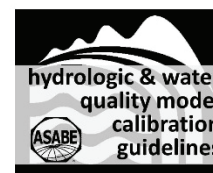# HYDROLOGICAL PROCESSES
AND MODEL REPRESENTATION:
IMPACT OF SOFT DATA ON CALIBRATION

J. G. Arnold, M. A. Youssef, H. Yen, M. J. White, A. Y. Sheshukov,
A. M. Sadeghi, D. N. Moriasi, J. L. Steiner, D. M. Amatya,
R. W. Skaggs, E. B. Haney, J. Jeong, M. Arabi, P. H. Gowda

hydrologic & water
quality model
calibration
guidelines

ASABE

ABSTRACT. *Hydrologic and water quality models are increasingly used to determine the environmental impacts of climate variability and land management. Due to differing model objectives and differences in monitored data, there are currently no universally accepted procedures for model calibration and validation in the literature. In an effort to develop accepted model calibration and validation procedures or guidelines, a special collection of 22 research articles that present and discuss calibration strategies for 25 hydrologic and water quality models was previously assembled. The models vary in scale temporally as well as spatially from point source to the watershed level. One suggestion for future work was to synthesize relevant information from this special collection and to identify significant calibration and validation topics. The objective of this article is to discuss the importance of accurate representation of model processes and its impact on calibration and scenario analysis using the information from these 22 research articles and other relevant literature. 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. In this article, model case studies are used to illustrate situations in which a model can show excellent statistical agreement with measured stream gauge data, while misrepresented processes (water balance, nutrient balance, sediment source/sinks) within a field or watershed can cause errors when running management scenarios. These errors may be amplified at the watershed scale where additional sources and transport processes are simulated. To account for processes in calibration, a diagnostic approach is recommended using both hard and soft data. The diagnostic approach looks at signature patterns of behavior of model outputs to determine which processes, and thus parameters representing them, need further adjustment during calibration. This overcomes the weaknesses of traditional regression-based calibration by discriminating between multiple processes within a budget. 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 budget that may not be directly measured within the study area, may be just an average annual estimate, and may entail considerable uncertainty. The advantage of developing soft data sets for calibration is that they require a basic understanding of processes (water, sediment, nutrient, and carbon budgets) within the spatial area being modeled and constrain the calibration.*

Keywords: Calibration, Field-scale models, Point models, Validation, Watershed models.

The authors are **Jeffrey G. Arnold, ASABE Fellow,** Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; **Mohamed A. Youssef, ASABE Member,** Associate Professor, Department of Biological and Agricultural Engineering, North Carolina State University, Raleigh, North Carolina; **Haw Yen,** Post-Doctoral Research Associate, Texas AgriLife Research, Temple, Texas; **Michael J. White,** Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; **Aleksey Y. Sheshukov, ASABE Member,** Assistant Professor, Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, Kansas; **Ali M. Sadeghi,** Soil Scientist, USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland; **Daniel N. Moriasi, ASABE Member,** Research Hydrologist, and **Jean L. Steiner,** Supervisory Soil Scientist, USDA-ARS Grazinglands Research Laboratory, El Reno, Oklahoma; **Devendra M. Amatya, ASABE Member,** Research Hydrologist, USDA Forest Service, Cordesville, South Carolina; **R. Wayne Skaggs, ASABE Fellow,** W. N. Reynolds and Distinguished University Professor, Department of Biological and Agricultural Engineering, North Carolina State University, Raleigh, North Carolina; **Elizabeth B. Haney,** Graduate Research Assistant, and **Jaehak Jeong, ASABE Member,** Assistant Professor, Texas AgriLife Research, Temple, Texas; **Mazdak Arabi,** Assistant Professor, Department of Civil and Environmental Engineering, Colorado State University, Ft. Collins, Colorado; **Prasanna H. Gowda, ASABE Member,** Research Agricultural Engineer, USDA-ARS Conservation and Production Research Laboratory, Bushland, Texas. **Corresponding author:** Jeffrey Arnold, USDA-ARS GSWRL, 808 East Blackland Road, Temple, TX; phone: 254-770-6508; e-mail: jeff.arnold@ars.usda.gov.

Water quality and hydrologic models are commonly used to assess the environmental impacts of land management and policy decisions. Models are increasingly being applied to large varied agricultural landscapes to address contemporary water resource issues in the context of climate change and sea level rise (Jayakrishnan et al., 2005). Moriasi et al. (2012) summarized the calibration approaches of 25 models in a special collection of 22 articles, each focus-

Transactions of the ASABE

Vol. 58(6): 1637-1660   2015 American Society of Agricultural and Biological Engineers   ISSN 2151-0032   DOI 10.13031/trans.58.10726   1637

ing on the individual model calibration and validation strategies. These models vary in scope from field-scale models that focus on flow, heat, and solute transport to large-scale watershed models that incorporate complex processes over spatially diverse subwatersheds. Regardless of scale, model accuracy is improved through calibration, and model uncertainty (thus utility) is evaluated via validation. Calibration and validation are important factors in the development of meaningful model predictions of potential future land use or climate effects. There are no universal standards for the calibration and validation of models in the current literature, as the procedure is generally dependent upon the processes in play at each model scale. Basic model processes include hydrology (water budget), erosion and sedimentation, plant growth, nutrient and carbon cycling, and contaminant fate and transport. Model processes, in varying degrees, are interconnected and impacted by land management, topography, climate, and scale. Point-scale models are used primarily to simulate physical and biological processes such as water and heat flow and reactive solute transport through a soil column and may be in finer time scales of less than an hour. At the field scale, the water balance processes include many variables that are interdependent with other processes, such as plant growth, soil properties, and weather. Nutrient cycling processes are complex and may vary greatly depending upon soil conditions even at the field scale. The carbon cycle appears simple at first glance; however, individual components, such as photosynthesis and soil carbon dynamics, are complex to simulate using a model. For example, photosynthesis is vegetation dependent and is influenced by resource (light, water, and nutrients) availability and environmental conditions. Soil carbon dynamics are influenced by the amount and biochemical composition of the soil organic matter, the quantity and quality (determined by C:N ratio and lignin content) of organic matter input (e.g., plant residue, animal manure, litter fall, root turnover) to the soil, and environmental factors (soil water and temperature) affecting biological activity. All of these factors influence the rate of soil organic carbon decomposition and the interaction between the mineral and organic forms of nitrogen and phosphorus (Youssef et al., 2005).

Model calibration techniques range from iterative manual methods to the use of fully automated calibration software. Field-scale models often simulate the main hydrological and biogeochemical processes, including infiltration and soil water distribution in the vadose zone, evapotranspiration, subsurface drainage, surface runoff, soil erosion, sediment transport, pesticide and nutrients dynamics, soil carbon cycling, and plant growth, for one or more plots or fields. Field-scale models are calibrated using manual or automated methods; both methods focus on hydrologic, chemical, or biologic parameters. Large-scale watershed models simulate complex hydrologic processes on a watershed or basin scale, in addition to the above processes in each plot or field. These processes include ditch/channel/riparian and reservoir processes, erosion and sediment movement and deposition, stream transport of nutrients, nutrient or pesticide degradation and transformation, water/air interactions, complex soil/plant/climate interactions,

algae and aquatic plants, and cycling in floodplains, estuaries, wetlands, and large hydraulic structures. Processes considered and calibration techniques used by each of the 25 models are described in the special collection (Moriasi et al., 2012).

This article describes the importance of realistically simulating all critical processes in the hydrological balance for calibration and validation of small- and large-scale models. For example, if surface runoff is overestimated, it is likely that evapotranspiration (ET) and/or subsurface and tile flow are underestimated, resulting in overestimation of sediment yields and underestimation of subsurface nitrate and other soluble contaminant yields. This will cause further error when parameterizing variables related to sediment and nutrient transport and result in unrealistic policy recommendations when running scenarios that target erosion and fertilizer management. Problems are even more compounded at the watershed scale when multiple fields or subbasins are simulated and output is routed through channels, flood plains, and reservoirs. Thus, it is important to reasonably simulate nutrient and sediment sources and sinks within a watershed in addition to their loads at a gauging station (outlet). If upland erosion is overpredicted, channel erosion must be underpredicted to match measured gauge loads. The management practices designed to reduce erosion from the landscape may then show significant impact on total sediment yields, while in reality the practices would have little impact at the basin outlet. It is also important at the watershed scale to accurately simulate proper source load allocations. For example, excellent calibration statistics can be obtained at a stream gauge outlet even though point sources are underestimated and the loads from agricultural lands are overestimated. This could result in policy scenarios that overestimate the impact of conservation or best management practices (BMPs) on agricultural and forested lands. For most models, little guidance is provided for considering process representation in the calibration and validation procedures.

The objectives of this article are to: (1) synthesize processes considered and calibration techniques used to account for processes within a field or watershed for the models in the special collection (Moriasi et al., 2012), (2) summarize other relevant literature related to process representation and calibration, (3) demonstrate the importance of proper process representation utilizing soft data and its impact on calibration/validation scenario analysis using case studies, and (4) provide recommendations for calibration/validation.

SYNTHESIS OF PROCESSES SIMULATED BY MODELS IN SPECIAL COLLECTION

The 25 hydrologic and water quality models in the special collection of calibration and validation concepts summarized by Moriasi et al. (2012) are categorized as: (1) flow, heat, and solute transport models, (2) field-scale models, and (3) watershed-scale models in order to facilitate discussion and compare concepts. Different approaches were taken for describing the processes in each category

due to general differences in the processes. For flow, heat, and solute transport models, the equations and solution techniques are similar in each model, so we described the general processes and then noted any models that differed from the general approach. Since all field-scale models represented the same basic processes of hydrology, soil erosion and sediment transport, vegetation, carbon, nitrogen and phosphorus, and pesticides, we divided the discussion by process and then described how each model simulated that process. Since watershed-scale models generally represent processes of more spatial and temporal complexity, the processes and calibration concepts of each model were summarized individually. Not only do watershed-scale models consider additional processes, they often employ significantly different levels of complexity. Table 1 shows the scale or dimension of each model, the processes that are represented, and typical calibration techniques.

FLOW, HEAT, AND SOLUTE TRANSPORT MODELS

Point-scale models are usually suited for modeling processes that are mostly one-dimensional, at soil profile or horizon column, and represent a single footprint at the surface. Models considered include SHAW, COUPMODEL, SWIM3, MACRO, VS2DI, HYDRUS, STANMOD, and TOUGH (table 1). These models are normally used to quantify the basic physical or chemical processes that occur

Table 1. Scale and calibration processes considered by 25 models in the special collection of 22 articles (Moriasi et al., 2012).

Model	Reference	Scale/Dimension	Model Processes	Typical Calibration Techniques
Flow, heat, and solute transport models				
SHAW	Flerchinger et al., 2012	1-D	Soil water, heat, and solute concentrations	Sensitivity analysis, manual or automated. Stepwise, trial and error, optimization algorithm
COUPMODEL	Jansson, 2012	1-D	Hydrology, erosion, pesticides, sediments, nutrients, plant growth	Manual and automated, stepwise systematic, use of double mass plot technique
SWIM3	Huth et al., 2012	1-D	Water and solute transport	Manual, water balance
MACRO	Jarvis and Larsbo, 2012	1-D	Macropore flow, pesticides	Forward, sequential and iterative procedure
VS2DI	Healy and Essaid, 2012	2-D	Water, heat, and solute transport	Manual, parameter estimation programs
HYDRUS	Šimůnek et al., 2012	1, 2, or 3-D	Water, heat, and solute transport, carbon dioxide flux	Gradient-based local optimization or automated
STANMOD	van Genuchten et al., 2012	1, 2, or 3-D	Solute transport in soils and groundwater	Weighted nonlinear least square method
MT3DMS	Zheng et al., 2012	1, 2, or 3-D	Multispecies solute transport	Manual and automated, variance-covariance matrix
TOUGH	Finsterle et al., 2012	1, 2, or 3-D	Multiphase, multicomponent fluid flows in porous and fracture geologic media	Weighted least squares objective function
Field-scale models				
DRAINMOD	Skaggs et al., 2012	Field	Hydrology, carbon and nitrogen cycles, plant growth	Manual, stepwise, iterative
ADAPT	Gowda et al., 2012	Field	Hydrology, erosion, nutrients, pesticides, subsurface tile drainage	Partition flows, annual water and nutrient balances
CREAMS/GLEAMS	Knisel and Douglas-Mankin, 2012	Field	Edge of field sediment, nutrient and pesticide losses	Manual, stepwise: hydrology, sediment, nutrients/pesticides
RZWQM2	Ma et al., 2012	Point	Hydrology, plant growth, nutrients, pesticides	Manual
EPIC	Wang et al., 2012	Field	Hydrology, carbon and nitrogen cycles, plant growth	Manual and automated, multi-site data
WEPP Hillslope	Flanagan et al., 2012	Hillslope	Hydrology, soil erosion across the landscape	
DAISY	Hansen et al., 2012	Point	Hydrology, carbon cycle, energy balance, nitrogen, crop production, pesticides	Stepwise: bioclimate, vegetation, soil and field management parameters
Watershed-scale models				
APEX	Wang et al., 2012	Small watershed/ whole farm	Hydrology and plant growth	Manual and automated, multi-site data
WEPP Watershed	Flanagan et al., 2012	Small watershed	Hydrology, soil erosion	Stepwise procedure: hydrology, erosion
SWAT	Arnold et al., 2012	Watershed to river basin	Hydrology, plant growth, sediment, nutrients, pesticides	Manual and automated, multi-site data
HSPF	Duda et al., 2012	Watershed to river basin	Hydrology, snowmelt, pollutant loadings, erosion, fate and transport	Iterative procedure of parameter evaluation and refinement
WAM	Bottcher et al., 2012	Watershed to river basin	Hydrology, sediments, nutrients	Manual and automated, multi-site data
KINEROS	Goodrich et al., 2012	Watershed	Streamflow and sediment	Simple manual to complex automated
MIKE-SHE (DAISY coupling)	Jaber and Shukla, 2012	Watershed to river basin	Surface and subsurface water dynamics, evapotranspiration, flow, water quality	Manual and automated

during the transport of water, exchange or transfer of heat, and/or movement of various nutrients, pollutants, pathogens, and carbon through the unsaturated zone (or vadose zone) in a soil column. The fate and transport processes occurring near the soil surface profiles are mostly dependent on soil properties but are also affected by the earth-atmosphere surface and the bottom of the column subsurface boundary conditions. These models have a varying degree of complexity and can be multidimensional. Analytical solutions can be sought for one-dimensional simplified formulations, especially under steady-state conditions, whereas a numerical approach is developed for multivariable non-linear physical processes. Commonly used, physically based approaches to solve processes in the soil profile are discussed in the following sections.

Water Flow Equation

All point-scale models consider water movement in the soil profile that generally occurs under two conditions, saturated or unsaturated. The saturated condition, however, is a special case of the unsaturated condition and occurs when the soil moisture content is at a maximum level for a particular soil type. Darcy's law (Darcy, 1856) was the first mathematical description of water movement in soil that showed the proportionality of the flux of water to the hydraulic gradient:

$$\vec{q} = -k\,\nabla H \tag{1}$$

where $\vec{q}$ is the soil water flux, k is the proportionality factor, which is known as the hydraulic conductivity, and ∇H is the gradient of the hydraulic head H in the multidimensional space.

Equation 1 is commonly used to evaluate a situation in which a steady-state flow or near steady-state prevails, that is, the flux remains constant at any point along the conducting water flow system. In the actual field conditions, however, most soil water transport processes occur under transient-state conditions, where the magnitude and the direction of the flux and hydraulic gradient vary with time. Therefore, considering the law of conservation of mass, in the multidimensional system, the relationship between change in soil moisture content (θ) and flux (q) or H can be expressed as:

$$\frac{\partial \theta}{\partial t} = -\nabla \cdot \vec{q} \quad \text{or} \quad \frac{\partial \theta}{\partial t} = \nabla \cdot k\,\nabla H \tag{2}$$

Equation 2, called the Richards equation (Richards, 1931), must be supplemented by the soil constituent relationships $k(\theta)$ and $H(\theta)$ that can be found from hydraulic experiments on the soil column. The Richards equation is widely used in point-scale models and is considered a cornerstone of water flow formulations that represent mostly the movement of water in unsaturated soils. It is a nonlinear partial differential equation and has a limited number of closed-form analytical solutions that were developed for special cases of homogeneous soils with simplified functional forms of $k(\theta)$ and $H(\theta)$, and certain initial and boundary conditions (Philip, 1969, 1972, 1992; Wooding, 1968). Most of these solutions consider infiltration applica-

tions and assume constant water content or flux values for steady-state water flow conditions. Šimůnek et al. (1999) reported that a large number of analytical models for one-, two-, and three-dimensional solute transport problems were recently incorporated into the comprehensive software package STANMOD (van Genuchten et al., 2012).

