Evidence of subsurface control on the coevolution of hillslope morphology and runoff generation

David G. Litwin[1,2], Ciaran J. Harman[2,3]

[1]Earth Surface Process Modelling, Helmholtz Center GFZ Potsdam, Potsdam, DE
[2]Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA
[3]Department of Earth and Planetary Science, Johns Hopkins University, Baltimore, MD, USA

Key Points:

- We test theoretical predictions about the relationship between hillslope length and relief, saturated area, and transmissivity
- Comparing two watersheds, we find lower transmissivity is associated with shorter hillslopes and larger variably saturated areas
- Hydrogeomorphic modelling suggests that subsurface properties drive the coevolution of differences between the sites

Corresponding author: David G. Litwin, `david.litwin@gfz-potsdam.de`

Abstract

Topography is a key control on runoff generation, as topographic slope affects hydraulic gradients and curvature affects water flow paths. At the same time, runoff generation shapes topography through erosion, which affects landscape morphology over long timescales. Previous modeling efforts suggest that subsurface hydrological properties, relative to climate, are key mediators of this relationship. Specifically, when subsurface transmissivity and water storage capacity are low, (1) saturated areas and storm runoff should be larger and more variable, and (2) hillslopes shorter and with less relief, assuming other geomorphic factors are held constant. While these patterns appear in simulations, it remains uncertain whether subsurface properties can exert such a strong control on emergent properties in the field. We compared emergent hydrological function and topography in two watersheds that have very similar climatic and geologic history, but very different subsurface properties due to contrasting bedrock lithology. We found that hillslopes were systematically shorter and saturated areas more dynamic at the site with lower transmissivity. To confirm that these differences were due to subsurface hydrology rather than differences in geomorphic process rates, we estimated all parameters of a coupled groundwater-landscape evolution model without calibration. We showed that the difference in subsurface properties has a profound effect on topography and hydrological function that cannot be explained by differences in geomorphic process rates alone. The comparison to field data also exposed model limitations, which we discuss in the context of future efforts to understand the role of hydrology in the long-term evolution of Earth's critical zone.

Plain Language Summary

In many humid landscapes, runoff is generated by water that flows through the shallow subsurface from ridges to valleys, eventually emerging and draining to rivers. The greater the capacity of the subsurface to move water, the more water can collect before surface runoff begins. Surface water may cause erosion, which shapes ridges and valleys over millions of years. We previously developed a computer model based on these principles and showed that the subsurface capacity to store and transmit water affects both runoff generation and topographic evolution. Lower capacity results in more surface runoff and shorter, lower relief hillslopes, when all other factors are held constant. Here we tested this by comparing two watersheds that differ primarily in their bedrock composition, which affects subsurface water storage and transmissivity. We found that the low transmissivity site had more dynamic surface runoff and shorter hillslopes, supporting our predictions. We set up computer models for both sites, which suggested that subsurface differences are necessary to explain observed differences in runoff and topography. Finally, we discuss some key limitations of the model that could be improved upon in future attempts to understand how hydrology affects the long-term evolution of Earth's surface.

1 Introduction

1.1 Background

It has long been understood that there is a close, two-way connection between runoff and the topographic form of landscapes. Topography influences flow paths of water over the surface and through the subsurface and supplies the elevation component of hydraulic head, while erosion by water shapes landscapes over long timescales. Horton (1945) first suggested that there is something valuable to learn about how places work hydrologically by considering this coupling. While Horton's work focused on the role of infiltration excess overland flow in determining contributing areas and drainage network topology, Carlston (1963) suggested that we should also be able to learn something about groundwater-driven runoff based on channel spacing. However, the vastly different timescales of runoff and evolution of channel networks via erosion has made it challenging to study the coevolution of hydrological and geomorphic states and fluxes. As a result, hydrologists studying runoff generation usually assume that landscape form is fixed, while geomorphologists studying landscape evolution usually assume hydrology can be reduced to a few parameters that capture how hydroclimate affects the efficiency of bedrock erosion and sediment transport.

