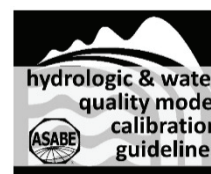


HYDROLOGIC AND WATER QUALITY MODELS: KEY CALIBRATION AND VALIDATION TOPICS



D. N. Moriasi, R. W. Zeckoski, J. G. Arnold, C. B. Baffaut,
R. W. Malone, P. Daggupati, J. A. Guzman, D. Saraswat, Y. Yuan,
B. W. Wilson, A. Shirmohammadi, K. R. Douglas-Mankin

ABSTRACT. *As a continuation of efforts to provide a common background and platform for development of calibration and validation (C/V) guidelines for hydrologic and water quality (H/WQ) modeling, ASABE members worked to determine critical topics related to model C/V, perform a synthesis of a previously published special collection of articles and other relevant literature, and provide topic-specific recommendations based on the synthesis as well as personal modeling expertise. This article introduces a special collection of nine research articles covering key topics related to calibration and validation of H/WQ models. The topics include: terminology, hydrologic processes and model representation, spatial and temporal scales, model parameterization, C/V strategies, sensitivity, uncertainty, performance measures and criteria, and documentation and reporting. The main objective of this introductory article is to introduce and summarize key aspects of these topics, including recommendations. Individually, the articles provide model practitioners with detailed topic-specific recommendations related to model calibration, validation, and use. Collectively, the articles present recommendations to enhance H/WQ modeling.*

Keywords. *ASABE, Calibration, Guidelines, Hydrologic models, Recommendations, Synthesis, Validation.*

Submitted for review in November 2014 as manuscript number NRES 11075; approved for publication by the Natural Resources & Environmental Systems Community of ASABE in December 2015.

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 **Daniel N. Moriasi, ASABE Member**, Research Hydrologist, USDA-ARS Grazinglands Research Laboratory, El Reno, Oklahoma; **Rebecca W. Zeckoski, ASABE Member**, Owner, Zeckoski Engineering, Charlotte, North Carolina; **Jeffrey G. Arnold, ASABE Fellow**, Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; **Claire Baffaut, ASABE Member**, Research Hydrologist, USDA-ARS Cropping Systems and Water Quality Research Unit, Columbia, Missouri; **Robert W. Malone, ASABE Member**, Agricultural Engineer, USDA-ARS National Laboratory for Agriculture and the Environment, Ames, Iowa; **Prasad Daggupati, ASABE Member**, Assistant Research Scientist, Texas Agrilife Research, Texas A&M University, College Station, Texas; **Jorge A. Guzman, ASABE Member**, Senior Engineer, Waterborne Environmental, Inc., Champaign, Illinois; **Dharmendra Saraswat, ASABE Member**, Associate Professor, Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana; **Yongping Yuan, ASABE Member**, Research Hydrologist, U.S. EPA Office of Research and Development, NERL-ESD-Landscape Ecology Branch, Las Vegas, Nevada; **Bruce N. Wilson, ASABE Fellow**, Professor, Department of Biosystems and Agricultural Engineering, University of Minnesota, St. Paul, Minnesota; **Adel Shirmohammadi, ASABE Fellow**, Professor, Department of Environmental Science and Technology, Associate Dean for Research, and Associate Director of Maryland Agricultural Experiment Station, University of Maryland, College Park, Maryland; **Kyle R. Douglas-Mankin, ASABE Member**, Senior Hydrologist, Everglades Program Team, U.S. Fish and Wildlife Service, A.R.M. Loxahatchee National Wildlife Refuge, Boynton Beach, Florida; **Corresponding author:** Daniel N. Moriasi, USDA-ARS Grazinglands Research Laboratory, 7207 W. Cheyenne St., El Reno, OK 73036-0000; phone: 405-262-5291, ext. 260; e-mail: daniel.moriasi@ars.usda.gov.

Hydrologic and water quality (H/WQ) models consist of interrelated assemblages of mathematical equations that represent complex physical, chemical, and biological processes governing the hydrology and fate and transport of sediments, nutrients, pesticides, and bacteria. H/WQ models are widely used to evaluate the impacts of changes in land use, land management, climate, and conservation practices on soil and water resources. Reliable models provide valuable information for making sound management, policy, and regulatory decisions. Comprehensive calibration and validation (C/V) of models is essential to obtain the “right outcomes” for the “right reasons” (Holling, 1978; Kirchner, 2006). Important components of proper C/V include selecting an appropriate model, setting up the model properly, establishing valid C/V strategies (Refsgaard and Henriksen, 2004), and capturing available datasets to represent the appropriate spatial and temporal scales (Kirchner, 2006).

H/WQ models are valuable because they represent processes operating at various spatial and temporal scales simultaneously in a complex and interrelated manner. For example, the hydrology component of the Root Zone Water Quality Model (RZWQM; Ahuja et al., 2000) consists of interrelated hydrologic processes such as precipitation (including rainfall, snow, or irrigation), evaporation and transpiration (ET), surface runoff, infiltration, percolation, lateral flow, streamflow, and groundwater. However, no matter how detailed a model, it remains a simplification of the natural system in which some processes are characterized by rates and thresholds whose values are unknown. Be-

cause many model parameters are typically available for calibration, the mere fitting of modeling results to observed data may result in good statistics of model performance but poor correspondence to the actual processes that the model is intended to represent. In such a case, a good fit is obtained for the wrong reasons. Application of such results can lead to incorrect conclusions with negative consequences (Dressel, 2010). The use of proper C/V practices can help prevent this circumstance while also allowing for a common foundation for interpretation of results and their comparison with other modeling studies (Refsgaard and Henriksen, 2004; Douglas-Mankin et al., 2010; Tuppard et al., 2011).

Research on and application of H/WQ models abounds in the literature. While calibration and validation are essential to ensuring model accuracy and precision, there is little consistency in how model practitioners conduct, document, and report this process. Generally, published articles emphasize modeling study results; however, without a clear, careful, and consistent C/V process, these results stand on uncertain footing. Many individual articles in the literature have addressed individual components of this topic (James and Burges, 1982; Beven and Binley, 1992; Refsgaard, 1997; Engel et al., 2007; Moriasi et al., 2007; Harmel et al., 2014).

