Effects of differential hillslope-scale water retention characteristics on rainfall-runoff response at the Landscape Evolution Observatory

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>van den Heuvel, Daniel B.; Troch, Peter A.; Booij, Martijn J.; Niu, Guo-Yue; Volkmann, Till H. M.; Pangle, Luke A.</td>
</tr>
<tr>
<td>DOI</td>
<td>10.1002/hyp.v32.13</td>
</tr>
<tr>
<td>Publisher</td>
<td>WILEY</td>
</tr>
<tr>
<td>Journal</td>
<td>HYDROLOGICAL PROCESSES</td>
</tr>
<tr>
<td>Rights</td>
<td>Copyright © 2018 John Wiley &amp; Sons, Ltd.</td>
</tr>
<tr>
<td>Download date</td>
<td>08/08/2019 08:55:47</td>
</tr>
<tr>
<td>Version</td>
<td>Final accepted manuscript</td>
</tr>
<tr>
<td>Link to Item</td>
<td><a href="http://hdl.handle.net/10150/628591">http://hdl.handle.net/10150/628591</a></td>
</tr>
</tbody>
</table>
Effects of differential hillslope-scale water retention characteristics on rainfall-runoff response at the Landscape Evolution Observatory


aBiosphere 2, University of Arizona, Tucson, Arizona, USA
bDepartment of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona, USA
cWater Engineering and Management Group, Faculty of Engineering Technology, University of Twente, Enschede, the Netherlands
dDepartment of Geosciences, Georgia State University, Atlanta, Georgia, USA

1present address: Water Engineering and Management Group, Faculty of Engineering Technology, University of Twente, Enschede, the Netherlands

*Correspondence to: P.A. Troch, Department of Hydrology and Atmospheric Sciences, 1133 E James E. Rogers Way. PO Box 210011, Tucson, AZ 85721-0011. Email: patroch@email.arizona.edu. Phone: 1-520-626-1277)

Keywords: Rainfall-runoff response, landscape evolution, soil water retention, soil characteristics, hydrological 3D modeling, Biosphere 2

Abstract

Hillslopes turn precipitation into runoff and thus exert important controls on various Earth system processes. It remains difficult to collect reliable data necessary for understanding and modeling these Earth system processes in real catchments. To overcome this problem, controlled experiments are being conducted at the Landscape Evolution Observatory (LEO) at Biosphere 2, The University of Arizona. Previous experiments have revealed differences in hydrological response between two landscapes within LEO, even though both landscapes were designed to be identical. In an attempt to discover where the observed differences stem from, we use a fully
three-dimensional hydrological model (CATchment HYdrology, CATHY) to show the effect of soil water retention characteristics and saturated hydraulic conductivity on the hydrological response of these two hillslopes. We also show that soil water retention characteristics can be derived at hillslope scale from experimental observations of soil moisture and matric potential. It is found that differences in soil packing between the two landscapes may be responsible for the observed differences in hydrological response. This modeling study also suggests that soil water retention characteristics and saturated hydraulic conductivity have a profound effect on rainfall-runoff processes at hillslope-scale and that parametrization of a single hillslope may be a promising step in modeling rainfall-runoff response in real catchments.

1. Introduction

Over the past decade, several opinion papers on hillslope and catchment hydrology have argued for the need to explicitly include subsurface heterogeneity in rainfall-runoff modeling (McDonnell et al., 2007; Sivapalan, 2003; Sivapalan, Blöschl, Zhang, & Vertessy, 2003; Troch et al., 2013). However, quantifying model parameters that reflect this heterogeneity is extremely difficult due to the range of spatial scales over which heterogeneity in soil properties manifests itself. Recently, hydrologists have raised the possibility that, when Earth system processes responsible for landscape evolution are better understood, some of these subsurface properties might be better quantifiable (Harman & Troch, 2014; Lin et al., 2006; McDonnell et al., 2007; Troch et al., 2015; Wagener, Sivapalan, Troch, & Woods, 2007). These Earth system processes are generally associated with various disciplines such as hydrology, ecology, geochemistry and geomorphology. Although the strong interdependence of these physical, chemical and biological processes is well-known and their influence on landscape evolution is widely acknowledged, it remains difficult to conduct reliable field experiments to collect the required data for model
In order to overcome this problem and be able to understand coupled Earth system processes associated with rainfall-runoff dynamics, the University of Arizona broke ground in 2007 on a large-scale interdisciplinary research project, the Landscape Evolution Observatory (LEO; see http://biosphere2.org/research/projects/landscape-evolution-observatory). The project’s goal is to understand how different interacting Earth systems processes determine the evolution of landscapes over time. This knowledge can then be used to shed light on past landscape changes and to predict future landscape evolution.

