AN ABSTRACT OF THE DISSERTATION OF

Christopher Brian Graham for the degree of Doctor of Philosophy in Water Resources Science presented on December 11, 2008
Title: A Macroscale Measurement and Modeling Approach to Improve Understanding of the Hydrology of Steep, Forested Hillslopes

Abstract approved:

Jeffrey J McDonnell

The dominant controls on flow generation in steep, forested hillslopes are poorly understood. This dissertation examined the dominant flow processes operating at the hillslope scale, using a combined macroscale measurement and model development and analysis framework. Irrigation experiments at two steep forested hillslopes were conducted to isolate individual hillslope flow components and reveal the dominant controls on flow routing to the stream. A new perceptual model of flow processes at the hillslope scale was developed from these field experiments that included three key components: 1) A connected preferential flow network located at the soil/bedrock interface controls lateral water and solute transport, 2) The bedrock surface controls the subsurface flow routing, and 3) Bedrock is permeable, and acts as a sink for precipitation at the hillslope scale. These components formed the basis of a new, low dimensional numerical model that was used to represent, quantitatively, qualitative experimental findings. A multiple criteria calibration using hydrometric and tracer experimental data was conducted to evaluate the model and determine parameter identifiability. The model was able to adequately reproduce both hydrometric response to precipitation, and a tracer application breakthrough.

The model was then used within a virtual experiment framework to test the dominant controls on the whole storm precipitation / discharge threshold relationship at the catchment scale. The modeling experiments showed that the macroscale precipitation / discharge threshold was controlled by a balance between antecedent evaporation

(evaporation rate times antecedent drainage time), and geologic factors (bedrock permeability and subsurface storage volume). Overall, these findings suggest that making measurements at the scale of the processes one wishes to understand, and constructing numerical models based on the dominant processes at the same scale may be a first step towards the development of macroscale laws of hillslope behavior.

A Macroscale Measurement and Modeling Approach to Improve Understanding of the Hydrology of Steep, Forested Hillslopes

by

Christopher Brian Graham

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented December 11, 2008

Commencement June 2009

Doctor of Philosophy dissertation of Christopher Brian Graham presented on December 11, 2008.

APPROVED:

Major Professor, representing Water Resources Science

Director of the Water Resources Science Graduate Program

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Christopher Brian Graham, Author

ACKNOWLEDGEMENTS

I would like to thank Jeff McDonnell, who has been as valuable a mentor as I could imagine. Jeff has provided endless enthusiasm support for my ideas, and has been a great editor and sounding board as I finished this thesis.

I would also like to thank Julia Jones and the Ecosystem Informatics IGERT program for funding both my studies and my internships to Georgia, USA, and New Zealand. Thanks also to Katherine Hoffman for providing logistical support for the trip and field work expenses.

The irrigation experiment at the HJ Andrews would not have been possible without the help of many at the HJA. I would like to thank Matthew Bergen and John Moreau for providing field assistance and John Selker and Sherri Johnson for loan of TDR and meteorological equipment. I especially thank the McKenzie River Ranger District for providing sprinkler water during the experiment, and Kari O'Connell and Cheryl Friesen for coordinating logistics.

I would like to thank Jake Peters, Jagath Ekanayake, John Payne and Ross Woods for providing logistical and emotional support in the wilds of Atlanta, Georgia and Reefton, New Zealand. Without their help, these experiments would not be possible. I also thank Jim Freer, Ilja Tromp-van Meerveld, Kevin McGuire, Willem van Verseveld and Ross Woods for constructing and developing the field sites in Oregon, Georgia and New Zealand.

I would like to thank the past and current members of the watershed hydrology lab, Holly Barnard, Luisa Hopp, Taka Sayama, Cody Hale, Adam Mazurkiewicz, Willem van Verseveld, David Alley, David Callery, April James, Kellie Vache, Kevin McGuire, Ilja Tromp-van Meerveld and others. My discussions with this group went a long way towards my ideas outline in these papers.

I would like to thank my family, who have been an endless support during this process. My parents have supported me both financially and emotionally, with nothing less than total confidence that I could finish. My brother, who kept things in perspective, and helped me relieve the stress on the frolfing fields. Thanks also to my boy, Buddy, and his long lost sister, Alice, who comforted and entertained me throughout.

I would especially like to thank Sarah, who spent 4 long years in Corvallis helping me finish, including three months in Atlanta, scene of countless bulldozer attacks on Walgreen ATMs, and another three in a hut in the middle of nowhere, with no phone, TV, radio, internet, but lots of large mosquitoes in the bathroom. Without Sarah holding me up, I would have been lost long ago.

CONTRIBUTION OF AUTHORS

Chapter 1: Jeff McDonnell provided critical reviews and editing support

Chapter 2: Jeff McDonnell provided critical reviews and editing support

Chapter 3: Holly Barnard and Willem van Verseveld provided technical and logistical support, and assisted in the editing. Jeff McDonnell provided critical reviews and editing support

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

1. Introduction

1.1. Introduction

Hillslopes are a fundamental landscape unit for many fields of environmental science (Wagener et al., 2004). Despite the importance of forested hillslopes as a filter in the landscape for hydrological, ecological, biogeochemical, and geomorphological signals and decades of field monitoring of heavily instrumented field sites, hillslope hydrology is poorly understood, (Bonell, 1993). Basic questions remain regarding the subsurface structure and dominant flow processes operating at the hillslope scale. Figure 1 illustrates the enormous complexity in subsurface flow processes that creates major challenges in understanding, conceptualizing and predicting the flow and transport in forest hillslope systems. As a result, we still lack the macroscale laws that might describe whole hillslope behavior (Dooge, 1986).

Some of the difficulties in making progress in hillslope hydrology are due to the standard methodologies used in the field. Our measurements are restricted to the point scale (e.g. time domain reflectometry, tensiometers) or at the base of a hillslope where we might construct a collection trench (e.g. for measuring lateral subsurface flow or soil water chemistry). While point scale measurements provide us very accurate information over the states and stores of water and nutrients at very small scales, they are often difficult to reconcile with hillslope trenching information that integrates over large spatial and temporal scales. We still lack a mechanistic linkage between the extremes of our point scale measures and the integrated hillslope flow response. While the two approaches have improved our understanding subsurface flow processes at the hillslope scale (Bonell, 1993; Kirkby, 1978; Weiler et al., 2006), the natural complexity and heterogeneity of these systems, combined with the inherent difficulty in measuring subsurface processes have been barriers to progress. Most hillslope studies rely on passive storm monitoring, and generally only measurements of the inputs, outputs and isolated points of internal state conditions are made. The boundary conditions are also rarely known, adding to the difficulty in measuring, modeling and understanding subsurface processes

This thesis explores a new way of conducting hillslope scale experiments, where inputs and boundary conditions are controlled, have many advantages over passive storm monitoring. Such macroscale experiments can test processes at the scale that they are operating (i.e. Brooks et al., 2004). Rather than measure flow processes at the scale that is convenient for our instrumentation (generally at the soil core scale), this thesis shows how these experiments integrate system response at the scale at which flow processes are occurring.

This thesis also uses hillslope scale irrigation to generate subsurface flow conditions. By generating steady state flow and storage through irrigation, the dominant flow processes can be revealed in a way not possible with the transient, dynamic conditions generally observed during storm monitoring. Finally, destructive sampling at the hillslope scale allows for further discovery of hydrological processes masked by an impenetrable soil profile.

Virtual experiments (i.e. Weiler and McDonnell, 2004) are another avenue for new process understanding. Virtual experiments incorporate numeric modeling informed by field experience designed to act as a learning tool, rather than for prediction. Using a carefully constructed and validated numeric model, we can develop numeric experiments to probe system response well beyond what is possible in the field. Heterogeneities in processes and parameters can be controlled, while the dominant system response can be identified. Virtual experiments have shown promise in determining the controls on a range of hillslope scale processes, including nutrient flushing (Weiler and McDonnell, 2006), canopy smoothing of precipitation (Keim et al., 2006), bedrock leakage (Ebel et al., 2007) and the effects of spatially variable hydrologic parameters (Fiori et al., 2007).

We focused our research on going beyond passive storm monitoring to open the black box of hillslope hydrology. We will describe irrigation experiments performed at well studied field sites in South Island, New Zealand, and Western Oregon, USA to determine the flow processes controlling water and solute transport at the hillslope scale. These experiments were used to develop a perceptual model of flow and transport in steep, forested catchments incorporating rapid lateral subsurface flow and leakage to near surface permeable bedrock. The new perceptual model was used to develop a simple numeric model of hillslope hydrology. The model is then used as a learning tool to investigate the causes and controls on the threshold behavior seen in many hillslopes and small catchments.

Description of Chapters

1.1.1. Chapter 2. Hillslope threshold response to storm rainfall: (1) A field based forensic approach

Chapter 2 outlines a series of hillslope scale irrigation experiments from the Maimai instrumented hillslope, NZ, a site of hydrological research for over 30 years (McGlynn et al., 2002). The goals of the experiment were to identify the location and nature of the preferential flow network, identify the role of bedrock topography on flow routing, and quantify the bedrock permeability and its influence on subsurface flow processes. A new perceptual model of hillslope flow processes was developed based on the findings.

1.1.2. Chapter 3. Hillslope threshold response to storm rainfall: (2) A virtual experimentation approach

Chapter 3 outlines the development of a distributed numeric hydrological model (MaiModel) using the dominant processes concept of Grayson and Blöschl (2000), informed by the new perceptual model of hillslope flow processes at the Maimai hillslope developed in Chapter 2. The conceptual, reservoir type model is calibrated against multiple criteria, including both hydrometric and tracer system response from previous combined tracer application and storm monitoring field campaign by Brammer (1996). The calibrated model is then used in a virtual experiment framework to determine the controls on the macroscale precipitation / discharge threshold relationship seen at this and other field sites (McDonnell, 1990; Tani, 1997; Tromp-van Meerveld and McDonnell, 2006). Alternative hypotheses of controls on the threshold behavior are explored with the model.

1.1.3. Chapter 4. Experimental closure of the hillslope water balance within a measurement uncertainty framework

Chapter 4 describes a field scale steady state irrigation experiment performed at the instrumented hillslope in Watershed 10, in the H.J. Andrews Experimental Forest. Using the closure of the hillslope water balance as a platform, here we present a rigorous uncertainty analysis of a hillslope scale irrigation experiment. The importance of uncertainty accounting is shown, as the water balance is closed, revealing large fluxes through storage in the shallow bedrock. Storage discharge dynamics reveal complex, hysteretic behavior.

1.2. References

Bonell, M., 1993. Progress in the understanding of runoff generation dynamics in forests. Journal of Hydrology, 150: 217-275.

Brammer, D., 1996. Hillslope hydrology in a small forested catchment, Maimai, New Zealand. M.S. Thesis, State University of New York College of Environmental Science and Forestry, Syracuse, 153 pp.

Brooks, E.S., Boll, J. and McDaniel, P.A., 2004. A hillslope-scale experiment to measure lateral saturated hydraulic conductivity. Water Resources Research, 40.

Dooge, J.C.I., 1986. Looking for hydrologic laws. Water Resources Research, 22: 46S-58S.

Ebel, B.A. et al., 2007. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 1. sprinkling experiments. Americal Journal of Science, 307(4): 678-708.

Fiori, A., Romanelli, M., Cavalli, D.J. and Russo, D., 2007. Numerical experiments of streamflow generation in steep catchments. Journal of Hydrology, 339: 183-192.

Grayson, R.B. and Blöschl, G., 2000. Summary of pattern comparison and concluding remarks. In: R.B. Grayson and G. Blöschl (Editors), In Spatial Patterns in Catchment Hydrology: Observations and Modeling. Cambridge University Press, Cambridge, UK, pp. 355–367.

Keim, R.F., Tromp-van Meerveld, I. and McDonnell, J.J., 2006. A virtual experiment on the effects of evaporation and intensity smoothing by canopy interception on subsurface stormflow generation. Journal of Hydrology, 327: 352-364.

Kirkby, M.J., 1978. Hillslope hydrology. Landscape systems. Wiley, Chichester, 389 pp.

McDonnell, J.J., 1990. A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11): 2821-2832.

McGlynn, B.L., McDonnell, J.J. and Brammer, D.D., 2002. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. Journal of Hydrology, 257: 1-26.

Sidle, R.C., Noguchi, S., Tsuboyama, Y. and Karin Laursen, 2001. A conceptual model of preferential flow systems in forested hillslopes: evidence of self-organization. Hydrological Processes, 15(10): 1675-1692.

Tani, M., 1997. Runoff generation processes estimated from hydrological observations on a steep forested hillslope with a thin soil layer. Journal of Hydrology, 200: 84-109.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006. Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope. Water Resources Research, 42.

Wagener, T., Wheater, H.S. and Gupta, H.V., 2004. Rainfall-runoff modelling in gauged and ungauged catchments. Imperial College Press, London, UK.

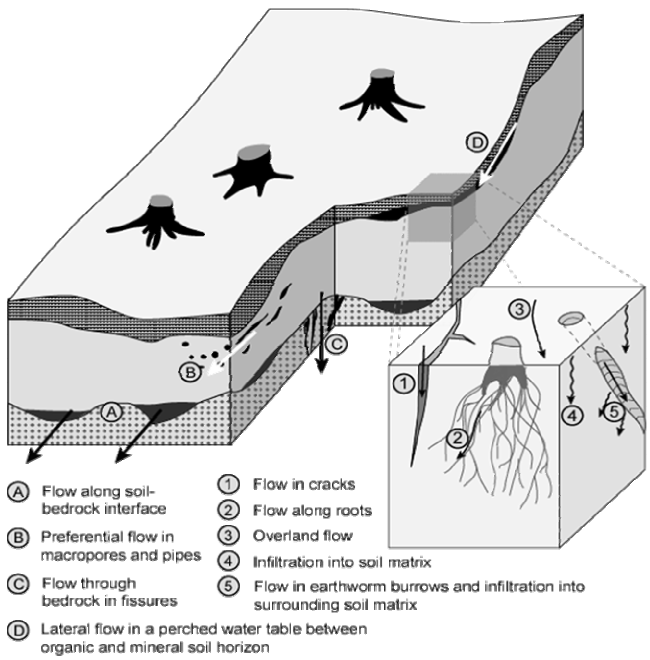
Weiler, M. and McDonnell, J., 2004. Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology. Journal of Hydrology, 285(1-4): 3-18.

Weiler, M. and McDonnell, J.J., 2006. Testing nutrient flushing hypotheses at the hillslope scale: A virtual experiment approach. Journal of Hydrology, 319(1-4): 339-356.

Weiler, M., McDonnell, J.J., Tromp-van Meerveld, I. and Uchida, T., 2006. Subsurface stormflow. Encyclopedia of Hydrological Sciences, 3.

1.3. Figures

Figure 1.1 The complexity of subsurface flow processes (from Sidle et al., 2001)



2. Hillslope threshold response to storm rainfall: (1) A field based forensic approach

2.1. Introduction

Hillslopes are fundamental units of the hydrologic landscape and the main filter for water and solute transport from the atmosphere to the stream. In forested regions of the world, quick lateral subsurface stormflow (often called interflow or throughflow) is the primary mechanism for stormflow generation in headwater catchments (Hursh, 1944). Much of the progress in identifying the different manifestations of subsurface stormflow behaviors was made in the 1960s and 1970s (Hewlett and Hibbert, 1967; Mosley, 1979; Whipkey, 1965). More recent work has tempered these discoveries by revealing the complexity, heterogeneity and uniqueness of natural hillslope drainage systems (McDonnell et al., 2007) and the wide range of scales of processes imposed by climate, geology and vegetation that control hillslope response (Sidle et al., 2007; Sivapalan, 2003; Zehe et al., 2007).

One common denominator in hillslope response to rainfall is the often-observed threshold relationship between total storm precipitation and lateral subsurface stormflow. Several (e.g. Lehmann et al., 2007) have recently argued that this is an emergent property at the hillslope scale—a property that subsumes much of the sub-grid complexity at the plot scale. While threshold relationships between storm rainfall and hillslope-scale runoff have been shown now in several environments (Buttle and McDonald, 2002; Hutchinson and Moore, 2000; Mosley, 1979; Spence and Woo, 2002; Tani, 1997; Uchida et al., 1999; Uchida et al., 2005) the physical cause of these thresholds has been difficult to demonstrate given the challenge of taking measurements at the hillslope scale. Recently, Tromp-van Meerveld and McDonnell (2006a; 2006b), based on field observations at the Panola Mountain Research Watershed in Georgia USA, developed the "fill and spill" theory to explain the precipitation threshold for lateral subsurface stormflow. The fill and spill hypothesis states that that connectivity of patches of subsurface saturation (at the interface between the soil and an impeding layer) is a necessary pre-condition for significant hillslope-scale storm response. These isolated patches of subsurface saturation are located in topographic hollows of the impeding layer, and connection of these patches is controlled by both the topography of the impeding layer and the permeability of the impeding layer. This hypothesis was supported by observed patterns of transient water table development and lateral subsurface stormflow at Panola, and since then by model analysis in two and three dimensions (James and McDonnell, 2008; Keim et al., 2006). However, the physical measurement of the dominant controls on the presence and magnitude of the observed thresholds have not been performed.

The fill and spill hypothesis for hillslope-scale threshold response depends on a number of factors, including the location and nature of lateral subsurface flow, and characteristics of the boundary where lateral flow occurs. In this hypothesis, it is implicitly assumed that flow occurs at the surface of an impeding layer, and that flow is routed by the bedrock topography. It is also assumed that the permeability of the impeding layer is of an intermediate value, permeable enough that storage is transient and requires filling during events, but not so permeable that storage cannot be filled during typical event rainfall intensities and durations. Despite the promise of fill and spill controlled thresholds as a way to define and quantify hillslope scale response to precipitation, these factors (the nature of the flow system and permeability of the impeding layer) have proved difficult to assess in single-realization field studies. The mapping, measuring and quantifying the flow network and bedrock permeability is extremely difficult with current field techniques and approaches.

So how can we explore the mechanistic controls on hillslope threshold response to storm rainfall (the fill and spill hypothesis) and develop a function that captures sub-grid scale variability into numerical macroscale behavior? Here we present a new field-based experiment aimed at defining hillslope-scale subsurface internal processes via limited destructive sampling of a well-researched site. We follow the tradition in soil science, where soil pits and excavations after tracer applications are a commonplace method for determining processes occurring at the soil pedon scale (e.g Flury et al., 1995; Zehe and Fluhler, 2001). In this study we excavated a whole slope section to develop new physical understanding of internal controls on threshold response and whole hillslope emergent behavior. Our work builds upon some destructive experimentation that has already been attempted in hillslope hydrology (e. g. Kitahara, 1993; Moran et al., 1989; Singh et al., 1991; Tippkötter, 1983), though previous experimentation was not based at the hillslope scale. Kitahara (1993) filled a network of macropores with plaster and removed the soil from surrounding the network, identifying the location and morphology of the preferential flow network. Numerous groups (e. g. Gibbens and Lenz, 2001; Heitschmidt et al., 1988) have also removed the soil from a network of tree roots, revealing the form and structure previously hidden by the soil profile. Additional pit scale irrigation and excavation experiments have been instrumental in revealing the structure and predominance of lateral and vertical preferential flow (e. g. Mosley, 1982; Noguchi et al., 2001; Weiler and Naef, 2003) but have been limited to the pedon scale and have not been attempted across a complete hillslope section.

We will show how destructive sampling at the hillslope scale can be especially useful at our well studied sites, where a history of observed field behaviors can be tested, *ex post facto*, using our forensic approach. Our research site is the Maimai Experimental Watershed on the South Island, New Zealand (see McGlynn et al., 2002) for review). Maimai was one of the early sites where the role of lateral subsurface preferential flow paths was observed (Mosley, 1979; Mosley, 1982). More recently, studies at Maimai have chronicled the initiation of subsurface stormflow through soil pipes (McDonnell, 1990), the patterns of subsurface stormflow (Woods and Rowe, 1996) and solute transport (Brammer, 1996) at the slope base, the relative role of hillslope vs. riparian zones in runoff initiation (McGlynn and McDonnell, 2003a) and nutrient and solute transport (McGlynn and McDonnell, 2003b). While the recognition of rainfall thresholds for generating hillslope response at Maimai dates back to the original work of Mosley (1979), the controls on this whole-hillslope response have been difficult to assess, even at this intensively studied site.

At Maimai, the key components of the fill and spill theory have not been resolved. Both the nature of the lateral subsurface flow network, and the permeability of the bedrock are poorly understood. The characteristics of the lateral flow network have been extrapolated from observations made at trench faces and limited, small scale excavations ($<1m^2$) (Weiler and McDonnell, 2007) while the upslope form, connectivity, extent of the lateral flow network remains unknown. While the bedrock permeability has been estimated using a catchments scale water balance (O'Loughlin et al., 1978; Pearce and Rowe, 1979), no direct measurements have been made. We posit that hillslope scale excavations are a powerful field method to reveal the existence and extent of the lateral flow network and a way to expose the bedrock surface for permeability measurements.

This paper details a hillslope scale irrigation - excavation experiment designed to identify the dominant flow pathways and the role of bedrock topography and permeability at the hillslope scale. Our work tests three sets of multiple working hypotheses directed at the first order controls on the fill and spill theory stemming from previous work at Maimai and other steep, forested hillslopes:

1. How can we characterize the lateral subsurface flow?

1a) Lateral subsurface storm flow is concentrated in the soil matrix and the preferential flow network is non existent or unimportant in generating flow at the hillslope scale (supported at the site by Sklash et al., 1986)

1b) A lateral preferential flow network exists, consisting of disconnected soil pipes located in the soil profile (supported at the site by McDonnell, 1990; Weiler and McDonnell, 2007; elsewhere by Noguchi et al., 1997; Noguchi et al., 2001)

1c) A lateral preferential flow network exists, consisting of a connected network located at the soil/bedrock interface (supported at the site by McDonnell, 1997; Mosley, 1979)

2. How does the boundary layer affect flow routing?

2a) The bedrock surface plays an indirect role in flow routing (supported at the site by Woods and Rowe, 1997)

2b) The bedrock surface determines flow routing (supported at the site by Freer et al., 1997; McDonnell, 1997; and elsewhere by Freer et al., 2002)

3. How does the permeability of the lower boundary affect flow processes?

3a) The bedrock is effectively impermeable (supported at the site by McDonnell, 1990; Mosley, 1979; O'Loughlin et al., 1978; Woods and Rowe, 1996)

3b) The bedrock permeability is high enough to have a significant impact on flow processes (supported elsewhere by Onda et al., 2001; Tromp-van Meerveld et al., 2006)

2.2. Site description

The experiments were performed at the Maimai Experimental Watershed, near Reefton, South Island, New Zealand (Figure 2.1). Maimai was established as a hydrological experimental field site in the late 1974 to examine the effects of forest management on water and sediment flux. The site has been continuously monitored since.

Soils are stony silt loam podzolized yellow brown earths (Rowe et al., 1994) overlain with a 15 cm thick high porosity organic humus layer (McDonnell et al., 1991). Hydraulic conductivity of the mineral soils range from 1.4E-6 – 8.33E-5 m/s (5-300 mm/hr), the mean porosity is 45%, and soil profiles average 0.6 m (McDonnell, 1990). The soil has a high density of preferential flow paths, including vertical cracks, live and dead root channels, and macropores in the soil profile and along the soil bedrock interface (Brammer, 1996; Mosley, 1979; Woods and Rowe, 1996). At the soil surface lies a high permeability (hydraulic conductivity >2.78E-4 m/s (1000 mm/hr) (McDonnell et al., 1991)) organic mat, where isolated, and short lived pseudo-overland flow has been observed. Due to the high annual rainfall (2.450 m mean annual rainfall (Woods and Rowe, 1996) and high storm frequency (average time between storms ~3 days), soils remain within 10% of saturation through most of the hydrologic year (Mosley, 1979). Considered poorly permeable the bedrock is Early Pleistocene Conglomerate of the Old Man Gravel formation, a moderately weathered, firmly compacted conglomerate with clasts of sandstone, schist and granite in a clay – sand matrix (Rowe et al., 1994). Deep seepage to the bedrock aquifer is estimated 100 mm/yr (Pearce and Rowe, 1979),

In the Maimai experimental forest, the first order catchments are highly responsive, with a runoff ratio (catchment discharge/rainfall) of 54% annually, of which 65% is quickflow (Pearce et al., 1986), as defined by Hewlett and Hibbert (1967). While also responding rapidly to precipitation, the hillslopes have a much lower runoff ratio, ~ 15%, and sustained baseflow for more than 4 days after an event has not been observed (Woods and Rowe, 1996). The difference between hillslope and catchment runoff ratios has not been explained. Reviews by McGlynn et al. (2002) and Rowe et al. (1994) provide additional details on the Maimai catchments.

Our experiments were performed at the hillslope instrumented by Woods and Rowe (1996). The relatively planar hillslope was chosen for their studies, downstream of the M8 catchment studied by earlier generations of scientists (McDonnell, 1990; Mosley, 1979; Mosley, 1982; Pearce et al., 1986; Sklash et al., 1986). The hillslope is representative of the Maimai slope lengths and gradients, with a maximum slope length of 50 m, and gradients above 35^o. Lateral subsurface flow is collected at the slope base by a 60 m long trench excavated into the conglomerate bedrock surface. Flow from the hillslope is routed to 30 trench sections (2 m long) and then into recording 2 L tipping buckets. Due to soil instability and a deep profile, the trench is split into two groups of 20 and 10 m wide trench sections, with a 20 m gap in between. Woods and Rowe (1996) monitored subsurface flow at the trench for 110 days in 1993. A key finding from their work was the recognition of the large spatial variability of lateral subsurface flow, something subsequently observed at field sites around the world (e. g. Freer et al., 2002; Hutchinson and Moore, 2000; Kim et al., 2004). While Woods and Rowe (1996) attributed the spatial variability of lateral subsurface flow to surface topography, subsequent analysis showed that subsurface topography of the soil-bedrock interface better explained the coarse patterns of flow distribution at the hillslope scale (Freer et al., 1997). Later work by Brammer (1996) monitored flow from the trench for 65 days and traced the flux of an applied line source bromide tracer at the instrumented hillslope 35 m upslope of the trench face and observed very fast subsurface stormflow tracer velocities, with 4% of tracer recovery in the first storm after application, less than 3 days later, and less than 9 hours after the storm began.

Analysis of data records from the Woods and Rowe (1996) and Brammer (1996) storm monitoring demonstrate a clear threshold for lateral subsurface flow at the monitored hillslope at Maimai. If one defines an individual storm as at least 1 mm rain preceded by 24 hours without 1 mm rain, 41 storms are identified in the Woods and Rowe dataset, with between 1 and 83 mm total precipitation (Figure 2.2). Total storm hillslope discharge was defined as the increase in discharge for the duration of the storm, including 24 hours after rainfall ceased. Total storm discharge ranged from 0 – 22.2 mm. For all events with less than the 23 mm total storm precipitation threshold, only one storm had measured discharge greater than 0 mm (0.19 mm discharge for a storm of 16.8 mm precipitation).

We reactivated trench sections 10-13 of the Woods and Rowe (1996) trench. These trench sections are located in a (surface and bedrock) topographic hollow where the majority of flow (>64%) was observed in both the Woods and Rowe (1996) and Brammer (1996) monitoring. Trench sections 10 – 13 drain upslope contributing areas of between 51-473 m^2, and peak flows ranged from 0.17-0.23 L/s (0.23 – 2.01 mm/hr) during storm monitoring. In this area of hillslope, pipe flow at the trench face was observed to dominate lateral subsurface flow by previous researchers. Overland flow has not been observed at the hillslope.

2.3. Methods

We performed 2 sets of irrigation experiments above trench sections 10-13. The first experiment was a injection of water and tracer at the soil bedrock interface 8 meters upslope of trench sections 12-13. The second experiment was line source surface application of water and tracer line 4 m upslope of the trench sections 10-11. The upslope application distance was constrained by the presence of a 25 year old *Radiata Pine* 5 m upslope of the trench face. Water was pumped 20 m from the nearby Powerline Creek to the application site with a small gas pump. Irrigation continued until steady state conditions were reached, as determined by steady discharge measured at the trench, and constant spatial patterns of flow at the trench face. For the deep injection experiment, the water was pumped directly into a trench excavated to the soil bedrock interface. The trench was 0.6 m deep and 1 m wide. For the surface application, water was pumped to a perforated gutter 1.7 m long. Water irrigated the soil surface evenly along the 1.7 m long by 0.1 m wide gutter, and the perforations were spaced 0.0025 m apart so that a constant shallow (<0.0025 m) water level was maintained in the gutter. Due to fluctuations in the water source (related to creek stage and pumping rate), the application rate was not constant during the 4 weeks of experimentation. However, steady application was possible over 2-3 hour application periods through careful monitoring of stream levels. Application rate was measured on site, and varied between 0.02 L/s to 0.30 L/s. Discharge was measured at the trench face using the Woods and Rowe (1996) guttering and tipping buckets, linked to a CR10 Campbell Scientific datalogger. As the excavations continued, much of the trench section was damaged, so subsequent trench discharge rates were not recorded. All hillslope discharge was routed to a common 5 L collection vessel where tracer concentration was measured.

2.3.1. Excavation and flow mapping

After steady state was reached in each of the experiments, the types and locations of dominant flow pathways were recorded at the trench face. To ease photographic recording of flowpaths, brilliant blue dye (C.I. Food Blue # 2; C.I. 42090; $C_{37}H_{34}N_2Na_2O_9S_3$) was added to the irrigation water, (e.g. Flury and Flühler, 1995). At steady state, the dominant flowpaths were labeled with orange tape, and vertical and lateral coordinates were recorded. Both matrix flow (as evidenced by wetness at the seepage face) and macropore or other preferential flowpaths were identified. A digital photograph was taken of each exposed trench slice, and of each noted flowpath for later analysis. After the flowpath types and locations were identified and recorded, 0.2 - 0.4 m of the soil profile was removed upslope from the trench face (Figure 2.3). As the soil was removed, the major flowpaths were traced upslope towards the application location to develop a near continuous map of lateral flow throughout the hillslope length. The bedrock surface was fully exposed after each slice removal and the new flow locations and flow features along the soil bedrock boundary were identified. For the direct soil bedrock interface injection experiment, 8 m soil was removed upslope in 37 slices. For the surface application, 4 m soil was removed in 18 slices.

2.3.2. Tracer injections

We measured tracer velocities between excavations by adding a Br^- solution to irrigation water. Tracer was added at every second or third steady state water application following soil removal (9 times during the direct soil bedrock interface application experiment (when 0, 1.30, 1.90 2.45, 2.84, 3.13, 4.19, 6.62 and 7.60 m soil removed) and 7 times during the surface application experiment (when 0.50, 0.99, 1.25, 1.53, 1.92, 2.15, 3.20 m soil removed). During the surface application experiment 4 additional tracer injections were added at different irrigation rates with 1.25 m soil removed. An ion selective electrode for Br^- (TempHion©, Instrumentation Northwest, Inc.) was placed in a 5 liter tank at the trench and readings were taken every minute. The Br^- solution was injected directly into the trench during experiment 1 and uniformly along the length of the gutter during experiment 2. The water application rate was held constant during the injection, and continued until Br^- concentration returned to within 200% of the background concentration, or as long as conditions would allow. Flow rates during the Br^- injection ranged from 0.03 to 0.11 L/s. Due to low flow conditions, irrigation water was recycled in some experiments, causing Br^- concentration to remain higher than background. In these cases, water application and tracer monitoring continued until steady concentration at the trench face was reached. While a mass recovery was not possible, due to deterioration of the trench face, a representative sample of the discharge was collected for all injections.

2.3.3. Bedrock permeability

The bedrock hydraulic conductivity was measured using a falling head test. A cylindrical pit was excavated into the Old Man Gravel bedrock. The pit was 0.25 m deep with radius 0.17 m, with a cross sectional surface area of 0.0934 m^2 and total surface area including the pit walls and bottom of 0.2777 m^2. The bedrock was relatively soft and no fracturing was observed as the pit was excavated. A 0.001 m resolution recording capacitance water level recorder (TruTrack, Inc., model WT-HR) was placed in the pit to record water height changes over time. Prior to the experiment, the pit was prewetted by maintaining a constant head of water for 5 hours. The pit was then filled to a depth of 0.17 m and drainage was monitored for 13 hours. Initial and final water levels were measured with a ruler to confirm capacitance rod function. The recession of the water table was fit to a quadratic. The hydraulic conductivity was calculated using Darcy's law assuming a unit head gradient at long time, and infiltration along either the pit bottom or the pit bottom and sides.

2.4. Results

2.4.1. Flow routing and locations

During both irrigation experiments, lateral subsurface flow at the hillslope trench was dominated by concentrated flow at the soil bedrock interface, including both sheet flow (thin (< 0.002 m), low volume diffuse flow spread over 0.05 – 0.20 m width) and concentrated interfacial flow (high volume flow in visible gaps at the base of the soil profile). During the first water application of Experiment 2, flow at the trench was restricted to within 0.05 m of the soil bedrock interface. At the trench face, flow was concentrated in 5 gaps connected by sheet flow along the bedrock surface. An estimated 70% of total lateral subsurface flow was in the concentrated flowpaths, with the remainder in sheet flow. The concentrated areas were generally voids between the bedrock surface and lower soil boundary, rather than decayed root channels or worm tunnels. These gaps were less than 0.005 – 0.010 m high and ranged from 0.01 – 0.10 m wide and often filled with roots (see Figure 2.4 for an example exposed trench face 0.5 m upslope of the trench).

After flow locations were recorded, 0.2 m soil was removed from the trench face, with the areas of concentrated flow traced upslope as the soil was excavated. This process was then repeated as the hillslope was excavated. As excavation progressed upslope in 0.2 – 0.4 m increments, the flowpaths remained continuous and connected, with some divergence and convergence, controlled by bedrock features such as small scale (< 0.1 m tall) valleys and ridges in the bedrock surface. A coat of brown organic staining was observed on the exposed bedrock surface along with a nearly ubiquitous matt of very fine to medium live roots along the bedrock surface (Figure 2.5). In some isolated locations water diverged from the bedrock surface and flowed through and above a thin (< 0.1 m) gleyed clay layer. These gleyed areas of soil appeared to be in topographic depressions in the bedrock surface, and suggest chronically saturated conditions.

Vertical preferential flowpaths were observed in the exposed vertical soil column in the immediate proximity of where the water and dye was applied. Such features were not active in the slices greater than 0.75 m downslope from the surface application. With the exception of limited matrix flow and some isolated macropores, the majority of water traveling from the soil surface to the bedrock was via thin, sub-vertical cracks in the soil, similar to those reported by previous researchers (McDonnell, 1990). These cracks were coated with a brown organic stain, similar to that seen on the bedrock surface. The vertical and subvertical cracks were planes of weakness in the soil structure, and slaked off while excavating.

For the direct soil bedrock interface injection experiment, flow was observed at the soil bedrock interface at all excavated slices, as well as during excavations between slices. Once excavations reached within 0.3 m of the direct soil bedrock interface application, some flow through the soil column was observed in the lower 0.25 m of soil. At this point both active macropore flow in the soil profile and saturated matrix flow were observed. The macropores were less than 10 cm long, and appeared disconnected from any larger preferential flow system.

Trench response more than 0.75 m downslope the surface application and 30 cm downslope of the direct soil bedrock interface application was identical for the two application regimes in terms of flowpath location (at the soil bedrock interface) morphology (areas of concentration controlled by bedrock depressions and obstructions connected by sheet flow) and flow response (rapid and sensitive to changes in application rate). Fluctuations in application rate, which varied from 0.07 - 0.25 L/s, did not have an impact on the locations of concentrated flow, though the relative magnitude of each flow path was sensitive to input rate.

Field observation and visual analysis of photographs of each trench section showed areas of organic staining in the lower profile (See Figure 2.4). This staining suggested areas of prolonged saturated conditions, and concentrated above the flow paths identified during the irrigation experiments. Stained areas were generally semicircular, with a diameter of up to 0.1 m, and located with the base on the bedrock surface. Additional staining was observed along the entire bedrock surface, while little was seen in the soil profile greater than 0.1 m above the bedrock surface.

2.4.2. Tracer breakthrough and velocity

Tracer breakthrough was similar for both the surface (and direct soil bedrock interface injections), with initial tracer breakthrough averaging 420 s (1020 s) after application (Figure 2.6, Table 2.1 and Table 2.2). Peak concentrations were reached in 1080 s (2700 s). The time to initial and peak concentration breakthrough were longer for the direct injection than the surface injection, as expected due to the longer travel distance. Breakthrough curves were skewed to the right, with a rapid peak and long tail. A skewed breakthrough curve indicates transport with a range of travel velocities, consistent with the combination of sheet flow and concentrated flow observed during excavation. Due to irrigation source water limitations, time constraints and pumping difficulties, the entire tail was not captured for the tracer experiments. Deterioration of the trench flow collecting system precluded an accurate mass balance for the tracer injections.

Despite these difficulties, the time to initial rise and time to peak tracer concentrations were well captured, giving an estimate of initial and peak travel velocities. Both initial breakthrough and peak concentration velocities were high for all bromide injections, at both the direct and surface experiments. For the surface applications, initial breakthrough velocities ranged from 6.7E-3 to 3.3E-2 m/s (Table 2.2). For Br- injections with greater than 1 m soil remaining downslope of the irrigation source, initial breakthrough velocity was uncorrelated with amount soil removed (R^2 = 0.12). Peak concentration velocities ranged from 2.1E-3 to 1.3E-2 m/s, with no correlation between tracer velocity and soil removal (R^2 = 0.02). For the direct soil bedrock interface application, initial and peak velocities ranged from 5.3E-3 to 6.7E-2 m/s and 1.9E-3 to 3.3E-2 m/s, respectively. For the direct application, initial breakthrough and peak concentration velocities were weakly correlated with soil removal (R^2 = 0.54 and 0.50 respectively), where velocity increased as the soil mass was removed (Table 2.1).

For the initial applications, while the trench system was still intact, we calculated the volume of water discharged from the trench before the peak concentration was reached, based on measurement of input rates and trench runoff. This represented the volume of water in the active flow paths, or the active pore volume. The active pore volumes in these experiments ranged from 0.04 m^3 - 0.18 m^3 (0.03 - 0.31 m^3). The active pore volumes averaged 4% (7%) of the total estimated pore volume, based on an average soil depth of 0.6 m and porosity of 0.45 reported at the site (McGlynn et al., 2002).

Tracer times to peak velocities were high, ranging from 1.9E-3 – 6.7E-2 m/s for the two sets of injections. Prior to excavation the tracer velocity was 3.3E-3 m/s for the surface application and 2.5E-3 m/s for the direct soil bedrock interface injection (Table 2.1). Our reported peak concentration velocities likely overestimate mean travel velocities, due to an observed left ward skew of the breakthrough (Figure 2.6).

Assuming plug, Darcy regime flow, a measured soil saturated hydraulic conductivity (K_{sat}) of 1.4E-6 – 8.33E-5 m/s (McDonnell, 1990), and the measured average hillslope gradient (s) of 56%, predicted Darcy ($v = K_{sat}s$) velocities were between

1.7 E-6 – 4.7 E-5 m/s more than 3 orders of magnitude less than that measured in our experiments.

2.4.3. Bedrock permeability

The drainage rate of pooled water in the bedrock permeability experiment decreased during the first 9 hours, with an initial rate of 1.8E-6 m/s (6.4 mm/hr), slowing to a steady rate of 8.6E-7 m/s (3.1 mm/hr) for the final 4 hours of the experiment. Assuming a unit head gradient at long time, the bedrock hydraulic conductivity was calculated from Darcy's Law ($Q = K_s A(\Delta h/L)$), where Q is pit drainage at late time, K_s is the hydraulic conductivity, A is the area over which drainage occurs, $\Delta h/L$ is the head gradient, assumed to near 1 at long time. Two estimates of A were made: (1) if drainage occurred only at the base of the pit, A = 0.0935 m^2 and (2) if drainage occurs over the entire surface area of the pit, A = 0.935+0.1842 m^2. Using assumption (1), the bedrock hydraulic conductivity was 8.6E-7 m/s (0.31 cm/hr). Using assumption (2) the hydraulic conductivity was 2.9E-7 m/s (0.10 cm/hr). The recession of the water table was also well fit (R^2 = 0.995) by a function of the form $z_t = z_{t-1}(1-k\Delta t)$, where k is a dimensionless recession coefficient. Using a least squares optimization, k = 0.086 s^{-1}.

2.5. Discussion

Our experiments represent the first hillslope scale destructive sample sampling at Maimai or any other previously-instrumented hillslope that we are aware of. This targeted destructive sampling was designed to isolate and illuminate the preferential flow network long hypothesized to dominate lateral subsurface flow at the site. This excavation allowed for the additional measurement of the permeability of the bedrock, a crucial control on the initiation of lateral subsurface flow and the partitioning of the water balance. The work was specifically designed to test three sets of competing alternative hypotheses related to the nature of the lateral subsurface flow and it relationship with the topography and permeability of the bedrock. The first set of hypotheses addresses the form and function of the preferential flow network; the second set of hypotheses addresses the role of the bedrock topography on the flow network; the third set of hypotheses addresses the permeability of the bedrock. The three sets of hypotheses are investigated in depth below. We then discuss the influence of our findings on the threshold relation between storm total precipitation and lateral hillslope discharge. A new perceptual model of flow at the site is developed, and its implications regarding model structure are discussed in detail Graham and McDonnell (this issue).

2.5.1. Preferential flow network hypotheses

We rejected hypotheses 1a and 1b, and accepted hypothesis 1c – that lateral subsurface flow is dominated by a connected preferential flow network located at the soil bedrock interface. Applied flow rates were consistent with lateral flow observed during medium to large stormflow, and the preferential flow network was able to accommodate the flow volumes. Lateral subsurface flow was observed solely at the soil bedrock interface, where water was transmitted as sheet flow and preferentially in voids restricted to within 5 cm above the bedrock surface, occupying only 4% of the available pore space. The active flow zone coincided with live and dead roots at the soil bedrock interface and organic staining on the bedrock surface and in the lower soil profile, indicating these flow paths are stationary and chronically saturated during natural events. Both the root density and organic staining were much reduced in the soil profile above the observed flow zone. There was no evidence for lateral macropore flow in the soil profile as hypothesized by Weiler and McDonnell (2007) and observed elsewhere (Tsuboyama et al., 1994), though vertical and subvertical cracks appeared responsible for routing water from the soil surface to depth, as observed at this site (McDonnell, 1990). While some macropores were seen in the soil profile, these were apparently disconnected to the flow occurring at depth, and not observed to be routing water except near the irrigation application source.