Potential benefits of a comprehensive C/V approach include:

- Better constraints on calibration parameter estimates (Khu et al., 2008).
- Reduction in uncertainty or documentation of uncertainty in model outputs.
- Minimization of equifinality (Henriksen et al., 2003; Khu et al., 2008) allowing model users to discriminate between different model solutions with identical or similar model metrics.
- Greater assurance that model structure and behavior are consistent with hydrological understanding of reality (Wagener and Gupta, 2005).
- Assessment of model processes that need better representation for certain conditions (region, climate, topography, soils, geologic formations, etc.).
- More unified assessment of modeling data compared to calibration and validation needs, thus targeting areas where resources need to be invested to improve modeling studies.
- Better coordination among model developers in building upon the body of modeling studies to address model needs.
- Better communication of modeling practices both to model practitioners and users of model results.

Moriasi et al. (2012) summarized model-specific C/V approaches for commonly used H/WQ models as part of the efforts initiated by ASABE in 2010 with the goal of developing calibration, validation, and documentation recommendations. The proposed follow-up phases were to identify critical C/V topics, synthesize the articles in the Moriasi et al. (2012) special collection along with other relevant information from existing literature, and provide topic-specific recommendations. The authors of the present

collection identified key C/V topics from the prior collection and addressed each in a separate article.

Recommendations from this special collection of articles and the communication subcommittee article (Harmel et al., 2014) will contribute to the discussion surrounding potential development of topic-specific ASABE engineering practices or standards (hereafter “standards”) for model C/V. In addition, the articles in this special collection provide model practitioners with invaluable topic-specific information needed to verify model accuracy and precision for a given modeling purpose. The objective of this article is to introduce and summarize key aspects of these topics, including main recommendations.

SUMMARY OF KEY CALIBRATION AND VALIDATION TOPICS

There are nine research articles in this special collection covering key topics related to the C/V of H/WQ models. The topics include: terminology (Zeckoski et al., 2015), hydrologic processes and model representation (Arnold et al., 2015), spatial and temporal scales (Baffaut et al., 2015), model parameterization (Malone et al., 2015), C/V strategies (Daggupati et al., 2015), sensitivity (Yuan et al., 2015), uncertainty (Guzman et al., 2015), model performance measures and criteria (Moriasi et al., 2015), and documentation and reporting (Saraswat et al., 2015). Each topic is introduced with its description and relevance to model C/V before presenting key recommendations. The main components of these articles are: (1) a clear statement of why the topic is important with respect to model C/V, (2) a synthesis of the Moriasi et al. (2012) special collection and other relevant literature with respect to the topic of focus, and (3) recommendations based on previous work summarized by Moriasi et al. (2012) as well as personal experience. For some topics, a case study is provided to demonstrate the application of the topic-specific C/V recommendations.

TERMINOLOGY

Zeckoski et al. (2015) describe the importance of consistent terminology and provide important recommendations. Confusion regarding the terminology used in H/WQ modeling is one of the major obstacles to establishing generally acceptable model evaluation guidelines. Inconsistent terminology can result in misinterpretation of the literature or misapplication of concepts and results from others’ research. Speaking a common language in H/WQ modeling community is an essential first step in communicating model C/V processes and the associated concerns.

Key Recommendations

Overall, Zeckoski et al. (2015) recommend that authors of H/WQ modeling literature adopt consistent usage of terminology. The article provides recommended terminology related to model description, model processes and techniques, sensitivity and uncertainty, calibration and validation, hydrologic modeling, plant growth modeling, general water quality modeling, sediment modeling, and nutrient

modeling. Many of the terms discussed are controversial, misunderstood, or ambiguous. For example, the terms “verification” and “validation” are used interchangeably by many researchers in the literature, and the terminology article sets forth a recommendation for consistent usage. The article also includes basic terms such as “hydrology” and “evapotranspiration” that are essential for proper understanding of the modeling literature and provides a common basis from which further discussions of H/WQ modeling can progress.

HYDROLOGICAL PROCESSES AND MODEL REPRESENTATION

Arnold et al. (2015) discuss the importance of accurate model process representation and its impact on calibration and scenario analysis. Models are divided into three categories: (1) flow, heat, and solute transport, (2) field scale, and (3) watershed scale. Processes simulated by models in each category are reviewed and discussed. Model case studies are used to illustrate situations in which a model can show excellent statistical agreement with measured data, while misrepresented processes (water balance, nutrient balance, sediment source/sinks) within a field or watershed can cause misinterpretation and/or poor decisions related to management scenarios. These errors may be amplified at the watershed scale, where additional sources and transport processes are simulated.

Key Recommendations

To account for calibration processes, the authors recommend a diagnostic approach using both hard and soft data, as suggested by Siebert and McDonnell (2002) and Yilmaz et al. (2008). The diagnostic approach looks at signature patterns of behavior to determine which processes, and thus parameters, need further adjustment during calibration. Hard data are defined as long-term, measured time series, typically at a point within a watershed. Soft data are defined as information on individual processes within a water, sediment, or nutrient budget that may not be directly measured within the study area (e.g., average annual estimate) and may entail considerable uncertainty. Use of both overcomes the weaknesses of traditional regression-based calibration by discriminating between multiple processes within a budget. Advantages of developing soft data for the calibration are that they (1) require a basic understanding of processes (water, sediment, nutrient, and carbon budgets) within the spatial area being modeled, and (2) constrain the hard data calibration within realistic bounds.

The approach recommended by Arnold et al. (2015) consists of four basic steps: (1) *Collect and assemble all hard data for the study area.* This may include time series data from stream gauges, groundwater wells, soil moisture monitors, and reservoir levels. (2) *Collect and assemble all soft data for the study area.* This step estimates water, nutrient, carbon, and sediment budgets and also focuses efforts on understanding basic processes. (3) *Perform manual or automated calibration.* Regress the measured time series (hard data) and simulated time series and use the soft data as constraints in the calibration. (4) *Repeat calibration.* Identify diagnostic signature indices and refine soft data.

A final recommendation is to build soft data processes

into automated calibration procedures. As an example first step, White et al. (2012) developed a screening tool called SWAT Check that assists SWAT modelers in ensuring that processes and water, sediment, and nutrient budgets are realistic. SWAT Check is a standalone program that reads SWAT output and alerts users of values outside typical ranges; creates process-based figures for visualizing water, sediment, and nutrient budgets; and detects and alerts users of common model application errors. The ultimate goal is to include these soft (process) data in automated calibration routines that are routinely used for calibration (Yen et al., 2014). This will increase the likelihood that simulation scenarios provide realistic and meaningful results.