LEO is unique in its field due to its fully controlled environment, state-of-the-art measuring equipment and hillslope-size scale. Other projects with similar research goals include the Critical Zone Observatory network in the USA (Anderson, Bales, & Duffy, 2008; Guo & Lin, 2016), the artificial catchment “Chicken Creek” located in Germany (Gerwin, Raab, Biemelt, Bens, & Hüttl, 2009; Hofer, Lehmann, Biemelt, Stähli, & Krafczyk, 2011) and the TERENO program (Bogena et al., 2016; Zacharias et al., 2011), also located in Germany. While these projects are similar in the sense that they also attempt to improve understanding of coupled processes in catchments, they take place at a different spatial scale. The two German projects comprise entire catchments, whereas the CZO network investigates pedon, hillslope and watershed scale Earth systems processes across climate gradients (from tropical sites in Puerto Rico to agricultural sites in Illinois). Also, these projects lack the control and observational capacity of LEO as they are not located within a controlled environment.

A successful modeling study where an attempt was made to model the behavior of an entire catchment using parameters of a single, representative hillslope was undertaken by Loritz et al. (2017) in the Attert experimental catchment, Luxembourg (Pfister, Humbert, & Hoffmann, ...
2000). This catchment comprises two different sub-catchments that are heavily instrumented and where elaborate field data are collected by the CAOS (Catchments As Organized Systems) research program (Zehe et al., 2014), but the catchment is not in a controlled environment. Loritz et al. (2017) used the 2-D physically based CATFLOW model (Zehe, Maurer, Ihringer, & Plate, 2001) and parametrizations were based on extensive field data, expert knowledge and process-based reasoning. While they were not able to simulate the entire range of spatial variabilities within a catchment using their physically based model, they could generate meaningful simulations of the streamflow in the catchment. They argue that some of the limitations found could be attributed to the chosen 2-D model and to our understanding of the dynamics within catchments, while others may be related to the concept of replacing a small catchment with a single hillslope. Their work indicates that we may not yet be able to set up a fully representative model for a catchment. However, their approach of replacing a small-scale catchment with single hillslope parametrization seems a promising step in modeling the behavior of small catchments. Their work is also related to the present study, in that we here attempt to model the behavior of a catchment using extensive field data collected from experiments at the hillslope scale, although the scale is much smaller than the catchment considered by Loritz et al. (2017).

LEO consists of three landscapes (hillslopes) that are identical in shape, soil, environment and technical equipment. Throughout this paper, these slopes will be referred to as the west, central and east landscapes, in accordance with their position within the complex. Extensive rainfall-runoff experiments conducted on LEO’s east landscape in 2014 were simulated using the CATchment HYdrology (CATHY) model prior to actual experiment execution, assuming soil homogeneity. However, unexpected observed overland flow led Niu et al. (2014) to conclude that the east landscape’s homogeneous soil might have become heterogeneous during the experiment.
In the spring of 2015, similar rainfall-runoff experiments were conducted nearly simultaneously on the central and west landscapes, revealing considerable differences in hydrological response times between the two landscapes. The central landscape seems to discharge water much faster than its western counterpart and preliminary hydrological analyses have shown that the hydrological response times of the east landscape resembled those of the central landscape. This issue is interesting and the variation in response larger than expected, as the three landscapes were assumed to be fully identical in geometry, soil composition and technical equipment installed. For instance, the landscapes were sequentially packed in the same fashion and laser scans were performed after each incremental installation. Soil depth maps of the three landscapes as presented by Pangle et al. (2015) leave the impression that these measures were fruitful, as the landscapes’ soil depths show only small deviations. Given the seemingly identicalness of these landscapes, the observed differences in hydrological response times were much larger than we had expected.

This study aims to elucidate why two identically designed and built hillslopes (central and west within LEO) differ substantially in rainfall-runoff response. To this end, models of these two landscapes are set up with CATHY in a similar fashion as was done for the east landscape in 2014. Since measurements and tests have left the strong impression that the landscapes’ geometries are identical and the measuring equipment functions properly, this work focuses on the role of the soil’s water retention characteristics and saturated hydraulic conductivity. While there may be other factors responsible for the observed differences, such as localized heterogeneities in soil parameters, difference in spatial distribution of applied rainfall and localized differences in initial wetness, we decided to focus first on the mentioned soil parameters. Many other studies have shown the importance of water retention characteristics and...
hydraulic conductivity at the hillslope scale, both from a scientific and engineering point of view (e.g. Antinoro, Arnone, & Noto, 2017; Bullied, Bullock, & Van Acker, 2011; Geroy et al., 2011; Jackisch et al., 2017). More specifically, we take into account the two parameters $\alpha^{-1}$ and $n$ from the Van Genuchten relation for soil moisture as a function of the matric potential ($\theta(\psi)$) (Van Genuchten, 1980) and the saturated hydraulic conductivity $K_s$. In this work we simulate both landscapes in CATHY using variations of the these three soil parameters. We then compare the simulated and observed values of the soil parameters in both landscapes in an attempt to explain the differences in rainfall-runoff response. We also derive those parameters through calibration and co-located in-situ measurements of soil moisture content and soil water potential.

