HYDROLOGICAL PROCESSES
Hydrol. Process. **24**, 3878–3893 (2010)

(wileyonlinelibrary.com) DOI: 10.1002/hyp.7819

Gypsies in the palace: experimentalist's view on the use of 3-D physics-based simulation of hillslope hydrological response

April L. James,[1]* Jeffrey J. McDonnell,[2] Ilja Tromp-van Meerveld[3] and Norman E. Peters[4]

[1] *Department of Geography and Environment, Nipissing University, North Bay, ON, P1B 8L7, CA*
[2] *Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis OR 97331, USA and School of Geosciences, University of Aberdeen, Aberdeen, Scotland, AB23 3UF, UK*
[3] *Department of Geography, Simon Fraser University, Burnaby, BC V5A-4Z1, Canada*
[4] *U.S. Geological Survey, Georgia Water Science Center, Atlanta, GA 30360, USA*

Abstract:

As a fundamental unit of the landscape, hillslopes are studied for their retention and release of water and nutrients across a wide range of ecosystems. The understanding of these near-surface processes is relevant to issues of runoff generation, groundwater–surface water interactions, catchment export of nutrients, dissolved organic carbon, contaminants (e.g. mercury) and ultimately surface water health. We develop a 3-D physics-based representation of the Panola Mountain Research Watershed experimental hillslope using the TOUGH2 sub-surface flow and transport simulator. A recent investigation of sub-surface flow within this experimental hillslope has generated important knowledge of threshold rainfall-runoff response and its relation to patterns of transient water table development. This work has identified components of the 3-D sub-surface, such as bedrock topography, that contribute to changing connectivity in saturated zones and the generation of sub-surface stormflow. Here, we test the ability of a 3-D hillslope model (both calibrated and uncalibrated) to simulate forested hillslope rainfall-runoff response and internal transient sub-surface stormflow dynamics. We also provide a transparent illustration of physics-based model development, issues of parameterization, examples of model rejection and usefulness of data types (e.g. runoff, mean soil moisture and transient water table depth) to the model enterprise. Our simulations show the inability of an uncalibrated model based on laboratory and field characterization of soil properties and topography to successfully simulate the integrated hydrological response or the distributed water table within the soil profile. Although not an uncommon result, the failure of the field-based characterized model to represent system behaviour is an important challenge that continues to vex scientists at many scales. We focus our attention particularly on examining the influence of bedrock permeability, soil anisotropy and drainable porosity on the development of patterns of transient groundwater and sub-surface flow. Internal dynamics of transient water table development prove to be essential in determining appropriate model parameterization. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS hillslope model; bedrock permeability; threshold; effective parameters; sub-surface flow; TOUGH2

Received 11 August 2009; Accepted 14 June 2010

INTRODUCTION

Hillslopes are fundamental units of the landscape and the essential building block of catchments (Troch *et al.*, 2003). They generate upslope response linking them to downslope components of the landscape. Their delivery of water and transport of energy sources (e.g. dissolved organic carbon (DOC)), nutrients and non-point source contaminants like mercury (Bushey *et al.*, 2008) link them to the sustainable health of downslope components of the landscape including riparian areas, wetlands, streams and lakes. Nevertheless, hillslopes exhibit considerable heterogeneity in soil and hydrological properties and complex responses to rainfall and snowmelt inputs. As a result, hillslope hydrology as a discipline still lacks the ability to predict hillslope behaviour in ungauged basins that may result, in part, from organization of observations of hillslope responses (across these heterogeneities) and their examination for common patterns of behaviour (McDonnell *et al.*, 2007). Clearly, there are many ways forward to new learning in hillslope hydrology. One recent approach to new learning has been the use of 3-D physics-based hydrological response simulation models, defined here as numerical models solving Darcy—Richards' equation (Richards, 1931) for variably saturated flow in porous media in three dimensions. These models can be considered a subset of physically based models that have been used over the last 40 years to examine near-surface hydrological processes (see Loague and Vanderkwaak, 2004, for a representative summary). Loague *et al.* (2006) argue that such models may provide a foundation for understanding coupled hydroecological and hydrogeomorphological systems at the hillslope scale, inform new understanding and prompt new experiments. While several papers have now advocated their use, few studies have applied complete hillslope-scale data sets to test physics-based models and parameterizations.

---

*Correspondence to: April L. James, Department of Geography and Environment, Nipissing University, North Bay, ON, P1B 8L7, CA.
E-mail: aprilj@nipissingu.ca

Recent case studies of 3-D physics-based model development and testing of hillslope or small catchment behaviour include the R-5 prairie grassland catchment (0·1 km^2) (Loague and Vanderkwaak, 2002; Loague *et al.*, 2005) and the steep forested unchanneled hollow CB1 catchment (860 m^2) (Ebel and Loague, 2006; Ebel *et al.*, 2007a). In both instances, physics-based models with effective parameterization are used to simulate near-surface hillslope hydrological response and results are compared with high-resolution temporal and spatially distributed observational datasets, making these case studies relatively unique. Both highlight key issues that come with using this approach, including boundary value problems, appropriate initialization of catchment storage or moisture conditions, challenges in parameter selection and additional process representation [e.g. evapotranspiration (ET) and preferential flow]. Further, these case studies provide detailed examination of how much of the sub-surface (i.e. how deep) and its complexity is required to model the near-surface hydrology. For instance, Loague *et al.* (2005) illustrate the significant effects of surface and alluvium permeability and soil–water retention curve parameters, mesh depth and inclusion of evaporatranspirative processes on the simulated near-surface response at the R-5 catchment. Simulation of the forested CB1 catchment (Ebel *et al.*, 2007a,b, 2008) emphasizes the need for greater characterization of the sub-surface and its heterogeneity to better simulate distributed pore water pressures in the near surface.

In parallel with physics-based modeling efforts, experimental work is increasingly focused on macroscale behaviour and, in particular, the intersection of patterns and thresholds at the scale of the entire hillslope. This contrasts with the 2-D and 3-D model view of hillslopes as superpositions of well-parameterized soil cores. Tromp-van Meerveld and McDonnell (2006a,b) showed how, at the Panola Mountain Research Watershed (Georgia, USA), experimental hillslope patterns of transient water table on the slope related to thresholds in rainfall amounts necessary to initiate lateral sub-surface flow at the hillslope scale. Connectivity of these patches of saturation was the necessary prerequisite for hillslope-scale lateral flow—a macroscale, emergent property of the hillslope. The experimental work performed at the Panola hillslope (Tromp-van Meerveld and McDonnell, 2006b; Tromp-van Meerveld *et al.*, 2007) suggests that the fill-and-spill phenomenon (development of transient saturation and its pattern) is a function of the contrast in soil and bedrock properties (e.g. hydraulic conductivity).

Many open questions remain in the area of reconciling physics-based model approaches in hillslope studies with field evidence of process behaviour. Here we take a decidedly experimentalist-centric view to physics-based modelling and ask the question: Can a 3-D physics-based approach, parameterized with field data, actually represent measured hillslope sub-surface stormflow response? To achieve this, we chronicle the often behind-the-scenes work that goes on with such a model to 'make it fit' and attempt to illustrate some of the issues involved—from both a field and modelling perspective. This work addresses our perceived impediment to moving forward in physics-based hillslope modelling, the missing framework for dialogue between the experimentalist (who from data analysis and direct observation has an understanding of how the system 'works') and the modeller (who understands how the code 'works'). Our work follows Cloke *et al.* (2003), who showed how this dialogue can aid model structural decisions and subsequent parameterization but only after the consequences of the implementation of these decisions are fully understood. We apply the physics-based simulator TOUGH2 (Pruess, 1991; Pruess *et al.*, 1999; Pruess, 2004) to explore 3-D model representation of sub-surface flow within the Panola hillslope. Our objectives are as follows: (1) to develop a parsimonious 3-D physics-based model to simulate hillslope hydrological response for the highly instrumented Panola experimental hillslope, (2) to test the ability of the 3-D model (both uncalibrated and calibrated) to simulate hillslope rainfall-runoff response, changing storage and internal transient sub-surface stormflow dynamics for a large rainstorm (63 mm) on 30 March 2002 and (3) to examine critical parameters and data requirements needed to characterize the hillslope and, in so doing, develop learning about the nature of hillslope hydrological response to rainfall.

Many physics-based hillslope modeling studies cannot test against high-resolution datasets, many are restricted to the simplification of 2-D and many often do not provide a transparent discussion of challenges encountered during implementation (see Finsterle *et al.*, 2008 for discussion of challenges in vadose zone modelling); challenges such as the evolution of a parsimonious model structure that leads to decisions to add complexity supported by experimental data, or issues arising from the attempt to reconcile appropriate parameter values from field data (e.g. K_{sat}) for a successful testing of an uncalibrated model. Our work with TOUGH2 at Panola seeks to provide a transparent illustration of the use of field observations (quantitative and qualitative) for testing of a 3-D physics-based model, the selection and exploration of an appropriate parameter space and the process of model rejection and process learning. Specifically, we examine the role of the soil-bedrock contact, the ratio of soil to bedrock permeability, soil anisotropy and soil drainable porosity—things that might get at the whole-slope emergent behaviour (as discussed by Tromp-van Meerveld and McDonnell, 2006b) and the representation of these in a state-of-the-art model. Our research builds on that of Hopp and McDonnell (2009) who worked with a virtual Panola hillslope to understand the interaction of hillslope and event-based factors (soil depth, slope angle, bedrock K_{sat} and storm size) that affect hillslope hydrological response. This is an extremely fertile area for new learning given the near impossibility of making the key measurements in the field to assess how bedrock permeability might influence the filling and spilling of subsoil depressions, what level of soil anisotropy might be necessary for development of connectivity of saturated patches and how soil drainable porosity modulates the overall sub-surface response—all likely ingredients of a new theory of hillslope hydrology still some years away (McDonnell *et al.*, 2007).

THE PANOLA MOUNTAIN RESEARCH WATERSHED EXPERIMENTAL HILLSLOPE

The Panola Mountain Research Watershed experimental trenched hillslope, henceforth called the Panola hillslope, is located within the Panola Mountain State Conservation Park, in the Piedmont of Georgia, USA (84°10′W, 33°37′N), approximately 25 km southeast of Atlanta. Experimental research at the Panola hillslope over the last 25 years has created a unique dataset that has been used to characterize its hydrological behaviour and provides an ideal framework for the development and testing of a physics-based model. Recent efforts have assembled earlier descriptive work (digital elevation models (DEMs), soil properties) with the high spatial and temporal resolution 2002 dataset (Tromp-van Meerveld and McDonnell, 2006a,b,c; Tromp-van Meerveld *et al.*, 2007) in a datanote (Tromp-van Meerveld *et al.*, 2008) with the purpose of making these data available to the greater community to promote future model testing and intercomparisons. The accompanying website, http://www.sfu.ca/PanolaData/index.htm provides a full description of the assembled data types, their organization and origins. The 2002 rainfall-runoff dataset is particularly rich in both spatial and temporal extent, including extensive measurements of internal hillslope hydrological response in the form of spatial water table development and dynamic hillslope-scale moisture storage.

