needed because discharge, sample preservation and storage, and laboratory analysis commonly receive much of the research focus (e.g., Sauer and Meyer, 1992; Kotlash and Chessman, 1998).

Because calibration and validation data have some level of uncertainty, as do all measured data, this uncertainty should be considered in model calibration and validation. Estimating and reporting the uncertainty in the measured data (observational uncertainty) used to calibrate and validate models is recommended because of its impact on the evaluation and interpretation of model results. Furthermore, as recommended by Harmel and Smith (2007) and Harmel et al. (2010), model evaluation would ideally compare predicted values to the uncertainty range of measured values as well as to actual values.

## ASSESSING UNCERTAINTY IN HYDROLOGIC MODELING

### SENSITIVITY ANALYSIS VS. UNCERTAINTY ANALYSIS

Uncertainty analysis (UA) is the process of quantifying uncertainty in model outputs (Saltelli et al., 2008), which is closely related to the process of apportioning uncertainty to its source (i.e., sensitivity analysis; SA). The purpose of conducting SA prior UA is to identify the important or influential variables that impact UA the most, and to assess the linearity of the model. Thus, SA has been combined with UA to decrease the time and computational resources required for a comprehensive UA, and in most cases SA is considered another step in conducting UA in hydrologic models.

### LOCAL VS. GLOBAL SENSITIVITY ANALYSIS

The most common type of SA is based on a simple local derivative,  $dy/dx$ , which measures the change in  $y$  for every change in  $x$  (fig. 1a). The sampling is performed locally,

and one model parameter is varied while the rest are assumed constant, hence the term “one-factor-at-a-time” (OAT). Although this method is relatively simple, it can only give information at the point at which measurements were taken, which is an issue when the inputs are uncertain or when the relationship between inputs and outputs changes. In most cases, no *a priori* knowledge of the linearity of the model is known, or the linearity may change depending on model input assumptions. Saltelli and Annoni (2010) showed that OAT methods are inefficient measures of sensitivity unless some form of averaging of the system derivative is performed.

The global method of SA can estimate the effects of each input factor on an output when all other inputs are varying, which allows identification of the interactions among factors (Saltelli and Annoni, 2010; Cariboni et al., 2006). Because global methods are based on the exploration of the whole input space (fig. 1b), they can identify critical or interesting regions in the input space (Saltelli et al., 2008). Global methods are model-independent, and no constraints are imposed based on the linearity of the models.

### SETUP OF SENSITIVITY AND UNCERTAINTY ANALYSES

Methods of uncertainty and sensitivity analysis (discussed later) involve multiple model simulations and evaluations. The economy of the analysis depends on the type of method used. A general step-by-step procedure for performing global UA and SA includes: (1) determination of probability distribution functions (PDFs) of input parameters, (2) generation of input samples based on the PDFs, (3) application of a screening method to shorten the list of important parameters, (4) refining parameters and their ranges for rigorous analysis, (5) model simulations to calculate desired outputs and decision variables, and (6) statistical analysis to obtain sensitivity indices, parameter rankings, predictive PDF, and confidence intervals. Saltelli et al. (2004) provided recommendations for estimating parameter distributions and for selecting sensitivity analysis methods. Input data such as precipitation, temperature, soils, and land cover are critical drivers for watershed simulation. When the analysis is carried out for spatial input data (e.g., land use, soils, etc.), a number of maps and data

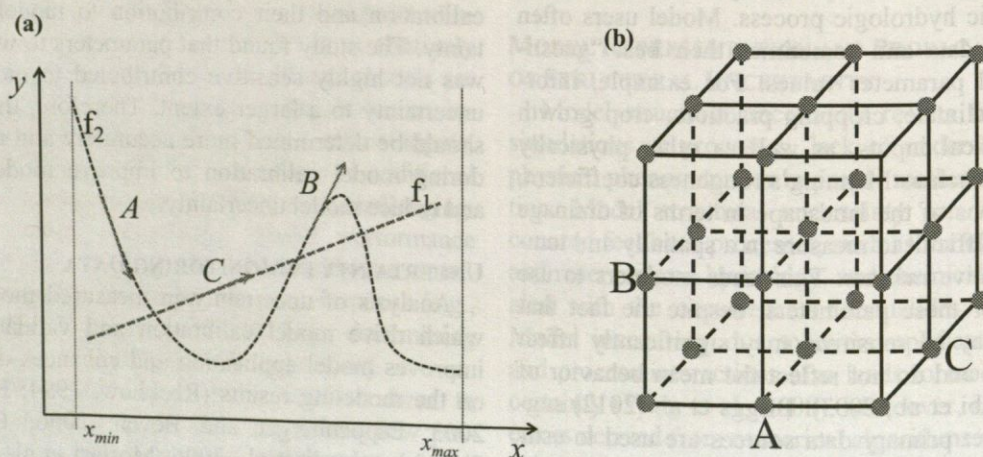


Figure 1. (a) Local sensitivity analysis expressed as a simple local derivative ( $dy/dx$ ) of a linear function ( $f_1$ ) and a non-linear function ( $f_2$ ). Note the change in slope of tangents A and B while tangent C remains constant. (b) Example of a sampling space of a three-input model. Each input is represented by minimum, mean, and maximum values. The dots show the sampling density (possible combinations) within the input space.



realizations are generated and used as variables (Lilburne and Tarantola, 2009; Zajac, 2010; Sexton et al., 2011). Kavetski et al. (2003), Ajami et al. (2007), and Yen et al. (2014) used an input error model to account for the uncertainty contributed by precipitation by introducing some "latent variables" (input error model parameters) to the system. These variables were then handled as model parameters during parameter estimation and optimization.

Various tools and software packages are available for parameter sample generation and post-simulation statistical analysis. These include SIMLAB (Saltelli et al., 2004), GLUE (Beven et al., 2001), PEST (Doherty, 2004), and MCAT (Wagener et al., 2001). A detailed summary of UA and SA tools, salient features, and application history can be found in Matott et al. (2009).

### **Sensitivity Analysis Methods**

The methods used for SA in HWQMs include simple derivative-based local methods (Lenhart et al., 2002), graphical methods (e.g., scatterplots), regionalized sensitivity analysis (RSA) coupled with tree-structured density estimation (TSDE) (Arabi et al., 2007), the method of elementary effects or Morris method (screening method) (Morris, 1991; Campolongo et al., 2007), correlation and regression indices (Helton and Davis, 2002), variance-based methods (ANOVA-like techniques) such as the Fourier amplitude sensitivity test (FAST; Cuiker et al., 1973) and its extended version (eFAST; Saltelli et al., 1999), and Sobol's method (Sobol, 1993). Among these, variance-based methods do not suffer from model non-linearity or non-monotonicity and produce robust sensitivity results. The FAST, eFAST, and Sobol methods have been implemented in a number of recent studies (Tang et al., 2007; van Werkhoven et al., 2008; Liu et al., 2009; Wagener et al., 2009; Yin et al., 2009; Chu-Agor et al., 2011, 2012). However, these methods are computationally expensive, limiting their applicability when the number of uncertain model parameters is very large, which has led to the use of combinations of these methods (Morris-FAST; Francos et al., 2003; Wang et al., 2006a; Muñoz-Carpena et al., 2007) and to modified and new sampling strategies (van Griensven et al., 2006). A summary of recent SA studies in surface hydrology and water quality monitoring can be found in Reusser et al. (2011). Surprisingly, Reusser et al. (2011) noted that FAST and eFAST have not been used as regularly as Sobol's method in rainfall-runoff models and watershed-scale models despite requiring a smaller number of model simulations compared to Sobol's method.

### **Uncertainty Analysis Methods**

Several methods have been used to quantify the uncertainty in model output due to the variability in input data. Ahsam and O'Connor (1994) used the Kalman filtering method. Shirmohammadi et al. (2006) provided a comprehensive description of several methods, including the mean value first-order reliability method (MFORM; Melching and Yoon, 1996), Monte Carlo simulation (MCS; Yu et al., 2001), Latin hypercube sampling (LHS) with constrained MCS (Sohrabi et al., 2003; Iman and Shortencarier, 1985; Iman et al., 1980), and GLUE (Beven and Binley, 1992). Sexton et al. (2011b) used MFORM and quantified the

margin of safety for a TMDL due to uncertainties in SWAT simulations as a result of variability in the input parameter values. Details of the various UA and SA methods can be found in Saltelli and Marivoet (1990), Melching (1995), Saltelli et al. (2000), Helton and Davis (2002), Saltelli et al. (2004), Beven (2001), Stow et al. (2007), and Manache and Melching (2008).