Recent advances in modeling and the availability of high performance computers have allowed the coupling of hydrologic and geomorphic models that consider the evolution of hydrologic and geomorphic states together. Litwin et al. (2022) used a shallow aquifer model to generate saturation excess runoff from steady recharge, and used the runoff to drive fluvial incision in a streampower-plus-diffusion landscape evolution model. They showed that the thickness and permeability of the subsurface were important controls on runoff, and as a consequence, the degree of drainage dissection and length of hillslopes. Litwin, Tucker, et al. (2023) extended this model to examine the emergence of variable source area hydrology, adding stochastic precipitation and a simple representation of the vadose zone to the prior model to capture more realistic hydrologic dynamics. Again, the thickness and permeability were key controls on both the morphology and hydrological function of the coevolved landscapes. They showed that landscapes with efficient subsurface drainage and large water storage capacity had less variable and smaller saturated areas than those that had poor subsurface drainage, and therefore generated less storm runoff. This difference in runoff response has implications for geomorphology as well. Decreasing the spatial extent of runoff decreases the extent of fluvial erosion, which decreases the degree of drainage dissection. Litwin, Tucker, et al. (2023) also found an emergent relationship between runoff and morphology. Specifically, the fraction of quickflow relative to total discharge scaled inversely with the dimensionless hillslope relief in the watershed. This relationship supported prior predictions (Dunne, 1978) that steeper landscapes (with more transmissive soils) generated more runoff via subsurface flow, while landscapes with gentle topography (and thinner less transmissive soils) generated more runoff via saturation excess.

While these numerical results indicate that the subsurface is a key link between topography and runoff generation, it is unclear whether these relationships appear outside idealized models. While field studies have shown that subsurface properties and topography have effects on hydrologic function (e.g., Prancevic & Kirchner, 2019; Jencso & McGlynn, 2011), relationships between subsurface properties and topography remain elusive (Luo et al., 2016; Sangireddy et al., 2016), let alone unambiguous evidence that the link between them is the result of coevolution (Yoshida & Troch, 2016). This lack of clear relationships is to be expected because hydrology, conditioned by climate, is only one connection between the subsurface and topography. Other controls include lithology and tectonic setting, which affect the styles and efficiencies of weathering and sediment transport, and vegetation, which alters subsurface properties and sediment transport efficiency through root growth, and hydrologic partitioning through evapotranspiration (Brantley et al., 2017; Collins & Bras, 2010). If the subsurface connects topography and runoff generation despite all of this complexity, catchment coevolution may be a useful tool for understanding and predicting hydrological function (Troch et al., 2015).

1.2 Approach

If a signature of coevolution between topography and hydrological function exists, we will be most likely to find it where we can isolate the hydrological effects from other influences. We selected two sites where contrasting lithology results in a strong contrast in subsurface properties, but climatic and tectonic histories are similar because of their proximity. Our first site, Druids Run, is underlain by serpentine bedrock that forms thin rocky soil, while the second site, Baisman Run, is underlain by schist that weathers to form deep, permeable soil and saprolite. Assuming that the present hydrological function is adjusted to the watershed geomorphology, we drew on insights from Litwin, Tucker, et al. (2023) to hypothesize that:

1. Saturated areas and storm runoff are larger and more variable at Druids Run than Baisman Run, and
2. Hillslopes are shorter and have less relief at Druids Run than Baisman Run.

First, we characterized the hydrological function and morphology of the two sites and evaluated whether they support these hypotheses. To determine whether these differences could be the result of coevolution, we fully parameterized the landscape evolution model used in Litwin, Tucker, et al. (2023) without calibration. We determined the importance of subsurface hydrological differences by performing a simple sensitivity analysis in which we swapped the geomorphic process variables between the two sites and observed whether geomorphic process rates could explain differences in emergent morphology and hydrologic function.

2 Materials and Methods

2.1 Site descriptions

Our study sites are located in the Piedmont physiographic province, north of Baltimore, Maryland. The climate is humid, with a mean annual precipitation of approximately 1150 mm and mean annual potential evapotranspiration of approximately 750 mm. There is no pronounced seasonality in precipitation, less than 5% falls of which falls as snow. Baisman Run is a 381 ha watershed in Oregon Ridge Park, defined by an outlet at (39.4795 N, 76.6779 W). Druids Run is a 107 ha watershed located in Soldiers Delight Natural Environment Area, and is defined by an outlet at (39.4171 N, 76.8523 W). The watersheds are 16 km apart, and are at approximately the same elevation (52 m and 56 m above sea level respectively). Both watersheds drain to the Chesapeake Bay; Baisman Run drains via the Gunpowder River and Druids Run via the Patapsco River. Baisman Run has been monitored extensively as part of the Baltimore Ecosystem Study, and more recently as part of several projects aimed at improving understanding of deeply weathered critical zones (Putnam, 2018; Cosans, 2022). Druids Run has no prior description or study. It is unnamed in the National Hydrography Dataset, so we unofficially named it in honor of a local group of druids that meet in the watershed.

