Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development

Fabrizio Fenicia,[1,2] Dmitri Kavetski,[3] and Hubert H. G. Savenije[2]

Received 27 October 2010; revised 19 September 2011; accepted 19 September 2011; published 17 November 2011.

[1] This paper introduces a flexible framework for conceptual hydrological modeling, with two related objectives: (1) generalize and systematize the currently fragmented field of conceptual models and (2) provide a robust platform for understanding and modeling hydrological systems. In contrast to currently dominant "fixed" model applications, the flexible framework proposed here allows the hydrologist to hypothesize, build, and test different model structures using combinations of generic components. This is particularly useful for conceptual modeling at the catchment scale, where limitations in process understanding and data availability remain major research and operational challenges. The formulation of the model architecture and individual components to represent distinct aspects of catchment-scale function, such as storage, release, and transmission of water, is discussed. Several numerical strategies for implementing the model equations within a computationally robust framework are also presented. In the companion paper, the potential of the flexible framework is examined with respect to supporting more systematic and stringent hypothesis testing, for characterizing catchment diversity, and, more generally, for aiding progress toward more unified hydrological theory at the catchment scale.

Citation: Fenicia, F., D. Kavetski, and H. H. G. Savenije (2011), Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, *Water Resour. Res., 47*, W11510, doi:10.1029/2010WR010174.

1. Introduction

1.1. Modeling Paradigms in Hydrology: Fixed Versus Flexible Model Structures

[2] A major challenge in hydrology and broader environmental studies is the development of more scientifically meaningful and operationally reliable models [e.g., *Beven*, 2001; *Singh and Woolhiser*, 2002; *Sivapalan*, 2009]. Many distinct model-building paradigms have been proposed, ranging from simple bucket models [*Sugawara*, 1995] to conceptual models such as TOPMODEL [*Beven*, 1997], VIC [*Wood et al.*, 1992], and HBV [*Lindström et al.*, 1997] to physically based distributed models such as SHE [*Abbot et al.*, 1986]. Approaches such as transfer function modeling [e.g., *Young*, 1998] and neural networks [e.g., *Kingston et al.*, 2005] have also been used. The current lack of unified theories of hydrology at the catchment scale has been noted by several commentators [e.g., *Sivapalan*, 2005; *McDonnell et al.*, 2007; *Troch et al.*, 2009; *Clark et al.*, 2011b]. This paper focuses on conceptual hydrological models, which seek to represent dominant catchment dynamics in a physically meaningful way while remaining parametrically parsimonious and computationally efficient. They are also increasingly forming the basis for semidistributed environmental models [e.g., *Ajami et al.*, 2004; *Uhlenbrook et al.*, 2004].

1.1.1. Fixed Model Structures

[3] Perhaps paradoxically, despite the proliferation of hydrological models (e.g., see collections in the work by *Singh and Woolhiser* [2002]), it could be argued that hydrological model development has been largely guided by a "one model fits all" paradigm, effectively seeking a single general model applicable to every catchment. For example, while acknowledging the diversity of environmental processes, *Linsley* [1982, p. 14] suggests that "these differences do not mean that a single model cannot be applied in all cases." A typical example of the development of a fixed model structure is the GR4J model [*Perrin et al.*, 2003]. Over a series of case studies, some involving hundreds of catchments [e.g., *Edijatno et al.*, 1999; *Perrin et al.*, 2001], this four-parameter model has been refined to provide, on average, a better performance than the considered alternatives. GR4J is now widely used in hydrological research [e.g., *Le Moine et al.*, 2007; *Thyer et al.*, 2009], and a modified version, GRP, is used by the French flood forecasting services [*Berthet et al.*, 2009].

[4] In principle, a fixed model structure has several appealing features. Repeated use of the same model facilitates the process of model application and improvement [*Linsley*, 1982], simplifies training, and prevents potential confusion arising because of differences in the behavior of different models. Large-scale applications of the same model to different catchments make it easier to identify

[1]Department of Environment and Agro-Biotechnologies, Centre de Recherche Public–Gabriel Lippmann, Belvaux, Luxembourg.
[2]Water Resources Section, Delft University of Technology, Delft, Netherlands.
[3]Environmental Engineering, University of Newcastle, Callaghan, New South Wales, Australia.

0043-1397/11/2010WR010174

and, possibly, interpret the dependencies of its parameters on catchment properties, benefitting model interpretation and regionalization. From an operational perspective, the use of a fixed model structure reduces effort and training requirements for practitioners and researchers [*Le Moine et al.*, 2007], standardizes model applications by different users, etc.

[5] On the other hand, a strong case can be made for a flexible modeling paradigm, which allows the hydrologist to adapt the model to the catchment of interest. While *Linsley* [1982, p. 14] suggests that "it seems axiomatic that the *fundamental* processes of hydrology are the same in all catchments," (italics added), the search for a single "universal" model structure *at a coarse lumped scale* seems difficult to reconcile with the subscale heterogeneity and variability of catchment systems [*McDonnell et al.*, 2007], the related perception of "uniqueness of place" [*Beven*, 2000], the requirement to "work for the right reasons" [*Kirchner*, 2006], and the impact of type, quality, and resolution of available data on identifiable complexity [e.g., *Jakeman and Hornberger*, 1993; *Schoups et al.*, 2008; *Kavetski et al.*, 2011]. From a more practical perspective, flexible models also allow adapting the model structure to the objectives of a particular application, such as the prediction of low versus high flows, forecasting at a particular lead time, etc. While such dependence of model structure on modeling objectives is arguably unsatisfactory from a physical point of view, it is often the case in practice, especially in engineering contexts [e.g., *Butts et al.*, 2004; *Refsgaard and Henriksen*, 2004].

[6] A simple illustration of the practical limitations of a fixed model structure is the need to add specialized modules for specific catchment conditions. For example, in many models, the simulation of snowmelt requires the addition of an external snow module, already implying that the overall model structure requires customization for a specific climatic region. For highly urbanized catchments, it may be necessary to add a representation of impervious areas [*Cuo et al.*, 2008]; models for catchments with significant groundwater exchange need corresponding external flux terms [*Le Moine et al.*, 2007] and so forth.

[7] Although a model could be designed such that certain components could be "turned off" a priori, more subtle adjustments are often suggested by the data. For example, TOPMODEL applications suggest exponential transmissivity profiles in some catchments versus parabolic and even linear profiles in others [*Ambroise et al.*, 1996]. In the dynamic TOPMODEL [*Beven and Freer*, 2001], the assumption of a quasi-steady state saturated zone is replaced by kinematic wave routing. Similarly, while many models assume the base flow–storage relation is linear (e.g., HBV [*Lindström et al.*, 1997]) or a power law (e.g., TOPMODEL [*Beven*, 1997] and GR4J [*Perrin et al.*, 2003]), recession analysis often suggests different shapes [*Lamb and Beven*, 1997; *Rupp and Woods*, 2008]. For example, the base flow response of the Maimai basin in New Zealand appears piecewise linear [*Fenicia et al.*, 2010]. It is dubious to disregard such direct independent insights, yet attempting to develop a single constitutive function that includes linear, exponential, sigmoid, and other shapes is parametrically cumbersome and, in many cases, will result in insensitive and nonidentifiable parameters. Moreover, *Uhlenbrook et al.* [1999] report a case study where several variants of the HBV model were examined, differing both in the overall connectivity of its comprising reservoirs and in the structure of individual reservoirs. Overall, given an arguably poor current understanding of hydrology at the catchment scale [e.g., *Sivapalan*, 2005; *McDonnell et al.*, 2007; *Clark et al.*, 2011b] and the effect of large observational data errors on model identification [e.g., *Beven*, 2008; *Renard et al.*, 2010], attempting to limit the analyses to a single lumped model structure could be currently premature and may overlook better alternative representations.

[8] The proliferation of hydrological models, differing in all aspects of their process conceptualization, spatial distribution, number of states and parameters, etc, is perhaps in itself an empirical manifestation that the "one model fits all" ideal may be unattainable, at least given current hydrological data. For example, TOPMODEL and SHE are frequently used in many European countries, NWS and MMS are standard in various U.S. agencies, ARNO/TOPKAPI is adopted in Italy, the Xinanjiang is used in China, and so forth [*Singh and Woolhiser*, 2002]. Even for the HBV model popular in Scandinavia, *Uhlenbrook et al.* [1999] detail Swiss, Swedish, and "new" variants.

[9] Given these considerations, at least in the context of rainfall-runoff modeling, we do not see a strong a priori reason why the dynamics of catchments that are widely different in their geomorphology and climatology should always reduce to a single common form at coarsely lumped scales. Should we then be surprised at the empirical difficulties in obtaining spatially transposable models [e.g., *Andréassian et al.*, 2009], or that some compromises are needed when deciding on a single "best" model structure for thousands of catchments [*Perrin et al.*, 2003]? On the other hand, temporal transposability is generally a necessity if the model is to be used for predictive purposes, and achieving this is not unprecedented in practice (although factors such as land use change may need to be included in the modeling process [e.g., *Croke et al.*, 2004]). In our opinion, it is not unreasonable to hypothesize, on average, less difference in behavior for the same catchment at different times than for different catchments altogether (especially across differences in size, climate, and physical attributes).

[10] Finally, consider model comparison studies, such as the recent Model Parameter Estimation Experiment (MOPEX) [*Duan et al.*, 2006], the earlier Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) [*Henderson-Sellers et al.*, 1993], and the comparison of 19 model structures on 429 catchments performed by *Perrin et al.* [2001]. Such experiments are valuable hypothesis-testing exercises, providing empirical indications of best performing models and thus potentially guiding model improvement. However, such comparative studies have arguably been quite rare, and in many cases their findings may have been confounded by multiple unaccounted differences and interactions between model components, numerical implementation aspects, software settings, calibrations schemes, etc. [*Clark et al.*, 2011b]. This stresses the need for a more systematic pursuit of hypothesis testing, even if searching for a best compromise model.

