INVITED COMMENTARY

HYDROLOGICAL PROCESSES
*Hydrol. Process.* **20**, 2887–2900 (2006)
DOI: 10.1002/hyp.6388

# Physics-based hydrologic-response simulation: Seeing through the fog of equifinality

Brian A. Ebel and
Keith Loague*

*Department of Geological and Environmental Sciences Stanford University, Stanford, CA 94305-2115, USA*

*Correspondence to:
Keith Loague, Department of Geological and Environmental Sciences Stanford University,
Stanford, CA 94305-2115, USA.
E-mail: keith@pangea.stanford.edu

**Models are to be used, not to be believed in!**
**Dooge (1972)**

**Our models should be designed expressly to maximize the possibility of discovering that of which we are ignorant.**
**Beck (2002)**

## Equifinality

**The need to characterize the distributed nature of near-surface hydrologic response has been stressed recently by Burt (2005), Bloschl (2006), and Sidle (2006). Serendipitously, 31 classic papers related to the processes governing streamflow generation have recently been reprinted in the *Benchmark Papers in Hydrology* series (Beven, 2006a). The limitations and benefits of comprehensive physics-based hydrologic-response simulation were discussed in two earlier commentaries (Loague and VanderKwaak, 2004; Loague *et al.*, 2006). Herein, the topic of *equifinality*, a perceived Achilles' heel for physics-based simulation, is addressed.**

**In the most general sense, equifinality is the case where different conditions lead to similar effects. Beven (2006b) thoroughly reviews the equifinality problem for hydrology and related disciplines (also see Haines-Young and Petch, 1983; Culling, 1987; Aronica *et al.*, 1998; Hankin and Beven, 1998; Schulz *et al.*, 1999; Beven, 2000, 2002a,b; Brazier *et al.*, 2000; Beven and Freer, 2001; Hreiche *et al.*, 2002; Binley and Beven, 2003). It is worth noting that the equifinality problem is not limited to the earth sciences. For example, equifinality is well-known in biology (Rogers, 2000), management (Doty *et al.*, 1993; Gresov and Drazin, 1997), and psychology (Gottlieb, 2001; McClure *et al.*, 2001).**

**In the context of hydrology and deterministic-conceptual simulation, equifinality frequently refers to more than one parameter set providing an equally good (or poor) representation of the integrated response. Discussing the issue of equifinality, Savenije (2001) wrote as follows: *Although we can consider equifinality as a nuisance since it implies that looking for more understanding through detailed distributed modelling is a dead-end track, it also offers an opening to the revival of larger-scale hydrological laws*. Commenting on the statement by Savenije (2001), Loague and VanderKwaak (2004) wrote: *Obviously, one should consider the intended use of a distributed model before characterizing the approach as a dead-end track (i.e. large***

*Received 28 April 2006*
*Accepted 24 May 2006*

***scales are not always the focus). It is important to remember that the space of possible parameter values (and their combinations), which are not independent, can be constrained in a physics-based model. Without the physics, there are no such constraints, and any combination is valid.*** **The objective of this commentary is to expand upon the brief remarks by Loague and VanderKwaak (2004) on the issue of equifinality, as the problem relates to comprehensive physics-based hydrologic-response simulation. An illustrative example, gleaned from recent physics-based simulations for a site with exhaustive field data, is presented, which addresses components of the equifinality problem.**

## Ill-Posed Problems and Model Performance

**Near-surface hydrologic response can be quantitatively analyzed for certain problems and scales, using physics-based mathematical models with meaningful (measurable) parameter estimates that are both spatially and temporally variable. Parameterization of a physics-based model, especially with sparse information, is no trivial task (see Bredehoeft, 2005). The inverse method of model parameterization, which involves determining the optimal parameter set by observing the dependent variable in the spatial and time domains, has been discussed by Neuman (1973), Carrera and Neuman (1986), Yeh (1986), and several others. When the number of unknowns is greater than the number of observations, the problem becomes *ill-posed* (Hadamard, 1902). For a given parameter set, the solution to an ill-posed problem is not necessarily unique or stable. Most, if not all, near-surface hydrologic-response problems are ill-posed. However, a problem being ill-posed does not imply, as pointed out by Allison (1979), that it is not worth solving.**

**The issues that are associated with ill-posed problems in hydrology fall into three categories (Neuman, 1973; Carrera and Neuman, 1986; Khatibi *et al*., 2000): (i) *identifiability*, (ii) *uniqueness*, and (iii) *stability*. Identifiability requires that all the model parameters contribute meaningfully to the solution and that there is only one correct parameterization for a given site. Uniqueness requires that only one set of parameter values can be estimated from a given set of observations and that the set of parameters determined from one set of observations (from one time period) must also represent the observed behaviour at a different time period. While identifiability and uniqueness seem very similar, Carrera and Neuman (1986) distinguish between the two by associating identifiability with the forward solution and uniqueness with the inversion from observed data to model parameterization. The stability constraint demands that small errors in the observed data do not produce large changes in the calibration. Stability can also be taken to mean that slight changes in parameter values do not cause the model solution to be drastically different.**

**The problems that are associated with identifiability, uniqueness, and stability are not new. Examples include studies from engineering (e.g. Chavent, 1974), geology (e.g. Bird and Rosenstock, 1984), geophysics (e.g. Roy, 1962; Anderssen, 1968), hydrogeology (e.g. Cooley and Sinclair, 1976; McElwee, 1982; Konikow, 1986), physics (e.g. Aktosun and Newton, 1985), and rainfall-runoff modelling (e.g. Gupta and Sorooshian, 1983). Because data cannot be realistically obtained from every point within the spatial domain of the subsurface without excavating the entire site (e.g. Wierenga *et al*., 1989), identifiability is essentially unachievable when comparing distributed parameters against point measurements (Chavent, 1974; Yeh, 1986). Perhaps equally challenging, as pointed out by Beck (2002), is the fact that many physically based models have an abundance of unobserved state variables (e.g. total head). For regional-scale problems with large elements, the parameters and state variables at the grid points often do not correspond with field-based estimates or observations, the *incommensurability* problem (Beven, 1989). Additional problems with physics-based modelling include model structural error, data and measurement errors, and imperfect representation of physical processes (Sorooshian and Gupta, 1983). All the issues associated with ill-posed problems in hydrologic-response simulation have been lumped together and termed equifinality. One of the tenets of equifinality is that many different parameter sets provide *behavioural* models (i.e. fit the data to an acceptable level), violating the principles of identifiability and uniqueness.**

One explanation for the equifinality problem, in physics-based simulation, is the nuances associated with identifiability and uniqueness that result from non-robust parameter estimation. The literature on parameter estimation techniques for distributed hydrologic systems is rich (e.g. Leliavsky, 1965; Yeh, 1986; Zimmerman *et al.*, 1998). However, parameter estimation is most often associated with the process of model calibration. It is well known that the calibration of deterministic-conceptual models is rife with problems (Neuman and Guadagnini, 1999). One calibration issue, paradoxically linked to the equifinality problem, is the effect of increased parameter space, resulting from greater levels of model complexity. As noted by Yeh (1986), increasing the degrees of freedom for the parameter space will generally reduce the simulation error but increase the parameter uncertainty. It is relatively common for the number of parameters to exceed what the field measurements justify (Kirkby, 1975). The optimum level of model parameterization depends on both the quantity and quality of the data. Obviously, some balance between process representation and parameter parsimony is necessary.