The remainder of this paper is organized as follows. Section 2 describes the physical model of LEO, the experiments, and the hydrological model CATHY. Section 3 presents the observed and simulated water retention characteristics and rainfall-runoff response of the different LEO hillslopes. Section 4 contains the discussion of this research and Section 5 comprises the conclusions drawn from this work.

2. Material and methods

2.1 Model landscapes

Construction of LEO within Biosphere 2 was finished in 2012. The result consists of three artificial landscapes measuring 30 m in length by 11.15 m in width. The average slope is 10° and the shape of the landscape is convergent. Figure 1 shows an artist impression of the complex.

Crushed basaltic tephra from the same crushed rock was used as a homogeneous soil layer of 1-m thickness. There was no vegetation on the soil during the experiments. The basaltic tephra is expected to evolve into structured soil over the course of multiple rainfall experiments, due to
geochemical weathering of the primary minerals and the precipitation of secondary clay minerals
at locations where soil solution reaches super-saturation. The bottom end of each landscape
features a 0.5-m wide section of gravel bordering a plastic plate with 2 mm diameter holes drilled
in it. The seepage face is located at the interface between the soil and gravel.

The landscapes sit in a controlled environment with over 1,800 subsurface sensors and
samplers per landscape. For measurement purposes, sensors are installed at five depth levels
throughout each landscape. This ensures high spatial resolution in horizontal and vertical
directions. A more detailed description of the sensors relevant in this study is provided in section
2.3.

Artificial rainfall can be applied to the landscapes using 14 sprinkler heads installed above
each slope. These sprinklers are equally distributed in space, are positioned approximately 3 m
above the soil surface and have a maximum rainfall capacity of 40 mm h⁻¹.

2.2 Rainfall-runoff experiments

Rainfall experiments were conducted on the central landscape on 11 May 2015 between 07:30
and 19:30 Local Time (LT) and on the west landscape on 18 May 2015 between 07:00 and 19:00
LT. Both rainfall events had a constant intensity of approximately 12 mm h⁻¹. On both landscapes
134 mm of rainfall was applied and no overland flow occurred. Prior to these events, test runs
had been carried out to bring the hillslopes to similar initial wetness conditions and to test all
equipment installed. Both landscapes were equally wet at the start of the experiment with water
storage values of approximately 105 mm. This value was derived from soil moisture content
measurements.
Throughout this work, we assume the relationship between the soil moisture and matric potential to be in accordance with the equation of Van Genuchten (Van Genuchten, 1980):

$$\theta(\psi) = \theta_r + \frac{\theta_s - \theta_r}{(1 + (|\alpha\psi|)^n)^{\frac{n-1}{n}}} \quad [1]$$

where $\theta$ is the volumetric soil water content (SWC) [L$^3$ L$^{-3}$], $\psi$ is the matric potential (MP) [L], $\theta_r$ the residual SWC (assumed to be zero at LEO’s landscapes (Pangle et al., 2015)), $\theta_s$ the saturated soil moisture, assumed equal to the soil porosity [L$^3$ L$^{-3}$], $\alpha$ a constant depending on the soil and the position of its capillary fringe [L$^{-1}$] and $n$ a constant depending on the soil packing [-].

2.4 Acquisition and processing of data

496 Decagon 5TM sensors (Decagon Devices, Inc.) measure the SWC. A calibration curve specific to the ground basalt material was used to convert the measured dielectric permittivity values to SWC values (95% confidence intervals of ±0.024). The number of SWC sensors decreases gradually with soil depth (154 sensors at –0.05 m and –0.2 m each, 76 sensors at –0.35 m, 78 sensors at –0.5 m and 34 sensors at –0.85 m). This allows for maintaining a 1 to 2-m resolution in the vertical and lateral direction of the landscapes. In addition, the soil’s MP is measured using 496 Decagon MPS-2 sensors. These sensors are co-located with the SWC sensors. The MP sensors can measure values within the range of –6 to –500 kPa. As a result, these sensors could not be used under wet conditions (SWC > 0.18 m$^3$ m$^{-3}$) because MP values were smaller than –6 kPa. In unsaturated cases, the MP sensors feature a manufacturer-reported accuracy of ±25% of the measured value. The co-location of the SWC and MP sensors allows for
in-situ measurements of the SWC versus the MP, thus allowing for deriving experimental Van Genuchten parameters $\alpha$ and $n$.

Total water storage values were retrieved from each landscape during the 12-hour experiments and during a period of 220 hours thereafter. First, the average value of all available SWC sensor readings was calculated for each depth at which SWC sensors are installed. These depth-specific averages were then weighted by the vertical distance between the sensors at the different soil depths.