The extensive analysis of rainfall-runoff response at the Panola hillslope has generated significant process-level understanding (Tromp-van Meerveld and McDonnell, 2006a,b,c; Tromp-van Meerveld *et al.*, 2007). The current conceptual model of Panola hillslope storm response includes a significant loss of water from the hillslope to the underlying bedrock as an important mechanism. There is no site-specific evidence for return flow from the bedrock into the shallow soils of the hillslope (Burns *et al.*, 2001). For high-intensity storms, there is some evidence of overland flow (displaced leaves) on small sections of the hillslope immediately below bedrock outcrops but that infiltrates quickly downslope. The 2002 rainfall-runoff experimental dataset, illustrated in Figure 1, includes observation of two large storms delivering significant (>1 mm) sub-surface flow. In this paper, we focus specifically on the simulation of the 30-March-2002 rainstorm that delivered 63 mm on the hillslope, producing a sizable runoff response and during which spatially variable transient water table development was observed.

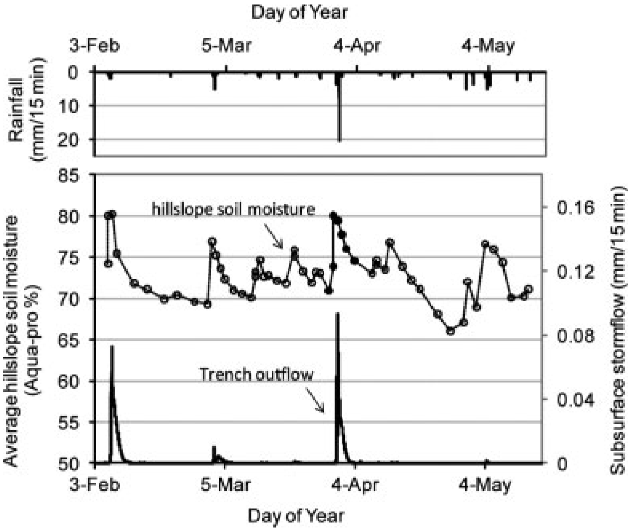

Figure 1. Observations of rainfall, hillslope average soil moisture and sub-surface stormflow recorded at the trench face (Tromp-van Meerveld *et al.*, 2008). Average hillslope soil moisture is presented in % AquaPro (uncalibrated soil moisture as collected by an AquaPro moisture probe; see Tromp-van Meerveld and McDonnell, 2006c, for more details). Seven surveys collected immediately before, during and after the 30-March-2002 rainstorm are indicated by black circles

MODEL DEVELOPMENT DESCRIPTION

TOUGH 2 process representation

The family of TOUGH codes (Pruess, 2004; Finsterle *et al.*, 2008) has been used in many near-surface vadose zone and groundwater applications (James and Oldenburg, 1997), sub-surface seepage studies (Finsterle *et al.*, 2003), in addition to deeper sub-surface flow and transport applications (Falta, 2003). It is well suited for our focus on hillslope sub-surface flow simulation where surface flow processes are of minimal to no importance. The integrated hillslope hydrological response and the 3-D sub-surface dynamics of variably saturated flow within the Panola hillslope were simulated using the TOUGH2 general purpose reservoir simulator (Pruess, 1991; Pruess *et al.*, 1999) and fluid property module EOS 9, which solves Richards' equation (Richards, 1931) for isothermal, variably saturated flow. TOUGH2's fluid property module EOS 9 performs a mass balance on the liquid phase (water) only, treating the gas phase in the unsaturated zone as passive and at constant pressure, and, under isothermal conditions, assumes constant liquid phase properties (viscosity and density). As a result, a single mass balance equation is solved for liquid phase flow only (Pruess, 1991):

$$\frac{\partial}{\partial t}\theta = div[K\nabla h] \tag{1}$$

where specific volumetric moisture content, θ, is a function of porous media density (ρ), liquid phase saturation (S_l) and liquid phase density (ρ_l),

$$\theta = \phi S_l \rho_l, \tag{2}$$

hydraulic conductivity K is a function of absolute permeability (k), relative liquid phase permeability (k_{rl}),

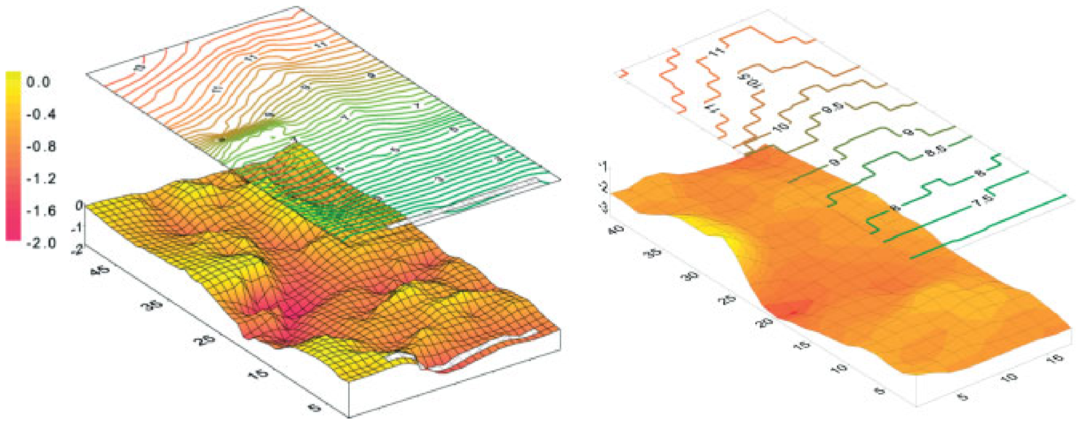

Figure 2. Observed surface elevations and bedrock topography data (a) and the TOUGH2 model grid (b). Units are in metres. Left-hand side image adopted from Freer *et al.* (2002)

liquid density (ρ_l) and viscosity (μ_l) and gravitational acceleration (g):

$$K = \frac{kk_{rl}\rho_l g}{\mu_l} \tag{3}$$

Relative permeability (k_{rl}) is described by a user-selected function (e.g. van Genuchten-Mualem model). Hydraulic head is expressed as a function of liquid phase fluid pressure (P_l) and density (ρ_l), gravity and elevation head, z.

$$h = z + \frac{P_l}{\rho_l g} \tag{4}$$

Liquid phase fluid pressure is a function of a reference phase pressure and capillary pressure described by a user-selected function (e.g. van Genuchten function). The TOUGH simulator does not include surface flow representation and all water infiltrates and moves through the sub-surface porous media. Space is discretized using an integral finite difference method (IFDM) (see Pruess, 1991) and equations are solved using Newton–Raphson iteration with automatically adjusted time steps. The use of the TOUGH2 code is further enhanced by iTOUGH2's (Finsterle, 2000) capabilities of parameter estimation, sensitivity analyses and uncertainty propagation analysis.

In development of the Panola hillslope model, we drew on the concept of parsimony (Mulligan and Wainwright, 2004), selecting a simplified representation of the hillslope structure supported by available data. The hillslope sub-surface was represented by two hydrogeologic-based units: a sandy-loam soil with variable thickness overlying weathered granite bedrock or saprolite. On the basis of the field observations of 'spill-and-fill' phenomenon (Tromp-van Meerveld and McDonnell, 2006b), the model was built to resolve the irregular surface and bedrock topography and the resulting variation in soil thickness (McDonnell *et al.*, 1996; Freer *et al.*, 1997) and the loss of water to a permeable lower bedrock formation (Tromp-van Meerveld *et al.*, 2007). The two geologic domains were prescribed as spatially (lateral and vertical) homogeneous in their properties and used an effective porous media to represent the potential mix of matrix and pipe flow. Although field observations indicate that porosity within the hillslope soil derives from both the soil matrix and macropores, with tree roots up to 6·4 cm in diameter observed in hillslope cores (McIntosh *et al.*, 1999) and soil pipes intersecting the trench face (Freer *et al.*, 1997) delivering significant pipeflow, no real information on which to base parameterization of dual continuum (matrix and macropore) flow is available. Laboratory infiltration experiments performed on a large undisturbed soil core (38 cm diameter by 38 cm long) collected at Panola hillslope suggested matrix flow as a dominant process (McIntosh *et al.*, 1999) and yet, not surprisingly, at the scale of the entire hillslope, observations indicate significant pipeflow at the trench face (McDonnell *et al.*, 1996; Freer *et al.*, 1997; Tromp-van Meerveld and McDonnell, 2006a). The underlying saprolite and weathered bedrock may also exhibit preferential flow (e.g. fracture flow).

Model domain

Figure 2a (left-hand side) presents topographic data of the ground surface elevation and bedrock topography, modified from Freer *et al.* (2002), while 2b (right-hand side) shows the corresponding model ground and bedrock surfaces. The hillslope was represented by a 3-D rectilinear grid 46 m in length and 20 m in width and set on a slope of 13°. The horizontal discretisation was 2·0 m, resulting in a maximum of 23 × 10 grid or 230 gridblocks per layer. The vertical discretisation was 12 cm. The irregular surface and sub-surface topography of the hillslope (Figure 2a), as measured by a 2 m × 2 m survey of ground surface elevations and bedrock topography (Freer *et al.*, 1997; Zumbuhl, 1998), was mimicked in the model (Figure 2b) by spatially varying the number of vertical layers (created by removing gridblocks from an initially rectangular grid). As a result, modelled soil depth varied from 12 cm (1 layer) to 192 cm (16 layers) compared to the actual soil depth variation of 0–186 cm. Beneath the soil layers, a hydrologically active zone of bedrock was defined by nine layers of increasing thickness. Each layer of bedrock was represented by a single gridblock extending over the entire areal extent of the hillslope sub-surface. At the soil–bedrock interface, all soil gridblocks were vertically connected to the top most bedrock layer gridblock. This configuration removed computation of lateral redistribution of water within the bedrock domain while still resolving vertical loss, a model simplification that decreases the total number of gridblocks for which the transition between unsaturated and saturated flow must be calculated, reducing the computation time.

Boundary conditions

Above the hillslope a Dirichlet-type atmospheric boundary condition with constant pressure was prescribed by assigning a very large gridblock volume. Rainfall was introduced into a thin sub-atmospheric layer (high permeability and porosity and no capillarity) immediately above the top soil layer and infiltrated into the soil layer below. Given that the dataset was collected during the dormant season, loss of water from the hillslope due to ET was not included in simulations. Below the bedrock domain, a free gravity-drainage Dirichlet-type boundary condition was prescribed. This lower boundary condition was used to emulate the loss of water from shallow bedrock to deeper storage and the active hillslope of the model domain. This water may ultimately contribute to recharge of the stream riparian area at the larger catchment scale but was not included in the hillslope simulation. The upslope boundary at the top of the hillslope and lateral boundaries on either side of the hillslope were defined as *no-flow* or *closed-boundaries*. These boundary conditions present some potential inaccuracies due to the irregular surface of the bedrock topography, particularly along the lower left-hand side (when looking upslope) of the hillslope where there exists a deep bedrock depression (Figure 2a). The model assumed these inaccuracies to be small. Infiltrating water could move vertically to depth and/or downslope under variably saturated conditions to the soil–trench interface.