SPATIAL AND TEMPORAL SCALES

Baffaut et al. (2015) describe how the components of the spatial and temporal scales of H/WQ models (i.e., the extent of the study area, the duration of the simulation period, and the spatial and temporal resolution of the input and calibration data and of the calculations) are affected by the critical processes being simulated. They also describe how these components along with the modeling objectives affect model parameterization, model performance, and the interpretation and use of model results. Consequently, selecting these components is an important step in the process of applying a model to a specific site and study, and modelers should consider the nature of the hydrological and biogeochemical processes simulated, the resolution of available input and calibration data, and the modeling objectives. The article discusses how processes, data, and modeling objectives affect the selection of a scale and presents four principles, two suggestions, and numerous examples.

Key Recommendations

Scale components are related to the simulated processes; as such, Baffaut et al. (2015) is a companion to the synthesis article on hydrological processes (Arnold et al., 2015). Baffaut et al. (2015) present tables that show what processes are relevant for each spatial and temporal extent and resolution. The authors recommend that prior to applying a model for a specific study at a given site, the modeler should first consider the relevant processes and select a model that simulates these processes at the appropriate spatial and temporal resolution. Furthermore, the spatial and temporal resolution should take into account the degree of variability in data and the modeling results necessary to meet the study objectives.

Spatial and temporal resolution of calibration and input data should be compatible with model spatial and temporal discretization levels, which are informed by the study objectives and the processes simulated. Finally, model validation for one scale does not imply validation for smaller or larger scales. Thus, model C/V at multiple scales should be performed in successive steps considering the dominant processes at the appropriate scales. In the absence of validation at the relevant scale, results should not be extrapolated from a model calibrated using data that are representative of a larger or smaller scale.

When modeling objectives or the processes simulated must consider multiple spatial resolutions, a complex mod-

el may not be appropriate. The modeler may want to consider breaking the problem into smaller questions and using models specifically intended for each scale. When modeling objectives address the interaction of spatial subunits or processes that operate over different spatial and temporal scales, one alternative is to use two models and, if feasible, develop a linkage to explicitly take into account the interactions that exist between the two scales. In this modeling approach, the outputs of small-scale models could override the equivalent output from a larger-scale model. While non-adherence of the four principles is likely to result in erroneous interpretation of the model results, these two suggestions aim at keeping models as simple as possible and limiting complexity.

PARAMETERIZATION

A detailed discussion on H/WQ model parameterization is provided by Malone et al. (2015). Parameters are typically measurable or quantifiable coefficients that define the characteristics of the catchment or flow domain and generally remain invariable through all or part of the simulation run (Beven and Freer, 2001; James, 2005; Barber, 2005). Parameterization includes imparting knowledge of the simulated field or watershed processes to the model and determining a set of acceptable parameter values for a model application (Arnold et al., 2012; Zeckoski et al., 2015). While model user manuals often provide value ranges for many parameters, this guidance is often inadequate for assigning values in many specific applications. Many applications of H/WQ models can contain a large number of adjustable parameters (e.g., MIKE SHE; Jaber and Shukla, 2012). A further complication is that many of the measured or estimated parameter values for a given field condition can include considerable uncertainty (e.g., Baroni et al., 2010). Therefore, determining a suitable set of parameter values (i.e., parameterization) for H/WQ models is a critical but difficult task. A persistent issue associated with calibration is that many parameter sets may produce acceptable simulations (non-uniqueness). The authors point out that although developing parameterization guidelines will not eliminate the problem of non-uniqueness, considering them will help reduce the acceptable parameter space.

Key Recommendations

The authors provide seven parameterization recommendations (table 1). These recommendations build upon and confirm previous general parameterization recommendations by Refsgaard and Storm (1996) and Engel et al. (2007).

CALIBRATION AND VALIDATION STRATEGIES

Daggupati et al. (2015) discuss various C/V strategies for H/WQ modeling. The authors recommend that a comprehensive model C/V strategy should consider and document: (1) goals of model use, (2) data used for model calibration and validation, (3) model input parameters and output variables be calibrated and validated, (4) strategies used to calibrate and validate the model, and (5) measures and criteria used to characterize model performance. The authors provide a generalized structure, process, and examples to assist the model practitioner in developing a C/V strategy for H/WQ modeling applications.

Key Recommendations

The authors provide C/V strategy recommendations in four important areas: output locations and processes, staging design, warm-up period, and spatio-temporal data allocations (data splitting). Daggupati et al. (2015) recommend single-site calibration for areas with uniform characteristics (e.g., soil, slope, vegetation, and meteorology) across the entire modeled area. A multi-site calibration method is recommended for large areas (or watersheds) with more varied, complex physical characteristics and/or when observed data for a given process are available at multiple locations within the study area.

The staging design refers to the systematic approach used by the model practitioner to organize the adjustment of input parameters and assessment of output variables. The many possible permutations may be distilled to just a few, distinct calibration philosophies, as follows: single-stage; stepwise single-pass; stepwise, iterative, limited parameter space; and stepwise, iterative, extensive parameter space. These calibration philosophies are shown in figure 1 and discussed in detail by the authors.

Daggupati et al. (2015) recommend using warm-up (initialization) periods of two to three years for hydrology and five to ten years for sediment and nutrients based on rec-

Table 1. Parameterization guidelines and associated references.

Parameterization guideline or typical practice	Reference
1. Use site-specific measurements where possible or estimate parameters based on knowledge of the site.	MACRO (Jarvis and Larsbo, 2012); WARMF (Herr and Chen, 2012); HYDRUS (Simunek et al., 2012); SHAW (Flerchinger et al., 2012); MIKE SHE (Jaber and Shukla, 2012); DRAINMOD (Skaggs et al., 2012).
2. Optimize and focus on uncertain or sensitive parameters.	MACRO (Jarvis and Larsbo et al., 2012); SWAT (Arnold et al., 2012); SHAW (Flerchinger et al., 2012); EPIC/APEX (Wang et al., 2012); ADAPT (Gowda et al., 2012); RZWQM (Ma et al., 2012); MIKE SHE (Jaber and Shukla, 2012).
3. Minimize the number of optimized parameters.	MIKE SHE (Jaber and Shukla, 2012).
4. Constrain optimized or estimated parameter values within accepted ranges and justify values, especially those that fall outside expected intervals.	ADAPT (Gowda et al., 2012); SWAT (Arnold et al., 2012); EPIC/APEX (Wang et al., 2012); COUPMODEL (Jansson, 2012); BASINS/HSPF (Duda et al., 2012).
5. Use multiple criteria to optimize parameter values (more than one model output or target is compared with observed data).	MT3DMS; SHAW (Flerchinger et al., 2012).
6. Use "hard" and "soft" data to optimize parameters ("soft" data are qualitative knowledge from experimentalists such as estimated ET).	BASINS/HSPF (Duda et al., 2012).
7. Use warm-up periods to reduce model dependence on estimates of initial condition state variables.	MIKE SHE (Jaber and Shukla, 2012); DAISY (Hansen et al., 2012).