While our irrigation rates were high when expressed as a precipitation rate (592 – 2117 mm/hr), our intent was to isolate the lateral subsurface flow component, rather than identify flow paths from the soil surface to depth. Measured lateral subsurface flow rates for natural storms at the gauged trench face for trench sections 12-13 (below the direct soil bedrock interface application) range from 0 – 0.40 L/s (Woods and Rowe, 1996), which bound our applied rates and measured discharge. Downslope of the surface application, measured throughflow for trench sections 10-11 during natural events were similar (0 – 0.38 L/s). Considering the relatively small contributing area between the application site and the collection trench, most of the water collected at the trench would pass the application site as lateral subsurface flow during natural events. The effect of a trench face on unsaturated flow paths has long been known (Atkinson, 1978), primarily diverging flow vectors from the trench face due to capillarity and other edge effects. Since our system was dominated by saturated flow, edge effects were not anticipated to be a large factor. In fact, no evidence of edge effects due to the trench face was seen while excavating upslope during the irrigation and no evidence of unsaturated matrix flow (staining of the dyed irrigation water in the soil profile) was seen upslope of the original trench.

Mosley (1979) identified bypass flow to the bedrock surface and downslope routing along the bedrock as one of the major lateral subsurface flow paths from irrigation experiments in the M8 catchment. During small scale experiments (application < 1 m upslope from his 1 m^2 pits), Mosley found very fast flow velocities (average 6 m/hr) along these and other flowpaths. These findings were seemingly contradicted by the age of the water (~4 months) and low percentage (<25%) of event water in pit discharge, as identified by analysis of naturally occurring oxygen and hydrogen isotopes in the rainfall (Pearce et al., 1986; Sklash et al., 1986). One possible source of mixing of event and stored, pre-event water is in the soil profile, as rainfall mixes in the large soil moisture reservoir before leaking onto the bedrock surface and rapidly routing downslope. This is consistent with the observed lack of downslope aging of water at M8 at the hillslope scale (Stewart and McDonnell, 1991).

This network differs from previous conceptual models in that it is connected, extensive, and it is located exclusively at the soil bedrock interface. Tani (1997) proposed a similar network after stormflow monitoring at Minamitani catchments, Japan, though bedrock interfacial flow was perceived to begin there after soil profile saturation. In an irrigation / excavation experiment at Hitachi Ohta, Japan, where irrigation was applied evenly on the surface 1 m upslope of a trench, Tsuboyama et al. (1994) showed that flow was dominated by matrix flow and laterally oriented pipes connected by organic rich areas of mesoporosity, while flow along the bedrock interface played a relatively minor role. The findings of the current experiment suggests that a very different flow network may have been observed at Hitachi Ohto had the irrigation been applied further upslope, allowing the irrigated water the time to reach the bedrock surface. At Panola

Georgia, storm monitoring by Tromp-van Meerveld and McDonnell (2006a) showed that flow from macropores located at the soil bedrock interface makes up 42% annually of trenchflow at a site where leakage to the bedrock dominates the water balance. The lack of upslope excavations has prevented the determination of the upslope nature of the flowpath network at Panola, though this research suggests that a connected preferential flow network at the soil bedrock interface is possible.

2.5.2. Bedrock surface flow routing hypotheses

We rejected hypothesis 2a and accepted hypothesis 2b – that the bedrock surface controls lateral subsurface stormflow routing. The bedrock micro and macrotopography were shown to be the major control of water routing at the hillslope scale. While occasionally the flow paths in the soil were observed above the soil bedrock interface on top of thin clay lenses, the majority of flow was in direct contact with the bedrock surface. During the excavations interfacial flow paths were observed to be routed primarily by features such as protruding cobbles and rills on the bedrock surface. Due to the steep slopes and generally planar bedrock, much of the flow routing was controlled by microtopographic features that were small, less than 0.1 m in relief. In one case, flow was observed to be routed from one collecting trench section to another by one such small rill on the bedrock surface 1 m upslope of the trench. This feature had a maximum relief of 0.05 m and routed 33% of the water from one trench section to the next, locally redirecting water fed by 50 m of upslope contributing area to the site. Since water reached the bedrock surface within one meter of application for both the direct soil bedrock interface and surface application, bedrock routing would be expected to dominate flow paths for the majority of water upslope of the collecting trench.

Freer et al. (1997) used a two meter DEM of the bedrock topography to determine hillslope scale flow routing at the Maimai hillslope and observed that it was a better predictor of the spatial pattern of hillslope trench flow than surface topography. Woods and Rowe (1997), however, showed that the difference was slight, and could be explained by uncertainty in the surface topography, where small errors in the DEM could result in large differences in the upslope contributing area at each two meter trench section. The findings from the present study suggest that small topographic features can have a disproportionately large impact on flow routing. Furthermore, the two meter DEM of bedrock topography used by Freer et al. (1997) was not likely of high enough resolution to reliably predict flow at the two meter trench section scale. From the present study, it seems that a very high resolution DEM (< 0.1 m grid spacing, with resolution greater than 0.05 m) of the bedrock surface is needed to predict flow as measured by two meter trench sections located at the hillslope base. While both the surface and subsurface two meter DEMs predicted the general pattern of flow (concentrated in the topographic hollow), neither is of sufficient precision to predict flow into each trench section.

2.5.3. Bedrock permeability hypotheses

We rejected hypothesis 3a and accepted hypothesis 3b – that the bedrock permeability is significant. The measured bedrock hydraulic conductivity falls in the semi-pervious range (Bear, 1972) and potentially a large component of the water balance. The bedrock at Maimai has been described as “poorly permeable” (O'Loughlin et al., 1978), “effectively impermeable” (McDonnell, 1990), and as “nearly impermeable” (McGlynn et al., 2002). However, to our knowledge, no direct measurements of bedrock permeability have ever been attempted at Maimai. Our falling head permeability measurement showed that bedrock Ksat was semi-pervious (2.9E-7 – 8.6E-7 m/s). While this was one point measurement of limited scale, the relatively high value suggests that losses to bedrock cannot be ignored.

There is evidence in the historic data record of significant losses to bedrock at the Maimai hillslope. At the nearby M8 catchment (a 3.8 ha zero order catchment whose outlet is 100 m upstream of the study hillslope in this paper) annual runoff ratios measured at a perennial stream average 54% (1404 mm) (McGlynn et al., 2002). Barring any lateral redistribution from nearby catchments and assuming no losses to deep groundwater, this suggests a maximum annual evaporation rate of 46% of precipitation (1196 mm). The hillslope scale runoff ratios have been reported to be 13% over 110 days of monitoring (Woods and Rowe, 1996), and 14% for a separate period of 90 days monitoring (Brammer, 1996). While these two periods of monitoring were not particularly long, the uniform temporal distribution of precipitation at the site suggests that seasonal changes in storage at the hillslope would be minor. Since both sites are experiencing similar evaporative conditions, the difference in runoff ratios is likely due to leakage to the bedrock, which is a likely sink at the hillslope scale, and a source at stream channel. This would suggest a minimum loss to bedrock at the hillslope scale of 41% of rainfall (1066 mm/year). Previous estimates at the M8 catchment place bedrock leakage at 100 mm/yr (O'Loughlin et al., 1978), or 3.9% of precipitation. The hillslope scale estimates are an order of magnitude larger than previous estimates of leakage at the catchment scale. We hypothesize that the majority of this “lost” water would reemerge at the stream channel, based on the higher annual runoff ratios measured at the stream channel. While not having an impact on the catchment water balance, water traveling through the bedrock and reemerging at the stream would have longer flow paths, more contact with the subsurface and longer mean residence times.

This finding of the relatively high permeability at a site where the bedrock has been previously considered “effectively impermeable” (McDonnell, 1990) suggests that a similar reassessment is warranted at other hillslopes. In fact, leakage to bedrock has been shown to be a significant subsurface flowpath at the hillslope scale at a number of research catchments that were considered impermeable prior to investigation, with a wide range of underlying bedrock types (e.g. Hornberger et al., 1991; Katsuyama et al., 2005; Montgomery et al., 1997; Onda et al., 2001; Tromp-van Meerveld et al., 2006). Low runoff ratios observed at the monitored hillslope in Panola, Georgia (underlain by Panola Granite) were initially attributed to transpiration losses, before direct measurement of the bedrock permeability estimated it at 1.4E-7 m/s (Tromp-van Meerveld et al., 2006), approximately double that measured at Maimai in this study. Waichler et al. (2005) showed through numeric modeling that bedrock leakage accounts for 15% of the water balance at three second and third order catchments in the HJ Andrews Experimental Forest in Oregon (underlain by Andesite), a volume that was previously assumed to be lost to evaporation. Subsurface flow through the bedrock has been shown to be significant in other steep, forested catchments (e.g. Onda et al., 2001 (Serpentinite Rocks); Montgomery et al., 1997 (Eocene Sandstone); Katsuyama et al., 2005 (weathered granite)), affecting mean residence times, tracer transport, and flow routing.

2.5.4. A new perceptual model of subsurface flow at Maimai

Based on the experimental results described above, the perceptual model of subsurface flow processes proposed by previous research at Maimai is modified to fit new findings of subsurface flow processes. While previous perceptual models have been proposed at this site, they have been constrained by isolated observations of limited water balance components (e. g. Mosley, 1979; Pearce et al., 1986), the spatial limitation of previous sprinkler experiments (e.g. McDonnell, 1990; Mosley, 1979), and the limiting nature of measurement techniques (Brammer, 1996; Woods and Rowe, 1996).

The new perceptual model of flow processes at the Maimai hillslopes is shown in Figure 2.7. While pseudo overland flow through the high permeability organic mat on the soil surface has been observed in topographic hollows and riparian areas (McDonnell et al., 1991), overland flow has not been observed at the hillslopes of Maimai even at the extremely high rainfall intensity of experiment 2. Water infiltrates into the soil matrix during small events where rainfall intensity is less than the hydraulic conductivity of the upper soil profile. When rainfall rates are higher than infiltration rates, or when saturation of the soil column occurs, bypass flow through the observed sub-vertical cracks delivers water from the soil surface to the bedrock (McDonnell, 1990). Water in the soil column mixes with pre-event water, and drains vertically to the bedrock, consistent with observation of vertical pressure heads at a nearby hillslope during and between storms (McDonnell, 1990).

Once at the soil - bedrock interface, water flows along the bedrock surface, as evidenced by both flow routing during experiments 1 and 2, and the observed root matting and organic staining along the soil bedrock interface during excavations. The disconnected macropore flow network as proposed by Weiler and McDonnell (2007) is not supported by the current experiments, where macropore flow in the soil profile was not observed. Bedrock flow routing is not inconsistent with the observed spatial variability of lateral subsurface flow at the hillslope trench (Brammer, 1996; Freer et al., 1997; Woods and Rowe, 1996), though storm monitoring alone could not conclusively prove such control.

Once at the soil bedrock interface, water moves either quickly downslope via a connected flowpath network of voids in the lower 0.5 m of the soil profile, or leaks into the bedrock, likely reemerging at the stream channel. High water velocities are consistent with those seen in both storm monitoring and irrigation experiments (Mosley, 1979; Mosley, 1982) and a hillslope scale tracer experiment (Brammer, 1996). Leakage into bedrock at the hillslope along with reemergence at the stream channel is consistent with the low runoff ratios (0.15) observed at the hillslope (Brammer, 1996; Woods and Rowe, 1996) combined with the high runoff ratios (0.60) seen at the first order catchment upstream (Pearce et al., 1986).

2.5.5. Implications on threshold for lateral subsurface flow initiation

The nature of the preferential flow network, its location and the permeability of the bedrock all have significant influence on the threshold for initiation of lateral subsurface flow. To demonstrate the influence of these factors on the threshold, we will compare the findings at Maimai with another well studied field site, the Panola hillslope in Georgia, USA. Panola is similarly instrumented, with a 20 m trench collecting lateral subsurface flow from a 960 m^2 hillslope. At Panola, the threshold for lateral subsurface flow initiation has been attributed to the filling of subsurface storage in the small bedrock surface depressions, which occurs despite leakage into the permeable bedrock. Upslope connection of filled subsurface storage has been observed after 54 mm rainfall, coinciding with the threshold for significant lateral subsurface flow (Tromp-van Meerveld and McDonnell, 2006a; 2006b). At the instrumented hillslope at Maimai, the threshold for flow is approximately 23 mm (Graham and McDonnell, this issue).

The threshold for initiation of lateral subsurface flow is directly dependent on the nature of the lateral subsurface flow network. Assuming no preferential flow network (hypothesis 1a), lateral subsurface flow would initiate in the soil matrix as soon as the head gradients began to develop downslope. While this would begin soon after rainfall, with the low hydraulic conductivity of the soil matrix, substantial amounts of lateral subsurface flow would not occur until saturated conditions had spread through most of the soil profile, due to the low hydraulic conductivity of the soil matrix. Assuming preferential flow was dominated by disconnected macropore flow in the soil profile (hypothesis 1b), lateral flow would not be initiated until the water table had risen above the inlet of the macropores. This flow network would need a much greater amount of precipitation to turn on than the situation where there was no network at all, as the water table would need to raise a considerable height to intersect a substantial number of macropores. Assuming the preferential flow network is a connected network at the soil bedrock interface (hypothesis 1c, supported at Maimai by these experiments), lateral subsurface flow in the preferential flow network would initiate as soon as saturated conditions were met at the base of the soil profile, and water began to drip into the observed voids along the soil bedrock interface. Of the three available hypotheses, the connected network at the soil bedrock interface leads to the smallest threshold for significant lateral subsurface flow. While the preferential flow network at Panola is still poorly understood, the significant portion of flow that emerges at the trench as macropore flow at the soil – bedrock interface suggests that a similar flow network is occurring at Panola, and the threshold should be similar at the two sites. Since the threshold is greater at Panola, another explanation is needed.

The threshold for initiation of lateral subsurface flow is directly dependent on whether the bedrock topography controls flow routing. Assuming the bedrock surface plays an indirect role in flow routing (hypothesis 2a), the filling of bedrock topographic storage should be incidental in lateral flow generation. However, assuming the bedrock surface is the direct control of flow routing (hypothesis 2b, supported at Maimai by these experiments), topographic hollows would need to be filled before lateral subsurface flow would initiate. Whereas Panola had a relatively shallow slope (14%), Maimai is very steep (38%), and bedrock topographic storage is likely much less at Maimai, assuming bedrock surface roughness are equal. In fact, no topographic pools larger than 0.1 m deep and of diameter greater than 0.05 m were observed at Maimai in the exposed bedrock surface after excavation. The threshold for initiation of lateral subsurface flow should be greater for Panola due to the shallower slope and greater potential storage at the bedrock surface.

The threshold for initiation of lateral subsurface flow is directly dependent on the permeability of the bedrock in a system where the lateral preferential flow network is at the soil bedrock interface. Assuming a (nearly) impermeable bedrock (hypothesis 3a), bedrock topographic storage would be filled quickly, and remain filled between events. This would lead to a much lower (if any) threshold at the site. Assuming the bedrock is permeable (hypothesis 3b, supported at Maimai by these experiments), flow along the bedrock surface will drain into the bedrock while moving downslope. At the extreme case, where the bedrock permeability is equal to the permeability, no lateral subsurface flow would occur at all, as flow paths would not be diverted downslope. This case was seen at Mettmann Ridge, Oregon, where an irrigation experiment at a similarly steep forested catchment resulted in little lateral subsurface flow above the bedrock surface due to the very high permeability of the underlying fractured bedrock sandstone (Montgomery et al., 1997). In the case of Maimai, where the bedrock hydraulic conductivity (2.8E-7 – 8.3E-7 m/s) was below the lower end of the range of the soil hydraulic conductivity (1.4E-7 – 8.3E-6 m/s), lower leakage to bedrock leads to more lateral subsurface. However, the bedrock hydraulic conductivity measured at Maimai is less than that of Panola (1.6E-7 m/s; Tromp-van Meerveld et al., 2006), another possible explanation of the higher threshold (54 mm) seen at Panola.

While our work has revealed implications regarding the relative value of the threshold when compared to Panola, it is still poorly understood how each factor directly impacts the threshold at each site. Additionally, the three factors mentioned above do not encompass all possible sources of the threshold, which also include geometry of the watershed, including the percent riparian area, slope and slope length, soil textural properties such as drainable porosity and hydraulic conductivity, or environmental factors such as storm frequency and potential evaporation rates. While analysis of long term data records can help tease out environmental effects (such as comparing the thresholds for flow for storms with different antecedent moisture conditions), determining the precise effect of geometry and bedrock and soil properties will require either extensive site intercomparison or physical and numeric modeling. Due to the wide range of factors that can potentially impact the threshold, it seems that virtual experiments are the way forward.

2.6. Conclusions

Field scale experimentation and destructive sampling demonstrated the form and function of the subsurface flow network at a well studied catchment. A hillslope excavation revealed a connected, extensive preferential lateral flow network at the soil bedrock interface capable of transmitting large volumes of water. The flow network was shown to be controlled by small scale features on the bedrock surface. Bromide tracer applications demonstrated high lateral velocities, reaching 2.5E-3 m/s. A falling head test determined the bedrock was permeable, with a saturated hydraulic conductivity of

2.8E-7 – 8.3E-7 m/s. These observations were combined with previous field observations to create a new perceptual model of flow processes at the site.

Our findings suggest that the major controls on subsurface flow paths are not the standard measured parameters, such as surface topography and soil depth, permeability and texture, but rather other, more difficult to measure parameters, such as the microscale bedrock topography, bedrock permeability, and the lateral subsurface velocities (hillslope scale anisotropy). These parameters are more difficult to measure because of their scale of operation and location, often buried beneath the soil profile. Numeric models using this critical information, and perhaps simplifying less dominant processes such as transport dynamics through the soil profile, may be the key to developing new parsimonious models whose structures capture the dominant processes at a site.

2.7. Acknowledgements

This work was funded through an NSF Ecosystem Informatics IGERT fellowship. I would like to thank Katherine Hoffman for providing logistical support. Tim Davies, John Payne and Jagath Ekanayake at Landcare provided logistical support while in New Zealand. David Rupp, Garrett Kopmann and Matt Floerl provided help with the field work. Sarah Graham kept me sane in the isolated confines of the Maimai hut. We would especially like to thank Jagath Ekanayake for housing and entertaining us while in New Zealand.

2.8. References

Atkinson, T.C., 1978. Techniques for measuring subsurface flow on hillslopes. In: M. Kirkby (Editor), Hillslope Hydrology. John Wiley & Sons, Chichester, UK, pp. 73-120.

Bear, J., 1972. Dynamics of fluids in porous media. American Elsevier Pub. Co., New York, 764 pp.

Brammer, D., 1996. Hillslope hydrology in a small forested catchment, Maimai, New Zealand. M.S. Thesis, State University of New York College of Environmental Science and Forestry, Syracuse, 153 pp.

Buttle, J.M. and McDonald, D.J., 2002. Coupled vertical and lateral preferential flow on a forested slope. Water Resources Research, 38(5).

Flury, M. and Flühler, H., 1995. Tracer characteristics of Brilliant Blue FCF. Soil Science Society of America Journal, 59: 22-27.

Flury, M., Leuenberger, J., Studer, B. and Flühler, H., 1995. Transport of anions and herbicides in a loamy and a sandy field soil. Water Resources Research, 30: 823-835.

Freer, J. et al., 1997. Topographic controls on subsurface storm flow at the hillslope scale for two hydrologically distinct small catchments. Hydrological Processes, 11(9): 1347-1352.

Freer, J. et al., 2002. The role of bedrock topography on subsurface storm flow. Water Resources Research, 38(12).

Gibbens, R.P. and Lenz, J.M., 2001. Root systems of some Chihuahuan Desert plants. Journal of Arid Environments, 49(2): 221-263.

Graham, C.B. and McDonnell, J.J., this issue. Hillslope threshold response to rainfall: (2) Development and use of a macroscale behavioral model.

Heitschmidt, R.K., Ansley, R.J., Dowhower, S.L., Jacoby, P.W. and Price, D.L., 1988. Some observations from the excavation of honey mesquite root systems. Journal of Range Mananagement, 41: 227-231.

Hewlett, J.D. and Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: W.E. Sopper and H.W. Lull (Editors), Forest Hydrology. Pergamon Press, New York, pp. 275-291.

Hornberger, G.M., Germann, P.F. and Beven, K.J., 1991. Throughflow and solute transport in an isolated sloping soil block in a forested catchment. Journal of Hydrology, 124: 81-99.

Hursh, C.R., 1944. Subsurface flow. Transactions of the American Geophysical Union, 25: 743-746.

Hutchinson, D.G. and Moore, R.D., 2000. Throughflow variability on a forested hillslope underlain by compacted glacial till. Hydrological Processes, 14(10): 1751-1766.

James, A. and McDonnell, J.J., 2008. Panola Modeling.

Katsuyama, M., Ohte, N. and Kabeya, N., 2005. Effects of bedrock permeability on hillslope and riparian groundwater dynamics in a weathered granite catchment. Water Resources Research, 41.

Keim, R.F., Tromp-van Meerveld, I. and McDonnell, J.J., 2006. A virtual experiment on the effects of evaporation and intensity smoothing by canopy interception on subsurface stormflow generation. Journal of Hydrology, 327: 352-364.

Kim, H.J., Sidle, R.C., Moore, R.D. and Hudson, R., 2004. Throughflow variability during snowmelt in a forested mountain catchment, coastal British Columbia, Canada. Hydrological Processes, 18(7): 1219-1236.

Kitahara, H., 1993. Characteristics of pipe flow in forested slopes. IAHS Publication, 212: 235-242.

Lehmann, P., Hinz, C., McGrath, G., Meerveld, H.J.T.-v. and McDonnell, J.J., 2007. Rainfall threshold for hillslope outflow: an emergent property of flow pathway connectivity. Hydrology and Earth System Sciences, 11(2).

McDonnell, J.J., 1990. A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11): 2821-2832.

McDonnell, J.J., 1997. Comment on "the changing spatial variability of subsurface flow across a hillslide" by Ross Woods and Lindsay Rowe. Journal of Hydrology (New Zealand), 36(1): 97-100.

McDonnell, J.J., Owens, I.F. and Stewart, M.K., 1991. A case study of shallow flow paths in a steep zero-order basin. Water Resources Bulletin, 27(4): 679-685.

McDonnell, J.J. et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, 43.

McGlynn, B.L. and McDonnell, J.J., 2003a. Quantifying the relative contributions of riparian and hillslope zones to catchment runoff. Water Resources Research, 39(11).

McGlynn, B.L. and McDonnell, J.J., 2003b. Role of discrete landscape units in controlling catchment dissolved organic carbon dynamics. Water Resources Research, 39(4).

McGlynn, B.L., McDonnell, J.J. and Brammer, D.D., 2002. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. Journal of Hydrology, 257: 1-26.

Montgomery, D.R. et al., 1997. Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. Water Resources Research, 33(1): 91-109.

Moran, C.J., McBratney, A.B. and Kopi, A.J., 1989. A rapid method for analysis of soil macropore structure. I. Specimen preparation and digital image production. Soil Science Society of America Journal, 53(3): 921-928.

Mosley, M.P., 1979. Streamflow generation in a forested watershed. Water Resources Research, 15: 795-806.

Mosley, M.P., 1982. Subsurface Flow Velocities Through Selected Forest Soils, South Island, New Zealand. Journal of Hydrology, 55: 65-92.

Noguchi, S., Nik, A.R., Kasran, B., Tani, M. and Sammori, T., 1997. Soil physical properties and preferential flow pathways in tropical rain forest, Bukit Tarek, Pennisular Malaysia. Journal of Forest Research, 2: 115-120.

Noguchi, S., Tsuboyama, Y., Sidle, R.C. and Hosoda, I., 2001. Subsurface runoff characteristics from a forest hillslope soil profile including macropores, Hitachi Ohta, Japan. Hydrological Processes, 15: 2131-2149.

O'Loughlin, C.L., Rowe, L.K. and Pearce, A.J., 1978. Sediment yields from small forested catchments, North Westland-Nelson, New Zealand. Journal of Hydrology (New Zealand), 17: 1-15.

Onda, Y., Komatsu, Y., Tsujimura, M. and Fujihara, J.-i., 2001. The role of subsurface runoff through bedrock on storm flow generation. Hydrological Processes, 15: 1693-1706.

Pearce, A.J. and Rowe, L.K., 1979. Forest management effects on interception, evaporation, and water yield. Journal of Hydrology, 18(2): 73-87.

Pearce, A.J., Stewart, M.K. and Sklash, M.G., 1986. Storm runoff generation in humid headwater catchments: 1. Where does the water come from? Water Resources Research, 22: 1263-1272.

Rowe, L.K., Pearce, A.J. and O'Loughlin, C.L., 1994. Hydrology and related changes after harvesting native forests and establishing Pinus radiata plantations. Part 1. Introduction to the study. Hydrological Processes, 8: 263-279.

Sidle, R.C., Hirano, T., Gomi, T. and Terajima, T., 2007. Hortonian overland flow from Japanese forest plantations - an aberration, the real thing, or something in between? Hydrological Processes, 21(23): 3237-3247.

Singh, P., Kanwar, R.S. and Thompson, M.L., 1991. Measurement and characterization of macropores by using AUTOCAD and automatic image analysis. Journal of Environmental Quality, 20: 289-294.

Sivapalan, M., 2003. Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection? Hydrological Processes, 17(5): 1037-1041.

Sklash, M.G., Stewart, M.K. and Pearce, A.J., 1986. Storm runoff generation in humid headwater catchments: 2. a case study of hillslope and low-order stream response. Water Resources Research, 22(8): 1273-1282.

Spence, C. and Woo, M.-K., 2002. Hydrology of subartic Canadian shield: bedrock upland. Journal of Hydrology, 262: 111-127.

Stewart, M.K. and McDonnell, J.J., 1991. Modeling base flow soil water residence times from deuterium concentrations. Water Resources Research, 27(10): 2681-2693.

Tani, M., 1997. Runoff generation processes estimated from hydrological observations on a steep forested hillslope with a thin soil layer. Journal of Hydrology, 200: 84-109.

Tippkötter, R., 1983. Morphology, spatial arrangement and origin of macropores in some hapludalfs, West Germany. Geoderma, 29: 355-371.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006a. Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope. Water Resources Research, 42.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006b. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resources Research, 42.

Tromp-van Meerveld, H.J., Peters, N.E. and McDonnell, J.J., 2006. 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.

Tsuboyama, Y., Sidle, R.C., Noguchi, S. and Hosoda, I., 1994. Flow and solute transport through the soil matrix and macropores of a hillslope segment. Water Resources Research, 30(4): 879-890.

Uchida, T., Kosugi, K. and Mizuyama, T., 1999. Runoff characteristics of pipeflow and effects of pipeflow on rainfall-runoff phenomena in a mountainous watershed. Journal of Hydrology, 222(1-4): 18-36.

Uchida, T., Tromp-van Meerveld, H.J. and McDonnell, J.J., 2005. The role of lateral pipe flow in hillslope runoff response: an intercomparison of non-linear hillslope response. Journal of Hydrology, 311: 117-133.

Waichler, S.R., Wemple, B.C. and Wigmosta, M.S., 2005. Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. Hydrological Processes, 19: 3177-3199.

Weiler, M. and McDonnell, J.J., 2007. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes. Water Resources Research, 43.

Weiler, M. and Naef, F., 2003. An experimental tracer study of the role of macropores in infiltration in grassland soils. Hydrological Processes, 17(2): 477-493.

Whipkey, R.Z., 1965. Subsurface storm flow from forested slopes. Bulletin of the International Association of Scientific Hydrology, 2: 74-85.

Woods, R. and Rowe, L., 1996. The changing spatial variability of subsurface flow across a hillside. Journal of Hydrology (New Zealand), 35(1): 51-86.

Woods, R. and Rowe, L., 1997. Reply to "Comment on "The changing spatial variability of subsurface flow across a hillside" by Ross Woods and Lindsay Rowe". Journal of Hydrology (New Zealand), 36(1): 51-86.

Zehe, E., Elsenbeer, H., Lindenmaier, F., K, S. and Blöschl, G., 2007. Patterns of predictability in hydrological threshold systems. Water Resources Research, 43.

Zehe, E. and Fluhler, H., 2001. Slope scale variation of flow patterns in soil profiles. Journal of Hydrology, 247: 116-132.

2.9. Tables

Table 2.1 Tracer breakthrough data for pit application

Soil Removed	Soil Remaining	Total Pore Volume	Input Rate	Time to Initial Rise	Time to Peak	Initial Rise Velocity	Peak Concentration Velocity	Active Pore Volume	Active Pore Space
(m)	(m)	(m3)	(L/s)	(s)	(s)	(m/s)	(m/s)	(m3)	%
0.00	8.00	16.0	9.75E-02	1200	3180	6.67E-03	2.52E-03	0.31	4.30
1.30	6.70	13.4	1.05E-01	1020	2520	7.84E-03	3.17E-03	0.26	4.38
1.90	6.10	12.2	6.27E-02	1800	3960	4.44E-03	2.02E-03	0.25	4.52
2.46	5.54	11.1	5.79E-02	1500	3240	5.33E-03	2.47E-03	0.19	3.76
3.13	4.87	9.7	5.92E-02	840	2880	9.52E-03	2.78E-03	0.17	3.89
3.13	4.87	9.7	2.94E-02	1560	4140	5.13E-03	1.93E-03	0.12	2.78
4.19	3.81	7.6	5.56E-02	1320	2820	6.06E-03	2.84E-03	0.16	4.57
4.20	3.80	7.6	4.33E-02	660	3120	1.21E-02	2.56E-03	0.13	3.95
6.62	1.38	2.8	6.04E-02	360	840	2.22E-02	9.52E-03	0.05	4.08
7.20	0.80	1.6	1.11E-01	120	240	6.67E-02	3.33E-02	0.03	3.70

Table 2.2 Tracer breakthrough data for surface application

Soil Removed	Soil Remaining	Total Pore Volume	Input Rate	Time to Initial Rise	Time to Peak	Initial Rise Velocity	Peak Concentration Velocity	Active Pore Volume	Active Pore Space
(m)	(m)	(m3)	(L/s)	(s)	(s)	(m/s)	(m/s)	(m3)	%
0.00	3.50	7.0	9.51E-02	480	1200	8.33E-03	3.33E-03	0.11	3.62
0.50	3.00	6.0	9.00E-02	480	1080	8.33E-03	3.70E-03	0.10	3.60
0.98	2.52	5.0	1.33E-01	420	840	9.52E-03	4.76E-03	0.11	4.92
1.25	2.25	4.5	4.50E-02	420	1920	9.52E-03	2.08E-03	0.09	4.27
1.25	2.25	4.5	9.00E-02	600	1080	6.67E-03	3.70E-03	0.10	4.80
1.25	2.25	4.5	1.70E-01	420	840	9.52E-03	4.76E-03	0.14	7.06
1.25	2.25	4.5	2.26E-01	300	660	1.33E-02	6.06E-03	0.15	7.37
1.25	2.25	4.5	3.05E-01	240	600	1.67E-02	6.67E-03	0.18	9.05
1.53	1.98	4.0	1.09E-01	480	1680	8.33E-03	2.38E-03	0.18	10.29
1.92	1.58	3.2	9.60E-02	420	1320	9.52E-03	3.03E-03	0.13	8.91
2.15	1.35	2.7	8.12E-02	300	1140	1.33E-02	3.51E-03	0.09	7.62
3.20	0.30	0.6	1.23E-01	120	300	3.33E-02	1.33E-02	0.04	13.69

2.10. Figures

Figure 2.1 Study site and excavation locations (grey, b). Excavated area with locations of surface and pit applications

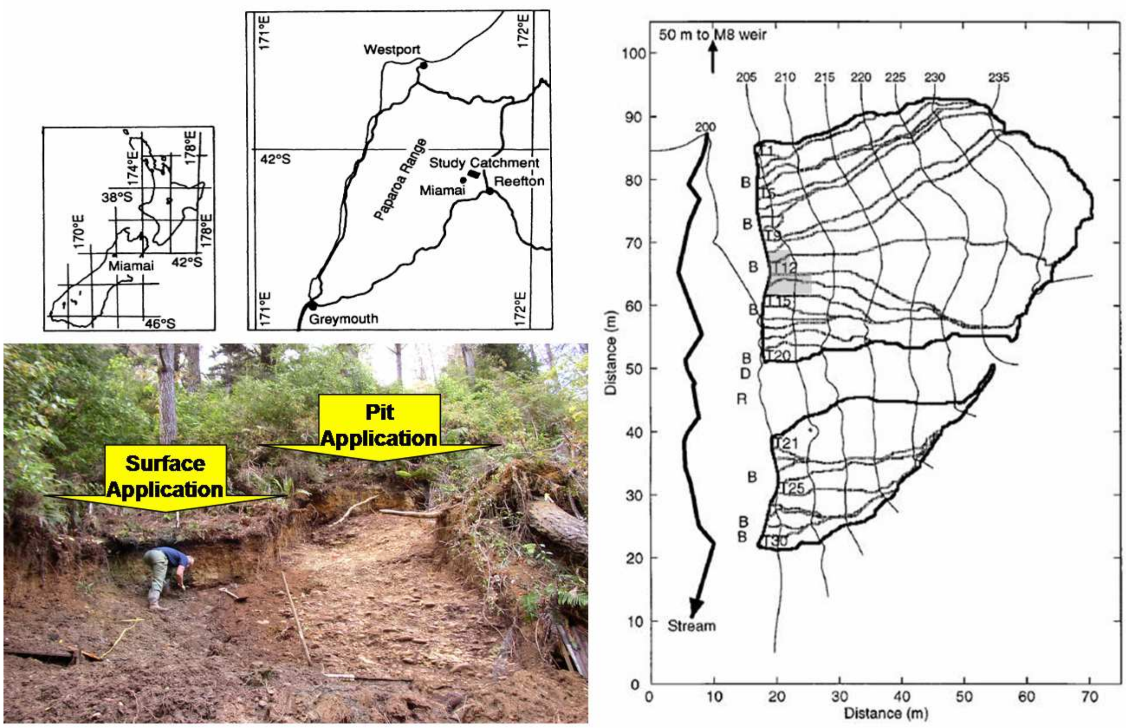


Figure 2.2 Precipitation discharge threshold relationship at Maimai hillslope (data from Woods and Rowe, 1996)

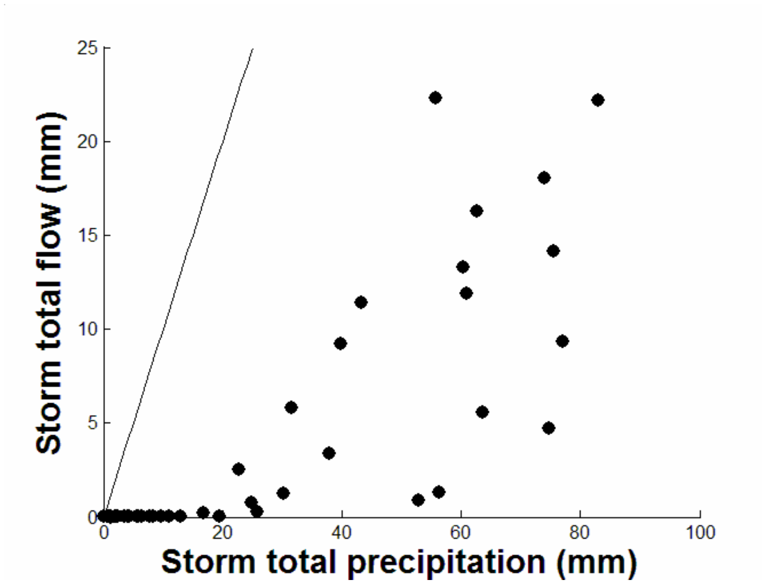


Figure 2.3 Schematic of excavation surfaces and bromide injections for surface (4 m upslope) and bedrock (8 m upslope) applications.

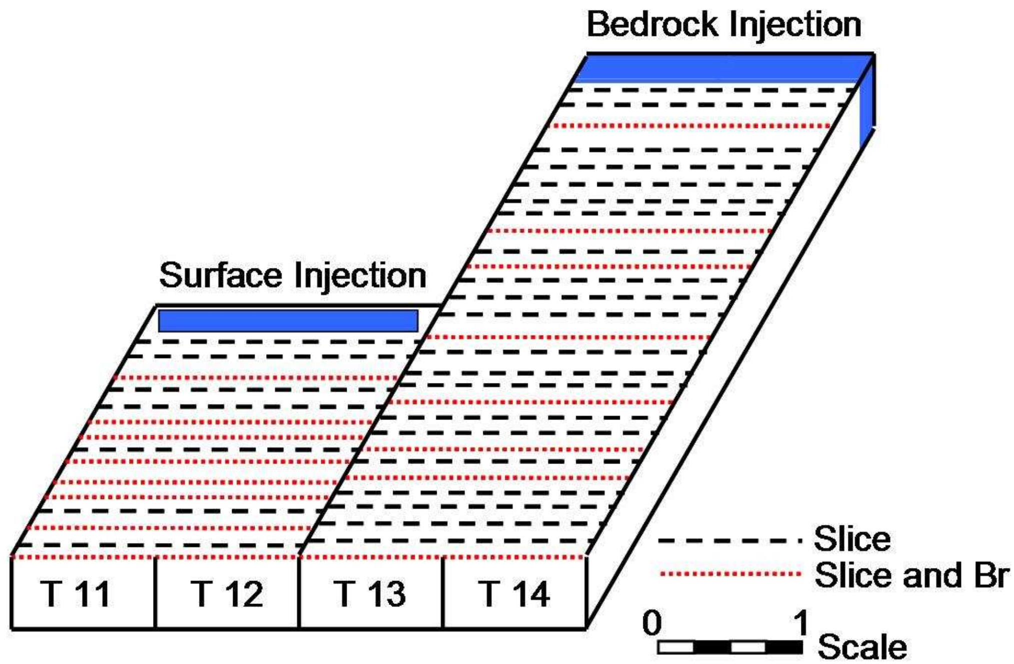


Figure 2.4 Trench excavated surface, with 0.5 m soil removed, 4 m downslope of surface application of dye and water. Note 4 areas of concentrated flow, coinciding with brown organic staining in lower soil profile.

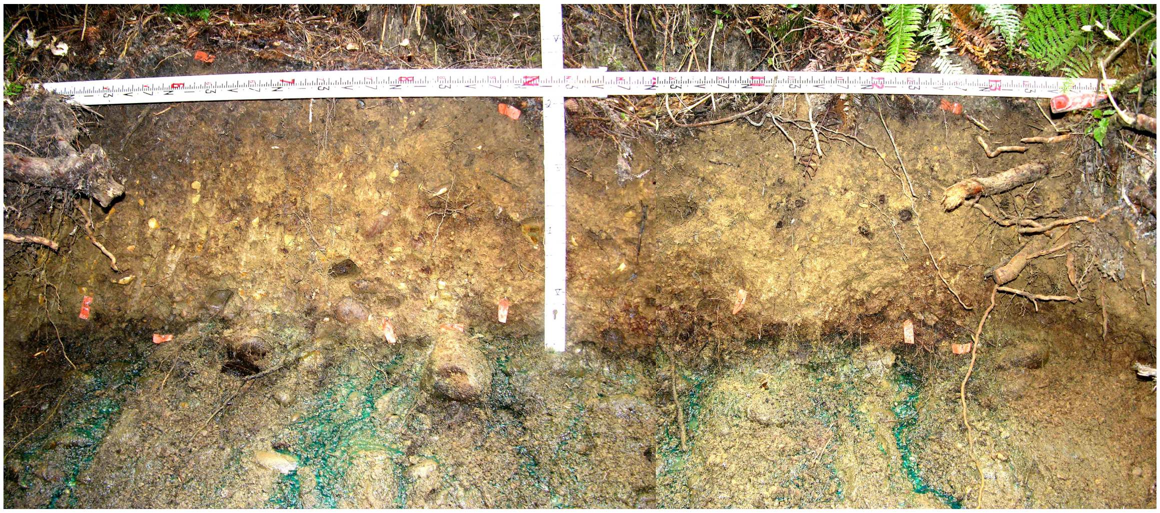


Figure 2.5 Preferential flow at soil bedrock interface with live roots (a). Organic staining on bedrock surface, indicating prolonged saturated conditions (b).

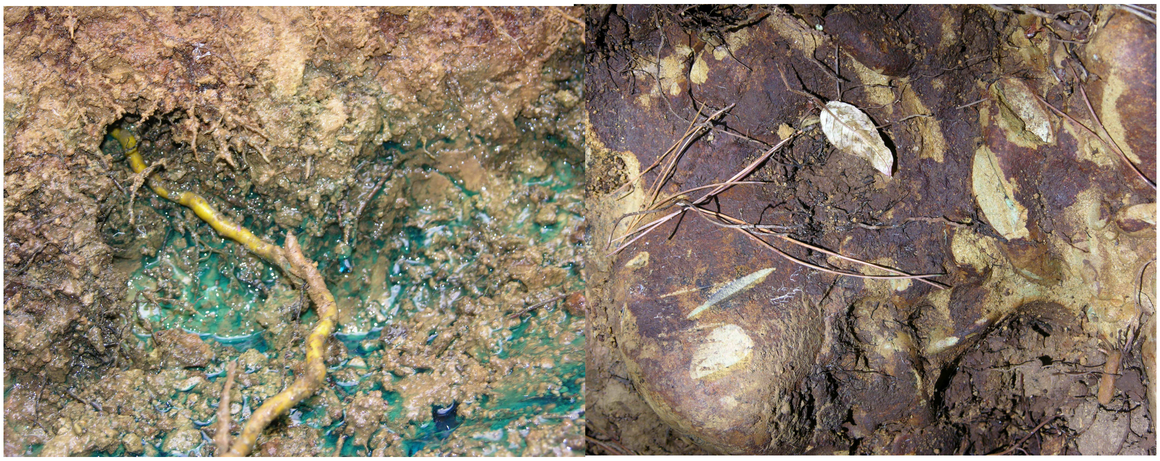


Figure 2.6 Tracer breakthrough at trench face of Br- tracer applied on soil surface. Tracer was applied 4 m upslope of the trench, and between 50 and 320 cm of soil was removed from the trench face upslope for each tracer application.