The UA methods implemented in HWQM simulation studies range from output distribution statistics (e.g., mean, range, interquartile ranges), output histograms and distribution functions, first-order and higher-order errors and mean estimates (e.g., Shirmohammadi et al., 2006) to quite sophisticated Markov chain Monte Carlo (MCMC) methods. Stow et al. (2007) recommended five UA methods for adaptive TMDL implementation: (1) RSA or generalized sensitivity analysis, (2) GLUE, (3) Bayesian Monte Carlo, (4) importance sampling, and (5) MCMC, all of which are essentially Monte Carlo methods. Among these, RSA (Hornberger and Spear, 1981) is not really a UA method but forms the basis of GLUE (Beven and Binley, 1992), which is the most widely used UA method in watershed modeling. RSA classifies model runs as behavioral or non-behavioral based on some fitting criteria. In GLUE, the concept of behavioral and non-behavioral runs is replaced by weighting of parameter sets based on likelihood measures of model runs. The likelihood measures used are generally goodness-of-fit criteria (e.g., root mean square error). Model runs can be classified as behavioral or non-behavioral using some subjective likelihood criteria and further rescaled so that the cumulative weight (of behavioral runs) is unity. However, considerable subjectivity is involved at various stages in GLUE (e.g., prior distributions, likelihood function, etc.), which can hamper the results. MCMC employs a full Bayesian framework and uses unassisted algorithms to choose a sample that approaches the posterior density function. This makes MCMC the fastest and most efficient method for UA. However, the efficiency of MCMC depends on the underlying algorithms to locate the optimal parameter combinations for the response surface. In HWQMs, response surfaces are quite complex; therefore, MCMC may need a large number of runs (Beven, 2001; Gallagher and Doherty, 2007). In addition, MCMC has been implemented in relatively few UA studies in hydrologic modeling (e.g., table 2). Mishra (2011), in his evaluation of the predictive uncertainty of HSPF in TMDL development using Monte Carlo based techniques, concluded that GLUE should be preferred over MCMC in watershed-scale water quality models. MCMC and GLUE both have advantages and disadvantages. A detailed discussion of this topic can be found in Vrugt et al. (2009), Jin et al. (2010), and Beven et al. (2011).

Ajami et al. (2007) proposed a framework, the integrated Bayesian uncertainty estimator (IBUNE), to account for model input, parameter, and structural uncertainty in hydrologic prediction. Yen et al. (2014) further considered the calibration and validation data in addition to model input, parameter, and structural uncertainty in watershed modeling. Ratto et al. (2001) proposed a GSA-GLUE approach for model calibration that essentially combines the two Monte Carlo techniques for global SA (FAST or Sobol) and



UA, respectively. This allows SA on raw model outputs as well as likelihood measures. The former helps in parameter fixing and research prioritization, while the latter helps in the parameter identification stage of model calibration. The GLUE part of this procedure estimates the predictive uncertainty. Because the same model runs are used for SA and UA, this is an efficient technique. Further, the tools used for sample generation, UA, and SA (e.g., SimLab and GLUE-Win) are compatible, easy to use, and freely available. Table 1 summarizes the major UA methods used in HWQMs.

## PROVIDING UNCERTAINTY RESULTS

In general, when modelers provide an uncertainty as-

essment, it is often interpreted in a negative context, thereby invalidating the modeling effort to some degree. A model, if properly applied, is the best representation of a real physical phenomenon or processes. This misconception may happen because of lack of proper communication of the meaning of uncertainty analysis output. Even during professional meetings, models are often criticized because of the existence of uncertainty. Uncertainty analysis should be used to show how confident modelers are with respect to the model output, indicating the range of expected model output rather than unique solutions. Knowing the level of confidence in the output is better than having no science-based knowledge of the limits of output acceptance (Har-

**Table 1. Summary of major uncertainty analysis methods used in hydrologic and water quality modeling.**

Method	Description	Requirement	Setup and Procedure	Examples
Monte Carlo simulation (MCS)	Involves drawing a number of stochastic parameter sets from the parameter PDFs, running the model for each set of parameters, and statistically analyzing the model output probability distribution.	MCS inputs should be representative of the probability distribution of the model parameter values if the predictive uncertainty is only based on all the MCS outputs, as is usually done. Thus, to evaluate the predictive uncertainty resulting from the input parameter uncertainty, prior knowledge of the parameter PDFs and a large number of simulations is critical for representative results.	Software: Advanced Risk and Reliability Assessment Modeling Information System (ARRAMIS), which is equipped with MCS.	MIKE SHE: Thorsen et al. (2001). DRAINMOD: Sabbagh and Fox (1999). AGNPS: Haan et al. (1998). SWAT: Shirmohammadi et al. (2006).
First-order analysis (FOA), or Gaussian approximation	Based on linearizing the relationship between a dependent random variable and a set of independent random variables by a Taylor series expansion to estimate the output variances due to the inputs.	Restricted by assumptions of linearity and coefficients of variation of input parameters less than 10% to 20%.	1. Determine base values. 2. Perform SA. 3. Select the most influential parameters for further analysis. 4. Estimate means, variances, correlations, and PDFs for the influential parameters. 5. Perform FOA.	AGNPS: Chaubey et al. (1999). CREAMS-WT: Haan and Zhang (1996).
Kalman filtering	Conditional probability density propagation.	Restricted to linear models in which system and measurement errors are normally distributed.	-	Peter (1979), Ahsam and O'Connor (1994).
Latin hypercube sampling (LHS)	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 an equal probability of occurrence.	Each variable should be divided into $n$ intervals of equal probability. Requires prior knowledge of parameter distributions. Restricted to linear and monotonic models.	Software: Advanced Risk and Reliability Assessment Modeling Information System (ARRAMIS), which is equipped with LHS.	SWAT: Shirmohammadi et al. (2006), Sohrabi et al. (2002, 2003).
Generalized likelihood uncertainty estimation (GLUE)	Based on MCS, likelihood measures, and the concept of Bayesian inference, it uses the behavior and non-behavior classification concept to distinguish between the fraction of MCS that matches the system behavior and the complement fraction that does not.	Measured values should be available to compare with the model predictions in order to calculate the likelihood of each model run corresponding to each parameter set.	Software: GLUEWIN (Ratto and Saltelli, 2001)	EPIC: Wang et al. (2005). DRAINMOD: Wang et al. (2006b). TOPMODEL: Beven and Binley (1992), Beven (1993), Freer et al. (1996), Lamb et al. (1998), Cameron et al. (1999). WEPP: Brazier et al. (2000). MIKE SHE: Christiaens and Feyen (2001, 2002).
Bayesian analysis	A probabilistic approach that uses the rules of probability theory to make predictions, compare alternative models, and express uncertainty in the model.	Bayesian approach requires prior values for unknown parameters and for other hyperparameters (e.g., missing data need to be estimated via maximum likelihood).	Multiple model structure	Ajami et al. (2007), van Oijen et al. (2013).
Mean value first-order reliability analysis method (MFORM)	A developed form of FOA that performs a Taylor series expansion of the model output function and allows the user to determine the variance of the dependent variable and the variance contributed by each input parameter.	Each basic variable should be standardized to receive equal consideration.	-	Melching and Bauwens (2001), Sexton et al. (2011b).



mel et al., 2014). Further, modelers should strive to make model output more useful by presenting a state-of-the-art assessment of uncertainty that may result in a range of model outputs or expected solutions when applied to complex systems.

In theory, modelers would not need uncertainty analysis if they were fully confident of model output. In practice, there is no perfect model that achieves that confidence level in complex systems. Higher uncertainty means lower confidence, and vice versa. Modelers perform uncertainty analysis to make sure the model outputs satisfy the expectations relative to actual system behavior. If the combined effect of all uncertainties provides a stronger signal than the model prediction, the model results might be meaningless and should not be considered in further analyses or decision-making. For example, if the change in sediment load is predicted to be 20% and the model uncertainty is found to be around 20%, then the model output is meaningless for the prediction of sediment load.

### USING UNCERTAINTY ANALYSIS OUTPUT

It is increasingly necessary to adequately address the uncertainty associated with model predictions to realize the potential of models in decision-making (Beven, 2007; Liu and Gupta, 2007; Wagener and Gupta, 2005; EPA, 2008). Without proper uncertainty analysis, and use of this information in explaining error origination and propagation, model output may not be useful. Attempts should be made to relate the level of uncertainty with physical explanations to communicate the confidence in the predictions. Quantifying and allocating the uncertainty to respective sources may be a challenging task, but it is worthwhile to correctly use the model output. This step may require developing new methods and strategies for the assessment, including additional and multi-level sensitivity analysis. For example, Liu and Gupta (2007) introduced an integrated hierarchical data assimilation framework to better understand the sources of uncertainty and reduce them in a stepwise manner.

### INCORPORATING MODEL AND MEASUREMENT UNCERTAINTY IN MODEL EVALUATION

The degree to which models adequately represent real-world processes is typically judged by pairwise comparison of measured data and model output (Legates and McCabe, 1999) and with graphical comparison of measured and predicted values. Willmott (1981), Legates and McCabe (1999), Moriasi et al. (2007), and Jain and Sudheer (2008) provided thorough discussions regarding common goodness-of-fit indicators for evaluating model performance. As traditionally applied, most quantitative indicators use a simple difference ( $O_i - P_i$ ) to represent the deviation between observed and predicted data (Legates and McCabe, 1999).

Engel et al. (2007) defined the content necessary to develop model application protocols (or quality assurance plans), which are needed to enhance the scientific validity of models and to increase the defensibility of model applications in light of regulatory, programmatic, and research implications. In discussing calibration and validation pro-

cedures, Engel et al. (2007) emphasized the need to assess model goodness-of-fit and to assess the uncertainty in model results and measured data; however, no methods for assessing goodness-of-fit considering both measurement uncertainty and model uncertainty were provided. Harmel and Smith (2007) developed a method to include uncertainty in measured calibration and validation data by modifying the traditional error term calculation in goodness-of-fit indicators, and Harmel et al. (2010) developed a method to modify goodness-of-fit indicator values considering both measurement and model uncertainty. This method is based on Haan et al. (1995), who stated that the degree of overlap between corresponding probability density functions for observed and predicted values is indicative of model predictive ability. Cibin et al. (2012) proposed a similar goodness-of-fit evaluation method, which can be used to modify the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) considering prediction uncertainty.

Such modified indicator values that incorporate uncertainty in measured calibration and validation data and model predictions were developed to provide valuable supplemental information to be used in conjunction with traditionally applied statistical and graphical model evaluation methods, including the model performance ratings of Moriasi et al. (2007). All of these methods applied and considered together provide a detailed evaluation of model performance.