There are many valuable approaches for examining hydrologic processes. However, at the catchment scale, the most sensible framework of investigation is the *measure and model* protocol, pioneered by Robert E. Horton (see Hall, 1987; Beven, 2004a,b,c,d). A fundamental principle behind Horton's approach is the rigorous employment of high-quality observed data to develop, parameterize, and evaluate mathematical hydrologic-response models. It follows that the trade-off associated with using increasingly complex models is the need for richer data sets.

For almost all hydrologic-response simulation efforts, the performance of the model is judged relative to some observation. This protocol is not without shortcomings, such as the incommensurability problem. There is, of course, a large difference between a *better fit* (to observed data) for a given model and a *better model* (see Sorooshian *et al.*, 1983; Klemes, 1997). Several authors have addressed the issue of model performance (e.g. Loague and Green, 1991; Konikow and Bredehoeft, 1992; Oreskes *et al.*, 1994; Pebesma *et al.*, 2005, 2006). However, there are still problems with quantitatively testing physics-based models, especially when the uncertainties associated with initial and boundary conditions are considered (see Stephenson and Freeze, 1974). For the example given below, both the Nash and Sutcliffe (1970) model efficiency statistic and observed versus simulated time-series plots are used to discuss (rank) model performance. We do not judge the acceptability of a simulation on the basis of a model performance cutoff criterion, as we are unaware of any such standards for hydrologic-response models. The widely employed Nash and Sutcliffe (1970) model efficiency, which is not without its own limitations (see Pebesma *et al.*, 2005, 2006; Heppner *et al.*, 2006b; Loague *et al.*, 2006; Mirus *et al.*, 2006), is calculated as

$$EF = \left[\sum_{i=1}^{n}(O_i - \overline{O})^2 - \sum_{i=1}^{n}(P_i - O_i)^2\right] \Big/ \sum_{i=1}^{n}(O_i - \overline{O})^2 \qquad (1)$$

where $P_i$ are the predicted values, $O_i$ are the observed values, $n$ is the number of samples, and $\overline{O}$ is the mean of the observed data.

## An Example, Coos Bay and the Integrated Hydrology Model

The study site used in this example is the CB1 experimental catchment (see Figure 1(a)), located along Mettman Ridge approximately 15 km north of Coos Bay in the Oregon Coast Range. The 860 $\mathrm{m}^2$ CB1 is an unchanneled hollow with an average slope of 43° (see Figure 1). In 1987 CB1 was clearcut; in 1996 the site experienced a major debris flow. The exhaustive CB1 data sets for three sprinkling experiments (1990 and 1992) and natural events (1990 through 1996) are described by Montgomery (1991), Anderson *et al.* (1993, 1997a,b, 2002), Anderson (1995), Montgomery *et al.* (1997, 2002), Torres (1997), Torres *et al.* (1998), Anderson and Dietrich (2001), Montgomery and Dietrich (2002), and Ebel *et al.* (2006a).

The comprehensive physics-based hydrologic-response model used in this example is InHM (Integrated Hydrology Model), which was designed

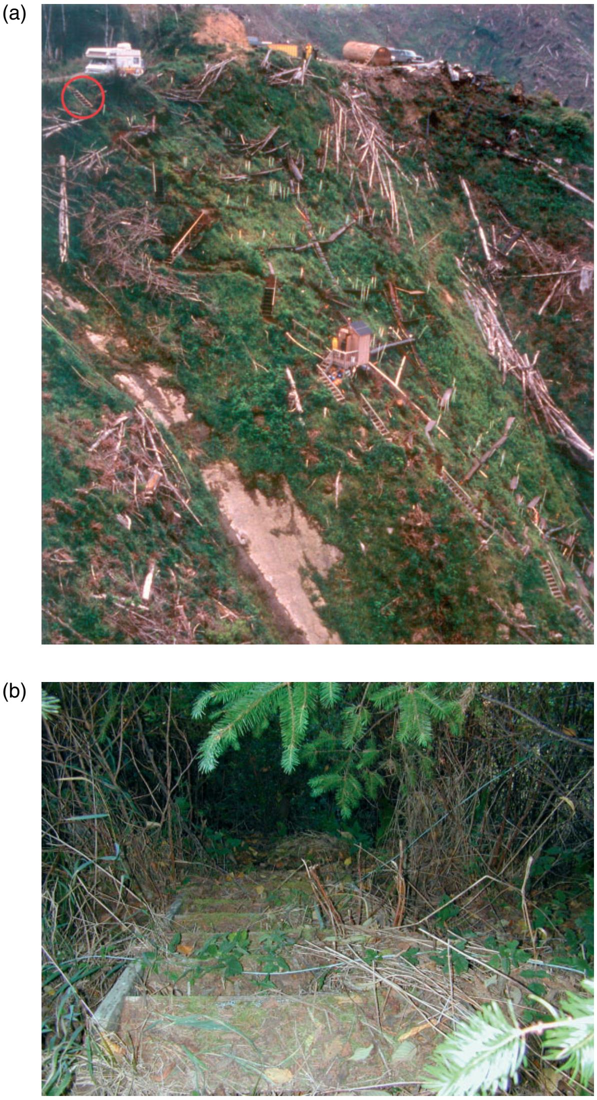

Figure 1. CB1 photographs. (a) 1990, showing most of the catchment, including the access staircase and hanging platforms (note, the east boundary of the catchment is to the right of the staircase in the photo). (b) 2003, showing the overgrown staircase at the top of the catchment [identified by the red circle in (a)]

and tested to quantitatively simulate, in a fully-coupled approach, 3D variably saturated flow and solute transport in porous media and macropores, and 2D flow and solute/sediment transport over the surface and in open channels (VanderKwaak, 1999; Heppner *et al.*, 2006a). For the example reported herein, the macropore, solute transport, and sediment transport components of InHM were not employed. The characteristics and past applications of InHM are discussed by VanderKwaak (1999), VanderKwaak and Loague (2001), Loague and VanderKwaak (2002, 2004), Loague *et al.* (2005, 2006), Heppner *et al.* (2006a, b), and Jones *et al.* (2006).