1.1.2. Flexible Model Structures

[11] The arguments above suggest that a fixed yet parsimonious model structure, attractive as it may be, may not be (at least currently) achievable in its pure form. In practice, modelers may need to include additional components to emphasize processes that are significant and/or dominant,

adjust constitutive functions to better match independent diagnostic evidence obtained from recession analysis, etc. Indeed, *Vaché and McDonnell* [2006] advocate "malleable" model structures in their hypothesis-testing study. Yet when hypothesizing and testing competing models, ad hoc methodologies may yield confounding results. What is needed, then, is a systematic and robust platform for carrying out model development, evaluation, and comparison [e.g., *Refsgaard and Henriksen*, 2004; *Young and Ratto*, 2009; *Clark et al.*, 2011b].

[12] Recognizing these challenges, a number of modular toolkits have been developed. For example, the Modular Modeling System (MMS) [*Leavesley et al.*, 1996] allows combining different submodels (e.g., linking a surface water model with a groundwater model such as MODFLOW). The Rainfall-Runoff Modeling Toolkit (RRMT) [*Wagener et al.*, 2002] allows a selection of different moisture accounting and routing modules, importantly, also providing a suite of diagnostic options to facilitate model evaluation and comparison. The Framework for Understanding Structural Errors (FUSE) [*Clark et al.*, 2008] provides a selection of alternative representations of the catchment soil store based on several existing models such as Sacramento [*Burnash*, 1995] and TOPMODEL [*Beven*, 1997], and it can be extended to include components such as vegetation and so on. Similar developments are taking place in the distributed modeling community, with models such as MIKE-SHE and MIKE-11 being extended to accommodate different processes representations [*Butts et al.*, 2004; *Refsgaard and Henriksen*, 2004], and in the "data-based mechanistic" (DBM) community [*Young*, 1998; *Young and Ratto*, 2009]. Overall, these approaches facilitate the application of the method of multiple working hypotheses to hydrological modeling [*Clark et al.*, 2011b]. In this paper, we advocate these principles for conceptual hydrological models.

[13] In conceptual hydrological modeling, flexible and multimodel frameworks have been already used to investigate questions such the patterns of structural errors across multiple catchments [*Clark et al.*, 2008], mean residence time and catchment mixing mechanisms [*Fenicia et al.*, 2010], representations of plot-scale surface and groundwater dynamics [*Krueger et al.*, 2010], model component fidelity to its intended process [*Clark et al.*, 2011a], time scale control on model parameters and inferred complexity [*Kavetski et al.*, 2011], and others.

[14] Finally, there are increasing calls for a more "holistic" view of catchment-scale hydrology, moving beyond a "mere" description of heterogeneity and toward a more unified theory that recognizes the self-organization, optimality, and human influences that shape the traits and patterns of specific catchments as well as entire classes of environmental systems [*Sivapalan*, 2005; *McDonnell et al.*, 2007; *Troch et al.*, 2009; *Sivapalan*, 2009]. Here a flexible model structure can accommodate the diversity of the climatologic, geomorphologic, and anthropogenic factors operating at specific locations [*Sivapalan*, 2009].

1.1.3. Numerical Robustness Issues in Conceptual Hydrological Modeling

[15] Conceptual model design generally begins with a perceptual model, proceeding through a mathematical formulation of the hypothesized structure (e.g., linking model components, relating fluxes to storages, etc.) to the numerical implementation in a computer code [e.g., *Beven*, 2001; *Clark et al.*, 2008]. Given the extent of process approximations made in conceptual hydrological modeling, it has been often tempting to disregard careful numerical implementations in favor of computationally simplistic schemes (e.g., see review by *Clark and Kavetski* [2010]). However, unless adequately controlled, numerical errors can easily overwhelm data and structural uncertainty and lead to major errors in model identification and interpretation, in uncertainty and sensitivity analyses, and, quite troublingly, in operational predictions [e.g., *Kavetski et al.*, 2003; *Kavetski and Clark*, 2010; *Michel et al.*, 2003, 2005; *Schoups et al.*, 2010]. Numerical robustness is hence a key design consideration in the hydrological modeling methods presented in this paper [*Clark and Kavetski*, 2010].

1.2. Aims and Significance

[16] This two-part paper formulates and illustrates a flexible framework for conceptual hydrological modeling, with two related objectives: (1) to generalize and systematize the field of conceptual models and augment existing flexible frameworks and (2) to provide a practical platform for pursuing process-oriented insights into catchment-scale water cycle dynamics as well as more robust and reliable performance in operational contexts.

[17] The SUPERFLEX modeling framework presented here is based on generic building blocks, such as reservoirs, junctions, and constitutive functions, implemented using robust numerical techniques. We argue that the richer set of model structures and parameterizations, which can be obtained by mix and matching generic components, is particularly well suited to supporting hydrologists' analyses, decisions, and testing with respect to which processes to include in the analysis and how to represent them. It also supports the incorporation of catchment-specific experimental insights as part of iterative model improvement [*Fenicia et al.*, 2008a].

[18] The proposed flexible framework provides a unified platform that not only includes many existing precipitation-streamflow models, but also simplifies the generation of new hypotheses (model configurations) as opposed to combinations of existing models. This extends the capabilities currently available in models such as FLEX [*Fenicia et al.*, 2008b] and other flexible frameworks such as FUSE [*Clark et al.*, 2008], which tended to outline an overall model architecture and then "populate it" using subcomponents of existing models. For example, the current version of FUSE [*Clark et al.*, 2008] hypothesizes a two-layer soil store comprising subcomponents and constitutive functions adapted from successful existing models, and RRMT [*Wagener et al.*, 2002] assumes a four-component configuration with interception, root zone, and fast/slow routing, etc.

[19] Our presentation further stresses the key distinction between the model hypothesis and its numerical implementation, and illustrates computational design considerations in hydrological model building. This is an often-neglected aspect in hydrological modeling, unnecessarily risking major numerical artifacts. Addressing this is critical for meaningful advances in hydrological science and operations [e.g., *Kavetski et al.*, 2011].

[20] In the companion paper [*Kavetski and Fenicia*, 2011], the potential of the flexible framework to provide

scientific insights and operational reliability in diverse catchment conditions is appraised through a series of case studies and is compared to the insights and performance achievable using a fixed model structure (here the GR4H model, which is an hourly version of the widely used GR4J model [*Perrin et al.*, 2003]).

[21] The ability to robustly implement and systematically evaluate different model hypotheses (formulated at any level, including architecture, connectivity, individual process representation, etc.) within a single framework facilitates learning from model comparison and hypothesis testing, unencumbered by confounding differences in conceptualization philosophy, numerical algorithms, software implementation, etc. [*Clark et al.*, 2011b]. Elements of the proposed framework, named SUPERFLEX as it develops the earlier FLEX model [*Fenicia et al.*, 2008b], have already shown promise in recent applications that explored the mean residence time of the Maimai basin [*Fenicia et al.*, 2010] and the interaction between calibration data resolution and model inference in the Weierbach basin [*Kavetski et al.*, 2011]. In this paper, the focus is on challenging the prevalent "fixed" model paradigm and on motivating and formulating SUPERFLEX as a multihypothesis framework [*Clark et al.*, 2011b] that seeks to provide physically interpretable analysis of catchment dynamics and diversity.

2. Model Architecture, Governing Equations, and Constitutive Functions

2.1. Model Architecture and General System Functionalities

[22] The overall model architecture, which defines and couples the dominant systems processes, is a key hypothesis. It is known that the connectivities between catchment compartments and their dependencies on storage thresholds, soil properties, and surface and bedrock topography significantly influence the hydrological [*McGuire and McDonnell*, 2010], ecological [*Pringle*, 2003], and morphological [*Reid et al.*, 2007] behavior of natural systems. Therefore, process interactions and the influence of common controls (such as topography, as in TOPMODEL [*Beven*, 1997]) must be reflected in the hypothesized model structure [see also *Sivapalan*, 2005].

[23] From a functional (mechanistic) perspective, catchment dynamics include partition, storage, release, and transmission of water [e.g., *Wagener et al.*, 2007]. In the SUPERFLEX framework, a hypothesized model structure can be implemented using generic components intended to approximate these functions:

[24] 1. Reservoir element: represents storage and release of water;

[25] 2. Lag function element: represents the transmission and delay of fluxes;

[26] 3. Junction element: represents the splitting, merging, and/or rescaling of fluxes.

[27] These components constitute the main building blocks of many existing conceptual models. In SUPERFLEX these elements are generalized and can be arranged into different flow configurations representing alternative conceptual hypotheses of catchment function. The model hypotheses are further characterized by selecting their constitutive functions (e.g., relating fluxes to reservoir storage) and associated parameters.