### UNCERTAINTY ANALYSIS FOR VARIOUS HWQMS

Uncertainty across different models is expected to differ due to model characteristics, differences in spatial and temporal discretization, process representation, mathematical approaches, code-based model complexity, and model input requirements. HWQMs are applied to specific objectives within the boundary of the model development (level of physical phenomena representation), and modelers must be cautious when applying models outside these boundaries. Similarly, modelers must avoid application of models outside of their original principles (structural uncertainty; Refsgaard et al., 2006) and the spatial or temporal scale and range of applicability originally intended. For example, uncertainty in curve number based models may arise when the models are applied to geographical regions outside the U.S., at temporal scales of rainfall different from the original definition of "event," when used to estimate overland flow from hydrological processes different from saturation excess, or when physically based characteristics are misleadingly assigned to the curve number model (Ponce and Hawkins, 1996; Garen and More, 2005; Walter and Shaw, 2006). Because model parameters in HWQMs are expected to be sensitive to observations, uncertainty arises when parameters derived under stationary conditions are used to assess future scenarios where stationarity is unknown. Modelers should also consistently evaluate the parameters in physically based models to ensure that they fall within the domain of the representative elementary volume. The availability of databases representing spatially distributed



datasets combined with computational tools for model parameterization and cloud-based model execution are facilitating HWQM application, but they can also contribute to increased uncertainty. For example, delineation and parameterization of hydrological response units (HRUs) from the curve number definition (soil, terrain slope, and land use) are commonly based on georeferenced databases (NRCS SSURGO or STASTGO databases, USGS digital elevation models, and land use) in combination with rainfall time series to assess the impact of best management practices. In some cases, derivation of these HRUs is outside the spatial and temporal scales of application of at least one of the datasets or outside the range of applicability of the curve number model, resulting in a digital mirage. Tables 2 and 3 summarize common HWQMs and recent uncertainty analysis assessment as a function of uncertainty in parameters and data inputs, respectively.

## CONCLUSIONS AND RECOMMENDATIONS

Uncertainty is ubiquitous when assessing the hydrological response and fate and transport phenomena in complex ecosystems. Uncertainty in the inputs propagates to the model outputs in a non-linear manner as a function of the model representation and computational limits; thus, uncertainty in model outputs is always expected. At the watershed scale, observations are limited in space and time and possess convoluted errors in which decomposition is problematic. In most cases, the probability function associated with uncertainty in observational errors is unknown and non-stationary due to changes in instrumentation, protocols, network operation, and the dynamics of the system. Mathematical operations and data transformations commonly conducted on observed data to fulfill specific HWQM requirements (e.g., spatiotemporal discretization and representation, computational limits, conceptual model simplification, etc.) can exacerbate the observational uncertainty and drive changes in the spatiotemporal patterns of the observations.

Uncertainty analysis in HWQMs aims to identify and quantify the uncertainty in model outputs that is linked to model inputs or structural uncertainty in a probabilistic manner. The different methodologies that have been developed for local or global uncertainty analysis have different purposes and different computational requirements. Uncertainty analysis applied to HWQMs allows assessment of model outputs in a more realistic manner because in most cases the variability of the ecosystem under investigation is not well known. In addition, uncertainty analysis allows modelers to assess model robustness and the range of model applicability.

In the present examination of considerations related to uncertainty in HWQM modeling, the following recommendations emerged:

- Uncertainty in the measured data used for model calibration and validation impacts the model results and the evaluation of those results. Therefore, uncertainty

in calibration and validation data should be a vital consideration in modeling and should not be viewed as trivial.

- Input data resolution (spatial and temporal) should be carefully examined in decisions on model spatial discretization. Of the input data (e.g., topography, land use, soil, land management, etc.), precipitation data is probably the biggest driver of uncertainty, which is aggravated due to non-stationary conditions.
- Building a conceptual model linked to the objectives of the problem prior to model application is recommended. Errors in the conceptual model due to incomplete system representation propagate throughout the modeling process regardless of the model's mathematical sophistication, accuracy of input parameter values, and observed data. Thus, improving the knowledge of various fundamental processes and linkages is vital for reducing the uncertainty that emanates from oversimplification, overrepresentation, or the lack of fundamental processes in the model.
- Model structure selection should be realistic in that tradeoffs may exist between different structures. Even though it sounds trivial, one important way to avoid uncertainty is for the model structure to match the represented processes. For example, when knowledge of the processes in the study area is limited, using a lumped model might be better because a distributed model may add counterproductive uncertainty due to the added complexity.
- Parameter adjustment should be systematic, and the limit should be related to the interactions of hydrologic and fate and transport processes. A hierarchical sequence of parameter calibration is recommended from flow to transport (e.g., flow volume, surface runoff, groundwater, temperature, suspended sediments, conservative ions, nutrients, phytoplankton, and dissolved oxygen) to minimize uncertainty propagation. In addition, care should be taken that auto-calibration tools do not adjust parameters unrealistically (i.e., out of the range of applicability), leading to correct outputs from unrealistic inputs.
- Understanding the uncertainty associated with model predictions is vital to appropriate application of models in decision-making. Specifically, model outputs may not be useful without uncertainty analysis and proper explanation of the uncertainty sources, error propagation, and the degree of confidence in model predictions.
- The statistical significance of the amount of input data and the data quality (data screening) should be analyzed when defining conceptual models. Multiple techniques exist to increase the amount of input data (e.g., Monte Carlo methods) or to provide spatial inferences of quantities from observed data points (e.g., kriging) that may increase the uncertainty in inputs and thus the uncertainty in model outputs.



**Table 2. Summary of recent parameter uncertainty analysis studies in watershed and hydrologic modeling.**

Model	Parameters	Method <sup>(a)</sup>	Runs	Evaluated Output <sup>(b)</sup>	Example
ADAPT <sup>(c)</sup>	NA	NA	NA	NA	NA
	7	MCMC	>250	NA	HSPF: Gallagher and Doherty (2007)
BASIN/HSPF <sup>(c)</sup>	40	MCS, GLUE, and MCMC	12,000 for MCS, 10,000 for GLUE, 100,000 for MCMC	Various outputs describing fecal coliform concentrations.	Mishra, 2011)
CoupModel <sup>(c)</sup>	30	GLUE	NA	ME, RMSE, and R <sup>2</sup> of various nitrogen leaching output	Conrad and Fohrer (2009)
	49	GLUE	30,000 based on MCS	ME, R <sup>2</sup>	Wu et al. (2011)
CREAM/GLEAMS <sup>(c)</sup>	NA	NA	NA	NA	NA
Daisy <sup>(c)</sup>	NA	NA	NA	NA	NA
Distributed rainfall-runoff model	3 (5 storm events)	MCS, LHS, RPEM, and HPEM	<1,000	Different for different methods	Yu et al. (2001)
DRAINMOD <sup>(c)</sup>	8	GLUE	1,992	MSE and its variants for flow	Wang et al. (2006b)
EPIC and APEX <sup>(c)</sup>	9	GLUE	1,500	Confidence interval for crop yield and soil organic carbon	Wang et al. (2005)
HYDRUS <sup>(c)</sup>	7	GLUE	30,000	NSE	Hansson and Lundin (2006)
HYMOD	5	MCMC	NA	Multiple	Vrugt et al. (2008)
HYMOD, NAM, and SAC-SMA	5 for HYMOD, 10 for NAM, 14 for SAC-SMA	Combination of MCMC and GLUE	NA	Exponential likelihood function	Blasone et al. (2008)
IHDM	4	GLUE	500 (× 5 storm events)	NSEC, sum of squares of residuals, scaled maximum absolute error	Beven and Binley (1992)
	23	GLUE	98,304	Multiple likelihood functions	Yatheन्द्रadas et al. (2008)
KINEROS/AGWA <sup>(c)</sup>	9	MC	1,000	Coefficients of variation, probability of exceedance	Hantush and Kalin (2005)
MACRO <sup>(c)</sup>	NA	NA	NA	NA	NA
Micro-pollutants in streamflow	5 most sensitive parameters from SA	GLUE	10,000	Multiple	Vezzaro and Mikkelsen (2012)
MIKE SHE <sup>(c)</sup>	8	GLUE	15,000	NSEC-based likelihood for streamflow	Vázquez et al. (2009)
MT3DMS <sup>(c)</sup>	NA	NA	NA	NA	NA
P export and discharge	8 parameters	GLUE	30,000	RMSE of TP concentrations	Smith and Wheatler (2004)
RZWQM <sup>(c)</sup>	NA	NA	NA	NA	NA
Sediment nutrient dynamics model for section of Chattahoochee River south (STAND)	5 management related parameters, and 14 sediment and nutrient related parameters	MC technique: RSA-UCPR-TSDE	Not reported	Multiple criteria	Osidade et al. (2003)
SHAW <sup>(c)</sup>	NA	NA	NA	NA	NA
STANMOD <sup>(c)</sup>	NA	NA	NA	NA	NA
	13 parameters, 12 HRU configurations	Single-factor between-subjects analysis	500 (× 12 HRU combo)	Percent relative error, coefficient of variability	Migliaccio and Chaubey (2008)
SWAT <sup>(c)</sup>	35	GLUE	5,000	NSEC for discharge and sediment concentration and their combinations	Muleta and Nicklow (2005)
	11	Bayesian model averaging	1,000		Zhang et al. (2009a)
	21	MC, FOEA	<100	Multiple	Shen et al. (2008)
SWAT-R	8	MCMC	1,000	Multiple criteria on streamflow	Wu and Liu (2012)
SWIM <sup>(c)</sup>	NA	NA	NA	NA	NA
TOUGH2 <sup>(c)</sup>	NA	NA	NA	NA	NA
	27	CDFs and other statistics of outputs	<8,000	Multiple	Muñoz-Carpena et al. (2007)
VFSMOD	18	CDFs and other statistics of outputs	<14,977	Multiple	Fox et al. (2010)
VS2DI <sup>(c)</sup>	NA	NA	NA	NA	NA
WAM <sup>(c)</sup>	NA	NA	NA	NA	NA
WARMF <sup>(c)</sup>	NA	NA	NA	NA	NA
WASMOD	3	GLUE and Bayesian method	50,000	Multiple based on flow	Jin et al. (2010)
WEPP <sup>(c)</sup>	NA	NA	NA	NA	NA

<sup>(a)</sup> FOEA = first-order error estimate, GLUE = generalized likelihood uncertainty estimation, HPEM = Harr's point estimation method, MC = Monte Carlo, MCMC = Markov chain Monte Carlo, MCS = Monte Carlo simulation, NA = not applicable, RPEM = Rosenblueth's point estimation method, and UCPR = uniform covering by probabilistic rejection.  
<sup>(b)</sup> MSE = mean square error, and NSE = Nash-Sutcliffe efficiency coefficient.  
<sup>(c)</sup> Models that are part of the model use, calibration, and validation special collection (Moriassi et al., 2012).