Several numerical methods based on finite difference, finite volumes, or finite element approximations have been developed to solve the water flow equation in unsaturated porous media and utilize the H-based form, θ-based form, or H-θ mixed form of equation 2. In most numerical schemes, the storage term is linearized using the Newton-Raphson or Picard methods, and equation 2 is solved iteratively. Such an approach is implemented in HYDRUS (Šimůnek et al., 1998), SWAP (van Dam et al., 1997), and MODFLOW-2000 (Thoms et al., 2006).

The SHAW model (Flerchinger and Saxton, 1989), for example, is one of the point-scale models that uses a modified form of the Richards equation and simulates heat and water movement through a plant residue-soil system. This one-dimensional model considers a profile that extends from the vegetation canopy to a specified depth within the soil profile. Preferential or macropore flow represents rapid flow of water and chemicals through porous media such as soil profiles and has recently attracted the attention of scientists and modelers working in the environmental field. Incorporation of this rapid movement of pollutants from the upper soil profile into lower behavior has created a multitude of difficulties in modeling vadose zone transport. One of the approaches used in dual-porosity models considers that main flow occurs in preferential flow paths while water is immobile in the soil matrix. For example, the model MACRO (Jarvis and Larsbo, 2012) uses a capacitance approach to calculate macropore flow, while the Richards equation is used to model micropore flow in the soil matrix. The more sophisticated approach is sought to describe the infiltrating flow under very dry conditions where flow breaks into fingers, but the Richards equation has proved to be unconditionally stable (Egorov et al., 2003). This approach looks at the flow process being non-equilibrium and assumes a kinetic relationship between water content and matrix head (Nieber et al., 2005; Chapwanya and Stockie, 2010).

Reactive Contaminant Transport

Many models, such as SHAW, STANMOD, VS2DI (Healy and Essaid, 2012), and SWIM3 (Huth et al., 2012), can simulate transport of reactive contaminants simultaneously with water flow. In these models, after water fluxes $\vec{q}$ are quantified in equations 1 and 2, a physically based model of reactive contaminant transport can be described as an advective-dispersive transport problem with the diffusive flux considered based on Fick's law (Fick, 1855) and solute subject to sorption and degradation on the soil particles. The flux $\vec{q}_c$ explains the mass of solute (C) or any other pollutants diffusing across a unit cross-sectional area per unit time as:

$$\vec{q}_c = -D\nabla C + \vec{q}C \tag{3}$$

where D is the solute dispersion coefficient and can be a function of θ. Thus, by substituting equation 3 into the general expression of the flux of the solute change in the vertical direction per unit volume and unit time, a flow equation will result that encompasses both dispersion and mass flow components for solute transport in the soil profile:

$$\frac{\partial(\theta C + C_s)}{\partial t} = \nabla \cdot D\nabla C - \nabla \cdot (\vec{q}C) \tag{4}$$

where C_s is the concentration of contaminant in the sorbed phase and subject to the sorption law. The solution of equation 4 requires both initial and boundary conditions.

Heat Transfer

The heat transfer in the soil column is usually modeled by the convection-conduction equation:

$$c\frac{\partial T}{\partial t} - L\frac{\partial \theta_s}{\partial t} = \nabla \cdot \lambda \nabla T - c_w \nabla \cdot (\vec{q}T) \tag{5}$$

where T is the temperature, c is the heat capacity, λ is the thermal conductivity, and L is the latent heat of evaporation or freezing. The second term in the left side of equation 5 accounts for the heat released during water phase change (freezing or evaporation) and was included in models such as SHAW.

Other Transport Processes

Many existing models, such as COUPMODEL, SWIM3, MACRO, VS2DI, HYDRUS, STANMOD, MT3DMS (Zheng et al., 2012), and TOUGH (Finsterle et al., 2012), can also model a variety of other physical and chemical processes: nutrient and carbon cycle, gas diffusion, ice buildup, soil freezing and melting, colloid detachment and movement, transport of charged particles, pesticide transport, etc. For example, the PRZM model (Carsel et al., 1985) was originally developed to assess pesticide fate and transport within the crop's root zone but was recently coupled with the VADOFT model (Mullins et al., 1993) to solve the Richards equation in the vadose zone. Despite major recent modifications, the model still lacks a macropore and preferential flow component (Sadeghi et al., 1995).

Surface Boundary Conditions

Most of the point-scale models (DAISY, SHAW, SWAP, HYDRUS, and COUP, among others) are subject to complex non-linear upper boundary conditions at the earth-atmosphere surface. These conditions evaluate water and energy budgets and provide values for surface flow, infiltration rates, and water vapor and heat fluxes. Two main conditions that work at various temporal scales and that point-scale models account for are the water budget and energy budget conditions, which can generally be written as:

$$(P - E)A + Q_i - Q_o = \frac{dW}{dt} \tag{6}$$

$$R = L \cdot E + H + G \tag{7}$$

where P is the rate of precipitation, E is the rate of evapora-

tion, A is the footprint surface area, Q_i and Q_o are the surface inflow and outflow rates, W is the water volume stored at the surface footprint, R is the net incoming radiation flux, and H and G are the specific fluxes of sensible and conductive heat. The equations at the surface boundary (eqs. 6 and 7) link inputs taken from climate models (P and R) with the variables dependent on processes inside the column, land cover, input from adjacent columns, etc. Most one-dimensional point-scale models simulate flow and transport in the column only vertically and assume that excess water at the surface boundary is removed by surface runoff. Interconnectivity of flow between vertical columns is considered in multi-dimensional point-scale models by modeling two- or three-dimensional processes in the column (eqs. 1 through 5) as well as overland processes at the surface. More advanced linkage of surface processes with inputs from soil is included in hillslope or field-scale models.

FIELD-SCALE MODELS

The field-scale models in the special issue (CREAMS/GLEAMS, DRAINMOD, ADAPT, RZWQM, EPIC, WEPP Hillslope, and DAISY) were developed to simulate the physical, chemical, and biological processes occurring in the soil-water-plant system and have been used to simulate the effects of management practices on agricultural production and soil and water resources with the goal of increasing productivity, reducing cost, and enhancing sustainability. Simulated agricultural management practices include crop rotations, fertilizer management practices, tillage and plant residue management practices, land application of animal manure, irrigation, and drainage water management. These models usually simulate the water balance (fig. 1), plant growth, nutrient cycling in soil (fig. 2), carbon dynamics (fig. 3), pesticide dynamics, and agriculture land management. Depending on the model, predictions can be a subset of crop/biomass yield, edge-of-field surface, subsurface and tile flow, sediment nitrogen, phosphorus, and pesticide losses, and nitrogen and soluble pesticide losses via surface runoff and leaching.

Over the years, field-scale models have been expanded and enhanced. GLEAMS was developed by improving CREAMS to better represent soil layers, crop rotations, and chemical transport. EPIC was also developed by improving GLEAMS to better represent cropping systems and management while accounting for the impact of erosion on crop productivity (Williams et al., 2008; Wang et al., 2012). ADAPT was developed as a water management model for high water table soils by incorporating DRAINMOD-based routines for tile drainage and water table simulation into GLEAMS (Gowda et al., 2012). RZWQM development in the mid-1980s and early 1990s was based on existing models including CREAMS, GLEAMS, Opus, and PRZM. Later, the model was further improved by incorporating the SHAW model for conducting full energy balance at the soil surface, DRAINMOD-based routines for tile drainage, and DSSAT crop modules for simulating crop growth and yield (Ma et al., 2012). The current version of RZWQM is considered an agricultural system model that simulates crop yield, water and nitrogen balances, and pesticide fate as influenced by management practices, soil properties, and
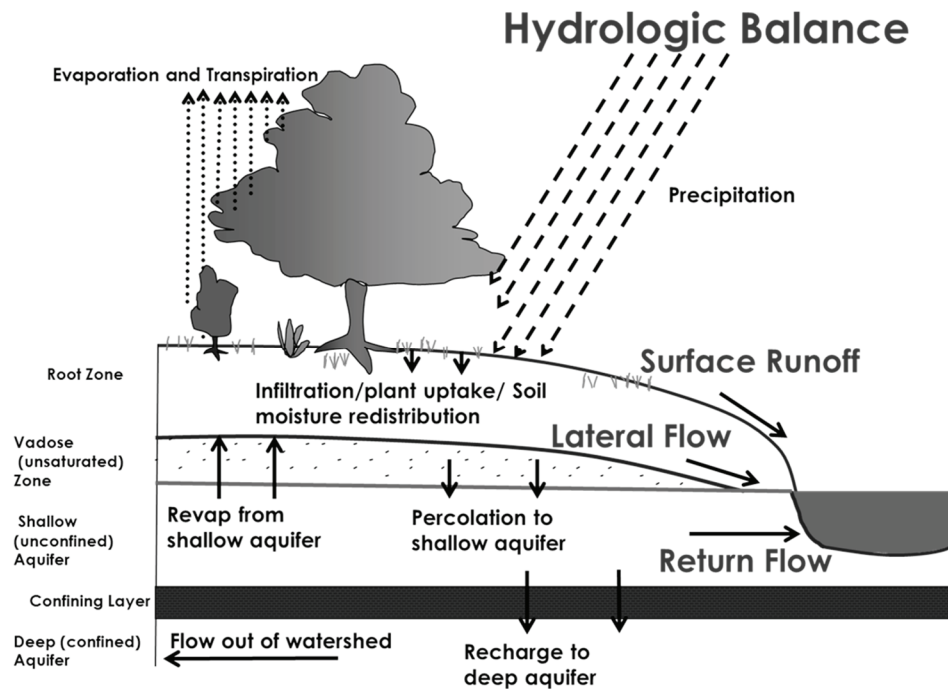
Figure 1. Hydrologic processes in the water balance (adapted from Arnold et al., 2000).
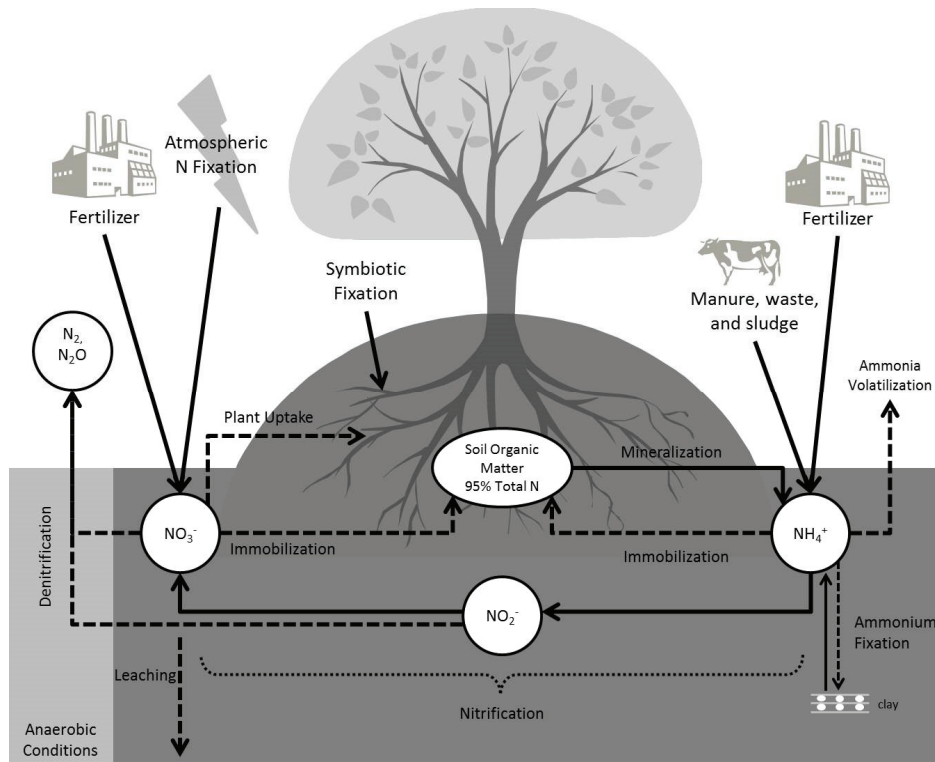

Figure 2. Processes in the nitrogen cycle.

climatic conditions. DRAINMOD was developed in the late 1970s as a hydrological model for simulating the performance of agricultural drainage and related water management systems. Over the years, the model has been expanded and enhanced to become a system model for simulating hydrology, soil carbon and nitrogen dynamics, and vegetation growth for agricultural and forest ecosystems on poorly or artificially drained shallow water table soils (Skaggs

et al., 2012). The WEPP Hillslope model was developed in the early 1980s to replace the USLE (Flanagan et al., 2012). Original water balance, plant growth, and nutrient components in WEPP were modified from the EPIC model.

The term "field" as a spatial scale could be defined as a spatial unit with homogeneous characteristics, including weather, soil, topography, cropping system, and manage-
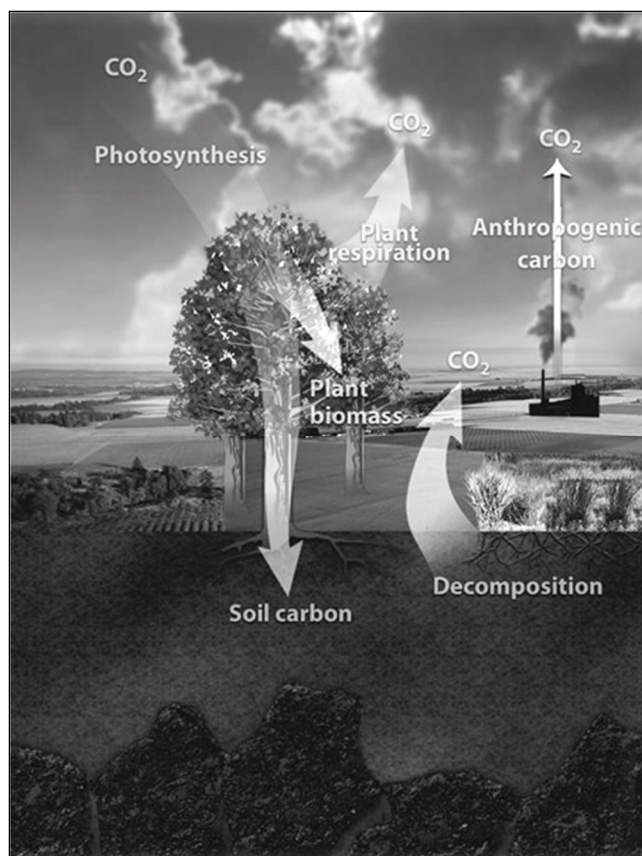
Figure 3. The carbon cycle (courtesy of the U.S. Department of Energy, Office of Biological and Environmental Research).

ment practices. With the exception of DAISY, the field-scale models reviewed in this article are in fact "point" scale models that simulate flow and transport in one dimension. As such, the inputs and outputs of these models represent field-average values, and the size of the field should be determined to preserve the definition of the "field" as a spatial unit. These models can be linked or incorporated into larger-scale models to simulate relatively small watersheds while taking advantage of the process-based nature of most of the field-scale models. For example, DAISY has been linked to MIKESHE (Gregersen et al., 2007), and watershed-scale models have been developed based on DRAINMOD (Fernandez et al., 2005; Amatya et al., 2004).

The temporal scale of these models ranges from daily to sub-hourly depending on the model and the simulated processes. For example, CREAMS/GLEAMS and ADAPT use a daily time step (Gowda et al., 2012; Knisel and Douglas-Mankin, 2012). DRAINMOD simulates water balance on an hourly basis but switches to a daily time step when drainage fluxes are slow and to a sub-hourly time step during precipitation events (Skaggs et al., 2012) and for nitrogen dynamics and heat balance (Youssef et al., 2005). RZWQM simulates water and energy balances using a sub-hourly time step and uses a daily time step for simulating plant growth, nutrient cycling, and pesticide dynamics (Ma et al., 2012). The sub-hourly time scale has been used in these models to satisfy the convergence and stability requirements of numerical solutions to the partial differential equations used to describe water flow, mass transport, and heat transfer.