Hillslope discharge is measured with two different types of sensors: calibrated NovaLynx 26-2501-A tipping bucket gauges and magnetic flow meters (SeaMetrics PE102 Flow Meter). The latter have a 1% relative error at 0.11-11.4 L min$^{-1}$. Both sensors register the discharge at 15-min intervals at each of six separate seepage sections at the down end of the slopes (section partitions located at -4 m, -2 m, -1 m, +1 m, +2 m and + 4 m relative to the center of the seepage face). The two sensors differ in their reliability for respectively low and high discharge flows. The NovaLynx sensors are set up for measuring low flows and will typically underestimate higher flow values. In turn, the PE102 sensors tend to be less reliable in measuring low flows, as they are calibrated for higher discharge values (more than 0.11 L min$^{-1}$). In the first 12-hour portion of the experiment during which rainfall was still applied, data from the NovaLynx tipping buckets were used. Data from the PE102 Flow Meters were used for the remainder of the experiment.

Hillslope-average values for SWC and MP were derived from their respective sensors located throughout the slopes. We averaged SWC and MP values for each interval of 15 minutes.

Because different water retention characteristics and saturated hydraulic conductivity in the landscapes might be responsible for the reported discrepancy in hydrological response times, we
derived soil water retention curves from SWC and MP sensor data. The Van Genuchten model (Van Genuchten, 1980; also see Sect. 2.4) was fitted to the observations by minimizing the sum of squared errors between observations and the fitted model. Porosity values of 0.395 for both landscapes were assumed as reported by Pangle et al. (2015) and the residual soil moisture content was set to zero. Values of the Van Genuchten parameters ($\alpha$ and $n$) were subsequently derived from the empirical curves.

The drainage tube associated with the central landscape’s seepage face section at –2 m was clogged over a period of approximately 13 hours (between 12 and 25 hours in the experiment), resulting in inaccurate discharge measurements. To correct for this data gap, a linear regression relationship between discharge measurements from the clogged section and a comparable unclogged section (located at +2 m relative to the seepage face center) was established during a period in which both were considered accurate (between 45 and 60 hours in the experiment).

2.5 Hydrological model

Because no overland flow occurred during the rainfall experiments, only the subsurface module of CATHY was used in this study. In the case of LEO, CATHY implements a numerical solution to the Richards equation (Richards, 1931), accounting for variably saturated porous media (Camporese, Paniconi, Putti, & Orlandini, 2010; Niu et al., 2014):

$$S_w S_s \frac{\partial \psi}{\partial t} + \phi \frac{\partial S_w}{\partial t} = \nabla \left[ K_s K_r (\psi) \left( \nabla \psi + \eta_z \right) \right]$$  \[2\]

where $S_w = \theta/\phi$ represents the relative soil saturation [$L^3 L^{-3}$], $\phi$ is the porosity [–], $S_s$ is the aquifer specific storage coefficient [$L^{-1}$], $\psi$ is the time [T], $\nabla$ is the gradient operator [$L^{-1}$], $K_s$ is
the saturated hydraulic conductivity tensor \([LT^{-1}]\), \(K_r(\psi)\) is the relative hydraulic conductivity function \([-]\), and \(\eta_z\) is a unit vector (0, 0, 1) with \(z\) measured vertically upward [L].

In this study, CATHY was set up in a similar fashion as described by Niu et al. (2014). The 30 \(\times\) 11.15 \(\times\) 1 m slopes were discretized into a grid of 60 \(\times\) 24 cells and 8 vertical layers. To better resolve infiltration and seepage, higher spatial resolutions (0.05 m) were assigned to the surface and bottom layers of the slopes. Unlike time stepping, this spatial grid in the slopes does not vary based on the number of iterations necessary to reach convergence and is thus constant throughout modeling the experiments.

Since evaporation (E) is not directly measured at LEO, we estimated E for modeling purposes through closure of the water balance expressed as \(\frac{dS}{dt} = P - E - Q\). Because of the availability of frequent measurements of \(S\), \(P\) and \(Q\), we could derive \(E\) for each time step. The effective precipitation \((P - E)\) was used in CATHY as the atmospheric boundary condition.

In order to find which parameters differ the most among the two landscapes and thus could be responsible for the different landscape responses, 350 preliminary model runs were conducted. In these simulations, the values of the input parameters \(\alpha^{-1}\), \(n\) and \(K_s\) were randomly varied within broad ranges of respectively \([-1.0 \text{ to } -0.05 \text{ (m)}]\), \([1.1 \text{ to } 3.3]\) and \([2.0 \cdot 10^{-5} \text{ to } 3.0 \cdot 10^{-4} \text{ (m s}^{-1})]\). These simulations were used to refine the parameter ranges used in the final calibration procedure. In that procedure, 1000 simulations of each landscape were obtained with CATHY.