Baisman Run is underlain by the Loch Raven Schist (Crowley et al., 1975), a Cambrian-Devonian mica schist that has weathered to form deep, permeable soil and saprolite. Depth to weathered bedrock is greater than 200 cm in most of the watershed, below saprolite tens of meters thick at the ridge crests (Cosans, 2022). Above the saprolite, primary soils include Manor loam and channery loam, Glenelg loam and channery loam, and Manor-Brinklow complex in steeper slopes. Agriculture was historically present in the eastern headwaters, where there is now suburban development, and a homestead and tree farm were historically present in the Pond Branch sub-watershed (Cleaves et al., 1970). The remainder of the watershed has been relatively undisturbed since the 1950s and today supports a mature deciduous forest.

Druids Run is primarily underlain by the Soldiers Delight Ultramafite (Guice et al., 2021). Soils are primarily classified as chrome silt loam, and are generally thin with a strong permeability contrast at the base of the A horizon (at an average depth of 46 cm). Ridgetop soil is rocky and can be as thin as 5 cm, and exposed bedrock is common near channel heads. In valley bottoms, alluvium and organic material accumulate to thicknesses around 1 m. The Soldiers Delight Ultramafite is host to a “serpentine barrens” ecosystem, which consists primarily of grasses and shrubs with some areas supporting hardwood and conifer trees. The Soldiers Delight area was mined for chromite in the 19th and 20th century. Several small pits are present near ridge crests in Druids Run, and placer

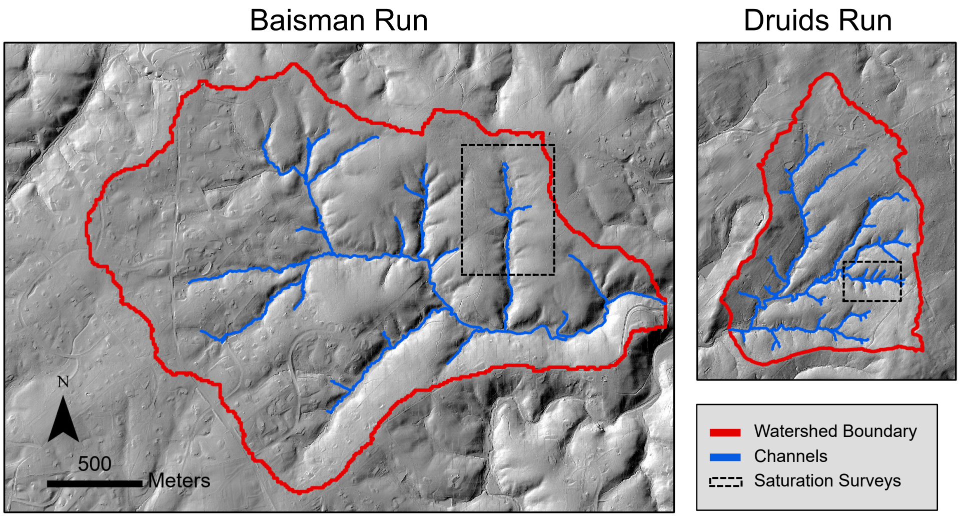

Figure 1. Hillshades of Baisman Run and Druids Run with the watershed boundary and channel network delineated with the DrEICH algorithm. Areas where we conducted saturation surveys (see Figure 4) are shown in dashed black boxes. The two sites are to scale, revealing the difference in their size and drainage dissection.

mining may have occurred in the valley bottoms, but the effects of this appear to be minimal in this watershed. Some structures and two small ponds are present in the upper portion of Druids Run, but most of the watershed is free from development.