Recently, Ebel *et al.* (2006b) simulated the CB1 sprinkling experiments with InHM. Here, we report both integrated (discharge) and distributed (tensiometers) response results from five of the approximately 200 *preliminary* 3D transient simulations for sprinkling experiment number 3,

**where the sprinkling rate was approximately 1·7 mm/hr. Three tensiometers (out of 101) were selected to represent the distributed response for this example. The simulation results shown herein, compared with the pressure-head data from the selected tensiometers, are neither the best nor the worst from the complete analysis (Ebel *et al.*, 2006b).**

**Table I summarizes the differences between the model parameterizations of InHM for CB1 used in the preliminary simulations. Inspection of Table I reveals the range of boundary conditions, initial conditions, and material values that were considered by Ebel *et al.* (2006b). The major difference between the five simulation scenarios presented in this commentary (see Table I) is the soil-water content and permeability characteristic functions for the near-surface colluvial layer. The sensitivity of the simulated hydrologic response to the nonlinear characteristic functions should not be a surprise, on the basis of what has been known (from soil physics) for 75 years (e.g. Richards, 1931).**

**Table II provides model performance results for the integrated CB1 response (discharge *EF* values), on the basis of comparisons of observed versus InHM simulated discharges at the upper weir, for five selected boundary-value problem (BVP) characterizations of experiment number 3. Table II also provides *EF* values for the distributed response (pressure-head *EF* values), on the basis of observed versus InHM simulated pressure heads at three different locations, for the same five scenarios for which the integrated response is considered. The evaluation period for both the discharge and pressure-head data (observed versus simulated) is from 10 AM on May 27th to midnight on June 5th (1992). The locations of the upper weir (i.e. discharge) and the three (9-1C, 6-3A, 3-3B) tensiometers (i.e. pressure head) at CB1 are given in Figure 2. Figure 3 shows the observed and five**

Table I. Summary of the different CB1 boundary-value problems considered by Ebel *et al.* (2006b)

| Characteristic | Scenarios considered |
|---|---|
| Subsurface boundary condition | |
| Upgradient | Impermeable (drainage divide)[a,b,c,d,e] |
| Downgradient | Radiation (explicit flux), radiation (implicit flux), local head (Heppner *et al.*, 2006b)[a,b,c,d,e] |
| Base | Impermeable[a,b,c,d,e]; leaking at unity gradient |
| Sides | Impermeable (drainage divide)[a,b,c,d,e] |
| Surface boundary condition | |
| Upgradient | Impermeable (drainage divide)[a,b,c,d,e] |
| Downgradient | Critical depth (weir)[a,b,c,d,e] |
| Sides | Impermeable (drainage divide)[a,b,c,d,e] |
| Surface | Specified flux (applied sprinkling): spatially/temporally constant, spatially constant/temporally variable, spatially/temporally variable[a,b,c,d,e] |
| Initial condition specification | Drainage from saturation[a], initial head equal to scaled surface elevation, five-month *warmup* simulation using measured precipitation; two-month *warmup* simulation using measured precipitation[b,c,d,e] |
| Hydrogeologic units | All layers/formations constant thickness; spatially variable colluvial soil; spatially variable colluvial soil and saprolite[a,d]; spatially variable colluvial soil, saprolite, and weathered bedrock[b,c,e] |
| Colluvial soil permeability | Spatially constant (varied across measured range)[a,b,c,d,e]; spatially variable |
| Saprolite permeability | Spatially constant (varied across measured range)[a,b,c,d,e] |
| Weathered bedrock permeability | Spatially constant (varied across measured range)[a,b,c,d,e] |
| Bedrock permeability | Spatially constant (varied across measured range)[a,b,c,d,e] |
| Colluvial soil-water retention curves | Sandy loam[b,e], loamy sand[a,d], medium sand[c], coarse sand |

[a−e] Simulation scenarios [note, a–e correspond to scenarios 1–5 in Table II and throughout the text (i.e. a corresponds to scenario 1, b corresponds to scenario 2, c corresponds to scenario 3, d corresponds to scenario 4, e corresponds to scenario 5)].

Table II. Model efficiency (*EF*) results for observed versus InHM simulated discharges and pressure heads for CB1 sprinkling experiment number three

| Simulation scenario[b] | *EF*[a] values | | | |
|---|---|---|---|---|
| | Discharge[c] | Pressure-head | | |
| | | 9-1C[d] | 6-3A | 3-3B |
| 1 | 0·82 | −1[e] | −3 | −13 |
| 2 | 0·78 | −7 | −3 | −3 |
| 3 | 0·76 | 0·84 | 0·89 | 0·8 |
| 4 | 0·74 | −3 | −5 | −17 |
| 5 | 0·66 | −11 | −5 | −5 |

[a] Model efficiency (Nash and Sutcliffe, 1970).
*EF* = 1 when the observed and simulated values (e.g. discharge, pressure head) are identical. *EF* < 1 for catchment scale rainfall-runoff simulations (to the best of our knowledge). *EF* < 0 infers that the simulation is worse than using the observed mean as the estimate.
[b] See Table I for generalized characterizations of the individual boundary-value problems.
[c] See Figure 2 for the location of the upper weir.
[d] See Figure 2 for the locations of tensiometers 9-1C, 6-3A, and 3-3B.
[e] The negative *EF* values are rounded off to the nearest integer.

InHM simulated hydrographs for the third sprinkling experiment. Figure 4 shows the observed and five InHM simulated pressure-head values at the three locations for the third sprinkling experiment.

Inspection of Table II and Figures 3 and 4 helps to see through the fog of equifinality. The hydrographs in Figure 3 clearly show that none of the five scenarios capture all the nuances in the observed discharge data. The best of the five scenarios does not stand out in Figure 3. The discharge *EF* values in Table II, ranging between 0·82 and 0·66, are similar to what has been reported in the literature for physics-based hydrologic-response simulation at the catchment scale (see Loague *et al.*, 2005; Heppner *et al.*, 2006b; Mirus *et al.*, 2006). The discharge *EF* value for scenario 1 is better than the values for scenarios 2 through 5 (see Table II). Therefore, at first blush, scenario 1 could mistakenly be taken as the best of the five scenarios. Notwithstanding the less than perfect model performance, the discharge results presented here demonstrate that the parameterizations for the different scenarios can be described as behavioural. The discharge results presented here also illustrate that equifinality is a problem, supporting the multiple behavioural conclusion that has been reported by past investigators (e.g. Zak and Beven, 1999; Brazier *et al.*, 2000; Beven and Freer, 2001; Feyen *et al.*, 2001), when model performance assessment is based solely on the catchment's integrated response.

The pressure-head *EF* values in Table II are positive only for scenario number 3. In Table II, the general order (from best to worst) for the three tensiometers is 6-3A, 9-1C, and 3-3B. Unlike for the integrated response, where all five scenarios are somewhat reasonable, the distributed response suggests (based upon the *EF* values) that only scenario number 3 is reasonable. For scenario number 3, the *EF* values for the three different tensiometer locations (i.e. 0·84, 0·89, and 0·80) are each greater than the 0·76 discharge *EF* value. The especially low pressure-head *EF* values for tensiometer 3-3B for scenarios 1 and 4 (i.e. −13 and −17, respectively) are traceable to misrepresenting the initial conditions. Figure 4 clearly shows, for the *quasi steady-state* portion of experiment number 3, that all five scenarios underestimate the pressure head at each tensiometer. In Figure 4, the order (from best to worst) for the five scenarios is 3, 1, 4, 2, and 5 for each of the tensiometers, with scenarios 1, 3, and 4 clearly separated from scenarios 2 and 5.