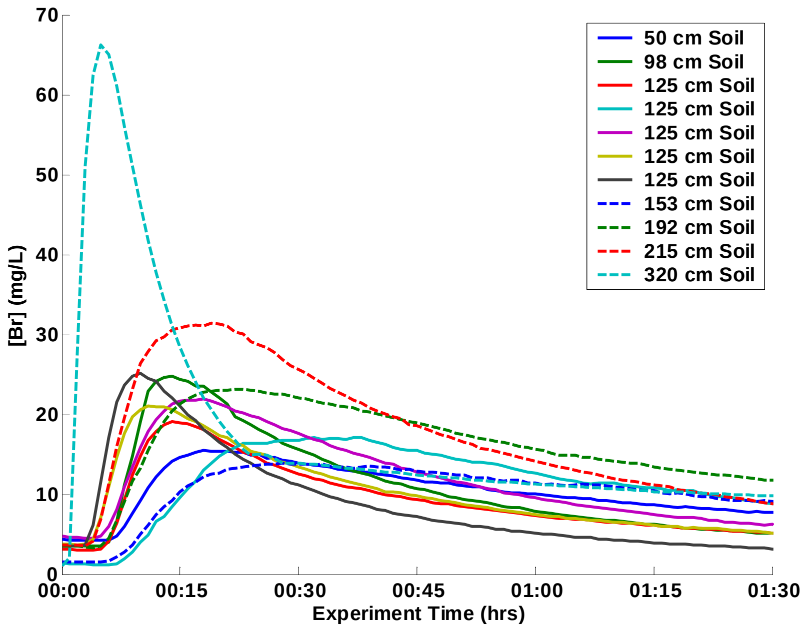
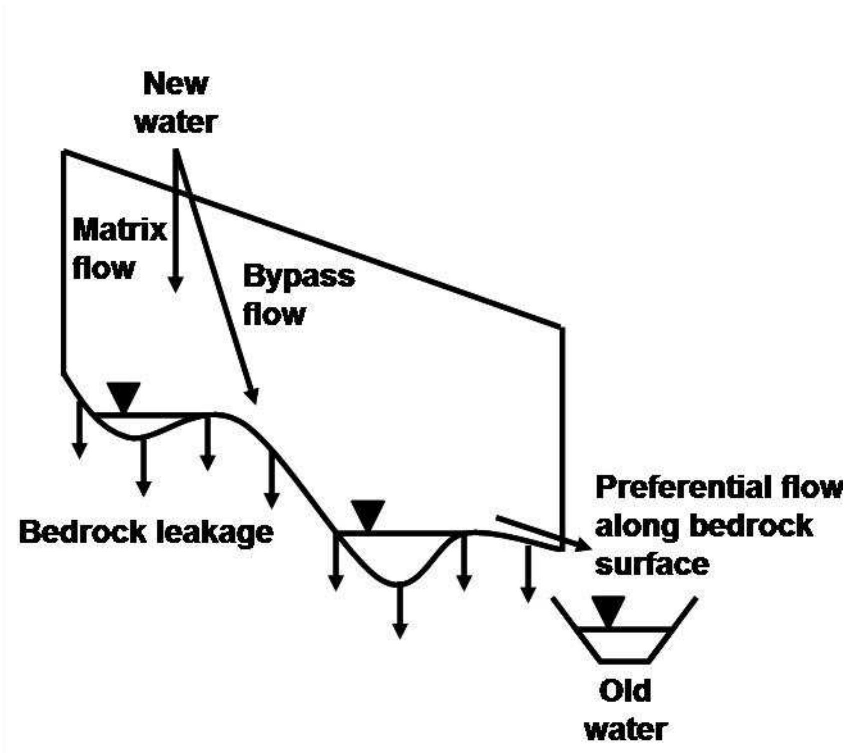


Figure 2.7 New perceptual model of lateral subsurface flow at the hillslope scale



3. Hillslope threshold response to storm rainfall: (2) Development and use of a macroscale behavioral model

3.1. Introduction

Hillslope hydrology still lacks the compact organization of empirical data and observations of responses that might facilitate extrapolation to and prediction of hillslope behavior in different places. Hillslope hydrology models based on our current small scale theories emphasize the explicit resolution of more and more of the unknown and unknowable heterogeneities of landscape properties and the resulting process complexities. (McDonnell et al., 2007) While the utility of a search for macroscale laws was enunciated over twenty years ago, (Dooge, 1986) few studies have been able to even observe macroscale behavior given the enormous logistical challenge for characterizing whole-hillslope response. The heterogeneity in hillslope soil, bedrock, and topographic conditions and complexity of the spatial and temporal rainfall and throughfall input are still extraordinarily difficult to quantify and resolve.

Graham et al. (this issue) presented a new macroscale perceptual model of subsurface flow processes at the well studied Maimai experimental watershed. (McGlynn et al., 2002) This work was based on whole-hillslope forensic analysis of subsurface flow paths and detailed hillslope scale irrigation aimed at identifying the dominant subsurface flow pathways and the role of bedrock topography and bedrock permeability at the hillslope scale. The complexities of hillslope response and heterogeneity of the hillslope site at Maimai could be summarized three key process statements: 1) A connected preferential flow network located at the soil/bedrock interface dominates lateral water and solute transport (with very high flow and transport velocities averaging 2.5E-3 m/s). 2) The bedrock surface controls the subsurface flow routing (where macroscale features in the highly textured bedrock influence the filling and spilling of small depressions and resultant threshold flow), and 3) Vertical loss to the permeable bedrock is large (up to 35% of the precipitation input) and delays lateral flow initiation and reduces lateral flow amount.

Here we take the dominant processes revealed by Graham et al. (this issue) and apply the dominant processes modeling concept of Grayson and Blöschl (2000) to construct, test and use a macroscale rainfall-runoff model for the Maimai hillslope. Within this philosophy, only the dominant flow processes, namely the three listed above from Graham et al. (this issue) are incorporated into the model structure. This is motivated by the difficulty in modeling all the complex and heterogeneous hydrological processes at a given site, in terms of parameter and process identification, and the computational limitations to such descriptions. The dominant processes approach is also motivated by the common finding that only a small number of processes may dominate hillslope flow and transport at the hillslope scale. We translate these processes into a simple, low dimensional conceptual mathematical model. This follows similar model development work at Maimai and elsewhere (e.g. Seibert and McDonnell, 2002; Son and Sivapalan, 2007) where only the most important factors are described in the model structure. (e.g. Weiler and McDonnell, 2007), or the model is built with enough flexibility so that the structure can be updated with new information in an iterative fashion (e.g. Fenicia et al., 2008; Vache and McDonnell, 2006). In this way the experimentalist works directly with the modeler, both in the experimental design to determine the dominant flow processes, and in model design to accurately implement the experimental findings.

We evaluate our new model using a multiple objective criteria framework (Gupta et al., 1998) incorporating extensive hydrometric and tracer data available from at Maimai site. Most importantly, we then use this new model as a learning tool to shed new light on whole-hillslope threshold responses to storm rainfall. Analysis of long term data records of flow at several field sites around the world has shown that this hillslope threshold response (i.e. the non-linear discharge response to precipitation) is a fundamental constitutive relation in hydrology (Buttle et al., 2004; Mosley, 1979; Peters et al., 1995; Weiler et al., 2006; Whipkey, 1965). This relationship can be expressed as a simple function with two variables:

$$Q = a\left(P - P_T\right) \qquad (1)$$

where Q is storm total hillslope discharge, a is the slope of the excess precipitation / discharge line, P is storm total precipitation and P_T is the precipitation threshold. The controls on the precipitation threshold and the slope of the excess precipitation / discharge line are not known. Both have been sown to vary with field site, possibly due to geologic and geometric catchment characteristics (Tromp-van Meerveld and McDonnell, 2006b; Uchida et al., 2005). Additionally environmental factors such as antecedent moisture conditions have been hypothesized to influence the threshold (Tani, 1997; Tromp-van Meerveld and McDonnell, 2006a). The nature and balance of the relative controls remains unclear and unresolved.

We use our new model to test alternative hypotheses of watershed scale threshold response to storm rainfall. In the “fill and spill” hypothesis subsurface storage at the base of the soil profile must be filled to connect the upslope areas with the base of the hillslope (Spence and Woo, 2002; Tromp-van Meerveld and McDonnell, 2006b). Alternatively, in the “soil moisture deficit” hypothesis unrequited storage in the soil profile must be filled before flow is initiated before lateral subsurface flow is initiated. The soil moisture deficit is supported by an apparent change in the threshold under different antecedent moisture conditions. (Tani, 1997; Tromp-van Meerveld and McDonnell, 2006b) Despite numerous observations of the threshold behavior in headwater watershed discharge records, the dominant controls and relative influence of fill and spill and soil moisture deficit factors on the threshold response to precipitation remains poorly understood. While things like storm spacing, intensity and duration effects, and evaporative demands may be able to be extracted from a long headwater flow record (where storms of different sizes but similar properties of intensity, duration, storm spacing etc…) the effects of the geologic factors such as the bedrock permeability or roughness are largely impossible to discern due to the high variability and uncertainty of these subsurface processes within and between catchments. We use our model as a learning tool to explore how subsurface processes represented in our model structure may link to those properties that can be extracted from a long terms data record, such as the threshold for initiation of storm runoff, and the relationship between the excess precipitation and runoff.

3.2. Study site and model development

3.2.1. Site physical and process description

We use the experimental work of Graham et al. (this issue) at the Maimai Experimental Catchments as the basis for model development and the virtual experiments aimed at understanding the controls on thresholds. The Maimai Experimental Catchments, South Island, New Zealand, have been a site of continuing hydrological research for over 30 years (see review in McGlynn et al., 2002). While isotopic work has shown that the majority of hillslope discharge and streamflow at Maimai is pre-event water stored for weeks to months, (McDonnell, 1990; Mosley, 1979; Pearce et al., 1986; Sklash et al., 1986) tracer experiments have demonstrated the ability of the hillslopes to rapidly transmit quantities of applied water at great velocities over long distances. (Brammer, 1996; Mosley, 1979; Mosley, 1982) Graham et al. (this issue) showed that lateral preferential flow is confined to the soil bedrock interface where flow velocities are very high (up to 5.6 m/s), routed by the bedrock topography and modulated by storage on the bedrock surface. The preferential flow paths seen at the soil bedrock interface have been shown to be well connected upslope for distances up to 8 m, and appear to be stationary in time and space Graham et al. (this issue). The bedrock, while previously considered effectively impermeable (McDonnell, 1990, pg. 2821; Mosley, 1979, pg. 795), has been shown to be semi-permeable, with bedrock hydraulic conductivity on the order of 2.8 - 8.3E-7 m/s, leading to the potential of substantial flowpaths through the bedrock. Overland flow has not been observed at this site except in limited areas near the stream channel. Vertical preferential flow from the soil surface to depth during extreme events has been hypothesized to occur in vertical cracks seen throughout the catchment dissecting the soil profile (McDonnell, 1990; Graham et al., this issue). Mixing of old and new water is thought to occur in both the soil column as well as in transient groundwater that forms at the soil bedrock interface, leading very low amounts of new water observed in trench discharge and streamflow. (Pearce et al., 1986; Sklash et al., 1986)

3.2.2. Description of the numerical model

The numerical model (called MaiModel) was built to incorporate the dominant processes that control subsurface flow at the Maimai hillslope as described by Graham et al. (this issue). Key components of MaiModel are

- Preferential flow pathways are connected, and located at the soil bedrock interface.
- Lateral subsurface travel velocities are high.
- Subsurface storage on the bedrock surface is explicitly designated.
- The bedrock is permeable.

- During high rainfall and/or saturated conditions rainfall can bypass the soil profile to depth.

In general terms, MaiModel consists of two reservoir types, soil storage and bedrock pool storage, which are fully distributed across the model domain. Two bulk reservoirs are included for system losses of evapotranspiration and bedrock leakage. Water is transmitted vertically from the soil surface, through the soil storage reservoir to the bedrock pool reservoir, with no lateral communication between soil reservoirs. Lateral subsurface flow is restricted to flow along the bedrock surface among the bedrock pools, consistent with the current experimental evidence of water routing at the soil bedrock interface. Bedrock leakage is controlled by water table height, and there is no reemergence of water once it has percolated into the bedrock. Evapotranspiration is driven by a set potential evaporation rate, and controlled by soil moisture.

3.2.3. Model structure

In MaiModel precipitation is split into transpiration losses and soil reservoir recharge. A map of tree locations by Woods and Rowe (unpublished data, 1996) was used to identify areas of interception. Interception rates are based on published values from a similar aged radiata pine (*Pinus radiata*) forest, showing an interception rate of 38% under the canopy. (Putuhena and Cordery, 2000) Interception was confined to areas of crown cover, which were estimated as the area within 3 m of the tree stems.

Throughfall enters each soil reservoir and is fully mixed with pre-event soil moisture, following Weiler and McDonnell (2007) and Vache and McDonnell (2006). Assuming a unit head gradient, vertical drainage to the bedrock pool reservoir (Q_{soil}) is equal to the soil relative hydraulic conductivity, using the (Brooks and Corey, 1964)formula:

$$Q_{soil} = k_{soil} A S^{\beta} \tag{2}$$

where k_{soil} is the saturated hydraulic conductivity, A is the grid cell area, β is the Brooks and Corey exponent determined by the soil texture, and S is dimensionless water content:

$$S = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{3}$$

where θ is water content, and θ_r and θ_s are the residual and saturated water content, respectively. S is restricted to a range between zero and one. $Q_{soil} = 0$ if θ is less than θ_r. Water drains vertically from the soil elements to bedrock pool elements, and does not drain downslope into adjacent soil elements.

Neither infiltration overland flow nor saturation excess overland flow has been observed on the side slopes of the Maimai catchment, even during high intensity irrigation experiments. (McGlynn et al., 2002) Bypass flow via sub-vertical fractures in the soil has been hypothesized as a mechanism for rapid water transport to depth. (McDonnell, 1990; Mosley, 1979; Woods and Rowe, 1996) Flow along these fractures was also observed in the irrigation experiments at very high irrigation rates. (Graham et al., this issue) If θ is greater than the soil pore volume, bypass flow to the bedrock pool is initiated by:

$$Q_{bypass} = fz_{soil} - \theta \tag{4}$$

where f is the porosity, z_{soil} is the soil depth, and θ is the water volume in the reservoir. In this way, θ is restricted to within the available porosity.

Evaporation from the soil reservoir is computed similarly to Seibert (1997)

$$ET = PET(tod)S \tag{5}$$

where S is the dimensionless water content and *PET(tod)* is potential evaporation, defined as

$$PET(tod) = PET_{daily} \sin\left(2\pi\left(tod - \frac{2}{24}\right)\right) \tag{6}$$

where PET_{daily} is the daily potential evaporation rate, *tod* is the Julian time of day. Evaporation depends on the soil reservoir water content, and is not constant across the domain temporally or spatially. We assume that *PET* peaks at 14:00 hr and reaches a minimum at 02:00 hr of each day.

Water drains from the soil element to pools of set volume on the bedrock surface. Water in the pool will either drain into the bedrock or be routed downslope along the bedrock to adjacent bedrock pools, following the fill and spill hypothesis of. (Tromp-van Meerveld and McDonnell, 2006b) All pools are connected, and flow routing is based on topography, with lateral subsurface flow (Q_{LSS}) split between the (up to) eight adjacent downslope elements, weighted by downslope gradient following the M8 flow routing algorithm Quinn et al. (1991).

Lateral subsurface flow (Q_{LSS}) is governed using the Dupuit - Forchheimer assumption for a sloping aquifer. (Freeze and Cherry, 1979) Q_{LSS} is a function of the grid element width (w), local gradient (s), the lateral hydraulic conductivity k_{LSS}, and the volume of stored water (S_{pool}) greater than the subsurface storage volume (V_{pool}):

$$Q_{LSS} = wsk_{LSS}\left(S_{pool} - V_{pool}\right) \tag{7}$$

As per field measurement of infiltration into the bedrock at the site, (Graham et al., this issue) bedrock leakage is a function of water table height and a bedrock leakage coefficient ($c_{bedrock}$) and grid element cross sectional area (A):

$$Q_{bedrock} = c_{bedrock} S_{pool} A \tag{8}$$

The tracer is injected into the soil elements as point injections and fully mixed in the soil and bedrock pool reservoirs. Tracer fluxes are limited to advective transport vertically between soil and bedrock pool elements, and laterally between bedrock pool elements. While tracer can percolate into the bedrock, no tracer is lost to evapotranspiration.

3.2.4. Multiple criteria model calibration

We used a multiple criteria model calibration to determine which parameter sets produced behavioral models, models that acceptably reprodced aspects of the system behavior (Beven and Freer, 2001). We capitalize on two extensive data sets from field campaigns at the site for model parameterization and calibration. Woods and Rowe (1996) built a 1 m grid DEM survey consisting of 755 survey points over an area of 2830 m^2, which was used in the model for flow routing. In addition, the interception module is based on a map of tree locations by Woods and Rowe (unpublished data, 1996). For 65 days beginning March 10, 1995, (Brammer, 1996) monitored hillslope discharge at the trench system built by Woods and Rowe (1996). A larger section of hillslope is modeled than drains into the collection trench, to minimize edge effects. In addition to monitoring hillslope discharge, Brammer (1996) added a Br^- tracer 35 m upslope of the trench as a 20 m wide line source injected directly into the soil profile 0.1 m below the soil surface (Figure 3.1). Precipitation, hillslope discharge and tracer breakthrough were monitored at the trench for 45 days after tracer application. Rainfall and trench discharge were recorded in 10 minute intervals, while tracer breakthrough at the trench was measured in grab samples during and between storms. Reanalysis of the Brammer (1996) tracer concentrations and trench discharge show that 15% of the tracer was recovered over 45 days, and the runoff ratio for the duration of the monitoring was 21%. The tracer breakthrough and trench hydrograph time series were used for model calibration. For more details about the hillslope gauging system, see Woods and Rowe (1996). For more details on the tracer injection, see (McGlynn et al., 2002).

MaiModel was calibrated using a Monte Carlo analysis with multiple criteria including hydrometric and tracer breakthrough data. Using the 40 day Brammer (1996) hyetograph as input, 10,000 simulations were run with five model parameters varied in calibration: soil hydraulic conductivity (k_{soil}), bedrock leakage coefficient ($c_{bedrock}$), lateral hydraulic conductivity (k_{LSS}), active pore space ($(\theta_s-\theta_r)$*z, f_{active}), and the subsurface storage volume (V_{pool}). Changing the residual water content, field capacity water contents and soil depth had the same impact on the active pore space, so one factor, θ_r, was chosen for calibration. The active pore space is presented as the variable in further analysis. Monte Carlo analyses were performed varying each parameter randomly across ranges of 0 – 1,000% of field measurements or the physically possible range, to ensure that the entire parameter space was interrogated. Field parameter measurements and ranges used in the model calibration are presented in Table 1.

The second subset of model parameters was assigned to field measurements due to either parameter uncertainty or model insensitivity. A spatially detailed soil depth map was unavailable, so soil depth was set at the average uniform soil depth for the modeled domain. While modeling variations in soil depths has been shown to be important for prediction of hillslope dynamics at other field sites (e. g. Tromp – van Meerveld & Weiler, submitted), previous work at Maimai has indicated that the soil surface and bedrock surface topography are similar at the hillslope scale, and the soil depth is relatively uniform across the hillslope (Woods and Rowe, 1997). In preliminary calibration runs, MaiModel was found to be insensitive to the Brooks and Corey moisture release coefficient (β), so it was set at a value appropriate for the silt loam soil texture measured in the field (Carsel and Parrish, 1988). As mentioned above, porosity, residual and field capacity water contents are interrelated with respect to model function, so residual water content was set as a variable during the calibration, and the porosity and field capacity water content were set to field measured values (from McDonnell, 1990). No measurements of PET were available at the site for the time of record. At the nearby town of Reefton (10 km northwest of the hillslope), a measured evaporation rate of 714 mm/year has been reported (Baker and Hawke, 2007). The model parameter PET_{daily} was set to 6 mm/day to result in a modeled evaporation rate of 714 mm/year (1.95 mm/day).

Model performance with respect to the hillslope discharge hydrograph was assessed by the Nash Sutcliffe (Nash and Sutcliffe, 1970) efficiency factor (E):

$$E = 1 - \frac{\sum (Q_o - Q_m)}{\sum (Q_o - \overline{Q_o})} \qquad (9)$$

where Q_o is observed discharge and Q_m is modeled discharge. An E value greater than 0 indicates the modeled results fit measured discharge better than the mean discharge. An E of 1.0 is a perfect fit. For calibration purposes, an E of over 0.8 was considered an acceptable fit. E calculations were made for six subsets of the time series, including the entire 40 days after tracer application, and for the five largest storms of the data record (storms B1 – B5). Only parameter sets with acceptable E for both the 40 day record and each individual storm were considered behavioral.

Due to temporally irregular measurements of tracer breakthrough at the hillslope, model tracer breakthrough was compared on a storm by storm basis from the Brammer (1996) monitoring. Both the spatial pattern of tracer breakthrough along the trench face and storm cumulative breakthrough were compared for each of the five recorded storms. Parameter sets with correlation coefficient greater than 0.8 for both the spatial and temporal breakthrough comparisons were considered acceptable. Cumulative tracer breakthrough for the 40 day time series was also used as a model evaluation criterion. Due to uncertainties in tracer recovery, modeled tracer recovery of within 2.5% of measured values was considered acceptable. After Monte Carlo calibration, the parameter sets meeting all criteria were analyzed, and the model run with the highest minimum storm E was chosen for additional virtual experiments.

3.2.5. Virtual experiment design

All virtual experiments were performed using a rainfall time series with multiple replications of storm B5 (April 26 through April 27, 1995) from the calibration hyetograph as input. Storm B5 was chosen for the virtual experiments due to its moderate size, variable intensity, and the relatively good model fit from calibration. This 50.6 mm storm had a duration of 24 hours, average intensity of 2.1 mm/hr, and maximum 10 minute (60 minute) intensity of 30 (11.8) mm/hr occuring 70 (240) minutes after the start of the storm. This storm was chosen for the virtual experiments as it was the largest of the five storms monitored after tracer application and exhibited the highest rainfall intensity. Measured discharge was 23.0 mm (runoff ratio = 44%), with peak discharge of 3.4 mm/hr, 400 minutes after the start of the storm. The hyetograph and hydrograph were skewed to the left (skew = 3.4 and 2.9 respectively). 87.2 and 87.4 mm rainfall fell in the previous 7 and 14 days. Analysis of 2 years of precipition records at the site (from McDonnell et al., 2009) indicates that this storm falls in the upper 25% and 5% of storms with respect to rainfall total precipitation and average intensity, respectively. The calibrated model had an efficiency of 0.95 for this storm.

To allow for the impact of antecedent moisture conditions in the virtual experiment hyetographs, storm B5 was replicated 9 times, bracketed between 1 to 21 days antecedent drainage time before each replication. With the soil reservoirs initially set at saturation, the rainfall time series consisted of 10 days drainage followed by the B5 hyetograph. The B5 hyetograph was then repeated 9 times with 1, 1, 3, 3, 5, 7, 7, 14 and 21 days drainage between storms (storms V1-V9). The model was run with the virtual experiment hyetograph 11 times, with the total storm precipitation scaled by a factor of 0.1, 0.2, 0.5, 0.75, 0.9, 1, 1.1, 1.5, 2, 5 and 10 (5-506 mm per event). Total storm precipitation was scaled by altering the duration of the storm to reach the desired total storm rainfall amount. For events smaller than the base case storm (50.6 mm), the rainfall ceased once the desired rainfall amount was reached. For storms larger than the base case storm, the storm time series was repeated until the desired precipitation amount was reached. Depending on the size (duration) of the virtual experiment events, the duration of the VE hyetograph ranged from 95 to 185 days.

For each simulation recorded output included water balace components (discharge, bedrock leakage, ET and soil moisture storage), and tracer fluxes. Total event precipitation and total storm discharge were calculated, binned by antecedent moisture and plotted. The threshold and excess precipitation / discharge slope for the calibrated model were determined based on all storms with a runoff ratio greater than 1% using a least squares regression.

For the virtual experiments that focused on the effect of fill and spill factors, only storm V5, with 3 day drainage was used for analysis. The simulations with the scaled virtual experiment hyetographs were repeated while also scaling the two fill and spill parameters, bedrock permeability and subsurface storage volume. Both bedrock permeability and subsurface storage volume were scaled by a factor of 0.1, 0.2, 0.5, 0.75, 1, 1.5, 2, 5, and 10 (9 factors), and scaled concurrently, to create 81 simulations for each event size (11). 11 * 81 = 891 simulations were run covering the ranges of parameter values and event sizes. The process used to determine the threshold and slope for the calibrated model was then repeated using model runs with the varied bedrock permeability and subsurface storage volumes for each storm (antecedent moisture).

For soil moisture deficit experiments, the potential evaporation rate was varied, and the drainage time between storms from the virtual experiment hyetograph was analysed as the second variable. The potential evaporation rate was scaled in a similar fasion as the fill and spill parameters (11 sets), and precipitation was scaled as before (nine hyetographs), resulting in 9*11 = 99 simulations.

A final 3 sets of experiments tested the hypotheses that neither or both fill and spill or soil moisture deficit are the cause of the threshold. In the first set of simulations $c_{bedrock}$ and V_{pool} were set to zero, eliminating the fill and spill mechanism of threshold development. In the second set of simulations, PET_{daily} was set to zero, eliminating the soil moisture deficit mechanism. In a final set of simulations, PET_{daily}, $c_{bedrock}$ and V_{pool} were set to zero, to determine if a third mechanism beyond “fill and spill” and “soil moisture deficit” could be responsible for the thresholds.

3.3. Results

3.3.1. Multi-criteria model calibration

The calibrated model was able to reproduce both the hydrometric and tracer response to precipitation. The calibrated model fit the measured hydrograph well, at both the 40 day and individual event time scale (Figure 3.2). While the smaller events were generally overpredicted, the large events were well modeled. The hydrograph recessions were generally underpredicted, with the model exhibiting a faster recession than that measured. Peak discharge for each event was well represented.

Our 9 calibration objective criteria provided different levels of model discrimination (Table 3.2, Figure 3.3). The hydrographs and cumulative tracer breakthrough were effective in reducing the number of acceptable parameter sets, reducing the total number of behavioral parameter sets between 96 – 99%. The spatial and temporal tracer breakthroughs were not as effective at parameter range reduction. Of the 10,000 model runs, 294 (2.9%) parameter sets led to acceptable model fits of trench discharge during the entire 40 day record (where an acceptable fit was a Nash-Sutcliffe efficiency (E) greater than 0.8). The best fit parameter set had an E of 0.95. For storms B1, B3, B4 and B5, a similar number of model runs (parameter sets) were found to be acceptable (260 - 416 (2.6-4.2%)), with maximum E of 0.97 – 0.98. Fewer parameter sets led to acceptable model fits of storm B2, (E > 0.8 for 102 (1.0%) parameter sets), though a maximum E of 0.97 was obtained. 300 (3.0%) model runs had between 11.5 and 16.5% (+- 2.5% of measured) of tracer breakthrough during the flow time series. The Brammer (1996) experimental tracer temporal and spatial breakthrough were reproduced by a minority of parameter sets; 1462 (14.6%) correctly modeled temporal tracer breakthrough, while 4827 (48.3%) correctly modeled the spatial breakthrough.

Of the 10,000 parameter sets, 13 (0.1%) met all nine calibration criteria. Four criteria (storms 1, 2, and 3, and the cumulative tracer breakthrough) were sufficient to determine the final group of acceptable parameter sets. Of the 294 simulations that modeled the entire 40 day trench hydrograph acceptably, 39 (13%) modeled all 5 storms acceptably. All parameter sets that had an acceptable fit for the five individual storms also adequately represented the 40 day hydrograph. Of the 39 simulations with acceptable fit for all storms, 13 (33%) had modeled the cumulative tracer breakthrough adequately.

3.3.2. Examination of parameter uncertainty within the calibrated runs

Each objective criterion served to reduce the range of each model parameter (Table 2). To compare the reduction in the uncertainty of each parameter after each objective criterion, the ratio of the parameter values in behavioral models to the range of the initial parameter distribution was computed. A ratio of 100% for a given parameter and objective criterion indicates the criterion did not reduce the uncertainty in the parameter, while a ratio of 10% indicates the parameter is restricted to 10% of the initial range for behavioral models. The initial range of the five calibrated parameters was set to 0 - 1,000% the measured value (in the case of bedrock leakage coefficient and soil hydraulic conductivity), the range physically possible (drainable porosity) or over a wider range than produced behavioral model runs (lateral subsurface flow coefficient, and subsurface storage volume). These values are relative as the initial parameter ranges chosen were different for each parameter, generally spanning the range observed in the field (k_{soil}, $c_{bedrock}$, k_{LSS}), the range of physically possible values (f_{active}), and the range possible determined through pre-experiment sensitivity analysis (V_{pool}). Nevertheless, they serve as a method to determine the relative strength of each objective criterion to reduce parameter uncertainty.

The amount of parameter uncertainty reduced by each objective criterion varied for each parameter (Table 3.2, Figure 3.3). In the case of the active porosity, parameter sets that led to acceptably modeled tracer fluxes sampled from the entire initial range. Parameter sets that acceptably modeled the 40 day and individual hydrographs, however, reduced the range of drainable porosity from 47 – 68%. Soil hydraulic conductivity, on the other hand, was insensitive to all of the objective criteria. While relatively few parameter sets resulted in acceptable cumulative tracer breakthrough (3.0%), this objective did little to reduce the uncertainty for any of the calibrated parameters except the bedrock leakage coefficient (43.6%). Using only the 13 parameter sets that fit all criteria, the parameter uncertainty was significantly reduced (35.9 – 62.6%) for all parameters except the soil hydraulic conductivity (3.7%).

Of the 13 parameter sets deemed behavioral for all objective criteria, one was chosen as the base case for the virtual experiments. The 13 parameter sets that met all of the model evaluation criteria were then ranked according to their goodness of fit to each of the objective criteria. The parameter set with the highest average rank was chosen for the virtual experiments. For the best fit parameter set, model efficiency for the five storms ranged from 0.92 (storm B2) – 0.97 (storm B5), with a 40 day efficiency of 0.95. Modeled cumulative tracer breakthrough was within 2% of the measured value. Modeled spatial and temporal patterns of tracer breakthrough all fell in acceptable ranges. Best fit parameter values were close to measured values. The drainable porosity was 0.1 m^3/m^3 (compared to 0.05 m^3/m^3 measured in the field (McDonnell, 1990)), bedrock leakage coefficient, 4.25 E-5 1/s (compared to 2.63 E-5 1/s reported by Graham et al., this issue) and lateral hydraulic conductivity, 7.1E-3 m/s (compared to 2.1E-3 – 5.6E-3 m/s observed in the field). Soil hydraulic conductivity was calibrated to 2.67E-4 m/s, nearly an order of magnitude greater than observed values between 2.7 E-6 – 8.3E-5 m/s (Mosley 1979). The average subsurface storage volume (1.7 mm) was not measured in the field.

3.3.3. Virtual experiments with the calibrated model

3.3.3.1. Base case parameterization

Using the base case model parameterization, the 11 scaled virtual experiment hyetographs were applied, with storm totals precipitation ranging from 5 – 506 mm (for an example set of hydrographs, see Figure 3.4). For the scaled realizations of storm V5, with 3 days antecedent drainage, total discharge ranged from 0 – 272.7 mm (runoff ratio = 0 – 57%). Trench flow was not observed for the 2 smallest events (5 and 10.1 mm rainfall), while the 25 mm storm yielded 3.6 mm of trench discharge (runoff ratio = 14%). Therefore a threshold for lateral subsurface stormflow appeared to exist between 10 – 25 mm. Total storm discharge increased linearly (R^2 = 0.999) after the threshold, with a slope of 0.59 mm discharge / mm precipitation. The calculated threshold, equal to the x axis intercept, was 17.7 mm precipitation. In the analysis below, the threshold refers to the x axis intercept (reported in mm rainfall), and the slope is the slope of the excess precipitation / discharge line (reported in mm discharge / mm precipitation).

3.3.3.2. Soil moisture deficit

For the application of the virtual experiment hyetograph using the base case parameterization and the base case storm (B5, 50.6 mm), total storm discharge was dependent on the antecedent drainage time. Total storm hillslope discharge decreased from 23.5 to 0.0 mm (runoff ratios decreased from 46 to 0%) for the storms with between 1 and 21 days of antecedent drainage (Table 3). For the simulated storm (V1) with the shortest antecedent drainage, 1 day, the water balance was split between discharge (23.5 mm or 46%), bedrock leakage (26.4 mm or 52%), evaporation (10.9 mm or 22%). The soil storage reservoir acted as a source for the additional water for all simulations, as soil moisture storage decreased from event to event. For the storm (V5) with 3 days antecedent drainage (the average time between storms at Maimai, as deduced from the long term data record), discharge decreased to 18.2 mm (36%), along with bedrock leakage to 23.7 mm (27%). Evaporation increased (15.9 mm or 32%), and the soil storage reservoir again acted as a source for the additional water. For the storm with 21 days antecedent drainage time (V9), longer than any observed at Maimai in the 2 year data record, (McDonnell et al., in review) discharge was reduced to 0 mm (0%), and bedrock leakage to 5.2 mm (10%). Evaporation increased to 64.4 mm (129%). For the storms with long antecedent drainage times, rainfall went to filling soil storage and then lost to evaporation. For storms with shorter antecedent drainage times, soil moisture deficit was quickly filled and precipitation was routed to the bedrock surface and lost to hillslope discharge and bedrock leakage.

The rainfall threshold for producing subsurface stormflow for storm V5 (3 days antecedent drainage) was 17.7 mm (Figure 3.5). Calculated thresholds for the other events (time between events) ranged from 9.1 (1 day antecedent drainage) to 60.8 mm (21 days antecedent drainage) (Table 4). The threshold was linearly related to the time between storms (Figure 3.6) of the form

$$P_T = 9.7mm + t_{drain}\left(2.5\, {}^{mm}\!/_{day}\right) \qquad (10)$$

(R2 = 0.984). The slope appeared independent of the time between events, varying from 0.56 – 0.57 mm / mm (Figure 3.7). The 11 virtual experiment hyetographs were then run with the potential evaporation rate scaled between 10 and 1,000% of the calibration value of 6 mm/day (0.6 – 60 mm / day). For the following virtual experiments we will focus on storm V5, with 3 days antecedent drainage, as similar patterns were seen for all events.

As the potential evaporation rate was increased, actual modeled evaporation increased, and pre-event soil moisture decreased. These losses were balanced by a decrease in both discharge and bedrock leakage. Varying the potential evaporation rate, the total storm hillslope discharge fell from a high of 23.0 mm (45%) at low PET (0.6 mm/hr) to a base case 18.2 mm (36%) and a low of 0 mm (0%) at very high PET (60 mm/hr). Bedrock leakage also fell from 30.2 mm (59%) to 23.8 mm (45%) to 0 mm for the low PET, base case parameterization and high PET simulations, respectively. Evaporation rates increased concurrently, from a low of 1.6 mm (3%) to a base case of 15.9 (32%) to a high of 77.4 (155%). Once the PET increased above 200% of the base case (>12 mm / hr), evaporative losses from the soil profile were greater than the total storm precipitation, leading to a progressively depleted soil moisture profile, and providing a water source for the high evaporation rates. For the simulations with high potential evaporation rates, rainfall went to filling soil storage and then lost to evaporation. For simulations with lower potential evaporation rates, soil moisture deficit was quickly filled and precipitation was routed to the bedrock surface and lost to hillslope discharge and bedrock leakage.

The precipitation / discharge threshold and the slope of the excess precipitation / discharge line for the soil moisture deficit virtual experiments were calculated in the same way as above, and only the summary results are presented here. For the calibrated PET (6 mm / day), the precipitation discharge threshold for the 3 day drainage event is 17.7 mm, as before, and reached an upper bound of 104.8 mm (PET = 60 mm/day) and lower bound of 12.4 mm (PET = 0.6 mm) (Figure 3.6). The slope of the excess precipitation discharge line decreased with increasing PET, from an upper limit of 0.62 mm / mm (PET = 0.6 mm/day) to a lower limit of 0.21 mm / mm (PET = 60 mm/day) (Figure 3.7). For simulations with no evaporation (PET = 0), the threshold for event 5 was 11.8 mm and the slope 0.62 mm / mm. For event 5, the threshold for initiation of hillslope discharge is positively linearly correlated with the potential evaporation rate (R^2 = 0.996), of the form

$$P_T = 11.4mm + PET\left(1.1 \frac{mm}{\frac{mm}{day}}\right) \qquad (11)$$

The slope of the excess precipitation discharge line (a) was also linearly correlated with PET (R^2 = 0.999) of the form

$$a = 0.63 \frac{mm}{mm} - PET\left(0.007 \frac{\frac{mm}{mm}}{\frac{mm}{day}}\right) \qquad (12)$$

Our analysis suggests that the effects of PET and the time between events are multiplicative (Figure 3.8, Table 3.4). The minimum threshold (5.8 mm) occurred with minimum antecedent drainage and minimum PET. The maximum threshold (240 mm) occurred with the maximum antecedent drainage and maximum PET. The product of PET and the antecedent drainage time (a measure of the total potential evaporative demand before the event) was positively correlated with the modeled threshold

$$P_T = 12.2mm + t_{drain} PET\left(0.35 \frac{mm}{mm}\right) \qquad (13)$$

The slope of the excess precipitation / discharge line was solely correlated with PET (Equation 12), with antecedent drainage time having no effect.

3.3.3.3. Fill and spill

To determine the influence of fill and spill factors on the precipitation discharge relationship threshold, the bedrock leakage coefficient and subsurface storage volumes were varied using the same factors as the PET experiments. Again, event V5, with 3 days antecedent drainage, was used for the analysis. For simulations where the bedrock leakage coefficient was increased, bedrock leakage increased while hillslope discharge decreased. Due to the physical disconnect between the soil profile and the bedrock in the model structure, changing the bedrock permeability did not impact either soil moisture storage or evaporation rates. At the lowest simulated value of the bedrock leakage coefficient (10% the calibrated value), leakage to bedrock was 11.9 mm (24%), while hillslope discharge was 35.8 mm (71%). The remainder of the water balance was composed of evaporation and decreased subsurface storage in the bedrock pools. For the base case parameterization, leakage into the bedrock accounted for 23.7 mm (47%), discharge for 18.2 mm (36%). For the high permeability scenario (bedrock leakage coefficient was 1,000% base case), leakage accounted for 37.8 mm (75%), while discharge was 0.6 mm (1%). The simulations with varied subsurface storage volumes showed a similar pattern, with increasing bedrock leakage (15.6 mm (31%); 23.7 mm (47%); 38.6 mm (77%)), and decreasing hillslope discharge (27.8 mm (55%); 18.2 mm (36%); 0 mm (0%)) for the low storage, base case calibrated value and high storage cases. Evaporation and changes in subsurface storage between events made up the remaining 14 – 23%. For simulations with no subsurface storage, the threshold was 6.45 mm and the slope 0.70 mm / mm, while for simulations with no bedrock leakage, the threshold was 6.45 mm and the slope 0.95 mm / mm.

The threshold for flow increased nonlinearly with increased $c_{bedrock}$ and V_{pool}, while the slope of the excess precipitation / discharge line decreased linearly with both (Table 3.5). Analysis of varying the two parameters shows that the threshold (slope) varied from a maximum (minimum) of 59.2 mm (0.0002 mm / mm) for the highest values of bedrock leakage coefficient and subsurface storage volumes to a minimum (maximum) of 6.45 mm (0.91 mm / mm) for the simulations with no leakage or subsurface storage (Table 3.5). While both parameters impacted storm response, the bedrock leakage coefficient seemed to have more impact on the slope, while the subsurface storage volume had more impact on the threshold. A similar pattern of bedrock leakage coefficient and subsurface pool storage influence on the precipitation / discharge relationship was seen in the other events, with different antecedent moisture conditions.

3.3.3.4. Alternative hypotheses

Finally, a set of simulations was run to determine if there was a third alternative for the thresholds beyond fill and spill and soil moisture deficit (hypothesis 4). With the potential evaporation rate, bedrock permeability and subsurface storage volumes set to zero, the virtual experiment hyetographs were run. For event 5, there was flow for all simulations, and the projected threshold was 0 mm. The slope of the precipitation / discharge line was 1.00 mm discharge / mm precipitation. A similar lack of threshold and identical slope was seen for the other events. These results suggest that the threshold was due entirely to the fill and spill and soil moisture deficit factors, and the slope approaches unity as these factors are reduced.

3.3.4. Thresholds at the watershed scale:

The newfound relationship between the soil moisture deficit and fill and spill factors and the precipitation / discharge threshold was tested against two long term data records. The first was from 3.8 ha watershed nearby the modeled hillslope. Analysis of the precipitation discharge relationship at this catchment, which has similar geology (bedrock leakage and subsurface storage) and climatic conditions (antecedent drainage time and PET) should provide a test for the fill and spill and soil moisture deficit correlated threshold relationship seen in the virtual experiments. The second test was for a set of watersheds at the H.J. Andrews experimental forest (HJA), ranging from 8-101 ha. The HJA will be a stronger test of the soil moisture deficit factors due to the higher antecedent drainage times for events towards the end of the summer season.

3.3.4.1. M8 Catchment, Maimai, New Zealand

Upstream of the instrumented hillslope described by (Graham et al., this issue) and used for numerical modeling in this paper is the first order, 3.8 ha M8 watershed. The watershed was gauged for nearly 30 years (1974 – 2003) and a subset of the data record (1985-1986) was used for watershed scale threshold analysis. Evapotranspiration was estimated during this period by estimation using a temperature index model (McDonnell et al., submitted).

For the analysis, the two year M8 hyetograph was split into 140 storm events. A storm event was defined as a precipitation event greater than 1 mm, preceded by 24 hours of less than 1 mm rain. Storm runoff was defined as the rise in stream discharge above baseflow (streamflow at the initiation of the rain event) from the initiation of the event until the beginning of the next event. Total storm precipitation during the two years of monitoring was 4437.5 mm, with individual event precipitation ranging from 1.1 to 220.1 mm, with an average storm size of 30.7 mm (Figure 3.9). Total storm runoff for the monitored time period was 1916.5 mm, where individual event storm runoff ranging from 0 to 109.5 mm, with an average of 13.7 mm. During the monitoring, the average time between storms was 3.2 days and the maximum was 17.2 days. Estimated PET averaged 2.4 mm / day, with a maximum 60 minute rate of 9.3 mm / day. The product of the antecedent drainage time and estimated PET ranged from 0 to 45.6 mm, with an average of 7.1 mm. Predicted discharge was calculated for each storm using 4 methods: the soil moisture deficit dependent thresholds, the bulk annual threshold, the annual storm runoff ratio (43%), and the average storm runoff ratio (64%).

When applying the model derived soil moisture deficit threshold relationship, storm flow was predicted at the M8 watershed using equations 12 and 13, assuming the calibrated fill and spill factors ($c_{bedrock}$, V_{pool}) from the hillslope were constant between sites. The root mean square error (RMSE) of the measured vs. modeled discharge was 6.2 mm, and total storm runoff of 1837 mm for the 141 monitored storms over 2 years (96% of the measured storm runoff). Using measured slope (0.59) and threshold (8.5 mm) from the lumped discharge / precipitation record resulted in a RMSE of 6.5 mm, and total storm runoff of 2030 mm (106%). Using the annual measured storm runoff ratio to predict storm discharge led to a poorer fit, with a RMSE of 9.0 mm, though by definition, the total storm runoff was 1916 mm, or 100%. Using the average storm runoff ratio yielded a higher RMSE (15.5 mm), and low total storm runoff (1109 mm, or 58%).