The 20-m-long-by-1·5-m-wide trench located at the base of the hillslope was excavated in 1995, and is protected from rainfall by a roof (Freer *et al.*, 1997; Burns *et al.*, 1998, 2002). Soil on the excavated trench face is exposed to the atmospheric. During rainstorms, sub-surface storm flow (sssf) is generated from both the soil matrix and individual soil macropores. The TOUGH2 model represented the trench as a Dirichet-type boundary condition (as described above), creating a capillary barrier at the soil–trench interface, where hillslope gridblocks were connected directly to a trench gridblock represented by a much coarser porous medium as described in Oldenburg and Pruess (1993). As a result, water would remain in the hillslope (the finer soil) due to stronger capillary forces until near-saturation conditions. The numerical simulation of the capillary barrier effect is affected by the weighting scheme by which relative and absolute permeabilities are evaluated at the interface of any two gridblocks. The transient, two-phase nature of flow for this hillslope runoff simulation required that relative and absolute permeabilities be upstream weighted as summarized by Pruess (1991). In upstream weighting, relative and absolute permeabilities are set by the upstream gridblock. Although the capillary barrier problem, which depends on the contrast in properties of two materials (in this case, the hillslope soil and the trench face), is better represented by harmonic weighting (where relative and absolute permeabilities are set as a function of both upstream and downstream gridblocks and weighted towards the smaller value), numerical studies have shown that upstream weighting also captures the capillary barrier effect (Oldenburg and Pruess, 1993). As with the atmospheric boundary condition described above, no evaporation was included at the soil-trench boundary condition.

Model parameters and hydrogeologic properties

Model parameters describing the soil and bedrock geologic domains were based on detailed site-specific and literature-derived data. McIntosh *et al.* (1999) performed detailed analysis of an undisturbed hillslope soil core (38 cm diameter by 38 cm long) collected from a nearby hillslope at Panola and four small soil cores (5 cm diameter by 30 cm long). From gravimetric analysis of the four small soil cores (9·7 cm × 7·7 cm long), soil porosity was estimated to be 0·47 and 0·48 at middle and bottom depths respectively (McIntosh *et al.*, 1999). Figure 3 illustrates matrix potential (soil tension) curves generated in the lab on a tension table for soil cores taken directly from the Panola hillslope at depths of 15, 40 and 70 cm (Tromp-van Meerveld *et al.*, 2008). These data indicate soil core porosities ranging from 0·64 at depths of 15 cm to 0·52 at 40 cm and 70 cm depths. Similar estimates of porosity (0·65 to 0·51) were obtained from time domain reflectometry (TDR) within the large undisturbed hillslope core of McIntosh *et al.* (1999). Constant head method testing of the small hillslope

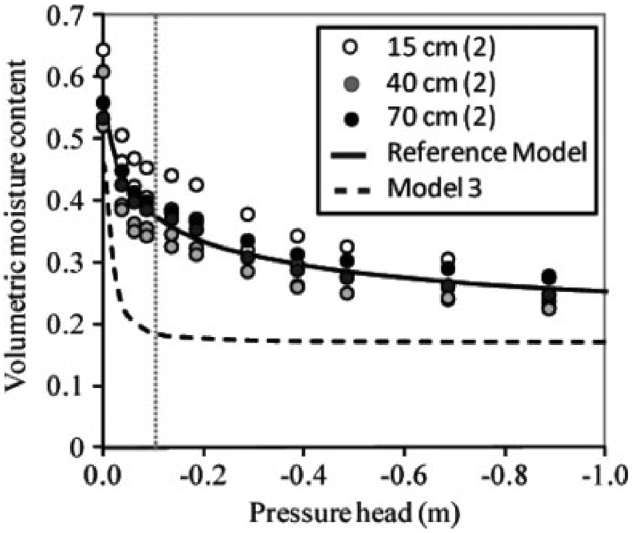

Figure 3. Soil moisture-retention curves for (a) laboratory generated measurements from hillslope soil core data collected at depths of 15, 40 and 70 cm (two cores at each depth), (b) Reference Model (van Genuchten) with drainable porosity of 0·20 and (c) Model 3 (van Genuchten) with drainable porosity of 0·30. See Table I for corresponding model parameters

Table I. Hydrogeologic properties of the Panola TOUGH2 model

Layer	ρ_b (g cm^{-3})	Porosity (m^3 m^{-3})	k (m^2)	S_{lr}	m	p_e (Pa^{-1})
Soil	1·39[a]	0·58[b]	$2{\cdot}5 \times 10^{-11}$[c]	0·09[d]	0·20[e]	$7{\cdot}2 \times 10^{-3}$[e]
Bedrock	2·25[f]	0·33[f]	$2{\cdot}5 \times 10^{-13}$[g]	0·09[f]	0·20[h]	$2{\cdot}0 \times 10^{-3}$[h]
Soil (Model 3)	1·39[a]	0·47[i]	$5{\cdot}1 \times 10^{-11}$[i]	0·36[i]	0·60[i]	$7{\cdot}2 \times 10^{-3}$[e]

[a] Gravimetric method performed on four small soil cores (5 cm diameter by 30 cm long) from a nearby hillslope location; average of two reported values (McIntosh *et al.*, 1999).

[b] Soil porosity (0·58) is the average value from Tromp-van Meerveld (unpublished data).

[c] Absolute permeability of hillslope soil, converted from saturated hydraulic conductivity values estimated with the constant head method for small cores (McIntosh *et al.*, 1999).

[d] General value of residual liquid saturation for a sandy-loam (Carsel and Parrish, 1988).

[e] Hillslope soil van Genuchten parameters determined from manual fit of matrix potential (soil tension) curves (Figure 2).

[f] Bulk density of weathered granite bedrock estimated from Jones and Graham (1993); bedrock porosity and residual liquid saturation estimated from Graham *et al.* (1997).

[g] Bedrock absolute permeability set at two orders of magnitude smaller than soil permeability as supported by both site-specific estimates and literature values of weathered granite bedrock (Graham *et al.*, 1997).

[h] Bedrock van Genuchten parameters estimated from moisture-retention curves of weathered granite bedrock of Graham *et al.* (1997) and Katsura *et al.* (2005).

[i] Model 3 parameterization (see text for explanation).

ρ_b is bulk density; k is absolute permeability; p_e is the air entry pressure ($1/\alpha$) of the van Genuchten function; m is the pore-size distribution index of the van Genuchten function ($m = 1 - 1/n$); n and m are often referred to as shape parameters (Mertens *et al.*, 2005); S_{lr} is the residual liquid saturation (in units of volumetric moisture content).

cores measured a saturated hydraulic conductivity (K_{sat}) of 64 cm/h at 10 cm depth (McIntosh *et al.*, 1999). Although no distinct layering is visible within the soil profile, K_{sat} values for greater depths within the soil profile suggest evidence of an exponential decline with depth (Tromp-van Meerveld, unpublished data; Zumbuhl, unpublished data).

In the simulations presented, a parsimonious model parameterization with an irregular soil depth but uniform soil K_{sat} profile was prescribed. The above data were used to characterize a single reference soil type with a set of spatially uniform parameter values (Table I), with K_{sat} of ~64 cm/h (absolute permeability, k, of $2{\cdot}5 \times 10^{-11}$ m^2). Soil moisture tension was modelled using a single representative van Genuchten (1980) curve with no hysteresis (Figure 3; Table I) and relative permeability using the van Genuchten–Mualem model (Mualem, 1976; van Genuchten, 1980) (Table I). Table I includes values of pore-size distribution index or van Genuchten shape parameter m and air entry pressure p_e.

The following site-specific and literature-based information was used to estimate field-based parameterization of the weathered granite bedrock underlying the Panola hillslope soils. Typical values of K_{sat} for weathered granite range between 0·36 and 72 cm/h (Morris and Johnson, 1967). Although few direct measurements of saprolite and bedrock K_{sat} exist for the Panola hillslope, sprinkler experiments conducted on lower sections of the hillslope estimate bedrock K_{sat} as 0·58 cm/h (k of $2{\cdot}3 \times 10^{-13}$ m^2) (Tromp-van Meerveld *et al.*, 2007). Table I describes the parameterization of the bedrock domain for the Reference Model simulation used in the TOUGH2 model. The model K_{sat} of bedrock was 0·64 cm/h (k of $2{\cdot}5 \times 10^{-13}$ m^2), 2 orders of magnitude smaller than soil permeability, consistent with sprinkler experiments derived values of Tromp-van Meerveld *et al.* (2007). The modelled bedrock also had lower porosity than the sandy-loam soil and a flatter moisture- retention curve with lower volumetric moisture content for a given matrix potential than for the overlying sandy-loam soil (Table I).

Sensitivities of simulated water balance components (hillslope storage and trenchflow) to soil and bedrock parameterization were evaluated using iTOUGH2 (Finsterle, 2000). iTOUGH2 calculated the sensitivity coefficient of each parameter (j) for each observation (water balance component) (i) using the Perturbation method with either forward of centred finite differences to calculate the Jacobian matrix of sensitivity coefficients (Finsterle, 2000). Sensitivities were scaled with respect to the standard deviation of the observation (σ_z) and expected parameter variation (σ_p) to obtain unitless sensitivity coefficients. As an example (using forward finite difference), a unitless sensitivity coefficient for parameter j and observation i is

$$\tilde{J}_{ij} = \frac{\partial z_i}{\partial p_j} \times \frac{\sigma_{p_j}}{\sigma_{z_i}} = \frac{z_i(p; p_j + \partial p_j) - z_i(p)}{\partial p_j} \times \frac{\sigma_{p_j}}{\sigma_{z_i}} \qquad (5)$$

where

z_i = observation i
p_j = parameter j
$\delta p = \alpha\ \mathrm{p}_j$,

and α = small variation (e.g. here 10%)

To get the total relative sensitivity d_j (unitless), for an individual parameter, $\tilde{J}_{ij}$ was summed over all observations i in a time series (Finsterle, 2000). Grid-searches of parameter space were then used to examine modelled hillslope behaviour (water balance components) for the most sensitive parameters.

Infiltration

Throughfall available for infiltration during the 30-March-2002 rainstorm was estimated by subtracting estimated interception from the rainfall time series (see

Tromp-van Meerveld *et al.*, 2008 for details): the first 1 mm of precipitation was assumed to be lost to canopy storage after which 95% of precipitation was assumed to be available for infiltration. The throughfall time series (mm/15 min) was then converted into mass injection rates (kg/s) and added as a source term to the thin atmospheric layer immediately above the soil domain on a 15-min time interval. All throughfall was assumed to infiltrate into the sandy-loam soil.

Model initial conditions

The Panola hillslope is a characteristically well-drained hillslope, remaining unsaturated for most of the water year. However, saturation within the hillslope occurs during high-magnitude rainstorms. As a result, the hillslope is in a continuous state of change in moisture storage due to the temporally varying processes of rainfall, ET, lateral redistribution and deep percolation. Spatial surveys of soil moisture (Tromp-van Meerveld and McDonnell, 2006c) provide an estimate of changing average hillslope soil moisture in response to these temporally varying forces and indicate fairly wet conditions during this period (Figure 1). However, the actual volumetric moisture content within the hillslope during this period remains uncertain due to the lack of a site-specific calibration for the AquaPro moisture probe calibration (see Tromp-van Meerveld and McDonnell, 2006c for details).