2.2. Storage-Release Processes: The Generic Reservoir Element

[28] A "generic" reservoir conceptualizes catchment-scale processes involving storage and release of water. It is shown schematically in Figure 1a and, mathematically, is described using ordinary differential equations (ODEs):

$$\mathrm{d}\mathbf{S}(t)/\mathrm{d}t = \mathbf{g}_S[\mathbf{S}(t), \mathbf{X}(t)|\boldsymbol{\theta}], \tag{1a}$$

$$\mathbf{Q}(t) = \mathbf{g}_Q[\mathbf{S}(t), \mathbf{X}(t)|\boldsymbol{\theta}], \tag{1b}$$

where $\mathbf{S}(t)$ are conceptual storage values at time t, $\mathbf{X}(t)$ is the (time-dependent) forcing and $\mathbf{Q}(t)$ is the response, $\mathbf{g}()$

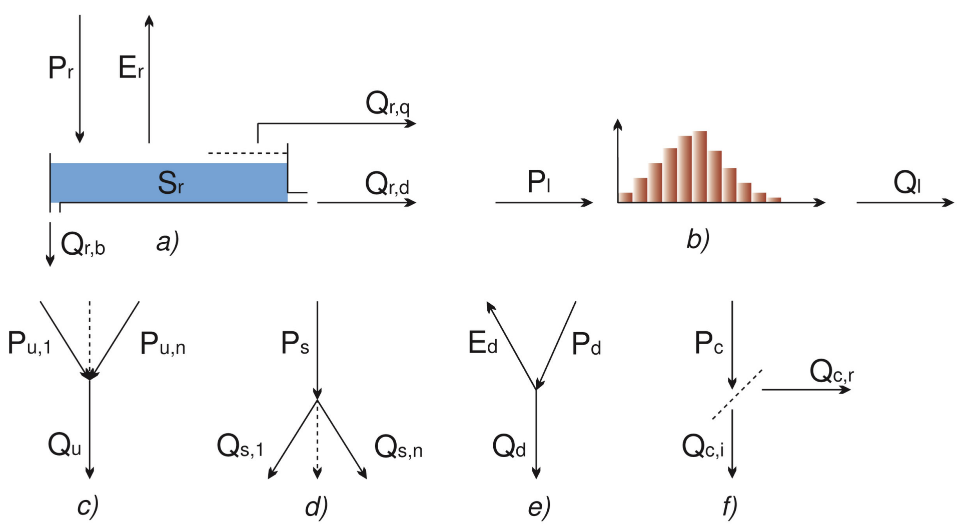

Figure 1. Generic building blocks of the flexible framework: (a) generic reservoir and (b) lag function. Connection elements: (c) union and (d) splitter. Splitters can be used to represent (e) the subtraction of evaporation from rainfall and (f) the threshold-type occurrence of Hortonian flow.

are the input-output fluxes associated with the component, and $\boldsymbol{\theta}$ are parameter values (section 2.5).

[29] Depending on its internal structure (see section 2.4), a reservoir can exhibit strong nonlinearities, including threshold behavior. Widely used in hydrological modeling, reservoirs can parsimoniously represent the dynamics of interception, soil moisture, groundwater, etc. (e.g., see *Clark et al.* [2008] for an overview). In addition, they can represent snow accumulation and melting dynamics [*Smith and Marshall*, 2010].

[30] While most conceptual hydrological models use multiple single-state reservoirs, more complex coupling can be supported within ODE (1), including mass and/or energy balances. For example, *Fenicia et al.* [2010] detail the inclusion of tracer balance as part of using a flexible model framework to investigate the influence of mixing assumptions on mean residence time estimation using experimental tracer data. Given the similarities between the conservation laws governing hydraulic, thermal, and other systems, suitably enhanced reservoir-type components can also be used to formulate other natural laws, such as conservation of energy, momentum, and mass (e.g., as in the temperature "reservoir" model of *Westhoff et al.* [2007]).

2.3. Flow Routing: Lag Function (Convolution Operator) Elements

[31] In many catchments, using a single reservoir, or even a few reservoirs, may be insufficient to represent the routing of water through extended flow networks. Instead, a different mathematical representation based on the convolution operator can be exploited (Figure 1b). Many existing models, including HBV [*Lindström et al.*, 1997], GR4J [*Perrin et al.*, 2003], and FUSE [*Clark et al.*, 2008], use convolutions to represent delays arising from channel routing. In models such as HBV and FUSE, a single convolution is applied to the output from the system of reservoirs. Here, in order to provide further flexibility, we allow lag functions to be included at any location within the system. This is the approach taken in GR4J and in the earlier FLEX model [*Fenicia et al.*, 2007], where the output of the lag function is fed into reservoirs farther downstream. For example, this can be useful when representing drainage from the unsaturated zone to the saturated zone in situations where the water table is several meters below the ground surface and the travel time is therefore not instantaneous.

[32] There is a close mathematical connection between reservoirs and lag function elements. In particular, the lag function represents the response of a (possibly multistate) linear system to a unit impulse [e.g., *Oppenheim et al.*, 1999]. Hence, the distinction between a "reservoir" and a "lag function" is, at least for linear systems, essentially semantic, reflecting different mathematical formulations of the same concept. For example, *Nash* [1958] showed that the impulse-response (lag) function associated with a sequence of N identical linear reservoirs (the "Nash cascade") is the Gamma function. As a result, many models, including rainfall-runoff models such as FUSE [*Clark et al.*, 2008] and HYMOD [e.g., *Wagener et al.*, 2001] as well as tracer models [e.g., *Weiler et al.*, 2003; *Hrachowitz et al.*, 2010], use the Gamma function in their convolutions. Other models, such as HBV [*Lindström et al.*, 1997], use triangular lag functions, sometimes allowing for curvature in their limbs, as in GR4J [*Perrin et al.*, 2003]. Transfer function modeling also underlies the DBM philosophy [*Young*, 1998; *Young and Ratto*, 2009] for characterizing the input-output behavior of a general system, with suitable enhancements to represent system nonlinearities [*Young*, 2003; *Young and Garnier*, 2006]. Hence, a flexible framework should support the specification of Gamma, triangular, and other lag functions.

[33] More general parametric and nonparametric forms of lag functions can also be accommodated. These allow representing the behavior of multiple linear reservoirs, including serial and parallel connections, and not necessarily with equal time constants. However, the impulse response of nonlinear systems (reservoirs) is more complex than for linear systems because of state dependencies and cannot be described using standard linear transfer functions [e.g., *Vidyasagar*, 2002]. In this respect, the nonlinear reservoir conceptualization is more general. Yet it is also more computationally expensive than a linear store, especially in multireservoir models.

2.4. Process Connectivities and Interactions: Junction Elements

[34] As elaborated in section 2.1, the overall system architecture is a key hypothesis within model development. Within a single model configuration, multiple reservoirs and lag functions are coupled using junction elements. Junctions are "zero-state" components, which may also contain parameters. For example, the junction component combines several fluxes, e.g., outputs from different reservoirs (Figure 1c). A splitter component can describe the separation of a single upstream flux into two or more downstream fluxes, e.g., as in GR4J [*Perrin et al.*, 2003] and HYMOD [*Wagener et al.*, 2001] (see Figure 1d). It can also be used to remove evaporation from rainfall, e.g., as in the interception module of GR4J (Figure 1e). With suitable specification of constitutive functions, e.g., with the splitting fractions dependent on the input, splitters can also simulate Hortonian overland flow occurring when precipitation intensity exceeds the soil infiltration capacity [e.g., *Beven*, 2004] (Figure 1f).

[35] More generally, the model architecture should also be reflective of spatial characteristics of the catchment. For example, separate storage and lag function elements can be used as part of a discretization of the catchment into areal elements [e.g., *Uhlenbrook et al.*, 2004] or into process-oriented classes of elements (e.g., as in the topographic index classes of TOPMODEL [*Beven*, 1997]). While such semidistributed approaches are beyond the scope of this presentation, they represent an important direction of ongoing studies.

2.5. Constitutive Functions

[36] The properties of individual components and hence the entire model are defined by the constitutive functions that represent hypothesized storage-discharge relations of reservoirs, shapes of lag functions, and characteristics of junction elements. We envisage, on the basis of experiences gained through the application of the framework, that constitutive functions will form part of an extendable "library" (e.g., some typical options are listed in Table 1, illustrated in Figure 2a, and explored as part of the empirical case studies in the companion paper).

[37] For example, a typical conceptualization of saturation–excess runoff is $g_S(S, P|\boldsymbol{\theta}) = [1 - A(S|\boldsymbol{\theta})]P$ and $g_Q(S, P|\boldsymbol{\theta})$

Table 1. Examples of Constitutive Functions[a]

Function	Name
$f_p(x\|m) = x^m$	power function
$f_r(x\|m) = 1 - (1-x)^m$	reflected power function [*Moore*, 1985]
$f_m(x\|m) = (1+m)\dfrac{x}{x+m}$	Monod-type kinetics, adjusted so that $f_m(1\|m) = 1$
$f_h(x\|m) = 1 - \dfrac{(1-x)(1+m)}{1-x+m}$	reflected hyperbolic function, scaled to the unit square
$f_e(x\|m) = 1 - e^{-x/m}$	Tessier function (note that $f_e(x\|m) \to 1$ as $x \to \infty$)
$f_\lambda(x\|m,\lambda) = \dfrac{\left(1+e^{-m(1-\lambda)}\right)\left(e^{-mx}-1\right)}{\left(1+e^{-m(x-\lambda)}\right)\left(e^{-m}-1\right)}$	modified logistic curve, scaled to the unit square

[a]Additional shapes can be accommodated within the flexible framework and its software implementation. Here $e = 2.71\ldots$ denotes the natural logarithm base.

$= A(S|\boldsymbol{\theta})P$, where $A(S|\boldsymbol{\theta})$ is the "saturated area" fraction, e.g., $A(S|\boldsymbol{\theta}) = \left(S|\theta_1\right)^{\theta_2}$. This formulation can be interpreted from an "infiltration capacity distribution" perspective [*Moore and Clarke*, 1981] and underlies the saturation–excess runoff mechanism in TOPMODEL [*Beven*, 1997] and VIC [*Wood et al.*, 1992].