This site is assumed to have similar geology and environmental characteristics as the Woods and Rowe (1996) hillslope where MaiModel was calibrated. Using the precipitation discharge relationship developed from the model, we were better able to predict both the individual storm discharge as well as the annual water balance than with any of the methods using analysis of the catchment discharge characteristics. While the annual threshold analysis performed nearly as well, if this were an ungauged catchment, the data needed to develop the threshold relationship would not be available.

3.3.4.2. H. J. Andrews, Oregon, USA

Further analysis of the watershed scale precipitation discharge threshold relationship was performed using the precipitation discharge record at the H. J. Andrews Experimental Forest (HJA) in western Oregon, USA. At the HJA, continuous discharge and precipitation records have been maintained at ten watersheds, ranging from 9 to over 100 ha for up to 50 years, though we do not have evaporation estimates for the duration. Of the ten gauged catchments, five (WS1, WS2, WS3, WS9 and WS10, from 9 – 101 ha) are at low enough elevation that their annual hydrographs are dominated by rainfall, rather than seasonal snowmelt. WS1-3 have been gauged since 1958, while WS9-10 have been gauged since 1969. 2246 (WS1-3) and 1718 (WS9-10) rainfall events were extracted from these records (storms begin when 1 mm rain falls, and end after 24 hours of no precipitation), ranging up to 731 mm. Storm runoff and total storm precipitation was extracted from the discharge record as per the M8 watershed procedure. A plot of total storm precipitation vs. discharge shows little evidence of a threshold (Figure 3.10a). However, if the storms are binned according to the antecedent drainage time, a threshold appears to exist for events with larger than five days, and the threshold increases with increasing antecedent drainage time (Figure 3.10b & c). The threshold appears to be consistent between the gauged catchments (~50 mm for greater than 5 days drainage, ~80 for greater than 10 days drainage), despite the wide range of catchment sizes.

While we were unable to determine the exact functional relationship between the soil moisture deficit factors (PET and antecedent drainage time) without estimates of the evaporation rates, this analysis shows that there is a clear influence of antecedent drainage time on the threshold. The small apparent threshold for the short antecedent drainage events indicates that the fill and spill factors are relatively insignificant.

3.4. Discussion

We used the dominant processes concept of Grayson and Blöschl (2000) to construct a simple, reservoir based numerical model based on the Maimai hillslope. The model, with simple unsaturated storage and flow conceptualization, was able to generally reproduce observed hydrometric and tracer behavior. The calibrated model was able to reproduce the 40 day hydrograph, as well as each individual storm. Additionally, the model was able to reproduce breakthrough of a line tracer application 35 m upslope – one characterized by both rapid initial breakthrough and extended hillslope discharge. The model was also able to capture the precipitation / discharge threshold relationship observed in the data record. The model was then used to determine the relative importance of fill and spill and soil moisture deficit factors on the threshold relationship. Below we highlight some of the issues associated with the calibration of the model, the results from the virtual experiments, and the application of the new understanding of threshold controls at the watershed scale, both at a nearby first order watershed and at a different set of watersheds ranging from 8.5-101 ha.

3.4.1. On the value of data for model construction and testing

The model objective criteria that we used did not have equal strength in either limiting the range of individual parameters, or in reducing the number of behavioral parameter sets. In general, the model criteria that were effective in reducing the range of individual parameters were also effective in reducing the number of behavioral parameter sets. Of notable exception was the cumulative tracer breakthrough criterion. While able to reject parameter sets (only 3% of the parameter sets acceptably met the criterion), it did little to reduce the range of acceptable individual parameters, except for the bedrock leakage coefficient ($c_{bedrock}$). While the ranges of most parameters were reduced through calibration, soil hydraulic conductivity was not, with acceptable models sampling from 94% of the original parameter space. The other parameters were well identified through calibration, with a reduction of the original parameter space by 56-86% (Table 2.1).

3.4.1.1. Modeled hydrograph

The six storm hydrograph criteria were responsible for reduction in both the range of the individual parameters and the total number of acceptable parameter sets (Figure 3.3). As in previous studies, different subsections of the hydrograph provided different amounts of power in parameter identifiability (i.e. (Seibert and McDonnell, 2002; Son and Sivapalan, 2007; Tetzlaff et al., 2008) While 2.9% of the parameter sets met the 40 day criterion, only 0.4% of the parameter sets resulted in simulations that met the hydrograph criteria for all five of the events. Seibert and McDonnell (2002) found that a storm event with the largest peak precipitation rate and discharge served as the most stringent criterion in their calibration of a similar reservoir model. In our case, the highest peak of precipitation and discharge occurred in storm B4, which was a relatively weak criterion.

B2, the event with the lowest peak discharge, and longest duration, was the most effective in both narrowing the parameter ranges and rejecting parameter sets. Whereas the other storms were relatively simple, with a single peaked hyetograph and hydrograph, B2 was more complex, with a double peaked hyetograph and hydrograph. B2 was especially effective in reducing the parameter space for the lateral subsurface storage and bedrock leakage coefficient variables. This sensitivity was likely due to the complex filling and draining of subsurface storage, a factor that was masked in the higher shorter, single peaked events, where the subsurface storage is filled early in the event, then monotonically drained. The prolonged nature of the B2 event, along with the refilling during the second peak, required a more precise definition of the subsurface storage processes. This suggests that it is not the size of the event, but perhaps the complexity that is important for model calibration. Breaking up the calibration hydrograph into 5 distinct time periods, centered on the significant rain events, proved to be a strong tool for both parameter identifiability and parameter set rejection.

3.4.1.2. Modeled tracer breakthrough

The modeled tracer breakthrough served as another source for parameter identification and parameter set rejection. Other researchers have shown the importance of using tracers (such as isotopic signatures of rainfall) in addition to hydrometric data for model calibration. (Fenicia et al., 2008; Son and Sivapalan, 2007; Soulsby and Dunn, 2003; Vache and McDonnell, 2006) Tracers are attractive as model objective criteria because tracer and pressure response to precipitation is often quite different (i.e. the rapid catchment response dominated by pre-event water Sklash and Farvolden, 1979). Tracer breakthroughs also serve to integrate hillslope scale response, in contrast to point measurements of hydraulic conductivity, water table height, soil moisture status or other similar objective criteria. While isotopic tracers and mean residence times of tracers have been used for model calibration, the use of an applied, chemical tracer is relatively rare for model calibration (although Weiler and McDonnell (2007) successfully modeled the Brammer tracer injection with a macropore based conceptual model of the Maimai hillslope).

While the temporal and spatial patterns of tracer breakthrough were not stringent criteria in the MaiModel calibration (eliminating only 48 and 15% of the parameter sets, respectively), the cumulative tracer breakthrough eliminated 97% of the parameter sets, and 66% of the simulations that were deemed behavioral for all storms. The calibration runs that modeled all of the sub-hydrographs and did not match the measured tracer breakthrough had a modeled cumulative tracer breakthrough ranging from 8-28%, compared to a measured value of 14%. Of the simulations that had acceptable fits for the hydrographs but missed the cumulative tracer breakthrough, 23% were below the acceptable limits, and 77% were greater. 54% were more than twice the acceptable range from the measured value. This wide range of modeled tracer flux for models that acceptably fit the hydrograph demonstrates the importance of measurements of both particle and pressure response at the hillslope scale for model calibration and validation.

While the cumulative tracer breakthrough was effective in reducing the total number of behavioral parameter sets, it did little to reduce the ranges of the individual parameters, with the exception of $c_{bedrock}$. Two possible explanations of the relative weakness of the tracer breakthrough on the parameter ranges are 1) the tracer breakthrough is due to a combination of parameters, or 2) the cumulative tracer breakthrough is too weak a test, and a time series of tracer breakthrough is needed. Further analysis of the tracer breakthrough against the individual parameters suggests that the first option is more likely. The cumulative tracer breakthrough was compared with the products of each pair of calibrated parameters (10 pairs in total). The cumulative tracer breakthrough was strongly constrained by the product of the bedrock leakage coefficient and the subsurface storm volume, with a reduction of 94% of the widest possible range of the product (Figure 3.11). This suggests that it is both the subsurface storage volume and the rate of drainage that controls the cumulative tracer breakthrough, more than each parameter individually that is important. The cumulative tracer breakthrough was not dependent on any other individual parameter, or product of parameters.

3.4.1.3. Soil hydraulic conductivity

Of the five calibrated parameters, all but the soil saturated hydraulic conductivity (k_{soil}) were significantly better defined through calibration. Of the nine calibration criteria, the number of behavioral parameter sets that matched each criteria was somewhat correlated to the reduction in the parameter space for each criteria ($0.48 < R^2 < 0.67$). While the range of k_{soil} was reduced 6% from the initial parameter range specified, (1,000% of the maximum measured hydraulic conductivity), this 6% is due more likely through chance than an actual narrowing of the possible parameter set, as acceptable parameter sets were evenly distributed over the calibration range. It is unclear whether the small reduction in the k_{soil} range is a validation of the model assumptions, especially that vertical percolation through the soil profile is relatively unimportant in the whole hillslope scale behavior, or a result of these assumptions being codified into the model.

3.4.2. Improved understanding of thresholds at the hillslope scale

The calibrated model was able to reproduce the precipitation discharge threshold relationship seen at the Maimai hillslope trench. From analysis of measured hillslope discharge by Woods and Rowe (1996) and Brammer (1996), an average threshold of approximately 19 mm was necessary for flow at the hillslope at Maimai (Figure 3.9). This threshold was close to the modeled threshold (17.7 mm) for the calibrated base case model. In the numerical simulations, the threshold was found to be due to both fill and spill factors (subsurface storage and bedrock leakage) as well as soil moisture deficit factors (potential evaporation rate and antecedent drainage time) (Figure 3.12). A linear relationship between thresholds and the product of the t_{drain} and PET was observed, with a slope of 0.38 mm / mm. An increase of 1 mm in the product of t_{drain} and PET prior to storm initiation yields an increase of 0.38 mm in the threshold over the range modeled. The relationship between the product of the bedrock leakage coefficient and the subsurface storage volume was positively concavely nonlinear. There appears to be an upper bound on the impact of the fill and spill factors on the threshold, while a similar bound has not been observed in the soil moisture deficit. Logically, a bound must exist for the soil moisture deficit, once evaporation depletes the entire soil profile and a storm greater than the available storage would overcome the threshold. This bound was not met by the current virtual experiments.

Additional simulations were performed to determine the precise impact of both sets of factors on the threshold. A simulation where the fill and spill factors were removed (V_{pool} and $c_{bedrock}$ were set to zero) and another where evaporation was removed (PET = 0 mm/day). These simulations showed that 11.8 mm (66%) of the threshold were due to fill and spill factors, while 6.4 mm (36%) was due to soil moisture deficit factors. Since these do not add up to the modeled 17.7 mm threshold for the base case scenario, it appears that the two processes are less than additive, with some of the effects of soil moisture deficit reducing the threshold response due to fill and spill, or vice versus. The lack of threshold after the fill and spill and soil moisture deficit mechanisms were eliminated indicates that these two are solely responsible for the simulated threshold.

The slope of the excess precipitation / discharge line was also found to be positively correlated with both fill and spill factors and PET, while not the antecedent drainage time. An increase in both the subsurface storage and bedrock leakage coefficient were shown to increase the slope, as an increase in each increased the rate of leakage into the bedrock, both directly (increased leakage coefficient = increased leakage rate) and indirectly (increased storage = increased driver on leakage and increased late time storage and leakage). An increase in the PET increased the slope, as rainfall stored in the soil profile was lost to evaporation during and after the storm. With fill and spill and soil moisture deficit removed, the slope was unity, indicating these are the only factors affecting the slope in the numeric model. 94% of the reduction in the slope was due to fill and spill mechanisms, while 6% of the reduction is due to the potential evaporation rate. The small impact of the PET on the slope is due to the limited time that PET can impact the discharge after the threshold is reached, as hillslope drainage lasted less than 4 days for all simulations.

3.4.3. Improved understanding of thresholds at the watershed scale

Our macroscale hillslope model has shown two causal mechanisms for the storm precipitation – discharge relationship seen at the hillslope and small catchment scale. This relationship depends on both climatic (event spacing and evaporative losses) and geologic (bedrock permeability and subsurface storage) factors. The geologic factors are difficult to determine, with bedrock permeability difficult to measure, and subsurface storage depending on the dominant lateral subsurface flow processes and the bedrock topography, two difficult to determine components. The climatic factors, however, are often available when a long term data set is present. While long term evaporation records remain uncommon, new analysis of long term precipitation records may provide a way forward towards better prediction of catchment storm response.

The soil moisture deficit influence on the precipitation discharge threshold and slope were previously suggested in the analysis of the long term precipitation discharge record of the instrumented hillslope at Panola, Georgia. (Tromp-van Meerveld and McDonnell, 2006a) At Panola, it appeared that storms where the soil volumetric water content at 70 cm depth was less than 40% prior to the event had a higher threshold for flow than those that were relatively wetter. However, the data record at Panola had too few storms with sufficiently dry antecedent moisture conditions to determine the precise relationship between antecedent moisture and thresholds. At Minamitani, Japan, the threshold for flow at an instrumented hillslope and nearby second order catchment was shown to be somewhat dependent on the flow rate at the initiation of the event. (Tani, 1997) As at Panola, with the limited number of events above and below the threshold for each initial flow rate, the precise nature of this dependency is unclear. In fact, none of the instrumented hillslopes we know of have a sufficient data record with enough storms above and below the threshold to determine the precise relationship between the threshold and antecedent moisture conditions. While it remains difficult and expensive to maintain gauging for a sufficiently long duration, especially at instrumented hillslopes, numerical modeling can serve to generate new knowledge about hillslope and small catchment processes. This newfound relationship between antecedent drainage time and the precipitation discharge threshold can now be explored at other sites.

Analysis using the predictions for event storm runoff at the nearby M8 catchment, based on soil moisture deficit factors, was shown to better predict whole storm discharge than the annual threshold or runoff ratio analyses. The root mean square error was minimized and the difference between the measured and modeled annual storm runoff using the soil moisture deficit method of discharge prediction when compared to predictions made using the bulk annual threshold, annual runoff ratio and average runoff ratio methods. These predictions were made over a range of storms with different average and maximum rainfall intensities, durations and precipitation patterns, yet the storm runoff was very well predicted based on two simple factors revealed through the numerical modeling. Due to the short data record at the instrumented hillslope, a similar analysis of hillslope threshold dependence on antecedent drainage time was not possible.

Additional analysis at five small research watersheds in western Oregon (9 - 101 ha) showed a dependence on antecedent drainage time for the threshold. While evaporation estimates were not available for the duration of the 50 year data record, events with long antecedent drainage were shown to exhibit a much higher threshold for flow. This threshold appears to be quite high for events with longer than 9 days of antecedent drainage (~ 80 mm), for a series of watersheds that are very responsive to rainfall (annual storm runoff ratios approach 38%) (McGuire et al., 2005).

However, at the catchments at the H J Andrews and the M8 catchment, there is little evidence for a threshold for flow for events with short (< 5 days) antecedent drainage time, perhaps due to minimal effect of bedrock leakage. At the Maimai and Panola hillslopes, bedrock leakage is a sink and not recovered at the hillslope. In the virtual experiments, using the calibrated model, a threshold of 8 mm was predicted for events with only one day antecedent drainage. At the M8 and H J Andrews catchments, however, bedrock leakage is likely recovered at the watershed outlet, though some evidence of deep seepage at the HJA does exist. (Waichler et al., 2005) At the catchment scale, the fill and spill mechanism should then not have a large impact on the precipitation discharge threshold. Therefore, during events where the soil has not had a chance to dry due to evaporative losses, a small threshold would be expected at these catchments, as seen in this analysis.

This functional dependence of the threshold on fill and spill and soil moisture deficit factors may be a means for prediction of flow at ungauged hillslopes and basins. At a site where the physical properties are similar to either Maimai, or some basin where the geologic dependent threshold and slope has been determined, the base case threshold can be determined, and the effects of the climatic factors would be determined from the storm spacing and evaporative demand. The geologic precipitation discharge threshold and slope can be determined by analysis of the system response to precipitation at the lower extreme of PET and storm spacing. The analyses of the HJA watersheds suggest that the antecedent drainage dependence of the thresholds suggest that this relationship may apply to other steep forested hillslopes and catchments. Special attention needs to be placed on locations with different geology, catchment geometry and dominant flow processes.

3.4.4. Constraints in numerical modeling and opportunities for future work

Our numerical model was built to determine the flow response at a site where lateral subsurface flow dominates, and a clear threshold for flow exists. The model was successfully applied to predict lateral subsurface flow at the site, and determine the sources of the observed threshold. The model was then validated with new analysis of the data record at a nearby site. This work suggests further work using carefully calibrated models to answer other, pressing questions on the controls on hillslope discharge, such as: What are the controls on baseflow generation? What are the controls on water aging? When can we predict the dominant subsurface flow mechanisms (i.e.

lateral vs. vertical flow; preferential vs. matrix flow)? The numerical model used in this paper is not well suited for all of these questions, though the dominant processes concept is. For example, there is no bedrock flow incorporated in MaiModel, which makes it poorly suited to determine the controls on baseflow, which is known to be highly dependent on subsurface flow through the bedrock. At the Maimai hillslope, bedrock flowpaths are not believed to greatly affect hillslope discharge, due to the steep slopes and consolidated nature of the bedrock. However, using a simple conceptual model allows for the flexibility for model evolution dictated by the field conditions (e.g. Fenicia et al., 2008). The addition of bedrock flow would be a relatively simple thing, if the field processes demanded it. Careful model construction including the dominant processes under consideration, along with stringent calibration and testing, can lead to new understanding of flow processes.

3.5. Conclusions

Graham et al. (this issue) developed a new perceptual model of hillslope subsurface flow processes at a well studied field site. We determined that lateral subsurface flow is dominated by flow in a well connected preferential flow network at the interface between the soil profile and permeable bedrock. This paper used this new perceptual model as the basis for a numerical model designed to model flow and transport based on these dominant processes. The model was able to reproduce both hydrometric and tracer data, using few (5) tunable parameters. A series of virtual experiments aimed at revealing the controls on the threshold response of hillslope discharge to precipitation were performed using the numerical model. We found that both fill and spill (geologic, including bedrock permeability and storage) and soil moisture deficit (climatic, including storm spacing and potential evapotranspiration rates) factors influenced threshold magnitude. While the climatic controls were shown to have a large potential impact on flow dynamics, in a climate like that of the study hillslope, where storm spacing was short (average time between storms = 3 days) and the PET demand was low (<6 mm / day), the geologic controls dominated (66% of the threshold and 94% of the slope of the excess precipitation / discharge relationship were determined by the geologic components). The relationship between the climatic factors and the precipitation discharge threshold and slope were applied to a nearby catchment and demonstrated to better predict storm discharge than either the annual runoff ratio or the bulk threshold relationship.

3.6. Acknowledgements

This work was funded through an NSF Ecosystem Informatics internship. We would like to thank Jurriaan Spaaks, Ciaran Harmon, Luisa Hopp, Taka Sayama and Kellie Vache for helpful comments on the modeling strategy. Special thanks go to Ross Woods, Lindsay Rowe and Dean Brammer for the use of their experimental data for model development and calibration.

3.7. References

Baker, T.M. and Hawke, R.M., 2007. The effects of land application of farm dairy effluent on groundwater quality – West Coast. Journal of Hydrology (New Zealand), 45.

Brammer, D., 1996. Hillslope hydrology in a small forested catchment, Maimai, New Zealand. M.S. Thesis, State University of New York College of Environmental Science and Forestry, Syracuse, 153 pp.

Brooks, R.H. and Corey, A.T., 1964. Hydraulic properties of porous media. Hydrology Paper 3, Colorado State University, Fort Collins, CO.

Buttle, J.M., Dillon, P.J. and Eerkes, G.R., 2004. Hydrologic coupling of slopes, riparian zones and streams: an example from the Canadian Shield. Journal of Hydrology, 287(1-4): 161-177.

Carsel, R.F. and Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. Water Resources Research, 24(5): Pages: 755-769.

Dooge, J.C.I., 1986. Looking for Hydrologic Laws. Water Resources Research, 22: 46S-58S.

Fenicia, F., McDonnell, J.J. and Savenije, H.H.G., 2008. Learning from model improvement: On the contribution of complementary data to process understanding. Water Resources Research, 44.

Freeze, R.A. and Cherry, J.A., 1979. Groundwater. Prentice-Hall, Englewood Cliffs, NJ, 604 pp.

Graham, C.B., McDonnell, J.J. and Woods, R., this issue. Hillslope threshold response to storm rainfall: (1) A field based forensic approach. Journal of Hydrology.

Grayson, R.B. and Blöschl, G., 2000. Summary of pattern comparison and concluding remarks. In: R.B. Grayson and G. Blöschl (Editors), In Spatial Patterns in Catchment Hydrology: Observations and Modeling. Cambridge University Press, Cambridge, UK, pp. 355–367.

Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1998. Towards improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.

McDonnell, J.J., 1990. A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11): 2821-2832.

McDonnell, J.J., Freer, J. and Vache, K., 2009. A reference dataset for event-based watershed model development and testing: The Maimai catchment. Water Resources Research, in review.

McDonnell, J.J., Freer, J. and Vache, K.B., in review. A reference dataset for event-based watershed model development and testing: The Maimai catchment. Water Resources Research.

McDonnell, J.J. et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, 43.

McGlynn, B.L., McDonnell, J.J. and Brammer, D.D., 2002. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. Journal of Hydrology, 257: 1-26.

McGuire, K.J. et al., 2005. The role of topography on catchment-scale water residence time. Water Resources Research, 41.

Mosley, M.P., 1979. Streamflow generation in a forested watershed. Water Resources Research, 15: 795-806.

Mosley, M.P., 1982. Subsurface Flow Velocities Through Selected Forest Soils, South Island, New Zealand. Journal of Hydrology, 55: 65-92.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, I, A discussion of principles. Journal of Hydrology, 10: 282-290.

Pearce, A.J., Stewart, M.K. and Sklash, M.G., 1986. Storm runoff generation in humid headwater catchments: 1. Where does the water come from? Water Resources Research, 22: 1263-1272.

Peters, D.L., Buttle, J.M., Taylor, C.H. and LaZerte, B.D., 1995. Runoff production in a forested, shallow soil, Canadian Shield basin. Water Resources Research, 31(5): 1291-1304.

Putuhena, W.M. and Cordery, I., 2000. Some hydrological effects of changing forest cover from eucalypts to Pinus radiata. Agricultural and Forest Meteorology, 100(1): 59-72.

Quinn, P.F., Beven, K.J., Chevallier, P. and Planchon, O., 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes, 5: 59-79.

Seibert, J., 1997. Estimation of parameter uncertainty in the HBV model. Nordic Hydrology, 28: 4-5.

Seibert, J. and McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. Water Resources Research, 38(11).

Sklash, M.G. and Farvolden, R.N., 1979. The role of groundwater in storm runoff. Journal of Hydrology, 43: 45-65.

Sklash, M.G., Stewart, M.K. and Pearce, A.J., 1986. Storm runoff generation in humid headwater catchments: 2. a case study of hillslope and low-order stream response. Water Resources Research, 22(8): 1273-1282.

Son, K. and Sivapalan, M., 2007. Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water Resources Research, 43.

Soulsby, C. and Dunn, S.M., 2003. Towards integrating tracer studies in conceptual rainfall-runoff models: recent insights from a sub-arctic catchment in the Cairngorm Mountains, Scotland. Hydrological Processes, 17: 403-416.

Spence, C. and Woo, M.-K., 2002. Hydrology of subartic Canadian shield: bedrock upland. Journal of Hydrology, 262: 111-127.

Tani, M., 1997. Runoff generation processes estimated from hydrological observations on a steep forested hillslope with a thin soil layer. Journal of Hydrology, 200: 84-109.

Tetzlaff, D., Uhlenbrook, S., Eppert, S. and Soulsby, C., 2008. Does the incorporation of process conceptualization and tracer data improve the structure and performance of a simple rainfall-runoff model in a Scottish mesoscale catchment? Hydrological Processes.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006a. Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope. Water Resources Research, 42.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006b. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resources Research, 42.

Uchida, T., Tromp-van Meerveld, H.J. and McDonnell, J.J., 2005. The role of lateral pipe flow in hillslope runoff response: an intercomparison of non-linear hillslope response. Journal of Hydrology, 311: 117-133.

Vache, K.B. and McDonnell, J.J., 2006. A process-based rejectionist framework for evaluating catchment runoff model structure. Water Resources Research, 42.

Waichler, S.R., Wemple, B.C. and Wigmosta, M.S., 2005. Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. Hydrological Processes, 19: 3177-3199.

Weiler, M. and McDonnell, J.J., 2007. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes. Water Resources Research, 43.

Weiler, M., McDonnell, J.J., Tromp-van Meerveld, I. and Uchida, T., 2006. Subsurface stormflow. Encyclopedia of Hydrological Sciences, 3.

Whipkey, R.Z., 1965. Subsurface storm flow from forested slopes. Bulletin of the International Association of Scientific Hydrology, 2: 74-85.

Woods, R. and Rowe, L., 1996. The changing spatial variability of subsurface flow across a hillside. Journal of Hydrology (New Zealand), 35(1): 51-86.

Woods, R. and Rowe, L., 1997. Reply to "Comment on "The changing spatial variability of subsurface flow across a hillside" by Ross Woods and Lindsay Rowe". Journal of Hydrology (New Zealand), 36(1): 51-86.

3.8. Tables

Table 3.1 Calibration parameter ranges and sources. V_{pool} was not measured in field and was constrained by a pre-calibration sensitivity analysis.

Parameter	Range	Source
$c_{bedrock}$	0 - 0.0284 1/s	10,000% maximum observed in field (Graham et al., this issue)
k_{soil}	0 - 3 m/hr	10,000% maximum observed in field (McDonnell, 1990)
k_{LSS}	0-30 m/hr	Range observed in field (Graham et al., this issue)
V_{pool}	0 - 0.1 cm	Sensitivity analysis
f_{active}	0 - 100%	Spans range of field measured porosity (McDonnell, 1990)

Table 3.2 Parameter range reduction and parameter set rejection due to calibration using each objective criterion. The total number of calibration simulations was 10,000. Initial ranges of parameters are listed in Table 3.1.

	$c_{bedrock}$	k_{soil}	k_{LSS}	V_{pool}	f_{active}	Criteria	Behavioral parameter sets
Initial	100	100	100	100	100		
40 day discharge	58	2	51	45	62	E > 0.8	294
Storm 1 discharge	72	2	47	2	63	E > 0.8	260
Storm 2 discharge	75	6	56	83	68	E > 0.8	102
Storm 3 discharge	60	2	43	24	47	E > 0.8	416
Storm 4 discharge	56	2	47	18	58	E > 0.8	402
Storm 5 discharge	56	2	30	34	47	E > 0.8	342
Total cumulative tracer breakthrough	44	1	5	6	0	11.5% ≤ T ≤ 16.5%	4,827
Tracer breakthrough temporal	0	1	1	1	1	$R^2 > 0.8$	1,462
Tracer breakthrough spatial	0	0	1	0	0	$R^2 > 0.8$	300
Final	86	6	56	85	78		13

Table 3.3 Storm discharge (mm) and calculated threshold (mm) and excess precipitation slope (mm/mm) for calibrated model and events V1-V10.

	Time Between Storms (days)									
Event Precipitation (mm)	10 (V1)	1 (V2)	1 (V3)	3 (V4)	3 (V5)	5 (V6)	7 (V7)	14 (V8)	21 (V9)	28 (V10)
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
25	0	5	8	4	4	1	0	0	0	0
38	3	15	15	11	10	8	5	0	0	0
46	5	18	18	13	13	10	8	1	0	0
51	6	18	18	14	14	11	8	1	0	0
56	9	21	21	17	17	14	11	3	1	0
76	17	31	31	26	26	23	20	12	6	3
101	29	41	41	37	37	34	31	23	16	10
253	98	110	110	106	106	103	100	91	85	78
506	212	225	225	220	220	218	215	206	200	193
Threshold	35	9	7	18	18	24	30	48	61	75
Slope	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45

Table 3.4 Threshold (mm) and excess precipitation slope (mm/mm) for each antecedent drainage time and PET. For the highest PET modeled and the second two longest antecedent drainage times the discharge was not non-zero for enough events to determine threshold and slope values.

		PET (mm/day)									
		0	0.6	1.2	3	4.5	6	9	12	30	60
t_{drain} (days)	10	17	18	20	25	30	35	46	57	135	207
	1	6	6	6	7	8	9	13	16	55	147
	1	6	6	6	7	7	8	9	11	35	78
	3	12	12	13	15	16	18	21	25	52	114
	3	12	12	13	15	16	18	21	25	56	117
	5	14	15	16	19	21	24	31	38	78	176
	7	16	17	18	22	26	30	39	48	101	192
	14	19	21	23	32	40	48	64	76	186	240
	21	20	23	28	41	54	64	80	101	206	NaN
	28	21	26	31	50	64	75	100	137	218	NaN

		PET (mm/day)									
		0	0.6	1.2	3	4.5	6	9	12	30	60
t_{drain} (days)	10	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.33	0.14
	1	0.5	0.48	0.48	0.47	0.46	0.45	0.44	0.42	0.36	0.24
	1	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.34	0.21
	3	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.32	0.16
	3	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.32	0.17
	5	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.32	0.19
	7	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.42	0.31	0.16
	14	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.41	0.32	0.12
	21	0.5	0.48	0.48	0.47	0.46	0.45	0.43	0.40	0.31	NaN
	28	0.5	0.48	0.48	0.47	0.46	0.45	0.42	0.42	0.28	NaN

Table 3.5 Threshold (mm) and excess precipitation slope (mm/mm) for each V_{pool} / $c_{bedrock}$ combination. For eight events with high V_{pool} and $c_{bedrock}$ the discharge was not non-zero for enough events to determine threshold and slope values.

		V_{pool} (mm)									
		0	0.18	0.36	0.89	1.34	1.78	2.67	3.56	8.90	17.80
$c_{bedrock}$ (1/s)	0	6	6	6	6	6	6	6	6	6	6
	4.3E-06	6	7	7	8	9	9	10	11	18	29
	8.5E-06	6	7	8	9	10	11	13	15	25	42
	2.1E-05	6	8	9	12	14	15	18	20	34	59
	3.2E-05	6	9	10	13	15	17	19	22	37	49
	4.3E-05	6	9	11	14	16	18	21	23	40	NaN
	6.4E-05	6	10	12	16	18	19	22	25	43	NaN
	8.5E-05	6	11	13	17	19	21	24	27	45	NaN
	2.1E-04	6	13	16	20	23	25	29	35	NaN	NaN
	4.3E-04	6	16	18	23	26	29	32	NaN	NaN	NaN

		V_{pool} (mm)									
		0	0.18	0.36	0.89	1.34	1.78	2.67	3.56	8.90	17.80
$c_{bedrock}$ (1/s)	0.0	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72
	4.3E-06	0.70	0.70	0.70	0.70	0.69	0.69	0.68	0.68	0.64	0.58
	8.5E-06	0.68	0.68	0.68	0.67	0.66	0.66	0.65	0.64	0.57	0.45
	2.1E-05	0.62	0.62	0.61	0.60	0.58	0.57	0.54	0.52	0.37	0.18
	3.2E-05	0.58	0.57	0.57	0.54	0.53	0.51	0.47	0.43	0.26	0.07
	4.3E-05	0.54	0.53	0.52	0.50	0.47	0.45	0.40	0.36	0.19	NaN
	6.4E-05	0.47	0.46	0.45	0.41	0.38	0.35	0.30	0.26	0.10	NaN
	8.5E-05	0.42	0.40	0.39	0.35	0.31	0.28	0.23	0.19	0.06	NaN
	2.1E-04	0.22	0.20	0.18	0.14	0.11	0.09	0.06	0.04	NaN	NaN
	4.3E-04	0.11	0.09	0.08	0.05	0.04	0.03	0.01	NaN	NaN	NaN

3.9. Figures

Figure 3.1 Map of modeled hillslope with tracer application site 35 m upslope of lateral subsurface flow collection trench (from McGlynn et al. 2002). Trench sections T1-T20 were used for analysis.

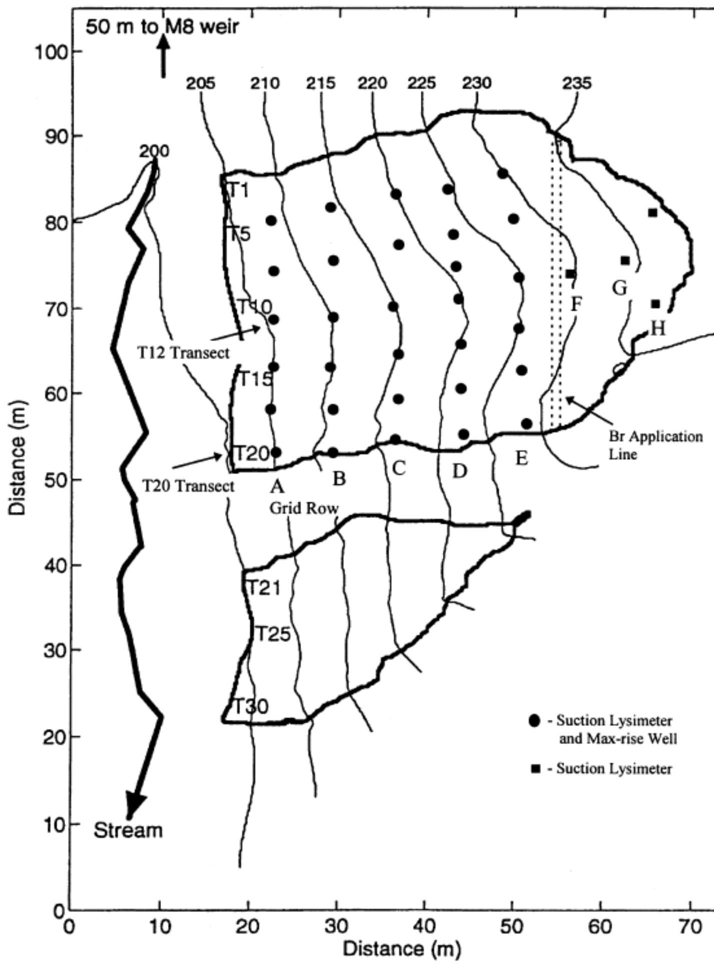


Figure 3.2 Measured (black) and modeled (grey) hydrograph with 13 simulations that matched all objective criteria for the entire data record (a), storm B2 (b), and tracer breakthrough. B2 was the most difficult storm to simulate, likely due to the complex double hydrograph.

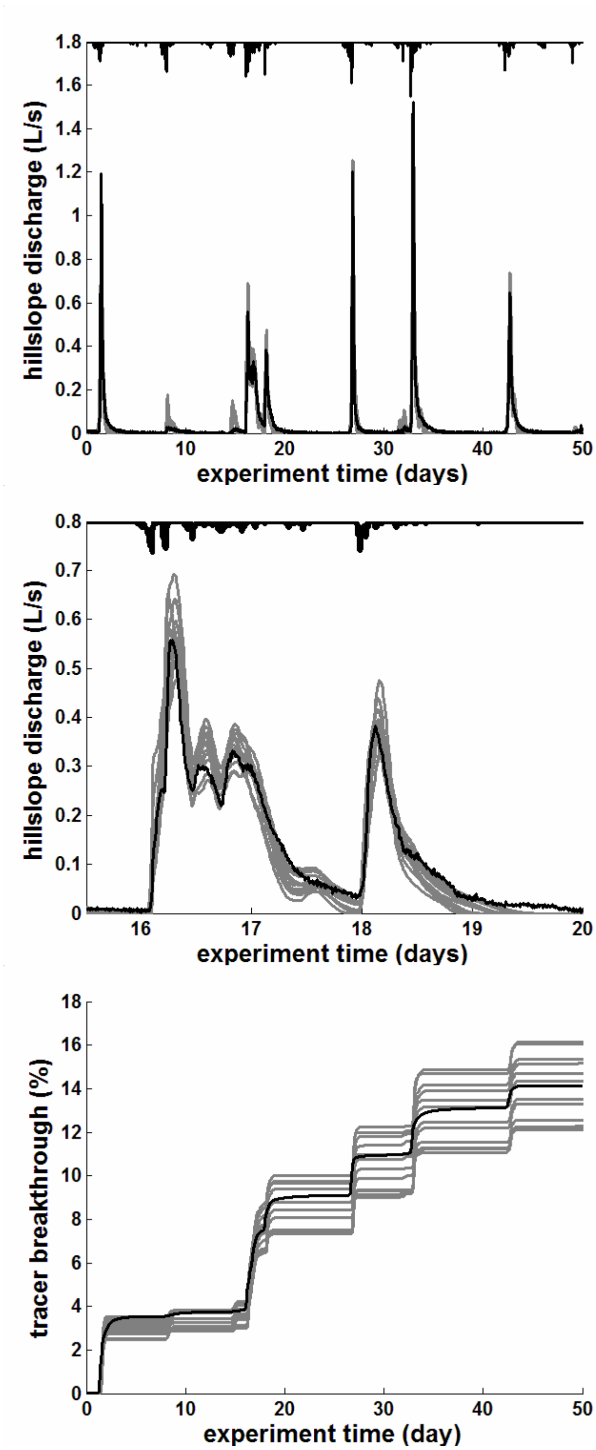


Figure 3.3 Summary of MaiModel calibration. Acceptable parameter sets (bars) and the reduction in parameter uncertainty for each model criteria. Parameters: k_{soil} (□); k_{LSS} (○); V_{pool} (●); $c_{bedrock}$ (▲); f_{active} (+). Storm B2 had the lowest number of acceptable parameter sets, and the highest reduction in the parameter space for each variable. While the temporal tracer breakthrough was not effective in rejecting parameter sets, it had some success in reducing the parameter space for $c_{bedrock}$.

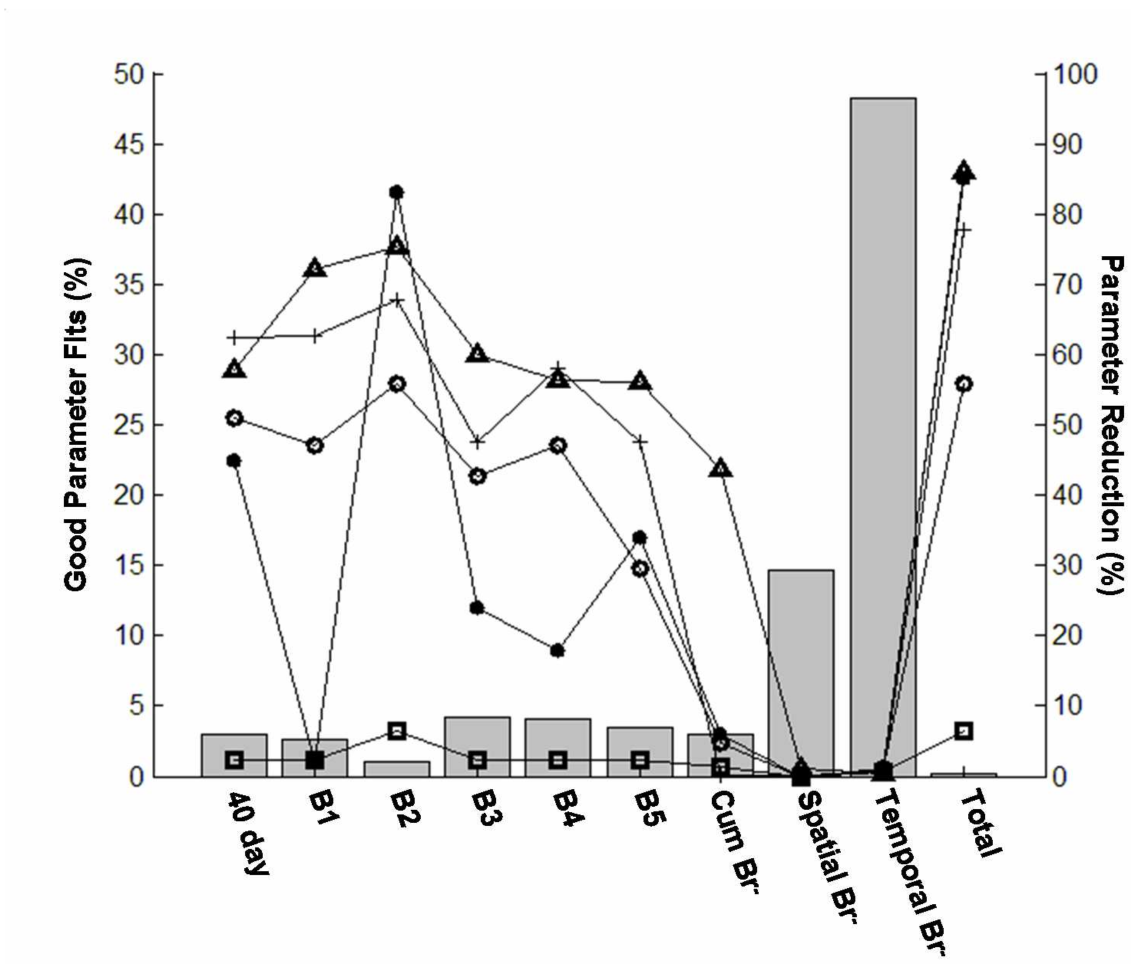


Figure 3.4 Modeled hydrographs for virtual experiment. A series of increasing duration hydrographs are applied to MaiModel with calibrated parameters and 3 days antecedent drainage (storm V5). Events with between 10 and 76 mm rainfall are simulated with scaled realizations of the hyetograph from measured storm B5. The base case event has 50.6 mm total precipitation.