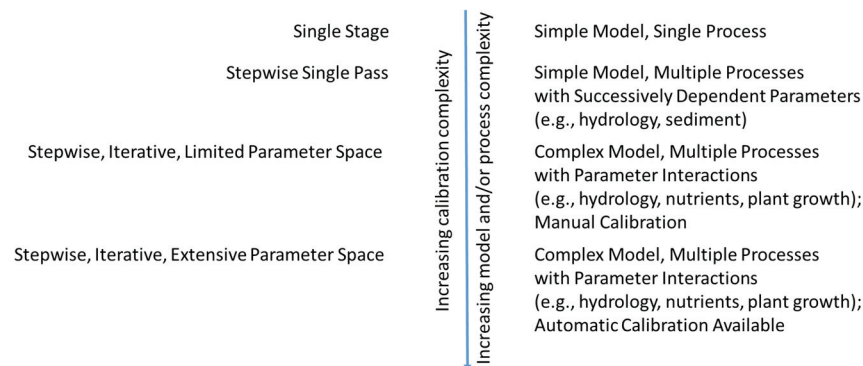


Figure 1. Staging design approaches.

ommendations from model developers and seasoned model users. The authors also discuss the importance of allocating available data for C/V. The most commonly used method to allocate data for C/V is the temporal split-sample method, in which measured data are split into two periods; however, other approaches (i.e., proxy basin, differential split-sample, and proxy-basin differential split-sample) are also discussed.

SENSITIVITY

A detailed discussion on sensitivity analysis (SA) methods and measures as well as independent tools for sensitivity analysis for H/WQ models is presented by Yuan et al. (2015). Although uncertainty (Guzman et al., 2015) and sensitivity analyses are often carried out in tandem, they essentially serve different purposes, as the former focuses on propagation of uncertainty from model inputs to outputs (Saltelli et al., 2000) while the latter quantifies and/or explores the strength of relationships between model inputs and outputs (Lane and Ferreira, 1980; Saltelli et al., 2004). SA is frequently conducted to guide data collection for model calibration and validation. In addition, SA can reveal interactions among combinations of parameters that may contribute to equifinality; therefore, SA enhances data collection and model C/V. Yuan et al. (2015) present results of a thorough review and synthesis of SA studies for the H/WQ models in Moriasi et al. (2012) and other commonly used H/WQ models. For each model reviewed, information is summarized on SA methods used, input parameters analyzed, outputs evaluated, ranking of influential parameters,

number of simulations needed to perform the SA, and sensitivity measures and indices used to evaluate the sensitivity of input parameters. The summary of sensitive parameters identified from past SA studies for each H/WQ model should prove useful for future model applications.

Key Recommendations

The authors recommend that model practitioners perform SA to identify key model parameters for use in model calibration, validation, and verification, particularly since results of SA are site and condition specific. Sensitivity analysis methods were broadly categorized by parameter sampling method, purpose, and assessment measure used (fig. 2). Parameter sensitivity can be assessed in various ways, ranging from simple visual inspection of input vs. output plots, to robust and sophisticated variance-based sensitivity indices.

The authors recommend that model users select appropriate SA methods depending on the intended model application and the users' assumptions about the certainty and linearity of the parameters. Whereas local SA (varying individual parameters in a small vicinity of a base point) may be suitable for simple and linear systems, global SA (varying all parameters within their entire uncertainty ranges simultaneously) may be needed to account for model non-linearity, non-monotonicity, and parameter interactions for complex systems.

UNCERTAINTY

Guzman et al. (2015) discuss different sources of uncer-

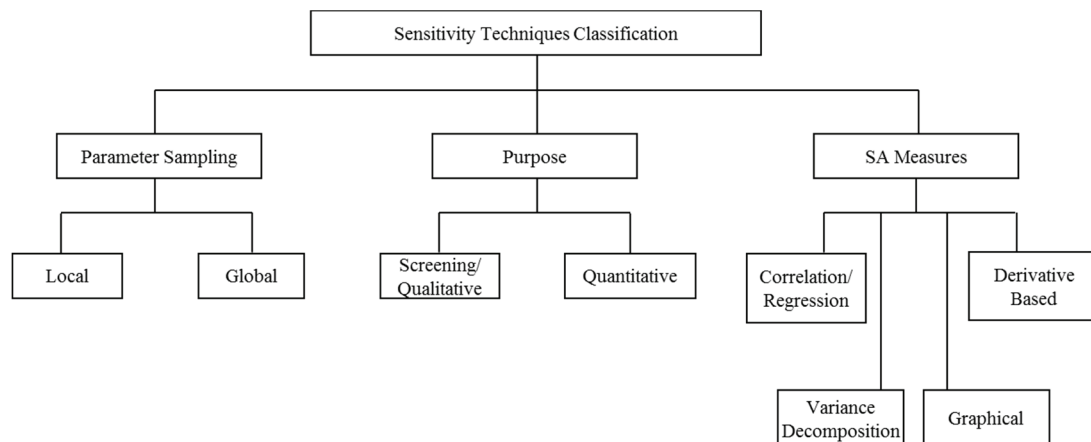


Figure 2. Sensitivity analysis techniques.

tainty and how uncertainty impacts the ability of H/WQ models to accurately evaluate the response of complex systems and may lead to misguided assessments and risk management decisions if uncertainty is ignored. Sources of uncertainty include model inputs and derived/computed parameters (e.g., observations and physical properties), a model's mathematical inability to properly represent fundamental processes and mechanisms, a modeler's lack of capacity to properly simplify and represent the system under investigation, and the measured data used for calibration and validation (Vicens et al., 1975; Loague and Corwin, 1996; Loucks et al., 2005; Harmel and Smith, 2007). Uncertainty in H/WQ models can propagate non-linearly throughout model runs, causing model output to substantially deviate from the expected response or masking specific cause-and-effect relationships in the natural system being investigated. 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, or system dynamics. Mathematical operations and transformations are commonly conducted on observed data to fulfill specific H/WQ model requirements (e.g., spatio-temporal discretization and representation, computational limits, conceptual model simplification, etc.). This can exacerbate observational uncertainty and change the spatio-temporal patterns of input data; thus, uncertainty in model outputs is always expected.