The use of drainage simulations to estimate initial storage within a hillslope is a common practice in both groundwater and hillslope model applications (Binley and Beven, 1992; Loague *et al.*, 2005). To estimate pre-storm moisture conditions within the hillslope, we simulated the draining of the hillslope from near saturation over a 48-day initialization period of natural throughfall (9 February 2002 through 29 March 2002) shown in Figure 4. We restricted the model initialization period to 48 days of natural throughfall because of the computation expense to run the model through an earlier large rainstorm on 8 February 2002. Figure 4 shows the changing total hillslope storage during the 48-day initialization period compared to changes recorded by field surveys. Original field survey data (Figure 1 in units of % AquaPro; see also Tromp-van Meerveld and McDonnell, 2006a) were converted to volumetric moisture content (θ) using an approximate conversion (θ = (%AquaPro)/2·2; Tromp-van Meerveld, personal communication), and integrated in space to estimate hillslope-scale storage (m^3) for the individual surveys, assuming an effective porosity of 0·58. For both the simulation and field surveys, changes in hillslope storage are referenced to the final day of the initialization period (29 March 2002). By 19 February 2002, modelled changes in hillslope storage appear similar to observations. By the end of the initialization simulation, simulated trenchflow was small (0·035 l/15 min), which is fairly consistent with the no-flow observations at the Panola hillslope during inter-storm periods. After 48 days, the average hillslope degree of saturation was 0·63, with individual gridblock saturations ranging from 0·60 to 0·66 and no presence of full saturation or transient groundwater. Normalized by soil depth, the spatial variation in soil moisture across the hillslope was very small (<1%), consistent with field observations. These initial conditions created by simulating 48 days of natural rainfall were used here to represent a wet hillslope with no trench flow prior to the 30-March-2002 rainstorms. A similar initialization simulation excluding natural rainfall resulted in a significantly dryer hillslope (average saturation of 0·54). Initial conditions were recreated for each change in parameterization (see Section on Investigating effective model parameterization).

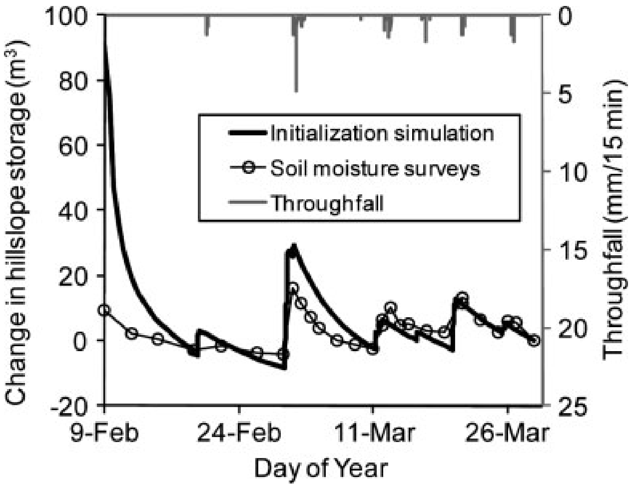

Figure 4. Reference model simulation of the observed change in hillslope soil water storage during the 48-day initialization period prior to the 30-March-2002 rainstorm. Observed and simulated changes in hillslope storage are zeroed with respect to the survey collected on Day 88, immediately prior to the 30-March-2002 rainstorm

30-MARCH-2002 RAINSTORM

The 63-mm 30-March-2002 event was chosen for model testing because it generated sub-surface stormflow at the trench, and extensive water level and soil moisture time series were available to initialize the model antecedent moisture conditions. It is also a temporally complex rainstorm, with three distinct sub-events: a small low-intensity event in the morning (12 mm), a second low-intensity event in the afternoon (12 mm) followed by a very intense thunderstorm in the evening of the same day (37 mm) (Tromp-van Meerveld and McDonnell, 2006b). The range in rainfall intensities and complexity of this rainstorm make it an ideal event for model testing. The observed trench flow [e.g. Figure 5a, panel (i)] exhibits a double peak hydrograph generated from both matrix and pipeflow, summed together here and represented as total trench flow. The small early discharge peak derives from the left-hand side of the trench (when looking upslope) where soils are shallow and have a low water storage capacity (Figure 2). A total of ~11% of the throughfall volume (6·4 mm) was delivered by trench outflow (Table II). Soil moisture surveys collected prior to, during and after the 29-March-2002 rainstorm

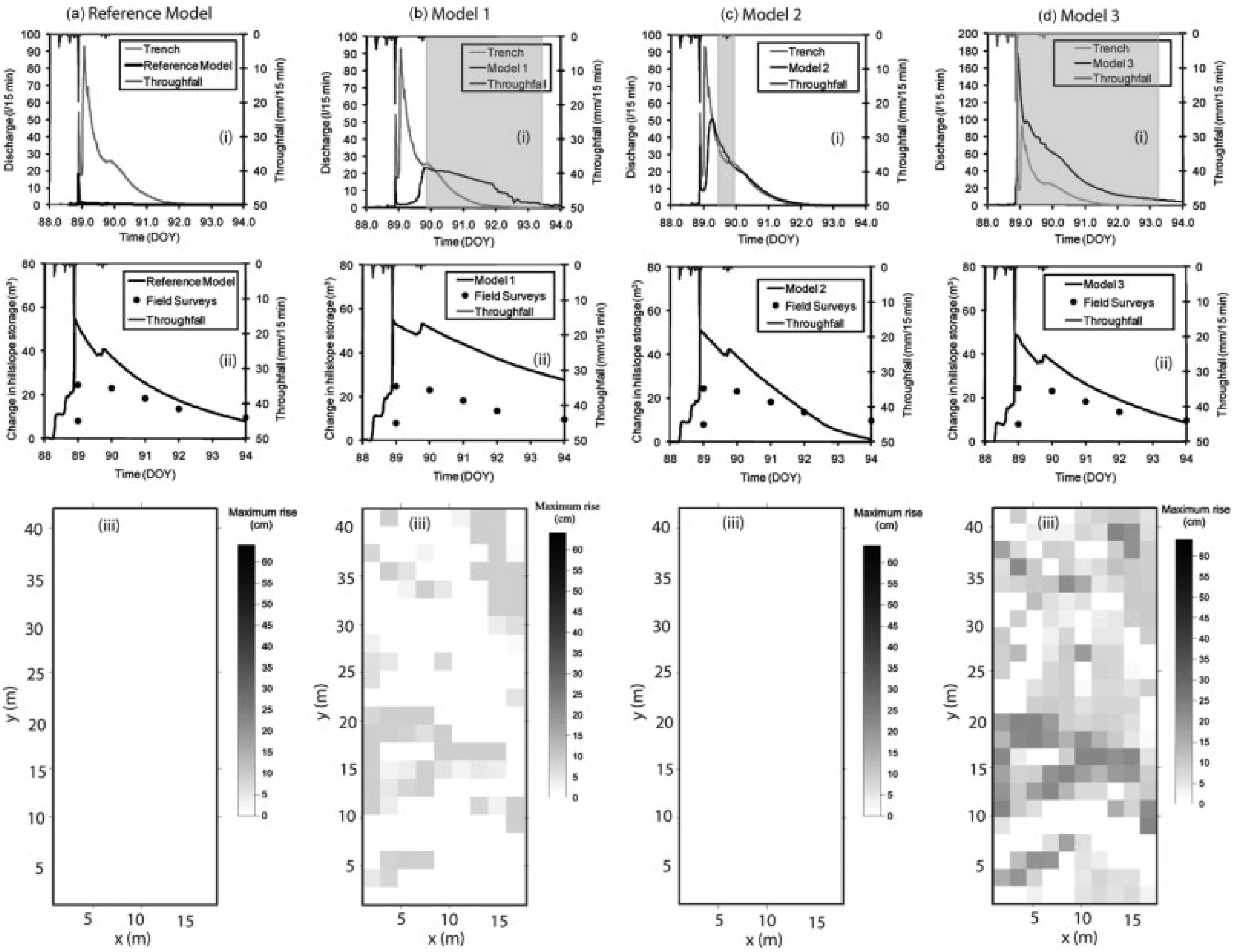

Figure 5. Comparison of observed and simulated hydrological response for the 63 mm 30-March-2002 rainstorm for (a) Reference Model, (b) Model 1, (c) Model 2 and (d) Model 3. For each model, panels (i), (ii) and (iii) illustrate instantaneous trench flow, change in hillslope storage (zeroed at day 88, 12 : 00 a.m.) and simulated maximum water table rise (cm) respectively in plan view of the hillslope. Shading in panel (i) indicates the time period for which transient saturation at the soil–bedrock interface is simulated somewhere on the hillslope. In panel (iii), *x*-axis is distance from left-hand side of trench and *y*-axis indicates distance upslope from trench

Table II. Mass balance for the 63-mm rainstorm on 30 March 2002

Scenario	$k_{\text{soil}}/k_{\text{bed}}$ ratio	Soil k_h/k_v ratio	Drainable porosity	Initial hillslope average saturation	Infiltration[a] (%)	Trenchflow[b] (%)	Change in storage[c] (%)	Loss to bedrock (%)[d]
Observed	110 : 1	Unknown	~0·15–0·20[e]	~0·58 (converted from AquaPro)	100	11	18	71
Reference	100 : 1	1 : 1	0.20	$0\cdot 63^{\text{f Ref sim}}$	100	<0·5	13	86
Model 1	5 000 : 1	1 : 1	0.20	$^{\text{f}}{}_{\text{sim}}$	100	9	46	45
Model 2	6 300 : 1	2 : 1	0.20	$^{\text{f}}{}_{\text{sim}}$	100	9	2	89
Model 3	10 000 : 1	1 : 1	0.30	$^{\text{f}}{}_{\text{sim}}$	100	30	14	56

[a] Estimate of observed infiltration as throughfall = [(Rainfall × 0·95) − 1] in mm.
[b] Observed or simulated total cumulative flow at trenchface.
[c] Observed or simulated change in storage after 6 days (evaluated on Day 94·0).
[d] Indirect calculation of loss to bedrock from mass balance (both observed and simulated).
[e] Estimated from moisture-retention curves presented in Figure 3.
[f] Simulated by initialization period; rerun per model.
Mass components expressed in % of total estimated throughfall (58·9 mm).

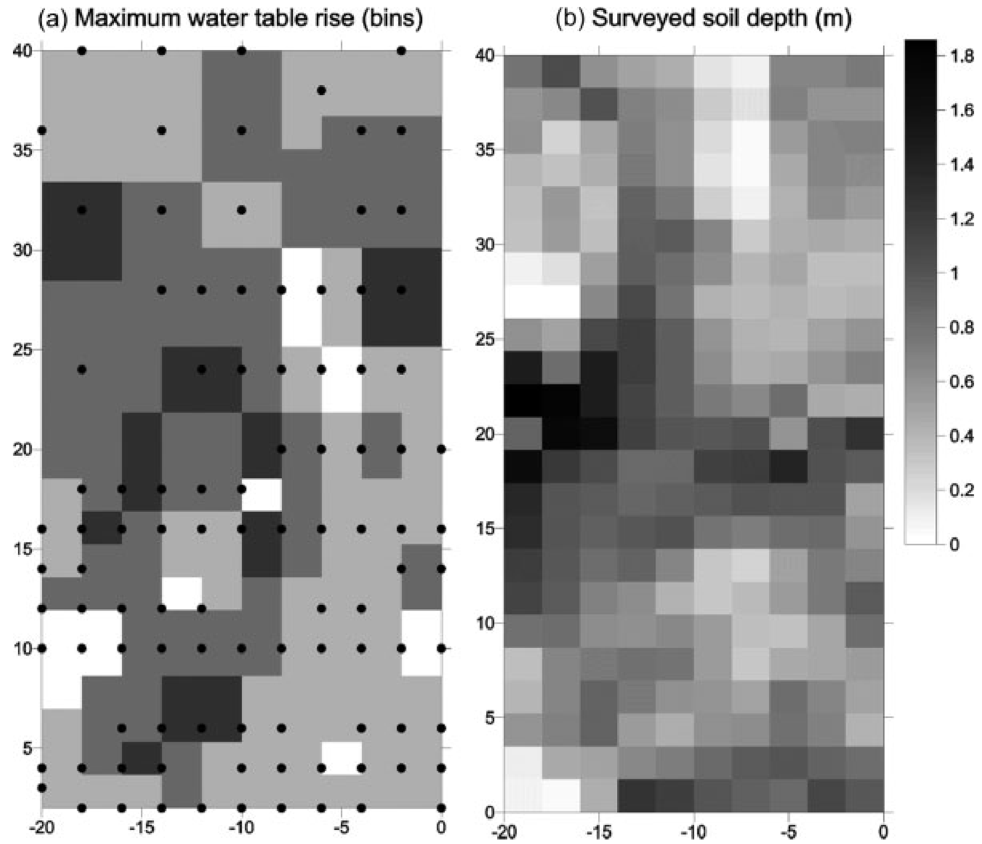

Figure 6. Observations of maximum water table rise (height of full saturation in crest-stage recorders) observed during 30-March-2002 rainstorm (a) and (b) surveyed soil depth (*m*). Black dots in (a) indicate the location of 107 crest-stage recorders. Observed values are binned in categories of high (>40 cm), medium (10–40 cm), low (<10 cm) and no rise (0 cm) from darkest to white

[Figure 5a, panel (ii)] allow estimation of changes in hillslope storage to which simulated hillslope storage can be compared. Storage within the hillslope peaks during the day of the storm. Six days after the rainstorm [4 April 2002 or Julian day 94) hillslope storage remains higher than pre-storm conditions, accounting for ~10·5 mm or 18% of throughfall (assuming mean soil porosity of 0·58). As a result, bedrock infiltration was estimated to be 42 mm or 71% of throughfall during the 6-day period.