2.2 Hydrological data

We combined existing hydrological data with new measurements of precipitation, streamflow, and saturated areas. Instantaneous precipitation rates were measured from June 2022 to February 2023 at a weather station located in an open field approximately 0.8 km north of Baisman Run. An identical unit was installed in an open area in Druids Run, which recorded instantaneous precipitation from April 2022 to February 2023. The stream gage at Baisman Run is operated and maintained by the U.S. Geological Survey (Gage 01583580). We established a new stream gage at Druids Run for this project.

The Druids Run stream gage is located at an existing concrete culvert crossing the stream channel. In April 2022 we installed a PVC housing on the concrete structure approximately 2 m from the culvert inlet. We measured water stage with a Solinist Levelogger pressure transducer within that housing, and corrected for atmospheric pressure with a Solinst Barologger. The pressure transducer operated until the device failed in October 2022. Periodic discharge measurements were made to construct a rating curve. Low flows were measured with salt dilution gaging recorded with a HOBO conductivity logger, and high flows were measured using an OTT MF Pro electromagnetic current profiler. A power law model fit the stage-discharge data well, as shown in Figure S1.

We surveyed limited areas of both watersheds manually for saturation conditions between April 2022 and March 2023. At Baisman Run, the surveys were conducted in the headwaters of the Pond Branch sub-watershed. At Druids Run, they were conducted in a headwater catchment near the eastern watershed boundary. We measured saturation at points along predefined transects, and returned to the approximate (but not exact) positions for each survey. We selected transects to balance capturing a range of hillslopes, zero- and first-order channels, while covering a small enough area to avoiding significant changes in saturation over the course of a measurement campaign. Saturation was measured by walking the transects, and pushing a rebar rod approximately 2 cm into the ground and moving the rod up and down in the shallow hole. Points along these transects were recorded as not saturated if no squishing sound was heard (N), soil-saturated if a squishing sound was heard (Ys), ponded (Yp), or flowing (Yf) if water was observed on the surface. Three close locations were measured at each point on the transect, and the highest category in this hierarchy was recorded as the value (e.g., if two points did not squish, but one did, the recorded class would still be Ys). This procedure was repeated under different discharge and moisture conditions.

2.3 Hydrological analysis

Valuable information about contributing areas can be extracted from rainfall and runoff timeseries. The event runoff ratio, defined as the ratio of the total event runoff to event precipitation, is an indicator of the proportion of the watershed that is contributing runoff during storms (e.g., O'Loughlin, 1986). To calculate event runoff, we first separated the discharge timeseries into baseflow and quickflow using the graphical approach described by Hewlett and Hibbert (1967). Baseflow is equal to discharge and quickflow is zero until discharge increases at a rate faster than 0.000546 m^3 s^{-1} km^{-2} h^{-1} (Hewlett & Hibbert, 1967). Baseflow continues to increase at this rate until discharge declines and is equal to baseflow. Storm events are periods where quickflow is greater than zero and the rise is associated with precipitation. We defined event precipitation as the total precipitation falling between a fixed time t_0 before the runoff event begins and a fixed time t_1 before the runoff event ends. By inspection of the timeseries, we found that $t_0 = 2$ hours and $t_1 = 1$ hour were appropriate for Druids Run, and $t_0 = 6$ hours and $t_1 = 2$ hours were appropriate for Baisman Run. We excluded runoff events shorter than 6 hours because these generally had small discharge responses relative to noise in the timeseries.

While the runoff ratio provides a signature of contributing area, the saturation dataset provides a direct means to assess the variability of saturated areas. The saturation surveys yielded categorical data that vary with topographic position and catchment discharge. To develop quantitative insights from the dataset, we first created a binary classification of whether points were not saturated (N) or saturated (Ys, Yp, Yf). We then used logistic regression to generalize our discrete measurements to predictions of how saturation probability varies with topographic (wetness) index (Beven & Kirkby, 1979) and discharge:

$$\log\left(\frac{p}{1-p}\right) = \alpha_0 + \alpha_1 \log\left(\frac{A}{v_0|\nabla z|}\right) + \alpha_2 \log\left(\frac{Q_b}{A_{tot}}\right) \tag{1}$$

where p is the probability of saturation, $TI = \frac{A}{v_0|\nabla z|}$ is the topographic index (note that here we do not include the log transform in the definition), Q is the discharge at the start of the saturation measurement campaign, A_{tot} is the watershed area, and α_0, α_1, and α_2 are model parameters.