Given both the discharge and pressure-head *EF* values in Table II, one would be hard pressed to justify the acceptability of scenarios 1, 2, 4, or 5. On the basis of both the integrated and distributed results, scenario number 3 is the best

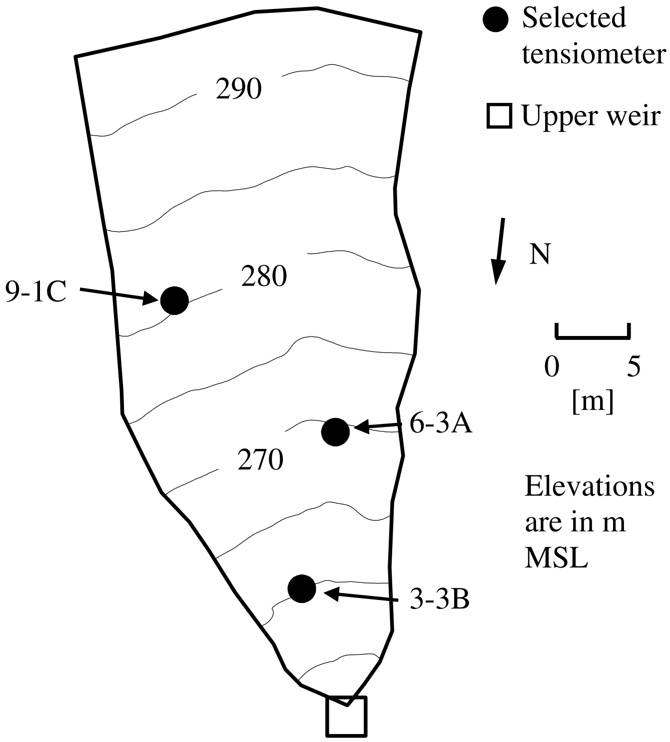

**Figure 2. Locations of the upper weir and tensiometers 9-1C, 6-3A, and 3-3B within CB1. The measurement depths (from the surface) for tensiometers 9-1C, 6-3A, and 3-3B are, respectively, 0·5, 0·7, and 0·5 m**

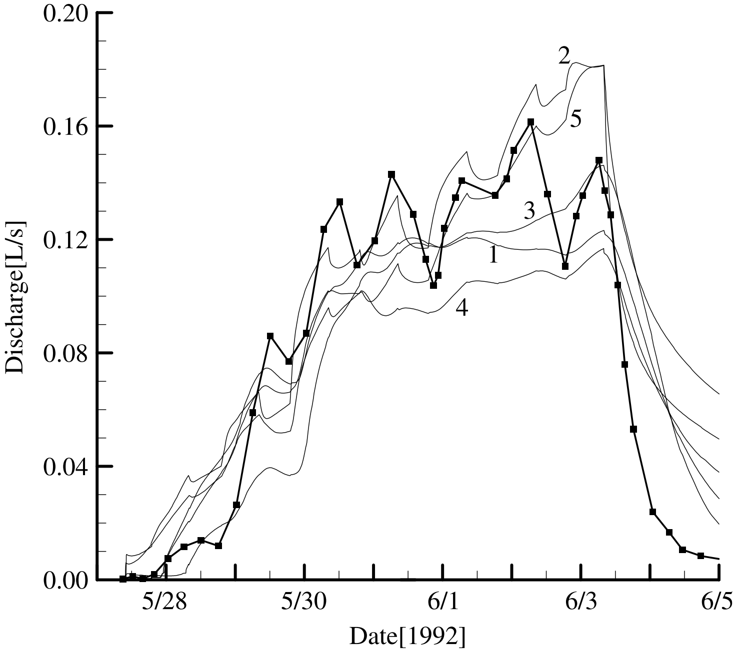

**Figure 3. Observed versus InHM simulated discharges for the third CB1 sprinkling experiment. The observed discharges are shown by the bold line with boxes. The five InHM simulations (numbered lines) correspond to the scenarios in Table II**