The maximum transient water table rise occurring within the hillslope during rainstorms was determined by 135 crest-stage gauges installed across the hillslope (Tromp-van Meerveld *et al.*, 2007). The crest-stage gauges consisted of a combination of powdered cork and a 6·35-mm-diameter dowel in a 19-mm-diameter PVC pipe screened over the lowest 200 mm. Each gauge was augered to refusal, which presumably was the bedrock surface. The gauges were located on an approximately 2 m × 2 m grid across the lower 16 m of the study hillslope and an irregular but approximately 4 m × 4 m grid across the remainder of the hillslope (Figure 6a). The highest recorded water table rise ($\leq$64 cm) occurred in locations where the soil profile was deep enough to accommodate sub-surface stormflow but there is no significant correlation between soil profile thickness (Figure 6b) and the maximum water table rise observed during this storm. The timing of water table response observed by 29 individual recording wells indicates that the transient saturation began ~6·9 h after the start of the rainstorm (DOY 88·61) and ended approximately 7 days later.

EVENT SIMULATION RESULTS

30-March-2002 simulation

Initial simulation results of rainfall response using best-guess parameters (Table I), herein referred to as the Reference Model simulation, are compared with the observations in Figure 5a and Table II. The overall trend in dynamic storage within the hillslope is moderately well reproduced during most of the 6-day period [Figure 5a (ii)]. The magnitude of storage change is underestimated during Day 90 and 91, but is well simulated during Day 92 and 94. Approximately 86% of throughfall infiltrated bedrock, over-estimated by ~15% and, consequently, modelled sub-surface flow in the trench was small (0·5%) compared to observed (11%). An early response from areas close to the trench face accounts for the simulated peak shown in Figure 5a (i). No 'fill-and-spill' delivery from upslope saturation in bedrock depressions occurred in this simulation. In fact, although the overall trend in storage appears to be well represented and the magnitude of storage either well simulated or over-estimated, no water table developed [Figure 5a (i) and (iii)], indicating that, in the modelled system, the dynamic redistribution

Table III. Sensitivity of simulated trenchflow and hillslope storage to model parameters

Parameter/observation	Total sensitivity (unitless) (trench flow and hillslope storage)	Sensitivity (unitless) (hillslope storage only)	Sensitivity (unitless) (trench flow only)
log(k)—horizontal soil	44	**2**	**42**
log(k)—vertical soil	**350**	**316**	**34**
log(k)—vertical bedrock	14	13	1
Porosity—soil	**475**	**458**	17
Porosity—bedrock	2	1·8	0·2
log(p_e)—soil	75	67	8
m—soil	**413**	**373**	**40**
log(p_e)—bedrock	23	22	1
m—bedrock	46	44	2
Change in hillslope storage	**1294**	**1294**	—
Total trench outflow	147	—	147

k, absolute permeability; m, pore-size distribution index or van Geneuchten shape parameter; p_e, air entry pressure ($1/\alpha$).
Each parameter was prescribed ±10% variation. Bold numbers indicate highest sensitivities.

of infiltrated water differed significantly from the observations. In the Reference Model simulation, travel times of infiltrating water laterally and vertically through the soil profile above the soil–bedrock interface and into the underlying bedrock unit combined such that full saturation at the soil–bedrock interface did not form. The small amount of trench flow delivered downslope was delivered by unsaturated flow.

Sensitivity analysis

To examine the range of the possible behaviour of the Panola hillslope model, we evaluated the sensitivity of the hillslope water balance components to the Reference Model parameterization using iTOUGH2. Although the soil domain was parameterized as being isotropic (with soil vertical and horizontal permeabilities the same) in the Reference Model, the sensitivity analysis allowed each to vary separately. The hillslope water balance showed strongest sensitivity to soil porosity, m (pore-size distribution index or van Genuchten shape parameter), and vertical soil permeability with hillslope storage the most sensitive water balance component (Table III, bold numbers). Small changes in both soil porosity and m had a strong effect on hillslope storage. Small decreases in soil vertical permeability retarded travel times to depth and reduced loss to the underlying bedrock. The sensitivity of simulated total trench flow to these parameters was significantly smaller and the influence (and order of importance) of model parameters was less apparent than for simulated hillslope storage. For trench flow, soil horizontal and vertical permeability and the van Genuchten shape parameter m were the most sensitive parameters.

Of particular interest was that trench flow and hillslope storage were not sensitive to bedrock permeability. The experimental evidence from the Panola hillslope (Tromp-van Meerveld and McDonnell, 2006b; Tromp-van Meerveld *et al.*, 2007) suggests that the fill-and-spill phenomenon is a function of the contrast in permeability between soil and bedrock layers and thus is likely to be a controlling model parameterization. This result did not match with our initial hypothesis of the strong potential influence of contrasting hydraulic conductivity of the soil and bedrock domains, in either keeping water stored within the hillslope (and not lost to depth) or in the development of trenchflow downslope. However, further exploration of parameter space defined by soil and bedrock permeabilities showed that this initial evaluation of sensitivity was strongly dependent on the Reference Model parameterization (Table I) from which the linear Perturbation Method (with 10% variations in parameter values) was made. A grid-search of parameter space, defined by a changing ratio of soil to bedrock vertical permeability (Figure 7a, 19 successful simulations), revealed a non-linear effect of bedrock permeability on the components of the hillslope water balance (Figure 7b, c and d). As shown in Figure 7b, significant increases in trenchflow were generated for soil-bedrock permeability ratios greater than ~2000. Figure 7c and d shows a corresponding increase in hillslope storage and a decrease in loss of water to bedrock.

Investigating effective model parameterization

In this section, we further explore model parameterization of the soil and bedrock domains and resulting simulation of the integrated hillslope hydrological response and the timing and distribution of the maximum water table development at the soil–bedrock interface.

Soil–bedrock permeability ratio. Model 1, one of the 19 simulations summarized in Figure 7 (open circled symbol), illustrates the changes in the internal hillslope dynamics and trench flow delivery mechanisms as a consequence of a larger contrast in soil and bedrock permeability. In Model 1, soil permeability (Figure 7a, y-axis) was identical to the Reference Model value and bedrock permeability (x-axis) was decreased by a factor of 50 (from $2{\cdot}5 \times 10^{-13}$ m^2 ($\log(k_{\mathrm{bed}}) = -12{\cdot}6$) to $5{\cdot}0 \times 10^{-15}$ m^2 ($\log(k_{\mathrm{bed}}) = -14{\cdot}3$)). Although the simulated cumulative trench flow was within 2% of observations (Table II), the timing and peak magnitude of

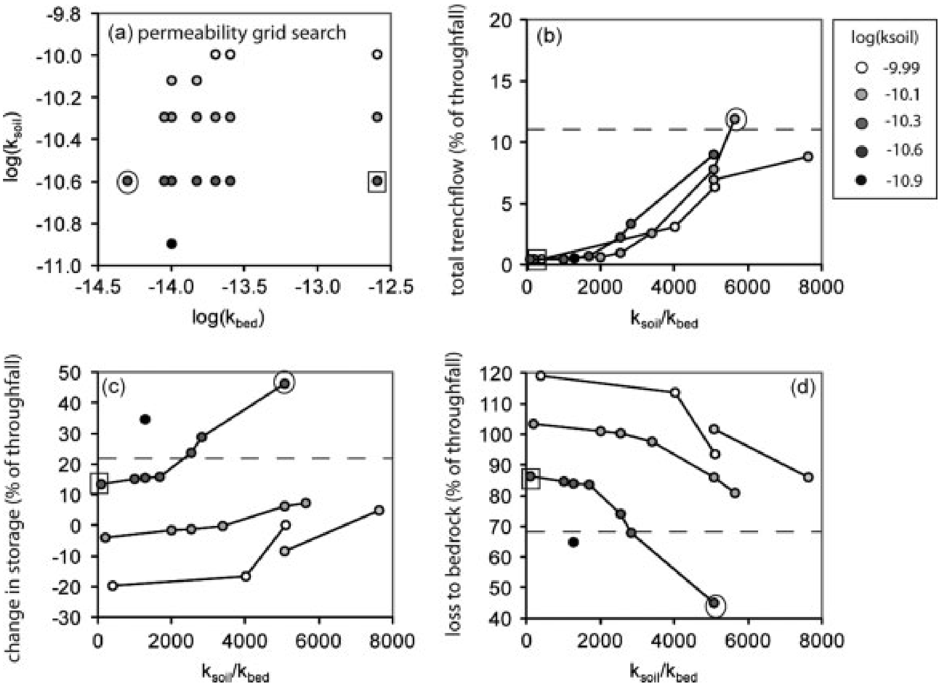

Figure 7. Sensitivity analysis showing non-linear response of hillslope water balance components with changing soil and bedrock permeability: (a) grid-search of parameter space defined by bedrock and soil vertical permeability, (b) total trench flow, (c) change in hillslope storage and (d) loss to bedrock as functions of the ratio of bedrock to vertical soil permeability. The water balance is evaluated at 12 : 00 a.m. Julian Day 94 (4 April 2002), approximately 6 days after the 30-March-2002 rainstorm. Dashed lines indicate the observed water balance values for the 30-March-2002 rainstorm. The Reference Model is highlighted by an open square and Model 1 by an open circle

delivery, in particular, the second peak in the hydrograph, were strongly lagged and dampened, respectively [Figure 5b (i)]. In this simulation, saturated sub-surface stormflow was generated at the soil–bedrock interface, the timing and spatial extent of which is visualized in Figure 5b (i) and (iii), respectively. Maximum water table heights (maximum of 9·4 cm) were lower and not as extensive as crest-stage observations (Figure 6a). After the initial storm response, the total change in storage remained >20% higher than observations throughout the simulation [Figure 5b (ii)].