2.4 Hillslope length and relief

We conducted geomorphic analyses using a lidar-derived digital elevation model with 0.76 m resolution, which was collected in 2015 and is publicly available from Baltimore County. We conducted all topographic analyses using LSDTopoTools (S. Mudd et al., 2022). To determine hillslope length and relief, we began by identifying the channel networks at both sites using the DrEICH algorithm (Clubb et al., 2014). DrEICH uses χ-analysis (Perron & Royden, 2013) to locate channel heads at the transition point from linear channel segments to nonlinear hillslope segments in χ-elevation space. χ-analysis is discussed in more detail in Section 2.7.2. We adjusted the DrEICH model parameters such that the predicted channel network matched the observed network in the subwatersheds where we surveyed saturated areas. We then used the channel network to identify hilltops, which are defined as edges shared by watersheds with the same Strahler stream order (Hurst et al., 2012). Finally, we calculated hillslope length as the steepest descent distance from each hilltop point to the nearest channel point, and hillslope relief as the hilltop elevation above the nearest channel point (Grieve et al., 2016).

2.5 Landscape evolution model

To test the link between hydrological and geomorphic features, we used the landscape evolution model described in Litwin, Tucker, et al. (2023). The model accounts for topographic evolution due to baselevel change, water-driven erosion using the stream-power erosion equation, and hillslope sediment transport using a nonlinear hillslope diffusion equation. We decided to use a linear diffusion formulation, as the hillslopes at Baisman Run and Druids Run generally remain convex until they reach valley bottoms, and the topography shows no evidence of shallow landsliding or other mass movements. The subsurface maintains constant and spatially uniform properties through evolution, implicitly assuming that the production of permeable material keeps pace with surface erosion. The overland flow that drives fluvial erosion is generated by exfiltration and precipitation on saturated areas in places where the shallow aquifer reaches the land surface. The shallow aquifer model uses the Dupuit-Forcheimer assumptions to calculate flow over a sloping impermeable base. The aquifer receives recharge from the vadose zone, which is represented as a single 1-dimensional profile in which discrete depth increments fill and drain by the plant-available water capacity in the increment. Recharge is calculated by determining the amount of water in the vadose profile that infiltrates below the water table depth at each point in the aquifer. The climate is treated as a simple random process, following Eagleson (1978), with exponentially distributed storm depth, duration, and interstorm duration, and constant evapotranspiration at the climatological mean rate during the interstorm periods.

We ran the model under the same initial and boundary conditions used in Litwin, Tucker, et al. (2023). The domain is square, and the bottom boundary is fixed to baselevel, while the remaining three side boundaries are zero-flux. This allows for the establishment of a drainage network with higher order streams than the same size domain where all boundaries are set to a fixed baselevel. In the absence of a known initial condition, we begin with a flat surface at baselevel. We ran the model for 50 Ma to approach dynamic equilibrium between erosion and uplift. While this timescale is long relative to periodic changes in climate and baselevel in the Eastern Piedmont (e.g., Cleaves, 1989), we know that both sites have experienced the same forcings through their evolution, such that a single climate and baselevel change rate should still provide insights into their evolution.

2.6 Hydrological parameters

2.6.1 Transmissivity, hydraulic conductivity, and permeable thickness

The maximum transmissivity, which we will just call the transmissivity, is defined as the depth-integrated saturated hydraulic conductivity. It appears in our model as the product of the effective saturated hydraulic conductivity k_s and permeable thickness b. We developed a novel method to use the saturation survey data to estimate a catchment-averaged transmissivity, building on an existing approach. We then divided that value into estimates of k_s and b.