**Table 3. Summary of recent studies on effect of data and input uncertainties in watershed and hydrologic modeling.**

Model	Analysis	Details	Example
ADAPT <sup>[a]</sup>	NA	NA	NA
AGNPS	Effect of spatial rainfall variability on parameter uncertainty.	17 rainfall distribution scenarios.	Chaubey et al. (1999)
BASIN/HSPF <sup>[a]</sup>	Effect of temporal and spatial data resolution on output uncertainty.	(1) Climate data (temperature and rainfall) at five different temporal resolutions were used to assess the impact on model output. (2) Five different resolutions of land use data were used to assess the impacts of spatial resolution.	HSPF: Patil et al. (2011)
CoupModel <sup>[a]</sup>	NA	NA	NA
CREAM/GLEAMS <sup>[a]</sup>	NA	NA	NA
Daisy <sup>[a]</sup>	NA	NA	NA
DRAINMOD <sup>[a]</sup>	NA	NA	NA
Empirical soil loss model	Effect of DEM resolution on soil loss.	Seven different DEM resolutions were used.	Wu et al. (2005)
EPIC and APEX <sup>[a]</sup>	Accounting for model parameter, input, and model structural uncertainty.	Developed an auto-calibration and uncertainty analysis tool for APEX (APEX-CUTE). Examined effects of model input error, model parameters, and model structural uncertainty on sediment prediction.	Wang and Yen (2014) and Wang et al. (2014a, 2014b)
HL-DHMS	Effect of rainfall characteristics and initial moisture conditions on spatial parameter sensitivity.	Six combinations of rainfall and initial soil moisture characteristics were used for spatial SA. For each case, spatial SA indices were calculated.	van Werkhoven et al. (2008)
HYDRUS <sup>[a]</sup>	NA	NA	NA
KINEROS/AGWA <sup>[a]</sup>	Effect of rainfall intensity on model output uncertainty.	Ten storm events of 120 min duration.	Hantush and Kalin (2005)
KINEROS2	Effect of land use misclassification on output (runoff) uncertainty.	100 realizations based on land use misclassification error were generated.	Miller et al. (2007)
MACRO <sup>[a]</sup>	Effect of variability in input parameter values on uncertainty in model output.	(1) Developed PDFs for each input parameter. (2) Used LHS-CMS to determine the output distribution due to variability in input data.	Sohrabi et al. (2002)
MIKE SHE <sup>[a]</sup>	Effect of grid size on parameters and model performance.	Three grid sizes were used (300, 600, and 1200 m).	Vázquez et al. (2002)
MT3DMS <sup>[a]</sup>	NA	NA	NA
RZWQM <sup>[a]</sup>	NA	NA	NA
SAC-SMA	Parameter SA at different time scales.	Four different SA methods were used to evaluate the effects of three different time scales on parameter sensitivity.	Tang et al. (2007)
SHAW <sup>[a]</sup>	NA	NA	NA
STANMOD <sup>[a]</sup>	NA	NA	NA
SWAT <sup>[a]</sup>	Effect of DEM resolution on output uncertainty.	Seven DEM resolution scenarios.	Chaubey et al. (2005)
	Effect of spatial distribution of rainfall on temporal and spatial uncertainty.	(1) Effect of rainfall input method in conjunction with subwatershed delineation (nine scenarios). (2) Effect of raingauge density on hydrology and water quality (ten scenarios).	Cho et al. (2009)
	Effect of spatial resolution of DEM, land use, and soil on water quality model output uncertainty.	Seven resolution scenarios.	Cotter et al. (2003)
	Sensitivity of SWAT to spatial scale.	Calibrated parameter values for six different discretization scenarios were compared.	Muleta et al. (2007)
	Impact of time scale of objective function (e.g., daily, annual, etc.) on model performance.	Model performance measures were compared for daily and monthly time scales.	Sudheer et al. (2007)
	Impact of parameterization, precipitation, choice of surface runoff method (model structure), and calibration and validation data on flow and water quality.	Different calibration cases were considered. The best solutions and predictive uncertainties of flow and water quality were compared.	Yen et al. (2014)
SWIM <sup>[a]</sup>	NA	NA	NA
TOPMODEL	Effect of DEM resolution and elevation data uncertainty.	Seven DEM resolution scenarios and thousands of realizations based on elevation errors were analyzed.	Wu et al. (2007)
TOUGH2 <sup>[a]</sup>	NA	NA	NA
VS2DI <sup>[a]</sup>	NA	NA	NA
WAM <sup>[a]</sup>	NA	NA	NA
WARMF <sup>[a]</sup>	NA	NA	NA
WEPP	NA	NA	NA

<sup>[a]</sup> Models that are part of the model use, calibration, and validation special collection (Moriiasi et al., 2012).



- Uncertainty analysis should be used to assess model outputs in a positive context. There is need to improve the manner in which modelers communicate the meaning of uncertainty analysis. As a matter of fact, well developed models show larger certainty metrics than uncertainty in their outputs.
- In determining model uncertainty, a proper and less cumbersome method should be recommended and used. For example, instead of using Monte Carlo simulation alone, the use of Latin hypercube sampling with Monte Carlo simulation can limit the number of simulations while providing relatively accurate output distribution.