Key Recommendations

The authors recommend building a conceptual model linked to project objectives before model development or application. Errors in the conceptual model due to incomplete system understanding or representation are bound to propagate throughout the modeling process regardless of the model's mathematical sophistication, accuracy of input parameter values and observed data, etc. Uncertainty analysis in H/WQ models aims to estimate uncertainty in model outputs linked to model inputs or structural uncertainty.

Therefore, the authors recommend that uncertainty analysis be carried out for H/WQ modeling studies. There are different uncertainty analysis methods for different purposes and computational requirements, which involve multiple model simulations and evaluations. Irrespective of the method used, the authors recommend the following general step-by-step procedure for performing global parameter uncertainty/sensitivity analyses: (1) determine probability distribution functions (PDFs) of input parameters, (2) generate input samples based on PDFs, (3) apply a screening method to short list the important parameters, (4) refine parameters and their ranges for rigorous analysis, (5) perform model simulations to calculate desired outputs and decision variables, and (6) perform statistical analysis to obtain sensitivity indices, parameter rankings, predictive PDFs, and confidence intervals.

In addition, 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 (Harmel et al., 2006, 2010). Estimating and reporting the uncertainty in the measured data (observational uncertainty) used to calibrate and validate models is rec-

ommended because of its impact on the evaluation and interpretation of model results (Harmel et al., 2014).

PERFORMANCE MEASURES AND CRITERIA

A detailed discussion on model performance measures and criteria is provided by Moriasi et al. (2015). Performance measures and evaluation criteria are important aspects of H/WQ modeling and should be set before C/V (ASCE, 1993; USEPA, 2002). Results of a synthesis of the performance measures and criteria of the articles in Moriasi et al. (2012) are provided, including a detailed review of their strengths and weaknesses to better determine recommended measures. Reported performance data for each simulated component during the C/V periods are recorded. A statistical meta-analysis is performed on existing model performance data to help guide the development of performance criteria covering various constituents at field and watershed spatial scales and daily to annual temporal scales. General guidelines for model performance evaluation are established based on the results of the synthesis of the performance measures and criteria and the statistical meta-analysis of the reported performance data. The guidelines are in the form of recommended measures and criteria. A case study is provided to illustrate the application of the recommended measures and the corresponding developed criteria.

Key Recommendations

The authors emphasize that before using recommended measures and criteria to assess model performance, modelers should consider recommendations for all the C/V topics covered in this special collection. Use of multiple measures, including graphical and statistical measures, is recommended. Graphical methods may include time series and scatter plots for shorter periods and coarse temporal resolution (e.g., monthly calibration for one to three years), cumulative distributions or duration curves for longer periods and finer resolutions, and maps for field- and watershed-scale models, whenever possible. Recommended statistical measures include coefficient of determination (R^2) together with the gradient and intercept of the corresponding regression line, Nash Sutcliffe efficiency (NSE), index of agreement (d), root mean square error (RMSE), ratio of RMSE to observations standard deviation (RSR), and percent bias (PBIAS). During low-flow simulations, logarithmic or relative derivatives of NSE or d need to be used (Krause et al., 2005).

The authors recommend that model performance can be judged "satisfactory" for flow simulations if monthly $R^2 > 0.70$ and $d > 0.75$ for field-scale models and daily, monthly, or annual $R^2 > 0.60$, $NSE > 0.50$, and $PBIAS \leq \pm 15\%$ for watershed-scale models. Model performance at the watershed scale can be evaluated as "satisfactory" if monthly $R^2 > 0.40$ and $NSE > 0.45$ and daily, monthly, or annual $PBIAS \leq \pm 20\%$ for sediment; monthly $R^2 > 0.40$ and $NSE > 0.35$ and daily, monthly, or annual $PBIAS \leq \pm 30\%$ for phosphorus; and monthly $R^2 > 0.30$ and $NSE > 0.35$ and daily, monthly, or annual $PBIAS \leq \pm 30\%$ for nitrogen. For RSR, the authors recommend that the criteria proposed by Moriasi et al. (2007) be used until new criteria are devel-

oped. Although the intent of this study was to develop generalizable performance evaluation criteria for all models, sufficient data for meta-analysis were available only for SWAT, HSPF, WARMF, and ADAPT. Therefore, the authors also recommend that the performance evaluation criteria developed in this study be used primarily for these models and used with caution for other models. These ratings, which apply to both calibration and validation, may be adjusted to be more or less strict based on considerations of the quality and quantity of available measured data, spatial and temporal scales, and project scope, magnitude, and intended purpose. A framework for determining recommended performance measures and their corresponding criteria as more information becomes available is provided. Finally, $R^2 < 0.18$, $NSE < 0.0$, $PBIAS \geq \pm 30\%$ for flow, $PBIAS \geq \pm 55\%$ for sediments, and $PBIAS \geq \pm 70\%$ for nutrients, and $d < 0.60$ are considered to indicate unacceptable model performance.

DOCUMENTATION AND REPORTING

Saraswat et al. (2015) present a detailed discussion of documentation and reporting of the H/WQ model C/V process. Proper documentation is a critical element and increases a modeling effort's scientific credibility. The authors discuss eight recommended elements of model documentation and provide examples for each element.

Key Recommendations

The authors make eight recommendations to facilitate communication of model C/V practices and results (table 2). The authors also recommend that the modeling community move toward fully reproducible model calibration, validation, and use (fig. 3). The current norm of limited data and model availability through publication should evolve toward a full sharing of data files, model code, and supplemental materials to allow full reproducibility of model results. These efforts will facilitate greater collaboration among practitioners, broaden opportunities for cross-regional syntheses of modeling studies, minimize duplicative efforts, and provide the transparency needed to advance the science of H/WQ modeling. The authors invite participation of peers to discuss their concerns, pose questions, and provide suggestions to facilitate this paradigm shift.