Input values of the Van Genuchten parameters \(\alpha^{-1}\) and \(n\), and \(K_s\) were varied each time within adapted ranges of respectively \([-0.9 \text{ to } -0.05 \text{ (m)}]\), \([1.2 \text{ to } 3.0]\) and \([3.0 \cdot 10^{-5} \text{ to } 2.8 \cdot 10^{-4} \text{ (m s}^{-1})]\). Each model run was thus conducted with a randomized set of parameters, assuming soil homogeneity.

Model efficiency was calculated for each model run. We used an efficiency coefficient based on the Nash-Sutcliffe Efficiency coefficient (NSE) (Nash & Sutcliffe, 1970) and the more recent
Kling-Gupta coefficient (KGE) (Gupta, Kling, Yilmaz, & Martinez, 2009). These coefficients are respectively expressed as follows:

\[ NSE = 1 - \frac{\sum_{t=1}^{T}(Y_o^t - Y_m^t)^2}{\sum_{t=1}^{T}(Y_o^t - \bar{Y}_o)^2} \quad [3] \]

in which \( Y_o^t \) is the observed value of quantity \( Y \) at time \( t \) [T], \( \bar{Y} \) is the temporal mean of \( Y \) [T], and \( Y_m^t \) is the modeled value of quantity \( Y \) at time \( t \), and:

\[ KGE = 1 - \sqrt{(R - 1)^2 + \left( \frac{\sigma_m}{\sigma_o} - 1 \right)^2 + \left( \frac{\bar{Y}_m}{\bar{Y}_o} - 1 \right)^2} \quad [4] \]

in which \( R \) is the correlation between the observed and modeled series of quantity \( Y \) [–] and \( \sigma \) is the standard deviation of the modeled and observed values.

Since the model performance considering only the storage was very similar to the model performance considering only the discharge, we decided to use an aggregate measure for model efficiency. CATHY’s model efficiency coefficient therefore takes into account both total storage and total discharge and is expressed as follows:

\[ E = \frac{1}{4} \left( NSE_Q + NSE_S + KGE_Q + KGE_S \right) \quad [5] \]

where subscript \( Q \) denotes the time series of the total discharge for each landscape (measurements from the six locations together) and subscript \( S \) denotes the time series of the storage for each landscape.

The top 2% of model runs in terms of model efficiency coefficient \( E \) were retained as behavioral, meaning that we considered these model runs sufficiently fit to draw conclusions from them. This resulted in 20 parameter sets for each landscape. These were used to set up ranges of each
parameter for which CATHY is considered behavioral. The optimal set of parameters was used to obtain simulated plots of the landscapes’ storage over time and discharge over time. With the 19 other behavioral parameter sets the uncertainty around storage and discharge simulations was obtained. The uncertainty bounds are defined by the minimum and maximum model results per time step.

Furthermore, the obtained parameter values found through model calibration were used in conjunction with the Van Genuchten model to compose simulated soil water retention curves. The 19 other retained parameter sets were used to quantify the uncertainty. The results for the central and west cases were subsequently compared to each other and to the empirical water retention curves found through experiments.

3. Results

3.1 Observed discharge and storage dynamics

The uncorrected water storage and discharge observations from both landscapes are shown in Figure 2. Storage in both landscapes increased steadily and at the same pace for the duration of the rainfall event (0-12 h). However, shortly after the rainfall has stopped, the storage dynamics between the central and west slope started to differ significantly. Both landscapes’ storage decreased due to the discharge of water through their seepage faces, but the central landscape did so much faster. Consequently, the storage difference between the slopes increased over time, up to 40 mm after approximately 90 hours. This observation is echoed by the discharge rates. As rainfall stopped, discharge from both landscapes continued to increase, but the west landscape discharged water much slower than the central landscape. The discharge inconsistency of the central landscape between 12 and 25 hours in the experiment is explained by clogging of the seepage face. After the clogs were removed, discharge observations increased abruptly. The data
presented in Figure 2 were not used for modeling purposes as they were not corrected for the
clogging episode. The data set used for model calibration did include the linearly interpolated
data.

3.2 Hydrological modeling of discharge and storage dynamics

Figures 3a and 4a (central and west landscape) compare the observed water storage as a function
of time with model simulations. The simulated water storage with the highest model efficiency
coefficient $E$ is shown, as well as model uncertainty generated from results from the 19 other
behavioral model runs. Observed and simulated discharge rates as a function of time are shown in
the lower panels (Figures 3b and 4b). CATHY mostly succeeds in simulating the slopes’ water
storage over time as observed values are almost always within uncertainty margins. The
simulations of discharge (Figures 3b and 4b) show some retardation in incipient discharge flows
as the simulated onset of seepage flow lagged behind the observed onset by about three hours in
both landscapes.