## REFERENCES

- Ahsam, M., & O'Connor, K. M. (1994). A reappraisal of the Kalman filtering technique as applied in river flow forecasting. *J. Hydrol.*, 161(1), 197-226. [http://dx.doi.org/10.1016/0022-1694\(94\)90129-5](http://dx.doi.org/10.1016/0022-1694(94)90129-5).
- Ajami, N. K., Duan, Q., & Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Res.*, 43(1), 1-19. <http://dx.doi.org/10.1029/2005WR004745>.
- Arabi, M., Govindaraju, R. S., Engel, B., & Hantush, M. (2007). Multi-objective sensitivity analysis of sediment and nitrogen processes with a watershed model. *Water Resources Res.*, 43(6), 1-11. <http://dx.doi.org/10.1029/2006WR005463>.
- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., van Griensven, A., Van Liew, M. W., Kannan, N., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1494-1508. <http://dx.doi.org/10.13031/2013.42256>.
- Beck, M. B. (1987). Water quality modeling: A review of the analysis of uncertainty. *Water Resources Res.*, 23(8), 1393-1442. <http://dx.doi.org/10.1029/WR023i008p01393>.
- Beck, M. B. (2009). Grand challenges of the future for environmental modeling. White paper. Arlington, Va.: National Science Foundation. Retrieved from <http://cfnnet.org/archives/249>.
- Beeson, P. C., Doraiswamy, P. C., Sadeghi, A. M., Di Luzio, M., Tomer, M. D., Arnold, J. G., & Daughtry, C. S. T. (2011). Treatments of precipitation inputs to hydrologic models. *Trans. ASABE*, 54(6), 2011-2020. <http://dx.doi.org/10.13031/2013.40652>.
- Beeson, P. C., Sadeghi, A. M., Lang, M. W., Tomer, M. D., & Daughtry, C. S. T. (2014). Sediment delivery estimates in water quality models altered by resolution and source of topographic data. *J. Environ. Qual.*, 43, 26-36. <http://dx.doi.org/10.2134/jeq2012.0148>.
- Beven, K. J. (1989). Changing ideas in hydrology: The case of physically based models. *J. Hydrol.*, 105(1-2), 157-172. [http://dx.doi.org/10.1016/0022-1694\(89\)90101-7](http://dx.doi.org/10.1016/0022-1694(89)90101-7).
- Beven, K. J. (1993). Prophecy, reality, and uncertainty in distributed hydrological modeling. *Adv. Water Resources*, 16(1), 41-51. [http://dx.doi.org/10.1016/0309-1708\(93\)90028-E](http://dx.doi.org/10.1016/0309-1708(93)90028-E).
- Beven, K. J. (2001). *Rainfall-Runoff Modeling: The Primer*. Chichester, U.K.: John Wiley and Sons.
- Beven, K. J. (2006). A manifesto for the equifinality thesis. *J. Hydrol.*, 320(1-2), 18-36.
- Beven, K. J. (2007). Towards integrated environmental models of everywhere: Uncertainty, data, and modeling as a learning process. *Hydrol. Earth Syst. Sci.*, 11(1), 460-467. <http://dx.doi.org/10.5194/hess-11-460-2007>.
- Beven, K. J., & Binley, A. M. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Proc.*, 6(3), 279-298. <http://dx.doi.org/10.1002/hyp.3360060305>.
- Beven, K. J., Smith, P. J., & Wood, A. (2011). On the colour and spin of epistemic error (and what we might do about it). *Hydrol. Earth Syst. Sci.*, 15, 3123-3133. <http://dx.doi.org/10.5194/hess-15-3123-2011>.
- Blasone, R. S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., & Zyvoloski, G. A. (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling. *Adv. Water Resources*, 31(4), 630-648. <http://dx.doi.org/10.1016/j.advwatres.2007.12.003>.
- Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrol. Proc.*, 9(3-4), 251-290. <http://dx.doi.org/10.1002/hyp.3360090305>.
- Brazier, R. E., Beven, K. J., Freer, J., & Rowan, J. S. (2000). Equifinality and uncertainty in physically based soil erosion models: Application of the GLUE methodology to WEPP: The Water Erosion Prediction Project for sites in the U.K. and U.S. *Earth Surface Proc. Landforms*, 25(8), 825-845. [http://dx.doi.org/10.1002/1096-9837\(200008\)25:8<825::AID-ESP101>3.0.CO;2-3](http://dx.doi.org/10.1002/1096-9837(200008)25:8<825::AID-ESP101>3.0.CO;2-3).
- Briggs, A. H., Weinstein, M. C., Fenwick, E. A. L., Karnon, J., Sculpher, M. J., & Paltiel, A. D. (2012). Model parameter estimation and uncertainty: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health* 15(6), 835-842.
- Cameron, D. S., Beven, K. J., Tawn, J., Blazkova, S., & Naden, P. (1999). Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *J. Hydrol.*, 219(3-4), 169-187. [http://dx.doi.org/10.1016/S0022-1694\(99\)00057-8](http://dx.doi.org/10.1016/S0022-1694(99)00057-8).
- Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environ. Model. Software*, 22(10), 1509-1518. <http://dx.doi.org/10.1016/j.envsoft.2006.10.004>.
- Chaubey, I., Haan, C. T., Grunwald, S., & Salsbury, J. M. (1999). Uncertainty in model parameters due to spatial variability of rainfall. *J. Hydrol.*, 220(1-2), 48-61. [http://dx.doi.org/10.1016/S0022-1694\(99\)00063-3](http://dx.doi.org/10.1016/S0022-1694(99)00063-3).
- Chaubey, I., Cotter, A. S., Costello, T. A., & Soerens, T. S. (2005). Effect of DEM data resolution on SWAT output uncertainty. *Hydrol. Proc.*, 19(3), 621-628. <http://dx.doi.org/10.1002/hyp.5607>.
- Cho, J., Bosch, D., Lowrance, R., Strickland, T., & Vellidis, G. (2009). Effect of spatial distribution of rainfall on temporal and spatial uncertainty of SWAT output. *Trans. ASABE*, 52(5), 1545-1555. <http://dx.doi.org/10.13031/2013.29143>.
- Christiaens, K., & Feyen, J. (2001). Analysis of uncertainties associated with different methods to determine soil hydraulic properties and their propagation in the distributed hydrological MIKE SHE model. *J. Hydrol.*, 246(1-4), 63-81. [http://dx.doi.org/10.1016/S0022-1694\(01\)00345-6](http://dx.doi.org/10.1016/S0022-1694(01)00345-6).
- Christiaens, K., & Feyen, J. (2002). Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrol. Proc.*, 16(2), 373-391. <http://dx.doi.org/10.1002/hyp.335>.
- Chu-Agor, M. L., Muñoz-Carpena, R., Kiker, G., Emanuelsson, A., & Linkov, I. (2011). Exploring sea level rise vulnerability of coastal habitats through global sensitivity and uncertainty analysis. *Environ. Model. Software*, 26(5), 593-604. <http://dx.doi.org/10.1016/j.envsoft.2010.12.003>.
- Chu-Agor, M. L., Muñoz-Carpena, R., Kiker, G. A., Aiello-Lammens, M., Akçakaya, R., Convertino, M., & Linkov, I.



- (2012). Simulating the fate of Florida snowy plovers with sea level rise: Exploring potential population management outcomes with a global uncertainty and sensitivity analysis perspective. *Ecol. Model.*, 224, 33-47. <http://dx.doi.org/10.1016/j.ecolmodel.2011.10.021>.
- Chu, M. L., Ghulam, A., Knouft, J. H., & Pan, Z. (2013). A hydrologic data screening procedure for exploring monotonic trends and shifts in rainfall and runoff patterns. *JAWRA*, 50(4), 928-942. <http://dx.doi.org/10.1111/jawr.12149>.
- Cibin, R., Chaubey, I., & Engel, B. (2012). Simulated watershed-scale impacts of corn stover removal for biofuels on hydrology and water quality. *Hydrol. Proc.*, 26(11), 1629-1641. <http://dx.doi.org/10.1002/hyp.8280>.
- Conrad, Y., & Fohrer, N. (2009). A test of CoupModel for assessing the nitrogen leaching in grassland systems with two different fertilization levels. *J. Plant Nutrition Soil Sci.*, 172(6), 745-756. <http://dx.doi.org/10.1002/jpln.200800264>.
- Constanza-Robinson, M. S., Estabrook, B. D., & Fouhey, D. F. (2011). Representative elementary volume estimation for porosity, moisture saturation, and air-water interfacial areas in unsaturated porous media: Data quality implications. *Water Resources Res.*, 47(7), W07513. <http://dx.doi.org/10.1029/2010WR009655>.
- Cotter, A. S., Chaubey, I., Costello, T. A., Soerens, T. S., & Nelson, M. A. (2003). Water quality model output uncertainty as affected by spatial resolution of input data. *JAWRA*, 39(4), 977-986. <http://dx.doi.org/10.1111/j.1752-1688.2003.tb04420.x>.
- Cukier, R. I., Fortuin, C. M., Schuler, K. E., Petschek, A. G., & Schaibly, J. H. (1973). Study of sensitivity of coupled reaction systems to uncertainties in rate coefficients: I. Theory. *J. Chem. Physics*, 59(8), 3873-3878. <http://dx.doi.org/10.1063/1.1680571>.
- Doherty, J. (2004). PEST: Model independent parameter estimation, User manual (5th ed.). Retrieved from [www.pesthomepage.org/Downloads.php](http://www.pesthomepage.org/Downloads.php).
- Engel, B., Storm, D., White, M., Arnold, J., & Arabi, M. (2007). A hydrologic/water quality model application protocol. *JAWRA*, 43(5), 1223-1236. <http://dx.doi.org/10.1111/j.1752-1688.2007.00105.x>.
- EPA. (2002). The twenty needs report: How research can improve the TMDL program. EPA 841-B-02-002. Washington, D.C.: U.S. Environmental Protection Agency.
- EPA. (2008). Integrated modeling for integrated environmental decision making. EPA 100/R-08/010. Washington, D.C.: U.S. Environmental Protection Agency.
- Fox, G. A., Muñoz-Carpena, R., & Sabbagh, G. J. (2010). Influence of flow concentration on parameter importance and prediction uncertainty of pesticide trapping by vegetative filter strips. *J. Hydrol.* 384(1-2): 164-173.
- Francos, A., Elorza, F. J., Bouraoui, F., Bidoglio, G., & Galbiati, L. (2003). Sensitivity analysis of distributed environmental simulation models: Understanding the model behavior in hydrological studies at the catchment scale. *Reliability Eng. Syst. Safety*, 79(2), 205-218. [http://dx.doi.org/10.1016/S0951-8320\(02\)00231-4](http://dx.doi.org/10.1016/S0951-8320(02)00231-4).
- Freer, J., Beven, K., & Ambrose, B. (1996). Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Res.*, 32(2), 2161-2173.
- Gallagher, M., & Doherty, J. (2007). Parameter estimation and uncertainty analysis for a watershed model. *Environ. Model. Software*, 22(7), 1000-1020. <http://dx.doi.org/10.1016/j.envsoft.2006.06.007>.
- Garen, D. C., & Moore, D. S. (2005). Curve number hydrology in water quality modeling: Uses, abuses, and future directions. *JAWRA*, 41(2), 377-388. <http://dx.doi.org/10.1111/j.1752-1688.2005.tb03742.x>.
- Graham, D. N., & Butts, M. B. (2005). Flexible, integrated watershed modelling with MIKE SHE. In V. P. Singh, & D. K. Frevert (Eds.), *Watershed Models* (pp. 245-272). Boca Raton, Fla.: CRC Press.
- Grayson, R. B., Blöschl, G., Western, A. W., & McMahon, T. A. (2002). Advances in the use of observed spatial patterns of catchment hydrology response. *Adv. Water Resources*, 25(8-12), 1313-1334. [http://dx.doi.org/10.1016/S0309-1708\(02\)00060-X](http://dx.doi.org/10.1016/S0309-1708(02)00060-X).
- Groisman, P. Y., & Legates, D. R. (1994). The accuracy of United States precipitation data. *Bull. American Meteorol. Soc.*, 75(2), 215-227. [http://dx.doi.org/10.1175/1520-0477\(1994\)075<0215:TAOUSP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2).
- Guzman, J. A., Chu, M. L., Stark, P. J., Moriasi, D. N., Steiner, J. L., Fiebrich, C. A., & McCombs, A. G. (2014). Upper Washita River experimental watersheds: Data screening procedure for data quality assurance. *J. Environ. Qual.*, 43(4), 1250-1261. <http://dx.doi.org/10.2134/jeq2013.08.0325>.
- Haan, C. T. (1989). Parametric uncertainty in hydrologic modeling. *Trans. ASAE*, 32(1), 137-146. <http://dx.doi.org/10.13031/2013.30973>.
- Haan, C. T., & Zhang, J. (1996). Impact of uncertain knowledge of model parameters on estimated runoff and phosphorus loads in the Lake Okeechobee basin. *Trans. ASAE*, 39(2), 511-516. <http://dx.doi.org/10.13031/2013.27529>.
- Haan, C. T., Allred, B., Storm, D. E., Sabbagh, G. J., & Prahhu, S. (1995). Statistical procedure for evaluating hydrologic/water quality models. *Trans. ASAE*, 38(3), 725-733. <http://dx.doi.org/10.13031/2013.27886>.
- Haan, C. T., Storm, D. E., Al-Issa, T., Prabhu, S., Sabbagh, G. J., & Edwards, D. R. (1998). Effect of parameter distributions on uncertainty analysis of hydrologic models. *Trans. ASAE*, 41(1), 65-70. <http://dx.doi.org/10.13031/2013.17158>.
- Hansson, K., & Lundin, L.-S. (2006). Equifinality and sensitivity in freezing and thawing simulations of laboratory and *in situ* data. *Cold Region Sci. Tech.*, 44(1), 20-37. <http://dx.doi.org/10.1016/j.coldregions.2005.06.004>.
- Hantush, M., & Kalin, L. (2005). Uncertainty and sensitivity analysis of runoff and sediment yield in a small agricultural watershed. *Hydrol. Sci.*, 50(6), 1151-1171.
- Harmel, R. D., & Smith, P. K. (2007). Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *J. Hydrol.*, 337(3-4), 326-336. <http://dx.doi.org/10.1016/j.jhydrol.2007.01.043>.
- Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., & Arnold, J. G. (2006). Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE*, 49(3), 689-701. <http://dx.doi.org/10.13031/2013.20488>.
- Harmel, R. D., Smith, D. R., King, K. W., & Slade, R. M. (2009). Estimating storm discharge and water quality data uncertainty: A software tool for monitoring and modeling applications. *Environ. Model. Software*, 24(7), 832-842. <http://dx.doi.org/10.1016/j.envsoft.2008.12.006>.
- Harmel, R. D., Smith, P. K., & Migliaccio, K. L. (2010). Modifying goodness-of-fit indicators to incorporate both measurement and model uncertainty in model calibration and validation. *Trans. ASABE*, 53(1), 55-63. <http://dx.doi.org/10.13031/2013.29502>.
- Harmel, R. D., Smith, P. K., Migliaccio, K. W., Chaubey, I., Douglas-Mankin, K. R., Benham, B., Shukla, S., Muñoz-Carpena, R., & Robson, B. J. (2014). Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. *Environ. Model. Software*, 57, 40-51. <http://dx.doi.org/10.1016/j.envsoft.2014.02.013>.
- Helton, J. C., & Davis, F. J. (2002). Illustration of sampling-based methods for uncertainty and sensitivity analysis. *Risk Analysis*, 22(3), 591-622. <http://dx.doi.org/10.1111/0272-4332.00041>.