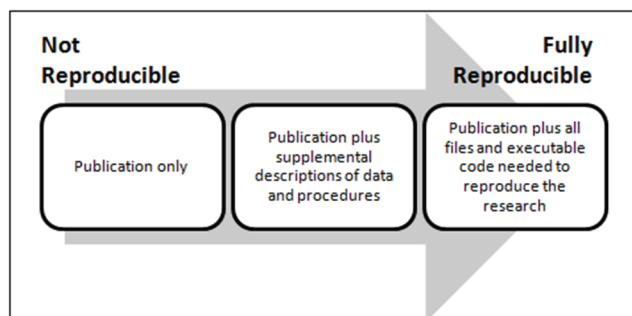


Figure 3. Moving toward reproducibility in hydrologic and water quality modeling (after Peng, 2011).

SUMMARY AND CONCLUSIONS

This special collection provides a description of critical calibration and validation topics for H/WQ models, and this introductory article summarizes the key aspects and relevant recommendations (table 3). The articles provide topic-specific recommendations, which together with those from the communication subcommittee article (Harmel et al., 2014) will contribute to discussion of potential development of ASABE modeling guidelines. The goal of this ASABE-led process is to enhance the field of hydrologic and water quality modeling. In the next phase of this process, the authors of this special collection and other interested parties will form groups to discuss and possibly write, review, and revise topic-specific guidelines.

ACKNOWLEDGEMENTS

This article introduces the ASABE 2015 Special Collection "Hydrologic and Water Quality Model Calibration Guidelines" in this issue of *Transactions of the ASABE*. The authors would like to thank all authors who contributed to the various topic-specific articles in this special collection for their invaluable contributions. In addition to authors of this article, these authors are S. Ale, D. M. Amatya, M. Arabi, J. V. Bonta, M. L. Chu, S. M. Dabney, G. W. Feyereisen, S. Finsterle, J. R. Frankenberger, M. W. Gitau, P. H. Gowda, T. R. Green, E. B. Haney, R. D. Harmel, J. Hernandez, J. Jeong, M. K. Jha, Y. Khare, I. Kisekka, N. Pai, P. B. Parajuli, Z. Qi, A. M. Sadeghi, V. Shedekar, A. Y. Sheshukov, R.W. Skaggs, M. D. Smolen, J. L. Steiner, X. Wang, M. J. White, G. Yagow, H. Yen, and M. A. Youssef.

Table 2. Documentation recommendations from Saraswat et al. (2015).

1.	Define and document the study purpose and end-user expectations.
2.	Describe the model used to enable readers assess the model suitability for the intended use and increase understanding of the modeling study.
3.	Describe the study area to clarify the setting in which the model was used.
4.	Document the methods used to collect observed data to enable readers to understand the relative uncertainty of the data.
5.	Document input data, such as study area boundaries, soil, land use, topography, management, and weather, required to set up and run the model for the study area.
6.	Document model calibration parameters and how they were obtained.
7.	Describe the C/V strategy used (Daggupati et al., 2015).
8.	Describe the performance measures and criteria used (Moriassi et al., 2015). A detailed description is essential when the model is applied after the C/V process is completed (Daggupati et al., 2015).

Table 3. Summary of key topic-specific model calibration and validation recommendations in this special collection.

Topic	Condensed Key Recommendations
Terminology (Zeckoski et al., 2015)	H/WQ modeling literature should adopt consistent usage of terminology. Recommended terminology includes controversial, misunderstood, or ambiguous terms in addition to basic terms related to H/WQ modeling.
Hydrologic processes and model representation (Arnold et al., 2015)	During calibration, use a diagnostic approach with both hard and soft data to account for and constrain modeled processes. This approach consists of four basic steps briefly described herein and in greater detail by the authors. Build soft data processes into automated calibration procedures to assist model users in ensuring that processes are realistic.
Spatial and temporal scales (Baffaut et al., 2015)	(1) Select a model that simulates processes relevant to the study objectives and at the appropriate spatial and temporal resolutions, (2) select appropriate temporal and spatial scales, (3) the available calibration and input data should determine spatial and temporal modeling scales, and (4) model C/V that involves multiple scales should be performed in successive steps considering the dominant processes. Suggestions to simplify modeling include breaking the project into smaller questions using appropriate scales when multiple spatial resolutions are considered and using two models when addressing interactions of processes that operate over differing spatial scales.
Model parameterization (Malone et al., 2015)	(1) Use the most uncertain and sensitive parameters, (2) minimize the number of optimized parameters, (3) where possible use site-specific measured or estimated parameter values, (4) in the absence of measured or estimated data, use “soft” data to optimize parameters during calibration, (5) use multiple criteria to help optimize parameter values, (6) constrain parameter values within justified ranges, and (7) use a warm-up period to reduce model dependence on initial condition state variables.
Calibration and validation strategies (Daggupati et al., 2015)	(1) Use single-site calibration for areas with uniform characteristics (e.g., soil, slope, vegetation, meteorology) and multi-site calibration for large areas with more varied, complex physical characteristics and/or when observed data for a given process are available at multiple locations within the study area, (2) apply an appropriate systematic calibration approach depending on the complexity of the application, ranging from single stage to stepwise, iterative, extensive parameter space, (3) use a two to three year warm-up (initialization) period for hydrology and five to ten years for sediment and nutrients, and (4) appropriately allocate calibration and validation data.
Sensitivity (Yuan et al., 2015)	Perform sensitivity analysis (SA) to identify key model parameters. Select appropriate SA methods depending on the intended purpose and assumptions about parameter certainty and linearity. Use local SA for simple linear problems and global SA for model non-linearity, non-monotonicity, and parameter interactions for complex systems.
Uncertainty (Guzman et al., 2015)	Perform uncertainty analysis for H/WQ modeling studies using the following general step-by-step procedure: (1) determine probability distribution functions (PDFs) of input parameters, (2) generate input samples based on PDFs, (3) apply a screening method to short list the important parameters, (4) refine parameters and their ranges for rigorous analysis, (5) perform model simulations to calculate desired outputs and decision variables, and (6) perform statistical analysis to obtain sensitivity indices, parameter rankings, predictive PDFs, and confident intervals. Estimate and report the uncertainty in the measured data (observational uncertainty) used to calibrate and validate models because of its impact on the evaluation and interpretation of model results.
Performance measures and criteria (Moriassi et al., 2015)	Consider recommendations for all modeling topics covered in this special collection before using recommended measures and criteria to assess model performance. Use multiple recommended graphical and statistical measures. General recommended criteria for the statistical measures can be adjusted based on several factors listed in this article and discussed in detail by the authors.
Documentation and reporting (Saraswat et al., 2015)	The following elements should be properly documented and reported: (1) study purpose and end-user expectations, (2) description of model used, (3) study area, (4) methods used to collect observed data, (5) input data needed to set up and run the model, (6) calibration parameters and how they are obtained, (7) calibration and validation strategy used, and (8) performance measures and criteria used. The modeling community is encouraged to move toward fully reproducible model calibration, validation, and use described by the authors.