Our method of estimating transmissivity is similar to that described by O'Loughlin (1986), as it is built on a steady state hillslope water balance and the assumption that places with the same topographic index TI saturate at the same time (Beven & Kirkby, 1979). The approach begins by considering recharge that is supplied at a rate $r(x, y)$ to the saturated zone. At hydrologic steady state, the total water outflow along a topographic contour segment with length v_0 is equal to the integral of recharge over the area upslope of the contour A_c. The maximum amount of recharge that can be moved through the subsurface before saturation occurs depends on the transmissivity T and the local hydraulic gradient, which is assumed to be equal to the topographic gradient ∇z. As a result, the criterion for saturation is:

$$\int_{A_c} r(x,y)dA \geq T|\nabla z|v_0. \tag{2}$$

At saturation, any additional recharge will become overland flow. Because in general the recharge is not known, O'Loughlin (1986) equated the total watershed recharge with the watershed baseflow Q_b:

$$\int_{A_{tot}} r(x,y)dA = Q_b, \tag{3}$$

where A_{tot} is the watershed area. From this expression, we derived an average recharge rate $\bar{r} = Q_b/A_{tot}$. Dividing Equation 2 by the average recharge rate equation and rearranging the terms, we derived an expression for the discharge-normalized transmissivity:

$$\frac{1}{|\nabla z|v_0}\int_{A_c}\left(\frac{r}{\bar{r}}\right)dA \geq \frac{T}{Q_b/A_{tot}}. \tag{4}$$

By further assuming that the integral in the above expression is approximately unity, we found an expression that relates the topographic index to transmissivity and baseflow discharge:

$$\frac{A}{|\nabla z|v_0} \geq \frac{T}{Q_b/A_{tot}}. \tag{5}$$

We will call the topographic index where saturation begins to occur TI^*, which is a function of discharge Q_b. Using a log transform, we derived an expression for the log of transmissivity:

$$\log(T) = \log(TI^*) + \log\left(\frac{Q_b}{A_{tot}}\right). \tag{6}$$

To find T using this expression and our saturation surveys, consider a logistic regression model with the form:

$$\rho(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \log\left(\frac{A}{v_0|\nabla z|}\frac{Q_b}{A_{tot}}\right) \tag{7}$$

where β_0 and β_1 are parameters of the regression model. This logistic regression model is very similar to that in Equation 1, but has one fewer parameter, and consequently enforces that the odds of saturation are log-linearly dependent on the product of Q_b and TI. At the critical value of topographic index TI^*, we will call the odds of saturation ρ^*:

$$\rho^* = \beta_0 + \beta_1 \log\left(TI^*\frac{Q_b}{A_{tot}}\right). \tag{8}$$

Finally, we rearranged Equation 8 to match the form of Equation 6, and solved for the transmissivity:

$$T = e^{(\rho^*-\beta_0)/\beta_1}. \tag{9}$$

The main difference between this approach and that described by O'Loughlin (1986) is that their approach equates the event runoff ratio with the proportion of the watershed that is saturated, while we have direct estimates of the saturated area. This should make our approach more robust, though it is still limited to the steady-state theory from which it was derived. Finally, we partitioned transmissivity between permeable thickness b and an effective saturated hydraulic conductivity k_s based on permeable thickness values taken from the USDA Soil Survey (Staff & Natural Resources Conservation Service, United States Department of Agriculture., 2023) and insights gained from prior subsurface investigations of Baisman Run.

2.6.2 Drainable porosity and plant available water content

Drainable porosity n_e relates the depth of water stored or released to the change in hydraulic head. Estimates usually require either hydraulic well tests or laboratory analyses. In the absence of hydraulic test data or permission to take soil samples from Druids Run, we assumed that the drainable porosity was the same at both sites. Plant available water content (n_a) is the amount of water, below the field capacity, that is available for plant use. The values were estimated based on the USDA Soil Survey data for the dominant soil types at the two sites.

2.6.3 Climatological parameters

We fit three independent exponential distributions for storm depth d_s, duration t_r, and interstorm duration t_b by calculating the mean values of these quantities from a precipitation dataset previously collected from 2014-2018 at the weather station at Baisman Run (Cosans, 2022). Because the two sites are very close together, this one timeseries was used to calculate storm statistics at both sites. The distributions are:

$$f(d_s) = \frac{1}{\langle d_s \rangle} \exp\left(-\frac{d_s}{\langle d_s \rangle} \right) \tag{10}$$

$$f(t_r) = \frac{1}{\langle t_r \rangle} \exp\left(-\frac{t_r}{\langle t_r \rangle} \right) \tag{11}$$

$$f(t_b) = \frac{1}{\langle t_b \rangle} \exp\left(-\frac{t_b}{\langle t_b \rangle} \right) \tag{12}$$