3.3 Observations and simulations of soil water retention characteristics

Since the landscapes’ geometry, soil type and technical equipment are believed to be very similar
and any minor deviations herein not thought to be able to bring about such large differences in
hydrological response, the most plausible reason for the discrepancy in hydrological behavior is
likely to lie within the water retention characteristics and possibly the hydraulic conductivity of
the packed soil. Observed soil water retention curves were derived from all measurements of
SWC and MP throughout both landscapes. The results are shown in Figure 5, as well as the fit
according to the Van Genuchten model. While the curves of the two landscapes have a similar
shape, the central landscape’s soil has lower MP than the west landscape’s soil at identical
moisture conditions, suggesting that at any state of wetness, the water held within pore space
within the central landscape will be under less suction pressure than in the west landscape. Also, the values of the fitted parameters $\alpha^{-1}$ and $n$ differ. The central landscape’s observations are best fitted with $\alpha^{-1} = -0.323$ m and $n = 2.22$, whereas $\alpha^{-1} = -0.364$ m and $n = 1.94$ for the west landscape. Figure 5 also shows a clear hysteresis effect. When rainfall is applied to the landscape, SWC increases rapidly which explains why some data points are relatively far apart. After the experiment, the landscapes slowly dry up as they lose water through discharge. During the wetting phase, the matric potential at a given SWC is higher than during the drying phase. In addition to observations, approximately 1000 simulations were conducted with CATHY for each landscape (Table I). The best 20 simulations of the central landscape are achieved with parameter ranges of $\alpha^{-1} = [-0.257 \text{ m} \text{ to } -0.137 \text{ m}], n = [1.73 \text{ to } 2.09]$ and $K_s = [1.64 \cdot 10^{-4} \text{ m s}^{-1} \text{ to } 1.99 \cdot 10^{-4} \text{ m s}^{-1}]$, yielding a model efficiency coefficient of 0.956 on average. For the west landscape we found ranges of $\alpha^{-1} = [-0.573 \text{ m} \text{ to } -0.370 \text{ m}], n = [1.97 \text{ to } 2.60]$ and $K_s = [1.05 \cdot 10^{-4} \text{ m s}^{-1} \text{ to } 1.37 \cdot 10^{-4} \text{ m s}^{-1}]$ with an average model efficiency coefficient of 0.930. The single optimal parameter values are also included, as well as the corresponding average model efficiency coefficients.

The soil parameter values that yield the best model performance were used to compose ‘simulated’ soil water retention curves for the two landscapes (Figure 6). While the curves are similar in shape, the west landscape’s MP (Fig. 6a,c) is higher than the central landscape’s MP (Fig. 6b,c) under similar wetness conditions. The observed soil water retention curves match well with the simulated ones when the landscapes are dry (Fig. 6a-c). It appears that model performance is worse during wet periods, but this could not be tested extensively because of MP sensor saturation during wet conditions (SWC $> 0.18$ m$^3$ m$^{-3}$).

The calibration results reveal interesting differences between the two landscapes. Especially the values of $\alpha^{-1}$ show a remarkable variation; it seems that the optimal value of $\alpha^{-1}$ in the west
landscape is more than twice its value in the central slope in absolute terms and optimal parameter ranges as obtained through calibration do not overlap for the central and west landscape cases. The optimal values of the soil’s pore size distribution index $n$ also differ, as they are related to $a^{-1}$, but to a lesser extent as the optimal ranges for both cases show overlap. Any differences in the hydraulic conductivity $K_s$ between both landscapes are considered minor when compared to the difference in the values of $a^{-1}$.

4. Discussion
Post-experiment observations have indicated a clear difference in the hydrological response of LEO’s central and west landscapes. The west landscape retains artificial rainfall applied to the slope much longer than the central landscape. This observation is attributed to post-experiment discharge rate of the central landscape increasing much faster than the west landscape’s discharge rate. Simulations of the same experiments on both landscapes conducted with CATHY yield very similar results. Simulated discharge rates of the west landscapes are much lower than those of the central landscape and match well with observations.

Moreover, observed soil water retention curves of both landscapes indicate a substantial difference in soil water characteristics among the two landscapes. Measurements at the west landscape show much higher absolute MP values when compared to the central landscape at similar soil wetness. Simulated soil water retention curves composed with parameters derived from behavioral model runs paint an analogous picture. They match reasonably well with observations under dry conditions and therefore may support our hypothesis that different soil water retention characteristics are mostly responsible for the difference in hydrological response times in the central and west landscapes. Because MP values under saturated conditions were too
low for the MP sensors to measure, we were not able to compare observations with simulations over the entire range of soil moisture values reached during the experiments.