[38] The form of the constitutive functions can be hypothesized parametrically a priori or estimated empirically from independent data analysis (e.g., a topography-based saturation function [*Beven*, 1997]). More general nonparametric forms can also be accommodated. For example, (1) recession analysis can be used to specify the storage-discharge relationship [e.g., *Lamb and Beven*, 1997; *Rupp and Woods*, 2008; *Fenicia et al.*, 2010], and (2) spatial analysis, e.g., using digital elevation maps, can be used to derive topographic-based saturation functions for TOPMODEL-like conceptualizations of saturation overland flow and so on [*Beven*, 1997; see also *Moore*, 1985]. This provides the flexible framework with capabilities to exploit independent information in addition to the rainfall-runoff time series typically used in hydrological calibration.

[39] "Composite" functional relationships can also be constructed. For example, the "leaky catchment" structure of *Wagener et al.* [2002], which represents a reservoir with multiple outlets at different storage "depths" (Figure 2b), can be accommodated in the general formulation using storage-discharge relations with (smoothed) piecewise linear or curvilinear segments (Figure 2c; see application of *Fenicia et al.* [2010]).

3. Computational Implementation

[40] Linear reservoir models and some simple nonlinear cases can be solved analytically. However, general exact solutions are rare. For example, while the power reservoir $\mathrm{d}S/\mathrm{d}t = P - kS^\alpha$ can be solved analytically when $P = 0$, we are not aware of analytical solutions for $P > 0$. When considering multiple coupled equations (e.g., for a coupled multicompartment system), analytical solutions are even more restrictive. Numerical approximations are hence necessary (e.g., see the recent review by *Clark and Kavetski* [2010]). Here we apply time stepping schemes over fixed discrete steps $\Delta t = t_{n+1} - t_n$, where the subscript n indexes the time step. The models are forced with data $\bar{\mathbf{X}}_{n+1/2}$ representing the observed average of $\mathbf{X}(t)$ over each such "data resolution" step.

[41] The importance of mathematically well-behaved hydrological models was recognized decades ago by *Ibbitt and O'Donnell* [1971] and *Moore and Clarke* [1981]. Despite an apparent simplicity, the numerical design of a conceptual hydrological model requires careful attention to avoid harmful numerical artifacts [*Kavetski et al.*, 2011]. For example, explicit methods, such as the explicit Euler (EE) scheme $\mathbf{S}_{n+1}^{(\mathrm{EE})} = \mathbf{S}_n + \Delta t\,\mathbf{g}(\mathbf{S}_n, \bar{\mathbf{X}}_n)$, are quite widely used in conceptual hydrological models (e.g., see the review by *Clark and Kavetski* [2010]). However, though superficially attractive because of algorithmic simplicity and low cost per step, fixed step explicit integration is only conditionally stable and can generate very large errors. It

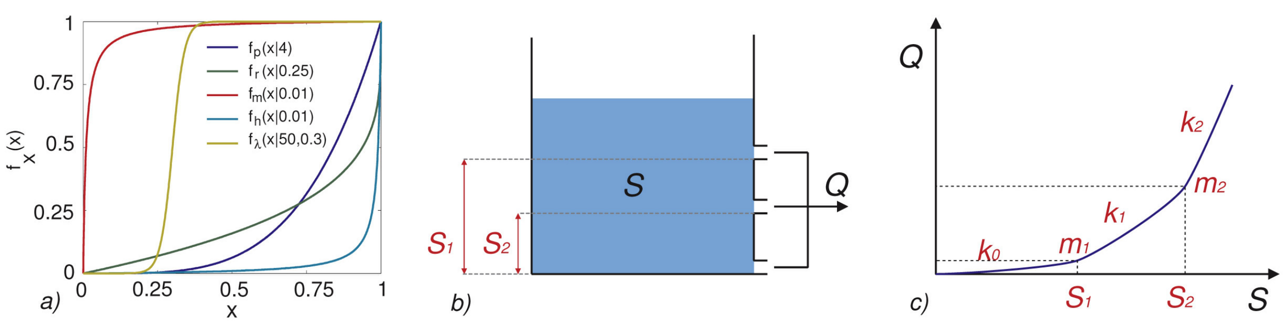

Figure 2. (a) Representative shapes of constitutive functions listed in Table 1 (parameter values indicated in brackets). (b) Reservoir with multiple outflow thresholds and (c) the corresponding "composite" constitutive relationship. Here parameters m_1 and m_2 control the smoothness of the transitions and would normally not be calibrated [*Kavetski and Kuczera*, 2007].

should not be used in practical computing unless implemented with adaptive substepping and numerical error control [e.g., *Kahaner et al.*, 1989]. Consequently, the fixed step EE method is not used in this paper.

[42] Sections 3.1 and 3.2 outline more reliable numerical strategies as well as their limitations.

3.1. Time Stepping Implementation

3.1.1. Fixed Step Implicit Approximations

[43] The implicit Euler (IE) method,

$$\mathbf{S}_{n+1}^{(\mathrm{IE})} = \mathbf{S}_n + \Delta t\, \mathbf{g}\left(\mathbf{S}_{n+1}^{(\mathrm{IE})}, \bar{\mathbf{X}}_{n+1/2}\right), \tag{2}$$

is an unconditionally stable approximation of equation (1) [e.g., *Kahaner et al.*, 1989]. Despite requiring potentially expensive iterative solutions (since equation (2) is, in general, nonlinear), the implicit Euler scheme is generally robust even for large step sizes. It is widely used in standard engineering software, e.g., the MODFLOW package for groundwater simulations, the ECLIPSE tool in the petroleum industry, geotechnical consolidation codes, etc. [e.g., *Clark and Kavetski*, 2010, and references therein]. Importantly, when implemented using fixed steps and a tight Newton-Raphson tolerance, the implicit Euler solution is smooth with respect to its forcings and parameters, resulting in a smooth objective function of the hydrological model. This significantly facilitates calibration and prediction. Given these considerations, we implemented the fixed step implicit Euler scheme as a solution option. Note that for closer correspondence with rainfall observation systems that report rainfall depths over discrete time intervals, equation (2) uses the step-averaged forcing rather than an "end-of-step" forcing.

3.1.2. Adaptive (Explicit or Implicit) Solutions

[44] From a strict numerical analysis perspective, unconditional stability is neither necessary nor sufficient to guarantee an accurate approximation. This is easily shown for systems with rapidly varying forcing. Implicit schemes are also computationally inefficient compared to explicit schemes when the governing equations are not "stiff" (e.g., see Appendix A of *Clark and Kavetski* [2010] for an outline or *Lambert* [1991] for a full exposition). Consequently, the numerical ODE literature discourages using fixed step methods at all, whether explicit or implicit [e.g., *Shampine and Reichelt*, 1997]. From this perspective, adaptive substepping within each "data resolution" step is a safer solution option because it produces near-exact solutions of the hypothesized model equations. This could be accomplished, e.g., using an adaptive fifth-order explicit Runge-Kutta method or adaptive implicit or semi-implicit methods [e.g., *Press et al.*, 1992; *Butcher*, 2008].

[45] However, in the context of model parameter optimization, adaptive substepping has an undesirable property. Unless very tight truncation temporal error tolerances are imposed in the ODE integrator, the objective function of the hydrological model will be discontinuous at the microscale, degrading the performance of gradient-based optimization and other derivative-based analysis (e.g., linearized covariance estimation). As discussed by *Kavetski and Clark* [2010], the trade-off between the fixed step implicit Euler versus adaptive schemes is not necessarily clear-cut and may depend on the modeling context.

[46] By formulating its governing equations in continuous-time state-space form, hydrological software can be much more readily upgraded to a different numerical solver if necessary. This should be contrasted to developing a model directly in discrete-time form, where it is not clear where the process conceptualizations end and numerical approximations begin, how to change the model step size, etc. See also *Young and Garnier* [2006] for a fuller discussion of the advantages of continuous-time models of physical phenomena, and see *Kavetski et al*. [2011] for a case study on the impact of data resolution on identifiable model complexity.

3.1.3. Alternative Implementation Approaches Such as Operator Splitting

[47] Operator-splitting (OS) numerical approximations integrate each flux sequentially in a predetermined order to estimate the overall solution [e.g., *Press et al.*, 1992] (see also *Kavetski et al.* [2003] and *Schoups et al.* [2010] for a discussion in hydrology). OS approaches are extremely useful for solving large systems of nonlinear partial differential equations (PDEs) with multiple components, such as those coupling advection, reaction, and diffusion, especially when dealing with large spatially distributed systems. In such cases, specialized numerical and/or semianalytical techniques can be applied to each flux separately [*Steefel and MacQuarrie*, 1996], making it possible to approximate problems that are otherwise computationally intractable.

[48] In conceptual hydrological modeling, OS schemes are appealing in the specific case of implementing a fixed model structure, especially in a context where computational speed is a strongly dominant factor and hence a customized implementation is needed (e.g., see the evaluations for an exponential-type reservoir model by *Schoups et al.* [2010]). However, in the more general case of hypothesis testing in hydrological modeling, especially when flexible configurations need to be supported, OS approaches have several important limitations:

[49] 1. They correspond to the physically unsatisfying assumption that hydrological processes operate in a certain predetermined order, for example, evaporation first, followed by percolation, surface runoff, etc.

[50] 2. Even if the individual fluxes are integrated analytically, which is possible for certain flux forms (although usually for single reservoirs only, as detailed by *Schoups et al.* [2010]), the final solution will still contain numerical ("splitting") errors, which are a consequence of approximation listed in point 1.