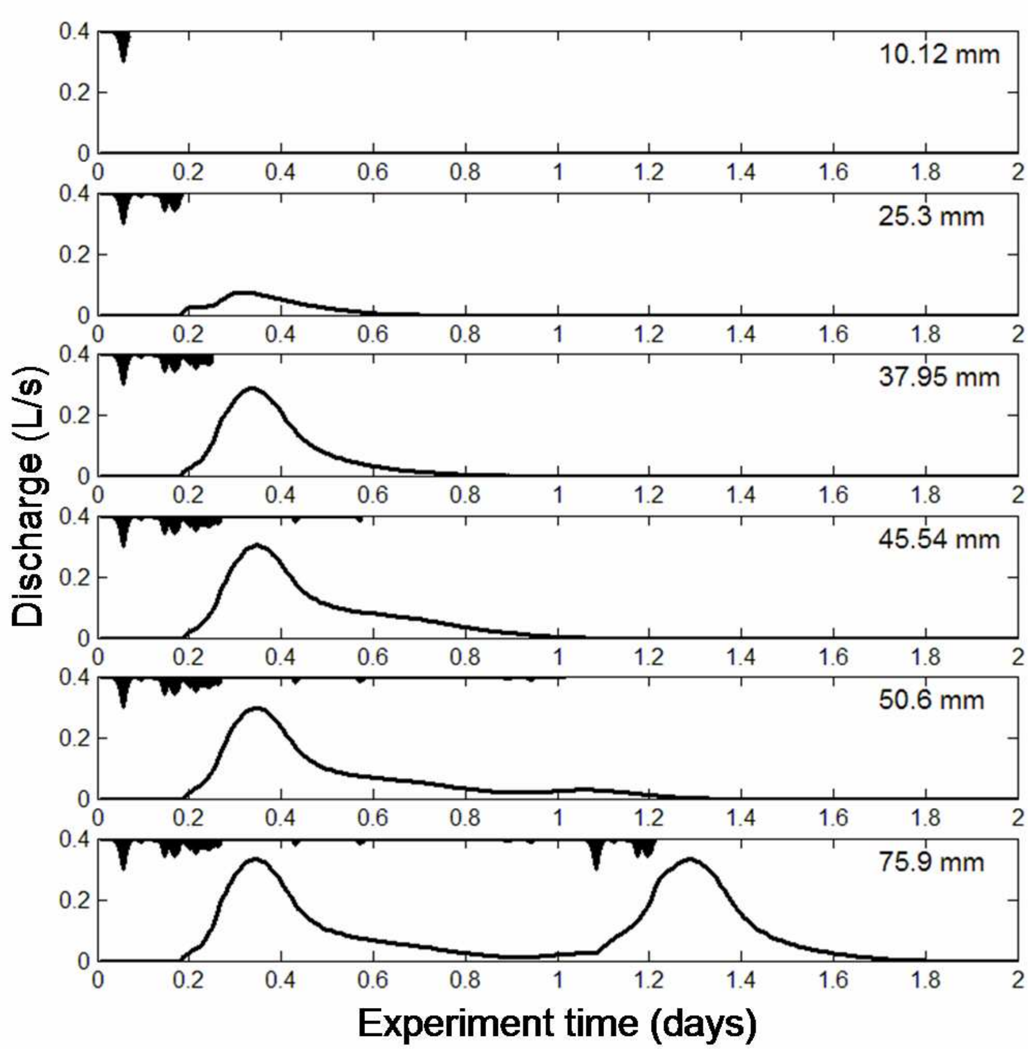


Figure 3.5 Whole storm precipitation vs. discharge for modeled events, using calibrated parameters and 3 days antecedent drainage (storm V5). The estimated threshold is 18 mm, and slope is 0.45 mm/mm (points taken from column 5, Table 3.3).

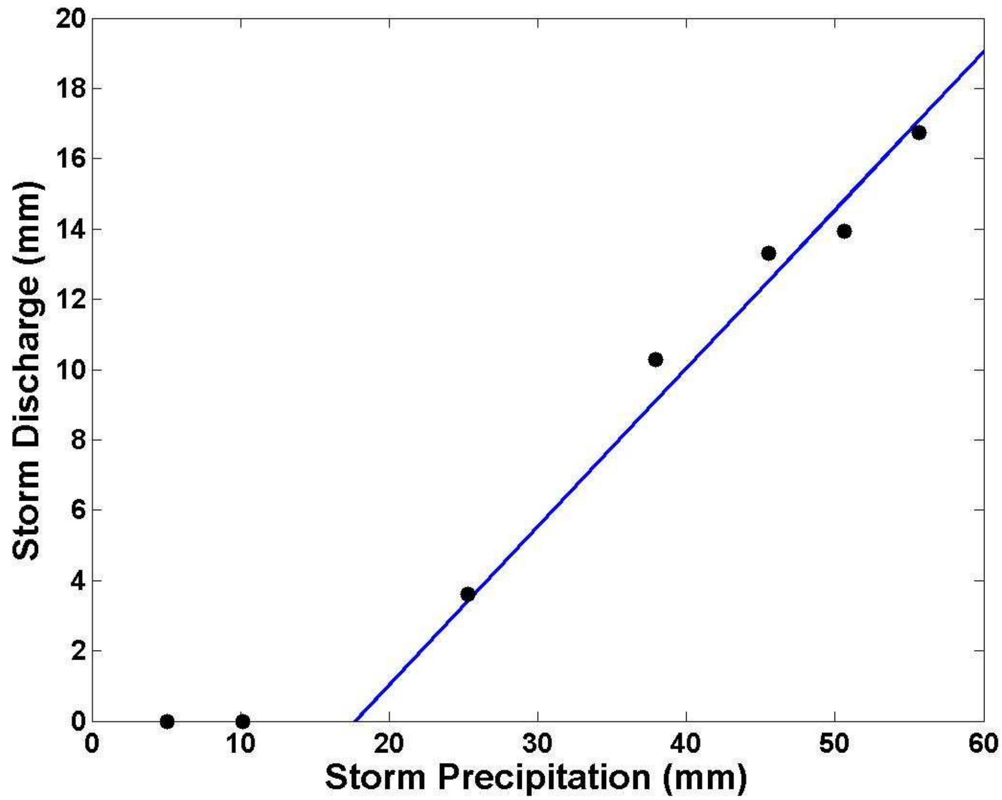


Figure 3.6 Dependence of precipitation / discharge threshold on soil moisture deficit (antecedent drainage time and PET) and fill and spill (bedrock leakage coefficient and bedrock pool storage volumes) factors.

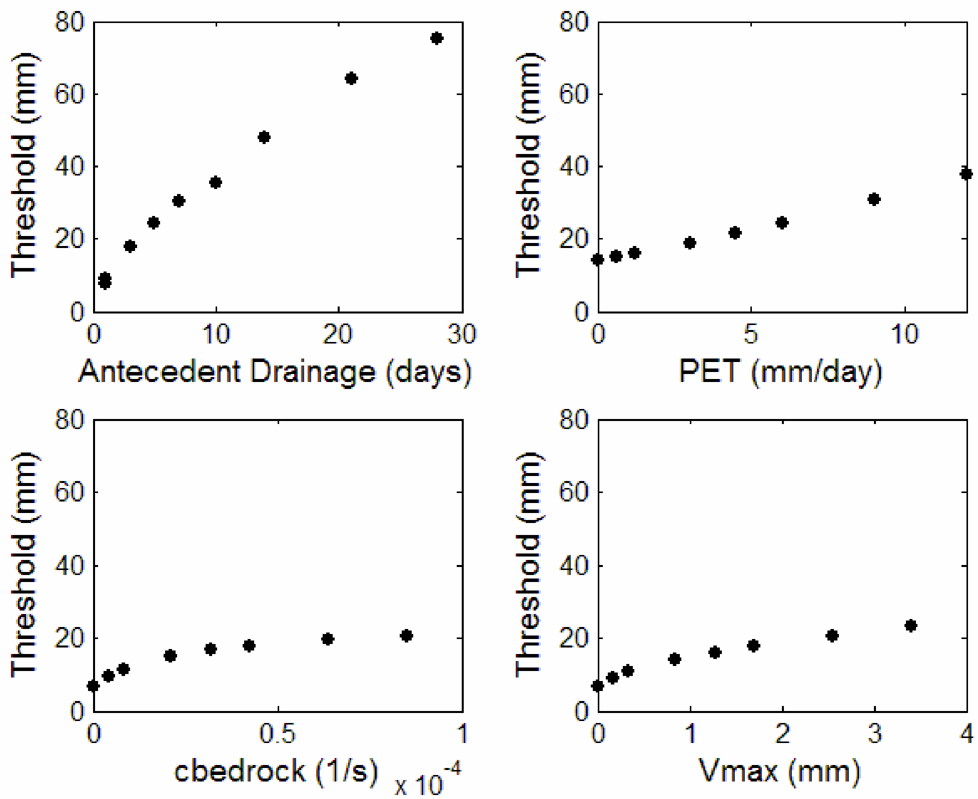


Figure 3.7 Dependence of slope of excess precipitation / discharge on soil moisture deficit (antecedent drainage time and PET) and fill and spill (bedrock leakage coefficient and bedrock pool storage volumes) factors.

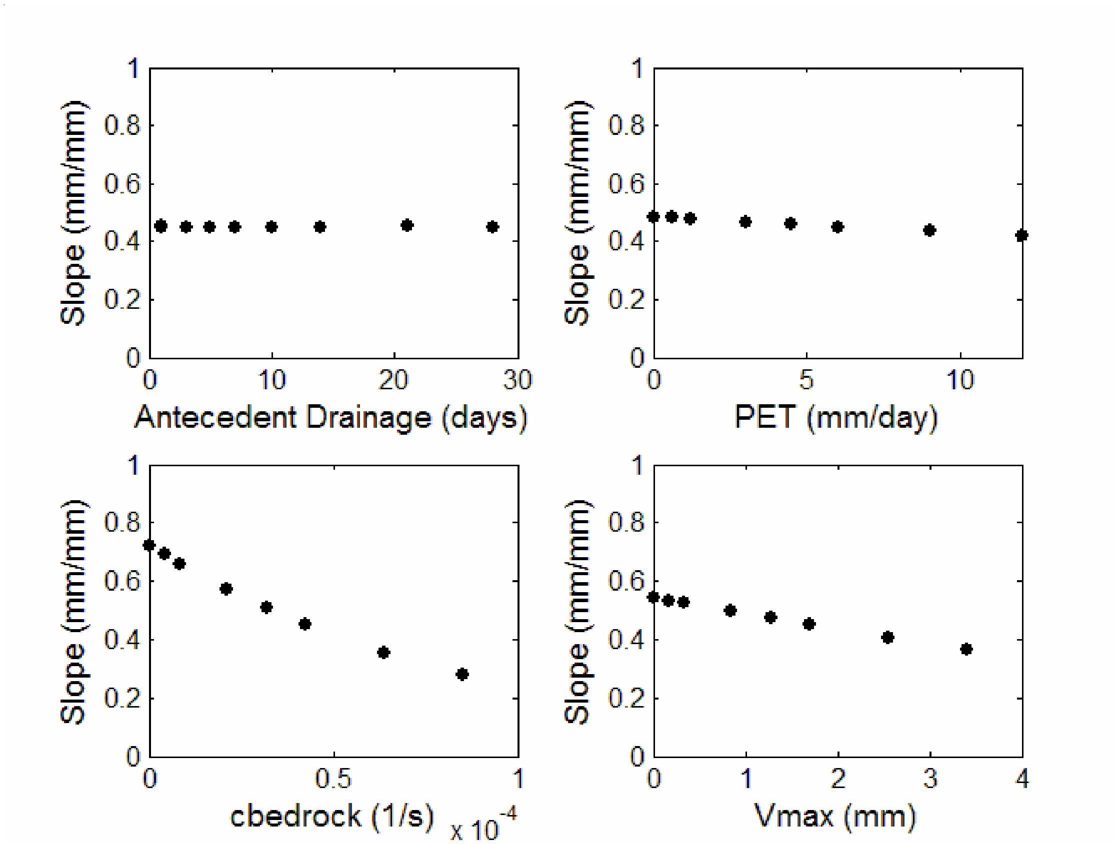


Figure 3.8 Threshold and slope vs. products of fill and spill factors, and soil moisture deficit factors. Antecedent PET is (t_{drain}*PET)

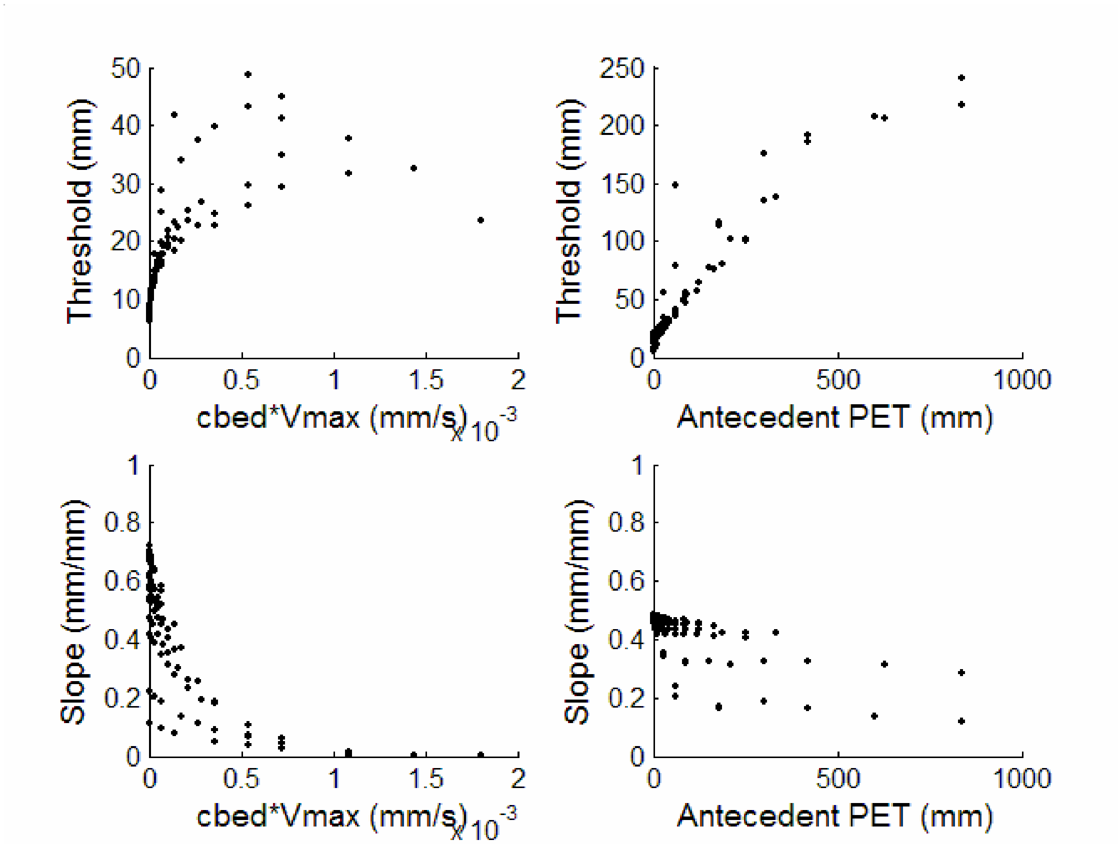


Figure 3.9 Measured whole storm precipitation / discharge dynamics at two instrumented field sites: a) Maimai hillslope (0.09 ha), 200 days of monitoring; b) M8 catchment (3.8 ha), two years of monitoring. Hillslope threshold estimated at 20 mm, while catchment threshold estimated at 8.5 mm.

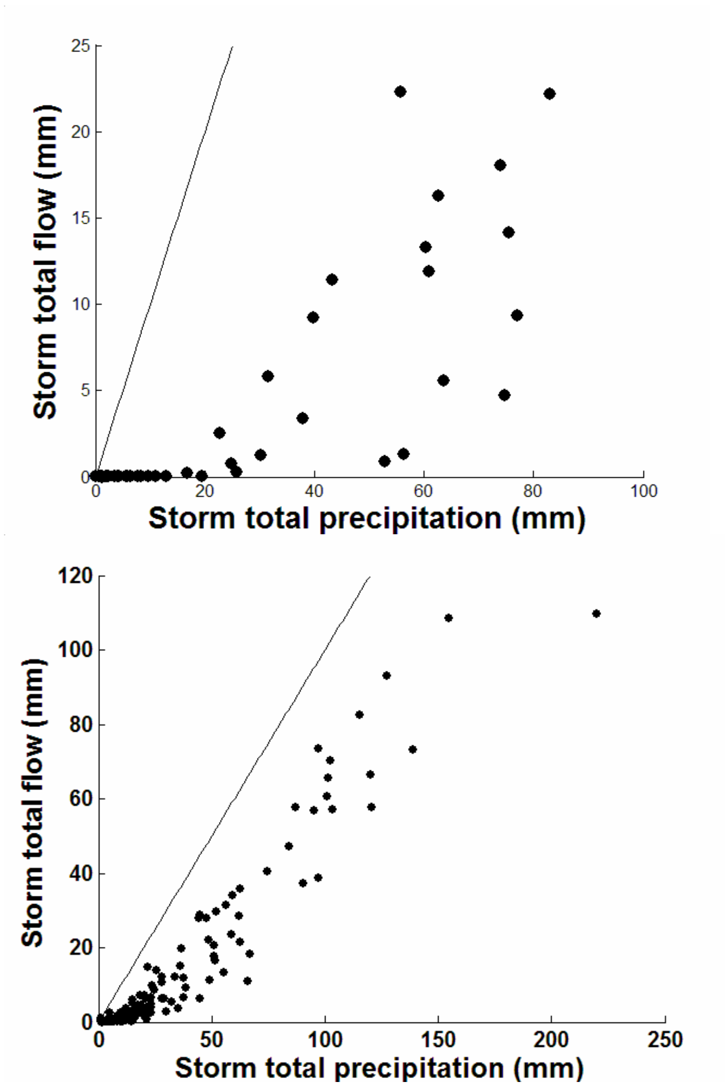


Figure 3.10 Measured whole storm precipitation / discharge dynamics at five instrumented research catchments: WS1 (101.3 ha), 2 (96 ha), 3 (60 ha), 9 (9 ha) and 10 (10 ha) at the HJA Andrews Experimental Forest. Storms are binned according to antecedent drainage time: a) All events with less than 5 days antecedent drainage; b) all events with between five and ten days antecedent drainage; c) all events with greater than ten days of antecedent drainage. Estimated thresholds for the three groups are 0 mm, 56 mm, and 83 mm.

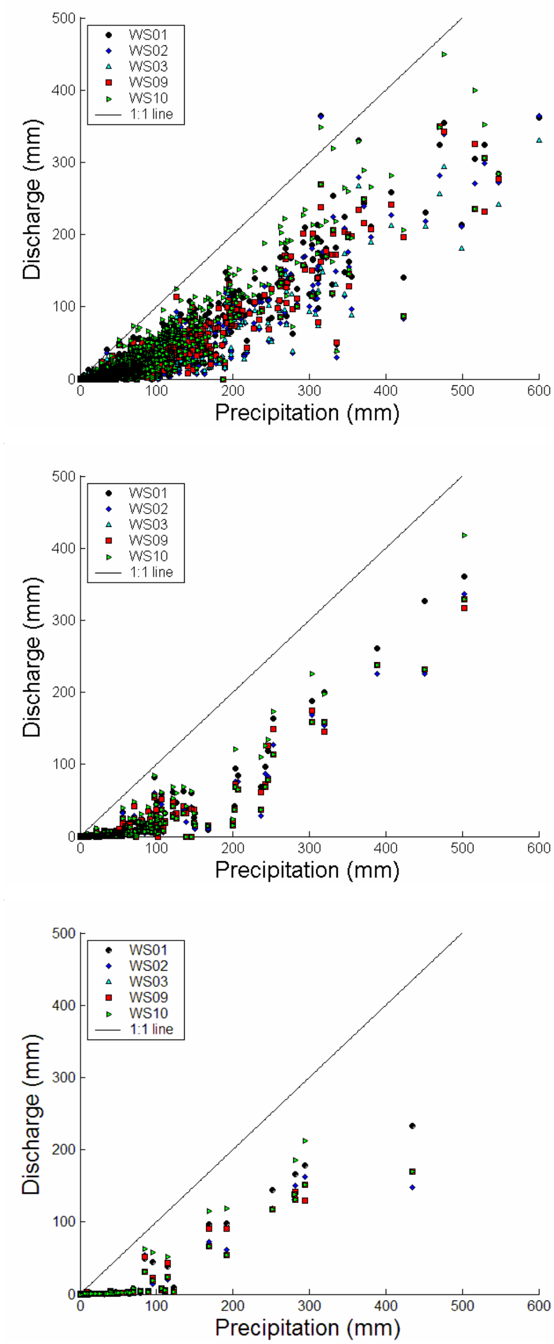


Figure 3.11 Dotty plots of $c_{bedrock}$ and V_{max} vs. 40 day E and cumulative tracer breakthrough. The lines denote the range of acceptable model fits. The x axis spans the range of the parameters. The product of the two parameters, a measure of the speed of pool drainage, is more identified than either parameter individually.

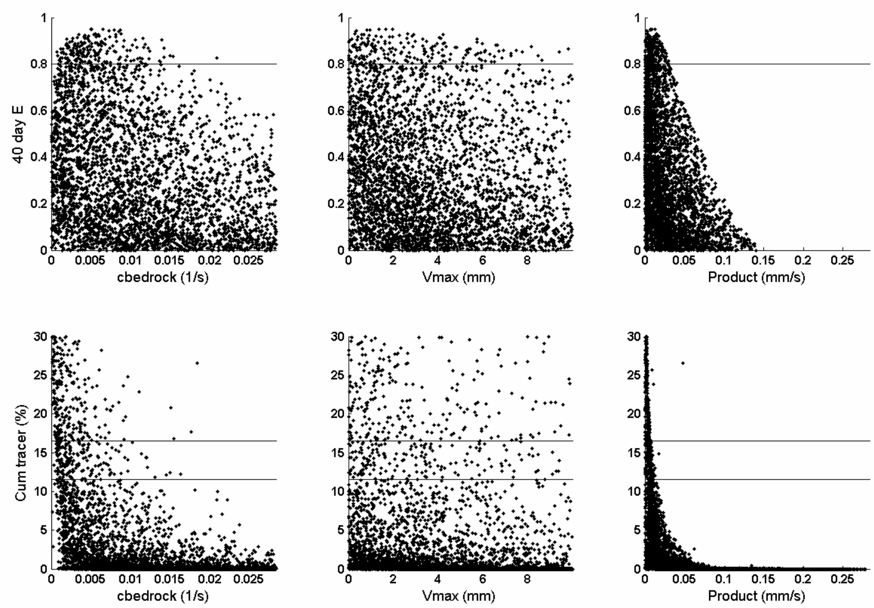
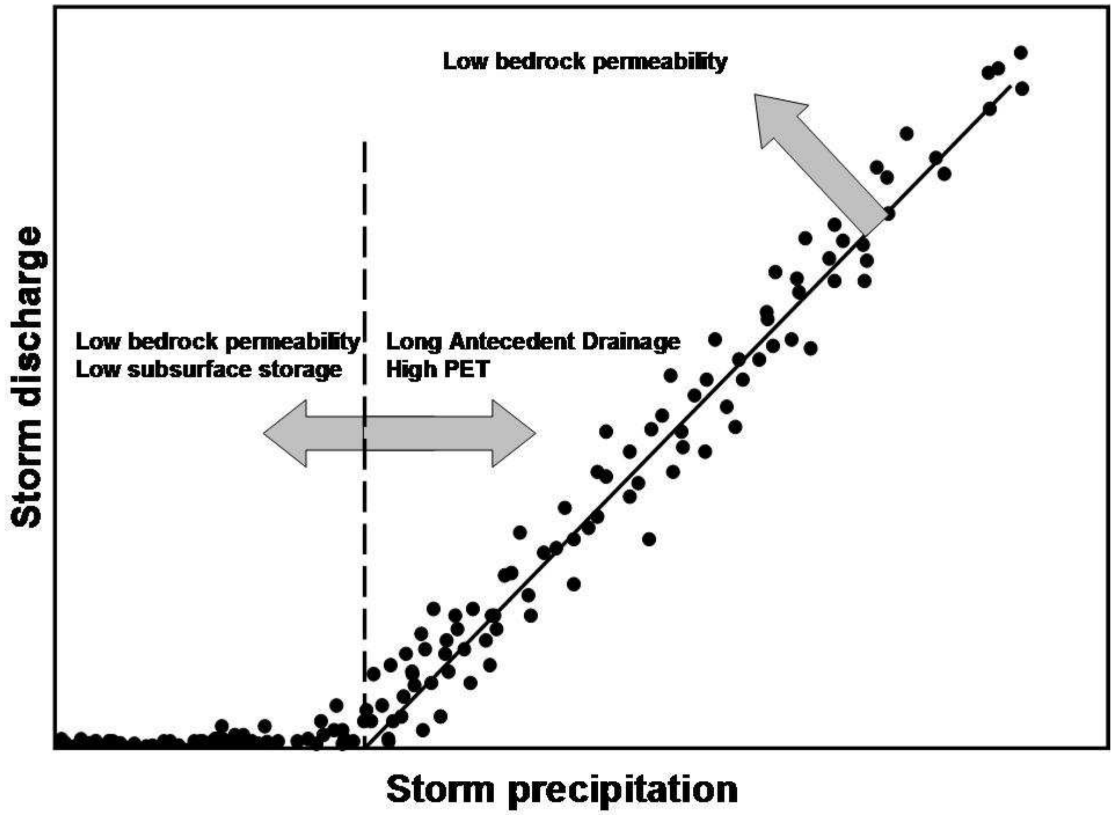


Figure 3.12 Schematic of fill and spill and soil moisture deficit control of precipitation discharge threshold and slope of the excess precipitation discharge line.



4. Experimental closure of the hillslope water balance within a measurement uncertainty framework

4.1. Introduction

Hillslope hydrology is an uncertain science. Uncertainties in initial conditions and boundary conditions are difficult to measure (Eberhardt and Thomas, 1991), and often coarsely estimated (Beven, 2006a). Internal state conditions are often based on limited point measurements extrapolated to much larger scales, with the corresponding difficulties in extrapolation of the uncertainty in these measurements. The difficulty in performing repeatable experiments at the hillslope scale makes the conclusions of field studies difficult to asses. The enormous heterogeneity of hydrological parameters, such as hydraulic conductivity, soil depth, macroscale soil structure, and soil texture compounds the measurement uncertainty (McDonnell et al., 2007). Despite, or perhaps because of these difficulties, a rigorous accounting of measurement uncertainty is rarely performed in conjunction with field experimentation. While many have made the case for explicit uncertainty analysis in numerical modeling, (Beven and Binley, 1992; Wagener, 2003) Beven (2006b) has extended this call to experimentalists, asking rhetorically “should it not be required that every paper in both field and modelling studies should attempt to evaluate the uncertainty in the results?” To date quantitative measurement uncertainty analysis remains rare in experimental hillslope hydrology.

Reporting of model uncertainty, whether due to equifinality issues, model parameterization and structure uncertainties (Beven, 2002), or the problems associated with extrapolating models beyond the calibration ranges, has become standard in catchment and hillslope hydrology (Andréassian et al., 2007; Beven, 2006b; Hall et al., 2007; Mantovan and Todini, 2006; Montanari, 2007; Sivakumar, 2008). Methods, such as the general likelihood uncertainty estimation (GLUE; Beven and Binley, 1992) and various Bayesian methodologies (Gupta et al., 1998; Kavetski et al., 2002; Kuczera and Mroczkowski, 1998; Mantovan and Todini, 2006; Vrugt et al., 2003) have been developed and are now widely used to quantify the uncertainty of predictions from calibrated model hydrological models. While this work is important, lack of clear and thoughtful analysis of uncertainty in field studies leaves catchment modelers to either determine on their own the error structure of field data, or to ignore it entirely. Some attempts to incorporate input uncertainty have been developed, such the use of fuzzy measures (i.e. Bárdossy, 1996; Özelkan and Duckstein, 2001) and soft data (i.e. Seibert and McDonnell, 2002). These efforts, however, are still dependent on experimentalist reports of measurement uncertainty. Nevertheless, as experimentalists, we rarely analyze and report the uncertainties in our field work, especially when reporting flux rates, mass balances, field parameter measurements and other potential inputs into numerical models.

Rigorous uncertainty analysis consists of both a thoughtful assessment of error for both the measurements themselves and the propagation of the measurement uncertainty through the functional uses of the data (Taylor, 1997). Error propagation needs to be performed when using measured data for both individual estimates and for aggregated measures such as daily averages or whole experiment total fluxes and stores. Random (precision) and systematic (accuracy) uncertainty types need be identified for each measurement instrument used and process assumption made, as the two are dealt with in error propagation differently.

Here we present a full measurement uncertainty analysis associated with a hillslope water balance experiment at the H. J. Andrews Experimental Forest (HJA) in western Oregon. The water balance, or continuity equation, is perhaps the most basic equation in hydrology (Rodriguez-Iturbe, 2000):

$$P = Q + \Delta S + ET + DS \qquad (1)$$

where the inputs to the hillslope (Precipitation, P) are balanced by the outputs (Surface runoff, Q; Changes in storage , ΔS; evapotranspiration, ET; deep seepage to groundwater DS). Despite the mathematical simplicity of the equation, measuring each of the components of the water balance is very difficult and closure of the water balance at the hillslope scale is rarely done (Beven, 2001). Surface runoff takes many forms, including overland flow, shallow lateral subsurface flow and bedrock return flow, each of which is difficult to measure without significant hillslope trenching infrastructure (Freer et al., 2002; Peters et al., 1995; Woods and Rowe, 1996) or overland flow trough deployment (Bonell and Gilmour, 1978). Quantifying fluxes into and out of storage requires extensive monitoring, as storage fluxes are dynamic (Ridolfi et al., 2003), hysteretic (Ewen and Birkinshaw, 2007), and vary with depth and topographic location (Western et al., 1998). The fluxes of evaporation and transpiration are generally lumped and their estimation is very data intensive, and often estimated as the residual of the other components in Equation 1 (e.g. Montgomery et al., 1997). Deep seepage is very difficult to measure at the hillslope scale, requiring deep wells into the hillslope aquifer. Deep seepage is often thought to be negligible, though recently flow through the bedrock has been shown to be a significant flow pathway (Katsura et al., 2008; Katsuyama et al., 2005; Onda et al., 2001; Tromp-van Meerveld et al., 2006).

Closure of the water balance is often done at larger time and space scales (Winter, 1981), especially in paired watershed experiments (Stednick, 1996). This closure is generally performed at an annual scale, where changes in storage and deep seepage fluxes are considered negligible, and the residual is assigned to ET, which is usually not measured. Closure is trivial at the smaller lab soil core scale, where inputs and outputs can be strictly controlled and monitored (e.g. McIntosh et al., 1999; Rasmussen et al., 2000). Closure of the water balance at the soil pedon scale has been performed using soil lysimeters, where lateral flow is restrained and vertical seepage can be measured (e.g. Gee et al., 1994). Intermediate to these scales, closure of the water balance at the hillslope scale has rarely been performed. Such a hillslope scale accounting is needed in hillslope hydrology because too often assumptions are made about the residuals, while only one or two components are measured (i.e. lateral subsurface flow and changes in storage).

For rigorous closing of the hillslope water balance we need to both minimize the uncertainty in the measurements of the components, and also attempt to quantify the uncertainties that remain. One method for minimizing uncertainty in the measurement of the water balance components is the use of controlled irrigation experiments, rather than passive storm monitoring (Eberhardt and Thomas, 1991). Irrigation experiments have the benefit of experimentalist control of the inputs, and directed measurements of the outputs. Hillslope irrigation experiments have been used in hydrology to determine solute transport characteristics (Hornberger et al., 1991), the role of flow through fractured bedrock (Montgomery et al., 1997), the identification of hillslope scale hydraulic conductivity (Brooks et al., 2004) and the partitioning of hillslope runoff processes (Scherrer et al., 2007).

The overall objective of this experiment was to close the water balance, identifying the relative partitioning of and uncertainties around the measured individual water balance components of evaporation, transpiration, lateral subsurface flow, bedrock return flow and fluxes into and out of storage. Within the overall objective, we address specific questions of:

1) How do uncertainties in individual measurements propagate through the functional uses of the measurements into the water balance components?
2) How does the quantification of individual water balance components improve our understanding of key hillslope processes, especially
    i) The storage discharge relationship at the hillslope scale
    ii) The role of bedrock flow at the hillslope scale
3) How do measurement uncertainties impact our process conceptualization of hillslope flow processes?

4.2. Site description

The study hillslope is located in WS10, in the HJA Experimental Forest in the western Cascades, Oregon, USA (44.20^{0}N, 122.25^{o}W). The HJA is part of the Long Term Ecological Research program, and has a data record of meteorological and discharge records from 1958 to the present. The climate is Mediterranean, with dry summers and wet winters characterized by long, low intensity storms: dry periods of 25 days and storms lasting 20 days have a 1 year return interval. WS10 has been the site of extensive research of hillslope hydrologic processes (Harr, 1977). A ten meter wide recording trench is situated at the hillslope base to collect lateral subsurface flow (McGuire et al., 2007).

Frequent debris flows at WS10 (most recently 1996) have scoured the stream channel to bedrock removing the riparian area in the lowermost reach. Soils are gravelly clay loams, classified as Typic Dystrochrepts, with poorly developed structure, high hydraulic conductivities (up to 2.8E-4 m/s, decreasing rapidly with depth), and high drainable porosity (15 – 30%) (Ranken, 1974). Soils are well aggregated, tending towards massive structure at depth. Soils on the study hillslope range from 0.1 m adjacent to the stream, to 2.4 m at the upper limit of the irrigated area. Soils have distinct pore size distribution shifts at 0.3, 0.7 and 1.0 m, resulting in transient lateral subsurface flow at these interfacial zones (Harr, 1977; van Verseveld, 2006). Soils are underlain by Saprolite, which thins towards the stream. Beneath this bedrock is mainly unweathered andesite and coarse breccias (James, 1978; Swanson and James, 1975). Additional site description can be found in McGuire et al. (2007).

4.3. Methods

4.3.1. Hillslope delineation

The irrigated area was chosen so that it would drain downslope into the 10 m wide collecting trench (Figure 4.1). We determined the irrigation area by analysis of surface topography. Bedrock topography on this planar slope was also planar (van Verseveld, 2006) in the region upslope of the collecting trench. The irrigated area was 9.4 m wide at the base, tapering to 8.2 m at the top and extended 20 m upslope of the trench (172 m^2). This trapezoidal area was narrower at the top than the base, reflecting increasing uncertainty in flow paths with increasing distance from the trench. Irrigation width and length measurements are estimated to have an uncertainty of 0.5 m due to difficulties in measurement and determination of wetted area. This leads to a total area uncertainty of 8.3 m^2.

4.3.2. Irrigation application

A rectangular grid of 36 (9 rows of 4) micro-sprinklers (with approximately 1 m irrigation radius) was installed on the hillslope, with sprinkler heads spaced 2 m apart. Sprinklers were controlled with an automatic timer to maintain a consistent application rate throughout the experiment with the exception of four minor malfunction periods. Sprinkler rate was measured by an array of 72 (0.05 and 0.1 m diameter) cups that were sampled every 4-12 hours during days 12 through 19 of the experiment. Additionally, three tipping bucket rain gauges (Trutrack, Rain-SYS-1mm) recorded irrigation rates throughout the experiment. The cups and tipping buckets were placed randomly in the sprinkled area, between 0.1 and 0.8 m from the sprinkler heads. The uncertainty in the rainfall application was determined by propagation of the uncertainty in the surface area of the measuring cups (cup radius ± 0.001 m) and the volumetric measurements (± 0.001 L).

4.3.3. Lateral subsurface flow

Hillslope lateral subsurface flow was measured with a 10 meter wide trench consisting of sheet metal anchored 0.05 m into bedrock and sealed with cement, installed at the intersection of the study hillslope and the exposed bedrock stream channel (McGuire et al., 2007). The trench system is assumed to be nearly water tight, as no evidence of leakage was seen during the experiment. Bulk lateral subsurface flow was routed to a stilling well with a 30^o V-Notch Weir, where a 0.25 m capacitance water level recorder (Trutrack, model PLUT-HR, measurement ± 0.0025 m) measured stage height at 10 minute intervals.

A rating curve for the stage / discharge relationship was developed using 32 manual measurements of discharge covering the range of values experienced during the irrigation experiment (R^2 = 0.97). The relative error between the manual measurements and the stage predicted discharge measurement averaged 8.76%, ranging from 1-20%. The relative error was weakly correlated with stage (R^2 = 0.34). The absolute error averaged 12 L/s, and was not correlated with stage. The absolute error was used as the systematic uncertainty in lateral subsurface flow, while the instrument precision (0.0025 m) was the random uncertainty.

4.3.4. Watershed discharge

Discharge from the second order stream draining WS10 has been monitored with a broad crested weir 100 m downstream of the hillslope since 1969 as part of the long term monitoring at the HJA. A 90º V notch weir had been installed for higher precision measurement of summer low flows, with the stage measured with a Model 2 Stevens Instruments Position Analog Transmitter (PAT, ± 0.0003 m) recorder controlled by a data logger (Campbell Scientific CR10X). A stage/discharge relationship was established based on 31 manual measurements of discharge taken over 0 – 2 times the range of discharge seen during the irrigation experiment. The absolute percent difference between measured and stage estimated discharge averaged 3.6%, with no correlation between relative error and stage. The absolute error was positively correlated with stage (R^2 = 0.43). The discharge during the experiment did not exceed the calibration range, so problems of rating curve indefinition are not expected (Clarke, 1999). The percent error in the rating curve was used as the systematic error, while the precision of the PAT was defined as the random error for uncertainty analysis.

During the course of the experiment, WS10 discharge receded, since the previous rainfall at the site was 10 days prior to the experiment. To determine the increase in watershed discharge due to the irrigation experiment, we created a master recession for WS10 using data from the summers of 2002 through 2004 from the WS10 gauging (data record available at andrewsforest.oregonstate.edu). Due to the variation in timing of the spring rainfall cessation, the summer WS10 discharge recession began at different dates in different years, ranging from mid June to late July. Recession from the three summers was aligned to begin with similar discharge rates. Discharge ($Q_M(t)$) was then modeled by:

$$Q_{\mathrm{M}}(t) = Q_o e^{-(t-t_o)/T_c} \qquad (2)$$

where t is the Julian day, T_c is the recession coefficient, and Q_0 is discharge at the time t_o (Chapman, 1999; Sujono et al., 2004). The recession coefficient T_c = 28.5 day led to a very good fit to the average of the three year's recession (R^2 = 0.97; Figure 4.2) We applied this function to our 24.4 day experiment period to determine the increase in WS10 discharge due to irrigation, using t_o as Julian Day 200, and Q_o as 1118 L/hr.

Uncertainty in the master recession has two components, uncertainty in initial discharge, Q_o, and uncertainty in the recession coefficient T_c. The uncertainty in Q_o was taken as the uncertainty in the measurement of watershed discharge at time zero, equal to 39 L/hr (0.035 x 1118 L/hr). The uncertainty in T_c was determined by fitting an exponential to the 3 year average watershed recession, then varying the uncertainty in T_c until 80% of the average recession readings fell within the error bounds. This led to an estimate of the uncertainty in T_c of 18.6% (Figure 4.2). These uncertainties were considered systematic uncertainties, since they were not based on measurements made during the experiment.

4.3.5. Transpiration and canopy reference evapotranspiration

Transpiration was estimated from sap flux measurements of the dominant trees located within or bordering the sprinkled area (n = 9). Sap flux was measured using the constant-heat method (Granier, 1987):

$$Q_n = 0.0119 sa_n \left(\frac{\Delta T_{M,n} - \Delta T_n}{\Delta T_n} \right)^{1.231} \tag{3}$$

where sa_n is the sapwood area of the nth tree, $\Delta T_{M,n}$ is the maximum daily temperature difference between thermistors installed into the sapwood 4 cm apart, and ΔT_n is the instantaneous temperature difference.

Temperature was measured every 15 s using a homemade Copper-constantan thermocouples hooked in series to measure temperature difference (± 0.2 C) following methods outlined in Moore, (2004) and stored in a CR-10x datalogger (Campbell Scientific, Logan, UT) as 15 min means. 0.02 m probes were used for the sap flux measurements. Sapwood depths were determined by visually examining and measuring tree cores from the height of the sap flux sensors on each tree (± 0.001 m). For trees with sapwood depths greater than 0.02 m, corrections for radial variations in sap flux were estimated from measured radial sap flux profiles of trees of the same species and age at another location following methods outlined in Domec et al. (2006) and Moore et al. (2004). Uncertainties in the transpiration measurements were treated as systematic, as an on-site calibration of the technique or equipment was not performed.

Evapotranspiration was estimated using the standard Penman Monteith equation for canopy reference evapotranspiration, estimated using measured meteorological data taken at the site (Monteith and Unsworth, 2008):

$$CRET = \frac{\Delta (R_n - G) + \rho c_p \frac{VPD}{r_a}}{\lambda \left(\Delta + \gamma \left(1 + \frac{r_c}{r_a} \right) \right)} \tag{4}$$

where Δ is the partial derivative of the saturated vapor pressure curve with respect to temperature, R_n is the net incoming radiation, G is the ground heat flux, ρ is the dry air density, c_p is the specific heat capacity of air, VPD is vapor pressure deficit, r_c is the canopy resistance, r_a is atmospheric resistance, λ is the latent heat of vaporization, and γ is the psychrometer constant. The parameters ρ, c_p, λ and γ were assumed to be constant, and values taken from Monteith and Unsworth (2008) were used. The *VPD* is the product of the relative humidity and saturated vapor pressure, $e_s(T)$, and Δ is the first derivative of the $e_s(T)$ curve with respect to temperature. The saturated vapor pressure

$e_s(T)$ was calculated using a empirically derived exponential function of temperature (Murray, 1967; Tetens, 1930). The atmospheric resistance, r_a is a function of wind speed (Unsworth and Monteith, 2008):

$$r_a = \frac{\ln\left({(z-d)}/{z_o} \right)}{ku} \quad (5)$$

where z is the height of the canopy (22 m), d is the zone of zero displacement (0.65z), k is von Karman's constant, z_o is the roughness length (0.1z) and u is the wind speed measured at the hillslope. The canopy resistance, r_c, is a function of the forest type and structure, ranging from 100 – 250 s/m (Tan and Black, 1976). In this case, r_c was assumed to be a 175 s/m, and the uncertainty was assigned to encompass the range observed by Tan and Black (1976), ± 75 s/m. Ground heat flux was expected to be small, and estimated as 10% of net radiation, with a similarly large uncertainty, in this case 100% (G = 0-20%R_n).

Net radiation (*Rn*; Campbell Scientific Inc., model Q-7.1, ± 6%), relative humidity (*RH*; Campbell Scientific Inc., model HMP 35C, ± 2-3%), air temperature (*T*; Campbell Scientific Inc., model HMP 35C, ± 0.4^{o}C) and wind speed (*u*; R. M. Young Wind Monitors, model 05305, ± 0.2 m/s) were measured at 15 minute intervals throughout the experiment. Since a calibration of the meteorological equipment was not performed, the measurement uncertainty presented by the manufacturers was propagated as systematic error through the functional uses of the measured data.