REFERENCES

- Ahuja, L. R., Hanson, J. D., Shaffer, M. J., & Ma, L. (Eds.). (2000). Root Zone Water Quality Model: Modeling Management Effects on Water Quality and Crop Production. Highlands Ranch, Colo.: Water Resources Publications.
- Arnold, J. G., Moriassi, 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., Van Liew, M. W., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1491-1508. <http://dx.doi.org/10.13031/2013.42256>.
- Arnold, J. G., Youssef, M. A., Yen, H., White, M. J., Sheshukov, A. Y., Sadeghi, A. M., Moriassi, D. N., Steiner, J. L., Amatya, D. M., Skaggs, R. W., Haney, E. B., Jeong, J., Arabi, M., & Gowda, P. H. (2015). Hydrological processes and model representation: Impact of soft data on calibration. *Trans. ASABE*, 58(6), 1637-1660. <http://dx.doi.org/10.13031/trans.58.10726>.
- ASCE. (1993). Criteria for evaluation of watershed models. *J. Irrig. Drain. Eng.*, 119(3), 429-442. [http://dx.doi.org/10.1061/\(ASCE\)0733-9437\(1993\)119:3\(429\)](http://dx.doi.org/10.1061/(ASCE)0733-9437(1993)119:3(429)).
- Baffaut, C., Dabney, S. M., Smolen, M. D., Youssef, M. A., Bonta, J. V., Chu, M. L., Guzman, J. A., Shedekar, V., Jha, M. K., & Arnold, J. G. (2015). Hydrologic and water quality modeling: Spatial and temporal considerations. *Trans. ASABE*, 58(6), 1661-1680. <http://dx.doi.org/10.13031/trans.58.10714>.
- Barber, K. (Ed.) (2005). Canadian Oxford Dictionary (2nd ed.). Don Mills, Ontario, Canada: Oxford University Press.
- Baroni, G., Facchi, A., Gandolfi, C., Ortuani, B., Horeschi, D., & van Dam, J.C. (2010). Uncertainty in the determination of soil hydraulic parameters and its influence on the performance of two hydrological models of different complexity. *Hydrol. Earth Syst. Sci.* 14(2), 251-270. <http://dx.doi.org/10.5194/hess-14-251-2010>.
- Beven, K., & Binley, A. (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., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.*, 249(1), 11-29. [http://dx.doi.org/10.1016/S0022-1694\(01\)00421-8](http://dx.doi.org/10.1016/S0022-1694(01)00421-8).
- Daggupati, P., Pai, N., Ale, S., Douglas-Mankin, K. R., Zeckoski, R. W., Jeong, J., Parajuli, P. B., Saraswat, D., & Youssef, M. A. (2015). A recommended calibration and validation strategy for hydrologic and water quality models. *Trans. ASABE*, 58(6), 1705-1719. <http://dx.doi.org/10.13031/trans.58.10712>.
- Douglas-Mankin, K. R., Srinivasan, R., & Arnold, J. G. (2010). Soil and Water Assessment Tool (SWAT) model: Current developments and applications. *Trans. ASABE*, 53(5), 1423-1431. <http://dx.doi.org/10.13031/2013.34915>.