[51] 3. When the hydrological model has coupled compartments, it is usually impossible to jointly integrate their fluxes analytically. But integrating them numerically foregoes a major potential advantage of OS in the context of rainfall-runoff modeling: the use of analytical solutions for individual fluxes.

[52] 4. OS can result in very complicated computer code, with multiple branches depending on the availability of analytical solutions and so on. This is especially problematic for a flexible framework, which must support a variety of flow networks, connectivities between storage elements, and different forms of constitutive fluxes. For most configurations, analytical solutions will not be available, and numerical techniques will again be needed.

[53] Given these considerations, we have not used operator-splitting solutions for the flexible models in this study.

3.2. Ensuring Smoothness in the Model Components

[54] System analyses and predictive applications can benefit significantly when the model response surface is smooth and well behaved [e.g., *Ibbitt and O'Donnell*, 1971; *Kavetski and Clark*, 2010]. For example, avoiding artificial discontinuities with respect to the model inputs is useful when exploring system sensitivities with respect to changes in future forcing. In terms of model building, this requires continuity and microscale smoothness with respect to forcing, states, and parameters in the model's constitutive functions, governing equations, and numerical approximations [*Kavetski et al.*, 2006; *Kavetski and Kuczera*, 2007; *Clark et al.*, 2011a]. In calibration, this also requires the objective function to be continuous and smooth with respect to the model response.

[55] In order to support overall model smoothness, the building blocks of the SUPERFLEX framework are implemented using smoothed techniques. This includes the storage-flux relationships in the reservoir elements, the lag functions in basin routing elements, and any parameterizations in the junction elements. Many smoothing techniques can be exploited for this purpose, including smoothed piecewise linear approximants [e.g., *Koutsoyiannis*, 2000; *Kavetski and Kuczera*, 2007] and more general smoothed piecewise curvilinear functions such as (cubic) splines and Bezier curves [e.g., *Kahaner et al.*, 1989]. These are not demonstrated in this work but are quite readily implemented. Note that smoothness should be enforced with respect to states, parameters, and forcing functions [*Kavetski and Kuczera*, 2007].

[56] Smoothing is more than a numerical approximation and as such represents a genuine alteration of the model equations. Since many environmental processes exhibit marked threshold behavior [*Zehe and Sivapalan*, 2009; *Spence*, 2010], the degree of smoothing can be controlled so that it facilitates mathematical analysis but does not materially affect the model predictions [*Kavetski and Kuczera*, 2007]. However, while many environmental dynamics have pronounced thresholds at a point spatiotemporal scale, they have much smoother behavior when integrated over catchment-scale areas [e.g., *Moore*, 1985]. This may explain improvements in hydrological model performance when stronger smoothing is applied. For example, see Figure 1 of *Kavetski et al.* [2006] and the discussion by *Clark et al.* [2011a, section 7.4].

4. Discussion

4.1. A Broader View on Modular Frameworks

[57] Many of the concepts and rationale behind flexible ("modular") model development are not new or unique to hydrology. It has long been recognized that at least in practical contexts, the degree of detail, resolution, and comprehensiveness of a model will usually depend on the study objectives, data constraints, spatial and temporal scales of application, and other factors [e.g., *Wagener et al.*, 2001; *Leavesley et al.*, 2002; *Refsgaard and Henriksen*, 2004; *Clark et al.*, 2011b].

[58] Several modular frameworks have been developed and applied in hydrological modeling. Depending on their "granularity" and underlying rationale, they could be loosely classified into "model-interfacing frameworks" versus "flexible process representation frameworks." In the former, the building blocks are often entire hydrological models in their own right, and the objective is to exploit existing models to build increasingly larger-scale integrated environmental models [e.g., *Kumar et al.*, 2006]. However, our aim in this paper is not to build larger-scale models but to provide finer-scale flexibility in hypothesizing and testing the overall catchment system architecture, component connectivity, and representation of individual processes. Hence, a key criterion is the "granularity" of the model hypotheses (see discussions by *Clark et al.* [2011b]).

[59] When viewed from this hypothesis-based perspective, SUPERFLEX and the earlier FLEX model [*Fenicia et al.*, 2008b] share their motivation with the FUSE framework [*Clark et al.*, 2008]. FUSE relaxed several significant restrictions of modular methodologies where the overall model architecture is fixed to contain certain functional components (e.g., interception, soil store, groundwater, and routing) in a predetermined order and the flexibility consisted in specifying the constitutive functions for these components. For example, see the schematic of the Cold Region Model [*Pomeroy et al.*, 2007, Figure 4]; an analogous structure is adopted in the RRMT [*Wagener et al.*, 2002]. FUSE allows considerably more flexibility in the architecture of the model. However, its current implementation remains restricted to a two-layer hypothesis of the soil store, which is then "populated" using components and constitutive functions from existing models such as TOPMODEL, ARNO/VIC, and others [e.g., see *Clark et al.*, 2008, Figure 3]. Conversely, this paper argues that an even richer set of model hypotheses can be constructed using generic model components such as those listed in section 2. For example, the model structures applied in the case studies in the companion paper vary, in a controlled way, in the number of reservoirs, their connectivities, and the constitutive relationships used (e.g., see Figure 2 of the companion paper).

[60] Yet we do not argue for "black box" models such as neural networks [e.g., *Kingston et al.*, 2008], which are not easily physically interpretable and are prone to overparameterization. To the contrary, we seek maximum incorporation of physical insights into hydrological model development and scrutiny of response signatures and individual components against all available observed data (e.g., as advocated by *Seibert and McDonnell* [2002], *Freer et al.* [2004], *Gupta et al.* [2008], *Fenicia et al.* [2008b], *Yilmaz et al.* [2008], and *Clark et al.* [2011a]).

[61] The SUPERFLEX framework introduced in this paper is a generalization of the FLEX model [*Fenicia et al.*, 2008b]. It is a more comprehensive application of the formalism of nonlinear differential equations, transfer functions, and connection elements to build conceptual hydrological models in a way that seeks to combine the perceptual understanding that is often available from experimental work with statistical inference of model parameters and structures. This helps pursue the dialog between "modeler" and "experimentalist" advocated by *Seibert and McDonnell* [2002]. An important practical advance relative to the earlier FLEX applications is the implementation of all components using robust numerical techniques (section 1.1.3). A broader objective relative to previous conceptual modeling methods, including FLEX, is the finer granularity of the supported

model decisions and wider range of possible model architectures and constitutive relationships between the state variables, fluxes, and parameters. As discussed, this allows implementing both the hypotheses implemented in existing hydrological models and multiple alternative hypotheses such as those presented in Part 2 of this paper [*Kavetski and Fenicia*, 2011]. This allows a broader and more thorough exploration of the model hypothesis space.

[62] While this paper focuses on conceptual hydrological models, useful connections can also be made to the class of DBM methods [*Young*, 1998], which use transfer function modeling to approximate the input-output behavior of a general system. Physical interpretability is an explicit objective of the DBM philosophy [*Young*, 2003]; model parsimony is sought using Dominant Mode Analysis [*Young and Ratto*, 2009]. DBM has been applied in many areas of science, engineering, and economics [*Young*, 1998, 2003; *Young and Ratto*, 2009]. When using lag functions and linear reservoirs, the framework in this paper is mathematically analogous to DBM models: continuous-time transfer function models are an alternative representation of linear differential equations [*Young and Ratto*, 2009]. However, in terms of model structure development, the conceptualization and mathematical handling of nonlinearities are different. In hydrological applications of DBM, system nonlinearities have been handled either (1) through a nonlinear transformation of the rainfall or runoff to estimate the "effective input," followed by an assumption that the subsequent routing dynamics can be approximated using linear components, or (2) using state- and/or time-dependent parameters in the transfer function representation [e.g., *Young*, 2002, 2005]. Conversely, SUPERFLEX can represent system nonlinearities using nonlinear differential equations or using lag functions.

[63] Consider the direct representation of nonlinearities in any system component using nonlinear differential equations (conceptualized as reservoirs with nonlinear discharge relationships), as carried out here and in similar conceptual models such as FUSE [*Clark et al.*, 2008] and PDM [*Moore and Clarke*, 1981]. In our opinion, this formalism may be better suited to hydrologists, both in the process of model building based on independent field insights (e.g., through the choice of components that more closely resemble the perception of physical processes by hydrologists and experimentalists) and in the process of model interpretation. Relevant considerations here are the incorporation of independently determined storage-discharge relationships (e.g., using recession analysis [*Lamb and Beven*, 1997; *Rupp and Woods*, 2008; *Kirchner*, 2009; *Clark et al.*, 2009]) and the generality of the framework to accommodate hydrological hypotheses where spatial variability is represented using systems of coupled nonlinear PDEs. Note that while nonlinear DE models do not require time-varying parameters to represent system nonlinearities, stochastic parameter variation may still be exploited to represent structural uncertainty (e.g., such as that due to lumping [*Kuczera et al.*, 2006]). The combination of DBM-type model structure identification techniques (e.g., Dominant Mode Analysis [*Young and Ratto*, 2009]) and Bayesian-type techniques for incorporating independent insights into nonlinear ODE and PDE models could represent a major step forward in hydrological modeling. Although it is a considerable challenge, we are seeing useful progress in this broad direction [e.g., *Bulygina and Gupta*, 2009].

[64] Finally, we stress that the focus of this paper is not on a particular computer code or on software design aspects. Rather, the focus is on the conceptual elements supporting controlled flexibility in hydrological models and on exploring their potential to improve hypothesis testing over a diverse range of environmental conditions. With suitable modifications, existing modular hydrological software may be extended to implement structures such as those applied in the companion paper and other model hypotheses. This is, indeed, the key appeal of a well-designed "plug-and-play" modular framework and software [e.g., *Leavesley et al.*, 2002; *Pomeroy et al.*, 2007; *Clark et al.*, 2008].