4.3.6. Soil moisture

Soil moisture (volumetric water content) was measured at 24 locations within the irrigated area, at 5 depths in each location (0 – 0.15 m, 0.15 – 0.30 m, 0.30 – 0.60 m, 0.60 - 0.90 m, and 0.90 - 1.20 m) with a time domain reflectrometry (TDR) array (Environmental Sensors, Inc., model PRB-A, ± 3%) (Figure 4.1). Measurement sites were in a 4 by 8 grid (parallel and perpendicular to the stream channel, respectively), with sensor spacing of 2 m in each direction. Soil water content was measured hourly through the experiment. Of the 120 measurements (locations and depth), 57 of the probe segments gave consistent results. The remaining 63 measurement segments had data recording problems due to probes incompletely in the soil profile, poor electrical connections and poor contact between the probe and soil caused inconsistent readings from. Only the data from consistently working rods were analyzed. To determine the total soil storage, the profile average soil moisture was multiplied by the estimated soil depth (1.2 ± 0.1 m) and the irrigated area (172 ± 8 m^2).

Two significant sources of error lie in the soil moisture data. The first is the uncertainty is in measurement of the soil moisture measurements. This uncertainty is constrained to within 3% for each measurement, and 3% for the background, pre-experiment water content, taken from manufacturer calibration. The second is the subsurface volume represented by the soil moisture measurements themselves. We assumed that soil moisture outside the TDR grid, but within the sprinkled area, reacted similarly to the area measured by the probes. There is a possibility of some flux of water outside of the sprinkled area due to capillary effects and subsurface flow paths controlled by bedrock topography diverting water from the sprinkled area. Additional storage could have occured in the bedrock itself, which was likely unsaturated prior to the experiment. Our computed subsurface storage volumes were considered a minimum value of total subsurface storage. Since calibration of the soil moisture probes was not conducted in the field, the factory calibration uncertainty in soil moisture readings were treated as systematic.

4.4. Uncertainty accounting and estimation

We subdivided our uncertainty analysis into three categories: identification and quantification of measurement uncertainty of the instruments, propagation of the measurement uncertainty through the functional uses of the data, and propagation of measurement uncertainty through aggregated measures. We define and describe the mathematical treatment of these terms below.

4.4.1. Individual measurement uncertainty

Measurement error is the uncertainty in the precision and accuracy of the field instrument. The uncertainty in field measurements can be determined in a number of ways, including field calibration, manufacturer calibration, and expert opinion. We utilized all of these sources for this experiment. Hillslope and watershed discharge were determined from field calibration of the stage discharge relationship. Much of the meteorological data were not calibrated in the field, so the factory calibration uncertainties were used for each of the individual readings. Some variables used in the extended analysis (pre-irrigation hillslope and watershed discharge) were not measured directly throughout the experiment, but were based on historical data. We estimated the uncertainty of these variables based on expert opinion.

Of these sources of error, there are two types: random and systematic error, as expressed as the precision and accuracy of the measurements. Random errors include measurement errors that deviate randomly from the true observed value. These errors are assumed to be evenly distributed above and below the true value and to some extent cancel each other out when aggregated to longer time periods and spatial scales. Systematic errors, on the other hand, can affect all measurements in the same direction (i.e. under or over prediction), and thus do not diminish with increasing the length of the data set. For most of the field instruments, the manufacturer presents only one uncertainty estimate. In this case, when a field calibration has not been made, this value is treated in the uncertainty analysis as both the systematic and random error. In cases where a calibration has occurred, such as hillslope and catchment discharge, the systematic error (accuracy) is taken from the uncertainty of the calibration, while the random error (precision) is taken from the equipment measurement uncertainty. Table 4.1 lists source and type of uncertainty for each measurand.

4.4.2. Error propagation

Uncertainty in field measurements needs to be transferred through the functional uses of the data. Because many of the measured variables, such as temperature and wind speed, are used nonlinearly to calculate the water balance components, such as CRET, a given uncertainty in the measured quantity affects the water balance component nonlinearly as well. To account for this nonlinearity, uncertainties in measurements are propagated using the standard error propagation formula (Taylor, 1997).

When propagating error, we first assumed that the individual instruments were independent from each other. If q is a function of N variables:

$$q = f\left(x_1,\ldots x_N\right) \qquad (6)$$

where x has some random uncertainty $\delta_r x_n$ and some systematic uncertainty $\delta_s x_n$, then uncertainty in each of the measured values is propagated through *q* by:

$$\delta q = \sqrt{(\delta_s q)^2 + (\delta_r q)^2} \tag{7}$$

where δq is the propagated error in *q*, and

$$\delta_r q = \sqrt{\sum_{n=1}^{N} \left(\frac{\partial q}{\partial x_n} \delta_r x_n \right)^2} \tag{8}$$

and

$$\delta_s q = \sqrt{\sum_{n=1}^{N} \left(\frac{\partial q}{\partial x_n} \delta_s x_n \right)^2} \tag{9}$$

For example, transpiration is calculated following the empirical relationship developed by Granier (1987) in Equation 3. Measured variables include the sapwood area (*sa*), the temperature difference between two thermocouples inserted into the sapwood (ΔT) and the maximum daily temperature difference (ΔT_M). The uncertainty in transpiration is then:

$$\delta Q = \sqrt{(\delta_s Q)^2 + (\delta_r Q)^2} \tag{10}$$

where

$$\delta_r Q = \sqrt{\left(\frac{\partial Q}{\partial s} \delta_r sa \right)^2 + \left(\frac{\partial Q}{\partial \Delta T} \delta_r \Delta T \right)^2 + \left(\frac{\partial Q}{\partial \Delta T_M} \delta_r \Delta T_M \right)^2} \tag{11}$$

$$\delta_s Q = \sqrt{\left(\frac{\partial Q}{\partial s} \delta_s sa \right)^2 + \left(\frac{\partial Q}{\partial \Delta T} \delta_s \Delta T \right)^2 + \left(\frac{\partial Q}{\partial \Delta T_M} \delta_s \Delta T_M \right)^2} \tag{12}$$

Complete propagation of error through the various formulae used in the calculation of the water balance components are presented in the appendix.

4.4.3. Aggregated error

When aggregating measurements from individual time steps to longer time scales (i.e. daily averages, whole experiment total fluxes) the type of measurement error, whether random or systematic, determines how the aggregation of the error is performed. Random errors, when aggregated, diminish with increasing size of the data set according to $1/\sqrt{T}$, where T is the number of data points (Taylor, 1997). For random errors, the error of the aggregate is the sum of the squares. If, for instance

$$\hat{q} = \sum_{t=1}^{T} q(x_t) \tag{13}$$

and the error in x is random, the aggregated error is:

$$\delta_r \hat{q} = \sqrt{\sum_{t=1}^{T} \left(\frac{\partial q}{\partial x_t} \delta_r x_t \right)^2} \tag{14}$$

for measurand x from time 1:T.

Systematic errors must be aggregated differently. As persistent offsets or multipliers to the data, they act in an additive manner (Moncrieff et al., 1996) and do not diminish with increasing data set size. When propagated, the error of the aggregate is the square of the sum:

$$\delta_s \hat{q} = \sqrt{\left(\sum_{t=1}^{T} \frac{\partial q}{\partial x_t} \delta_s x_t \right)^2} \tag{15}$$

for measurand x from time 1:T. For values that aggregate over long time periods, with large T, the aggregated random error is dwarfed by the systematic error, as T becomes much larger than $\sqrt{T}$. The appendix presents formulas for the aggregation of measurement uncertainties.

4.5. Results

4.5.1. Water Balance Components

4.5.1.1. Inputs

Irrigation application was relatively constant for the 24.4 day experiment with the exception of four malfunctions in the timer apparatus which caused the irrigation to remain either on or off for a short period of time. Irrigation began at 0530 hr on Julian Day 208 (27-July-05), and ended at 1412 hr, Julian Day 232 (20-August-05). On midnight day 210 irrigation turned off for 9 hours and thereafter constant for the next 18 days. Sprinkler malfunctions occurred on days 228, 229 and 230.

The weighted irrigation rate based on the 72 collection cups was 3.8 ± 3 mm/hr (Figure 4.3). With a measured irrigated area of 172 ± 8 m^2, the corresponding total application was 654 ± 33 L/hr. Irrigation rates varied spatially due to both variations in the individual sprinkler heads application rates, and temporary obstructions (including vegetation and equipment) between sprinklers and measuring cups (SD = 3.3 mm/hr). This variability is more a measure of the spatial variability of application than a measure of application rate uncertainty.

4.5.1.2. Outputs - lateral subsurface flow

Lateral subsurface flow measured at the trench responded quickly to irrigation, with a detectable rise in discharge within an hour of irrigation initiation (Figure 4.4). Lateral subsurface flow rose from a pre-irrigation daily average rate of 30 ± 1 L/hr to a steady state daily average value of 284 ± 20 L/hr within 5 days. Before, during and after the experiment, a clear diel pattern in flow was evident. Steady state discharge was maintained for 13 days, after which a series of sprinkler malfunctions increased discharge by over 30% for 3 days. At the end of the irrigation, on Julian Day 232, the instantaneous lateral subsurface flow was 270 ± 16 L/hr. After irrigation ceased, lateral subsurface flow returned to within 200% of pre experiment levels within 24 hours. Lateral subsurface flow for the duration of the experiment was 102543 ± 7451 L. Lateral subsurface flow for the periods of the experiment plus five and ten days was 106156 ± 8979 L and 107760 ± 10507 L respectively (Table 4.2).

4.5.1.3. Outputs - WS10 discharge

WS10 discharge responded to irrigation similarly to the lateral subsurface flow measured at the hillslope trench (Figure 4.5). The pre-irrigation WS10 recession slowed within one hour after the onset of irrigation. After 5 hours, WS10 discharge then increased for the next 6 days of the experiment. After day 6, WS10 discharge began to recede parallel to the master recession curve. This was due to combined steady input from the irrigated hillslope and continued recession from the remaining area of the watershed. Comparison to the master recession indicates an increase in discharge due to the sprinkling of 461 ± 115 L/hr during the period of steady state input. The recession remained parallel to the master recession until the series of sprinkler malfunctions caused an increase in discharge similar to that seen at the hillslope. After cessation of the sprinkling, WS10 drainage was slower than observed at the trenched hillslope. WS10 discharge did not return to the master recession before a rain event 10 days after the end of irrigation. Total increased discharge measured at the watershed outlet for the duration of the irrigation was 227829 ± 56847 L, with 250489 ± 67776 L and 260677 ± 78247 L for the irrigation plus 5 and 10 days drainage, respectively (Table 4.2). The uncertainty in the aggregated measures increased due to increased uncertainty in background watershed discharge at late time.

4.5.1.4. Outputs - transpiration and canopy reference evapotranspiration

Transpiration from the dominant trees in the sprinkled area and canopy reference evapotranspiration both showed a strong diel pattern, during and after the irrigation experiment (Figure 4.6). Sap flux averaged 0.8 ± 0.1 L/hr for the 9 instrumented trees, for a total sap flux of 9 ± 1 L/hr for the stand of trees on the instrumented hillslope. The maximum instantaneous stand flux rate of 25 ± 2 L/hr typically occurred around early afternoon (1400 hr). The mean and maximum flux rates remained constant before, during and after the experiment, suggesting that trees in the plot were not water stressed at the onset of irrigation. Analysis of transpiration and water use patterns of vegetation during the experiment is discussed in Barnard et al. (submitted). Transpiration for the duration of the experiment totaled 5,448 ± 343 L. Transpiration for the period of the experiment +5 and +10 days was 6,456 ± 409 and 7,318 ± 470 L, respectively (Table 4.2).

CRET showed a steady decline from a high at the initiation of irrigation (36.0 ± 1.3 L/hr – daily average) through the end of the experiment (33.5 ± 1.2 L/hr) and to 5 and 10 days after the end of irrigation (30.4 ± 1.2 and 26.4 ± 1 L/hr, respectively) (Figure 4.7). This decline was influenced primarily by incoming net radiation, which declined throughout the monitoring period, due a decline in the daylight hours. Since the soil remained wet and water supply was not likely the limiting factor for evapotranspiration, we assume that actual evapotranspiration equals canopy reference evapotranspiration during the experiment and afterwards. Total evaporative losses for the duration of the experiment were estimated as 30055 ± 12692 L. Evaporative losses for the period extending 5 and 10 days afterwards were 35767 ± 15170 L and 41114 ± 17721 L, respectively (Table 4.2).

4.5.1.5. Change in storage - soil moisture

Soil moisture followed the same general pattern as the hillslope and WS10 discharge: a quick response to irrigation, then steady state, and a recession after the stoppage of irrigation on Julian Day 232 (Figure 4.8). Initial soil volumetric water content averaged 8.6 ± 0.3% at the onset of irrigation. TDR readings showed an initial increase in soil moisture in the upper 0.6 m in the first 30 min of irrigation. Soil moisture at 0.6 -0.9 m increased after 90 min, and sensors below 0.9 m increased after 150 min. Soil moisture reached a steady state within 5-6 days (Julian Day 213-214), with the shallower depths reaching steady state more quickly than at depth. Average profile volumetric water content during steady state was 20.2 ± 0.6%. Steady state conditions persisted until day 228, when the first of the sprinkler malfunctions caused an increase in soil moisture. After irrigation ceased on day 232, the soil profile drained quickly for the first 8 - 12 hours from a high of 21.5 ± 0.6%, followed by a slower, more sustained drainage for the duration of monitoring. The upper soil layers drained most rapidly, with slower drainage at depth. Average profile soil volumetric moisture dropped to 16.6 ± 0.5% within 5 days, and 14.4 ± 0.4% within 10 days. None of the five soil profiles returned to pre-irrigation levels by day 250, over 3 weeks after irrigation ceased.

Flow into soil moisture storage ($Q_{storage}$) was calculated as:

$$Q_S = \Delta SAd \tag{16}$$

where ΔS is the difference in average soil column water content before the experiment and at the measurement time, A was the area sprinkled, and d was the average soil depth. The average soil depth on the sprinkled hillslope is 1.2 ± 0.1 m, and the area sprinkled was 172 ± 8 m^2. This formula assumes that the depth of soil storage is equivalent to the soil depth (i.e. the bedrock is saturated), and the aerial extent is equivalent to the sprinkled area (i.e. no lateral spreading parallel to the trench). At steady state, the total soil storage is estimated at 41693 ± 4208 L. At the end of irrigation, after the series of sprinkler malfunctions, total storage was 44,376 ± 4479 L. Total profile storage declined to 34,324 ± 3465 L and 29,784 ± 3006 L after 5 and 10 days, respectively (Table 4.2).

4.5.2. Partitioning of the water balance

4.5.2.1. The water balance at steady state

The steady state water balance was split into 4 components: lateral subsurface flow collected by the hillslope trench, deep seepage and bypass flow that missed the hillslope trench and was measured at the WS10 outlet, transpiration, and evapotranspiration. Soil moisture storage was omitted from the steady state water balance because it was assumed to be constant during the steady state conditions. Steady state rates were calculated as the average values for the period Julian Day 219 through 226. The uncertainty in each of the averaged values was calculated using the error propagation formulas in the appendix. This period was after the system reached steady state, as evidenced by steady lateral subsurface flow, soil moisture storage, and WS10 discharge, and did not include the period of sprinkler malfunctions.

At steady state, the irrigation rate was 659 ± 33 L/hr (Table 1). The discharge at the WS10 outlet, corrected with the master recession, was 461 ± 115 L/hr (70 ± 17%), which includes 284 ± 20 L/hr (43 ± 3%) increased lateral subsurface flow measured at the hillslope trench. Evapotranspiration was estimated at 50 ± 21 L/hr (8 ± 3%), of which 9 ± 1 L/hr (1 ± 0.2%) was measured as sapflow. Flow measured at the watershed, hillslope trench, sapflow and evapotranspiration account for 511 ± 121 L/hr, or 76 ± 18% of the irrigated water during steady state conditions.

4.5.2.2. Cumulative water balance

Cumulative flow volumes for each component were compared to the total irrigated volume for the duration of irrigation, the duration + 5 days, and the duration +10 days (Table 1). The uncertainty in each of the aggregated values was calculated using the error propagation formulas in the appendix. Total irrigation volume applied was 394000 ± 19700 L. Total flow measured at the watershed, corrected for the master recession, was 58 ± 14, 64 ± 17 and 66 ± 20% of irrigation, which includes 26 ± 2, 27 ± 2 and 27 ± 3% lateral subsurface flow from the hillslope for the duration of irrigation, +5 days and +10 days, respectively. Canopy reference evaporation accounted for 8 ± 3, 9 ± 4 and 10 ± 4% of the irrigated amount for the three time periods, respectively, which includes 1 ± 0.2, 2 ± 0.2, and 2 ± 0.2% from forest transpiration. Net flow into soil moisture accounted for 7

± 0.4, 4 ± 0.4 and 3 ± 0.4% of irrigation for the three time periods, respectively, decreasing as the soil drained after sprinkling ceased. The total mass accounted for are then 72 ± 16, 77 ± 18 and 79 ± 21% from these 5 sources for the 3 time periods.

4.6. Discussion

4.6.1. New process understanding

4.6.1.1. Hillslope scale storage discharge relationship

The hysteretic nature of the soil moisture release curve has been acknowledged for nearly 80 years (Jaynes, 1990). Hysteretic loops also exist in the storage discharge relationship at the hillslope scale (Beven, 2006a; Ewen and Birkinshaw, 2007; Kendall et al., 1999; Seibert et al., 2003). This hysteresis, a signal of the non-singular relationship between hillslope storage and hillslope discharge, has been attributed to the connection – disconnection of subsurface saturated areas (i.e. (Tromp-van Meerveld and McDonnell, 2006) aggregated hysteresis in the core scale soil characteristics (Beven, 2006a), the activation of preferential flow pathways (McDonnell, 1990), and the transition between different flow processes (Ewen and Birkinshaw, 2007).

We observed a hysteretic relationship between storage and lateral subsurface flow was observed in this experiment. Soil moisture storage and hillslope and watershed discharge all responded very quickly to irrigation, while transpiration and evaporation remained steady and relatively unchanging throughout. We interpret the muted response of transpiration and canopy reference evapotranspiration to the irrigation to be due to the lack of water stress experienced by the vegetation (for detailed discussion see Barnard et al. (submitted). Lateral subsurface flow increased by 34% within 1 hour of irrigation, WS10 discharge increased 10% within 32 hours, and profile average soil moisture storage increased by over 200% within the first 3 hours of irrigation. During the wetup, hillslope and watershed discharge were well correlated with soil moisture storage. The sprinkler malfunction on day 210, where irrigation ceased for 9 hours, is seen in both the lateral subsurface flow and the soil moisture, especially in the shallow depths. This suggests a tight connection between the shallow soil and discharge, consistent with the findings of lateral subsurface flow at permeability discontinuities in the soil profiles (noted at 30 and 70 cm depth by Harr, 1977).

Following the termination of irrigation, the coupling between storage, hillslope subsurface flow and WS10 discharge weakened. While lateral subsurface flow recessed very quickly, declining 90% within 50 minutes, the WS10 discharge recessed more slowly. WS10 discharge remained more than 10% above steady state for more than 5 days, though at this point the discharge was within the uncertainty bounds of the background. The soil moisture storage exhibited a bimodal recession, with a quick, short drop in profile average soil moisture, followed by a slow recession for the duration of monitoring. Though the profile average soil moisture storage dropped quickly after the end of irrigation, the drop was not very large (<10%). The recession of soil moisture after this initial drop, when the largest pores were emptied, was very gradual. Average profile storage remained over 300% pre-event levels at the end of monitoring, 10 days after the end of the experiment (Figure 4.8). The magnitude of the rapid recession was negatively correlated with the soil depth on the hillslope, consistent with the findings of Ranken (1974), who found that the macroporosity declines with depth at the site. During the recession, soil moisture storage was not correlated with either the WS10 discharge or the hillslope subsurface flow, demonstrating a complex hysteretic relationship between the soil moisture storage and hillslope and watershed discharge.

A strong counterclockwise hysteretic relationship between storage and lateral subsurface flow was observed during the irrigation experiment (Figure 4.9). At the WS10 scale, however, no clear hysteretic relationship was observed (Figure 4.10). This was due to the rapid recession of the lateral subsurface flow, and the slower recession in the discharge, when compared with the soil moisture. Although somewhat masked by the strong diel signal seen in the discharge, there appears to be a singular relationship between watershed discharge and hillslope storage. Paradoxically, the sprinkler malfunction during day 208 is easily seen in the lateral subsurface flow, but not in the WS10 discharge. One possible explanation for this anomalous behavior is the transition between vertical and lateral subsurface flow. During irrigation, the infiltration capacity of the underlying bedrock is reached, and lateral subsurface flow is initiated. Later, as irrigation ceases, lateral subsurface flow ceases as vertical fluxes drop lower than the infiltration rate of the bedrock. This would result in a quick reduction in lateral subsurface flow as observed at the hillslope during the sprinkler malfunction and after irrigation had ceased. Infiltration into the bedrock, and correspondingly high stream discharge (as measured at WS10 outlet) would remain relatively high, as the soil drained, now predominantly vertically. This would result in a singular relationship between storage and stream discharge, as observed.

This suggests that the hysteretic relationship between storage and flow could be a measure of the relative contributions of lateral and vertical flow. If a system is dominated by lateral flow, with minimal bedrock leakage at the site of monitoring (either a system underlain by relatively impermeable bedrock, or at larger scales, where streamflow is thought to be much greater than deep seepage), then little hysteresis would be expected to be observed. On the other hand, a system where leakage is a significant component of the water balance (such as this hillslope, see below), a strong hysteretic pattern would be expected, as high bedrock infiltration rates are exceeded only during large events or high intensity rainfall. The observed hysteretic response observed by others at the watershed scale (Beven, 2006a; Ewen and Birkinshaw, 2007) then suggests that their watersheds are not watertight, and deep seepage may be a significant component of the water balance.

4.6.1.2. Bedrock flow contribution to hillslope hydrology

Transient flow through the near surface bedrock has been observed at a number of field sites (Katsura et al., 2008; Katsuyama et al., 2005; Montgomery et al., 1997; Tromp-van Meerveld et al., 2006). While the reemergence of water lost to bedrock at the hillslope scale has sometimes been observed downstream (Montgomery et al., 1997), generally, the fate of this water, its interaction with shallow lateral subsurface flow paths, and the time in which it takes to reach the stream channel are unclear and poorly understood.

Previous WS10 hillslope storm monitoring of lateral subsurface flow by McGuire et al. (2007) reported that the hillslope area defined by the collection trench placement, (upslope contributing area of 1.7% of the watershed), contributes 2% of the annual catchment discharge. Assuming small uncertainty in the upslope contributing area, this indicates that the majority of the water falling on the hillslope is observed in the trench. During the irrigation experiment, however, lateral subsurface flow was underrepresented in the water balance (26 ± 2% of irrigation). While irrigation was confined to the area near the trench (< 20 m upslope), and a 1 m buffer was placed on each side to minimize flow bypassing to the right or left of the trench, the majority of water measured in the stream at the WS10 outlet was not observed in the trench. The hillslope trench, while not assumed to be watertight, is designed to minimize leakage, and thus is expected to capture the vast majority of the lateral subsurface flow at the hillslope base.

Two possible sources of this bypass are hypothesized: leakage below the hillslope trench, and flow routing around (likely down valley) of the hillslope trench. While it is often assumed that the bedrock is effectively impermeable during hillslope experimentation and monitoring (e.g. Freer et al., 2002; Mosley, 1979), recent evidence has shown that significantly permeable bedrock is the rule, rather than the exception at steep, forested hillslopes (Katsura et al., 2008; Katsuyama et al., 2005; Montgomery et al., 1997; Tromp-van Meerveld et al., 2006). Leakage at the hillslope scale and reemergence at the catchment scale is a possible explanation for the high discrepancies often seen between hillslope and catchment runoff ratios (e.g. Woods and Rowe, 1996). At steady state, the difference in the hillslope and WS10 discharge, (the amount of water bypassing the hillslope trench) averaged 177 ± 116 L/hr. If we assume that this water bypassed the hillslope trench by infiltrating into the bedrock, across the wetted cross sectional area, this would correspond to a leakage rate of 1.1 ± 0.6 mm/hr. While significant, these rates are well below the measured hydraulic conductivities of other steep, forested hillslopes (~5 mm/hr (Graham et al., in review; Tromp-van Meerveld et al., 2006)).

An alternative explanation for the low recovery at the hillslope trench is bypass down valley of the trench. While the irrigated area was delineated to drain downslope and into the trench, some uncertainties remain in locations of the dominant flowpath in the subsurface. Studies elsewhere have shown that bedrock topography is often a first order control on flow routing (Freer et al., 2002; Graham et al., in review). Analysis of soil depth measurements at the site indicates a relatively planar bedrock surface, parallel to the soil surface. However, the spatial scale of the bedrock features that can control routing can be very small, on the order of cm (Graham et al., in review). The map of soil depth was made on a 1 m grid (van Verseveld, 2006), and likely did not capture these small scale features. Some evidence of down valley flow routing around the hillslope trench was observed in the form of bank seepage downstream of the trench, though the observed seepage was a small fraction (estimated <10%) of the lateral subsurface flow.

The first explanation, bedrock leakage, is inconsistent with the findings of McGuire et al. (2007), who showed that the runoff ratio for the hillslope was consistent with that of the watershed. However, their estimate of the runoff ratio for the hillslope was dependent on an accurate assessment of the upslope contributing area, an easily calculated but very imprecise measure at the hillslope scale. Woods and Rowe (1997) demonstrated that small uncertainties in topography measurements greatly affected the upslope contributing area for a hillslope trench system. Additionally, if bedrock rather than surface topography controls flow routing, further uncertainties arise.

4.6.1.3. On the consequences of water balance losses

In attempting to close the water balance, we measured six components – precipitation lateral subsurface flow, watershed discharge, changes in sol moisture storage, evaporation and transpiration. Often, only two (precipitation and discharge, or precipitation and evaporation) or three (precipitation, discharge and evaporation) components are measured, and the residual is attributed to either evaporation or changes in storage. However, two more water balance components that are difficult to measure are generally ignored – deep seepage and changes in bedrock storage.

During and immediately after the irrigation experiment, a significant amount of the irrigation was not accounted for. The amount of unaccounted water decreased as the measurement time increased, declining from 28 ± 16% at the end of the experiment, to 23 ± 18% after five days drainage, to 20 ± 21% after ten days drainage. The uncertainty in the remainder correspondingly increased, primarily due to increased uncertainty in the WS10 master recession. A natural rainfall event occurred 10 days after the experiment, preventing further monitoring. Accounting for this “missing” water is necessary for the closing of the water balance. Two possible explanations are presented: storage in the bedrock and deep seepage below the watershed 10 weir.

In the calculation of the storage component of the water balance above, it was assumed that the storage was confined to the soil horizon. When accounting for the discrepancy between the hillslope and watershed recovery, however, we hypothesized that leakage through the bedrock was a significant flowpath. If we assume that flow through the bedrock is significant, and further assume that the bedrock was unsaturated prior to irrigation (reasonable due to 10+ days of antecedent drainage time before irrigation), it is reasonable to argue that increased water storage during the irrigation experiment was non-trivial. Storage in the bedrock that would drain and contribute to WS10 discharge, but not lateral subsurface flow (as that was dominated by shallow flow in the soil profile), would account for the rapid decline in lateral subsurface flow and the slow recession in the WS10 discharge. As the bedrock drains, releasing water into the stream channel, a component that is not measured (bedrock storage) decreases, while a component that is measured (WS10 discharge) remains high. This would lead to an increased mass recovery through time. This explanation does not account for the remaining 20% that was unaccounted for after ten days of drainage, when it would be assumed that the bedrock had returned to the level prior to the experiment (also with 10 days drainage), though the high uncertainty in this value precludes a strong statement.

Deep seepage to groundwater and flow in the bedrock aquifer underneath the WS10 gauging station could account for the remaining water missing from the water balance. Similar to the hypothesized leakage under the hillslope trench due to permeable bedrock, leakage either underneath or around the watershed weir could account for a significant portion of the water balance during this experiment. Recent modeling work in other HJA watersheds has suggested that deep seepage is a significant part of the water balance. Waichler et al. (2005) modeled the nearby Watersheds 1, 2, and 3 at the H. J. Andrews using a distributed conceptual model of hillslope processes (DHSVM), and concluded that evapotranspiration could not account for the differences between measured inputs (P) and outflows (Q). The discrepancy was attributed to deep seepage bypass flow past the catchment gauging. This bypass was a significant portion, 12%, of the annual water balance, and especially concentrated at the wet, winter months. These estimates are similar to the observed missing water after 10 days drainage.

The possibility that up to 20% of precipitation is being held in potentially large, low permeability bedrock, and another 12% is being transmitted past the weirs via deep bedrock flowpaths has considerable implications for catchment scale water transit times. If water held as groundwater is subsequently draining below weirs at watershed outlets, the isotopic and chemical signature of this old water is likely not being expressed in the stream discharge, resulting in an under-estimate of water age at this scale. Consequently, reported estimates of a mean transit time of ~one year at this site are likely skewed since they only consider surface waters (McGuire et al., 2005). Additionally, the subsurface flow under the WS10 weir would result in an underestimate of the flows from the watershed, as suggested by Waichler et al. (2005).

4.6.2. Relationship of our hillslope irrigation to natural events

Analysis of the 50 year (1954 – 2004) rainfall record at the HJA reveals how the irrigation scheme employed during this experiment compares to natural events at the field site. Winter rainfall is characterized by long, low intensity storms. With storms defined as continuous rainfall with no breaks longer than 24 hours, storms longer than 24 days have a return period of 10 years. Our irrigation intensity averaged 3.8 mm/hr. The return period for an event that averages 3.8 mm/hour for 2 days is 8 years, while the return period for an event averaging 3.8 mm/hour for 4 days is greater than the length of the data record (>50 years). Thus, the likelihood of an event that averages the applied rate for the duration (24 days) is exceedingly unlikely. In fact, the applied water volume (2290 mm) is nearly equivalent to the annual rainfall (2300 mm). The rainfall amount before steady state conditions were achieved (~5 days at 3.8 mm/hr, or 456 mm), was also high, falling in the outer range seen in the field. In the 50 years of gauging, only 14 storms (0.6%) had more than 450 mm total precipitation. On the other hand, for events falling during winter and spring (n = 1184), the average time between events was 2.7 days. From the watershed discharge and soil moisture measurements made in this experiment, this is likely not a long enough drainage time to return to pre-event conditions. This can also be seen in the continuous high baseflow seen in the WS10 discharge throughout the winter. The conditions during the irrigation were similar to those seen after fall wet up, during the longer, low intensity events.

Three of the water balance components are likely overestimated due to the nature of the irrigation application: evapotranspiration, hillslope subsurface flow and catchment discharge. Evapotranspiration was overestimated if compared to the rate occurring during an event, while underestimated in comparison with the annual water budget. While the irrigation occurred in late July through August, when potential transpiration is at its peak, the time of highest natural rainfall is generally when the trees are dormant, and the evapotranspiration drivers are at their minimum. This suggests that the steady state transpiration and CRET rates estimated during this experiment are likely an overestimate of rates seen during the long, late season events observed in the field. However, the fraction of the annual water balance occupied by ET is likely higher than that measured during the irrigation experiment, due to evapotranspiration from water stored in the soil profile during the long, dry summers at the site.

During the experiment, canopy transpiration was a small component of ET, ~10%. This was likely an underestimation of normal summertime conditions. ET rates were assumed to be high, equal to the CRET, due to the very wet conditions near the ground surface during the irrigation experiment. This provided a constant source of water for evaporation during the experiment. Normally, during the late summer, the upper soil dries out, and evaporation is suppressed, especially at the HJA, where the high permeability soils allow for rapid drainage. During this time, when bare soil evaporation is suppressed, transpiration likely remains at a high level, as evidenced by the high transpiration rates at the initiation of the experiment. The large trees on the plot would be expected to draw from either the tightly held water near the surface, or from the wetter areas at depth. During the winter, on the other hand, when the soil is wet and evaporation repression due to soil moisture deficit should not be a factor, the potential rates are lower for both transpiration and bare soil evaporation due to lower radiation and vapor pressure deficit drivers.

Lateral subsurface flow is also likely overestimated in the irrigation experiment, especially for small events, due to the length of time (~5 days) needed for steady state conditions to be reached with regards to lateral subsurface flow. A threshold relationship between precipitation and discharge has been observed at the site. During irrigation, this threshold was quickly reached, leading to higher runoff ratios than would be expected for small storms.

4.6.3. Measurement uncertainty and undermining the science

After propagation from the various sources of measurement, the uncertainty in many of the water balance components was large, up to 20% (WS10 increased discharge) of the total component flux for the duration of the experiment. While perhaps alarming, these are due to real uncertainties in the assumptions and measurements used in the calculations. For instance, the uncertainty in the exponent in the WS10 master recession curve alone was 15%, though this value was chosen to encompass only 80% of the 3 year watershed discharge measurements. To encompass 90%, the increase in the uncertainty of the recession coefficient would have to increase to 50%, with a corresponding increase in overall WS10 discharge uncertainty.

The evaporation measures showed a similarly large uncertainty, up to 43% (Experiment + 5 days drainage). This high uncertainty is due to two factors: measurement uncertainty, and the way we classify the uncertainty. The measurement uncertainty was small for most of the sensors, between 1-5% of the readings. For wind speed, however, the instrument accuracy was within ± 0.2 m/s. Wind speeds on the hillslope were within 0 -0.4 m/s for 94% of the monitoring period. This leads to enormous uncertainty in the evaporation measurements, though the effect was somewhat mitigated by the relative unimportance of wind speed in the final calculations (see appendix for the wind speed contribution to uncertainty).

The other, larger factor in the uncertainty is the treatment of the errors. Since a calibration of the meteorological sensors was not performed in the field, it was impossible to determine whether the sensors were accurate. Therefore, the accuracy of the sensors was taken as the factory level uncertainty, and propagated as systematic error. Systematic errors are propagated as the square of sums, rather than the sum of squares. For a given, aggregation of N measurements,

$$\sum_{n=1}^{N}(\delta x_n)^2 \approx \frac{1}{\sqrt{N}}\left(\sum_{n=1}^{N}\delta x_n\right)^2 \tag{17}$$

where δx_n is the uncertainty in measurement x_n. For the aggregated estimate of evaporation where measurements were taken every 15 minutes, or 96 times daily, 2304 readings were aggregated to determine the total experiment flux. For an instrument with equal systematic and random error then, the aggregated systematic error for the duration of the experiment would be $\sqrt{2304}$ or 48 times larger than the random error. While some systematic errors are unavoidable, such as the wetted area measurements, or the calculation of the background watershed discharge, elimination or reduction of systematic errors should be the focus of experimental design. Random errors, while still a concern, are shown to be a much smaller component of the uncertainty.

4.6.3.1. On undermining field hydrology

Beven (2006b) has been challenged about undermining hydrological science by overemphasizing model uncertainty. Nevertheless, a generally positive response to his paper suggests that a rigorous, honest assessment of model uncertainty is considered a positive development by the scientific hydrological community, despite the concerns it may confuse or discourage shareholders (Andréassian et al., 2007; Hall et al., 2007; Mantovan and Todini, 2006; Montanari, 2007; Sivakumar, 2008). A similar concern might be raised for a rigorous analysis of uncertainty in experimental hydrology, especially in field campaigns, where measurement and process uncertainty have the potential to be large compared to the measurements. Indeed, in this experiment, the uncertainties in the residual of the water balance are of greater magnitude than the residual in some cases, calling into question whether a residual exists at all. Did we measure all of the water and not notice it? How can our measured fluxes have a total uncertainty of over 82,000 liters (82 m^3), or 20% of the application? How does this impact our conclusions (namely that the system responds quickly, that flow through the bedrock is a significant component at the hillslope scale, and that deep seepage and bypass flow through the bedrock may be a significant component at the watershed scale)?

The presented uncertainty, while significant, did not impact our conclusions on some of the observed processes. The dynamics of hillslope and catchment response with respect to storage have little to do with the uncertainty in the measurements, as the uncertainty does not include the possibility that no response occurred. Additionally, the uncertainty in the instantaneous measurements of hillslope and watershed response suggests that the evidence of bedrock flow is robust. The evidence for deep seepage at the watershed scale, and bypass of the weir is tempered, however. The uncertainty in the deep seepage / deep storage (the residual of the water balance) after 10 days drainage encompasses the estimate, which precludes the conclusion that we have strong evidence that deep seepage / deep storage exists. However, the uncertainty also allows for the possibility that deep seepage is a much larger proportion of the water balance (up to 41%) than previously predicted (12%; Waichler et al., 2005).

The rigorous analysis of the uncertainties allows for identification of weaknesses in study design and implementation. While it is unpleasant to identify the weaknesses in one's experiment and to quantify the uncertainties in the results, this analysis allows for a better understanding of the strengths and weaknesses of the work. In this case, the need for onsite calibration of field instrumentation is highlighted, to turn the potentially systematic measurement errors into less significant random errors. The development of better methodology for determining the master recession, perhaps using correlations developed with nearby instrumented catchments, would also serve greatly to reduce the uncertainty in the increase in watershed discharge, and increase the strengths of the observations.

Finally, the explicit presentation of the uncertainty will help in model development and evaluation. As the identification and incorporation of input uncertainty into hydrological models becomes standard practice, the uncertainties in reference data sets will be required. Without this explicit uncertainty analysis, modelers are often required to either completely trust the data sets that they calibrate their models to, or arbitrarily assign an uncertainty bound based on estimates of measurement precision, typical bounds for similar systems, stochastic assignment of errors or the modelers inherent level of trust of the experimentalist. While these methods are all valid for certain circumstances, a quantitative analysis is likely preferred.

4.7. Conclusions

This experiment demonstrated the relative importance of the four water balance components at this scale: soil moisture storage, deep seepage to groundwater, lateral subsurface flow, both parallel and perpendicular to the stream channel, and evapotranspiration. Additionally this experiment highlighted some of the difficulties in closing the water balance at the hillslope scale. Deep seepage to groundwater, which bypassed both the hillslope trench and a gauging station downstream, is a large component of the water balance, accounting for up to 24% of the irrigated water at steady state. While much of this water was later measured at the watershed outlet, the loss to groundwater was estimated at 12% 10 days after the end of the experiment. Measurement of evaporation and transpiration allowed us to determine the maximum possible amount of evapotranspiration and to estimate this deep seepage. In order to quantify the deep seepage component of the water balance more precisely, methods need to be developed to directly measure fluxes through and below the soil profile.

4.8. Acknowledgements

This work was supported through funding from the National Science Foundation (grant DEB 021-8088 to the Long-Term Ecological Research Program at the H. J. Andrews Experimental Forest) and from the Ecosystem Informatics IGERT program. Renee Brooks, Barbara Bond, Kate Lajtha, Bob McKane and Julia Jones provided help in developing the experimental design and research objectives. We thank Matthew Bergen and John Moreau for providing field assistance and John Selker and Sherri Johnson for loan of TDR and meteorological equipment. We especially thank the McKenzie River Ranger District for providing sprinkler water during the experiment, and Kari O'Connell and Cheryl Friesen for coordinating logistics.

4.9. Appendix: Derivation of propagated error formulas

The derivations of the functional propagation of the measurement uncertainty for the five components of the water balance are below. For notational simplicity, the systematic and random errors are not differentiated when the functional response to the errors is the same, such as in the soil moisture, meteorological and transpiration measurements. For the hillslope and watershed discharge measurements, the two are differentiated. The measurement uncertainty in variable x is expressed as δx.