- Herr, J. W., & Chen, C. W. (2012). WARMF: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1387-1396. <http://dx.doi.org/10.13031/2013.42249>.
- Heuvelink, G. B. M. (1998). Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling Agroecosyst.*, 50(1-3), 255-264. <http://dx.doi.org/10.1023/A:1009700614041>.
- Hiebeler, D. E., & Michaud, I. J. (2012). Quantifying spatial and temporal variability of spatially correlated disturbances. *Ecol. Model.*, 240, 64-73.
- Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Sci.*, 14(4), 382-417.
- Hornberger, G. M., & Spear, R. C. (1981). An approach to the preliminary analysis of environmental systems. *J. Environ. Mgmt.*, 12(1), 7-18.
- Iman, R. L., & Shortencarier, M. J. (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., Davenport, J. M., & Zeigler, D. K. (1980). Latin hypercube sampling (program user's guide). Technical Report SAND79-1473. Albuquerque, N.M.: Sandia National Laboratories.
- Jain, S. K., & Sudheer, K. P. (2008). Fitting of hydrologic models: A close look at the Nash-Sutcliffe index. *J. Hydrol. Eng.*, 13(10), 981-986. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2008\)13:10\(981\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2008)13:10(981)).
- Jha, M. K., & Gassman, P. W. (2014). Changes in hydrology and streamflow as predicted by modeling experiment forced with climate models. *Hydrol. Proc.*, 28(5), 2772-2781. <http://dx.doi.org/10.1002/hyp.9836>.
- Jha, M. K., Gassman, P. W., Secchi, S., Gu, R., & Arnold, J. G. (2004). Impact of watershed subdivision level on flows, sediment loads, and nutrient losses predicted by SWAT. *JAWRA*, 40(3), 811-825. <http://dx.doi.org/10.1111/j.1752-1688.2004.tb04460.x>.
- Jha, M. K., Gassman, P. W., & Panagopoulos, Y. (2013). Regional changes in nitrogen loadings in the upper Mississippi River basin under predicted mid-century climate. *Regional Environ. Change*, 15(3), 449-460. [10.1007/s.10113-013-05539-y](http://dx.doi.org/10.1007/s.10113-013-05539-y).
- Jin, X., Xu, C., Zhang, Q., & Singh, V. P. (2010). Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrologic model. *J. Hydrol.*, 383(3-4), 147-155. <http://dx.doi.org/10.1016/j.jhydrol.2009.12.028>.
- Kavetski, D., Franks, S. W., & Kuczera, G. (2003). Confronting input uncertainty in environmental modeling. In Q. Duan, H. Gupta, V. S. Sorooshian, A. N. Rousseau, & R. Turcotte (Eds.), *Calibration of Watershed Models* (pp. 49-68). Washington, D.C.: American Geophysical Union.
- Kelly, M., Tuxen, K. A., & Stralberg, D. (2011). Mapping changes to vegetation pattern in a restoring wetland: Finding pattern metrics that are consistent across spatial scale and time. *Ecol. Indicators*, 11(2), 263-273. <http://dx.doi.org/10.1016/j.ecolind.2010.05.003>.
- Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Res.*, 42(3). <http://dx.doi.org/10.1029/2005WR004362>.
- Kirkby, M. J. (1976). Test of the random network model, and its application to basin hydrology. *Earth Surface Proc.*, 1(3), 197-212. <http://dx.doi.org/10.1002/esp.3290010302>.
- Kotlash, A. R., & Chessman, B. C. (1998). Effects of water sample preservation and storage on nitrogen and phosphorus determinations: Implications for the use of automated sampling equipment. *Water Res.*, 32(12), 3731-3737. [http://dx.doi.org/10.1016/S0043-1354\(98\)00145-6](http://dx.doi.org/10.1016/S0043-1354(98)00145-6).
- Lamb, R., Beven, K., & Myrø, S. (1998). Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Adv. Water Resources*, 22(4), 305-317. [http://dx.doi.org/10.1016/S0309-1708\(98\)00020-7](http://dx.doi.org/10.1016/S0309-1708(98)00020-7).
- Lapedes, D. N. (1978). *McGraw-Hill Dictionary of Scientific and Technical Terms*. (2nd ed.). New York, N.Y.: McGraw-Hill.
- Legates, D. R., & McCabe Jr., G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resource Res.*, 35(1), 233-241. <http://dx.doi.org/10.1029/1998WR900018>.
- Lenhardt, T., Eckhardt, K., Fohrer, N., & Frede, H.-G. (2002). Comparison of two different approaches of sensitivity analysis. *Physics Chem. Earth*, 27(9-10), 645-654. [http://dx.doi.org/10.1016/S1474-7065\(02\)00049-9](http://dx.doi.org/10.1016/S1474-7065(02)00049-9).
- Lilburne, L., & Tarantola, S. (2009). Sensitivity analysis of spatial models. *Int. J. Geographic Info. Sci.*, 23(2), 151-168. <http://dx.doi.org/10.1080/13658810802094995>.
- Liu, J., Williams, J. R., Wang, X., & Yang, H. (2009). Using MODAWEC to generate daily weather data for the EPIC model. *Environ. Model. Software*, 24(5), 655-664. <http://dx.doi.org/10.1016/j.envsoft.2008.10.008>.
- Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Res.*, 43(7), 1-18, W07401. <http://dx.doi.org/10.1029/2006WR005756>.
- Loague, K., & Corwin, D. L. (1996). Uncertainty in regional-scale assessment of nonpoint-source pollution. In *Application of GIS to the Modeling of Nonpoint-Source Pollution in the Vadose Zone* (pp. 131-152). Special Pub. 48. Madison, Wisc.: SSSA.
- Loe, L. E., Bonenfant, C., Meisingset, E. L., & Myrø, A. (2012). Effects of spatial scale and sample size in GPS-based species distribution models: Are the best models trivial for red deer management? *European J. Wildlife Res.*, 58(1), 195-203. <http://dx.doi.org/10.1007/s10344-011-0563-5>.
- Loucks, D. P., van Beek, E., Stedinger, J. R., Dijkman, J. P., & Villars, M. T. (2005). *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*. Paris, France: UNESCO.
- Manache, G., & Melching, C. S. (2008). Identification of reliable regression and correlation based sensitivity measures for importance ranking of water quality model parameters. *Environ. Model. Software*, 23(5), 549-562. <http://dx.doi.org/10.1016/j.envsoft.2007.08.001>.
- Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Res.*, 45(6), W06421. <http://dx.doi.org/10.1029/2008WR007301>.
- Melching, C. S. (1995). Chapter 3: Reliability estimation. In V. P. Singh (Ed.), *Computer Models of Watershed Hydrology*. Highlands Ranch, Colo.: Water Resources Publications.
- Melching, C. S., & Bauwens, W. (2001). Uncertainty in coupled nonpoint sources and stream water quality models. *J. Water Resources Planning Mgmt.*, 127(6), 403-413. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(2001\)127:6\(403\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(2001)127:6(403)).
- Melching, C. S., & Yoon, C. G. (1996). Key source of uncertainty in QUAL2E model of Passaic River. *J. Water Resources Planning Mgmt.*, 122(2), 105-113. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1996\)122:2\(105\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1996)122:2(105)).
- Michev, D. I. (1996). Finite volume and finite volume element methods for nonsymmetric problems. PhD diss. College Station, Tex.: Texas A&M University, Department of Mathematics.
- Migliaccio, K. W., & Chaubey, I. (2008). Spatial distributions and stochastic parameter influence on SWAT flow and sediment