4.2. The Challenge of "Uniqueness of Place"

[65] The concept of "uniqueness of place" alludes to the diversity of nature and the consequent difficulty of formulating general hydrological models [e.g., *Beven*, 2000]. This concept in itself has generated debates, e.g., the exchange between Beven and Pappenberger and Abbot et al. [*Beven and Pappenberger*, 2003], where Abbot et al. challenged "uniqueness of place" as an artifact of the crudeness of model representations, such as the lumping of distinctly different hydraulic controls such as flood plains, barriers, etc. We agree that knowledge of the heterogeneities within a catchment could support models based on independently derived physical laws (which, as vividly pointed out by Abbott et al., are equally valid in central Africa as on the North Pole). However, it is not clear whether sufficiently accurate and precise distributed information, especially for the subsurface of a catchment, can become feasibly and routinely available in the foreseeable future [e.g., *Beven*, 2000; *Kirchner*, 2006]. Hence, "uniqueness of place" could be viewed as arising when the current lack of reliable prior parameterization data and current limitations in process representation are "handled" by formulating lumped conceptual models directly at the catchment scale. In other words, apparent "uniqueness of place" could be the price of the apparent "simplicity" of lumped models. Yet catchment-scale behavior could also be simpler than suggested purely by subscale complexities [e.g., *McDonnell et al.*, 2007; *Savenije*, 2009; *Sivapalan*, 2009]. In either case, we believe the extent of resulting "uniqueness" and/or "generality" is amenable to careful quantitative analysis and is far from a theoretical or practical impasse (see *Sivapalan* [2009] for examples).

[66] The community response to the challenge of "catchment uniqueness" has taken several forms. For example, a focus on specific "idiosyncrasies" of individual catchments has resulted in a largely ad hoc development of a plethora of fixed structure models differing uncontrollably in many distinct respects such as level of process conceptualization, spatial discretization, implementation details, etc. [*McDonnell and Woods*, 2004]. Many models with alternative conceptualizations have also been applied to the same areas (e.g., see section 4.4.1 in the companion paper [*Kavetski and Fenicia*, 2011] for a brief review of models previously developed for the Maimai area). As pointed out by *McDonnell and Woods* [2004], little has been done to compare and generalize these findings and use them to explain the behavior of different basins and to provide reliable practical guidelines for hydrological model development based on process understanding.

A flexible model framework can help organize insights from multiple existing models.

[67] Another approach to the catchment diversity problem has been the ongoing search for good "compromise" models, e.g., the multicatchment studies used to derive the GR4J model [e.g., *Edijatno and Michel*, 1989; *Edijatno et al.*, 1999; *Perrin et al.*, 2003; *Le Moine et al.*, 2007]. In these studies, the comparison was carried out between several fixed model structures applied to many different catchments. This may simplify application in operational purposes but, in our opinion, is not the most effective strategy for elucidating differences and similarities of catchments. Instead of comparing the errors of a fixed structure in different catchments, it could be equally or more insightful to examine differences between models that provide a good description of these basins. To avoid either masking or exaggerating intercatchment differences, such analysis requires rigorously accounting for data uncertainty [e.g., *Renard et al.*, 2010] and scrutinizing each individual application using multiple diagnostics [e.g., *Gupta et al.*, 2008]. Unfortunately, as critiqued by *Pappenberger and Beven* [2006], recognition of uncertainties and their interactions with model development and scrutiny remains far from established and routinely used in environmental research, let alone practice. Indeed, previously reported findings and this study both suggest that fixed structure applications would be unlikely to pass a stringent degree of scrutiny. For example, large multicatchment studies [e.g., *Perrin et al.*, 2001; *Le Moine et al.*, 2007] by necessity relied on quite limited performance measures, most often the Nash-Sutcliffe index applied to streamflow series alone, i.e., omitting more robust uncertainty analysis and cross validation using different data types. Even the range of Nash-Sutcliffe performance values reported in applications of a single fixed model to multiple basins is often discouraging (e.g., *Perrin et al.* [2001] report ranges from 0.0 to 0.9).

[68] On the other hand, previous analyses of different models in different basins have offered limited insights in terms of process understanding (e.g., see the critique by *Dunn et al.* [2008]). In particular, intercomparison studies such as MOPEX [*Duan et al.*, 2006] have not found a single "best" model structure that outperformed competing alternatives under all conditions and have been inconclusive in examining why this happens, how to account for it, and what can be learned in terms of process representation (but note that their primary objective was to determine strategies for a priori parameter estimation, rather than diagnose intermodel and intercatchment differences). As argued by *Clark et al.* [2011b], the lack of reliable generalizable insights is, at least partially, a consequence of multiple uncontrolled differences in the model structure (including both process representation and numerical implementation), objective function, and calibration algorithms, which make it impossible to unambiguously disaggregate the influence of these distinct factors.

[69] The recent study of *Krueger et al.* [2010] provided important insights into model hypothesis testing using the generalized likelihood uncertainty estimation (GLUE) method. Given its focus on small areas (six fields of about 100 m^2 in the Rowden Experimental Research Platform, United Kingdom), the comparison focused on a single-bucket model where alternative runoff and base flow components were evaluated. This yields important insights into plot-scale variability, yet these insights do not seem immediately transferable to much larger catchments, to catchments with different climatology, or to catchment with different geomorphology.

[70] It is clear that rigorously understanding and quantifying "uniqueness of place" is a difficult task. However, we anticipate that comparison studies exploiting advances in data collection as well as more robust and informative modeling and analysis tools are likely to bring progress in this respect. The case study in the companion paper suggests that depending on their specific characteristics, different catchments may require different conceptual model representations. More importantly, it indicates a likely connection between catchment-scale properties and appropriate model structure. While further research is clearly needed, flexible frameworks combined with a fieldwork-based understanding of processes may provide a more systematic basis for comparing distinct hydrological model hypotheses for distinct types of catchment systems.

5. Conclusions

[71] In this paper, we propose a flexible model framework for conceptual hydrological modeling to explore important contemporary challenges of catchment-scale hydrology. Within this framework, model structures are hypothesized and constructed using generic components such as reservoirs and lag functions, assembled (connected) into a coupled system model using junctions and fluxes, and parameterized using constitutive functions relating internal states and fluxes.

[72] Importantly, all model components are built with attention toward important mathematical aspects of model design, including the use of robust numerical approximations and numerically smooth constitutive functions (even when representing threshold behavior). This prevents avoidable numerical artifacts from corrupting the model equations and obscuring the model analysis, interpretation, and predictive use.

[73] The flexible framework proposed here aims to organize and systematize the fragmented field of conceptual modeling and provides a platform for more systematic and robust hypothesis testing. In addition to their ability to reproduce many existing models within a single mathematically consistent and robust framework, a flexible framework can be used to generate a large variety of alternative hypotheses describing different catchment functions and their connectivities within the overall system architecture. Importantly, by operating within a single framework, hypothesis testing can proceed unobscured by uncontrolled and/or unaccounted differences in overall model philosophy and/or component implementation and software, all of which can arise when simply combining together disparate and separately developed process modules. In more pragmatic applications, the flexibility of the framework permits tailoring, in a controlled way, specific model configurations to specific modeling contexts (e.g., specific catchments, specific data availability, etc.).

[74] We stress that the generation and testing of multiple model alternatives should be carefully controlled. For example, while it is reasonable to hypothesize that catchments with different hydrological regimes may require

different model structures, it is equally important to avoid spuriously "overfitting" a model structure. In this respect, the application of a flexible modeling framework is subject to the same advantages and limitations with respect to justifiable complexity as any other modeling endeavor.

[75] The fixed model structure, on the other hand, may be advantageous in many operational contexts, in particular, when there are insufficient time and/or human resources to set up and trial multiple model structures. Yet even for proponents of fixed model structures, a flexible framework offers a much more systematic platform for the ongoing model comparison and refinement needed for improving a single fixed model. Indeed, essentially regardless of their research or operational aims and philosophical motivation, large-scale multimodel and multicatchment studies benefit from a controlled way of trialing different competing hypotheses and representations and improving them as new data and independent insights become available.

[76] Flexible modeling frameworks are relatively new in catchment-scale hydrology, and their advantages and limitations remain to be investigated in more comprehensive applications. As argued here and in the companion paper, they hold significant promise for addressing important contemporary challenges, including incorporating fieldwork insights as part of ongoing dialogue between modelers and experimentalists, exploring potential theoretical relationships between catchment properties, climatology, and conceptual model structure and complexity, more stringently examining the concept of "uniqueness of place," and pursuing more unified theories of catchment function at the catchment scale.

[77] **Acknowledgments.** We thank Martyn Clark for his feedback on an earlier version of the manuscript. We also thank our colleagues, including Vazken Andréassian, Keith Beven, Jim Freer, Hoshin Gupta, George Kuczera, Jeff McDonnell, Charles Perrin, Laurent Pfister, Jérôme Juilleret, Gerrit Schoups, Murugesu Sivapalan, Mark Thyer, and Peter Young, for valuable discussions on the topic of this work. We also thank the five anonymous reviewers for their valuable comments, criticisms, and suggestions. This work was funded by the National Research Fund of Luxembourg (grant FOR 1598 CAOS: "From Catchments as Organized Systems to Models based on Dynamic Functional Units" and AM2c grants for the Mobility of Researchers). We also thank the STITPRO Foundation for financial support.

References

Abbot, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen (1986), An introduction to the European hydrological system–Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system, *J. Hydrol.*, *87*, 61–77.