4.9.1. Increase in lateral subsurface flow

Lateral subsurface flow, as measured at the hillslope trench, appeared to be steady at a constant rate of 12 L/hr at the beginning of the experiment. The increase in lateral subsurface flow was the expressed as:

$$Q_{hill,\exp} = Q_{hill} - Q_o \tag{A1}$$

where $Q_{hill,\exp}$ is the increase in lateral subsurface flow due to the experiment, Q_{hill} is the instantaneous lateral subsurface flow measurement, and Q_o is the background, pre-event discharge. The uncertainty in $Q_{hill,\exp}$ is a function of the uncertainty in Q_{hill} and Q_o:

$$\delta Q_{hilll,\exp} = \sqrt{\left(\frac{\partial Q_{hilll,\exp}}{\partial Q_{hill}}\delta Q_{hill}\right)^2 + \left(\frac{\partial Q_{hilll,\exp}}{\partial Q_o}\delta Q_o\right)^2} \tag{A2}$$

where

$$\frac{\partial Q_{hilll,\exp}}{\partial Q_{hill}}\delta Q_{hill} = \frac{\partial\left(Q_{hill} - Q_o\right)}{\partial Q_{hill}}\delta Q_{hill} = \delta Q_{hill} \tag{A3}$$

$$\frac{\partial Q_{hilll,\exp}}{\partial Q_o}\delta Q_o = \frac{\partial\left(Q_{hill} - Q_o\right)}{\partial Q_o}\delta Q_o = -\delta Q_o \tag{A4}$$

The total uncertainty for instantaneous measurements of lateral subsurface flow is then:

$$\delta Q_{hilll,\exp} = \sqrt{\left(\delta Q_{hill}\right)^2 + \left(\delta Q_o\right)^2} \tag{A5}$$

For determining the total increase in lateral subsurface flow over N measurements (i.e. total experimental lateral subsurface flow):

$$\hat{Q}_{hill,\exp} = \sum_{n=1}^{N} Q_{hill,\exp,n} = \sum_{n=1}^{N}\left(Q_{hill,n} - Q_o\right) \tag{A6}$$

Since the uncertainties in the individual measurements are assumed to be random, and the uncertainty in the background discharge is considered systematic, the total uncertainty for summed measurements of lateral subsurface flow is then:

$$\delta\hat{Q}_{hill,\exp} = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial Q_{hill,\exp,n}}{\partial Q_{hill,n}}\delta Q_{hill,n}\right)^2 + \left(\sum_{n=1}^{N}\frac{\partial Q_{hill,\exp,n}}{\partial Q_o}\delta Q_o\right)^2} \tag{A7}$$

or

$$\delta\hat{Q}_{hill,\exp} = \sqrt{\sum_{n=1}^{N}\left(\delta Q_{hill,n}\right)^2 + \left(N\delta Q_o\right)^2} \tag{A8}$$

4.9.2. Increase in watershed discharge

The increase in watershed discharge is expressed as

$$Q_{WS,\exp} = Q_{WS} - Q_o e^{-t/T_c} \tag{A9}$$

where $Q_{WS,\exp}$ is the increase in watershed discharge due to the experiment, Q_{WS} is the instantaneous watershed discharge measurement, and $Q_o e^{-t/T_c}$ is the expected watershed discharge derived from the mater recession analysis. The uncertainty in instantaneous measurements of $Q_{WS,\exp}$ is a function of uncertainty in Q_{WS}, Q_o and T_c.

$$\delta Q_{WS,\exp} = \sqrt{\left(\frac{\partial Q_{ws,\exp}}{\partial Q_{WS}}\delta Q_{WS}\right)^2 + \left(\frac{\partial Q_{ws,\exp}}{\partial Q_o}\delta Q_o\right)^2 + \left(\frac{\partial Q_{ws,\exp}}{\partial T_c}\delta T_c\right)^2} \tag{A10}$$

where

$$\frac{\partial Q_{ws,\exp}}{\partial Q_{WS}}\delta Q_{WS} = \frac{\partial\left(Q_{WS} - Q_o e^{-t/T_c}\right)}{\partial Q_{WS}}\delta Q_{WS} = \delta Q_{WS} \tag{A11}$$

$$\frac{\partial Q_{ws,\exp}}{\partial Q_o}\delta Q_o = \frac{\partial\left(Q_{WS} - Q_o e^{-t/T_c}\right)}{\partial Q_o}\delta Q_o = -e^{-t/T_c}\delta Q_o \tag{A12}$$

$$\frac{\partial Q_{ws,\exp}}{\partial T_c}\delta T_c = \frac{\partial\left(Q_{WS} - Q_o e^{-t/T_c}\right)}{\partial T_c}\delta T_c = \frac{Q_o t}{T_c^2} e^{-t/T_c}\delta T_c \tag{A13}$$

The total uncertainty for instantaneous measurements of watershed discharge is then:

$$\delta Q_{WS,\exp} = \sqrt{\left(\delta Q_{WS}\right)^2 + \left(-e^{-t/T_c}\delta Q_o\right)^2 + \left(\frac{Q_o t}{T_c^2} e^{-t/T_c}\delta T_c\right)^2} \tag{A14}$$

For determining the total increase in lateral subsurface flow over N measurements:

$$Q_{WS,\exp} = \sum_{n=1}^{N} Q_{WS,\exp,n} = \sum_{n=1}^{N}\left(Q_{WS,n} - Q_o e^{-t/T_c}\right) \tag{A15}$$

Since the uncertainties in the individual measurements are assumed to be random, and the uncertainty in the background discharge is considered systematic, a function of the parameters Q_o and T_c, the total uncertainty for summed measurements of lateral subsurface flow is then:

$$\delta \hat{Q}_{WS,\exp} = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial Q_{WS,\exp,n}}{\partial Q_{WS,n}}\delta Q_{WS,n}\right)^2 + \left(\frac{\partial \sum_{t=1}^{T} Q_{WS,\exp,n}}{\partial Q_o}\delta Q_o\right)^2 + \left(\frac{\partial \sum_{t=1}^{T} Q_{WS,\exp,t}}{\partial Q_o}\delta T_c\right)^2} \qquad \text{(A16)}$$

or

$$\delta \hat{Q}_{WS,\exp} = \sqrt{\sum_{n=1}^{N}\left(\delta Q_{WS,n}\right)^2 + \left(\sum_{t=1}^{T} e^{t/T_c}\delta Q_o\right)^2 + \left(\sum_{t=1}^{T} Q_o t e^{t/T_c^2}\delta T_c\right)^2} \qquad \text{(A17)}$$

4.9.3. Transpiration

Stand level transpiration is the sum of the 9 instrumented trees. Transpiration is measured from each tree using the empirical formula (Granier, 1987):

$$Q_n = 119E - 6sa_n\left(\frac{\Delta T_{M,n} - \Delta T_n}{\Delta T_n}\right)^{1.231} \qquad \text{(A18)}$$

where sapwood area, sa, the daily maximum temperature difference, ΔTM and the instantaneous temperature difference, ΔT, for n trees. Stand level transpiration is then

$$Q_T = \sum_{n=1}^{N} Q_{T,n} \qquad \text{(A19)}$$

All variables exhibit random and systematic uncertainty. Since the probes were not calibrated onsite, the random and systematic uncertainties are equivalent. The error in Q_T is then

$$\delta Q_T = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial Q_{T,n}}{\partial s_{A,n}}\delta sa_n\right)^2 + \sum_{n=1}^{N}\left(\frac{\partial Q_{T,n}}{\partial \Delta T_n}\delta \Delta T_n\right)^2 + \sum_{n=1}^{N}\left(\frac{\partial Q_{T,n}}{\partial \Delta T_{M,n}}\delta \Delta T_{M,n}\right)^2} \qquad \text{(A20)}$$

where

$$\frac{\partial Q_{T,n}}{\partial sa_n}\delta sa_n = 119E - 6\left(\frac{\Delta T_{M,n} - \Delta T_n}{\Delta T_n}\right)^{1.231}\delta sa_n \qquad \text{(A21)}$$

$$\frac{\partial Q_{T,n}}{\partial \Delta T_M}\delta \Delta T_M = \frac{119E - 6(1.231)sa_n}{\Delta T_n}\left(\frac{\Delta T_{M,n} - \Delta T_n}{\Delta T_n}\right)^{0.231}\delta \Delta T_M \qquad \text{(A22)}$$

$$\frac{\partial Q_{T,n}}{\partial \Delta T_n}\delta\Delta T_n = -\left(\frac{119E-6(1.231)sa_n\Delta T_M}{\Delta T_n^2}\right)\left(\frac{\Delta T_{M,n}-\Delta T_n}{\Delta T_n}\right)^{0.231}\delta\Delta T_n \quad \text{(A23)}$$

Without onsite calibration of temperature sensors or sapflow estimates, we must assume that the uncertainties in temperature measurements are stationary. Sapwood depth measurements are by definition stationary, as the values used are repeated for all sapflow estimates. The uncertainties between trees, however, are considered random. Therefore, the aggregated stand level transpiration uncertainty is:

$$\delta\hat{Q}_T = 119E-6\sqrt{\sum_{n=1}^{N}\left(\sum_{t=1}^{T}\frac{\partial Q_{T,n}}{\partial sa_n}\delta sa_n\right)^2+\sum_{n=1}^{N}\left(\sum_{t=1}^{T}\frac{\partial Q_{T,n}}{\partial \Delta T_n}\delta\Delta T_n\right)^2+\sum_{n=1}^{N}\left(\sum_{t=1}^{T}\frac{\partial Q_{T,n}}{\partial \Delta T_{M,n}}\delta\Delta T_{M,n}\right)^2} \quad \text{(A24)}$$

4.9.4. CRET

Canopy reference evapotranspiration was calculated using the Penman-Monteith equation and is dependent on five measurements: Net radiation, *Rn*, relative humidity, *RH*, air temperature *T*, and wind speed, *u*, and two estimated parameters *a*, the proportion of incoming net radiation reflected from the ground surface and r_c, the canopy resistance:

$$CRET = \frac{\Delta\left(R_n - G\right)+\rho c_p\frac{VPD}{r_a}}{\lambda\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)} = \frac{\Delta\left(T\right)R_n\left(1-a\right)+\rho c_p\frac{RHe_s\left(T\right)}{r_a\left(u\right)}}{\lambda\left(\Delta\left(T\right)+\gamma\left(1+\frac{r_c}{r_a\left(u\right)}\right)\right)} \quad \text{(A25)}$$

where $\Delta(T)$, the derivative of the saturated vapor pressure curve, is an exponential function of temperature, $e_s(T)$, the saturation vapor pressure curve, is an exponential function of temperature, and r_a is a function of wind speed. Additionally, G is assumed to be aR_n, where a is a cefficient with uncertainty δa. The uncertainty in CRET is then:

$$\delta CRET = \sqrt{\left(\frac{\partial CRET}{\partial R_n}\delta R_n\right)^2+\left(\frac{\partial CRET}{\partial RH}\delta RH\right)^2+\left(\frac{\partial CRET}{\partial T}\delta T\right)^2+\left(\frac{\partial CRET}{\partial u}\delta u\right)^2+\left(\frac{\partial CRET}{\partial a}\delta a\right)^2+\left(\frac{\partial CRET}{\partial r_c}\delta r_c\right)^2} \quad \text{(A26)}$$

where

$$\frac{\partial CRET}{\partial R_n}\delta R_n = \frac{(1-a)\Delta}{\lambda\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)}\delta R_n \tag{A27}$$

$$\frac{\partial CRET}{\partial RH}\delta RH = \frac{\rho c_p \frac{es(T)}{r_a}}{\lambda\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)}\delta RH \tag{A28}$$

$$\frac{\partial CRET}{\partial T}\delta T = \frac{\left(\Delta'(R_n-G)+\frac{\rho c_p RH}{r_a}\Delta\right)\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)-\Delta'\left(\Delta(R_n-G)+\frac{\rho c_p RH}{r_a}e_s(T)\right)}{\lambda\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)^2}\delta T \tag{A29}$$

where

$$e_s(T) = e_s(T^*)e^{A\frac{T}{T-T'}} \tag{A30}$$

$$e_s(T)' = \Delta = e_s(T^*)\left(A\frac{T'}{(T-T')^2}\right)e^{A\frac{T}{T-T'}} \tag{A31}$$

$$e_s(T)'' = \Delta' = e_s(T^*)T'e^{A\frac{T}{T-T'}}\left(\frac{AT'}{(T-T')^3}-\frac{2}{(T-T')^4}\right) \tag{A32}$$

where $e_s(T)$, T^* and T' are all constants.

$$\frac{\partial CRET}{\partial u}\delta u = \frac{\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)\left(\rho c_p\frac{VPD}{ur_a}\right)-\left(\Delta(R_n-G)+\rho c_p\frac{VPD}{r_a}\right)\left(\frac{\gamma r_c}{ur_a}\right)}{\lambda\left(\Delta+\gamma\left(1+\frac{r_c}{r_a}\right)\right)^2}\delta u \tag{A33}$$

where

$$r_a = \frac{\ln\left((z-d)/z_o\right)}{ku} \tag{A34}$$

$$\frac{\partial CRET}{\partial a}\delta a = \frac{-R_n\Delta}{\lambda\left(\Delta + \gamma\left(1+\frac{r_c}{r_a}\right)\right)}\delta a \tag{A35}$$

$$\frac{\partial CRET}{\partial r_c}\delta r_c = -\frac{\frac{\gamma}{r_a}\left(\Delta\left(R_n - G\right) + \rho c_p \frac{VPD}{r_a}\right)}{\lambda\left(\Delta + \gamma\left(1+\frac{r_c}{r_a}\right)\right)^2}\delta r_c \tag{A36}$$

Since there was no onsite calibration, all uncertainty is treated as systematic.

4.9.5. Soil Moisture

Soil Moisture was measured in 57 TDR probe sections, and the relative change in soil moisture storage (*ΔS*) is a weighted average of the individual measurements (S_n) taken at times i and j:

$$\Delta S_{i,j} = \overline{z}A\frac{1}{\sum_{n=1}^{N} z_n}\sum_{n=1}^{N} S_{n,i}z_n - \overline{z}A\frac{1}{\sum_{n=1}^{N} z_n}\sum_{n=1}^{N} S_{n,j}z_n = \overline{z}A\frac{1}{\sum_{n=1}^{N} z_n}\sum_{n=1}^{N} z_n(S_{n,i} - S_{n,j}) \tag{A37}$$

where $\overline{z}$ is the average soil depth of the wetted area, A is the wetted area, and z_n is the depth sampled by the probe section. No uncertainty is assumed in the sampling depth (z_n). The uncertainty in the individual readings of soil moisture storage are assumed to be random. The error δΔS is then:

$$\delta\Delta S_{i,j} = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial\Delta S_{i,j}}{\partial S_{n,i}}\delta S_{n,i}\right)^2 + \sum_{n=1}^{N}\left(\frac{\partial\Delta S_{i,j}}{\partial S_{n,j}}\delta S_{n,j}\right)^2 + \left(\frac{\partial\Delta S_{i,j}}{\partial\overline{z}}\delta\overline{z}\right)} \tag{A38}$$

where

$$\frac{\partial\Delta S_{i,j}}{\partial S_{n,i}} = \frac{\partial\Delta S_{i,j}}{\partial S_{n,j}} = \overline{z}A\frac{1}{\sum_{n=1}^{N} z_n}z_n \tag{A39}$$

and

$$\frac{\partial\Delta S_{i,j}}{\partial\overline{z}} = A\frac{1}{\sum_{n=1}^{N} z_n}\sum_{n=1}^{N} z_n(S_{n,i} - S_{n,j}) \tag{A40}$$

The error in the change in soil moisture profile soil moisture is then:

$$\delta\Delta S_{i,j} = \sqrt{\sum_{n=1}^{N}\left(\bar{z}A\frac{1}{\sum_{n=1}^{N} z_n} z_n \delta S_{n,i}\right)^2 + \sum_{n=1}^{N}\left(\bar{z}A\frac{1}{\sum_{n=1}^{N} z_n} z_n \delta S_{n,j}\right)^2 + \left(A\frac{1}{\sum_{n=1}^{N} z_n}\sum_{n=1}^{N} z_n (S_{n,i} - S_{n,j})\delta\bar{z}\right)^2} \qquad \text{(A41)}$$

4.10. References

Andréassian, V. et al., 2007. What is really undermining hydrologic science today? Hydrological Processes, 21(20): 2819-2822.

Bárdossy, A., 1996. The use of fuzzy rules for the description of elements of the hydrological cycle. Ecological Modelling, 85(1): 59-65.

Barnard, H.R. et al., submitted. Examining hillslope transpiration controls of diel sub-surface flow using a steady-state irrigation experiment. Ecohydrology.

Beven, K., 2001. On hypothesis testing in hydrology. Hydrological Processes, 15(9): 1655-1657.

Beven, K., 2002. 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., 2006a. Searching for the Holy Grail of scientific hydrology: Q_t=(S,R,Δt) as closure. Hydrology and Earth System Sciences, 10(5): 609-618.

Beven, K. and Binley, A., 1992. Future of distributed models: model calibration and uncertainty prediction. Hydrological Processes, 6(3): 279-298.

Beven, K.J., 2006b. On undermining the science? Hydrological Processes, 20: 3141-3146.

Bonell, M. and Gilmour, D.A., 1978. The development of overland flow in a tropical rainforest catchment. Journal of Hydrology, 39(3-4): 365-382.

Brooks, E.S., Boll, J. and McDaniel, P.A., 2004. A hillslope-scale experiment to measure lateral saturated hydraulic conductivity. Water Resources Research, 40.

Chapman, T., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes, 13(5): 701-714.

Clarke, R.T., 1999. Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. Journal of Hydrology, 222(1-4): 185-190.

Domec, J.-C., Meinzer, F.C., Gartner, B.L. and Woodruff, D., 2006. Transpiration-induced axial and radial tension gradients in trunks of Douglas-fir trees. Tree Physiology, 26(3): 275-284.

Eberhardt, L.L. and Thomas, J.M., 1991. Designing environmental field studies. Ecological Monographs, 61(1): 53-73.

Ewen, J. and Birkinshaw, S.J., 2007. Lumped hysteretic model for subsurface stormflow developed using downward approach. Hydrological Processes, 21: 1496-1505.

Freer, J. et al., 2002. The role of bedrock topography on subsurface storm flow. Water Resources Research, 38(12).

Gee, G.W. et al., 1994. Variations in Water Balance and Recharge Potential at Three Western Desert Sites. Soil Science Society of America Journal, 58(1): 63-72.

Graham, C.B., McDonnell, J.J. and Woods, R., in review. Hillslope threshold response to storm rainfall: (1) A field based forensic approach. Journal of Hydrology.
Granier, A., 1987. Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. Tree Physiology, 3: 309-320.
Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1998. Towards improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.
Hall, J., O'Connell, E. and Ewen, J., 2007. On not undermining the science: coherence, validation and expertise. Discussion of Invited Commentary by Keith Beven Hydrological Processes, 20, 3141-3146 (2006). Hydrological Processes, 21: 985-988.
Harr, R.D., 1977. Water flux in soil and subsoil on a steep forested slope. Journal of Hydrology, 33: 37-58.
Hornberger, G.M., Germann, P.F. and Beven, K.J., 1991. Throughflow and solute transport in an isolated sloping soil block in a forested catchment. Journal of Hydrology, 124: 81-99.
James, M.E., 1978. Rock weathering in the central western Cascades. M.S. Thesis, University of Oregon, Eugene, 119 pp.
Jaynes, D.B., 1990. Soil water hysteresis: models and implications. In: M.G. Anderson and T.P. Burt (Editors), Process Studies in Hillslope Hydrology, Chichester, pp. 93–126.
Katsura, S.y., Kosugi, K., Mizutani, T., Okunaka, S. and Mizuyama, T., 2008. Effects of bedrock groundwater on spatial and temporal variations in soil mantle groundwater in a steep granitic headwater catchment. Water Resources Research, 44.
Katsuyama, M., Ohte, N. and Kabeya, N., 2005. Effects of bedrock permeability on hillslope and riparian groundwater dynamics in a weathered granite catchment. Water Resources Research, 41.
Kavetski, D., Franks, S.W. and Kuczera, G., 2002. Confronting input uncertainty in environmental modelling. In: Q. Duan, H.V. Gupta, S. Sorooshian, A.N. Rousseau and R. Turcotte (Editors), In Calibration of Watershed Models. AGU Books, Washington, WA, pp. 49-68.
Kendall, K.A., Shanley, J.B. and McDonnell, J.J., 1999. A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt. Journal of Hydrology, 219(3-4): 188-205.
Kuczera, G. and Mroczkowski, M., 1998. Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. Water Resources Research, 34: 1481-1489.
Mantovan, P. and Todini, E., 2006. Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology. Journal of Hydrology.
McDonnell, J.J., 1990. A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11): 2821-2832.
McDonnell, J.J. et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, 43.

McGuire, K.J., McDonnell, J.J. and Weiler, M.H., 2007. Integrating tracer experiments with modeling to infer water transit times. Advances in Water Resources, 30: 824-837.
McGuire, K.J. et al., 2005. The role of topography on catchment-scale water residence time. Water Resources Research, 41.
McIntosh, J., McDonnell, J.J. and Peters, N.E., 1999. Tracer and hydrometric study of preferential flow in large undisturbed soil cores from the Georgia Piedmont, USA. Hydrological Processes, 13: 139-155.
Moncrieff, J.B., Malhi, Y. and Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Global Change Biology, 2(3): 231-240.
Montanari, A., 2007. What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. Hydrological Processes, 21(6): 841-845.
Monteith, J.L. and Unsworth, M.H., 2008. Principles of Environmental Physics. Academic Press, London, 418 pp.
Montgomery, D.R. et al., 1997. Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. Water Resources Research, 33(1): 91-109.
Moore, G.W., 2004. Ph. D. Thesis, Oregon State University, Corvallis, OR.
Moore, G.W., Bond, B.J., Jones, J.A., Phillips, N. and Meinzer, F.C., 2004. Structural and compositional controls on transpiration in 40- and 450-year-old riparian forests in western Oregon, USA. Tree Physiology, 24: 481-491.
Mosley, M.P., 1979. Streamflow generation in a forested watershed. Water Resources Research, 15: 795-806.
Murray, F.W., 1967. On the computation of saturation vapor pressure. Journal of Applied Meteorology, 6: 203-204.
Onda, Y., Komatsu, Y., Tsujimura, M. and Fujihara, J.-i., 2001. The role of subsurface runoff through bedrock on storm flow generation. Hydrological Processes, 15: 1693-1706.
Özelkan, E.C. and Duckstein, L., 2001. Fuzzy conceptual rainfall-runoff models. Journal of Hydrology, 253(1-4): 41-68.
Peters, D.L., Buttle, J.M., Taylor, C.H. and LaZerte, B.D., 1995. Runoff production in a forested, shallow soil, Canadian Shield basin. Water Resources Research, 31(5): 1291-1304.
Ranken, D.W., 1974. Hydrologic properties of soil and subsoil on a steep, forested slope. M.S. Thesis, Oregon State University, Corvallis, 114 pp.
Rasmussen, T.C., Baldwin, R.H., Dowd, J.F. and Williams, A.G., 2000. Tracer vs. pressure wave velocities through unsaturated saprolite. Soil Science Society of America Journal, 64: 75-85.
Ridolfi, L., D'Odorico, P., Porporato, A. and Rodriguez-Iturbe, I., 2003. Stochastic soil moisture dynamics along a hillslope. Journal of Hydrology, 272(1-4): 264-275.
Rodriguez-Iturbe, I., 2000. Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics. Water Resources Research, 36(1): 3-9.
Scherrer, S., Naef, F., Faeh, A., O. and Cordery, I., 2007. Formation of runoff at the hillslope scale during intense precipitation. Hydrology and Earth System Sciences, 11: 907-922.

Seibert, J., Bishop, K., Rodhe, A. and McDonnell, J.J., 2003. Groundwater dynamics along a hillslope: a test of the steady state hypothesis. Water Resources Research, 39(1): 1014, doi:10.1029/2002WR001404.

Seibert, J. and McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. Water Resources Research, 38(11).

Sivakumar, B., 2008. Undermining the science or undermining Nature? Hydrological Processes, 22(6): 893-897.

Stednick, J.D., 1996. Monitoring the effects of timber harvest on annual water yield. Journal of Hydrology, 176(1-4): 79-95.

Sujono, J., Shikasho, S. and Hiramatsu, K., 2004. A comparison of techniques for hydrograph recession analysis. Hydrological Processes, 18(3): 403-413.

Swanson, F.J. and James, M.E., 1975. Geology and geomorphology of the H.J. Andrews Experimental Forest, western Cascades, Oregon. Res. Pap. PNW-188, U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, OR.

Tan, C.S. and Black, T.A., 1976. Factors affecting the canopy resistance of a Douglas-fir forest. Boundary-Layer Meteorology, 10(4).

Taylor, J.R., 1997. An Introduction to Error Analysis. University Science Books, Sausalito, California, 328 pp.

Tetens, O., 1930. Uber einige meteorologische begriffe. Z. Geophys, 6: 203-204.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resources Research, 42.

Tromp-van Meerveld, H.J., Peters, N.E. and McDonnell, J.J., 2006. 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.

van Verseveld, W., 2006. Hydro-biogeochemical coupling at the hillslope and catchment scale. Ph. D. Thesis, Oregon State University, Corvallis, OR.

Vrugt, J.A., Dekker, S.C. and Bouten, W., 2003. Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage. Water Resources Research, 39(9).

Wagener, T., 2003. Evaluation of catchment models. Hydrological Processes, 17(16): 3375-3378.

Waichler, S.R., Wemple, B.C. and Wigmosta, M.S., 2005. Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. Hydrological Processes, 19: 3177-3199.

Western, A.W., Bloeschl, G. and Grayson, R.B., 1998. Geostatistical characterisation of soil moisture patterns in the Tarrawarra catchment. Journal of Hydrology, 205(1-2): 20-37.

Winter, T.C., 1981. Uncertainties in estimating the water balance of lakes. Journal of the American Water Resources Association, 17(1): 82-115.

Woods, R. and Rowe, L., 1996. The changing spatial variability of subsurface flow across a hillside. Journal of Hydrology (New Zealand), 35(1): 51-86.

Woods, R. and Rowe, L., 1997. Reply to "Comment on "The changing spatial variability of subsurface flow across a hillside" by Ross Woods and Lindsay Rowe". Journal of Hydrology (New Zealand), 36(1): 51-86.

4.11. Tables

Table 4.1 Measurands used in calculation of water balance components and their estimated random and systematic errors. The sources of the uncertainty estimates are either factory reported accuracy and precision or our estimates of measurement uncertainty.

Water Balance Component	Measurand	Random Uncertainty	Systematic Uncertainty
Precipitation	Precipitation rate		± 1 mm/hr (std dev rainfall rate)
	Rainfall volume	± 1 ml (estimate)	
	Wetted area (A)	± 8 m^2 (estimate)	± 8 m^2 (estimate)
Lateral subsurface flow	Stage (s)	± 0.25 mm (factory)	± 0.0034 L/s (calibration)
WS10 discharge	Stage (s)	± 0.3 mm (factory)	± 3.6% (calibration)
Transpiration	Temperature (T)	± 0.2 C (estimate)	± 0.2 C (estimate)
	Sapwood depth (sa)	± 1 mm (estimate)	± 1 mm (estimate)
CRET	Temperature (T)	± 0.4 C (factory)	± 0.4 C (factory)
	Wind speed (u)	± 0.2 m/s (factory)	± 0.2 m/s (factory)
	Ground heat flux coefficient (a)	± 100% (estimate)	± 100% (estimate)
	Incoming net radiation (R_n)	± 6% (factory)	± 6% (factory)
	Relative humidity (RH)	± 2-3% (factory)	± 2-3% (factory)
	Wetted area (A)	± 8 m^2 (estimate)	± 8 m^2 (estimate)
Soil Moisture	Volumetric water content (S)	± 3% (factory)	± 3% (factory)
	Wetted area (A)	± 8 m^2 (estimate)	± 8 m^2 (estimate)

Table 4.2 Water balance components with propagated uncertainty.

Water Balance Component	Steady State (L/hr)	Entire Irrigation Experiment (L)	Irrigation + 5 days (L)	Irrigation + 10 Days (L)
Irrigation	659 ± 33	394,000 ± 19,700	394,000 ± 19,700	394,000 ± 19,700
Hill	284 ± 20	102,543 ± 7,451	106,156 ± 8,979	107,760 ± 10,507
WS10	461 ± 115	252,125 ± 48,035	275,523 ± 56,935	295,781 ± 65,578
Transpiration	9 ± 1	5,448 ± 343	6,456 ± 409	7,318 ± 470
CRET	50 ± 21	30,055 ± 12,692	35,767 ± 15,170	41,114 ± 17,721
Δ Storage	0	25,837 ± 1,565	15,718 ± 1,560	11,438 ± 1,559
P-WS10-E-ΔS	148 ± 121	85,983 ± 53,469	66,992 ± 62,147	45,667 ± 70,746

4.12. Figures

Figure 4.1 Map of study site with outline of irrigated area. 24 TDR rods, meteorological station and instrumented trees are labeled.

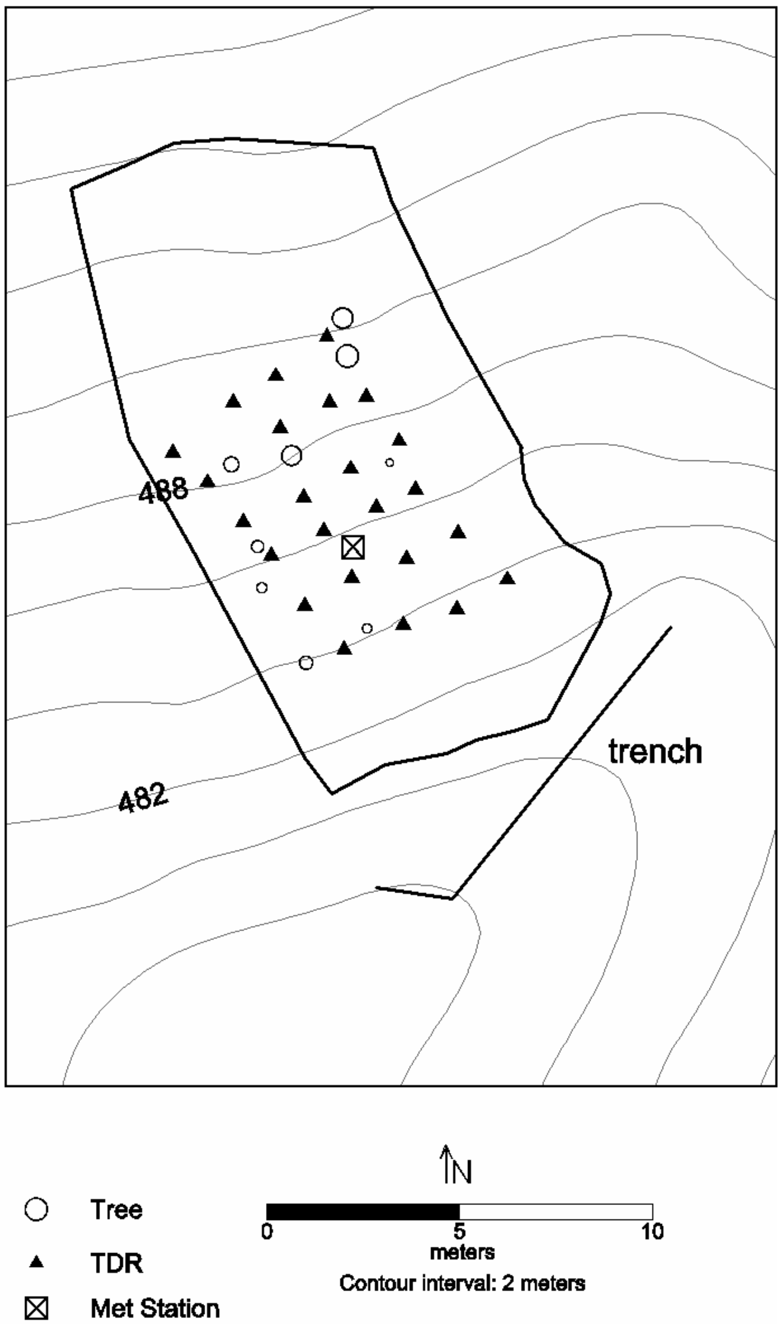


Figure 4.2 WS10 2002 – 2004 summer recession and calibrated master recession. Daily average discharge was used for calibration, and periods where rainfall diverted the discharge from the natural recession were eliminated.

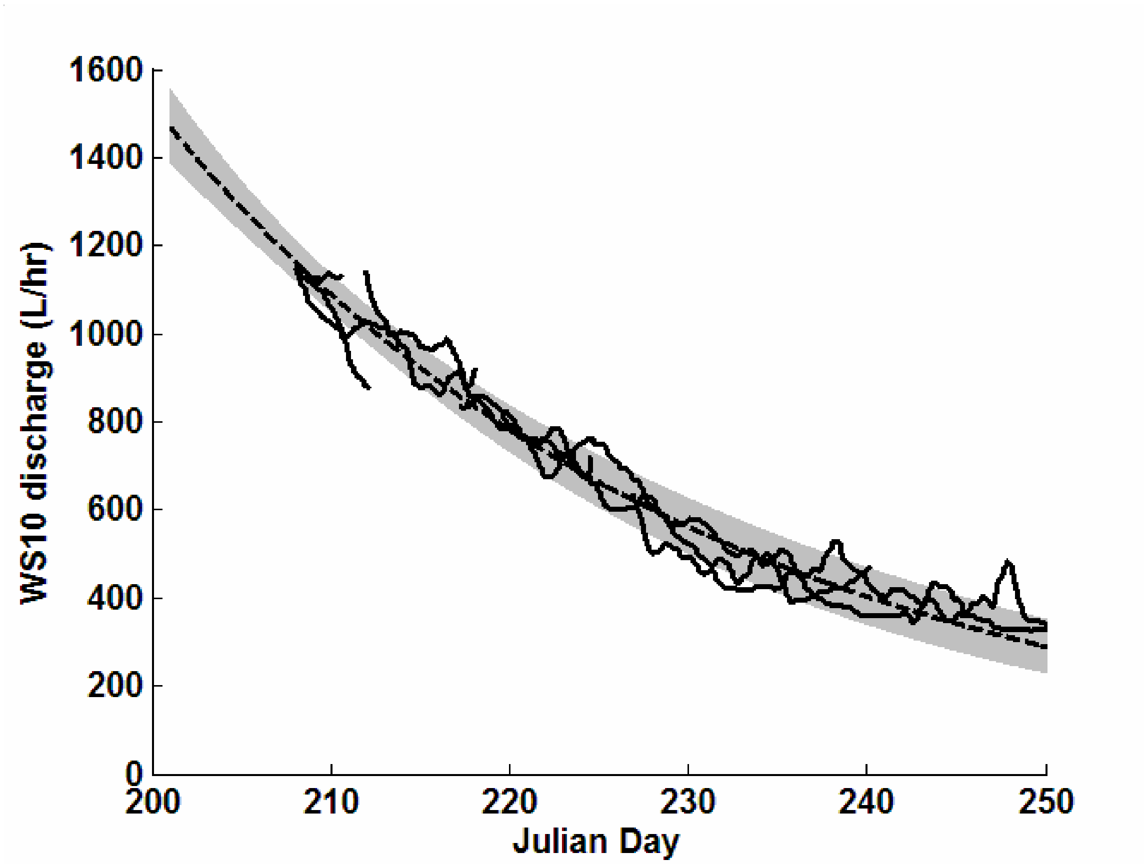


Figure 4.3 Precipitation rates from 72 cups. A weighted average was used to determine the experiment irrigation rate.

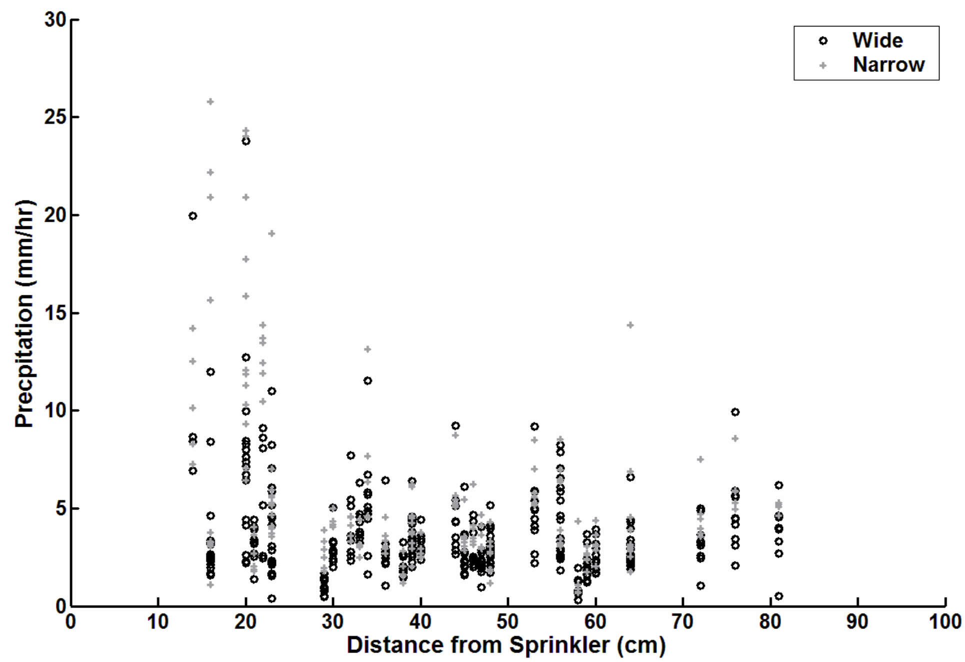


Figure 4.4 Irrigation and lateral subsurface flow (black) with uncertainty bounds (grey) measured at the hillslope trench. Dashed line is background flow rate.

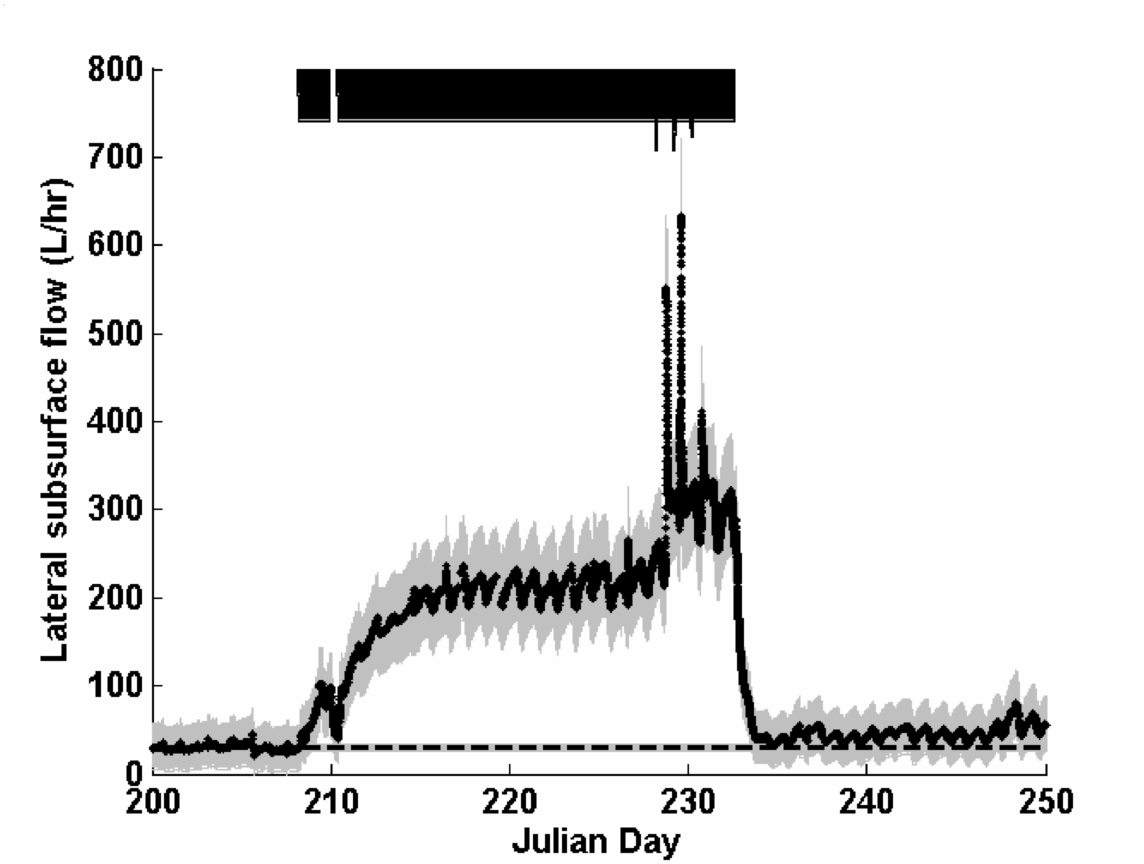


Figure 4.5 Irrigation, WS10 discharge and master recession, with uncertainty bounds. Note increased uncertainty at late time as irrigation progresses.

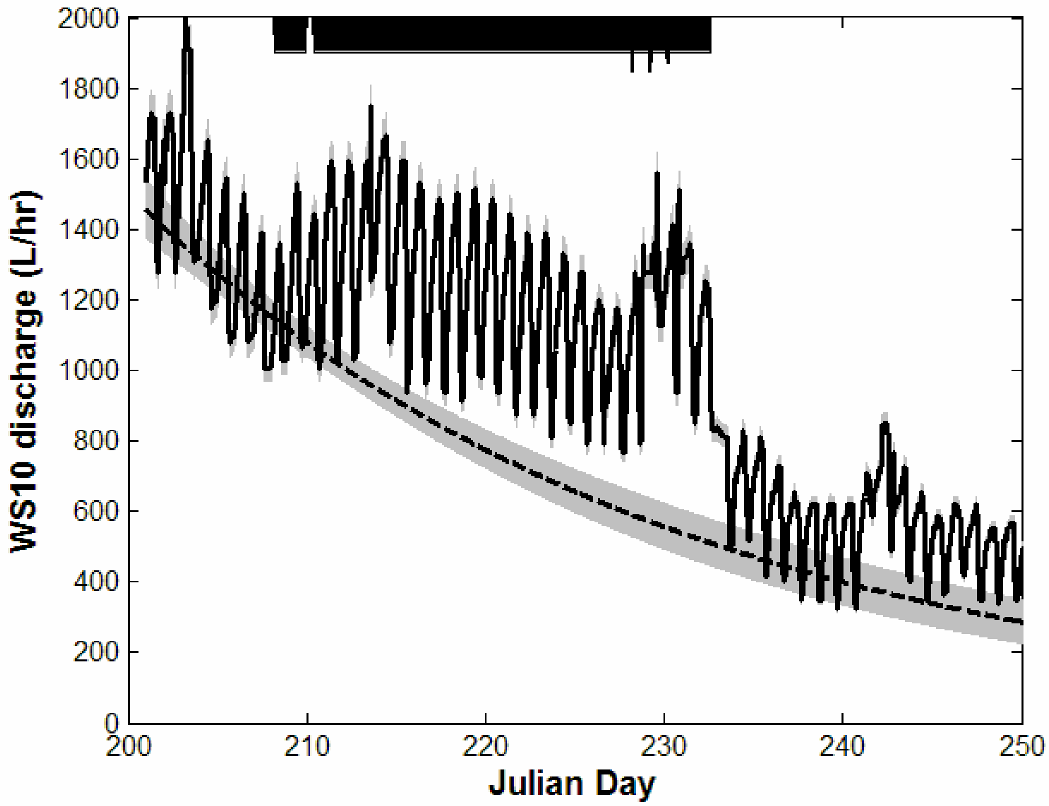


Figure 4.6 Transpiration with uncertainty bounds measured from 9 trees on site for duration of experiment (a) and during steady state (b). Transpiration declines throughout experiment.