- Dressel, W.F. (2010). *Hydrologic Modeling Benchbook: Dividing the Waters*. Reno, Nev.: National Judicial College, Dividing the Waters Program.
- Duda, P. B., Hummel, P. R., Donigan Jr., A. S., & Imhoff, J. C. (2012). BASINS/HSPF: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1523-1547. <http://dx.doi.org/10.13031/2013.42261>.
- 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>.
- Flerchinger, G. N., Caldwell, T. G., Cho, J., & Hardegree, S. (2012). Simultaneous Heat and Water (SHAW): Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1395-1411. <http://dx.doi.org/10.13031/2013.42250>.
- Gowda, P. H., Mulla, D. J., Desmond, E. D., Ward, A. D., & Moriasi, D. N. (2012). ADAPT: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1345-1352. <http://dx.doi.org/10.13031/2013.42246>.
- Guzman, J. A., Shirmohammadi, A., Sadeghi, A. M., Wang, X., Chu, M. L., Jha, M. K., Parajuli, P. B., Harmel, R. D., Khare, Y., & Hernandez, J. (2015). Uncertainty considerations in calibration and validation of hydrologic and water quality models. *Trans. ASABE*, 58(6), 1745-1762. <http://dx.doi.org/10.13031/trans.58.10710>.
- Hansen, S., Abrahamsen, P., Petersen, C. T., & Styczen, M. (2012). Daisy: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1317-1335. <http://dx.doi.org/10.13031/2013.42244>.
- 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, 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., 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., and 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>.
- Henriksen, H. J., Trolborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C., & Madsen, B. (2003). Methodology for construction, calibration, and validation of a national hydrological model for Denmark. *J. Hydrol.* 280(1-4), 52-71. [http://dx.doi.org/10.1016/S0022-1694\(03\)00186-0](http://dx.doi.org/10.1016/S0022-1694(03)00186-0).
- Herr, J. W., & Chen, C. W. (2012). WARMF: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1385-1394. <http://dx.doi.org/10.13031/2013.42249>.
- Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*. Chichester, U.K.: John Wiley and Sons (reprinted by Blackburn Press in 2005).
- Jaber, F. H., & Shukla, S. (2012). MIKE SHE: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1479-1489. <http://dx.doi.org/10.13031/2013.42255>.
- James, L. D., & Burges, S. J. (1982). Selection, calibration, and testing of hydrologic models. In C. T. Haan, H. P. Johnson, & D. L. Brakensiek (Eds.), *Hydrologic Modeling of Small Watersheds* (pp. 437-472). St. Joseph, Mich.: ASAE.
- James, W. (2005). *Rules for Responsible Modeling* (4th ed.). Compiled and published by CHI (Computational Hydraulics International), Guelph, Ontario, Canada.
- Jansson, P. (2012). COUP Model: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1335-1344.
- Jarvis, N., & Larsbo, M. (2012). MACRO (v5.2): Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1413-1423. <http://dx.doi.org/10.13031/2013.42251>.
- 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), W03S04. <http://dx.doi.org/10.1029/2005WR004362>.
- Khu, S., Madsen, H., & Pierro, F. (2008). Incorporating multiple observations for distributed hydrologic model calibration: An approach using a multi-objective evolutionary algorithm and clustering. *Adv. Water Resources* 31(10), 1387-1398. <http://dx.doi.org/10.1016/j.advwatres.2008.07.011>.
- Krause, P., Boyle, D., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.*, 5, 89-97. <http://dx.doi.org/10.5194/adgeo-5-89-2005>.
- Lane, L. J., & Ferreira, V. A. (1980). Chapter 6: Sensitivity analysis. In W. G. Knisel (Ed.), *CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems* (pp. 113-158). Conservation Report No. 26. Washington, D.C.: USDA-SEA.
- 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.
- 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.
- Ma, L., Ahuja, L. R., Nolan, B. T., Malone, R. W., Trout, T. J., & Qi, Z. (2012). Root Zone Water Quality Model (RZWQM 2): Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1425-1446. <http://dx.doi.org/10.13031/2013.42252>.
- Malone, R. W., Yagow, G., Baffaut, C., Gitau, M. W., Qi, Z., Amatya, D. M., Parajuli, P. B., Bonta, J. V., & Green, T. R. (2015). Parameterization guidelines and considerations for hydrologic models. *Trans. ASABE*, 58(6), 1681-1703. <http://dx.doi.org/10.13031/trans.58.10709>.
- 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>.
- Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE*, 58(6), 1763-1785. <http://dx.doi.org/10.13031/trans.58.10715>.
- Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226-1227. <http://dx.doi.org/10.1126/science.1213847>.
- Refsgaard, J. C. (1997). Parameterisation, calibration, and validation of distributed hydrological models. *J. Hydrol.* 198(1), 69-97. [http://dx.doi.org/10.1016/S0022-1694\(96\)03329-X](http://dx.doi.org/10.1016/S0022-1694(96)03329-X).
- Refsgaard, J. C., & Henriksen, H. J. (2004). Modelling guidelines: Terminology and guiding principles. *Adv. Water Resources*, 27(1), 71-82. <http://dx.doi.org/10.1016/j.advwatres.2003.08.006>.
- Refsgaard, J. C., & Storm, B. (1996). Construction, calibration, and validation of hydrological models. In M. B. Abbott & J. C. Refsgaard (Eds.), *Distributed Hydrological Modeling* (pp. 41-

- 54). Dordrecht, The Netherlands: Kluwer Academic.
- Saltelli, A., Chan, K. S., & 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.
- Saraswat, D., Frankenberg, J. R., Pai, N., Ale, S., Daggupati, P., Douglas-Mankin, K. R., & Youssef, M. A. (2015). Hydrologic and water quality models: Documentation and reporting procedures for calibration, validation, and use. *Trans. ASABE*, 58(6), 1787-1797. <http://dx.doi.org/10.13031/trans.58.10707>.
- Seibert, J., & McDonnell, J. J. (2002). The quest for an improved dialog between modeler and experimentalist. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, and R. Turcotte (Eds.), *Calibration of Watershed Models* (pp. 301-315). AGU Monograph, Water Science and Applications Series Volume 6. Washington, D.C.: American Geophysical Union.
- Šimunek, J., vanGenuchten, M. T., & Šejna, M. (2012). HYDRUS: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1261-1274. <http://dx.doi.org/10.13031/2013.42239>.
- Skaggs, R. W., Youssef, M. A., & Chescheir, G. A. (2012). DRAINMOD: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1509-1522. <http://dx.doi.org/10.13031/2013.42259>.
- Tuppad, P., Douglas-Mankin, K. R., Lee, T., Srinivasan, R., & Arnold, J. G. (2011). Soil and Water Assessment Tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. *Trans. ASABE*, 54(5), 1677-1684. <http://dx.doi.org/10.13031/2013.39856>.
- USEPA. (2002). Guidance for quality assurance project plans for modeling. EPA QA/G-5M Report EPA/240/R-02/007. Washington, D.C.: U.S. Environmental Protection Agency, Office of Environmental Information.
- 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>.
- 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>.
- Wang, X., Williams, J. R., Gassman, P. W., Baffaut, C., Izaurralde, R. C., Jeong, J., & Kiniry, J. R. (2012). EPIC and APEX: Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1447-1462. <http://dx.doi.org/10.13031/2013.42253>.
- White, M. J., Harmel, R. D., Arnold, J. G., & Williams, J. R. (2012). SWAT Check: A screening tool to assist users in the identification of potential model application problems. *J. Environ. Qual.*, 43(1), 208-214. <http://dx.doi.org/10.2134/jeq2012.0039>.
- Yen, H., Bailey, R. T., Arabi, M., Ahmadi, M., White, M. J., & Arnold, J. G. (2014). The role of interior watershed processes in improving parameter estimation and performance of watershed models. *J. Environ. Qual.*, 43(5), 1601-1613. <http://dx.doi.org/10.2134/jeq2013.03.0110>.
- Yilmaz, K. K., Gupta, H. V., & Wagener, T. (2008). A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resources Res.*, 44(9), W09417. <http://dx.doi.org/10.1029/2007WR006716>.
- Yuan, Y., Khare, Y., Wang, X., Parajuli, P. B., Kisekka, I., & Finsterle, S. (2015). Hydrologic and water quality models: Sensitivity. *Trans. ASABE*, 58(6), 1721-1744. <http://dx.doi.org/10.13031/trans.58.10611>.
- Zeckoski, R. W., Smolen, M. D., Moriasi, D. N., Frankenberger, J. R., & Feyereisen, G. W. (2015). Hydrologic and water quality terminology as applied to modeling. *Trans. ASABE*, 58(6), 1619-1635. <http://dx.doi.org/10.13031/trans.58.10713>.