Hillslope soil anisotropy. We examine the effect of soil permeability anisotropy on the hillslope water balance. Figure 5c (Model 2) illustrates a representative simulation with soil horizontal permeability of $2{\cdot}0 \times 10^{-10}$ m^2 ($\log(k_{\text{soil }x-y}) = -9{\cdot}7$; equivalent to $K_{\text{sat}} \sim 500$ cm/h) and a ratio of horizontal to vertical permeability of 2 : 1. Bedrock permeability was $1{\cdot}5 \times 10^{-14}$ m^2 ($\log(k_{\text{bed}}) = -13{\cdot}8$; equivalent to $K_{\text{sat}} = 0{\cdot}04$ cm/h), giving a ratio of vertical permeabilities (soil : bedrock) of 6300 : 1, compared to 100 : 1 and 5000 : 1 of the Reference Model and model 1, respectively. Model 2 estimates of trenchflow magnitude and timing were very similar to observations [Figure 5c (i); Table II]. However, for the Model 2 simulation, the maximum rise in the water table was too small, too short lived and the total storage decreased too rapidly within the last 3 days (Figure 5c). Evaluation of Model 2 by the timing and magnitude of trenchflow alone is encouraging but criteria provided by the internal hillslope measurements of water table development and integrated change in storage clearly identify the model as being poor. Further exploration of anisotropic parameter space defined by soil and bedrock permeability using Model 2 was unsuccessful in reproducing behaviour more representative of the Panola hillslope.

Hillslope soil drainable porosity. In Model 3, we examine the influence of soil drainable porosity on the hillslope water balance dynamics within the 3-D physics-based model. Drainable porosity of the soil domain is defined here as the difference between soil porosity and the volumetric moisture content, at $h = 0{\cdot}1$ m of pressure head (Weiler and McDonnell, 2004; Weiler *et al.*, 2005) and is varied as a function of soil porosity, van Genuchten shape parameter m to which hillslope storage and to a lesser extent, trenchflow showed significant sensitivity in earlier analysis (Table III) and residual saturation. In the reference case, soil drainable porosity was approximately 0·20, although possibly as low as 0·15 for some depths (e.g. 70 cm) (Table III; Figure 3). Figure 3 illustrates the alternative characteristic van Genuchten moisture-retention curve (dashed black line) used in Model 3. In comparison to the Reference Model moisture-retention curve, Model 3 represents a soil in which moisture content decreases to residual values at less negative pressure heads (drainable porosity of $\sim$0·30). This characterization was selected as an alternative representation of the effective functioning of a forested soil generating flow via both the soil matrix and macropores. An effective forest soil moisture-retention curve may be expected to deviate from the Panola small soil core data, allowing more water to drain quickly from the soil for a given pressure head

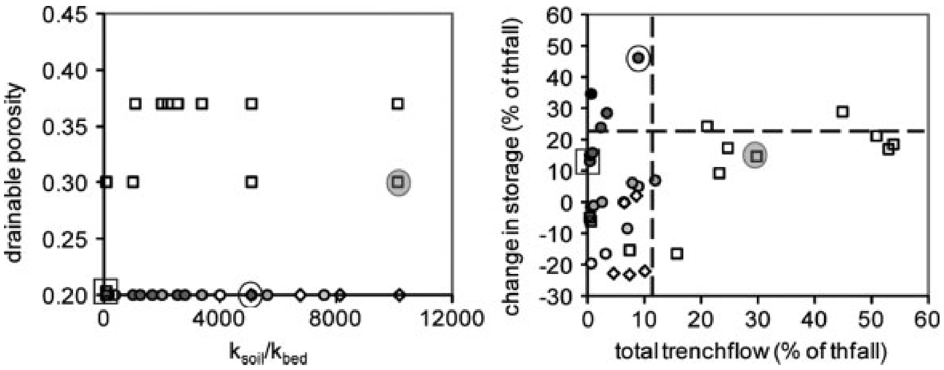

Figure 8. Effects of drainable porosity and soil–bedrock permeability on trenchflow and dynamic hillslope storage: (a) grid-search of parameter space defined by the ratio of soil to bedrock vertical permeability ($k_{\mathrm{soil}}/k_{\mathrm{bedrock}}$) and drainable porosity and (b) total trench flow versus change in hillslope storage. The water balance is evaluated at 12 : 00 a.m. Julian Day 94 (4 April 2002), approximately 6 days after the 30-March-2002 rainstorm. Dashed lines indicate the observed water balance values for the 30-March-2002 rainstorm. Model 3 is shown by the shaded circled symbol. Previous simulations from the grid search presented in Figure 7 are shown as circles, (Reference Model and Model 1 are highlighted by the boxed and open circled symbols respectively). Simulations examining the effects of anisotropy, represented by Model 2, are illustrated as diamonds (all with drainable porosity of 0·20)

close to saturation. A side-effect of the new representation of Model 3 was significantly more efficient during run times, with fewer very short time steps during the formation and depletion of the transient water table as compared to the Reference Model and Models 1 and 2.

Figure 5d illustrates Model 3 simulated water balance components and transient saturation formation at the soil–bedrock interface. The soil permeability was isotropic with a value of $5{\cdot}1 \times 10^{-11}$ m^2; ($\log(k_{\mathrm{soil}\ x-y})$ $= -10{\cdot}3$; equivalent to $K_{\mathrm{sat}} = 130$ cm/h), approximately twice that of the Reference Model. The bedrock permeability was the same as that in Model 1 ($5{\cdot}0 \times 10^{-15}$ m^2; ($\log(k_{\mathrm{bed}}) = -14{\cdot}3$; equivalent to $K_{\mathrm{sat}} = 0{\cdot}013$ cm/h), giving a ratio of vertical permeabilities (soil : bedrock) 2 orders of magnitude larger than that in the the Reference Model (Table II). Although overpredicting trenchflow ($\sim$2·8 times observations), the initial timing of the simulated discharge response at the trenchface was identical to observations [Figure 5d (i)]. Model 3 simulation produced a double peak in trenchflow; the second larger peak arrived 2·4 h earlier than observation. The trend in hillslope-scale dynamic storage [Figure 5d (ii)] was moderately well simulated through the 6-day period after the rainstorm. In this simulation, extensive transient saturation of depths within range of observations (maximum height of 21 cm) developed at the soil–bedrock interface and indicated saturated sub-surface stormflow moving downlope to the trench [Figure 5d (iii)]. The simulated saturation on the hillslope formed early during the rainstorm and persisted for $\sim$4 days.

Figure 8a shows a summary of all simulations performed within a parameter space defined by the ratio of soil to bedrock permeability ($k_{\mathrm{soil}}/k_{\mathrm{bed}}$) and soil drainable porosity. Figure 8b shows the variation in water balance components of total trenchflow and change in storage and the proximity to actual observations, represented by the crossed dashed lines. Note that there are several versions of Model 3 (cluster of square symbols right of crossed dashed lines) that, though over-estimate trenchflow, generate rapid timing of delivery (second peak of discharge ranging from several hours earlier to later than observed), have similar changes in storage to observations and develop spatially extensive transient saturation.

DISCUSSION

An evolving model description and process understanding

The simulations presented here inform Panola hillslope representation and illustrate some of the challenges in 3-D physics-based model development and the dichotomy of field-based measurements and effective model parameterization. Although parameterization was based on detailed site-specific field data, the uncalibrated Reference Model (Figure 5a) was unable to successfully simulate significant trenchflow and transient water table development, even though trends in dynamic storage, including the 48-day initialization period, were generally well simulated. Grid-search analysis of parameter space defined by bedrock and soil vertical permeability failed to provide a suitable calibrated model. For example, although the significant increase in soil to bedrock permeability ratio of Model 1 (Figure 5b) illustrated the ability to generate patterns of transient saturation and similar cumulative trenchflow (Table II), it did not mimic the rapid delivery of water to the trench face. With the introduction of anisotropy in soil permeability, Model 2 (Figure 5c) showed the best agreement between observed and simulated trenchflow magnitude and timing but further consideration of the patterns of transient water table showed very limited development of saturated conditions (simulated transient saturation was present for less than 0·5 days compared to 6 days of observation and rose only of 0·4 cm compared to 10–64 cm from observations), indicating that trenchflow delivery was dominated by unsaturated flow in contrast to the fill-and-spill mechanism observed in the field. This model deviated from observations for both internal storage dynamics and downslope delivery mechanisms.

Model 3, in which soil drainable porosity was increased, does not offer a fully calibrated fit to observational data. However, it was conceptually (if not exactly) successful in emulating the main behavioural characteristics of patterns and thresholds of the Panola hillslope, and, in this sense, offers a general calibration of phenomenon (Anderson and Bates, 2001) for the physics-based model. Significant sub-surface stormflow was simulated at the trench face with a double peak and similar timing. Spatially extensive transient saturation (up to 21 cm in depth) developed at the soil–bedrock interface, indicating that downslope water delivery and hillslope-scale dynamic storage are well simulated. The TOUGH2 Panola model demonstrates the importance of the ratio of soil to bedrock permeability in determining the water balance of the hillslope and the generation of sub-surface stormflow. This result is similar to the findings of Ebel *et al.* (2008) and Hopp and McDonnell (2009) who suggest that the characterization of this permeability contrast should be a key aim of field studies of hillslope function.

Discussion of model simplifications and process compromises

Our experience with development of the TOUGH2 Panola model is not unlike other 3-D physics-based model studies that ultimately build representation by observing model failure as much as success. Although the TOUGH2 model shows that it can capture key behavioural aspects of the Panola hillslope response as evaluated by three types of response data (hillslope-scale rainfall-runoff, internal dynamic storage and maps of transient saturation development), its shortcomings suggest a number of outstanding issues related to representation of sub-surface complexity. The best behavioural simulations (Model 3) were generated using effective K_{sat} within the range of those provided by soil core data (1–2 times Reference Model values; Table III), but required a significantly larger contrast between an effective bedrock and vertical soil permeability ($\sim$100 times the estimates from sub-hillslope scale sprinkler experiments). This discrepancy between field-scale sprinkler experiments and the TOUGH2 model parameterization is likely a result from too simple a model representation of the sub-surface. One possibility is that the uniform soil profile results in an over-estimation of the ratio of bedrock to soil permeability. Although no distinct layering is visible within the soil profile, K_{sat} values for greater depths within the soil profile suggest evidence of an exponential decline with depth (Zumbuhl, unpublished data). Further, simulations from two other Panola modelling efforts (Tromp-van Meerveld and Weiler, 2008; Hopp and McDonnell, 2009) were able to calibrate a spatially distributed conceptual (Hillvi) model and a 3-D physics-based model (HYDRUS-3-D), respectively, with bedrock K_{sat} estimates similar to those obtained from the sprinkler experiments ($\sim$0·58 cm/h). Both these models include some representation of decreasing K_{sat} with soil depth.