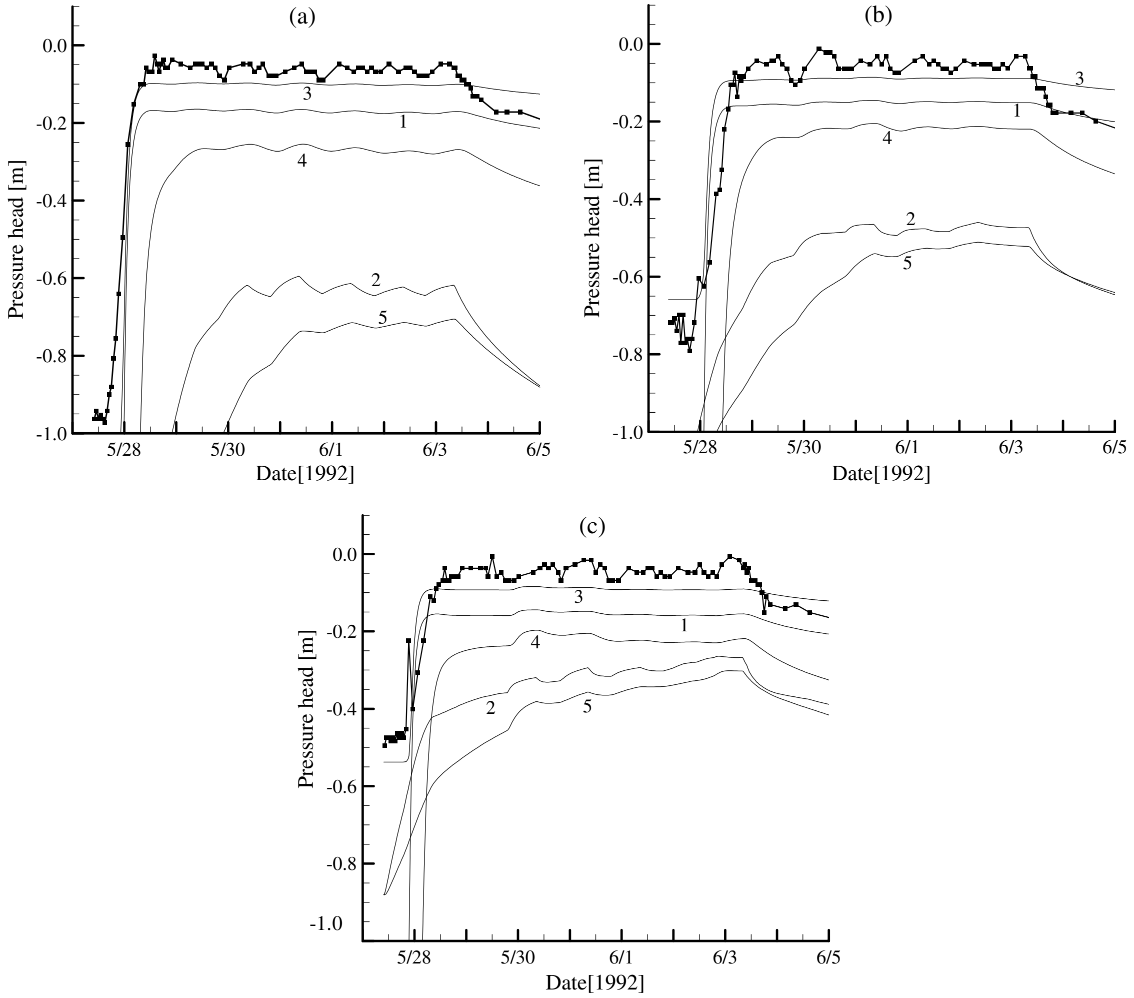

**Figure 4. Observed versus InHM simulated pressure heads for the third CB1 sprinkling experiment. The observed pressure heads are shown by the bold line with boxes. The five InHM simulations (numbered lines) correspond to the scenarios in Table II; simulated pressure heads < − 1·0 are not shown. (a) tensiometer 9-1C. The observed initial condition is −1·0 m; the simulated initial conditions for scenarios 1 through 5 are < − 2, < − 2, −1·1, < − 2, and < − 2 m, respectively. (b) tensiometer 6-3A. The observed initial condition is −0·7 m; the simulated initial conditions for scenarios 1 through 5 are < − 2, −1·3, −0·7, < − 2, and −1·3 m, respectively. (c) tensiometer 3-3B. The observed initial condition is −0·5 m; the simulated initial conditions for scenarios 1 through 5 are < − 2, −0·9, −0·5, < − 2, and −0·9 m, respectively**

at capturing the overall hydrologic response for experiment number 3. Selecting scenario number 1 as the best, on the sole basis of having the highest discharge *EF*, is an example of being *right for the wrong reason* (Loague and VanderKwaak, 2004). The example presented here provides a convincing demonstration (at least for the CB1 case) that the equifinality problem becomes tractable when both integrated and distributed response data are used to assess model performance.

In the straightforward example presented here, each component of the CB1 BVP contributed to the InHM simulated results (both integrated and distributed) for sprinkling experiment number 3. Figure 4 shows that evaluating the simulation results against the distributed pressure-head response helps to ensure that the parameterization is stable. With respect to small changes in the parameterization yielding dramatically different results, the relative sensitivity of individual parameters is important. For example, a small change in the saturated hydraulic conductivity for an impeding layer (e.g. at the soil-bedrock interface) could have a large effect on the integrated response, if the primary runoff-generation mechanism was subsurface stormflow, but have a minimal effect on the distributed response seen in the near-surface pressure-head values.

Field evidence has shown that the time invariance assumption (often accepted with uniqueness) is not realistic for many near-surface soil-hydraulic parameters. For example, permeability can change dramatically both on a seasonal basis (Loague and VanderKwaak, 2004) and due to land-use change (e.g. Mohanty *et al.*, 1994). Figure 1 illustrates the dramatic change in vegetative cover at CB1 between 1990 and 2003. The Douglas Fir (*Pseudotsuga menziesii*) seedlings planted at CB1 in 1989 are now 6 m (or higher), with a thick understory. It should be expected that the vegetation changes shown in Figure 1 have affected the near-surface soil-hydraulic properties.

Previous research has shown that simulations of hydrologic response that are only evaluated against the observed integrated response provide no guarantee of getting the distributed response correct (Ambroise *et al.*, 1996a,b; Beven, 1997; Wigmosta and Lettenmaier, 1999; Saulnier and Datin, 2004). It must also be pointed out that several researchers have found that the use of distributed observations did not appreciably reduce equifinality in their rainfall-runoff modelling efforts (e.g. Lamb *et al.*, 1998; Vertessy and Elsenbeer, 1999; Blazkova and Beven, 2002; Blazkova *et al.*, 2002; Christiaens and Feyen, 2002). However, the example presented here, using a comprehensive physics-based hydrologic-response model and a rich data set, illustrates that the equifinality problem is not insurmountable. It is worth pointing out that there are many more measurements of state variables for the CB1 sprinkling experiments, besides pressure heads at the other 98 tensiometers, that were not employed in the example presented here. For example, the CB1 observations include soil-water contents from 56 TDR waveguide pairs (Torres, 1997), total heads from 223 piezometers (Montgomery *et al.*, 1997), and tracer concentrations from 33 lysimeters (Anderson *et al.*, 1997b). The number of equally-likely parameterizations of InHM for CB1 can easily (but not trivially) be constrained to a much greater degree than has been illustrated here by using all the information within the data base.

## Epilogue

In response to the problems associated with equifinality, there have been many calls for simpler models (e.g. Beven, 1996a,b; Kirkby, 1996; Young *et al.*, 1996). Parsimony in science has been a virtue for some time (e.g. William of Occam advocated it in the fourteenth century). However, some systems are not simple (e.g. linear) enough to justify simple models for all applications. A model cannot adequately capture processes and process interactions that are not included in its formulation. For example, black (or any colour) box models cannot provide process-based diagnostics within the box. Our perception of the best use of physics-based hydrologic-response simulation was lucidly characterized by Kirkby (1996): *Models are thought experiments which help refine our understanding of the dominant processes acting…While most simulation models may be used in a forecasting mode, the most important role of models is as a qualitative thought experiment, testing whether we have a sufficient and consistent theoretical explanation of physical processes. The best model can only provide a possible explanation which is more consistent with known data than its current rivals. Every field observation, and especially the more qualitative or anecdotal ones, provides an opportunity to refute, or in some cases overturn, existing models and the theories which lie behind them.*

There is a plethora of important problems in the earth sciences where it is important to get the hydrology right (see Paola *et al.*, 2006). If the *Holy Grail* for a hydrologic model is an ability to effectively simulate coupled surface/near-surface responses for all problems at every scale, then the quest, at least for the foreseeable future, is well beyond reach. However, depending upon the problem/scale, there are many useful and/or promising simulation approaches. Near one end of the hydrologic-response model spectrum are the approaches discussed by McDonnell (2003) and Kirchner (2006). At the other delimiter are physics-based hydrologic-response models like TUSLO/STREAM (Freeze, 1971, 1972a,b), InHM, MODHMS (Panday and Huyakorn, 2004), and HydroGeoSphere (Sudicky *et al.*, 2005). It is not our intention to discourage the recent exhortations (e.g. Kirchner, 2006) for universal hydrologic laws that are simple and elegant. Perhaps the next generation of hydrologists will discover the equivalent of *string* (or $M$) theory from physics, which connects the micro and macro scales.

A comprehensive physics-based model (such as InHM) would be a poor choice if the objective of a simulation effort was to estimate an integrated response (e.g. discharge from an ungauged catchment). On the other hand, if the objective of the simulation effort was to quantitatively estimate the spatial and temporal distributions of pore pressures within the variably saturated 3D subsurface (i.e. for understanding slope failure initiation at the hillslope/catchment scale), then a comprehensive physics-based model would be a good choice.

Despite the ill-posed nature of many hydrologic-response simulations, these problems are still meaningful as tools for conceptual analysis and discovery. It is our position, relative to understanding and reducing the equifinality problem, that distributed measurements hold more promise than has been previously recognized in most past field measurement campaigns. On the basis of the quantity and quality of information that is available today, equifinality (from practical or mathematical perspectives) cannot be totally eliminated from hydrologic-response simulation for any model, simple or complex. One promising protocol for reducing parameter uncertainty for physics-based simulation is an improved dialogue between experimentalists and modelers (see Weiler and McDonnell, 2004), so that future field experiments and long-term observations will better capture the nonintuitive nuances associated with the distributed response for real systems.