The higher observed and simulated absolute MP at constant SWC in the west landscape indicates that the west slope soil may have more fine pores. This could have led to lower discharge during and after experiments and could have caused the landscape to retain more water compared to the central landscape. As the soil drains, differences in absolute MP between the landscapes become substantial. This difference in soil water retention characteristic is reflected by the strong difference in observed and modeled values for the parameter $\alpha^{-1}$. Differences in $n$ are somewhat smaller and related to $\alpha^{-1}$ through the Van Genuchten equation and differences in $K_s$ seem minor. A greater degree of compaction of the west slope soil may be an explanation for this, but there is currently no evidence to support this explanation.

Another possible explanation for the found difference in soil parameters could lie in different particle size distributions. When the landscapes at LEO were constructed, the soil was stored in one pile. Fine particles could have settled to the bottom of the pile. If soil from the upper layer of the pile was used to fill one landscape and soil from the bottom of the pile to fill the other, it is conceivable that the particle size distributions in the two landscapes differ. However, we have no evidence to support this hypothesis.

This work has also shown that CATHY is capable of simulating both the landscapes’ storage and discharge at the level of the entire landscape. In addition, simulated water retention curves resemble observed ones in shape. It therefore seems that physical experiments conducted at these LEO hillslopes can be simulated well using CATHY. However, despite these good fits, there are considerable differences among the Van Genuchten parameters estimated from the measured SWC and obtained through calibration. We think the use of homogeneous parameters is
acceptable to retrieve integrated simulations of storage and discharge as presented in this study.

To successfully model local behavior, a more complex structure of these parameters is probably necessary (Pangle et al., 2017). The experiments described by Niu et al. (2014) most likely caused soil heterogeneity in LEO’s east slope as a result of heavy rainfall and this made it necessary to modify CATHY to assume a certain degree of soil heterogeneity in order to obtain acceptable model fits. Similar modifications to the model may be used in the future to obtain localized simulations or even better simulations of integrated storage and discharge.

This study is related to the work of Loritz et al. (2017) in that we also attempted to model a landscape using a limited number of soil parameters obtained through measurements at the hillslope scale. Their findings are similar to the ones presented here in that they concluded that meaningful simulations could be obtained despite the limitations of the model and approach.

In addition, while hysteresis seems to be a relevant process for LEO during the infiltration phase of experiments, this process cannot be modeled well using CATHY. The phenomenon of hysteresis is included in only a limited number of physically based models (Paniconi & Putti, 2015). Accurate inclusion of hysteresis in the model might lead to slower simulated infiltration and lower discharge peaks. More effort is necessary to be able to include hysteresis effects in CATHY and other physically based models.

Furthermore, this research has also shown that differences between observations and simulations continue to exist even at large-scale experiments conducted in fully controlled environments. The implications of this observation are twofold. First, it shows that our hydrological models are not yet fully developed and are still not always able to predict or simulate hydrological processes and soil water behavior, particularly at a limited scale. We were able in this research to model
observed rainfall-runoff response at the scale of the entire landscape quite accurately using CATHY, but good simulations of local hydrological behavior probably require a more complex structure of soil parameters. In addition, simulated soil water retention characteristics resulting from model calibration were not entirely in accordance with observations. Predicting catchment behavior using water retention characteristics will therefore continue to be challenging, since real-world sites are not usually as heavy-instrumented and are not located in controlled environments. Second, this study underlines the importance of soil water retention characteristics in explaining catchment rainfall-runoff responses. Apparently small deviations in soil water retention characteristics can bring about large differences in actual hydrological response. The influences of soil water retention characteristics on hydrological response are therefore not to be underestimated, since apparently their effects are difficult to isolate even in a controlled environment.

Given that LEO is located within a controlled environment and its landscapes are heavily instrumented, it would be possible to calibrate distributed models such as CATHY using measurements from the spatially distributed sensors. This could be achieved through Markov Chain Monte Carlo (MCMC) or data assimilation approaches considering soil heterogeneity throughout the landscapes. This has not yet been conducted as it was outside the scope of this work. However, this could well be subject of future studies since LEO seems a suitable environment for such a calibration procedure.

It is important to realize that many conclusions drawn from this research are based on just one experiment. Most results obtained during this research seem to be in accordance with each other
and explain the difference in observations, but analysis of a larger number of experiments may further support the conclusions drawn here.

Furthermore, it is necessary to acknowledge that the results gathered and discussed in this research rely on some important model idealizations and abstractions, in spite of LEO’s controlled environment. While LEO is unique in its field because of this fully controlled environment, it has to deal with some model assumptions. For instance, the Van Genuchten model has been assumed throughout this research. Although renowned in its field, the Van Genuchten model is highly empirical and features idealizations to simplify the equation. In fact, one of its goals was to reduce the number of parameters involved at the cost of accuracy (van Genuchten, 1980). This has a direct influence on many of the results of this research, where the comparison of Van Genuchten parameters in both landscapes is an important component.