The behavioural TOUGH2 model also required a soil moisture-retention curve that mimicked a more efficient draining of water through the soil profile than indicated by the soil core data, via a quicker release of water and a more rapid increase in K_{sat} at high moisture contents (see also numerical experiments with Hillvi of Weiler and McDonnell, 2004). This is not inconsistent with the influence of macropores or preferential flowpaths delivering water to depth quicker, promoting the formation of transient saturation at the soil–bedrock interface. The inclusion of the irregular sub-surface topography generates a roughly similar pattern in water table development [compare Figures 5d (iii) and 6] and provides evidence that the soil–bedrock interface is one of the most significant components of heterogeneity of the hillslope system. But the current model does not reproduce the range of maximum water table depths ($\leq$64 cm in height) during the 30-March-2002 rainstorm. With the inclusion of decreasing K_{sat} with depth in their HYDRUS 3-D model, and a range of moisture-retention curves not dissimilar to soil core data for each layer within the soil profile, Hopp and McDonnell (2009) successfully simulated pore water pressures at six locations on the Panola hillslope during a 1996 rainstorm, again suggesting that representation of a layered soil domain may provide a key component of additional heterogeneity on the Panola hillslope currently not included in the TOUGH2 model. Other studies have suggested similar potential for the importance of spatial heterogeneity missing from model representation. In their 3-D physics-based simulation of the CB1 catchment, Ebel *et al.* (2008) concluded that additional characterization of the deeper sub-surface and its ability to generate fracture flow to the shallower near surface was needed to more accurately predict near-surface hydrological response and in particular distributed piezometric pore water pressures. In their comparison of synthetic homogeneous and heterogeneous hillslopes, Fiori and Russo (2007) found that increased degrees of heterogeneity enhanced sub-surface stormflow and streamflow generation, increased velocities and generated faster transient saturation within a modelled hillslope. Laine-Kaulio (2008) also found that physics-based simulations of observed flow velocities and chloride tracer concentrations within a shallow forested hillslope were most successful when incorporating decreased K_{sat} values with soil depth.

The TOUGH2 Panola model shortcomings suggest greater sub-surface complexity, beyond the irregular soil depth, is required to better emulate hillslope response. Preliminary TOUGH2 simulations using an approximate exponential soil K_{sat} profile in conjunction with the irregular soil depth resulted in simulations with extremely long run times (long periods of extremely small time steps) and difficulty in reaching convergence, often associated with phase appearance and disappearance (Pruess, 2004; Finsterle *et al.*, 2008); in this case, the formation and depletion of the transient water table, and as a result, were not pursued further. Further investigations with grid generation may be one possible solution to these numerical difficulties. Also, the current bedrock model layers do not resolve lateral spatial variability in moisture conditions within this lower domain (only vertical variability, as described in the section on Model domain). This model simplification may contribute to an 'over-conductive' bedrock domain, requiring lower effective K_{sat} compared to sprinkler experimental values, and dampen the development of transient saturation and its spatial variation within the soil domain above. Future Panola TOUGH2 model development should include a more complex bedrock representation, allowing both lateral and vertical dynamics. Future work may also choose to explore the effect of fracture flow within the bedrock domain. Ebel *et al.* (2008) improved Coos Bay hillslope pore water pressure simulation by adding zones of high and low K_{sat} and a simple representation of fracture flow. Sprinkler experiment estimates of K_{sat} at the Panola hillslope may themselves be an average of less conductive weathered granite matrix and more conductive bedrock fractures as observed on granite outcrops within the larger watershed (N. Peters, personal communication). There are numerous techniques currently being used in concert with a continuum approach to investigate representation of preferential flow in the vadose zone (e.g. macropore or fracture flow) and the amount of pore space active in transient storage, including use of heterogeneous property fields (e.g. K_{sat}, see Fiori and Russo, 2007), alternative moisture-retention functions (El-Kadi and Torikai, 2001; Podgorney and Fairley, 2008) and dual-continuum approaches (e.g. double-porosity, multiple interacting continua; see Wu and Qin, 2009 for a brief summary of approaches). The TOUGH2 family of codes is particularly well suited to examining many of these options, making these directions wholly feasible steps forward for the TOUGH2 Panola hillslope model.

3-D hillslope hydrological simulation: an experimentalist's view

The value of an uncalibrated model. Many studies have shown that calibration is necessary for a good model performance, particularly in the case of 3-D physics-based models (Loague and Vanderkwaak, 2002), but recent focus on being able to describe the hydrology of ungauged basins (Sivapalan *et al.*, 2003; Pomeroy *et al.*, 2007) highlights the importance of developing uncalibrated models that function well with limited site-specific data. Pomeroy *et al.* (2007) suggest '...a purpose-built physically based model based on a good understanding of the principles and characteristics of hydrology in a basin, with an appropriate structure and appropriate spatial resolution and parameter selection, should have a good chance of simulating the hydrological cycle including the water balance, streamflow, and other variables of interests such as soil moisture and snow accumulation...'.

The departure of effective parameterization from field-based characterization is well documented (Beven, 1995; Bloschl and Sivapalan, 1995; Grayson *et al.*, 1995) and is a real challenge in development of uncalibrated models (3-D physics-based or other) for the ungauged hillslope or catchment. Given the recent interest in 3-D physics-based hillslope representation, this issue remains extremely relevant and needs to be highlighted. Forested soils have shown significant differences between laboratory core-scale and larger scale field-based K_{sat}, the latter providing higher values attributed to the presence of macropores (Buttle and House, 1997; Laine-Kaulio, 2008). Use of pedrotransfer functions to determine K_{sat} and effective soil moisture-retention curves suffers from the same issue of scale (see general discussion in Beven (2001)). Mertens *et al.* (2005) found that optimized effective values of K_{sat} and moisture-retention curves generated using a 1-D MIKE-SHE model were overall different than *in situ* or laboratory values, concluding that they could not replicate field-scale behaviour with laboratory-scale parameterization. Clearly shallow forested hillslopes are not a superposition of soil cores, as soil cores are too small to include the large pores or preferential flow paths In their study of the Coos Bay catchment, Ebel *et al.* (2007b) recommend *in situ* measurements of soil–water retention as being most supportive of physics-based model parameterization. In the search for appropriate structure and parameter selection for hillslope models, recent modelling studies suggest that we need to include the representation of preferential flow or sub-surface heterogeneity (Weiler and McDonnell, 2007; Ebel *et al.*, 2008; Tromp-van Meerveld and Weiler, 2008), exploration of which is taking on many alternative forms, including decreasing K_{sat} profiles (Laine-Kaulio, 2008; Hopp and McDonnell, 2009), statistically generated property fields (Fiori and Russo, 2007) and dual-continuum approaches capable of generating rapid streamflow, trenchflow and sub-surface response.

So what data is most useful when modeling hillslope hydrological response?. 3-D physics-based models require a large number of parameters for which field and laboratory data can be of limited to no value. Loague *et al.* (2005) argue that full characterization of small catchment systems is simply not possible with the types of data that we can currently collect, even on detailed research hillslopes or catchments. Some of the data that are relatively easy to get (e.g. core-scale K_{sat}) are not necessarily very useful to model development (see above discussion) and distributed data that are most useful in constraining model parameterization (e.g. sub-surface topography, observations of state variables such as water table and pressure head) are difficult to obtain. Here, trench discharge data was insufficient to characterize the Panola hillslope TOUGH2 model, and water table data was found to be more useful than average soil moisture conditions in determining appropriate model parameterization, similar to findings of Tromp-van Meerveld and Weiler (2008). The uncalibrated Reference Model showed similar trends in average hillslope soil moisture during the 48-day initialization period and through the 30-March-2002 rainstorm while unsuccessfully replicating runoff response and internal transient water table development (see for instance Figures 4 and 5). Use of pressure head or water table data in addition to discharge to constrain behavioural parameter comes with additional issues of data uncertainty and scale of measurement (see Freer *et al*., 2004 for discussion), although one might argue that the discrepancy between model and measurement scales might be less of an issue for a hillslope than a full catchment (with different topographic positions). Although the use of distributed state variable data is considered critical to physics-based hillslope modelling studies, Ebel and Loague (2008) suggest that we may need even higher temporal resolution (e.g. $<$10 min) of data (pressure head, discharge and soil water) for a more in-depth evaluation of the ability of the physics-based model to generate a rapid response of the shallow sub-surface.

CONCLUSIONS

In this study, the TOUGH2 simulator was used to develop a 3-D physics-based model of the Panola hillslope that could ultimately allow examination of the physical controls (e.g. ratio of soil to bedrock ratio) on the patterns, thresholds and macroscale behaviour of this hillslope. Once calibrated, the TOUGH2 Panola model could simulate general behavioural characteristics of hillslope response, including dynamic hillslope-scale change in storage, development of a spatially extensive transient water table at the soil–bedrock interface and generation of significant sub-surface stormflow but an uncalibrated model was unsuccessful. Further, there were significant limitations to the best behavioural model, suggesting that additional model complexity (e.g. decreasing soil K_{sat} profile or spatial variation in bedrock conditions and/or properties) beyond the irregular soil depth is needed to improve model performance. Components of the Panola hillslope water balance were strongly controlled by the ratio of soil to bedrock K_{sat}, (in addition to soil K_{sat} anisotropy and drainable porosity), supporting recent modelling studies, which suggest that this permeability contrast is key to characterizing hillslope function. Our experience with the TOUGH2 Panola hillslope model offers a transparent illustration of the many challenges involved in physics-based modelling (e.g parameterization, process representation, numerics), and the evolution of improved model structure and process understanding using a variety of data types. As suggested by Ebel and Loague (2006), the value of physics-based simulation of hillslope and small catchment response will be the examination of their failure to replicate experimental observations and the changes it will bring about to the models themselves.

REFERENCES

Anderson MG, Bates PD. 2001. Hydrological science: model credibility and scientific integrity. In *Model Validation: Perspectives in Hydrological Sciences*, Anderson M, Bates P (eds). Wiley: Chichester, UK; 1–10.

Beven K. 1995. Linking parameters across scales: subgrid parameterizations and scale dependent hydrological models. In *Scale Issues in Hydrological Modelling*, Kalma JD, Sivapalan M (eds). Wiley: Chichester, UK; 263–281.

Beven K. 2001. *Rainfall-Runoff Modelling*, Wiley: Chichester, UK; 360.

Binley AM, Beven KJ. 1992. Three dimensional modelling of hillslope hydrology. *Hydrological Processes* **6**: 347–359.

Bloschl G, Sivapalan M. 1995. Scale issues in hydrological modelling: a review. *Hydrological Processes* **9**: 251–290. DOI: 10.1002/hyp.3360090305.

Burns DA, Hooper RP, McDonnell JJ, Freer JE, Beven CK. K. 1998. Base cation concentrations in subsurface flow from a forested hillslope: the role of flushing frequency. *Water Resources Research* **34**: 3535–3544. DOI: 10.1029/98WR02450.

Burns DA, McDonnell JJ, Hooper RP, Peters NE, Freer JE, Kendall C, Beven K. 2001. Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA). *Hydrological Processes* **15**: 1903–1924.

Burns DA, McDonnell JJ, Hooper RP, Peters NE, Kendall C. 2002. Transport of solutes in groundwater through a hillslope-riparian transition during two rain events at the Panola Moutain Research Watershed, Georgia. *EOS, Transactions of the American Geophysical Union* **83**(47): Fall Meet. Suppl., H51D-05.

Bushey JT, Driscoll CT, Mitchell MJ, Selvendiran P, Montesdeoca MR. 2008. Mercury transport in response to storm events from a northern forest landscape. *Hydrological Processes* **22**: 4813–4826. DOI: 10.1002/hyp.7091.

Buttle JM, House DA. 1997. Spatial variability of saturated hydraulic conductivity in shallow macroporous soils in a forested basin. *Journal of Hydrology* **203**: 127–142.

Carsel RF, Parrish RS. 1988. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research* **24**: 755–769.

Cloke HL, Renaud JP, Claxton AJ, McDonnell J, Anderson MG, Blake JR, Bates PD. 2003. The effect of model configuration on modelled hillslope-riparian interactions. *Journal of Hydrology* **279**: 167–181. DOI: 10.1016/s0022-1694(03)00177-x.