Knowing that no hydrologic-response model works equally well for all situations (e.g. scale, data availability), it is unproductive (at this time) to suggest the complete abandonment of any approach. Hopefully, the intellectual drift discussed by Church (2005) will not take place within the hydrologic-response modelling community. The difficult problems associated with the characterization and physics-based simulation of near-surface hydrologic response (e.g. preferential flow, old/new water, and the occasional failures of the Darcian flow and Fickian dispersion models) are no more intractable than the problems that have been successfully tackled in other disciplines. For example, the fundamental problems of turbulence, which for decades defied the efforts of several stellar scientists (e.g. Werner Heisenberg, Richard Feynman, Andrei Kolmogorov), are only now being unraveled (see Falkovich and Sreenivasan, 2006). Conceptual obstacles are not broken down by black-box models, but rather by the failure of rigorous, physics-based models. These failures lead to different conceptualizations that incorporate new (frequently more complex) physics. Incorporating more physics into the next generation of hydrologic-response models will require new carefully executed field measurements for both parameterization and evaluation. The current generation of physics-based models is most useful for concept development (e.g. process identification and characterization) at the hillslope/catchment scale (instead of, for example, real-time flood forecasting at the watershed/basin scale). A continued effort to improve data shortfalls will facilitate wrestling behemoths like equifinality (at least for certain situations) closer to the ground.

## Errata

In the second of our commentaries on *physics-based hydrologic-response simulation* (Loague *et al.*, 2006), the *EF* value for the long-term InHM simulation of the C3 catchment, given correctly in Mirus *et al.* (2006), was mistakenly reported (on page 1235) as 0·89; the correct value is 0·75. Relative to the plus/minus 20% data error impacts on model performance reported by Loague *et al.* (2006), the 0·75 *EF* value is just outside the reported range [(see Table I, Loague *et al.* (2006))].

## Acknowledgements

The effort reported here was supported by National Science Foundation grant EAR-0409133. We are grateful to Keith Beven for sharing an early version of his *equifinality manifesto* during a visit to *The Farm* in 2003. The developer of InHM

**is our long-time collaborator Joel VanderKwaak. The Coos Bay data set was the result of heroic field efforts by Suzanne Anderson, Bill Dietrich, David Montgomery, and Ray Torres. The thoughtful comments from Adrianne Carr, Chris Heppner, Ben Mirus, and Joel VanderKwaak on an earlier version of this commentary are greatly appreciated.**

## References

Aktosun T, Newton RG. 1985. Non-uniqueness in the one-dimensional inverse scattering problem. *Inverse Problems* 1: 291–300.

Allison H. 1979. Inverse unstable problems and some of their applications. *The Mathematical Scientist* 4: 9–30.

Ambroise B, Beven KJ, Freer J. 1996a. Toward a generalization of the TOPMODEL concepts: Topographic indices of hydrological similarity. *Water Resources Research* 32: 2135–2145.

Ambroise B, Freer J, Beven KJ. 1996b. Application of a generalised TOPMODEL to the small ringelbach catchment, Vosges, France. *Water Resources Research* 32: 2147–2159.

Anderson SP. 1995. Flow paths, solute sources, weathering, and denudation rates: The chemical geomorphology of a small catchment. PhD Dissertation, University of California, Berkeley.

Anderson SP, Dietrich WE. 2001. Chemical weathering and runoff chemistry in a steep headwater catchment. *Hydrological Processes* 15: 1791–1815.

Anderson SP, Dietrich WE, Brimhall GH. 2002. Weathering profiles, mass balance analysis, and rates of solute loss: Linkages between weathering and erosion in a small, steep catchment. *Geological Society of America Bulletin* 114: 1143–1158.

Anderson SP, Dietrich WE, Montgomery DR, Torres R, Loague K. 1993. A case for geochemical control of concentration-discharge relationships. *Chemical Geology* 107: 369 371.

Anderson SP, Dietrich WE, Montgomery DR, Torres R, Loague K. 1997a. Concentration discharge relationships in runoff from a steep, unchanneled catchment. *Water Resources Research* 33: 211–225.

Anderson SP, Dietrich WE, Montgomery DR, Torres R, Conrad ME, Loague K. 1997b. Subsurface flowpaths in a steep, unchanneled catchment. *Water Resources Research* 33: 2637–2653.

Anderssen RS. 1968. Note on conductivity models for the earth. *Journal of Geophysical Research* 73: 6535–6543.

Aronica G, Hankin B, Beven K. 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources* 22: 349–365.

Beck MB. 2002. We have a problem. In *Environmental Foresight and Models: A Manifesto*, Beck MB (ed). Elsevier Science: Amsterdam, 11–33.

Beven KJ. 1989. Changing ideas in hydrology: the case of physically-based models. *Journal of Hydrology* 105: 157–172.

Beven K. 1996a. Equifinality and uncertainty in geomorphological modeling. In *The Scientific Nature of Geomorphology*, Rhoads BL, Thorn CE (eds). Wiley: Chichester, 289–313.

Beven K. 1996b. A discussion of distributed hydrological modelling. In *Distributed Hydrological Modelling*, Abbott MB, Refsgaard JC (eds). Kluwer Academic Publishers: Dordrecht, 255–278.

Beven KJ. 1997. TOPMODEL: A critique. *Hydrological Processes* 11: 1069–1086.

Beven KJ. 2000. Uniqueness of place and process representations in hydrological modeling. *Hydrology and Earth System Sciences* 4: 203–213.

Beven K. 2002a. Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrological Processes* 16: 189–206.

Beven K. 2002b. Towards a coherent philosophy for modelling the environment. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* 458: 2465–2484.

Beven K. 2004a. Infiltration excess at the Horton hydrology laboratory (or not?). *Journal of Hydrology* 293: 219–234.

Beven K. 2004b. Erratum to Infiltration excess at the Horton hydrology laboratory (or not?). *Journal of Hydrology* 294: 294–295.

Beven K. 2004c. Robert Horton and abrupt rises of groundwater. *Hydrological Processes* 18: 3687–3696.

Beven K. 2004d. Robert Horton's perceptual model of infiltration processes. *Hydrological Processes* 18: 3447–3460.

Beven KJ (ed). 2006a. *Streamflow Generation Processes*. IAHS Press: Wallingford, Oxfordshire.

Beven K. 2006b. A manifesto for the equifinality thesis. *Journal of Hydrology* 320: 18–36.

Beven K, Freer J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249: 11–29.

Binley A, Beven KJ. 2003. Vadose zone model uncertainty as conditioned on geophysical data. *Ground Water* 41: 119–127.

Bird P, Rosenstock RW. 1984. Kinematics of present crust and mantle flow in Southern California. *Geological Society of America Bulletin* 95: 946–957.

Blazkova S, Beven KJ. 2002. Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). *Water Resources Research* 38: 14, 10·1029/2001WR000500.

Blazkova S, Beven KJ, Kulasova A. 2002. On constraining TOPMODEL hydrograph simulations using partial saturated area information. *Hydrological Processes* 16: 441–458.