Ajami, N. K., H. Gupta, T. Wagener, and S. Sorooshian (2004), Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *J. Hydrol.*, *298*, 112–135.

Ambroise, B., K. Beven, and J. Freer (1996), Toward a generalization of the TOPMODEL concepts: Topographic indices of hydrological similarity, *Water Resour. Res.*, *32*(7), 2135–2145, doi:10.1029/95WR03716.

Andréassian, V., C. Perrin, L. Berthet, N. Le Moine, J. Lerat, C. Loumagne, L. Oudin, T. Mathevet, M.-H. Ramos, and A. Valery (2009), Crash tests for standardized evaluation of hydrological models, *Hydrol. Earth Syst. Sci.*, *13*, 1757–1764.

Berthet, L., V. Andréassian, C. Perrin, and P. Javelle (2009), How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments, *Hydrol. Earth Syst. Sci.*, *13*, 819–831.

Beven, K. (1997), TOPMODEL: A critique, *Hydrol. Processes*, *11*, 1069–1085.

Beven, K. J. (2000), Uniqueness of place and process representations in hydrological modelling, *Hydrol. Earth Syst. Sci.*, *4*, 203–213.

Beven, K. J. (2001), *Rainfall-Runoff Modelling: The Primer*, John Wiley, Chichester, U. K.

Beven, K. (2004), Robert E. Horton's perceptual model of infiltration processes, *Hydrol. Processes*, *18*, 3447–3460.

Beven, K. (2008), On doing better hydrological science, *Hydrol. Processes*, *22*, 3549–3553.

Beven, K., and J. Freer (2001), A dynamic TOPMODEL, *Hydrol. Processes*, *15*(10), 1993–2011.

Beven, K., and F. Pappenberger (2003), Discussion of "Towards the hydraulic of the hydroinformatics era," *J. Hydraul. Res.*, *41*(3), 331–336.

Bulygina, N., and H. Gupta (2009), Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, *Water Resour. Res.*, *45*, W00B13, doi:10.1029/2007WR006749.

Burnash, R. J. C. (1995), The NWS river forecast system—Catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.

Butcher, J. (2008), *Numerical Methods for Ordinary Differential Equations*, 2nd ed., John Wiley, Chichester, U. K.

Butts, M. B., J. T. Payne, M. Kristensen, and H. Madsen (2004), An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation, *J. Hydrol.*, *298*, 242–266.

Clark, M. P., and D. Kavetski (2010), Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes, *Water Resour. Res.*, *46*, W10510, doi:10.1029/2009WR008894.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.

Clark, M. P., D. E. Rupp, R. A. Woods, H. J. Tromp-van Meerveld, N. E. Peters, and J. E. Freer (2009), Consistency between hydrological models and field observations: Linking processes at the hillslope scale to hydrological responses at the watershed scale, *Hydrol. Processes*, *23*, 311–319.

Clark, M. P., H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods (2011a), Hydrological field data from a modeller's perspective. Part 2: Process-based evaluation of model hypotheses, *Hydrol. Processes*, *25*, 523–543.

Clark, M. P., D. Kavetski, and F. Fenicia (2011b), Pursuing the method of multiple working hypotheses in hydrological modeling, *Water Resour. Res.*, *47*, W09301, doi:10.1029/2010WR009827.

Croke, B. F. W., W. S. Merritt, and A. J. Jakeman (2004), A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments, *J. Hydrol.*, *291*, 115–131.

Cuo, L., D. P. Lettenmaier, B. V. Mattheussen, P. Storck, and M. Wiley (2008), Hydrologic prediction for urban watersheds with the Distributed Hydrology-Soil-Vegetation Model, *Hydrol. Processes*, *22*(21), 4205–4213, doi:10.1002/Hyp.7023.

Duan, Q., et al. (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*, 3–17.

Dunn, S. M., J. Freer, M. Weiler, M. J. Kirkby, J. Seibert, P. F. Quinn, G. Lischeid, D. Tetzlaff, and C. Soulsby (2008), Conceptualization in catchment modelling: Simply learning?, *Hydrol. Processes*, *22*(13), 2389–2393.

Edijatno, and C. Michel (1989), Un modele pluie-debit journalier a trois parametres, *Houille Blanche*, *2*, 113–121.

Edijatno, N. De Oliveira Nascimento, X. Yang, Z. Makhlouf, and C. Michel (1999), GR3J: A daily watershed model with three free parameters, *Hydrol. Sci. J.*, *44*(2), 263–277.

Fenicia, F., H. H. G. Savenije, P. Matgen, and L. Pfister (2007), A comparison of alternative multiobjective calibration strategies for hydrological modeling, *Water Resour. Res.*, *43*, W03434, doi:10.1029/2006WR005098.

Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008a), Learning from model improvement: On the contribution of complementary data to process understanding, *Water Resour. Res.*, *44*, W06419, doi:10.1029/2007WR006386.

Fenicia, F., H. H. G. Savenije, P. Matgen, and L. Pfister (2008b), Understanding catchment behavior through stepwise model concept improvement, *Water Resour. Res.*, *44*, W01402, doi:10.1029/2006WR005563.

Fenicia, F., S. Wrede, D. Kavetski, L. Pfister, L. Hoffmann, H. Savenije, and J. J. McDonnell (2010), Impact of mixing assumptions on mean residence time estimation, *Hydrol. Processes*, *24*, 1730–1741.

Freer, J. E., H. McMillan, J. J. McDonnell, and K. J. Beven (2004), Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *J. Hydrol.*, *291*, 254–277.

Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, *Hydrol. Processes*, *22*, 3802–3813.

Henderson-Sellers, A., Z. L. Yang, and R. E. Dickinson (1993), The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS), *Bull. Am. Meteorol. Soc.*, *74*, 1335–1349.

Hrachowitz, M., C. Soulsby, D. Tetzlaff, I. A. Malcolm, and G. Schoups (2010), Gamma distribution models for transit time estimation in catchments: Physical interpretation of parameters and implications for time-variant transit time assessment, *Water Resour. Res.*, *46*, W10536, doi:10.1029/2010WR009148.

Ibbitt, R. P., and T. O'Donnell (1971), Designing conceptual catchment models for automatic fitting methods, *Int. Assoc. Sci. Hydrol. Publ.*, *101*(2), 461–475.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model?, *Water Resour. Res.*, *29*(8), 2637–2649, doi:10.1029/93WR00877.

Kahaner, D., C. Moler, and S. Nash (1989), *Numerical Methods and Software*, Prentice-Hall, Englewood Cliffs, N. J.

Kavetski, D., and M. P. Clark (2010), Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction, *Water Resour. Res.*, *46*, W10511, doi:10.1029/2009WR008896.

Kavetski, D., and F. Fenicia (2011), Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights, *Water Resour. Res.*, *47*, W11511, doi:10.1029/2011WR010748.

Kavetski, D., and G. Kuczera (2007), Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration, *Water Resour. Res.*, *43*, W03411, doi:10.1029/2006WR005195.

Kavetski, D., G. Kuczera, and S. W. Franks (2003), Semidistributed hydrological modeling: A "saturation path'' perspective on TOPMODEL and VIC, *Water Resour. Res.*, *39*(9), 1246, doi:10.1029/2003WR002122.

Kavetski, D., G. Kuczera, and S. W. Franks (2006), Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts, *J. Hydrol.*, *320*, 173–186.

Kavetski, D., F. Fenicia, and M. P. Clark (2011), Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment, *Water Resour. Res.*, *47*, W05501, doi:10.1029/2010WR009525.

Kingston, G. B., M. F. Lambert, and H. R. Maier (2005), Bayesian training of artificial neural networks used for water resources modeling, *Water Resour. Res.*, *41*, W12409, doi:10.1029/2005WR004152.

Kingston, G. B., H. R. Maier, and M. F. Lambert (2008), Bayesian model selection applied to artificial neural networks used for water resources modeling, *Water Resour. Res.*, *44*, W04419, doi:10.1029/2007WR006155.

Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, *42*, W03S04, doi:10.1029/2005WR004362.

Kirchner, J. W. (2009), Catchments as simple dynamic systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, *Water Resour. Res.*, *45*, W02429, doi:10.1029/2008WR006912.

Koutsoyiannis, D. (2000), Broken line smoothing: A simple method for interpolating and smoothing data series, *Environ. Modell. Software*, *15*, 139–149.

Krueger, T., J. Freer, J. N. Quinton, C. J. A. Macleod, G. S. Bilotta, R. E. Brazier, P. Butler, and P. M. Haygarth (2010), Ensemble evaluation of hydrological model hypotheses, *Water Resour. Res.*, *46*, W07516, doi:10.1029/2009WR007845.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters, *J. Hydrol.*, *331*, 161–177.

Kumar, S. V., et al. (2006), Land information system: An interoperable framework for high resolution land surface modeling, *Environ. Modell. Software*, *21*, 1402–1415.

Lamb, R., and K. Beven (1997), Using interactive recession curve analysis to specify a general catchment storage model, *Hydrol. Earth Syst. Sci.*, *1*, 101–113.

Lambert, J. D. (1991), *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*, John Wiley, Chichester, U. K.

Leavesley, G. H., P. J. Restrepo, S. L. Markstrom, M. Dixon, and L. G. Stannard (1996), The Modular Modeling System—MMS, User's Manual, *U.S. Geol. Surv. Open File Rep.*, 96–151, Denver, Col.

Leavesley, G. H., S. L. Markstrom, P. J. Restrepo, and R. J. Viger (2002), A modular approach to addressing model design, scale, and parameter estimation issues in distributed hydrological modelling, *Hydrol. Processes*, *16*, 173–187.