- prediction. *J. Hydrol.*, 13(4), 258-269.
- Miller, S. N., Guertin, D. P., & Goodrich, D. C. (2007). Hydrologic modeling uncertainty resulting from land cover misclassification. *JAWRA*, 43(4), 1065-1075.
- Mishra, A. (2011). Estimating uncertainty in HSPF based water quality model: Application of Monte Carlo based techniques. PhD diss. Blacksburg, Va.: Virginia Tech, Department of Biological Systems Engineering.
- Montgomery, R. H., & Sanders, T. G. (1986). Uncertainty in water quality data. *Developments in Water Sci.*, 27, 17-29. [http://dx.doi.org/10.1016/S0167-5648\(08\)70781-0](http://dx.doi.org/10.1016/S0167-5648(08)70781-0).
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, 50(3), 885-900. <http://dx.doi.org/10.13031/2013.23153>.
- Moriasi, D. N., Wilson, B. N., Douglas-Mankin, K. R., Arnold, J. G., & Gowda, P. H. (2012). Hydrologic and water quality models: Use, calibration, and validation. *Trans. ASABE*, 55(4), 1241-1247. <http://dx.doi.org/10.13031/2013.42265>.
- Morris, D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161-174. <http://dx.doi.org/10.1080/00401706.1991.10484804>.
- Muleta, M. K., & Nicklow, J. W. (2005). Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *J. Hydrol.*, 306(1-4), 127-145. <http://dx.doi.org/10.1016/j.jhydrol.2004.09.005>.
- Muleta, M. K., Nicklow, J. W., & Bekele, E. G. (2007). Sensitivity of a distributed watershed simulation model to spatial scale. *J. Hydrol. Eng.*, 12(2), 163-172. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:2\(163\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2007)12:2(163)).
- Muñoz-Carpena, R., Zajac, Z., & Kuo, Y. M. (2007). Global sensitivity and uncertainty analyses of the water quality model VFSMOD-W. *Trans. ASABE*, 50(5), 1719-1732. <http://dx.doi.org/10.13031/2013.23967>.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: Part I. A discussion of principles. *J. Hydrol.*, 10(3), 282-290. [http://dx.doi.org/10.1016/0022-1694\(70\)90255-6](http://dx.doi.org/10.1016/0022-1694(70)90255-6).
- NAS. (2001). *Assessing the TMDL Approach to Water Quality Management*. Washington, D.C.: National Academy Press. Retrieved from [www.nap.edu/openbook.php?record\\_id=10146&page=R1](http://www.nap.edu/openbook.php?record_id=10146&page=R1).
- NRC. (2001). *Assessing the TMDL approach to water quality management*. Washington, D.C.: National Research Council.
- Pappenberger, F., & Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Res.*, 42(5), W05302. <http://dx.doi.org/10.1029/2005WR004820>.
- Patil, A., Deng, Z. Q., & Malone, R. F. (2011). Input data resolution induced uncertainty in watershed modeling. *Hydrol. Proc.*, 25(14), 2302-2312. <http://dx.doi.org/10.1002/hyp.8018>.
- Peter, S. M. (1979). *Stochastic Models, Estimation, and Control: Volume 1*. New York, N.Y.: Academic Press.
- Ponce, V. M., & Hawkins, R. H. (1996). Runoff curve number: Has it reached maturity? *J. Hydrol. Eng.*, 1(1), 11-19. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(1996\)1:1\(11\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(1996)1:1(11)).
- Poveda, G. (2011). Mixed memory, (non) Hurst effect, and maximum entropy of rainfall in the tropical Andes. *Adv. Water Resources*, 34(2), 243-256. <http://dx.doi.org/10.1016/j.advwatres.2010.11.007>.
- Ratto, M., Tarantola, S., & Saltelli, A. (2001). Sensitivity analysis in model calibration: GSA-GLUE approach. *Comput. Physics Comm.*, 136(3), 212-224. [http://dx.doi.org/10.1016/S0010-4655\(01\)00159-X](http://dx.doi.org/10.1016/S0010-4655(01)00159-X).
- Reckhow, K. H. (1994). Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecol. Model.*, 72(1-2), 1-20. [http://dx.doi.org/10.1016/0304-3800\(94\)90143-0](http://dx.doi.org/10.1016/0304-3800(94)90143-0).
- Refsgaard, J. C., van der Sluijs, J. P., Brown, J., & van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Adv. Water Resources*, 29(11), 1586-1597. <http://dx.doi.org/10.1016/j.advwatres.2005.11.013>.
- Reusser, D. E., Buytaert, W., & Zehe, E. (2011). Temporal dynamics of model parameter sensitivity for computationally intensive models with the Fourier amplitude sensitivity test. *Water Resources Res.*, 47(7), W07551. <http://dx.doi.org/10.1029/2010WR009947>.
- Sabbagh, G. J., & Fox, G. A. (1999). Statistical method for evaluation of a water table management model. *Trans. ASAE*, 42(3), 713-719. <http://dx.doi.org/10.13031/2013.13232>.
- Saltelli, A., & Marivoet, J. (1990). Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. *Reliability Eng. Syst. Safety*, 28(2), 229-253. [http://dx.doi.org/10.1016/0951-8320\(90\)90065-U](http://dx.doi.org/10.1016/0951-8320(90)90065-U).
- Saltelli, A., Tarantola, S., & Chan, K. P. S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *American Soc. Qual.*, 41(1), 39-56.
- Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity Analysis*. Chichester U.K.: John Wiley and Sons.
- Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Chichester, U.K.: John Wiley and Sons.
- Sauer, V. B., & Meyer, R. W. (1992). Determination of error in individual discharge measurements. USGS Open File Report 92-144. Washington, D.C.: U.S. Geological Survey.
- Sexton, A. M., Sadeghi, A. M., Zhang, X., Srinivasan, R., & Shirmohammadi, A. (2010). Using NEXRAD and rain gauge precipitation data for hydrologic calibration of SWAT in a northeastern watershed. *Trans. ASABE*, 53(5), 1501-1510. <http://dx.doi.org/10.13031/2013.34900>.
- Sexton, A. M., Shirmohammadi, A., Sadeghi, A. M., & Montas, H. J. (2011a). Impact of parameter uncertainty on critical SWAT output simulations. *Trans. ASABE*, 54(2), 461-471. <http://dx.doi.org/10.13031/2013.36449>.
- Sexton, A. M., Shirmohammadi, A., Sadeghi, A. M., & Montas, H. J. (2011b). A stochastic method to characterize model uncertainty for a nutrient TMDL. *Trans. ASABE*, 54(6), 2197-2207. <http://dx.doi.org/10.13031/2013.40659>.
- Shen, Z., Hong, Q., Yu, H., & Liu, R. (2008). Parameter uncertainty analysis of nonpoint-source pollution in the Danning River watershed of the Three Gorges Reservoir region, China. *Sci. Total Environ.*, 405(1-3), 195-205.
- Shirmohammadi, A., Chaubey, I., Harmel, R. D., Bosch, D. D., Muñoz-Carpena, R., Dharmasri, C., Sexton, A., Arabi, M., Wolf, M. L., Frankenberger, J., Graff, C., & Sohrabi, T. M. (2006). Uncertainty in TMDL models. *Trans. ASABE*, 49(4), 1033-1049. <http://dx.doi.org/10.13031/2013.21741>.
- Smith, R. M., & Wheeler, H. S. (2004). Multiple-objective evaluation of a simple phosphorus transfer model. *Hydrol. Proc.*, 18(9), 1703-1720. <http://dx.doi.org/10.1002/hyp.1413>.
- Sobol, I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Math. Model. Comput. Exp.*, 1, 407-414.
- Sohrabi, T. M., Shirmohammadi, A., & Montas, H. J. (2002). Uncertainty in nonpoint-source pollution models and associated risks. *Environ. Forensics*, 3(2), 179-189. <http://dx.doi.org/10.1006/enfo.2002.0089>.
- Sohrabi, T. M., Shirmohammadi, A., Chu, T. W., Montas, H., & Nejadhashemi, A. P. (2003). Uncertainty analysis of hydrologic and water quality predictions for a small watershed using SWAT2000. *Environ. Forensics*, 4(4), 229-238. <http://dx.doi.org/10.1080/714044368>.
- Stow, C. A., Reckhow, K. H., Qian, S. S., Lamon III, E. C., Arhonditsis, G. B., Borsuk, M. E., & Seo, D. (2007).



- Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. *JAWRA*, 43(6), 1499-1507.
- Sudheer, K. P., Chaubey, I., Garg, V., & Migliaccio, K. W. (2007). Impact of time scale of the calibration objective function on the performance of watershed models. *Hydrol. Proc.*, 21(25), 3409-3419. <http://dx.doi.org/10.1002/hyp.6555>.
- Takle, E. S., Jha, M. K., & Anderson, C. J. (2005). Hydrological cycle in the upper Mississippi River basin: 20th century simulations by multiple GCMs. *Geophys. Res. Letters*, 32(18), L18407. <http://dx.doi.org/10.1029/2005GL023630>.
- Tang, Y., Reed, P., van Wekhoven, K., & Wagener, T. (2007). Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. *Water Resources Res.*, 43(6), W06415. <http://dx.doi.org/10.1029/2006WR005813>.
- Thorsen, M., Refsgaard, J. C., Hansen, S., Pebesma, E., Jensen, J. B., & Kleeschulte, S. (2001). Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. *J. Hydrol.*, 242(3-4), 210-227. [http://dx.doi.org/10.1016/S0022-1694\(00\)00396-6](http://dx.doi.org/10.1016/S0022-1694(00)00396-6).
- Tolson, B. A., & Shoemaker, C. A. (2008). Efficient prediction uncertainty approximation in the calibration of environmental simulation models. *Water Resources Res.*, 44(4), W04411. <http://dx.doi.org/10.1029/2007WR005869>.
- van Griensven, A., & Meixner, T. (2006). Methods to quantify and identify the sources of uncertainty for river basin water quality models. *Water Sci. Tech.*, 53(1), 51-59. <http://dx.doi.org/10.2166/wst.2006.007>.
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., & Srinivasan, R. (2006). A global sensitivity tool for the parameters of multi-variable catchment models. *J. Hydrol.*, 324(1-4), 10-23. <http://dx.doi.org/10.1016/j.jhydrol.2005.09.008>.
- van Oijen, M., Reyer, C., Bohn, F. J., Cameron, D. R., Deckmyn, G., Flechsig, M., Härkönen, S., Hartig, F., Huth, A., Kivistie, A., Lasch, P., Mäkelä, A., Mette, T., Minunno, F., & Rammer, W. (2013). Bayesian calibration, comparison, and averaging of six forest models using data from Scots pine stands across Europe. *Forest Ecol. Mgmt.*, 289, 255-268. <http://dx.doi.org/10.1016/j.foreco.2012.09.043>.
- van Wekhoven, K., Wagener, T., Reed, P., & Tang, Y. (2008). Rainfall characteristics define the value of streamflow observations for distributed watershed model identification. *Geophys. Res. Letters*, 35(11), L11403. <http://dx.doi.org/10.1029/2008GL034162>.
- Vázquez, R. F., Feyen, L., Feyen, J., & Refsgaard, J. C. (2002). Effect of grid size on effective parameters and model performance of the MIKE SHE code. *Hydrol. Proc.*, 16(2), 355-372. <http://dx.doi.org/10.1002/hyp.334>.
- Vázquez, R. F., Beven, K., & Feyen, J. (2009). GLUE-based assessment on the overall prediction of a MIKE SHE application. *Water Resources Mgmt.*, 23(7), 1325-1349. <http://dx.doi.org/10.1007/s11269-008-9329-6>.
- Vezzaro, L., & Mikkelsen, P. S. (2012). Application of global sensitivity analysis and uncertainty quantification in dynamic modeling of micropollutants in stormwater runoff. *Environ. Model. Software*, 27-28, 40-51. <http://dx.doi.org/10.1016/j.envsoft.2011.09.012>.
- Vicens, G. J., Rodriguez-Iturbe, I., & Schaake, J. C. (1975). A Bayesian framework for the use of regional information in hydrology. *Water Resources Res.*, 11(3), 405-414. <http://dx.doi.org/10.1029/WR011i003p00405>.
- Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., & Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Res.*, 44(12), W00B09. <http://dx.doi.org/10.1029/2007WR006720>.
- Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., & Robinson, B. A. (2009). Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environ. Res. Risk Assess.*, 23(7), 1011-1026. <http://dx.doi.org/10.1007/s00477-008-0274-y>.
- Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environ. Res. Risk Assess.*, 19(6), 378-387. <http://dx.doi.org/10.1007/s00477-005-0006-5>.
- Wagener, T., Lees, M. J., & Wheater, H. S. (2001). Monte Carlo Analysis Toolbox (MCAT), version 3. London, U.K.: Imperial College of Science, Technology, and Medicine.
- Wagener, T., van Wekhoven, K., Reed, P., & Tang, Y. (2009). Multi-objective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. *Water Resources Res.*, 45(2), W02501. <http://dx.doi.org/10.1029/2008WR007347>.
- Walter, M. T., & Shaw, S. B. (2006). Discussion: Curve number hydrology in water quality modeling: Uses, abuses, and future directions. *JAWRA*, 41(6), 1491-1492.
- Wang, X., & Yen, H. (2014). APEX-CUTE: An auto-calibration and uncertainty analysis tool for the APEX model. In D. P. Ames, N. W. Quinn, & A. E. Rizzoli (Ed.), *Proc. 7th Intl. Congress Environ. Model. Software*. Manno, Switzerland: International Environmental Modeling and Software Society.
- Wang, X., He, X., Williams, J. R., Izaurralde, R. C., & Atwood, J. D. (2005). Sensitivity and uncertainty analyses of crop yields and soil organic carbon simulated with EPIC. *Trans. ASAE*, 48(3), 1041-1054. <http://dx.doi.org/10.13031/2013.18515>.
- Wang, X., Potter, S. R., Williams, J. R., Atwood, J. D., & Pitts, T. (2006a). Sensitivity analysis of APEX for national assessment. *Trans. ASABE*, 49(3), 679-688. <http://dx.doi.org/10.13031/2013.20487>.
- Wang, X., Frankenberger, J. R., & Kladivko, E. J. (2006b). Uncertainties in DRAINMOD predictions of subsurface drain flow for an Indiana silt loam using the GLUE methodology. *Hydrol. Proc.*, 20(14), 3069-3084. <http://dx.doi.org/10.1002/hyp.6080>.
- Wang, X., Yen, H., Jeon, J., & Williams, J. R. (2014a). Accounting for conceptual soil erosion and sediment yield modeling uncertainty in the APEX model using Bayesian model averaging. *J. Hydrol. Eng.*, 20(3). [http://dx.doi.org/10.1061/\(ASCE\)HE.1943-5584.0001119](http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001119).
- Wang, X., Yen, H., Liu, Q., & Liu, J. 2. (2014b). An auto-calibration tool for the Agricultural Policy Environmental eXtender (APEX) model. *Trans. ASABE*, 57(4), 1087-1098. <http://dx.doi.org/10.13031/trans.57.10601>.
- Willmott, C. J. (1981). On the validation of models. *Physical Geographer*, 2(2), 184-194.
- Wu, S., Li, J., & Huang, G. H. (2005). An evaluation of grid size uncertainty in empirical soil loss modeling with digital elevation models. *Environ. Model. Software*, 10(1), 33-42.
- Wu, S., Li, J., & Huang, G. H. (2007). Characterization and evaluation of elevation data uncertainty in water resource modeling with GIS. *Water Resources Res.*, 22(8), 959-972.
- Wu, S. H., Jansson, P.-E., & Kolari, P. (2011). Modeling seasonal course of carbon fluxes and evapotranspiration in response to low temperature and moisture in a boreal Scots pine ecosystem. *Ecol. Model.*, 222(17), 3103-3119. <http://dx.doi.org/10.1016/j.ecolmodel.2011.05.023>.
- Wu, Y., & Liu, S. (2012). Automatic calibration, sensitivity, and uncertainty analyses of a complex model using the R package Flexible Modeling Environment (FME): SWAT as an example. *Environ. Model. Software*, 31, 99-101. <http://dx.doi.org/10.1016/j.envsoft.2011.11.013>.



- Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., & Yang, H. (2008). Comparing uncertainty analysis techniques for a SWAT application to the Chaohe basin in China. *J. Hydrol.*, 358(1-2), 1-23. <http://dx.doi.org/10.1016/j.jhydrol.2008.05.012>.
- Yantheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., & Stewart, A. (2008). Understanding uncertainty in distributed flash flood forecasting for semiarid regions. *Water Resources Res.*, 40(5), W05S19. <http://dx.doi.org/10.1029/2007WR005940>.
- Yen, H., Wang, X., Darrell, G., Fontane, R., Harmel, D., & Arabi, M. (2014). A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. *J. Environ. Model. Software*, 54, 211-221. <http://dx.doi.org/10.1016/j.envsoft.2014.01.004>.
- Yevjevich, V., & Jeng, R. I. (1969). Properties of non-homogeneous hydrologic time series. Hydrology Paper 32. Fort Collins, Colo.: Colorado State University.
- Yin, L., Wang, X., Pan, J., & Gassman, P. W. (2009). Evaluation of APEX for daily runoff and sediment yield from three plots in the upland Huaihe River watershed, China. *Trans. ASABE*, 52(6), 1833-1845. <http://dx.doi.org/10.13031/2013.29212>.
- Zajac, Z. B. (2010). Global sensitivity and uncertainty analysis of spatially distributed watershed models. PhD diss. Gainesville, Fla.: University of Florida, Agricultural and Biological Engineering.
- Zhang, K., Srinivasan, R., & Bosch, D. (2009a). Calibration and uncertainty analysis of the SWAT model using genetic algorithms and Bayesian model averaging. *J. Hydrol.*, 374(3-4), 307-317. <http://dx.doi.org/10.1016/j.jhydrol.2009.06.023>.
- Zhang, X., Liang, F., Srinivasan, R., & van Liew, M. (2009b). Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Resources Res.*, 45(2), W02403. <http://dx.doi.org/10.1029/2008WR007030>.