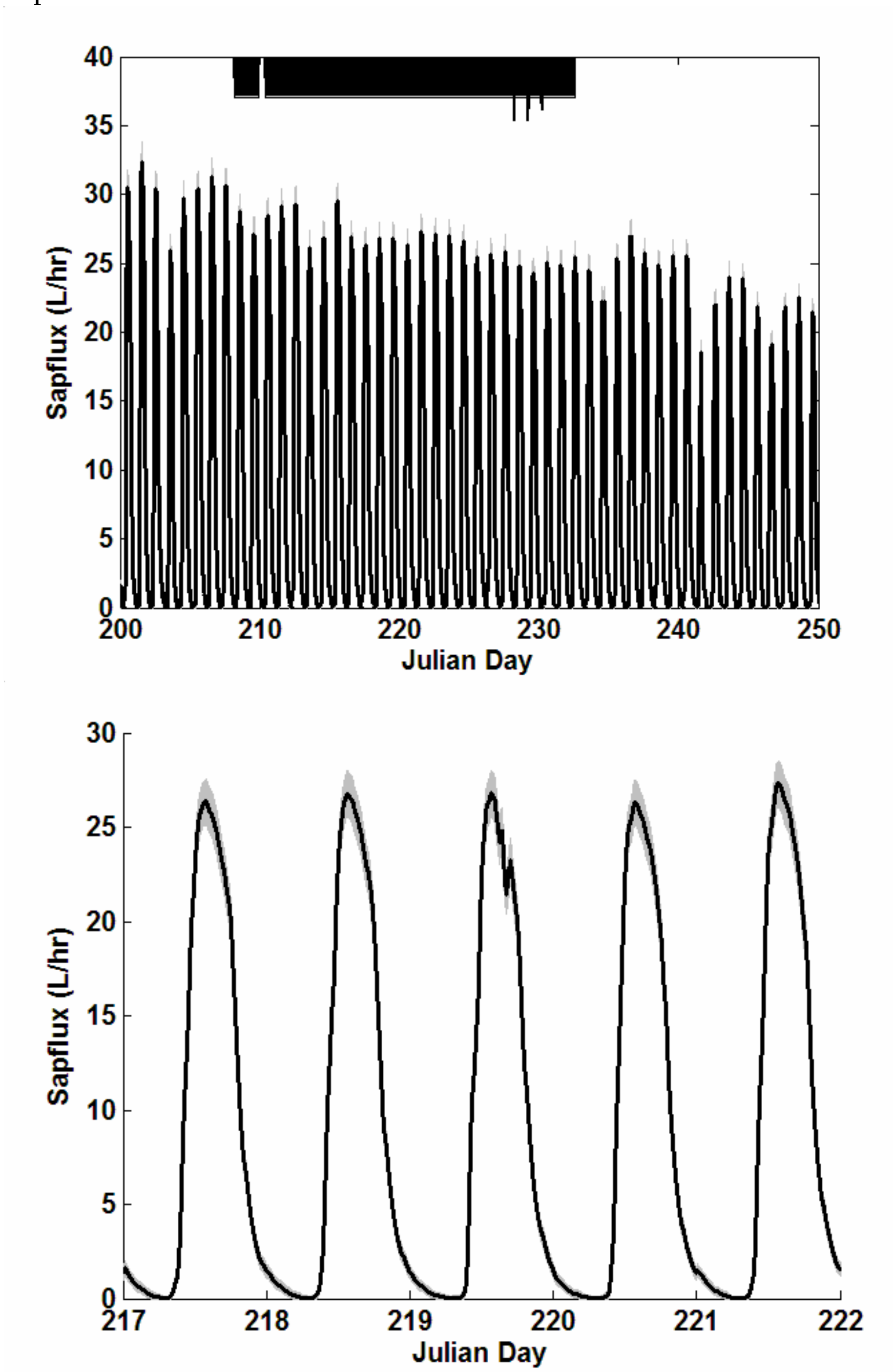
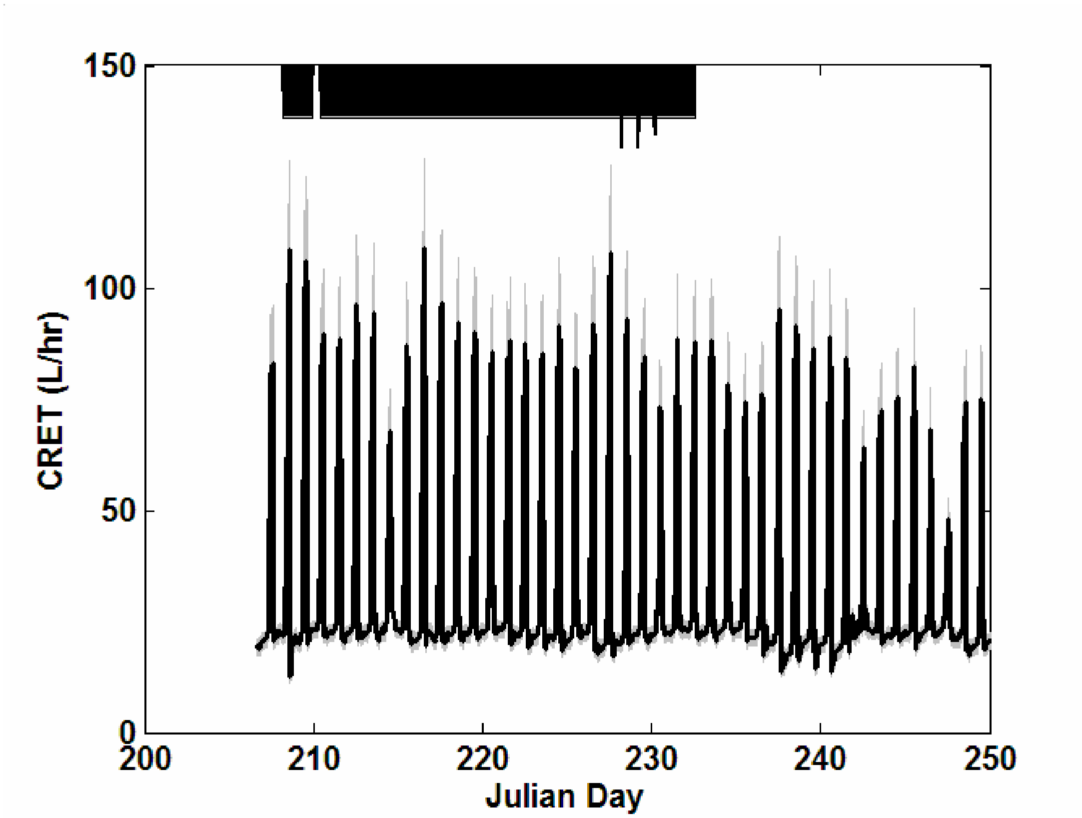


Figure 4.7 CRET with uncertainty bounds for experiment duration (a) and during steady state conditions. CRET was relatively insensitive to irrigation.



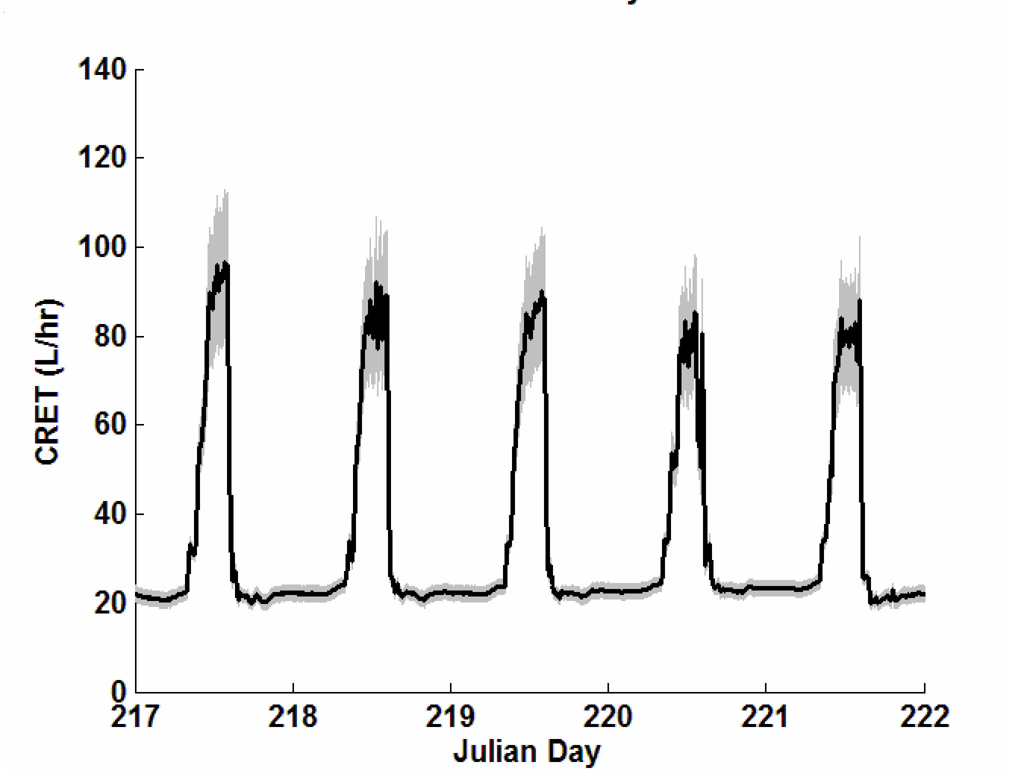
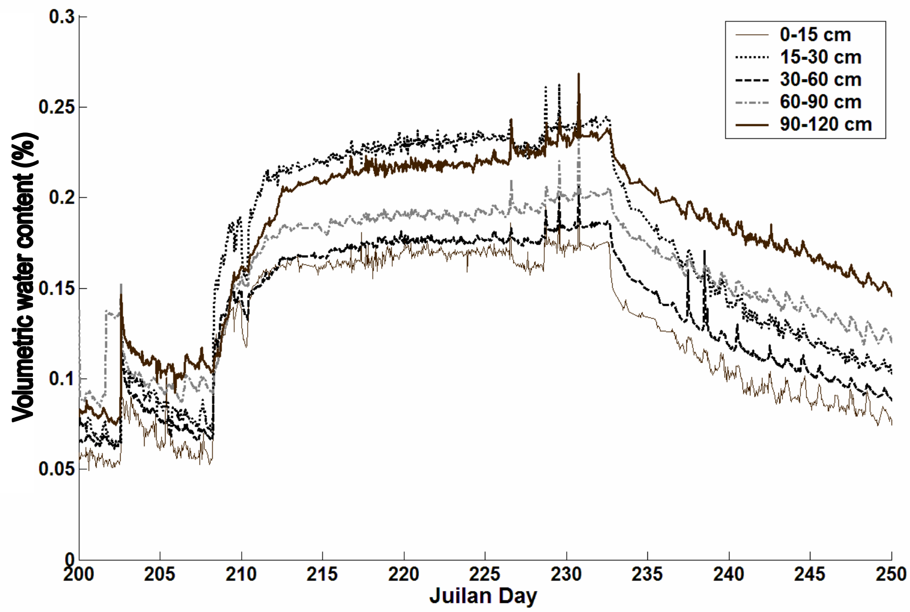


Figure 4.8 Volumetric water content (a) and profile averaged volumetric water content with error bounds (b). Note the rapid recession of the shallower depths (0-30 cm), and the rapid response to sprinkler malfunction on day 210.



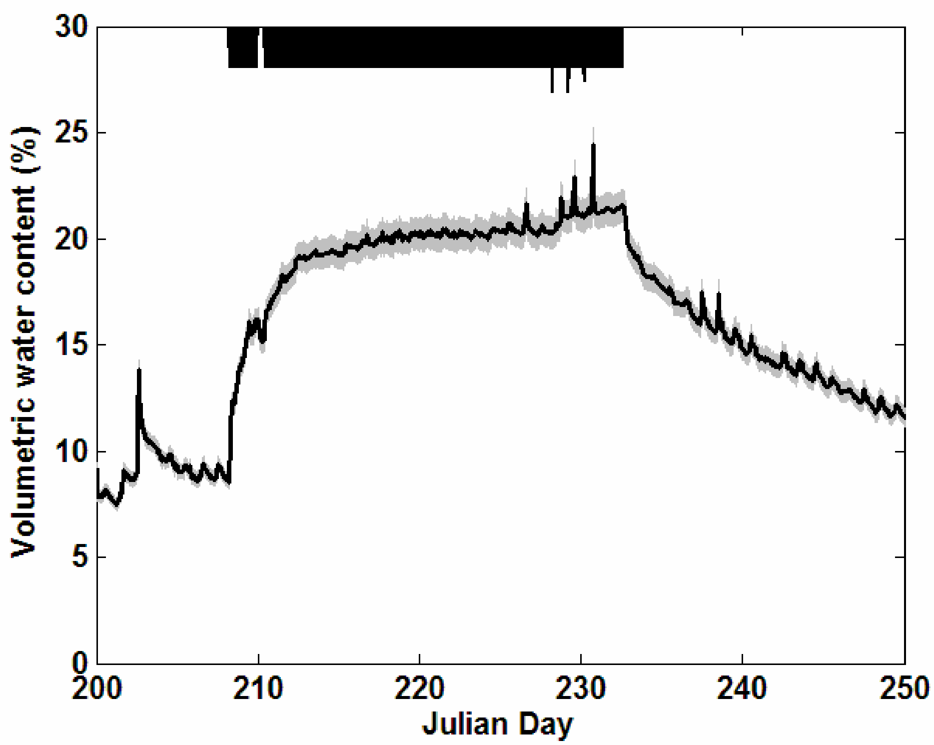


Figure 4.9 Lateral subsurface flow measured at the hillslope trench vs. hillslope average volumetric water content. Hysteretic loops in a counterclockwise fashion.

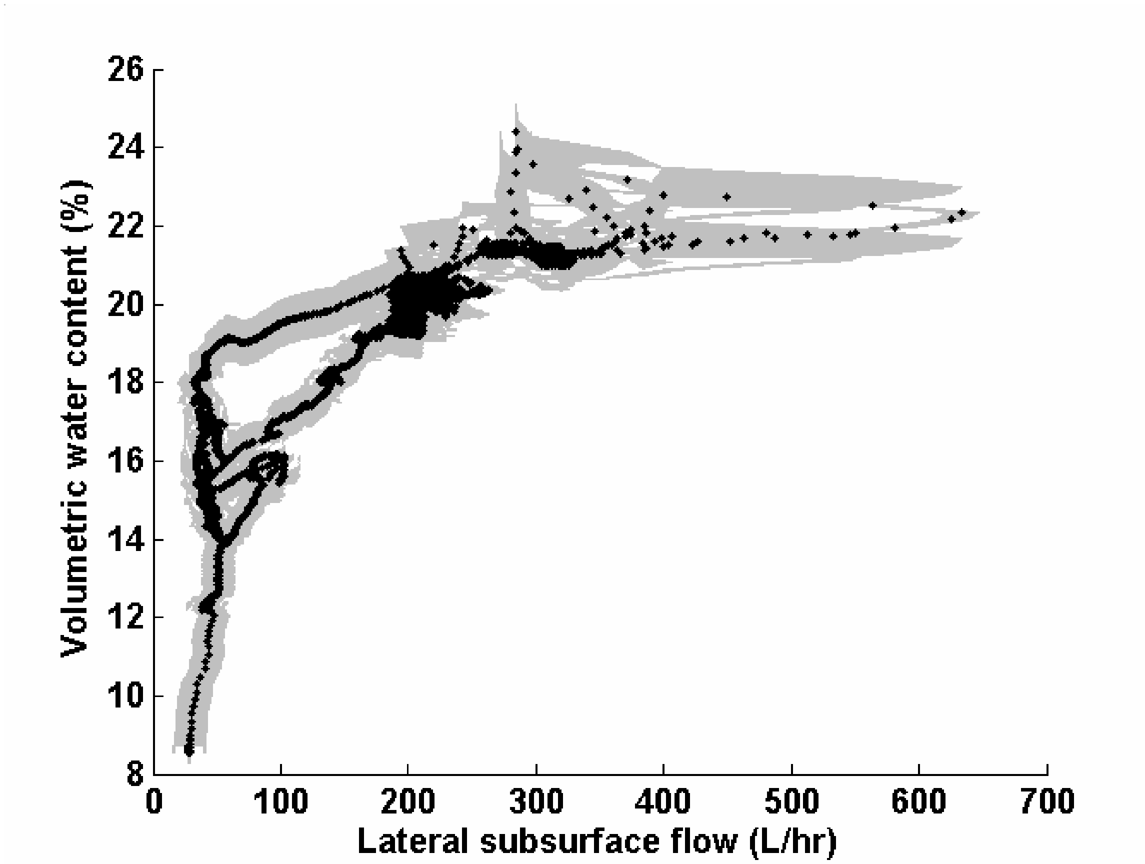
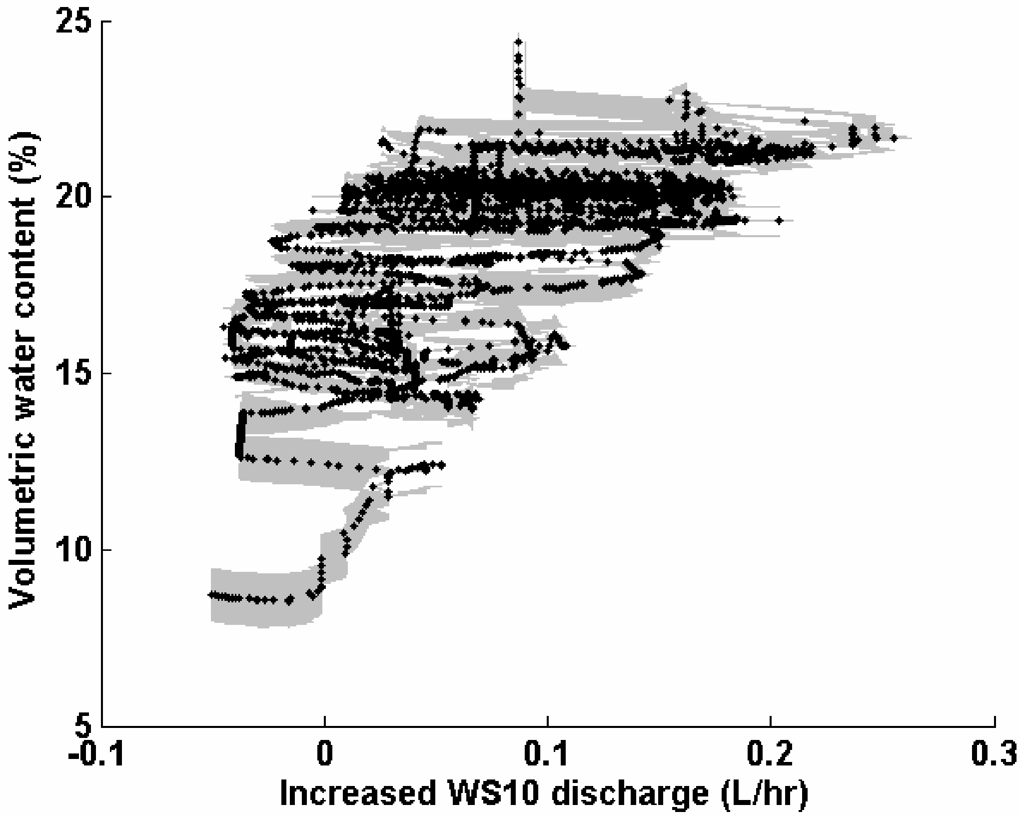


Figure 4.10 WS10 discharge minus background vs. hillslope average volumetric water content.



5. Conclusions

5.1. Summary of main thesis findings

Basic questions remain regarding the subsurface structure and dominant flow processes operating at the hillslope scale. In particular, the search for macroscale laws to describe hillslope scale properties has not progressed since Dooge (1986). This thesis has described a combined macroscale measurement and modeling approach, with an emphasis on uncertainty analysis, to determine dominant hillslope scale processes at two instrumented, steep forested hillslopes.

In Chapter 2, we showed that flow is dominated by rapid lateral subsurface flow isolated at the soil bedrock interface. Contrary to prior expectations, flow through macropores or the soil matrix was an insignificant component of lateral subsurface flow, and the soil profile served mainly as a conduit for vertical flow. Additionally, the bedrock at the site, long considered effectively impermeable, was shown to be semi-permeable. Water balance component analysis of the hillslope and a nearby instrumented catchment indicated that vertical percolation into the bedrock was a significant component of the water balance, similar to the observations at the WS10 hillslope in Chapter 5.

In Chapter 3, we demonstrated that a simple, low dimensional model of subsurface flow processes was able to capture the dynamics of hillslope flow and transport, including macroscale threshold behaviors. The model was then used to determine the controls on the threshold behavior seen in the modeling and at many instrumented hillslopes and catchments. Our model virtual experiments showed that the thresholds were controlled by both “fill and spill” (subsurface storage and bedrock permeability) and “soil moisture deficit” (evaporation rate and antecedent drainage time) factors. Application of the functional relationship between the soil moisture deficit factors and the threshold observed in a pair of long term data records demonstrated the value of the approach. The functional relationship, revealed from the virtual experiments, could provide a method for prediction of catchment and hillslope response in ungauged basins.

In Chapter 4, hillslope scale field experiments at a site with different geometry, geology and soil properties showed similarities in hillslope response to Maimai. As at Maimai, rapid hillslope response to irrigation was observed, both in lateral subsurface flow, and in flow through the bedrock. At the WS10 hillslope, flow through the bedrock accounted for up to one third of applied water, more than suggested by storm monitoring. Rigorous uncertainty analysis at WS10 showed that uncertainty in water balance components can be quite large, even with careful controls on inputs and measurement of outputs.

This work shows the value of a combined effort between hillslope experimentation and model development, both in terms of model development that concentrates on the dominant field processes, and in new process understanding from virtual experimentation.

5.2. Future research needs

This thesis shows two possible paths towards better macroscale process understanding. The demonstrated climatic and geologic controls on the precipitation / discharge threshold, and the influence of bedrock leakage and topography on flow routing and partitioning, could be a possible first step towards the joint goals of macroscale laws in hillslope hydrology (Dooge, 1986) and catchment classification (Wagener et al., 2007). However, further work is needed. The observed dominant flow processes at these two humid, steep forested hillslopes needs to be compared to sites with different geology, geometry and climate. Sites with different dominant flow processes should also be examined. While some functional intercomparison work has been performed (Uchida et al., 2006; Uchida et al., 2005), much more is needed.

The influence of scale also warrants further attention. While the soil moisture deficit dependent threshold relationship was shown at sites from hillslopes to small catchments (0.09 - 101.3 ha), the upper and lower bounds of scale are unknown. While this relationship appears to be a powerful tool for prediction of catchment response to precipitation, it is unclear whether it is applicable to much larger scales, where issues of spatial variability of rainfall, and different land use dynamics may be important. Finally, these field and virtual experiments have shown the value in the dialogue between modeler and experimentalist. To identify macroscale laws and develop catchment classification schemes, we need more combination of field and model experimentation, to serve as a platform for further progress.

5.3. References

Dooge, J.C.I., 1986. Looking for hydrologic laws. Water Resources Research, 22: 46S-58S.

Uchida, T., McDonnell, J.J. and Asano, Y., 2006. Functional intercomparison of hillslopes and small catchments by examining water source, flowpath and mean residence time. Journal of Hydrology, 327: 627-642.

Uchida, T., Tromp-van Meerveld, H.J. and McDonnell, J.J., 2005. The role of lateral pipe flow in hillslope runoff response: an intercomparison of non-linear hillslope response. Journal of Hydrology, 311: 117-133.

Wagener, T., Sivapalan, M., Troch, P. and Woods, R., 2007. Catchment classification and hydrologic similarity. Geography Compass, 1(4): 901-931.

Bibliography

Andréassian, V. et al., 2007. What is really undermining hydrologic science today? Hydrological Processes, 21(20): 2819-2822.

Atkinson, T.C., 1978. Techniques for measuring subsurface flow on hillslopes. In: M. Kirkby (Editor), Hillslope Hydrology. John Wiley & Sons, Chichester, UK, pp. 73-120.

Baker, T.M. and Hawke, R.M., 2007. The effects of land application of farm dairy effluent on groundwater quality – West Coast. Journal of Hydrology (New Zealand), 45.

Bárdossy, A., 1996. The use of fuzzy rules for the description of elements of the hydrological cycle. Ecological Modelling, 85(1): 59-65.

Barnard, H.R. et al., in prep. Examining hillslope transpiration controls of diel sub-surface flow using a steady-state irrigation experiment.

Bear, J., 1972. Dynamics of fluids in porous media. American Elsevier Pub. Co., New York, 764 pp.

Beven, K., 2001. On hypothesis testing in hydrology. Hydrological Processes, 15(9): 1655-1657.

Beven, K., 2002. 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., 2006a. Searching for the Holy Grail of scientific hydrology: $Q_t=(S,R,\Delta t)$ as closure. Hydrology and Earth System Sciences, 10(5): 609-618.

Beven, K. and Binley, A., 1992. Future of distributed models: model calibration and uncertainty prediction. Hydrological Processes, 6(3): 279-298.

Beven, K.J., 2006b. On undermining the science? Hydrological Processes, 20: 3141-3146.

Bonell, M. and Gilmour, D.A., 1978. The development of overland flow in a tropical rainforest catchment. Journal of Hydrology, 39(3-4): 365-382.

Brammer, D., 1996. Hillslope hydrology in a small forested catchment, Maimai, New Zealand. M.S. Thesis, State University of New York College of Environmental Science and Forestry, Syracuse, 153 pp.

Brooks, E.S., Boll, J. and McDaniel, P.A., 2004. A hillslope-scale experiment to measure lateral saturated hydraulic conductivity. Water Resources Research, 40.

Brooks, R.H. and Corey, A.T., 1964. Hydraulic properties of porous media. Hydrology Paper 3, Colorado State University, Fort Collins, CO.

Buttle, J.M., Dillon, P.J. and Eerkes, G.R., 2004. Hydrologic coupling of slopes, riparian zones and streams: an example from the Canadian Shield. Journal of Hydrology, 287(1-4): 161-177.

Buttle, J.M. and McDonald, D.J., 2002. Coupled vertical and lateral preferential flow on a forested slope. Water Resources Research, 38(5).

Carsel, R.F. and Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. Water Resources Research, 24(5): Pages: 755-769.

Chapman, T., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes, 13(5): 701-714.

Clarke, R.T., 1999. Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. Journal of Hydrology, 222(1-4): 185-190.
Domec, J.-C., Meinzer, F.C., Gartner, B.L. and Woodruff, D., 2006. Transpiration-induced axial and radial tension gradients in trunks of Douglas-fir trees. Tree Physiology, 26(3): 275-284.
Dooge, J.C.I., 1986. Looking for Hydrologic Laws. Water Resources Research, 22: 46S-58S.
Eberhardt, L.L. and Thomas, J.M., 1991. Designing environmental field studies. Ecological Monographs, 61(1): 53-73.
Ewen, J. and Birkinshaw, S.J., 2007. Lumped hysteretic model for subsurface stormflow developed using downward approach. Hydrological Processes, 21: 1496-1505.
Fenicia, F., McDonnell, J.J. and Savenije, H.H.G., 2008. Learning from model improvement: On the contribution of complementary data to process understanding. Water Resources Research, 44.
Flury, M. and Flühler, H., 1995. Tracer characteristics of Brilliant Blue FCF. Soil Science Society of America Journal, 59: 22-27.
Flury, M., Leuenberger, J., Studer, B. and Flühler, H., 1995. Transport of anions and herbicides in a loamy and a sandy field soil. Water Resources Research, 30: 823-835.
Freer, J. et al., 1997. Topographic controls on subsurface storm flow at the hillslope scale for two hydrologically distinct small catchments. Hydrological Processes, 11(9): 1347-1352.
Freer, J. et al., 2002. The role of bedrock topography on subsurface storm flow. Water Resources Research, 38(12).
Freeze, R.A. and Cherry, J.A., 1979. Groundwater. Prentice-Hall, Englewood Cliffs, NJ, 604 pp.
Gee, G.W. et al., 1994. Variations in Water Balance and Recharge Potential at Three Western Desert Sites. Soil Science Society of America Journal, 58(1): 63-72.
Gibbens, R.P. and Lenz, J.M., 2001. Root systems of some Chihuahuan Desert plants. Journal of Arid Environments, 49(2): 221-263.
Graham, C.B. and McDonnell, J.J., this issue. Hillslope threshold response to rainfall: (2) Development and use of a macroscale behavioral model.
Granier, A., 1987. Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. Tree Physiology, 3: 309-320.
Grayson, R.B. and Blöschl, G., 2000. Summary of pattern comparison and concluding remarks. In: R.B. Grayson and G. Blöschl (Editors), In Spatial Patterns in Catchment Hydrology: Observations and Modeling. Cambridge University Press, Cambridge, UK, pp. 355–367.
Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1998. Towards improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.
Hall, J., O'Connell, E. and Ewen, J., 2007. On not undermining the science: coherence, validation and expertise. Discussion of Invited Commentary by Keith Beven Hydrological Processes, 20, 3141-3146 (2006). Hydrological Processes, 21: 985-988.

Harr, R.D., 1977. Water flux in soil and subsoil on a steep forested slope. Journal of Hydrology, 33: 37-58.

Heitschmidt, R.K., Ansley, R.J., Dowhower, S.L., Jacoby, P.W. and Price, D.L., 1988. Some observations from the excavation of honey mesquite root systems. Journal of Range Mananagement, 41: 227-231.

Hewlett, J.D. and Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: W.E. Sopper and H.W. Lull (Editors), Forest Hydrology. Pergamon Press, New York, pp. 275-291.

Hornberger, G.M., Germann, P.F. and Beven, K.J., 1991. Throughflow and solute transport in an isolated sloping soil block in a forested catchment. Journal of Hydrology, 124: 81-99.

Hursh, C.R., 1944. Subsurface flow. Transactions of the American Geophysical Union, 25: 743-746.

Hutchinson, D.G. and Moore, R.D., 2000. Throughflow variability on a forested hillslope underlain by compacted glacial till. Hydrological Processes, 14(10): 1751-1766.

James, A. and McDonnell, J.J., 2008. Panola Modeling.

James, M.E., 1978. Rock weathering in the central western Cascades. M.S. Thesis, University of Oregon, Eugene, 119 pp.

Katsura, S.y., Kosugi, K., Mizutani, T., Okunaka, S. and Mizuyama, T., 2008. Effects of bedrock groundwater on spatial and temporal variations in soil mantle groundwater in a steep granitic headwater catchment. Water Resources Research, 44.

Katsuyama, M., Ohte, N. and Kabeya, N., 2005. Effects of bedrock permeability on hillslope and riparian groundwater dynamics in a weathered granite catchment. Water Resources Research, 41.

Kavetski, D., Franks, S.W. and Kuczera, G., 2002. Confronting input uncertainty in environmental modelling. In: Q. Duan, H.V. Gupta, S. Sorooshian, A.N. Rousseau and R. Turcotte (Editors), In Calibration of Watershed Models. AGU Books, Washington, WA, pp. 49-68.

Keim, R.F., Tromp-van Meerveld, I. and McDonnell, J.J., 2006. A virtual experiment on the effects of evaporation and intensity smoothing by canopy interception on subsurface stormflow generation. Journal of Hydrology, 327: 352-364.

Kendall, K.A., Shanley, J.B. and McDonnell, J.J., 1999. A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt. Journal of Hydrology, 219(3-4): 188-205.

Kim, H.J., Sidle, R.C., Moore, R.D. and Hudson, R., 2004. Throughflow variability during snowmelt in a forested mountain catchment, coastal British Columbia, Canada. Hydrological Processes, 18(7): 1219-1236.

Kitahara, H., 1993. Characteristics of pipe flow in forested slopes. IAHS Publication, 212: 235-242.

Kuczera, G. and Mroczkowski, M., 1998. Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. Water Resources Research, 34: 1481-1489.

Lehmann, P., Hinz, C., McGrath, G., Meerveld, H.J.T.-v. and McDonnell, J.J., 2007. Rainfall threshold for hillslope outflow: an emergent property of flow pathway connectivity. Hydrology and Earth System Sciences, 11(2).

Mantovan, P. and Todini, E., 2006. Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology. Journal of Hydrology.

McDonnell, J.J., 1990. A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11): 2821-2832.

McDonnell, J.J., 1997. Comment on "the changing spatial variability of subsurface flow across a hillslide" by Ross Woods and Lindsay Rowe. Journal of Hydrology (New Zealand), 36(1): 97-100.

McDonnell, J.J., Freer, J. and Vache, K., 2009. A reference dataset for event-based watershed model development and testing: The Maimai catchment. Water Resources Research, in review.

McDonnell, J.J., Owens, I.F. and Stewart, M.K., 1991. A case study of shallow flow paths in a steep zero-order basin. Water Resources Bulletin, 27(4): 679-685.

McDonnell, J.J. et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, 43.

McGlynn, B.L. and McDonnell, J.J., 2003a. Quantifying the relative contributions of riparian and hillslope zones to catchment runoff. Water Resources Research, 39(11).

McGlynn, B.L. and McDonnell, J.J., 2003b. Role of discrete landscape units in controlling catchment dissolved organic carbon dynamics. Water Resources Research, 39(4).

McGlynn, B.L., McDonnell, J.J. and Brammer, D.D., 2002. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. Journal of Hydrology, 257: 1-26.

McGuire, K.J., McDonnell, J.J. and Weiler, M.H., 2007. Integrating tracer experiments with modeling to infer water transit times. Advances in Water Resources, 30: 824-837.

McGuire, K.J. et al., 2005. The role of topography on catchment-scale water residence time. Water Resources Research, 41.

McIntosh, J., McDonnell, J.J. and Peters, N.E., 1999. Tracer and hydrometric study of preferential flow in large undisturbed soil cores from the Georgia Piedmont, USA. Hydrological Processes, 13: 139-155.

Moncrieff, J.B., Malhi, Y. and Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Global Change Biology, 2(3): 231-240.

Montanari, A., 2007. What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. Hydrological Processes, 21(6): 841-845.

Monteith, J.L. and Unsworth, M.H., 2008. Principles of Environmental Physics. Academic Press, London, 418 pp.

Montgomery, D.R. et al., 1997. Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. Water Resources Research, 33(1): 91-109.

Moore, G.W., Bond, B.J., Jones, J.A., Phillips, N. and Meinzer, F.C., 2004. Structural and compositional controls on transpiration in 40- and 450-year-old riparian forests in western Oregon, USA. Tree Physiology, 24: 481-491.

Moran, C.J., McBratney, A.B. and Kopi, A.J., 1989. A rapid method for analysis of soil macropore structure. I. Specimen preparation and digital image production. Soil Science Society of America Journal, 53(3): 921-928.
Mosley, M.P., 1979. Streamflow generation in a forested watershed. Water Resources Research, 15: 795-806.
Mosley, M.P., 1982. Subsurface Flow Velocities Through Selected Forest Soils, South Island, New Zealand. Journal of Hydrology, 55: 65-92.
Murray, F.W., 1967. On the computation of saturation vapor pressure. Journal of Applied Meteorology, 6: 203-204.
Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, I, A discussion of principles. Journal of Hydrology, 10: 282-290.
Noguchi, S., Nik, A.R., Kasran, B., Tani, M. and Sammori, T., 1997. Soil physical properties and preferential flow pathways in tropical rain forest, Bukit Tarek, Pennisular Malaysia. Journal of Forest Research, 2: 115-120.
Noguchi, S., Tsuboyama, Y., Sidle, R.C. and Hosoda, I., 2001. Subsurface runoff characteristics from a forest hillslope soil profile including macropores, Hitachi Ohta, Japan. Hydrological Processes, 15: 2131-2149.
O'Loughlin, C.L., Rowe, L.K. and Pearce, A.J., 1978. Sediment yields from small forested catchments, North Westland-Nelson, New Zealand. Journal of Hydrology (New Zealand), 17: 1-15.
Onda, Y., Komatsu, Y., Tsujimura, M. and Fujihara, J.-i., 2001. The role of subsurface runoff through bedrock on storm flow generation. Hydrological Processes, 15: 1693-1706.
Özelkan, E.C. and Duckstein, L., 2001. Fuzzy conceptual rainfall-runoff models. Journal of Hydrology, 253(1-4): 41-68.
Pearce, A.J. and Rowe, L.K., 1979. Forest management effects on interception, evaporation, and water yield. Journal of Hydrology, 18(2): 73-87.
Pearce, A.J., Stewart, M.K. and Sklash, M.G., 1986. Storm runoff generation in humid headwater catchments: 1. Where does the water come from? Water Resources Research, 22: 1263-1272.
Peters, D.L., Buttle, J.M., Taylor, C.H. and LaZerte, B.D., 1995. Runoff production in a forested, shallow soil, Canadian Shield basin. Water Resources Research, 31(5): 1291-1304.
Putuhena, W.M. and Cordery, I., 2000. Some hydrological effects of changing forest cover from eucalypts to Pinus radiata. Agricultural and Forest Meteorology, 100(1): 59-72.
Quinn, P.F., Beven, K.J., Chevallier, P. and Planchon, O., 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes, 5: 59-79.
Ranken, D.W., 1974. Hydrologic properties of soil and subsoil on a steep, forested slope. M.S. Thesis, Oregon State University, Corvallis, 114 pp.
Rasmussen, T.C., Baldwin, R.H., Dowd, J.F. and Williams, A.G., 2000. Tracer vs. pressure wave velocities through unsaturated saprolite. Soil Science Society of America Journal, 64: 75-85.
Ridolfi, L., D'Odorico, P., Porporato, A. and Rodriguez-Iturbe, I., 2003. Stochastic soil moisture dynamics along a hillslope. Journal of Hydrology, 272(1-4): 264-275.

Rodriguez-Iturbe, I., 2000. Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics. Water Resources Research, 36(1): 3-9.
Rowe, L.K., Pearce, A.J. and O'Loughlin, C.L., 1994. Hydrology and related changes after harvesting native forests and establishing Pinus radiata plantations. Part 1. Introduction to the study. Hydrological Processes, 8: 263-279.
Scherrer, S., Naef, F., Faeh, A., O. and Cordery, I., 2007. Formation of runoff at the hillslope scale during intense precipitation. Hydrology and Earth System Sciences, 11: 907-922.
Seibert, J., 1997. Estimation of parameter uncertainty in the HBV model. Nordic Hydrology, 28: 4-5.
Seibert, J., Bishop, K., Rodhe, A. and McDonnell, J.J., 2003. Groundwater dynamics along a hillslope: a test of the steady state hypothesis. Water Resources Research, 39(1): 1014, doi:10.1029/2002WR001404.
Seibert, J. and McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. Water Resources Research, 38(11).
Sidle, R.C., Hirano, T., Gomi, T. and Terajima, T., 2007. Hortonian overland flow from Japanese forest plantations - an aberration, the real thing, or something in between? Hydrological Processes, 21(23): 3237-3247.
Singh, P., Kanwar, R.S. and Thompson, M.L., 1991. Measurement and characterization of macropores by using AUTOCAD and automatic image analysis. Journal of Environmental Quality, 20: 289-294.
Sivakumar, B., 2008. Undermining the science or undermining Nature? Hydrological Processes, 22(6): 893-897.
Sivapalan, M., 2003. Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection? Hydrological Processes, 17(5): 1037-1041.
Sklash, M.G. and Farvolden, R.N., 1979. The role of groundwater in storm runoff. Journal of Hydrology, 43: 45-65.
Sklash, M.G., Stewart, M.K. and Pearce, A.J., 1986. Storm runoff generation in humid headwater catchments: 2. a case study of hillslope and low-order stream response. Water Resources Research, 22(8): 1273-1282.
Son, K. and Sivapalan, M., 2007. Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water Resources Research, 43.
Soulsby, C. and Dunn, S.M., 2003. Towards integrating tracer studies in conceptual rainfall-runoff models: recent insights from a sub-arctic catchment in the Cairngorm Mountains, Scotland. Hydrological Processes, 17: 403-416.
Spence, C. and Woo, M.-K., 2002. Hydrology of subartic Canadian shield: bedrock upland. Journal of Hydrology, 262: 111-127.
Stednick, J.D., 1996. Monitoring the effects of timber harvest on annual water yield. Journal of Hydrology, 176(1-4): 79-95.
Stewart, M.K. and McDonnell, J.J., 1991. Modeling base flow soil water residence times from deuterium concentrations. Water Resources Research, 27(10): 2681-2693.
Sujono, J., Shikasho, S. and Hiramatsu, K., 2004. A comparison of techniques for hydrograph recession analysis. Hydrological Processes, 18(3): 403-413.

Swanson, F.J. and James, M.E., 1975. Geology and geomorphology of the H.J. Andrews Experimental Forest, western Cascades, Oregon. Res. Pap. PNW-188, U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, OR.

Tan, C.S. and Black, T.A., 1976. Factors affecting the canopy resistance of a Douglas-fir forest. Boundary-Layer Meteorology, 10(4).

Tani, M., 1997. Runoff generation processes estimated from hydrological observations on a steep forested hillslope with a thin soil layer. Journal of Hydrology, 200: 84-109.

Taylor, J.R., 1997. An introduction to error analysis. University Science Books, Sausalito, California, 328 pp.

Tetens, O., 1930. Uber einige meteorologische begriffe. Z. Geophys, 6: 203-204.

Tetzlaff, D., Uhlenbrook, S., Eppert, S. and Soulsby, C., 2008. Does the incorporation of process conceptualization and tracer data improve the structure and performance of a simple rainfall-runoff model in a Scottish mesoscale catchment? Hydrological Processes.

Tippkötter, R., 1983. Morphology, spatial arrangement and origin of macropores in some hapludalfs, West Germany. Geoderma, 29: 355-371.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006a. Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope. Water Resources Research, 42.

Tromp-van Meerveld, H.J. and McDonnell, J.J., 2006b. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resources Research, 42.

Tromp-van Meerveld, H.J., Peters, N.E. and McDonnell, J.J., 2006. 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.

Tsuboyama, Y., Sidle, R.C., Noguchi, S. and Hosoda, I., 1994. Flow and solute transport through the soil matrix and macropores of a hillslope segment. Water Resources Research, 30(4): 879-890.

Uchida, T., Kosugi, K. and Mizuyama, T., 1999. Runoff characteristics of pipeflow and effects of pipeflow on rainfall-runoff phenomena in a mountainous watershed. Journal of Hydrology, 222(1-4): 18-36.

Uchida, T., Tromp-van Meerveld, H.J. and McDonnell, J.J., 2005. The role of lateral pipe flow in hillslope runoff response: an intercomparison of non-linear hillslope response. Journal of Hydrology, 311: 117-133.

Vache, K.B. and McDonnell, J.J., 2006. A process-based rejectionist framework for evaluating catchment runoff model structure. Water Resources Research, 42.

van Verseveld, W., 2006. Hydro-biogeochemical coupling at the hillslope and catchment scale. PhD Thesis, Oregon State University, Corvallis, OR.

Vrugt, J.A., Dekker, S.C. and Bouten, W., 2003. Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage. Water Resources Research, 39(9).

Wagener, T., 2003. Evaluation of catchment models. Hydrological Processes, 17(16): 3375-3378.

Wagener, T., Sivapalan, M., Troch, P. and Woods, R., 2007. Catchment classification and hydrologic similarity. Geography Compass, 1.
Waichler, S.R., Wemple, B.C. and Wigmosta, M.S., 2005. Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. Hydrological Processes, 19: 3177-3199.
Weiler, M. and McDonnell, J., 2004. Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology. Journal of Hydrology, 285(1-4): 3-18.
Weiler, M. and McDonnell, J.J., 2007. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes. Water Resources Research, 43.
Weiler, M., McDonnell, J.J., Tromp-van Meerveld, I. and Uchida, T., 2006. Subsurface stormflow. Encyclopedia of Hydrological Sciences, 3.
Weiler, M. and Naef, F., 2003. An experimental tracer study of the role of macropores in infiltration in grassland soils. Hydrological Processes, 17(2): 477-493.
Western, A.W., Bloeschl, G. and Grayson, R.B., 1998. Geostatistical characterisation of soil moisture patterns in the Tarrawarra catchment. Journal of Hydrology, 205(1-2): 20-37.
Whipkey, R.Z., 1965. Subsurface storm flow from forested slopes. Bulletin of the International Association of Scientific Hydrology, 2: 74-85.
Winter, T.C., 1981. Uncertainties in estimating the water balance of lakes. Journal of the American Water Resources Association, 17(1): 82-115.
Woods, R. and Rowe, L., 1996. The changing spatial variability of subsurface flow across a hillside. Journal of Hydrology (New Zealand), 35(1): 51-86.
Woods, R. and Rowe, L., 1997. Reply to "Comment on "The changing spatial variability of subsurface flow across a hillside" by Ross Woods and Lindsay Rowe". Journal of Hydrology (New Zealand), 36(1): 51-86.
Woods, R.A., Sivapalan, M. and Robinson, J.S., 1997. Modeling the spatial variability of subsurface runoff using a topographic index. Water Resources Research, 33(5): 1061-1073.
Zehe, E., Elsenbeer, H., Lindenmaier, F., K, S. and Blöschl, G., 2007. Patterns of predictability in hydrological threshold systems. Water Resources Research, 43.
Zehe, E. and Fluhler, H., 2001. Slope scale variation of flow patterns in soil profiles. Journal of Hydrology, 247: 116-132.