The field-scale models reviewed in this article can be categorized as process-based or more precisely as "hybrid" models. In hybrid models, some processes are empirically represented in the model structure in order to reduce input requirement and reduce uncertainty in model predictions while maintaining the key advantages of process-based models. The models differ in their level of detail in simulating different physical, chemical, and biological processes occurring in the soil-water-plant system.

Hydrology

The purpose and scope of the field-scale models determine how hydrologic processes affecting the water balance are represented in these models. The hydrological component of CREAMS/GLEAMS is a simple rainfall-runoff model that uses the empirical curve number method to estimate runoff in response to rainfall events (Knisel and Douglas-Mankin, 2012). The model is not applicable to lowland areas where subsurface water movement has a large influence over the water balance. The hydrological component of DRAINMOD was mainly developed for naturally poorly drained, high water table soils. In DRAIN-MOD, subsurface water movement is based on a mass balance approach assuming the soil profile above the groundwater table is drained to equilibrium, a simple and proven reliable approach for high water table soils where tile drainage is frequently used. In both RZWQM and DAISY, numerical solutions to the Richards equation are used to describe subsurface water movement; both models solve the equation in the vertical dimension, while DAISY has the option of modeling two-dimensional flow.

Most models use the Green-Ampt equation to estimate infiltration rates. DAISY and DRAINMOD simulate surface depression storage occurring when precipitation rate exceeds infiltration rate; both models generate surface runoff once surface storage capacity is exceeded. RZWQM does not represent surface storage and generates surface runoff once precipitation rate exceeds infiltration rate. Since these models are point scale, they do not route surface runoff to the field edge; model predictions of surface runoff are assumed to be at the field edge. This assumption is reasonable for small fields. EPIC and WEPP Hillslope have options to use curve number or Green-Ampt depending on the availability of subdaily rainfall.

Evapotranspiration (ET) is modeled as a function of potential (or crop reference) evapotranspiration (PET) and available soil water within the plant root zone. Different models use different methods for estimating PET; CREAMS/GLEAMS use the Priestly-Taylor method, and RZWQM uses the Shuttleworth-Wallace equation. EPIC has options for using Penman-Monteith, Priestley-Taylor, or Hargreaves. The original version of DRAINMOD uses the Thornthwaite method for estimating PET, with the option of daily PET inputs from any available source; the latest forestry version of the model estimates PET using the Penman-Monteith method with canopy conductance estimated as a function of climatologically regulated stomatal conductance and leaf area index (LAI) that is internally predicted by the forest growth component of the model (Tian et al., 2012a). DAISY has the most detailed represen-

tation of ET process. It simulates ET (evaporation/transpiration) from four different sources: (1) evaporation of liquid water in the snowpack, (2) evaporation of water intercepted by the canopy and the surface litter layer, (3) evaporation of ponded water on the soil surface, and (4) evapotranspiration of soil water in the root zone. It uses two approaches to simulate ET depending on the availability of high-resolution weather data. One approach couples surface water balance and surface energy balance and requires high-resolution weather data. In the other approach, which decouples surface water and energy balances, the concept of PET or crop reference ET is used and the model user can select the FAO Penman-Monteith, the Makkink, or the Hargreaves-Samani equation to estimate PET.

Tile drainage is implemented on large areas of cropland in the U.S. Midwest and Southeast. DRAINMOD was developed particularly for simulating crop production systems on high water table soils with subsurface drainage tile or open ditches. The model uses the steady-state Hooghoudt equation to calculate drainage flux during water table drawdown and uses Kirkham's equation to calculate drainage flux during surface water ponding. The rise and fall of the water table in response to rainfall/irrigation and drainage is determined based on a relationship between drainage volume and water table depth that is obtained from soil water characteristic data. Deep percolation (seepage) is estimated using Darcy's equation. DRAINMOD's approach has been adopted in other models including ADAPT and RZWQM.

Accurate representation of soil temperature is required for hydrologic simulations in cold regions characterized by freezing and thawing, snow accumulation, and snowmelt. It is also needed for simulating the biochemical processes regulating nutrient cycling and pesticide dynamics. DAISY, RZWQM, and DRAINMOD simulate soil temperature variation along the soil profile using numerical solutions to the heat equation, which describes heat transfer due to conduction and convection. Usually, measured air temperature defines the upper boundary condition, and the long-term average temperature for the location defines the lower boundary condition. Models determine the form of precipitation based on air temperature. Snowmelt is usually simulated using the empirical degree day method.

Macropores mainly occur in cracking soils. Earthworms and decomposing plant roots also create macropores in the biologically active root zone. Macropore flow can significantly alter the subsurface water movement and influence the transport of agrochemicals to groundwater and surface water. For example, macropore flow can carry sediment, phosphorus, and pesticides bound to the sediment through shallow subsurface tile drains to receiving surface water. RZWQM and DAISY model macropore flow, including the interactions between macropores and the surrounding soil matrix. Workman and Skaggs (1989) developed a DRAINMOD-based model that simulates macropore flow in drained agricultural land. In this model, the simple water balance approach of DRAINMOD was replaced with numerical solution of the Richards equation. However, this improvement of DRAINMOD was not incorporated in the model's distribution version.

Soil Erosion and Sediment Transport

The WEPP Hillslope model simulates sediment deposition and degradation across a hillslope in addition to sediment yield leaving the hillslope. EPIC and CREAMS/GLEAMS simulate erosion using a modified form of the Universal Soil Loss Equation (USLE) and estimate sediment transport as a function of particle/aggregate size and transport capacity. The erosion and sediment component of the model determines the enrichment ratios required for modeling the transport of chemicals bound to the sediment (Knisel and Douglas-Mankin, 2012). Because it was developed based on CREAMS/GLEAMS, the ADAPT model uses the same approach for modeling soil erosion and sediment transport. DAISY, RZWQM, and DRAINMOD do not simulate soil erosion and sediment transport. Previous research was conducted to link DRAINMOD and CREAMS to enable DRAINMOD to route surface runoff, estimate soil erosion, and model the transport of sediment and chemicals bound to sediment. This work, however, was not incorporated in the distribution version of DRAINMOD.

Vegetation

Vegetation plays a central role in the hydrology and biogeochemistry of agricultural and forest ecosystems. It influences and is influenced by key hydrological and biogeochemical processes that affect water, carbon, and nutrient balances in these ecosystems. Vegetation influences evapotranspiration, the largest component of the water balance in most climatic regions. Plant uptake is the largest sink of nitrogen and phosphorus in both agricultural and forest ecosystems. Crop residue, litter fall, and dead roots are the primary carbon sources for soil organic matter.

The two main goals of field-scale models are: (1) predicting the effects of management practices, soil type, and climatic conditions on crop growth and yield, and (2) predicting the effects of crop production on hydrology, water quality, and soil quality. In order for models to provide reliable predictions, they need to represent vegetation growth at a sufficient level of detail, depending on the purpose and scope of the model. Comparing the field-scale models reviewed in this article, one can make two important observations. The first observation is that two approaches, one is empirical and the other is process-based, were generally followed in representing vegetation or its influence on water and nutrient balances. The other important observation is that several of the models, which are still supported and actively updated, have been enhanced by replacing the simpler empirical representation of vegetation with a more process-based approach. This shift toward a more mechanistic representation of vegetation in hydrologic and water quality models is driven by the need for application of these models outside their original scope. A clear example is the application of models to predict the potential effect of climate change on crop production systems, including both crop yield and water quantity and quality.

EPIC uses a generic plant growth model (Williams et al., 1989) that computes daily photosynthetically active radiation (PAR) as a function of solar radiation, leaf area index

(LAI), and a light extinction coefficient. Constraints on optimal daily biomass accumulation include water (limiting and aeration), temperature, and nitrogen and phosphorus stress. EPIC uses a harvest index to partition aboveground biomass and yield that also considers a water limiting stress. Algorithms have been added to simulate competition within a plant community (Kiniry et al., 1992). The EPIC plant growth model has been parameterized for over 150 plants and is used in other models including SWAT, APEX, and WEPP Hillslope.

CREAMS/GLEAMS empirically represent processes related to vegetation, including N and P uptake by plants, N fixation by legume crops, and mineralization of crop residue (Knisel and Douglas-Mankin, 2012). The purpose of these models, which are not currently supported, was to assess agricultural nonpoint-source pollution and evaluate management practices for alleviating the negative environmental impacts of agriculture. This empirical approach for modeling vegetation was then believed to be adequate for the scope of the model.

A generic but detailed plant growth model was developed for DAISY. It simulates the photosynthesis process using two approaches (one empirical and one process-based) and considers the effect of source (light, soil water, and nitrogen) availability on process rate. It simulates plant respiration (growth and maintenance) and partitions fixed biomass into different plant organs. Plant development stage influences partitioning of biomass, leaf and root turnover, and senescence. This level of detail in representing plant growth is expected to provide better predictions of growth response to stressors as well as better predictions of nutrient uptake and plant organic material input to the soil.

RZWQM originally used a simple generic model that simulates the growth of annual crops, which was parameterized for corn, soybean, and winter wheat. Alternatively, a fully empirical approach based on growth curves was used to estimate water and nitrogen uptake. With the evolution of RZWQM as an agricultural system model, the process-based crop modules of DSSAT were incorporated into RZWQM (Ma et al., 2012).

DRAINMOD was originally developed to simulate the effect of drainage design and management on crop yield. The model adopted the stress-day index approach, a simple empirical approach for quantifying the effect of soil water-related stresses on crop yield, and the model predicted a relative yield (relative to a site-specific potential yield) in response to changes in the drainage design and management. Nitrogen versions of DRAINMOD were developed in response to major concerns regarding the negative water quality impacts, mainly N losses, of agricultural drainage. Plant uptake and crop residue input to the soil were empirically represented in the model. This approach led to inaccurate plant uptake and yield predictions, which affected the predicted nitrogen balance in drained cropland and increased the uncertainty in model predictions of N losses via tile drainage as affected by both farming and water management practices. The latest version of DRAINMOD has evolved into a fully integrated system model with a process-based vegetation component for raw crops, perennial grasses, and forests. The row crop component of DRAIN-MOD is based on the crop modules of DSSAT.

Carbon

The need for the explicit representation of the carbon cycle in the soil-water-plant system varies depending on the scope of the model. N and P mineralization, the largest source of N and P in forests and the second largest source of N and P in fertilized agriculture, is regulated by soil organic carbon dynamics during the decomposition process. C transformations, however, occur at much slower rates compared to hydrologic, N, and P processes. In many applications requiring relatively short-term simulations, hydrologic and water quality models can adequately represent N and P dynamics without the need for explicit representation of C. For other applications, a more comprehensive modeling approach that explicitly represents C cycling in the simulated system would be required. For example, assessment of the sustainability of an emergent land use change, such as replacing a forest with bio-energy grass, would require a whole-system model to run long-term simulations and predict the effects of this land use change on water, carbon, nutrients, and sediment budgets. In this case, the predicted change in C sequestration is an important factor in assessing the long-term sustainability of this land use change. In many studies involving assessment of the potential impacts of climate change, modeling the changes in C fluxes under different scenarios is required.

Generally, carbon dynamics are simulated by dividing the soil carbon and added fresh organic material into different pools or compartments that vary in their biochemical composition (N, P, and lignin contents) and rate of decomposition. The names and numbers of these compartments may vary among models. DRAINMOD and DAISY have components for simulating soil organic carbon dynamics. The other field-scale models reviewed in this article do not have C modeling components. EPIC (Causarano et al., 2007) and DRAINMOD adapted the CENTURY model approach for simulating soil carbon dynamics. This approach divides organic carbon into three soil pools (active, slow, and passive): two aboveground and belowground residue pools (metabolic and structural), and a surface microbial pool. The soil organic matter component of DAISY divides soil organic C into three main types: soil organic matter or humus (SOM), soil microbial biomass (SMB), and added organic material AOM). The SOM type is further divided into three pools: fast, slow, and inert. The other two types are further divided into fast and slow pools. Both DRAINMOD and DAISY simulate organic carbon decomposition using first-order kinetics. The organic carbon sources considered by the models include animal manure, plant residue, litterfall, and root turnover. Both models simulate the transport of dissolved organic carbon.

Nitrogen and Phosphorus

The negative impacts of N land P losses from agricultural lands on groundwater and surface water quality have long been recognized. The field-scale models reviewed in this article have incorporated a relatively detailed nitrogen cycle in the soil-water-plant system. Differences among models were found, especially in the method of representing various processes affecting the N balance. The simulat-

ed N cycle includes application of N fertilizer, animal manure, and plant residue, wet and dry atmospheric deposition, N mineralization and immobilization, plant uptake, nitrification, denitrification, ammonia volatilization, and N losses via surface runoff and leaching. Simulated forms of N include organic N, nitrate, and ammonium-N.

DAISY and DRAINMOD simulate N transport using numerical solutions of the advection-dispersion-reaction equation (Hansen et al., 2012; Youssef et al., 2005). The forestry version of DRAINMOD also simulates the fate and transport of dissolved organic nitrogen (Tian et al., 2012b). Empirical approaches are used in CREAMS/GLEAMS, ADAPT, and RZWQM to simulate N losses due to leaching and surface runoff. In RZWQM, all nitrogen biological transformations are assumed to follow first-order kinetics (Knisel and Douglas-Mankin, 2012; Gowda et al., 2012; Ma et al., 2012). In DRAINMOD, the Michaelis-Menten kinetics model is used to describe nitrification, denitrification, and urea hydrolysis (Youssef et al., 2005). In DAISY, nitrification is described using Michaelis-Menten kinetics, and denitrification is described using an index-type model (a function of nitrate, carbon, and soil anaerobic status; Hansen et al., 2012). Unlike DAISY and DRAINMOD, RZWQM does not simulate soil organic carbon dynamics. To realistically describe the interaction between organic and mineral N forms, RZWQM divides organic N into several pools (fast and slow residue pools; fast, intermediate, and slow humus pools; and three microbial pools) with different carbon-to-nitrogen ratios and decomposition rates. This approach, which is similar to soil organic carbon modeling, makes RZWQM as able as DRAINMOD and DAISY to capture the temporal change in net mineralization following fresh organic material input to the soil (Ma et al., 2012; Hansen et al., 2012; Youssef et al., 2005). For nitrogen transformations, EPIC uses first-order rate constants as functions of soil moisture and temperature. EPIC considers active, stable, and fresh organic pools and ammonium and nitrate mineral pools (He et al., 2006). CREAMS/GLEAMS and ADAPT do not simulate N immobilization (Knisel and Douglas-Mankin, 2012; Gowda et al., 2012).

Differences among the models in simulating plant uptake are mainly attributable to differences in modeling plant growth. Empirical representation of vegetation uses an empirical uptake function that defines cumulative plant uptake from planting until the end of the growing season. Daily potential N uptake is estimated from the empirical uptake function. For process-based modeling of plant growth, daily potential N uptake is estimated based on the biomass fixed by the photosynthesis process and the nitrogen content of different plant organs. Uptake from nitrate and ammonium forms is assumed to occur according to their relative proportions within the root zone. Nitrogen fixation by legumes is assumed to occur only after the depletion of mineral N in the root zone.

Among the field-scale models reviewed in this article, EPIC, CREAMS/GLEAMS, and ADAPT are the only models to simulate the fate and transport of phosphorus in the agro-ecosystem. Simulated P processes and transformations include P mineralization, plant uptake, sorption onto the soil matrix, runoff and leaching losses of soluble P, and losses of adsorbed P with sediment. The P cycle is simulated with similar algorithms as the N cycle.

Pesticides

The fate and transport of pesticides are simulated by CREAMS/GLEAMS, ADAPT, EPIC, RZWQM, and DAISY. More than one type of pesticide can be surface applied, incorporated or injected into the topsoil, or sprayed on plant foliage. RZWQM simulates the application of slow-release forms of pesticides. Pesticides intercepted by foliage and plant residues may be washed off during rainfall events. The models simulate three main processes affecting the fate of pesticides in the agro-ecosystem: degradation, sorption, and transport. Degradation is represented by first-order kinetics. RZWQM uses either equilibrium-based or kinetics-based sorption to describe the sorption of pesticides into the soil matrix. Both CREAM/GLEAMS and RZWQM track the fate of pesticide degradation products that might be harmful to the environment.