Bloschl G. 2006. Hydrologic synthesis: Across processes, places, and scales. *Water Resources Research* 42: W03S02, DOI:10·1029/2005WR004319.

Brazier RE, Beven KJ, Freer J, Rowan JS. 2000. Equifinality and uncertainty in physically based soil erosion models: Application of the GLUE methodology to WEPP–The water erosion prediction project–For sites in the UK and USA. *Earth Surface Processes and Landforms* 25: 825–845.

Bredehoeft J. 2005. The conceptualization model problem –surprise. *Hydrogeology Journal* 13: 37–46.

Burt TP. 2005. A third paradox in catchment hydrology and biogeochemistry: Decoupling in the riparian zone. *Hydrological Processes* 19: 2087–2089.

Carrera J, Neuman SP. 1986. Estimation of aquifer parameters under transient and steady state conditions: Uniqueness, stability, and solution algorithms. *Water Resources Research* 22: 211–227.

Chavent G. 1974. Identification of functional parameters in partial differential equations. In *Identification of Parameters in Distributed Systems*, Goodson RE, Polis M (eds). American Society of Mechanical Engineers: New York, 31–48.

Christiaens K, Feyen J. 2002. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrological Processes* 16: 373–391.

Church M. 2005. Continental drift. *Earth Surface Processes and Landforms* 30: 129–130.

Cooley RL, Sinclair P. 1976. Uniqueness of a model of steady-state groundwater flow. *Journal of Hydrology* 31: 245–269.

Culling WEH. 1987. Equifinality: Modern approaches to dynamical systems and their potential for geographical thought. *Transactions of the Institute of British Geographers* 12: 57–72.

Dooge JC. 1972. Mathematical models of hydrological systems. *Proceedings of the Symposium on Modelling Techniques in Water Resources Systems*, vol. 1. Environment Canada: Ottawa, 171–189.

Doty DH, Glick WH, Huber GP. 1993. Fit, equifinality, and organizational effectiveness: A test of two configurational theories. *Academy of Management Journal* 36: 1196–1250.

Ebel BA, Loague K, Dietrich WE, Montgomery DR, Torres R, Anderson SP, Giambelluca TW. 2006a. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 1. Sprinkling experiments. *American Journal of Science* (to be submitted).

Ebel BA, Loague K, VanderKwaak JE, Dietrich WE, Montgomery DR, Torres R, Anderson SP. 2006b. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 2. Comprehensive physics-based simulations. *American Journal of Science* (to be submitted).

Falkovich G, Sreenivasan KR. 2006. Lessons from hydrodynamic turbulence. *Physics Today* 59: 43–49.

Feyen L, Beven KJ, De Smedt F, Freer J. 2001. Stochastic capture zones delineated within the generalized likelihood uncertainty estimation methodology: conditioning on head observations. *Water Resources Research* 37: 625–638.

Freeze RA. 1971. Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. *Water Resources Research* 7: 347–366.

Freeze RA. 1972a. Role of subsurface flow in generating surface runoff, 1, Base flow contributions to channel flow. *Water Resources Research* 8: 609–623.

Freeze RA. 1972b. Role of subsurface flow in generating surface runoff, 2, Upstream source areas. *Water Resources Research* 8: 1272–1283.

Gottlieb G. 2001. The relevance of developmental-psychobiological metatheory to developmental neuropsychology. *Developmental Neuropsychology* 19: 1–9.

Gresov C, Drazin R. 1997. Equifinality: Functional equivalence in organization design. *Academy of Management Review* 22: 403–428.

Gupta VK, Sorooshian S. 1983. Uniqueness and observability of conceptual rainfall-runoff model parameters; the percolation process examined. *Water Resources Research* 19: 269–276.

Hadamard J. 1902. Sur les probl‘emes aux derivees partielles et leur signification physique. *Princeton University Bulletin* 13: 49–52.

Haines-Young RH, Petch JR. 1983. Multiple working hypotheses: Equifinality and the study of landforms. *Transactions of the Institute of British Geographers* 8: 458–466.

Hall FR. 1987. Contributions of Robert E. Horton. In *The History of Hydrology*, Landa ER, Ince S (eds). American Geophysical Union: Washington, 113–117.

Hankin BG, Beven KJ. 1998. Modelling dispersion in complex open channel flows: Equifinality of model structure (1). *Stochastic Hydrology and Hydraulics* 12: 377–396.

Heppner CS, Ran Q, VanderKwaak JE, Loague K. 2006a. Adding sediment transport to the integrated hydrology model (InHM): Development and testing. *Advances in Water Resources* 9: 930–943.

Heppner CS, Loague K, VanderKwaak JE. 2006b. Hydrologic response and sediment transport for the R-5 catchment: 2. Long-term simulations with InHM. *Earth Surface Processes and Landforms* (submitted).

Hreiche A, Mezher D, Bocquillon C, Dezetter A, Servant E, Najem W. 2002. Parallel processing for a better understanding of equifinality in hydrological models. In *Integrated Assessment and Decision Support*, Rizzoli AE, Jakeman AJ (eds). International Environmental Modelling and Software Society: Switzerland, 410–415.

Jones JP, Sudicky EA, Brookfield AE, park Y-J. 2006. An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow. *Water Resources Research* 42: W02 407, DOI:10·1029/2005WR004130.

Khatibi RH, Wormleaton PR, Williams JR. 2000. Parameter quality conditions in open-channel inverse problems. *Journal of Hydraulic Research* 38: 447–458.

Kirchner JW. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research* 42: W03S04, DOI:10·1029/2005WR004362.

Kirkby MJ. 1975. Hydrograph modeling strategies. In *Process in Physical and Human Geography*, Peel R Chisolm M, Hagett P (eds). Heinemann: London, 69–90.

Kirkby MJ. 1996. A role for theoretical models in geomorphology? In *The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology*, Rhoads BL, Thorn CE (eds). John Wiley and Sons: New York, 257–272.

Klemes V. 1997. Of carts and horses in hydrologic modeling. *Journal of Hydrologic Engineering* 2: 43–49.

Konikow LF. 1986. Predictive accuracy of a ground-water model–lessons from a postaudit. *Ground Water* 24: 173–184.

Konikow LF, Bredehoeft JD. 1992. Groundwater models cannot be validated. *Advances in Water Resources* 15: 75–83.

Lamb R, Beven KJ, Myrabo S. 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources* 22: 305–317.

Leliavsky S. 1965. *River and Canal Hydraulics*. Chapman and Hall: London.

Loague K, Green RE. 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology* 7: 51–73.

Loague K, VanderKwaak JE. 2002. Simulating hydrologic response for the R-5 catchment: Comparison of two models and the impact of the roads. *Hydrological Processes* 16: 1015–1032.

Loague K, VanderKwaak JE. 2004. Physics-based hydrologic response simulation: Platinum bridge, 1958 Edsel, or useful tool. *Hydrological Processes* 18: 2949–2956.

Loague K, Heppner CS, Abrams RH, VanderKwaak JE, Carr AE, Ebel BA. 2005. Further testing of the Integrated Hydrology Model (InHM): Event-based simulations for a small rangeland catchment located near Chickasha, Oklahoma. *Hydrological Processes* 19: 1373–1398.