5. Conclusions

From the analysis and modeling of storage and discharge time series, the conclusion is drawn that there is a significant difference in terms of hydrological response times between the central and west landscapes, mainly because of considerable difference in soil parameters. We believe that the west slope soil may have more fine pores, causing the post-experiment discharge from the west slope to be lower than from the central slope. This would explain the relative high observed absolute MP across the west slope at similar SWC. The hypothesis that soil water retention characteristics play an important role in the observed differences in hydrological response is supported by simulations of the experiments using CATHY. Calibration of the model shows that the central hillslope can be best simulated using a value $\alpha^{-1}$ of about $-0.20$ m, whereas a value around $-0.45$ m seems to be optimal in the west slope. The difference in this parameter, which may be related to the capillary fringe height, is substantial. While slight variations in optimal
values of $K_s$ were also found, these are considered to be minor. Therefore, an important conclusion of this study is that the central and west landscape feature different soil water characteristics in terms of the parameter $\alpha^{-1}$. While we do not yet have conclusive proof, the west slope soil may somehow be more compacted than the central slope soil. Differences in the packing of the soils may have resulted in differences in bulk density, which could explain the differences found in capillary fringe. Variations in the packing of the soils for the landscapes might originate from transport from the quarry to Biosphere 2, or during the construction of LEO. This was not tested in this research. Instead, it is left for future research to address this issue. A possibility would be to use isotope tracer experiments in an attempt to test this hypothesis. The most direct way to obtain evidence supporting or rejecting this hypothesis probably involves particle-size bulk-density measurements of repeated soil cores over time. In addition, we conclude that CATHY performs well in simulating subsurface flows at a hillslope scale. Its application may therefore be extended to small-scale catchment using parametrization from single hillslopes.

References


modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. Water Resources Research, 46(2). https://doi.org/10.1029/2008WR007536


Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I —
https://doi.org/10.1016/0022-1694(70)90255-6

https://doi.org/10.5194/hess-18-1873-2014

https://doi.org/10.1016/j.geomorph.2015.01.020


https://doi.org/10.1002/2015WR017780

https://doi.org/10.1023/A:1005567808533


https://doi.org/10.1002/hyp.1425


https://doi.org/10.1002/2015WR017032

https://doi.org/10.2136/sssaj1980.03615995004400050002x


Table I: Calibration results comprising optimal ranges and values for $\alpha^{-1}$, $n$ and $K_s$ for both landscapes. The (average) optimal model efficiency coefficient $E$ is also included, as well as the parameter values fitted to in-situ observations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Central range (simulated)</th>
<th>Central optimal (simulated)</th>
<th>Central optimal (observed)</th>
<th>West range (simulated)</th>
<th>West optimal (simulated)</th>
<th>West optimal (observed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha^{-1} (m)$</td>
<td>[−0.257 to −0.137]</td>
<td>−0.197</td>
<td>−0.323</td>
<td>[−0.573 to −0.370]</td>
<td>−0.444</td>
<td>−0.364</td>
</tr>
<tr>
<td>$n (-)$</td>
<td>[1.73 to 2.09]</td>
<td>1.88</td>
<td>2.22</td>
<td>[1.97 to 2.60]</td>
<td>2.25</td>
<td>1.94</td>
</tr>
<tr>
<td>$K_s (10^{-4} m s^{-1})$</td>
<td>[1.64 to 1.99]</td>
<td>1.79</td>
<td>−</td>
<td>[1.05 to 1.37]</td>
<td>1.19</td>
<td>−</td>
</tr>
<tr>
<td>Average $E$</td>
<td>0.956</td>
<td>0.965</td>
<td>−</td>
<td>0.930</td>
<td>0.941</td>
<td>−</td>
</tr>
</tbody>
</table>
Figure 1: Artist impression of the LEO project showing the three convergent landscapes within Biosphere2.
Figure 2: Timeseries of observed, uncorrected (a) water storage and (b) discharge of two model landscapes (center and west) of the Landscape Evolution Observatory over the course of the rainfall experiments.
Figure 3: Timeseries of (a) observed and simulated storage and (b) observed and simulated discharge in LEO’s central landscape. Shadings represent uncertainty margins.
Figure 4: Timeseries of (a) observed and simulated storage and (b) observed and simulated discharge in LEO’s west landscape. Shadings represent uncertainty margins.
Figure 5: Observed volumetric water content versus observed matric potential in both landscapes. The solid lines are the results from fitting the Van Genuchten model to the observations.
Figure 6: Observed volumetric water content versus observed matric potential for the central and west slopes. The Van Genuchten model was fitted to the observations using optimal values for $\alpha^{-1}$ and $n$ derived from CATHY simulations. The shaded areas represent uncertainty (variation resulting from the 20 best performing parameter sets).