WATERSHED-SCALE MODELS

Compared to point-scale and field-scale models, watershed-scale models are difficult to calibrate due to multiple fields draining to channels and streams, resulting in (1) spatial variation in land use and management, soil properties, climate, topography, geology, and (2) addition of processes for streams/channels, floodplains, aquifers, ponds, reservoirs, water regulating structures, wetlands, and their interactions with adjoining fields (fig. 4).

To account for these multiple processes and sources/sinks, watershed models divide watersheds and basins into subbasins, fields, or cells. Field-scale model developers acknowledge that all processes in a balance are never known for an individual field. This uncertainty is exacerbated at the watershed scale, as a watershed or basin may contain hundreds or thousands of subbasins or fields. Thus, scaling to a watershed or basin creates additional calibration challenges that include: (1) spatial landscape calibration (multiple subbasins that should be calibrated for water, nutrient, and sediment balances) and (2) sources and sinks (deposition and degradation in streams/channels, flood plains, riparian buffers, wetlands, impoundments, and aquifers). While most watershed model developers suggest that water, sediment, and nutrient balances should be calibrated spatially, that rarely happens in practice, primarily due to limited data availability. In the majority of calibration studies, the models are calibrated to the flow and load (or concentration) at one or more stream gauges within the watershed. Empirical data collected at a few locations do not meet the needs of modelers attempting to integrate across temporal and spatial scales (Wallenstein et al., 2012), thus limiting the applicability of models to simulate physical processes adequately. In recent years, there has been an increasing recommendation from the modeling community to parameterize process-based models using multi-criteria calibration as opposed to just using one output parameter, which is generally streamflow at the outlet of a watershed (Harmel et al., 2014; Dai et al., 2010; Boyle et al., 2003; Meixner et al., 2003; Shrestha and Rode, 2008).
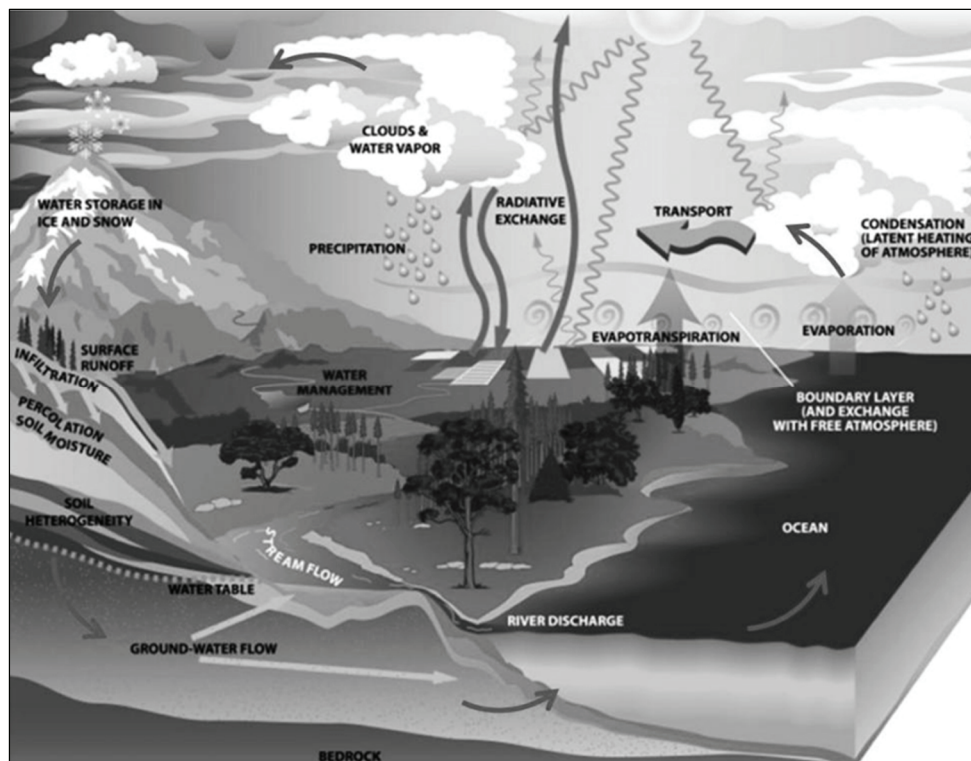
Figure 4. The hydrologic cycle (courtesy of NASA; http://science.nasa.gov).

A few publications in the special issue articles described attempts to incorporate spatial processes and sources/sinks within the watershed model calibration. HSPF (Duda et al., 2012) has routines to calibrate water balance, snow, hydraulic flow, in-stream sediment, and in-stream water quality. The other watershed models shown in table 1 calibrate parameters related to these processes; however, most calibration case studies used streamflow data. MIKE-SHE showed an example of calibrating reservoir levels, and SWAT and MIKE-SHE gave examples of calibrating surface runoff and baseflow separately. Recently, Dai et al. (2010) demonstrated the advantage of a bi-criteria calibration of MIKE-SHE using both the streamflow and field water table depth as a surrogate of soil moisture and ET for simulating the hydrology of a 155 ha coastal forested watershed.

MIKE-SHE

MIKE-SHE can simulate only limited surface water quality processes using the advection-dispersion equation and groundwater quality with a random-walk tracking method (Jaber and Shukla, 2012). Jaber and Shukla (2012) noted a need for additional efforts to enhance the coupling of MIKE-SHE with DAISY (Daisy, 2011) to simulate the nitrogen and carbon cycles on agricultural lands (Refsgaard and Hansen, 2010). In a similar effort, Dai et al. (2011) successfully applied MIKE-SHE linked with the Wetland-DNDC model (Cui et al., 2005), originally developed for uplands, to assess the spatial distribution of N cycles and greenhouse gas fluxes from forested wetlands in coastal South Carolina. MIKE-SHE-DNDC simulates flow in streams and channels but only total nutrient flux in drainage water, without considering nutrient movement in the

soil water and loss in the air and also without in-stream nutrient transport and transformation (Amatya et al., 2013). However, the use of MIKEBASIN (DHI, 2012) could alleviate this issue by dividing the watershed or basin into several subwatersheds or subbasins and routing the flow and nutrient load along the stream channel. Similarly, Jaber and Shukla (2012) demonstrated, with a case study of reservoirs, the capabilities of the coupled MIKE-SHE and MIKE11 with a 1-D hydrodynamic model to solve the complex processes and their interactions in the face of climate and land use changes. This was also shown earlier by Thompson et al. (2004), who successfully applied the MIKE-SHE/MIKE-11 model to evaluate surface hydrology and ditch/channel flow processes, including ditch water evaporation, using bi-criteria validation with groundwater table and ditch water level on a lowland wet grassland in England. The linked model was further applied to evaluate the hydrologic impacts of climate change using two scenarios on the same site (Thompson et al., 2009). These examples including the use of MIKE21 and MIKEFLOOD demonstrate the potential of the linked model in addressing complex watershed-scale hydrologic and hydraulic processes for both uplands and lowlands in the face of climate and land use changes.

To account for spatial landscape processes, WAM plots the output distributions by land use for each grid cell in the watershed. Although each cell may have different slopes and soils, a visual inspection can reveal if losses from different land uses are reasonable. The WEPP watershed model requires sediment deposition and degradation across the landscape and the particle sizes of eroded sediment for comprehensive calibration and validation, which are only

available for a limited number of research plots. The DAISY model, when coupled with MIKE-SHE, uses LAI and ET estimated from remote sensing, which improves spatial crop yield prediction, similar to the potential of MIKE-SHE/Wetland-DNDC for forested wetland conditions.

APEX

APEX, a direct extension of EPIC, can describe hydrology, forest growth, N fate processes, and plant competition in fields and in the more complex multi-subarea landscapes of whole farms and small watersheds (Wang et al., 2012), but mostly for upland conditions. APEX has a multi-run function that allows for tree growth prior to other vegetation establishment or allows for the input of tree start year with a given weight and height for stand development. Wang et al. (2012) reported the processes and their influential inputs and parameters, including flow and sediment routing in channels, for the APEX model. APEX can route N according to a specified path, and while it can also sum its results at the sub-watershed and watershed levels, particular N loadings can be tracked back to the source from which they originated, allowing for problem detection to occur more readily. For scenarios that simulate the effects of land management and BMP practices (filter strips, contour buffers, etc.), Wang et al. (2012) recommended coupling APEX with SWAT for large watershed and regional studies because of APEX's lack of detailed stream processes and databases needed in large-scale simulations. Such a coupling will also enhance the simulation of forested fields, including riparian buffers; however, the authors report a need for further research on effective and unbiased parameterization during the simultaneous calibration process of APEX and SWAT. Although Wang et al. (2012) enlisted the ongoing work on various flood routing subroutines in APEX, there is also a need to consider in-stream nutrient transport and transformation processes.

DRAINMOD

DRAINMOD-based watershed-scale models have been developed by linking the process-based, field-scale DRAINMOD hydrology model (Skaggs et al., 2012) with hydraulic transport subroutines for routing outflows through drainage canal and stream networks. These models, with various levels of complexity for flow and nitrogen transport, have been applied to poorly or artificially drained watersheds of several thousand hectares (Konyha and Skaggs, 1992; Amatya et al., 1997, 2004; Fernandez et al., 2002, 2005, 2006, 2007). Most recently, one of these models was applied to assess the hydrologic effects of conversion of forest lands into agricultural croplands (Kim et al., 2013) and to assess the hydrologic impacts of climate change (Amatya et al., 2006). In the context of multi-criteria validation, these models have the ability to calibrate the hydrology using the field water table, in-stream flows, and main outlet streamflows within a watershed. All of the currently available DRAINMOD watershed-scale models rely on empirically derived, published, or measured field N concentration data rather than simulating these data within the model, except for the model developed recently for agricultural lands using DRAINMOD-NII (Negm et al., 2014). Linking the recently developed DRAINMOD-

FOREST (Tian et al., 2012a, 2012b), a comprehensive model fully integrated for simulating the processes of hydrology, nitrogen, carbon, and productivity in forest ecosystems, with DRAINMOD-NII (Youssef et al., 2005) as a core N and C submodel with the available in-stream flow and nutrient transport and transformation routines of various complexities (Fernandez et al., 2002, 2005, 2006, 2007; Amatya et al., 2004) would correct this deficiency. While these models are limited to simulating hydrologic and transport processes on poorly drained lands, they are capable of simulating backwater effects in streams and channels, which are characteristic of coastal systems (Amatya et al., 2013).

KINEROS2

KINEROS2 (Goodrich et al., 2012) is an event-based model that simulates runoff, erosion, and sediment transport. Ideally, a measured soil water budget is preferred for calibrating and validating KINEROS2 for event runoff, erosion, and sediment transport. Therefore, the model developers recommend co-locating soil moisture measurements at the rain gauges with recording intervals of no longer than 1 h to define pre-storm soil moisture levels. However, Goodrich et al. (2012) also recognize the limitation of observed data in validating different processes of the hydrologic cycle and sediment transport (Al-Qurashi et al., 2008) for distributed and/or semi-distributed watershed models. KINEROS2 uses a stepwise, multi-scale calibration approach to improve calibration by using the dynamic version of WEPP option (Bulygina et al., 2007) instead of the traditional "lumped" calibration in which uncertainty is high and model performance is poor when moving across spatial scales (Goodrich et al., 2012).

HSPF

HSPF (Duda et al., 2012) simulates nonpoint-source runoff and pollutant loadings for a watershed and performs flow and water quality routing in stream reaches and well-mixed lakes and impoundments. The calibration and validation procedures in HSPF include database development and watershed segmentation, followed by calibration and validation of hydrology, sediment, and water quality in that order, iteratively, depending on available data. HSPF simulates runoff from four components: surface runoff from impervious areas directly connected to the channel network, surface runoff from pervious areas, interflow from pervious areas, and groundwater flow (Duda et al., 2012). According to Duda et al. (2012), a complete annual hydrologic calibration involves a successive examination of the following four characteristics of the watershed hydrology, in the following order: (1) annual water balance, (2) seasonal and monthly flow volumes, (3) baseflow, and (4) storm events. This is indicated as follows:

$$RO = P - AET - DP \pm \Delta SW \qquad (8)$$

where RO is runoff, P is precipitation, AET is actual evapotranspiration, DP is deep percolation, and ΔSW is change in soil moisture. The second step is the seasonal or monthly distribution of runoff, which is controlled by the infiltration parameter. In practice, incoming water is divided among

surface runoff, interflow, upper zone soil moisture storage, and percolation to lower zone soil moisture and groundwater storage (Duda et al., 2012). The third step involves the baseflow component. Finally, when an acceptable agreement has been attained for annual and monthly volumes and baseflow conditions, simulated hydrographs for selected storm events are effectively altered by adjusting surface detention and interflow parameters. Detailed budgets are considered by HSPF during calibration for snow, sediment erosion, in-stream sediment transport, nonpoint-source loading, and water quality (Duda et al., 2012). Duda et al. (2012) provided two case studies that demonstrate how budgets are taken into account during the calibration process.

WEPP

WEPP is a hydrologic and soil erosion model for predicting runoff, soil detachment and sediment deposition, and sediment yield at the hillslope profile and small watershed scales (Flanagan et al., 2012). The model simulates surface and subsurface water movement comprising percolation, deep seepage, subsurface lateral flow, and impervious subsurface layers, such as rock parent material below forest soils. Detailed observed soil moisture content, surface runoff, subsurface drainage, sediment loss, and sediment particle size characteristics for each storm event are needed to validate the hydrologic and sediment processes. For hillslope and profile scale calibration, WEPP performs detailed tests of the various model components related to hydrology, ET, plant growth, and erosion and sediment transport, when comprehensive datasets are available. Detailed parameterization is also performed, especially at the hillslope or profile scale. However, at the watershed scale, calibration becomes complex because of the considerable variation in soil types across a catchment, particularly as catchment size increases (Flanagan et al., 2012).

WAM

WAM simulates the constituents important to eutrophication processes in water bodies (water, total suspended solids, biological oxygen demand, and soluble and particulate nitrogen and phosphorus) within a watershed (Bottcher et al., 2012). Bottcher et al. (2012) stated that "when the model parameters represent actual physical quantities (e.g., fertilizer rates, planting dates, stream layout dimensions, land slope, irrigation rates, etc.), the adjustment of such parameters must be limited by the physical knowledge of the basin" (p. 1372). Therefore, in WAM, the first step is to verify the accuracy of the input parameters within realistic physical limits. The next step is the calibration of the statistical or empirical model parameters. As with HSPF, hydrologic and hydraulic processes are calibrated first in WAM to ensure that the correct flows and stages are simulated in the reaches. This is followed by calibration and validation of sediment and nutrients. WAM provides a detailed table of steps suggested to adequately calibrate and validate WAM, in which each calibration component is further separated into the simulation processes relevant to the calibrated components. A case study is provided illustrating how WAM is calibrated and validated, in which basin water balance is also provided.

SWAT

SWAT simulates weather, hydrology, soil temperature and properties, plant growth, sediments, nutrients, pesticides, bacteria and pathogens, and land management (Arnold et al., 2012). In SWAT, water balance is the driving force behind all the processes because it impacts plant growth and the movement of sediments, nutrients, pesticides, and pathogens. Simulation of watershed hydrology in SWAT is separated into the land and in-stream or routing phases. The land phase controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each subbasin, while the in-stream phase is the movement of water, sediments, etc., through the channel network of the watershed to the outlet. Because SWAT input parameters are process-based, they must be held within a realistic uncertainty range during the calibration process. Under ideal conditions, calibration and validation in SWAT are process and spatially based. For example, the streamflow process consists of the water balance in the land phase of the hydrology, including ET, lateral flow, surface runoff, return flow, tile flow (if present), channel transmission losses, and deep aquifer recharge (Arnold et al., 2012). Irrigation applications to the land, as well as point discharges of water, must be accounted for. However, Arnold et al. (2012) are aware of the data limitations that affect ideal calibration and provide recommendations on how SWAT should be calibrated and validated with limited measured data to validate these processes. For instance, streamflow is generally split between surface and baseflow components using the baseflow filter to ensure that overland processes are properly simulated. Hydrologic and water quality budgets are emphasized in SWAT, as indicated by the development of the SWAT Check program (White et al., 2012) to ensure that the processes are simulated realistically. As SWAT is used more extensively and intensively, more studies are focusing on nutrient budgets as well, such as the detailed study by Yen et al. (2014a) described earlier. The goal of all these efforts is to obtain good and reliable calibration and validation performance results for the right reasons in order to minimize uncertainty in simulated scenarios results.