Le Moine, N., V. Andréassian, C. Perrin, and C. Michel (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resour. Res.*, *43*, W06428, doi:10.1029/2006WR005608.

Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström (1997), Development and test of the distributed HBV-96 hydrological model, *J. Hydrol.*, *201*, 272–288.

Linsley, R. K. (1982), Rainfall-runoff models: An overview, in *Proceedings of the International Symposium of Rainfall-Runoff Modelling*, edited by V. P. Singh, pp. 3–22, Water Resour. Publ., Littleton, Colo.

McDonnell, J. J., and R. Woods (2004), On the need for catchment classification, *J. Hydrol.*, *299*, 2–3.

McDonnell, J. J., et al. (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resour. Res.*, *43*, W07301, doi:10.1029/2006WR005467.

McGuire, K. J., and J. J. McDonnell (2010), Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities, *Water Resour. Res.*, *46*, W10543, doi:10.1029/2010WR009341.

Michel, C., C. Perrin, and V. Andréassian (2003), The exponential store: A correct formulation for rainfall-runoff modelling, *Hydrol. Sci. J.*, *48*(1), 109–124.

Michel, C., V. Andréassian, and C. Perrin (2005), Soil Conservation Service curve number method: How to mend a wrong soil moisture accounting procedure, *Water Resour. Res.*, *41*, W02011, doi:10.1029/2004WR003191.

Moore, R. J. (1985), The probability-distributed principle and runoff production at point and basin scales, *Hydrol. Sci. J.*, *30*(2), 273–297.

Moore, R. J., and R. T. Clarke (1981), A distribution function approach to rainfall runoff modeling, *Water Resour. Res.*, *17*(5), 1367–1382, doi:10.1029/WR017i005p01367.

Nash, J. E. (1958), The form of the instantaneous unit hydrograph, *Bull. Int. Assoc. Hydrol. Sci.*, *3*(45), 114–121.

Oppenheim, A. V., R. W. Schafer, and J. R. Buck (1999), *Discrete-Time Signal Processing*, 2nd ed., Prentice Hall, Upper Saddle River, N. J.

Pappenberger, F., and K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resour. Res.*, *42*, W05302, doi:10.1029/2005WR004820.

Perrin, C., C. Michel, and V. Andréassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, *242*, 275–301.

Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, *279*, 275–289.

Pomeroy, J. W., D. M. Gray, T. Brown, N. R. Hedstrom, W. L. Quinton, R. J. Granger, and S. K. Carey (2007), The cold regions hydrological model: A platform for basing process representation and model structure on physical evidence, *Hydrol. Processes*, *21*, 2650–2667.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992), *Numerical Recipes in Fortran-77: The Art of Scientific Computing*, 820 pp., Cambridge Univ. Press, Cambridge, U. K.

Pringle, C. (2003), What is hydrologic connectivity and why is it ecologically important?, *Hydrol. Processes*, *17*, 2685–2689.

Refsgaard, J. C., and H. J. Henriksen (2004), Modelling guidelines—Terminology and guiding principles, *Adv. Water Resour.*, *27*, 71–82.

Reid, S. C., S. N. Lane, D. R. Montgomery, and C. J. Brookes (2007), Does hydrological connectivity improve modelling of coarse sediment delivery in upland environments?, *Geomorphology*, *90*(3–4), 263–282.

Renard, B., D. Kavetski, M. Thyer, G. Kuczera, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, *46*, W05521, doi:10.1029/2009WR008328.

Rupp, D. E., and R. A. Woods (2008), Increased flexibility in base flow modelling using a power law transmissivity profile, *Hydrol. Processes*, *22*, 2667–2671.

Savenije, H. H. G. (2009), The art of hydrology, *Hydrol. Earth Syst. Sci.*, *13*, 157–161.

Schoups, G., N. C. van de Giesen, and H. H. G. Savenije (2008), Model complexity control for hydrologic prediction, *Water Resour. Res.*, *44*, W00B03, doi:10.1029/2008WR006836.

Schoups, G., J. A. Vrugt, F. Fenicia, and N. C. van de Giesen (2010), Corruption of accuracy and efficiency of Markov chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models, *Water Resour. Res.*, *46*, W10530, doi:10.1029/2009WR008648.

Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, *38*(11), 1241, doi:10.1029/2001WR000978.

Shampine, L. F., and M. W. Reichelt (1997), The MATLAB ODE suite, *SIAM J. Sci. Comput.*, *18*(1), 1–22.

Singh, V. P., and D. A. Woolhiser (2002), Mathematical modeling of watershed hydrology, *J. Hydrol. Eng.*, *7*(4), 270–292.

Sivapalan, M. (2005), Pattern, process and function: Elements of a new unified hydrologic theory at the catchment scale, in *Encyclopedia of Hydrological Sciences*, vol. 13, edited by M. G. Anderson, pp. 193–219, John Wiley, Hoboken, N. J.

Sivapalan, M. (2009), The secret to 'doing better hydrological science': Change the question!, *Hydrol. Processes*, *23*, 1391–1396.

Smith, T. J., and L. A. Marshall (2010), Exploring uncertainty and model predictive performance concepts via a modular snowmelt-runoff modeling framework, *Environ. Modell. Software*, *25*, 691–701.

Spence, C. (2010), A paradigm shift in hydrology: Storage thresholds across scales influence catchment runoff generation, *Geogr. Compass*, *4*, 819–833.

Steefel, C. I., and K. T. B. MacQuarrie (1996), Approaches to modeling of reactive transport in porous media, in *Reactive Transport in Porous Media*, edited by P. C. Lichtner et al., pp. 83–125, Mineral. Soc. of Am., Washington, D. C.

Sugawara, M. (1995), Tank model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 165–214, Water Resour. Publ., Highlands Ranch, Colo.

Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resour. Res.*, *45*, W00B14, doi:10.1029/2008WR006825.

Troch, P. A., G. A. Carrilo, I. Heidbuchel, R. S., M. Switanek, T. H. M. Volkman, and M. Yaeger (2009), Dealing with landscape heterogeneity in watershed hydrology: A review of recent progress toward new hydrological theory, *Geogr. Compass*, *3*, 395–392.

Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe (1999), Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrol. Sci. J.*, *44*(5), 779–797.

Uhlenbrook, S., S. Roser, and N. Tilch (2004), Hydrological process representation at the meso-scale: The potential of a distributed, conceptual catchment model, *J. Hydrol.*, *291*, 278–296.

Vaché, K. B., and J. J. McDonnell (2006), A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resour. Res.*, *42*, W02409, doi:10.1029/2005WR004247.

Vidyasagar, M. (2002), *Nonlinear Systems Analysis*, 2nd ed., Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian (2001), A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*, 13–26.

Wagener, T., M. J. Lees, and H. S. Wheater (2002), A toolkit for the development and application of parsimonious hydrological models, in *Mathematical Models of Small Watershed Hydrology*, vol. 2, edited by V. P. Singh et al., pp. 91–139, Water Resour. Publ., Highlands Ranch, Colo.

Wagener, T., M. Sivapalan, P. Troch, and R. A. Woods (2007), Catchment classification and hydrologic similarity, *Geogr. Compass*, *1*, 901–931.

Weiler, M., B. L. McGlynn, K. J. McGuire, and J. J. McDonnell (2003), How does rainfall become runoff? A combined tracer and runoff transfer function approach, *Water Resour. Res.*, *39*(11), 1315, doi:10.1029/2003WR002331.

Westhoff, M. C., H. H. G. Savenije, W. M. J. Luxemburg, G. S. Stelling, and L. Pfister (2007), A distributed temperature model using high resolution temperature observations, *Hydrol. Earth Syst. Sci.*, *11*, 1469–1480.

Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian (1992), A land-surface hydrology parameterization with subgrid variability for general circulation models, *J. Geophys. Res.*, *97*(D3), 2717–2728, doi:10.1029/91JD01786.

Yilmaz, K. K., H. V. Gupta, and T. Wagener (2008), A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, *Water Resour. Res.*, *44*, W09417, doi:10.1029/2007WR006716.

Young, P. (1998), Data-based mechanistic modelling of environmental, ecological, economic and engineering systems, *Environ. Modell. Software*, *13*, 105–122.

Young, P. (2003), Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale, *Hydrol. Processes*, *17*, 2195–2217.

Young, P. C. (2002), Advances in real-time flood forecasting, *Philos. Trans. R. Soc., Ser. A*, *360*(1796), 1433–1450.

Young, P. C. (2005), Rainfall-runoff modeling: Transfer function modeling, in *Encyclopedia of Hydrological Sciences*, vol. 3, part 2, edited by M. G. Anderson, pp. 1985–2000, John Wiley, Hoboken, N. J.

Young, P. C., and H. Garnier (2006), Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems, *Environ. Modell. Software*, *21*, 1055–1072.

Young, P. C., and M. Ratto (2009), A unified approach to environmental systems modeling, *Stochastic Environ. Res. Risk Assess.*, *23*(7), 1037–1057.

Zehe, E., and M. Sivapalan (2009), Threshold behaviour in hydrological systems as (human) geo-ecosystems: Manifestations, controls, implications, *Hydrol. Earth Syst. Sci.*, *13*, 1273–1297.

F. Fenicia, Department of Environment and Agro-Biotechnologies, Centre de Recherche Public–Gabriel Lippmann, rue du Brill 41, L-4422 Belvaux, Grand-Duchy of Luxembourg. (fenicia@lippmann.lu)

D. Kavetski, Environmental Engineering, University of Newcastle, University Dr., Callaghan, NSW 2308, Australia.

H. H. G. Savenije, Water Resources Section, Delft University of Technology, Stevinweg 1, 2600 GA Delft, Netherlands.