Loague K, Heppner CS, Mirus BB, Ebel BA, Ran Q, Carr AE, BeVille SH, VanderKwaak JE. 2006. Physics-based hydrologic-response simulation: Foundation for hydroecology and hydrogeomorphology. *Hydrological Processes* 20: 1231–1237.

McClure J, Densley L, Liu JH, Allen M. 2001. Constraints on equifinality: Goals are good explanations only for controllable outcomes. *British Journal of Social Psychology* 40: 99–115.

McDonnell JJ. 2003. Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response. *Hydrological Processes* 17: 1869–1875.

McElwee CD. 1982. Sensitivity analysis and the groundwater inverse problem. *Ground Water* 20: 723–735.

Mirus BB, Ebel BA, Loague K, Wemple BC. 2006. Simulated effect of a forest road on near-surface hydrologic response: Redux. *Earth Surface Processes and Landforms* (in press).

Mohanty BP, Ankeny MD, Horton R, Kanwar RS. 1994. Spatial analysis of hydraulic conductivity measured using disc infiltrometers. *Water Resources Research* 30: 2489–2498.

Montgomery DR. 1991. Channel initiation and landscape evolution. PhD Dissertation, University of California, Berkeley.

Montgomery DR, Dietrich WE. 2002. Runoff generation in a steep, soil-mantled landscape. *Water Resources Research* 38: 1168, DOI:10·1029/2001WR000822.

Montgomery DR, Dietrich WE, Heffner JT. 2002. Piezometric response in shallow bedrock at CB1: Implication for runoff generation and landsliding. *Water Resources Research* 38: 1274, DOI:10·1029/2002WR001429.

Montgomery DR, Dietrich WE, Torres R, Anderson SP, Heffner JT, Loague K. 1997. Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. *Water Resources Research* 33: 91–109.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models, Part I–A discussion of principles. *Journal of Hydrology* 10: 282–290.

Neuman SP. 1973. Calibration of distributed parameter groundwater flow models viewed as a multi-objective decision process under uncertainty. *Water Resources Research* 9: 1006–1021.

Neuman SP, Guadagnini A. 1999. A new look at traditional deterministic flow models and their calibration in the context of randomly heterogeneous media. In *Proceedings of the International Conference on Calibration and Reliability in Groundwater Modeling*, ModelCARE99: Zurich; 189–197.

Oreskes N, Schrader-Frechette K, Belitz K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263: 641–644.

Panday S, Huyakorn PS. 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Advances in Water Resources* 27: 361–382.

Paola C, Foufoula-Georgiou E, Dietrich WE, Hondzo M, Mohrig D, Parker G, Power ME, Rodriguez-Iturbe I, Voller V, Wilcox P. 2006. Toward a unified science of the earth's surface: Opportunities for synthesis among hydrology, geomorphology, geochemistry, and ecology. *Water Resources Research* 42: W03S10, DOI:10·1029/2005WR004336.

Pebesma EJ, Switzer P, Loague K. 2005. Error analysis for the evaluation of model performance: Rainfall-runoff event time series data. *Hydrological Processes* 19: 1529–1548.

Pebesma EJ, Switzer P, Loague K. 2006. Error analysis for the evaluation of model performance: Rainfall-runoff event summary variables. *Hydrological Processes* (in press).

Richards LA. 1931. Capillary conduction of liquids through porous mediums. *Physics* 1: 318–333.

Rogers AR. 2000. On equifinality in faunal analysis. *American Antiquity* 65: 709–723.

Roy A. 1962. Ambiguity in geophysical interpretation. *Geophysics* 27: 90–99.

Saulnier GM, Datin R. 2004. Analytical solution to a bias in the TOPMODEL framework balance. *Hydrological Processes* 18: 1195–1218.

Savenije HHG. 2001. Equifinality, a blessing in disguise? *Hydrological Processes* 15: 2835–2838.

Schulz K, Beven K, Huwe B. 1999. Equifinality and the problem of robust calibration in nitrogen budget simulations. *Soil Science Society of America Journal* 63: 1934–1941.

Sidle RC. 2006. Field observations and process understanding in hydrology: Essential components in scaling. *Hydrological Processes* 20: 1439–1445.

Sorooshian S, Gupta VJ. 1983. Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness. *Water Resources Research* 19: 260–268.

Sorooshian S, Gupta VJ, Fulton JL. 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility. *Water Resources Research* 19: 251–259.

Stephenson GR, Freeze RA. 1974. Mathematical simulation of subsurface flow contributions to snowmelt runoff, Reynolds Creek, Idaho. *Water Resources Research* 10: 284–298.

Sudicky EA, Therrien R, Park Y-J, McLaren RG, Jones JP, Lemieux J-M, Brookfield AE, Colautti D, Panday S, Guvanasen V. 2005. On the challenge of integrated surface-subsurface flow and transport modelling at multiple catchment scales. *Geological Society of America Abstracts with Programs* 37: 28.

Torres R. 1997. Unsaturated zone processes and the hydrologic response of a steep, unchanneled valley. PhD Dissertation, University of California, Berkeley.

Torres R, Dietrich WE, Montgomery DR, Anderson SP, Loague K. 1998. Unsaturated zone processes and the hydrologic response of a steep, unchanneled catchment. *Water Resources Research* 34: 1865–1879.

VanderKwaak JE. 1999. Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems. PhD dissertation, University of Waterloo: Waterloo, Ontario.

VanderKwaak JE, Loague K. 2001. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resources Research* 37: 999–1013.

Vertessy RA, Elsenbeer H. 1999. Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization. *Water Resources Research* 35: 2173–2187.

Weiler M, McDonnell JJ. 2004. Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology. *Journal of Hydrology* 285: 3–18, DOI:10·1016/S0022-1694(03)00 271-3.

Wierenga PJ, Toorman AF, Hudson DB, Vinson J, Nash M, Hills RG. 1989. Soil physical properties and the Las Cruces trench site. Report NUREG/CR-5441. U.S. Nuclear Regulatory Commission, Washington, DC.

Wigmosta MS, Lettenmaier DP. 1999. A comparison of simplified methods for routing topographically-driven subsurface flow. *Water Resources Research* 35: 255–264.

Yeh WWG. 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research* 22: 95–108.

Young P, Parkinson S, Lees M. 1996. Simplicity out of complexity in environmental modelling: Occam's razor revisited. *Journal of Applied Statistics* 23: 165–210.

Zak S, Beven KJ. 1999. Equifinality, sensitivity and uncertainty in the estimation of critical loads. *Science of the Total Environment* 236: 191–214.

Zimmerman DM, de Marsily G, Gotway CA, Marietta MG, Axness CL, Beauheim RL, Bras RL, Carrera J, Dagan G, Davies PB, Gallegos DP, Galli A, Gómez-Hernández J, Grindrod P, Gutjahr AL, Kitanidis PK, Lavenue AM, McLaughlin D, Neuman SP, RamaRao BS, Ravenne C, Rubin Y. 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research* 34: 1373–1413.