Common Themes of Watershed Modeling Processes

The authors of the articles describing watershed-scale modeling (Goodrich et al., 2012; Duda et al., 2012; Arnold et al., 2012; Bottcher et al., 2012) in the model calibration special issue of *Transactions of ASABE* (Moriasi et al., 2012) identified several common themes. All emphasized that the most critical aspect of watershed modeling is that the model user should have a sound understanding of the watershed characteristics and processes that are important in the watershed under study to properly represent that watershed in the model. A critical assessment of the available knowledge and data is essential to select the appropriate model for the desired application and, indeed, to determine if a particular application can be conducted in a data-scarce watershed.

Precipitation data are particularly important as a driving force of watershed-scale processes. The spatiotemporal scale of precipitation data will determine whether applica-

tion of a detailed process-level model such as KINEROS2 is feasible or not. The degree to which a rain gauge or set of rain gauges realistically represents the actual precipitation in a watershed depends on the spatiotemporal scale of the gauge network and the characteristics of storms in the region. In snowmelt-driven watersheds, sparseness of representative temperature and solar radiation data present additional limitations (Duda et al., 2012). Because the processes are so interactive in heterogeneous watersheds, the coefficients of the process equations in the models are also highly interactive. All of the model developers recommended a specific sequence of calibration steps. The initial step may include using expert knowledge as well as available data in selecting initial parameter values. Preliminary assessment of the plant biomass and annual nutrient budgets is often helpful to ensure that the model is configured to produce believable values. The plant component is highly interactive with components that are generally calibrated in more detail, and achieving reasonable plant biomass and yield data may provide a calibration approach for data-sparse watersheds (e.g., Ávila-Carrasco et al., 2012). The general sequence of calibration is to calibrate the hydrology parameters first at multiple time scales, explicitly addressing baseflow and storm runoff conditions, followed by sediment, explicitly examining the ratios of upland and channel sources to ensure that the ratio is within the acceptable range for that watershed. Nutrient and pesticide calibration follows sediment calibration, since sediment-borne transport of contaminants is an important process to be evaluated.

OTHER LITERATURE RELATED TO PROCESSES AND CALIBRATION

In addition to the calibration strategies summarized here for the 25 hydrologic and water quality models in the special collection (Moriasi et al., 2012), other research has focused on incorporation of processes into calibration procedures. Yilmaz et al. (2008) suggest a diagnostic approach for model calibration and defined model diagnosis as the process by which we make inferences about possible causes of an observed undesirable symptom using targeted evaluations of the input-state-output response of the model. The diagnostic approach is used to: (1) identify signature patterns of behavior related to primary watershed functions using observed data, (2) extract diagnostic signature indices related to these behaviors, (3) test the ability of a model to reproduce these signature indices, (4) detect and group model components and parameters related to each signature index, and (5) resolve signature index match failures with modifications to model components and parameters. Yilmaz et al. (2008) gave examples of signature measures related to vertical soil moisture redistribution and long-term behavior of baseflow and major components of the water balance. In the long-term water balance example, the authors examined annual and monthly flow and ET processes and determined through the diagnostic approach that the symptom was caused by incorrect parameterization of potential ET input variables. Gupta et al. (2008) and Wagener

and Gupta (2005) noted that model evaluation strategies that rely on regression-based measures of performance (e.g., Nash-Sutcliffe efficiency) are weak at discriminating between the varied influences of multiple model components. Thus, the diagnostic approach can overcome the weakness of regression-based calibration methods by incorporating an understanding of key processes and an understanding of the input-state-output response of the model.

Often, the model user does not notice or understand the symptom and signature measure. In the case study illustrated by Yen et al. (2014a), the user may obtain excellent agreement (high Nash-Sutcliffe coefficient) at a stream gauge using a regression-based criterion with an optimization scheme and yet not realize that processes within the watershed are compensating some components with specific signature measures (e.g., denitrification and surface NO_3 are overestimated while tile flow are severely underestimated). Thus, an understanding of the dominant processes within a watershed is critical to proper calibration and parameterization of inputs. White et al. (2012) developed a diagnostic tool to analyze SWAT output and suggested signature measures related to water balance, nutrient balances, plant growth, and sediment sources/sinks and corresponding adjustments to input parameters.

As noted earlier, long-term time series of all major hydrologic processes are rarely if ever available. However, general information on individual processes may be obtained from the literature or even by visual inspection of the study site. Seibert and McDonnell (2002) suggested the use of "hard" and "soft" data for multi-criteria model calibration. Hard data are defined as measured time series, typically at a point (e.g., streamflow, groundwater levels, or soil moisture) that is commonly used in regression-based calibration techniques. In their study of a watershed in New Zealand, Seibert and McDonnell (2002) cited the willingness to use only hard data in model calibration as a hindrance to moving forward. They used general information on reservoir volume and percent new water as soft data to constrain the calibration. Soft data are defined as information on individual processes within a balance that may not be directly measured in the study area, may be an average annual estimate, and may entail considerable uncertainty. Examples of soft data include regional estimates of baseflow ratios or ET, average depths of groundwater tables, average annual runoff coefficients for various land uses, annual rates of denitrification from research plots found in the literature, event mean concentrations, nutrient/sediment export coefficients, sediment deposition from reservoir sedimentation studies, average crop/vegetation LAI, county crop yields, etc. Other researchers have used maps of surface saturated area to constrain parameter ranges for TOPMODEL (Franks et al., 1998) and fuzzy measures for ET (Franks and Beven, 1997). Seibert and McDonnell (2002) argued that soft data represent a new dimension to the model calibration process that could: (1) enable dialog between experimentalists and modelers, (2) be a formal check on the reasonableness and consistency of internal model structures and simulations, and (3) specify realistic parameter ranges often ignored in today's automatic calibration routines.

Winsemius et al. (2009) advanced the work of Seibert and McDonnell (2002) by presenting a framework for integrating hard and soft hydrological information in model calibration. Similarly, they defined hard information signatures as data for which the limits of acceptability may be objectively derived from the distribution of long series of observed values. Soft signatures are less effective in parameter conditioning, or their limits of acceptability cannot be objectively derived. A framework was developed by Winsemius et al. (2009) to integrate the Generalized Likelihood Uncertainty Estimation (GLUE) method (Beven and Binley, 1992) within a "limit of acceptability" approach. The framework consists of the following steps: (1) search for information content in the form of signatures in any data that are readily available, (2) divide the information into hard and soft data, (3) perform Monte Carlo simulations using the limits of acceptability of the hard information, (4) perform a new Monte Carlo simulation using the soft information as constraints, (5) determine what constraints are still lacking, and (6) after collection of new information, update the parameter distributions with new targets. In an application to the Luangwa River in Zambia, Winsemius et al. (2009) used soft data on the shape of the recession curve, spectral properties of daily streamflows, and monthly water balance as constraints in calibration. Additional sources of soft data suggested by Winsemius et al. (2009) include satellite-based ET estimates and estimates of large-scale water storage from the GRACE gravity information (Tapley et al., 2004). Yen et al. (2014a) used a similar framework to incorporate hard and soft data in the calibration procedure. They used the time series of streamflow and nitrate concentrations as hard data for regression-based calibration and literature values of annual denitrification and tile flow nitrates as soft data to constrain the optimization. Vaché and McDonnell (2006) suggested stream and soil water residence times as soft data for watershed diagnostics. Additional valuable sources of soft data on hydrology and water quality may come from the long-term experimental watersheds maintained by federal and other agencies in a given region.

There have been additional attempts reported in the literature to account for process components and spatial variability using soft data when calibrating models. In their study of simulating runoff response at 12 catchments using the HBV model, Zelelew and Alfredson (2012) found that varying up to a minimum of four to six influential parameters for high flow conditions and up to a minimum of six influential parameters for low flow conditions can sufficiently capture the catchments' response characteristics. Perrin et al. (2008) developed a new method called "discrete parameterization" that relies on the sole use of prior information on parameters gained from other catchments (soft data). The authors reported that although this method is not as efficient as a classical global search calibration approach, it provides more robust parameter sets when the flow time series (hard data) available for calibration are shorter than two years, which is generally the case in many poorly gauged catchment studies.

Several studies have calibrated both surface runoff and subsurface flow contributions to total streamflow. Most use daily streamflow and partition or filter quick surface response and delayed groundwater response. Zhang et al. (2011) developed a new scheme to simultaneously calibrate surface flow and baseflow in SWAT by combining evolutionary multi-objective optimization and baseflow separation techniques. Similar techniques have been used in SWAT applications by Arnold et al. (2000) in the Upper Mississippi River basin and by Santhi et al. (2008) in the Ohio/Tennessee basin. Vasquez and Feyen (2010) separated baseflow from interflow, which allowed for calibration of hydraulic conductivity for a MIKE-SHE application despite a lack of piezometric data.

The advantage of using filtered streamflow data to calibrate both surface and groundwater flow is that this method only relies on daily flow data (Arnold et al., 2000). The method also ensures that the basic components of the water balance (surface runoff, subsurface flow, and ET, assuming deep percolation is negligible) are realistic, that surface runoff is accurate for surface sediment transport and groundwater percolation, and that discharge is accurate for nitrate and soluble pesticide transport. It is difficult for the hydrograph separation technique to differentiate between interflow, baseflow, and tile flow. Soft data assumptions on tile flow contributions usually need to be made based on knowledge of drainage research studies from nearby or similar fields. The disadvantage of using the baseflow ratio for larger watersheds is that the baseflow ratio is representative of the entire drainage area, which typically includes multiple land uses, soils, and topography. Other attempts at using soft data to quantify processes in the water balance include that of Immerzeel and Droogers (2008), who used satellite-based ET estimates to spatially calibrate SWAT in the Krishna River basin in southern India. In addition, Hymer et al. (2000) used satellite-derived soil water maps to calibrate the SHAW flow and heat model.

Another approach to calibrating the nutrient balance is to calibrate the crop yield or plant biomass. Most of the field-scale models (EPIC, RZWQM2, DRAINMOD, ADAPT, and WEPP) suggest that crop yield should be calibrated, and numerous examples are available in the literature of yield and biomass calibration and validation (Ávila-Carrasco et al., 2012). RZWQM2 includes a yield and biomass case study in the special issue (Moriasi et al., 2012). Nair et al. (2011) suggested that crop yield comparison be added to the SWAT calibration procedure. Compared to traditional approaches that do not include crop yield calibration, Nair et al. (2011) produced improved prediction efficiencies, especially for the nutrient balance. Challinor et al. (2004) calibrated crop yields (peanut) across India for the period 1966-1989 and illustrated the impact of proper crop yield validation on a regional nutrient budget. Although crop yield and biomass removal is only one component of the nutrient balance, it can also be a good indicator of water uptake and residue remaining after harvest, as well as being a major component of the nutrient balance. The advantage of using crop yields in calibration is that the information is readily available in the U.S. by county for each crop and is also available from numerous research and test plots.

Case Studies Demonstrating Impacts of Processes

Field-Scale Modeling: Water Balance

This case study demonstrates the importance of accurate representation of the individual processes influencing the water balance at the field scale. It demonstrates how the differences in simulating these processes, which may not be captured by the regression-based calibration and validation, may lead to large differences in model predictions of the hydrologic response to a management practice change in a scenario analysis.

Agricultural drainage is essential for crop production on about 25% of the cropland in the U.S. It improves trafficability, providing timely access for performing field operations, and removes excess soil water from the root zone. However, drainage significantly alters the hydrology and N cycling in naturally poorly drained soils, causing increase in subsurface water movement and N leaching losses to groundwater and receiving surface waters.

Drainage water management (DWM), also referred to as controlled drainage, is a management practice developed for reducing nutrient export from drained cropland. DWM involves the use of an overflow control device to reduce drainage rates by raising the water level in the drainage outlet during periods when intensive drainage is not required. DWM works by reducing drainage volumes and enhancing denitrification. The performance of DWM depends on several factors, including climatological conditions, soil properties, cropping system and farming practices, and drainage system design. Thus, the effectiveness of DWM is expected to vary from location to location and from year to year. Field-scale models such as DRAINMOD and RZWQM have been used to predict long-term performance of DWM for different field conditions.

This case study involves the field-scale models DRAINMOD and RZWQM, which were calibrated and compared using a ten-year dataset from a drained corn and soybean field in Iowa (Thorp et al., 2007, 2009). The calibrated models were used to simulate the performance of DWM across the U.S. Midwest (Thorp et al., 2009), where the practice can potentially be applied to reduce N losses from millions of hectares of drained cropland to the Mississippi River and the Gulf of Mexico. The results of model calibration and comparison as well as the results of the DWM scenario analysis conducted using the two calibrated models are used to demonstrate the effect of individual hydrologic processes on model predictions of the water balance at the field scale and the associated predictions of the performance of DWM.

Thorp et al. (2007, 2009) evaluated and compared the RZWQM and DRAINMOD models using ten years of measured hydrologic, water quality, and crop yield data collected for a drained corn-soybean field in central Iowa. DWM was not implemented on the site during the calibration and validation period. On average over the ten-year period, the predictions by both models of the largest two components of the annual water balance (ET and subsurface drainage flow) were similar. According to the water balances predicted by the two models, the sum of the aver-

age annual surface runoff and vertical seepage to the underlying aquifer was less than 10% of average annual precipitation. Compared to the RZWQM predictions, DRAINMOD predicted 1.7 cm year^{-1} more surface runoff and 0.6 cm year^{-1} less vertical seepage. Nash-Sutcliffe modeling efficiency values for annual drainage flow predictions were 0.92 and 0.91 for DRAINMOD and 0.98 and 0.82 for RZWQM during the calibration and validation periods, respectively. These Nash-Sutcliffe efficiency values indicate very close agreement between measured annual drainage flow and predictions of annual drainage by the two models.

The two calibrated models were used to predict the long-term effects of implementing DWM across the U.S. Midwest on annual drainage and nitrogen loads from drained agricultural fields. Each of the two models simulated two scenarios, conventional drainage (not managed) and DWM, using 25 years of climatological records for 48 locations across the Midwest. Readers are referred to Thorp et al. (2009) for detailed descriptions of the simulated scenarios.

Compared to the DRAINMOD predictions, RZWQM predicted significantly larger reductions in annual drainage flow and N loads across the simulated 48 locations. On average, RZWQM predicted that implementing DWM would reduce annual drainage by 53% (range 35% to 68%) and reduce annual N loads by 51% (range 33% to 58%). On the other hand, DRAINMOD predicted that DWM, on average, would reduce annual drainage by 30% (range 19% to 45%) and reduce annual N loads by 32% (range 12% to 47%). For the purpose of this case study, we focus on the differences in the hydrologic predictions of the two models. As previously mentioned, the differences in model predictions of the annual water balance of the Iowa site during model calibration and validation (for conventional drainage) were very small and were mainly in predicting surface runoff and deep seepage. Analyzing the hydrologic results of the calibration and validation of the two models for the Iowa site, one could not foresee such a large difference in the water balance predictions of the two models in the DWM scenario analysis studies. Implementing DWM reduces flow via subsurface drains and raises the groundwater table. As a result, DWM is expected to increase evapotranspiration, surface runoff, and vertical seepage to the underlying aquifer. The relative significance of each pathway depends on site conditions.