Ebel BA, Loague K. 2006. Physics-based hydrologic-response simulation: seeing through the fog of equifinality. *Hydrological Processes* **20**: 2887–2900. DOI: 10.1002/hyp.6388.

Ebel BA, Loague K. 2008. Rapid simulated hydrologic response with the variably saturated near surface. *Hydrological Processes* **22**: 464–471. DOI: 10.1002/hyp.6926.

Ebel BA, Loague K, Dietrich WE, Montgomery DR, Torres R, Anderson SP, Giambelluca TW. 2007a. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 1. Sprinkling experiments. *American Journal of Sciences* **307**: 678–708. DOI: 10.2475/04.2007.02.

Ebel BA, Loague K, Montgomery DR, Dietrich WE. 2008. Physics-based continuous simulation of long-term near-surface hydrologic response for the Coos Bay experimental catchment. *Water Resources Research* **44**: W07417. DOI: 10.1029/2007WR006442.

Ebel BA, Loague K, VanderKwaak JE, Dietrich WE, Montgomery DR, Torres R, Anderson SP. 2007b. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 2. Physics-based simulations. *American Journal of Sciences* **307**: 709–748. DOI: 10.2475/04.2007.03.

El-Kadi AI, Torikai JD. 2001. Identifying variably saturated water-flow patterns in a steep hillslope under intermittent heavy rainfall. *Hydrogeology Journal* **9**: 231–242. DOI: 10.1007/s100400100129.

Falta R. 2003. Modeling sub-grid-block-scale dense nonaqueous phase liquid (DNAPL) pool dissolution using a dual-domain approach,. *Water Resources Research* **39**: 1360. DOI: 10.1029/2003WR002351.

Finsterle S. 2000. *iTOUGH2 User's Guide*, LBNL-40040. Lawrence Berkeley National Laboratory: Berkeley; 130.

Finsterle S, Ahlers CF, Trautz RC, Cook PJ. 2003. Inverse and predictive modeling of seepage into underground openings. *Journal of Contaminant Hydrology* **62–63**: 89–109. DOI: 10.1016/S0169-7722(02)00174-2.

Finsterle S, Doughty C, Kowalsky MB, Moridis GJ, Pan L, Xu T, Zhang Y, Pruess K. 2008. Advanced vadose zone simulations using TOUGH. *Vadose Zone Journal* **7**: 601–609. DOI: 10.2136/vzj2007.0059.

Fiori A, Russo D. 2007. Numerical analyses of subsurface flow in a steep hillslope under rainfall: The role of the spatial heterogeneity of the formation hydraulic properties. *Water Resources Research* **43**: W07445, DOI:10.1029/2006WR005365.

Freer JE, McDonnell JJ, Beven KJ, Brammer D, Burns DA, Hooper RP, Kendall C. 1997. Topographic controls on subsurface storm flow at the hillslope scale for two hydrologically distinct small catchments. *Hydrological Processes* **11**: 1347–1352.

Freer JE, McDonnell JJ, Beven KJ, Peters NE, Burns DA, Hooper RP, Aulenbach B. 2002. The role of bedrock topography on subsurface storm flow. *Water Resource Research* **38**: 1269, DOI:10.1029/2001WR000872.

Freer JE, McMillan H, McDonnell JJ, Beven KJ. 2004. Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures. *Journal of Hydrology* **291**: 254–277.

van Genuchten MT. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* **44**: 892–898.

Graham RC, Schoeneberger PJ, Anderson MA, Sternberg PD, Tice KR. 1997. Morphology, porosity, and hydraulic conductivity of weathered granitic bedrock and overlying soils. *Soil Science Society of America Journal* **61**: 516–522.

Grayson RB, Bloschl G, Moore ID. 1995. Distributed parameter hydrologic modelling using vector elevation data: Thales and Tapes-C. In *Computer Models of Watershed Hydrology*, Singh VP (ed.). Water Resources Publications: Highlands Ranch, CO; 669–696.

Hopp L, McDonnell JJ. 2009. Connectivity at the hillslope scale: identifying interactions between storm size, bedrock permeability, slope angle and soil depth. *Journal of Hydrology* **376**: 378–391. DOI: 10.1016/j.jhydrol.2009.07.047.

James AL, Oldenburg CM. 1997. Linear and Monte Carol uncertainty analysis for subsurface contaminant transport simulation. *Water Resources Research* **33**: 2495–2508.

Jones DP, Graham RC. 1993. Water-holding characteristics of weathered granitic rock in chaparral and forest ecosystems. *Soil Science Society of America Journal* **57**: 256–261.

Katsura S, Kosugi K, Yamamoto N, Mizuyama T. 2005. Saturated and unsaturated hydraulic conductivities and water retention characteristics of weathered granitic bedrock. *Vadose Zone Journal* **5**: 35–47. DOI: 10.2136/vzj2005.0040.

Laine-Kaulio H. 2008. *Subsurface flow in a forested till slope: Soil analysis, tracer experiments and physics-based modelling*. PhD thesis, Henlsinki University of Technology: Finland; 95.

Loague K, Heppner CS, Abrams RA, Carr AE, Vanderkwaak JE, Ebel BA. 2005. Further testing of the Integrated Hydrology Model (InHM): event-based simulations for a small rangeland catchment located near hickasha, Oklahoma. *Hydrological Processes* **19**: 1373–1398. DOI: 10.1002/hyp.5566.

Loague K, Heppner CS, Mirus BB, Ebel BA, Carr AE, BeVille SH, Vanderkwaak JE. 2006. Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology,. *Hydrological Processes* **20**: 1231–1237. DOI: 10.1002/hyp.6179.

Loague K, Vanderkwaak JE. 2002. Simulating hydrological response for the R-5 catchment: comparison of two models and the impact of the roads. *Hydrological Processes* **16**: 1015–1032. DOI: 10.1002/hyp.316.

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

McDonnell JJ, Freer JE, Hooper R, Kendall C, Burns DA, Beven KJ, Peters NE. 1996. New method developed for studying flow on hillslopes. *Eos, Transactions of the American Geophysical Union* **77**: 465–466.

McDonnell JJ, Sivapalan M, Vaché K, Dunn SM, Grant G, Haggerty R, Hinz C, Hooper R, Kirchner J, Roderick ML, Selker J, Weiler M. 2007. Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology. *Water Resources Research* **43**: W07301, DOI:10.1029/2006WR005467.

McIntosh J, McDonnell JJ, Peters NE. 1999. Tracers and hydrometric study of preferential flow in large undisturbed soil cores from the Georgia Piedmont,USA. *Hydrological Processes* **13**: 139–155.

Mertens J, Madsen H, Kristensen M, Jacques D, Feyen J. 2005. Sensitivity of soil parameters in unsaturated zone modelling and the relation between effective, laboratory and in situ estimates. *Hydrological Processes* **19**: 1611–1633. DOI: 10.1002/hyp.5591.

Morris DA, Johnson AI. 1967. *Summary of hydrologic and physical properties of rock and soil materials as analyzed by the Hydrologic Laboratory of the U.S. Geological Survey*. U.S. Geol. Surv. Water-Supply Paper 1839-D. 42.

Mualem Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research* **12**: 513–522.

Mulligan M, Wainwright J. 2004. Modelling and model building. In *Environmental Modelling, Finding Simplicity in Complexity*, Wainwright J, Mulligan M (eds). John Wiley & Sons Ltd: 7–73.

Oldenburg CM, Pruess K. 1993. On numerical modeling of capillary barriers. *Water Resources Research* **29**: 1045–1056.

Podgorney RK, Fairley JP. 2008. Investigation of episodic flow form unsaturated porous media into a macropore. *Vadose Zone Journal* **7**: 332–339. DOI: 10.2136/vzj2006.0107.

Pomeroy JW, Gray DM, Brown T, Hedstrom NR, Quinton WL, Granger RJ, Carey SK. 2007. The cold regions hydrological model: a platform for basin process representation and model structure on physical evidence. *Hydrological Processes* **21**: 2650–2667. DOI: 10.1002/hyp.6787.

Pruess K. 1991. *TOUGH2—A General Purpose Numerical Simulator for Multiphase Fluid and Heat Flow*, Lawrence Berkeley Laboratory, University of California: Berkeley, CA.

Pruess K. 2004. The TOUGH codes—a family of simulation tools for multiphase flow and transport processes in permeable media. *Vadoze Zone Journal* **3**: 738–746.

Pruess K, Oldenburg C, Moridis G. 1999. *TOUGH2 User's Guide, Version 2.0*, Lawrence Berkeley National Laboratory: Berkeley, CA; 197 p.

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

Sivapalan M, Takeuchi K, Franks S, Gupta V, Karambiri H, Lakshmi V, Liang X, McDonnell JJ, Mendiondo E, O'Connell P, Oki T, Pomeroy JW, Schertzer D, Uhlenbrook S, Zehe E. 2003. IAHS decade on predictions in ungauged basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal* **48**: 857–880.

Troch PA, Paniconi C, Emiel van Loon E. 2003. Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response. *Water Resources Research* **39**: 1316, DOI:10.1029/2002WR001728.

Tromp-van Meerveld HJ, James AL, McDonnell JJ, Peters NE. 2008. A reference dataset of hillslope rainfall-runoff response, Panola Mountain Research Watershed, USA. *Water Resources Research* **44**: W06502, DOI:10.1029/2007WR006299.

Tromp-van Meerveld HJ, McDonnell JJ. 2006a. Threshold relations in subsurface stormflow 1. A 147-storm analysis of the Panola hillslope. *Water Resources Research* **42**: W02410. DOI: 10.1029/2004 WR003778.

Tromp-van Meerveld HJ, McDonnell JJ. 2006b. Threshold relations in subsurface stormflow 2. The fill and spill hypothesis. *Water Resources Research* **42**: W02411. DOI: 10.1029/2004WR003800.

Tromp-van Meerveld HJ, McDonnell JJ. 2006c. On the interrelations between topography, soil depth, soil moisture, transpiration rates and species distribution at the hillslope scale. *Advances in Water Resources* **29**: 293–310. DOI: 10.1016/j.adv watres.2005.02.016.

Tromp-van Meerveld HJ, Peters NE, McDonnell JJ. 2007. Effect of bedrock permeability on subsurface stormflow and the water balance of a trenched hillslope at the Panola Mountain Research Watershed, Georgia, USA. *Hydrological Processes* **21**: 750–769. DOI: 10.1002/hyp.6265.

Tromp-van Meerveld I, Weiler M. 2008. Hillslope dynamics modeled with increasing complexity. *Journal of Hydrology* **361**: 24–40. DOI: 10.1029/2007WR006299.

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)00271-3.

Weiler M, McDonnell JJ. 2007. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes,. *Water Resources Research* **43**: W03403, DOI:10.1029/2006WR004867.

Weiler M, McDonnell JJ, Tromp-van Meerveld HJ, Uchida T. 2005. Subsurface stormflow. In *Encyclopedia of Hydrological Sciences*, Chapter 112, Anderson MG (ed.). John Wiley & Sons Ltd: Chichester, England; Hoboken, NJ.

Wu Y-S, Qin G. 2009. A generalized numerical approach for modeling mulitphase flow and transport in fractured porous media,. *Communications in Computational Physics* **6**: 85–108.

Zumbuhl A. 1998. *Spatial modeling of soil depth and landscape variability in a small forested catchment*. MS thesis, State University of New York (SUNY): Syracuse, NY.