Results of the DWM scenario analysis studies show that the average annual subsurface drainage flow predicted by the two models for the conventional (unmanaged) drainage scenario were similar (27.9 cm year^{-1} for DRAINMOD and 28.3 cm year^{-1} for RZWQM). However, the DRAINMOD and RZWQM predictions of annual subsurface drainage flow for the DWM scenario were considerably different. On average, DRAINMOD predicted annual subsurface drainage of 19.3 cm, compared to 13.3 cm predicted by RZWQM. RZWQM predicted more reduction in drainage flow associated with DWM than DRAINMOD mainly because RZWQM predicted more increase in surface runoff associated with DWM. On average, RZWQM predicted that DWM would increase surface runoff by 8.5 cm year^{-1} (from 2.6 to 11.1 cm year^{-1}). On the other hand, DRAIN-

MOD predicted that DWM would increase surface runoff by only 4.8 cm year^{-1} (from 3.9 to 8.7 cm year^{-1}). Despite the substantial difference in annual vertical seepage predicted by the two models for conventional drainage (4.0 cm year^{-1} for DRAINMOD and 12.0 cm year^{-1} for RZWQM), both models predicted that implementing DWM would increase vertical seepage for simulated soil conditions by only 1.0 to 2.0 cm year^{-1}.

The point of this case study is not to conclude that one model performed better than the other for this application. Rather, these results point out the importance of accurate simulation of all the processes impacting the water balance over the range of application of the model. While there was little difference in the values of all the hydrologic components predicted by the two models during the calibration period under conventional drainage, RZWQM predicted a substantially greater impact of DWM on subsurface drainage than did DRAINMOD. In this case, there was no opportunity to calibrate the models for DWM. Experimental data on the same site for this application would have likely have resulted in changes to the calibrated inputs for both models, with final predictions of the effects of DWM being more similar to each other and, more importantly, to the actual impacts. Realistically, it is often necessary to apply a model under conditions that do not allow calibration over the range of its specific application. This emphasizes the need for continued research to develop stronger, more reliable models and model components, as well as methods for determining field effective inputs and guidance for their reliable application.

WATERSHED-SCALE MODELING: NITROGEN BALANCE

The implementation of sophisticated watershed simulation models could incorporate a large number of model parameters to imitate real-world phenomena (Yen et al., 2014b; Vrugt et al., 2008; Yang et al., 2008). Therefore, various auto-calibration techniques have been developed to solve high-dimensional watershed calibration problems (Duan et al., 1992; Haario et al., 2006; Klepper and Hendrix, 1994; Tolson and Shoemaker, 2007; Vrugt et al., 2009; Boyle et al., 2003). The goal of auto-calibration is to estimate model parameters by minimizing the error statistics between observed and simulated data, such as stream gauge data, through mathematical processes. The commonly implemented statistical indices are organized as a reference of standards (Moriasi et al., 2007) so that users can have a moderately dependable guide in evaluating the performance of a specific set of model parameters. However, for watershed modeling, the main pitfall in using only stream gauge data at an individual point is that this approach does not account for processes within the catchment area. By only matching time-varying hydrologic or water quality responses at a single location, statistically well performed calibration results may not fully reflect actual watershed behavior with heterogeneous characteristics (Meixner et al., 2003).

Yen et al. (2014a) presented an excellent case study demonstrating the potential problems encountered when calibrating at a gauge and not accounting for processes within the watershed. SWAT was applied to the 248 km^2 Eagle Creek watershed in central Indiana, the source of drinking water for the Indianapolis metropolitan area. Two processes and their impact on calibration were examined: (1) the annual mass of denitrification and (2) the fraction of annual subsurface NO$_3$ loading at the watershed outlet.

To account for the nitrogen balance processes in the calibration procedure, Yen et al. (2014a) constrained the processes based on literature information (soft data). David et al. (2009) reported that denitrification rates in the Midwest U.S. are regularly less than 50 kg ha^{-1}. NO$_3$ loads from tile flow in the Raccoon River in Iowa were shown to contribute two-thirds of the total NO$_3$ load to the river (Schilling, 2002). Based on these studies, denitrification rates were constrained to 50 kg ha^{-1}, and tile flow NO$_3$ loads were constrained to be at least two-thirds of total NO$_3$. Auto-calibration of streamflow and nitrate loadings without constraints yielded Nash-Sutcliffe efficiencies from 0.84 to 0.95. However, denitrification in the basin was simulated as 257 kg ha^{-1}, and the ratio of NO$_3$ contributed by tile flow was only 13%. With constraints in place, the Nash-Sutcliffe efficiencies ranged from 0.66 to 0.94, with denitrification of 33 kg ha^{-1} and tile NO$_3$ contribution of 67%. The results showed that good statistical agreement with gauge data could be obtained without proper nitrogen balances and process representation. However, potential problems arise when management scenarios are simulated. Yen et al. (2014a) ran a no-till scenario on cultivated cropland throughout the watershed, which would increase ground cover, decreased erosion, and potentially enhanced infiltration. In the simulation without process constraints, with denitrification rates extremely high and tile flow unrealistically low, no-till decreased NO$_3$ loadings slightly by providing ground cover and reducing runoff and erosion. When denitrification and tile flow were constrained to realistic ranges, the no-till scenario caused an increase in tile flow and actually increased NO$_3$ loads at the watershed outlet. In this case, the unconstrained (incorrect) simulation was superior to the constrained (conceptually valid) simulation, using hard data statistics alone, further indicating the danger in maximizing singular statistical indicators of model performance.

WATERSHED-SCALE MODELING: SEDIMENT SOURCES AND SINKS

A case study was developed for this article to illustrate the importance of proper simulation of processes in the sediment budget for a watershed. Coon Creek watershed is located in southwest Wisconsin and drains into the Mississippi River basin. The Coon Creek watershed was established as one of the first demonstration watersheds during the formation of the Soil Erosion Service (Helm, 2009) in 1934. When conservation efforts began after 80 years of poor land management, severe gullying existed, and soils were depleted. Sediment from sheet, rill, and gully erosion exceeded stream transport capacity, and 2 m of deposition occurred in ten years in the main valleys. Pastures and woodlands were overgrazed, forming extensive woodland gullies. Small stream channels eroded, while extreme aggradation was occurring in main valleys. Present-day land management includes terracing, contour strip cropping,

minimum tillage, and cover crops. Pastures are well managed, and woodlands are not grazed. Many channels were established with streambank structures.

A watershed sediment budget is a quantitative assessment of the rates of erosion, transport, and deposition of sediment. This involves determining the temporal and spatial variations of transport and storage processes. Trimble (1999) estimated sediment budgets in Coon Creek during three periods: (1) introduction of agricultural characterized by total lack of conservation (1853-1938), (2) start of conservation efforts (1938-1975), and (3) continuation of conservation efforts (1975-present). Trimble's sediment budget diagram was published in *Science* in 1999, and modified in Trimble (2009), to include sediment sources and sinks in each period. Using the definitions of soft data of Seibert and McDonnell (2002), the sediment budget developed by Trimble would be considered soft data. Sediment degradation and deposition in valley bottoms and reservoirs were determined from dating core samples, erosion rates were extrapolated from nearby research data, and sediment transport from the basin was determined from limited hard data and the knowledge that the main channel was transport limited.

Since hard data (gauge data) were limited to four years of daily flow in the 1970s, flow parameters were calibrated by maximizing daily Nash-Sutcliffe efficiencies for the four-year period. Trimble spatially divided sediment sources and sinks into net upland sheet and rill erosion, upland gully erosion, tributaries, upland valleys, upper main valley, and lower main valley. In this study, spatial processes were simplified to include: (1) net upland sheet, rill, and gully erosion, and (2) net deposition in the tributaries and valleys. In the pre-conservation period of 1853-1938, Trimble (1999) estimated 11.1 t ha^{-1} year^{-1} from upland sheet, rill, and gully erosion, 10.0 t ha^{-1} year^{-1} net deposition in tributaries and valleys, and 1.1 t ha^{-1} year^{-1} sediment yield leaving the watershed.

To illustrate the need to realistically model the sediment budget during calibration, a SWAT simulation was developed using land management with no conservation, indicative of the 1853-1938 period. Climate data from 1960-2010 were used in all scenarios to remove the influence of climate variability and maintain focus solely on land management impacts. Soft data on sediment yield from the watershed were used to calibrate sediment parameters for the 1853-1938 (no conservation) period (scenario 1 in table 2). Next, the model was calibrated to account for erosion and deposition processes (scenario 2 in table 2). When calibrating only hard data at the watershed outlet, upland sources were underestimated by 32% and net deposition in the tributaries and valleys was underestimated by 36%. The real effect of improper process simulation became apparent

during scenario analysis. Two additional scenarios were performed by adding a full suite of conservation practices that are in place today, including terracing, contour strip cropping, minimum tillage, cover crops, no forest grazing, and streambank structures, to each of the two previously calibrated models. Scenario 3 consists of full practices using parameters from calibrating at the outlet only, while scenario 4 consists of full practices using parameters from calibrating erosion and deposition processes. The difference between the calibrated condition and respective conservation scenario for each of the two calibration strategies is minimal at the watershed outlet (0.7 vs. 0.8 t ha^{-1} year^{-1}) because Coon Creek is a transport-limited stream. However, the strategy of using only outlet data for calibration resulted in upland sources being underestimated by 60% and, instead of correctly predicting net deposition in the valley bottoms, net degradation was predicted (table 2).

RECOMMENDATIONS AND CONCLUSIONS

The objectives of this article were to: (1) synthesize processes considered and calibration techniques used to account for processes within the field or watershed for the models in the special collection (Moriasi et al., 2012), (2) summarize other relevant literature related to process representation and calibration, (3) demonstrate the importance of proper process representation and its impact on calibration/validation scenario analysis using case studies, and (4) provide recommendations for calibration/validation. The principal objective was to evaluate the impact of processes on calibration and scenario analysis. The models in the special collection were divided into three groups: (1) water, heat, and solute transport, (2) field scale, and (3) watershed scale, and the processes simulated in each model were synthesized.

Literature in the special collection revealed that point-scale and field-scale models calibrate to individual processes in the water, sediment, nutrient, and plant budgets but never to all major components. At the watershed scale, all procedures identified in the special collection calibrate to a time series of flow and constituent data at a stream gauge (hard data at a point) in the stream. The few examples existing in the general literature of process-based calibration using soft data at the watershed scale included: (1) calibrating both surface runoff and baseflow, (2) satellite-based ET estimates, (3) calibrating crop yields, (4) attempts at simulating the sediment budget, and (5) accounting for denitrification and nitrogen in tile flow (Yen et al., 2014a).

Case studies were developed for this article to illustrate the importance of (1) the water balance at the field scale and (2) sediment source and sink processes within a watershed modeling study. Both these examples and the example

Table 2. Land management and calibration scenarios for Coon Creek illustrating the impact of proper process representation.

Scenario	Calibration Method	Watershed Outlet (t ha^{-1})	Upland Sources Sheet Rill and Gully (t ha^{-1})	Net Deposition in Tributaries and Valleys (t ha^{-1})
1. No practice	Outlet only	1.1	7.5	6.4
2. No practice	Outlet and processes	1.1	11.1	10.0
3. Full practice	Parameters from scenario 1	0.8	0.5	-0.3
4. Full practice	Parameters from scenario 2	0.7	0.8	0.1
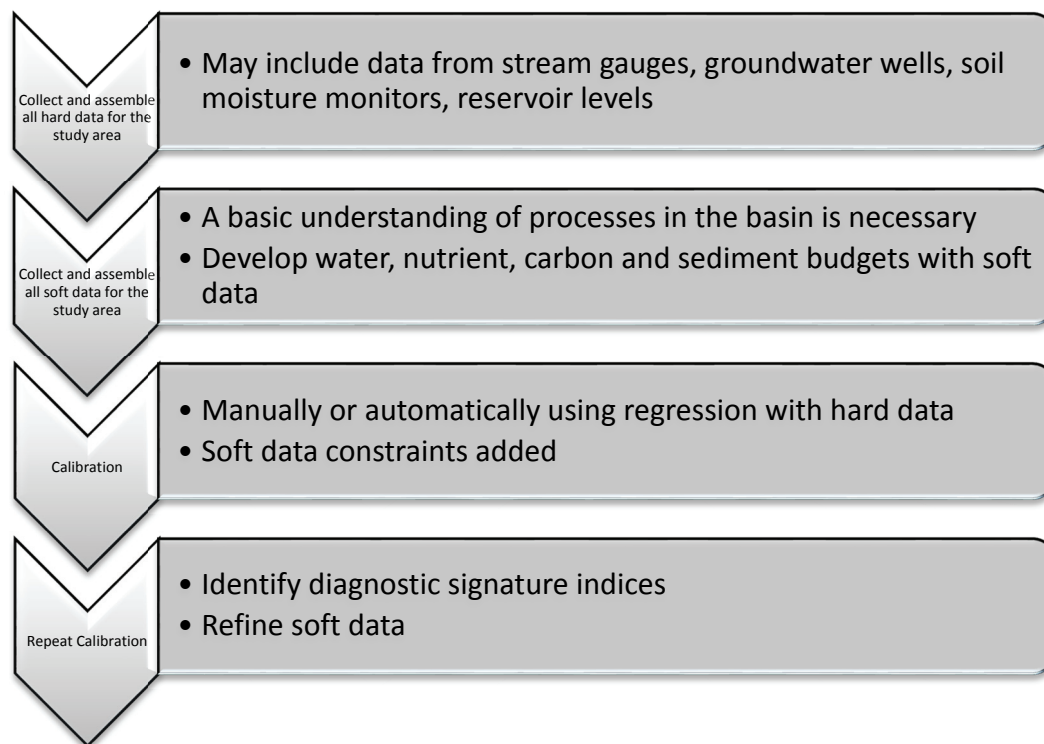
Figure 5. Recommended approach for inclusion of soft processes in calibration.

Flowchart steps:
- Collect and assemble all hard data for the study area
  - May include data from stream gauges, groundwater wells, soil moisture monitors, reservoir levels
- Collect and assemble all soft data for the study area
  - A basic understanding of processes in the basin is necessary
  - Develop water, nutrient, carbon and sediment budgets with soft data
- Calibration
  - Manually or automatically using regression with hard data
  - Soft data constraints added
- Repeat Calibration
  - Identify diagnostic signature indices
  - Refine soft data

of Yen et al. (2014a) demonstrated that a model could be calibrated to a gauge using proper auto-calibration methods and statistics. However, if the processes were not accurately represented, then land use and management practices and climate scenarios would not give meaningful results. Accurate scenario analysis to aid policy makers in making science-based decisions should be the major focus of any modeling study. To wholly focus model calibration at the outlet or a few locations to achieve optimal statistics and disregard internal model process is myopic, and often detrimental to the greater objective.

To account for processes during calibration, we recommend a diagnostic approach using both hard and soft data, as suggested by Yilmaz et al. (2008) and Siebert and McDonnell (2002). The diagnostic approach looks at signature patterns of behavior to determine which processes and thus parameters need further adjustment during calibration. This overcomes the weaknesses of traditional regression-based calibration by discriminating between multiple processes within a budget. The advantage of developing soft data for the calibration is that it (1) requires a basic understanding of processes (water, sediment, nutrient, and carbon budgets) within the spatial area being modeled and (2) constrains the calibration. The approach recommended here consists of four basic steps (fig. 5).

STEP 1. COLLECT AND ASSEMBLE ALL HARD DATA FOR THE STUDY AREA

Depending on the scale and project goals, this may include data from stream gauges, groundwater wells, soil moisture monitors, reservoir levels, and lysimetric, energy balance, and/or eddy flux measurements of ET.

STEP 2. COLLECT AND ASSEMBLE ALL SOFT DATA FOR THE STUDY AREA

Although long-term time series of all processes may not be known, it is important that model users have a basic understanding of processes within the basin. Soft data sources may include refereed literature; engineering, technical, and research reports; unpublished documents (theses and dissertations); and field surveys. Thus, it is important that budgets are developed with soft data to ensure that processes are realistically modeled.

Water Balance

Even at the field and point scales, measured time series data are rarely available for all processes in the water balance. Typically, runoff may be measured in point-scale and field-scale studies and leaching in column or lysimeter studies (point scale). If surface runoff and leaching are measured, the other dominant processes are lateral flow and ET. At the watershed scale, a common technique is to use digital filters on daily measured flow data to separate surface runoff and baseflow (Arnold et al., 1995). Although the technique does not define the baseflow ratio spatially, it does provide general bounds for the major components of the water balance within the watershed. In the future, remotely sensed data may provide spatial estimates of ET, average ET for various land uses and ecosystems (Sun et al., 2011), and groundwater heights that can be used directly in model calibration.

Nutrient and Carbon Balances

Nutrient and carbon balances are more difficult to estimate due to lack of measured data and understanding of several key processes. At the field scale, all processes are

never measured; however, for many research sites, nitrogen in runoff and tile flow, nitrogen removed in yield, and nitrogen measured in the soil profile are often collected. At the watershed scale, the best we can currently do is (1) remove as much uncertainty in inputs as possible (e.g., fertilizer applications and atmospheric deposition), (2) calibrate crop yields to reduce uncertainty in plant uptake and nutrient outputs, (3) calibrate to all available forms (organic and inorganic) of nitrogen and phosphorus measured at gauges in the watershed, and (4) constrain other processes (e.g., denitrification, mineralization, tile flow nitrates) based on similar research data. Additional difficulties at the watershed scale occur due to differences in land use and management and differences in landscape position within the watershed (Amatya et al., 2013).

Sediment Budget

As illustrated in the Coon Creek case study, proper simulation of sediment sources and sinks is crucial to scenario analysis and ultimately policy recommendations. It is highly recommended that a sediment budget is developed for watershed studies. A sediment budget is defined as an accounting of the sources and disposition of sediment as it travels from its point of origin to its eventual exit from the watershed (Reid and Dunn, 1996). There are numerous approaches to constructing sediment budgets; however, Reid and Dunn (1996) suggested a consistent series of steps for an approximate sediment budget:

1. Define the problem and determine the accuracy and spatial detail needed.
2. Acquire background information in similar ecoregions, including published literature, data and records from state and government agencies, academia, engineering and technical studies, and local theses

and dissertations. Reservoir sedimentation data are also useful in constructing sediment budgets (e.g., Moriasi et al., 2011). Lidar mapping or aerial photographs may also provide soft information on erosion rates and stream migration.

3. Subdivide the area into relatively uniform areas of soils, geology, vegetation, land use, and topography. This is often done in model parameterization; however, it may also be useful to look at USGS hydrologic landscape units (Winter, 2001).
4. Interpret aerial photographs to identify erosion and transport processes, measure or categorize process rates, and select sites for field work.
5. Conduct fieldwork. A field tour is used to identify active erosion and sediment transport processes, and guide aerial photographic interpretation. At this point, a flowchart could be developed of major sediment processes in the watershed (fig. 6).
6. Analyze the data.
7. Check results by comparing to other studies.

The typical time required to develop a sediment budget is approximately 2 to 30 person-days for field work and an additional 6 to 30 person-days of office time to analyze the budget (Reid and Dunn, 1996). With web-based aerial photographs now available, that time can be shortened significantly.

STEP 3. PERFORM CALIBRATION

Calibration can be performed manually or automatically using a regression-based approach with hard data and statistics, as suggested by Moriasi et al. (2007). Soft data constraints can then be added to ensure that processes are within reasonable limits (Yen et al., 2014a; Siebert and McDonnell, 2002).
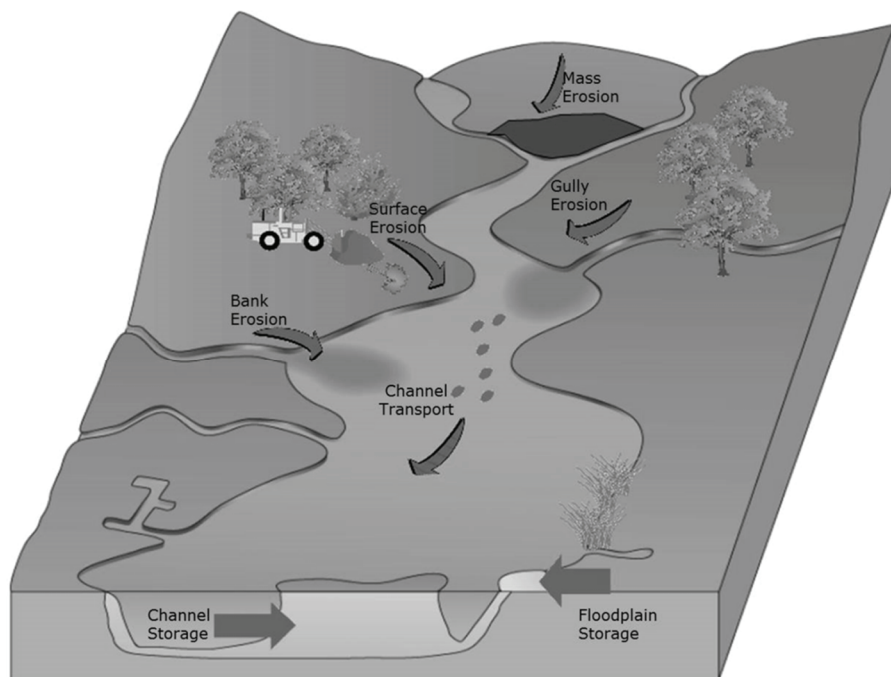
Figure 6. The erosion and sediment channel process (courtesy of the Integration and Application Network, University of Maryland Center for Environmental Science; http://ian.umces.edu/symbols).

STEP 4. IDENTIFY DIAGNOSTIC SIGNATURE INDICES, REFINE SOFT DATA, AND REPEAT CALIBRATION

A final recommendation is to build soft data processes into automated calibration procedures. As a first step, White et al. (2012) developed a screening tool called SWAT Check that assists model users in ensuring that processes are realistic and that water, sediment, and nutrient budgets are realistic. SWAT Check is a standalone program that (1) reads SWAT output and alerts users of values outside typical ranges, (2) creates process-based figures for visualizing water, sediment, and nutrient budgets, and (3) detects and alerts users of common model application errors. This software assists model users during calibration by ensuring that the model is realistically simulating the processes. The ultimate goal is to include these soft (process) data in automated calibration routines that are routinely used for calibration. This will ensure that land use and management scenario simulations provide meaningful results for environmental policy makers.

REFERENCES

Al-Qurashi, A., McIntyre, N., Wheater, H., & Unkrich, C. L. (2008). Application of the KINEROS2 rainfall-runoff model to an arid catchment in Oman. *J. Hydrol., 355*(1-4), 91-105. http://dx.doi.org/10.1016/j.jhydrol.2008.03.022.

Amatya, D. M., Skaggs, R. W., & Gregory, J. D. (1997). Evaluation of a watershed-scale forest hydrologic model. *J. Agric. Water Mgmt., 32*(3), 239-258. http://dx.doi.org/10.1016/S0378-3774(96)01274-7.

Amatya, D. M., Chescheir, G. M., Fernandez, G. P., Skaggs, R. W., & Gilliam, J. W. (2004). DRAINWAT-based methods for estimating nitrogen transport in poorly drained soils. *Trans. ASAE, 43*(3), 677-687. http://dx.doi.org/10.13031/2013.16100.

Amatya, D. M., Sun, G., Skaggs, R. W., Chescheir, G. M., & Nettles, J. E. (2006). Hydrologic effects of global climate change on a large drained pine forest. In T. Williams, & J. Nettles (Ed.), *Proc. Intl. Conf. Hydrology and Management of Forested Wetlands* (pp. 383-394). St. Joseph, Mich.: ASABE.

Amatya, D. M., Rossi, C. G., Saleh, A., Dai, Z., Youssef, M. A., Williams, R. G., Bosch, D. D., Chescheir, G. M., Sun, G., Skaggs, R. W., Trettin, C. C., Vance, E. D., Nettles, J. E., & Tian, S. (2013). Review of nitrogen fate models applicable to forest landscapes in the southern U.S. *Trans. ASABE, 56*(5), 1731-1757. http://dx.doi.org/10.13031/trans.56.10096.

Arnold, J. G., Allen, P. M., Muttiah, R. S., & Bernhardt, G. (1995). Automated base flow separation and recession analysis techniques. *Ground Water, 33*(6), 1010-1018.

Arnold, J. G., Muttiah, R. S., Srinivasan, R., & Allen, P. M. (2000). Regional estimation of base flow and groundwater recharge in the Upper Mississippi River basin. *J. Hydrol., 227*(1-4), 21-40. http://dx.doi.org/10.1016/S0022-1694(99)00139-0.

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.

Ávila-Carrasco, R., Mojarro Dávila, F., Moriasi, D. N., Gowda, P. H., Bautista-Capetillo, C., Echavarria-Cháirez, F. G., Garbrecht, J. D., Steiner, J. L., Howell, T. A., Kanemasu, E. T., Verser, J. A., Wagner K., & Hernandez, J. (2012). Calibration of SWAT2009 using crop biomass, evapotranspiration, and deep recharge: Calera watershed in Zacatecas, Mexico, case study. *J. Water Resource Prot., 4*(7), 439-450. http://dx.doi.org/10.4236/jwarp.2012.47051.

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.

Bottcher, A. B., Whiteley, B. J., James, A. I., & Hiscock, J. H. (2012). Watershed assessment model (WAM): Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1367-1383. http://dx.doi.org/10.13031/2013.42248.

Boyle, D. P., Gupta, H. V., & Sorooshian, S. (2003). Multicriteria calibration of hydrologic models. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, & R. Turcotte (Eds.), *Calibration of Watershed Models* (pp. 185-196). Washington, D.C.: American Geophysical Union. http://dx.doi.org/10.1029/WS006p0185.

Bulygina, N. S., Nearing, M. A., Stone, J. J., & Nichols, M. H. (2007). DWEPP: A dynamic soil erosion model based on WEPP source terms. *Earth Surface Proc. Landforms, 32*(7), 998-1012. http://dx.doi.org/10.1002/esp.1467.

Carsel, R., Mulkey, L. A., Lorber, M. N., & Baskin, L. B. (1985). The pesticide root zone model (PRZM): A procedure for evaluating pesticide movement to groundwater. *Ecol. Model., 30*(1-2), 49-69.

Causarano, H. J., Shaw, J. N., Franzluebbers, A. J., Reeves, D. W., Raper, R. L., Balkcom, K. S., Norfleet, M. L., & Izaurralde, R. C. (2007). Simulating field-scale soil organic carbon dynamics using EPIC. *SSSA J., 71*(4), 1174-1185. http://dx.doi.org/10.2136/sssaj2006.0356.

Challinor, A. J., Wheeler, T. R., Slingo, J. M., Craufurd, P. Q., & Grimes, D. I. (2004). Design and optimisation of a large-area process-based model for annual crops. *Agric. Forest Meteorol., 124*(1-2), 99-120. http://dx.doi.org/10.1016/j.agrformet.2004.01.002.

Chapwanya, M., & Stockie, J. M. (2010). Numerical simulations of gravity-driven fingering in unsaturated porous media using a nonequilibrium model. *Water Resources Res., 46*(9). http://dx.doi.org/10.1029/2009WR008583.

Cui, J., Li, C., Sun, G., & Trettin, C. (2005). Linkage of MIKE SHE to Wetland-DNDC for carbon budgeting and anaerobic biogeochemistry simulation. *Biogeochem., 72*(2), 147-167. http://dx.doi.org/10.1007/s10533-004-0367-8.

Dai, Z., Li, C., Trettin, C. C., Sun, G., Amatya, D. M., & Li, H. (2010). Bi-criteria evaluation of the MIKE SHE model for a forested watershed on the South Carolina coastal plain. *Hydrol. Earth Sys. Sci., 14*(6), 1033-1046. http://dx.doi.org/10.5194/hess-14-1033-2010.

Dai, Z., Trettin, C. C., Li, C., Li, H., Sun, G., & Amatya, D. M. (2011). Effect of assessment scale on spatial and temporal variations in CH_4, CO_2, and N_2O fluxes in a forested wetland. *Water, Air, Soil Poll., 223*(1), 253-265. http://dx.doi.org/10.1007/s11270-011-0855-0.

Daisy. (2011). Daisy model: A soil-plant atmosphere system model. Copenhagen, Denmark: University of Copenhagen, Department of Basic Sciences and Environment. Retrieved from http://code.google.com/p/daisy-model.

Darcy, H. (1856). *Les Fontaines Publiques de la Ville de Dijon*. Paris, France: Dalmont.

David, M. B., Del Grosso, S. J., Hu, X., Marshall, E. P., McIsaac, G. F., Parton, W. J., Tonitto, C., & Youssef, M. (2009). Modeling denitrification in a tile-drained, corn and soybean agroecosystem of Illinois, USA. *Biogeochem., 93*(1-2), 7-30. http://dx.doi.org/10.1007/s10533-008-9273-9.

DHI. (2012). MIKE BASIN: An integrated river basin planning and management tool. Horsholm, Denmark: Danish Hydraulic Institute, DHI Water and Environment.

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Res., 28*(4), 1015-1031. http://dx.doi.org/10.1029/91WR02985.

Duda, P. B., Hummel, P. R., Donigian 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.

Egorov, A., Dautov, R., Nieber, J. L., & Sheshukov, A. (2003). Stability analysis of gravity-driven infiltrating flow. *Water Resources Res., 39*(9). http://dx.doi.org/10.1029/2002WR001886.

Fernandez, G. P., Chescheir, G. M., Amatya, D. M., & Skaggs, R. W. (2002). WATGIS: A GIS-based lumped parameter water quality model. *Trans. ASAE, 45*(3), 593-600. http://dx.doi.org/10.13031/2013.8822.

Fernandez, G. P., Chescheir, G. M., Skaggs, R. W., & Amatya, D. M. (2005). Development and testing of watershed-scale models for poorly drained soils. *Trans. ASAE, 48*(2), 639-652. http://dx.doi.org/10.13031/2013.18323.

Fernandez, G. P., Chescheir, G. M., Skaggs, R. W., & Amatya, D. M. (2006). DRAINMOD-GIS: A lumped parameter watershed-scale drainage and water quality model. *Agric. Water Mgmt., 81*(1-2), 77-97. http://dx.doi.org/10.1016/j.agwat.2005.03.004.

Fernandez, G. P., Chescheir, G. M., Skaggs, R. W., & Amatya, D. M. (2007). Application of DRAINMOD-GIS to a lower coastal plain watershed. *Trans. ASABE, 50*(2), 439-447. http://dx.doi.org/10.13031/2013.22635.

Fick, A. (1855). Ueber diffusion (in German). *Ann. der Physik, 170*(1), 59-86. http://dx.doi.org/10.1002/andp.18551700105.

Finsterle, S., Kowalsky, M. B., & Pruess, K. (2012). TOUGH: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1275-1290. http://dx.doi.org/10.13031/2013.42240.

Flanagan, D. C., Frankenberger, J. R., & Ascough II, J. C. (2012). WEPP: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1463-1477. http://dx.doi.org/10.13031/2013.42254.

Flerchinger, G. N., & Saxton, K. E. (1989). Simultaneous heat and water model of a freezing snow-residue-soil system: I. Theory and development. *Trans. ASABE, 32*(2): 565-571. http://dx.doi.org/10.13031/2013.31040.

Flerchinger, G. N., Caldwell, T. G., Cho, J., & Hardegree, S. P. (2012). Simultaneous Heat and Water (SHAW) Model: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1395-1411. http://dx.doi.org/10.13031/2013.42250.

Franks, S. W., & Beven, K. J. (1997). Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. *J. Geophys. Res. Atmos., 102*(D20), 23991-23999. http://dx.doi.org/10.1029/97JD02011.

Franks, S. W., Gineste, P., Beven, K. J., & Merot, P. (1998). On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Res., 34*(4), 787-797.

Goodrich, D. C., Burns, I. S., Unkrich, C. L., Semmens, D. J., Guertin, D. P., Hernandez, M., Yatheendradas, S., Kennedy, J. R., & Levick, L. R. (2012). KINEROS2/AGWA: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1561-1574. http://dx.doi.org/10.13031/2013.42264.

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.

Gregersen, H. M., Ffolliott, P. F., & Brooks, K. N. (2007). Integrated watershed management: Connecting people to their land and water. Wallingford, U.K.: CAB International.

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: Elements of a diagnostic approach to model evaluation. *Hydrol. Proc., 22*(18), 3802-3813. http://dx.doi.org/10.1002/hyp.6989.

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Stat. Computing, 16*(4), 339-354. http://dx.doi.org/10.1007/s11222-006-9438-0.

Hansen, S., Abrahamsen, P., Petersen, C. T., & Styczen, M. (2012). DAISY: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1315-1333. http://dx.doi.org/10.13031/2013.42244.

Harmel, R. D., Smith, P. K., Migliaccio, K. W., Chaubey, I., Douglas-Mankin, K. R., Benham, B., Shukla, S., Munoz-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.

He, X., Izaurralde, R. C., Vanotti, M. B., Williams, J. R., & Thomson, A. M. (2006). Simulating long-term and residual effects of nitrogen fertilization on corn yields, soil carbon sequestration, and soil nitrogen dynamics. *J. Environ. Qual., 35*(4), 1608-1619. http://dx.doi.org/10.2134/jeq2005.0259.

Healy, R. W., & Essaid, H. I. (2012). VS2DI: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1249-1260. http://dx.doi.org/10.13031/2013.42238.

Helm, D. (2009). Hugh Hammond Bennett and the creation of the Soil Conservation Service. *J. Soil Water Cons., 64*(2), 68A-74A. http://dx.doi.org/10.2489/jswc.64.2.68A.

Huth, N. I., Bristow, K. L., & Verburg, K. (2012). SWIM3: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1303-1313. http://dx.doi.org/10.13031/2013.42243.

Hymer, D. C., Moran, M. S., & Keefer, T. O. (2000). Soil moisture evaluation using a hydrologic model and calibrated sensor network. *SSSA J., 64*(1), 319-326. http://dx.doi.org/10.2136/sssaj2000.641319x.

Immerzeel, W. W., & Droogers, P. (2008). Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol., 349*(3-4), 411-424. http://dx.doi.org/10.1016/j.jhydrol.2007.11.017.

Jaber, F. H., & Shukla, S. (2012). MIKESHE: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1479-1489. http://dx.doi.org/10.13031/2013.42255.

Jansson, P.-E. (2012). MIKESHE: COUPModel: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1337-1346. http://dx.doi.org/10.13031/2013.42245.

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.

Jayakrishnan, R., Srinivsan, R., Santhi, C., & Arnold, J. G. (2005). Advances in the application of the SWAT model for water resources management. *Hydrol. Proc., 19*(3), 749-762.

Kim, H. W., Amatya, D. M., Chescheir, G. M., Skaggs, R. W., & Nettles, J. E. (2013). Hydrologic effects of size and location of fields converted from drained pine forest to agricultural cropland. *J. Hydrol. Eng., 18*(5), 552-566. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000566.

Kiniry, J. R., Williams, J. R., Gassman, P. W., & Debaeke, P. (1992). A general, process-oriented model for two competing plant species. *Trans. ASAE, 35*(3), 801-810. http://dx.doi.org/10.13031/2013.28665.

Klepper, O., & Hendrix, E. (1994). A method for robust calibration of ecological models under different types of uncertainty. *Ecol. Model., 74*(3-4), 161-82. http://dx.doi.org/10.1016/0304-3800(94)90118-X.

Knisel, W. G., & Douglas-Mankin, K. R. (2012). CREAMS/GLEAMS: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1291-1302. http://dx.doi.org/10.13031/2013.42241.

Konyha, K. D., & Skaggs, R. W. (1992). A coupled, field hydrology - open channel flow model: Theory. *Trans. ASAE, 35*(5), 1431-1440. http://dx.doi.org/10.13031/2013.28750.

Ma, L., Ahuja, L. R., Nola, B. T., Malone, R. W., Trout, T. J., & Qi, Z. (2012). Root Zone Water Quality Model (RZWQM2): Model use, calibration, and validation. *Trans. ASABE, 54*(5), 1425-1446. http://dx.doi.org/10.13031/2013.42252.

Meixner, T., Gupta, H. V., Bastidas, L. A., & Bales, R. C. (2003). Estimating parameters and structure of a hydrologic model using multiple criteria. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, & R. Turcotte (Eds.), *Calibration of Watershed Models* (pp. 213-228). Washington, D.C.: American Geophysical Union. http://dx.doi.org/10.1029/WS006p0213.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. (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., Steiner, J. L., & Arnold, J. G. (2011). Sediment measurement and transport modeling: Impact of riparian and filter strip buffers. *J. Environ. Qual., 40*(3), 807-814. http://dx.doi.org/10.2134/jeq2010.0066.

Moriasi, D. N., Rossi, C. G., Arnold, J. G., & Tommer, M. D. (2012). Evaluating hydrology of the soil and water assessment tool (SWAT) with new tile drain equations. *J. Soil Water Cons., 67*(6), 513-524. http://dx.doi.org/10.2489/jswc.67.6.513.

Mullins, J. A., Carsel, R. F., Scarbrough, J. E., & Ivery, A. M. (1993). PRZM-2: A model for predicting pesticide fate in crop root and unsaturated soil zones. User manual for release 2.0. EPA/600/R-93/046. Athens, Ga.: U.S. EPA, Environmental Research Laboratory.

Nair, S. S., King, K. W., Witter, J. D., Sohngen, B. L., & Fausey, N. R. (2011). Importance of crop yield in calibrating watershed water quality simulation tools. *JAWRA, 47*(6), 1285-1297. http://dx.doi.org/10.1111/j.1752-1688.2011.00570.x.

Negm, L. M., Youssef, M. A., Skaggs, R. W., Chescheir, G. M., & Jones, J. (2014). DRAINMOD-DSSAT model for simulating hydrology, soil carbon and nitrogen dynamics, and crop growth for drained crop land. *Agric. Water Mgmt., 137*, 30-45.

Nieber, J. L., Dautov, R., Egorov, A., & Sheshukov, A. (2005). Dynamic capillary pressure mechanism for gravity-driven flows: Review and extension to dry conditions. *Transport in Porous Media, 58*(1-2), 147-172. http://dx.doi.org/10.1007/s11242-004-5473-5.

Perrin, C., Andreassian, V., Serna, C. R., Mathevet, T., & Moine, N. L. (2008). Discrete parameterization of hydrologic models: Evaluating the use of parameter sets libraries over 900 catchments. *Water Resources Res., 44*(8). http://dx.doi.org/10.1029/2007WR006579.

Philip, J. R. (1969). Theory of infiltration. *Adv. Hydrosci., 5*, 215-296. http://dx.doi.org/10.1016/B978-1-4831-9936-8.50010-6.

Philip, J. R. (1972). Flows satisfying mixed no-slip and no-shear conditions. *Zeitschrift für Angewandte Math. Physik ZAMP, 23*(3), 353-372.

Philip, J. R. (1992). What happens near a quasi-linear point source? *Water Resources Res., 28*(1), 47-52. http://dx.doi.org/10.1029/91WR02600.

Refsgaard, J. C., & Hansen, J. R. (2010). A good-looking catchment can turn into a modeller's nightmare. *Hydrol. Sci. J., 55*(6), 899-912.

Reid, L. M., & Dunne, T. (1996). *Rapid Evaluation of Sediment Budgets*. Reiskirchen, Germany: Catena Verlag.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *J. Appl. Physics, 1*(5), 318-333.

Sadeghi, A. M., Isensee, A. R., & Shirmohammadi, A. (1995). Atrazine movement in soil: Comparison of field observation and PRZM simulations. *J. Soil Contam., 4*(2), 151-161. http://dx.doi.org/10.1080/15320389509383489.

Santhi, C., Kannan, N., Arnold., J. G., Di Luzio, M., & Wang, S. (2008). Spatial calibration and temporal validation of flow for regional-scale hydrologic modeling. *JAWRA, 44*(4), 829-846. http://dx.doi.org/10.1111/j.1752-1688.2008.00207.x.

Schilling, K. E. (2002). Chemical transport from paired agricultural and restored prairie watersheds. *J. Environ. Qual., 31*(4), 1846-1851. http://dx.doi.org/10.2134/jeq2002.1184.

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, & R. Turcotte (Eds.), *Calibration of Watershed Models* (Vol. 6). Washington, D.C.: American Geophysical Union. http://dx.doi.org/10.1029/006WS22.

Shrestha, R. R., & Rode, M. (2008). Multi-objective calibration and fuzzy preference selection of a distributed hydrological model. *Environ. Model. Software, 23*(12), 1384-1395. http://dx.doi.org/10.1016/j.envsoft.2008.04.001.

Šimůnek, J., Šejna, M., & van Genuchten, M. T. (1998). The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably saturated media, version 2.0. IGWMC TPS-70. Golden, Colo.: Colorado School of Mines, Integrated Ground Water Modeling Center.

Šimůnek, J., van Genuchten, M. T., Šejna, M., Toride, N., & Leij, F. J. (1999). The STANMOD computer software for evaluating solute transport in porous media using analytical solutions of convection-dispersion equation, versions 1.0 and 2.0. IGWMC TPS-71. Golden, Colo.: Colorado School of Mines, Integrated Ground Water Modeling Center.

Šimůnek, J., van Genuchten, 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. M. (2012). DRAINMOD: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1509-1522. http://dx.doi.org/10.13031/2013.42259.

Sun, G., Alstad, K., Chen, J., Chen, S., Ford, C. R., Lin, G., Lu, N., McNulty, S. G., Noormets, A., Vose, J. M., Wilske, B., Zeppel, M., Zhang, Y., & Zhang, Z. (2011). A general predictive model for estimating monthly ecosystem evapotranspiration. *Ecohydrol., 4*(2), 245-255. http://dx.doi.org/10.1002/eco.194.

Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophys. Res. Letters, 31*(9), L09607. http://dx.doi.org/10.1029/2004GL019920.

Thompson, J. R., Sorenson, H. R., Gavin, H., & Refsgaard, A. (2004). Application of the coupled MIKE SHE / MIKE 11 modelling system to a lowland wet grassland in southeast England. *J. Hydrol., 293*(1-4), 151-179. http://dx.doi.org/10.1016/j.jhydrol.2004.01.017.

Thompson, J. R., Gavin, H., Refsgaard, A., Sorenson, H. R., & Gowing, D. J. (2009). Modeling the hydrological impacts of climate change on U.K. lowland wet grassland. *Wetlands Ecol. Mgmt., 17*(5), 503-523. http://dx.doi.org/10.1007/s11273-008-9127-1.

Thoms, R. B., Johnson, R. L., & Healy, R. W. (2006). User's guide to the variably saturated flow (VSF) process for MODFLOW. USGS Techniques and Methods 6-A18. Reston, Va.: U.S. Geological Survey.

Thorp, K. R., Jaynes, D. B., & Malone, R. W. (2007). Simulating the long-term performance of drainage water management across the Midwestern United States. *Trans. ASABE, 51*(3), 961-976. http://dx.doi.org/10.13031/2013.24534.

Thorp, K. R., Youssef, M. A., Jaynes, D. B., Malone, R. W., & Ma,

L. (2009). DRAINMOD-N II: Evaluated for an agricultural system in Iowa and compared to RZWQM-DSSAT. *Trans. ASABE, 52*(5), 1557-1573. http://dx.doi.org/10.13031/2013.29144.

Tian, S., Youssef, M. A., Skaggs, R. W., Amatya, D. M., & Chescheir, G. M. (2012a). DRAINMOD-FOREST: Integrated modeling of hydrology, soil carbon and nitrogen dynamics, and plant growth for drained forests. *J. Environ. Qual., 41*(3), 764-782. http://dx.doi.org/10.2134/jeq2011.0388.

Tian, S., Youssef, M. A., Skaggs, R. W., Amatya, D. M., & Chescheir, G. M. (2012b). Modeling water, carbon, and nitrogen dynamics for two drained pine plantations under intensive management practices. *Forest Ecol. Mgmt., 264*, 20-36. http://dx.doi.org/10.1016/j.foreco.2011.09.041.

Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Res., 43*(1), 1-16. http://dx.doi.org/10.1029/2005WR004723.

Trimble, S. W. (1999). Rates of soil erosion. *Science, 286*(5444), 1477-1478.

Trimble, S. W. (2009). Fluvial processes, morphology, and sediment budgets for the Coon Creek basin, Wisconsin, USA, 1975-1993. *Geomorphology, 108*(1-2), 8-23.

Vaché, K. B., & McDonnell, J. J. (2006). A process-based rejectionist framework for evaluating catchment runoff model structure. *Water Resources Res., 42*(2). http://dx.doi.org/10.1029/2005WR004247.

van Dam, J. C., Huygen, J., Wesseling, J. G., Feddes, R. A., Kabat, P., van Valsum, P. E. V., Groenendijk, P., & van Diepen, C. A. (1997). Theory of SWAP, Version 2.0: Simulation of water flow, solute transport, and plant growth in the soil-water-atmosphere-plant environment. WAU, Report 71, Technical Document 45. Wageningen, The Netherlands: DLO Winand Staring Centre, Department of Water Resources.

van Genuchten, M. T., Šimůnek, J., Leij, F. J., Toride, N., & Šejna, M. (2012). STANMOD: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1355-1368. http://dx.doi.org/10.13031/2013.42247.

Vasquez, R. F., & Feyen, J. (2010). Rainfall-runoff modelling of a rocky catchment with limited data availability: Defining prediction limits. *J. Hydrol., 387*(1-2), 128-140. http://dx.doi.org/10.1016/j.jhydrol.2010.04.004.

Vrugt, J. A., ter Braak, C. J., 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), 1-15. http://dx.doi.org/10.1029/2007WR006720.

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Intl. J. Nonlinear Sci. Num. Sim., 10*(3), 271-288.

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.

Wallenstein, M., Stromberger, M., & Bell, C. (2012). Bridging the gap between modelers and experimentalists. *EOS, 93*(32), 312. http://dx.doi.org/10.1029/2012EO320005.

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.

Williams, J. R., Jones, C. A., Kiniry, J. R., & Spanel, D. A. (1989). The EPIC crop growth model. *Trans. ASAE, 32*(2), 497-511. http://dx.doi.org/10.13031/2013.31032.

Williams, J. R., Arnold, J. G., Kiniry, J. R., Gassman, P. W., & Green, C. H. (2008). History of model development at Temple, Texas. *Hydrol. Sci. J., 53*(5), 948-960. http://dx.doi.org/10.1623/hysj.53.5.948.

Winsemius, H. C., Schaefli, B., Montanari, A., & Savenije, H. H. G. (2009). On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Res., 45*(12). http://dx.doi.org/10.1029/2009WR007706.

Winter, T. C. (2001). The concept of hydrologic landscapes. *JAWRA, 37*(2), 335-349. http://dx.doi.org/10.1111/j.1752-1688.2001.tb00973.x.

Wooding, R. A. (1968). Steady infiltration from a large shallow circular pond. *Water Resources Res., 4*(6), 1259-1273. http://dx.doi.org/10.1029/WR004i006p01259.

Workman, S. R., & Skaggs, R. W. (1989). Comparison of two drainage simulation models using field data. *Trans. ASAE, 32*(6), 1933-1938. http://dx.doi.org/10.13031/2013.31245.

Yang, J., Reichert, P., Abbaspour, K., 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.

Yen, H., Bailey, R. T., Arabi, M., Ahmadi, M., White, M. J., & Arnold, J. G. (2014a). 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.

Yen, H., Wang, X., Fontane, D. G., Arabi, M., & Harmel, R. D. (2014b). A framework for propagation of uncertainty contributed by input data, parameterization, model structure, and calibration/validation data in watershed modeling. *Environ. Model. Software, 54*, 211-221. http://dx.doi.org/10.1016/j.envsoft.2014.01.004.

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.

Youssef, M. A., Skaggs, R. W., Chescheir, G. M., & Gilliam, J. W. (2005). The nitrogen simulation model, DRAINMOD-NII. *Trans. ASAE, 48*(2), 611-626. http://dx.doi.org/10.13031/2013.18335.

Zelelew, M. B., & Alfredson, K. (2012). Sensitivity-guided evaluation of the HBV hydrological model parameterization. *J. Hydroinformatics, 15*(3), 967-990. http://dx.doi.org/10.2166/hydro.2012.011.

Zhang, X., Srinivasan, R., Arnold, J. G., Izaurralde, R. C., & Bosch, D. (2011). Simultaneous calibration of surface flow and baseflow simulations: A revisit of the SWAT model calibration framework. *Hydrol. Proc., 25*(14), 2313-2320. http://dx.doi.org/10.1002/hyp.8058.

Zheng, C., Hill, M. C., Cao, C., & Ma, R. (2012). MT3DMS: Model use, calibration, and validation. *Trans. ASABE, 55*(4), 1549-1559. http://dx.doi.org/10.13031/2013.42263.