$$\tag{13}$$

where the angled braces indicate the temporal mean of the quantity. Potential evapotranspiration (ET) was estimated based on the average annual value in Baltimore between 1981 and 2010, as reported by the Northeast Regional Climate Center at Cornell University. In our model, ET only occurs during interstorm periods, so the interstorm potential ET rate *pet* was estimated by rescaling the average potential ET rate with the interstorm time fraction. Our climatological approach is simplistic, neglecting covariance of storm depth, duration, and interstorm duration, seasonality, paleoclimatic variability, and so on. However, we do not expect any large differences in the climate between the two sites, so even a simple approach should allow us to make comparisons of how landscapes with different geomorphic and subsurface hydrologic properties respond to climatic conditions similar to those observed at our sites.

2.7 Estimating geomorphic parameters

The topographic parameters of our model are the uplift or baselevel change rate U, hillslope diffusivity D, streampower incision coefficient K, characteristic contour width v_0, and the streampower exponents m and n, as discussed below. The Piedmont is thought to be in geomorphic steady state (Pavich, 1989; Bazilevskaya et al., 2013), so the regional rate of baselevel change was estimated the long-term erosion rate estimated with cosmogenic 10Be. The remaining parameters were identified using topographic analysis.

2.7.1 Hillslope diffusivity

Hillslope diffusivity can be derived from the rate of baselevel change U and hilltop curvature C_{HT} (Roering et al., 2007; Hurst et al., 2012):

$$D = -\frac{U}{C_{HT}}. \tag{14}$$

In hillslope evolution contexts, it is typical to account for the ratio of the bulk densities of regolith (on which the diffusion process occurs) and parent material (on which baselevel change occurs) (Roering et al., 2007). Because we are working with an integrated channel and hillslope model, and we do not have good estimates for the bulk density of fluvially-eroded material, we will neglect the bulk density terms. In this context, D is an effective diffusivity that will match the simulated hilltop curvature with that from our topographic measurements. We calculated the hilltop curvature by taking the second derivative of a polynomial surface fit to a 10 m footprint around each hilltop point. Hilltop points are the same as those used for the hillslope length analysis. The footprint size was selected by calculating the hilltop curvature for footprints of varying sizes and selecting the size at which there is a break in the standard deviation of curvature, following the procedure described by Hurst et al. (2012).

2.7.2 Streampower parameters

We estimated the streampower law parameters using an integral approach called χ-analysis (Perron & Royden, 2013). While the parameters can be derived from slope-area analysis, slope estimates often have significant noise that can result in poor parameter estimates (Perron & Royden, 2013). The integral approach is more stable, as it only requires the elevation and the upslope area to calculate the model parameters. The typical χ-analysis needed slight modification to accommodate our landscape evolution model. Litwin et al. (2022) derived the fluvial incision term of the landscape evolution model with assumptions that yielded linear dependence on the dimensionless discharge Q^*, a slope exponent $n = 1$, and area exponent $m = 1/2$. We derived a more general form by assuming that the exponent that determines the channel width from area and the exponent that determines erosion rate from shear stress were free parameters:

$$E_f = KQ^{*n} \left(v_0 a\right)^m |\nabla z|^n \tag{15}$$

where E_f is the fluvial incision rate, K is the erodibility, v_0 is the characteristic contour width, a is the area per contour width, and ∇z is the elevation gradient. For simplicity, we will use the variable Q^* to refer to the temporally-averaged dimensionless discharge which is called $\langle Q^* \rangle$ in Litwin, Tucker, et al. (2023). Because χ-analysis is usually only applied to river channels, it is typical to neglect the hillslope diffusion term, and write the solution at equilibrium between uplift and fluvial incision along a channel distance coordinate x:

$$U = KQ^{*n} \left(v_0 a\right)^m \left|\frac{\partial z}{\partial x}\right|^n . \tag{16}$$

We then solved for $|\partial z/\partial x|$, and substituted area for area per contour width times the characteristic contour width $A = v_0 a$:

$$\left|\frac{\partial z}{\partial x}\right| = \left(\frac{U}{KQ^{*n}}\right)^{1/n} A^{-m/n}. \tag{17}$$

Next we normalized upslope area to a reference drainage area A_0, and integrate the equation above with respect to x:

$$z(x) = z(x_b) + \int_{x_b}^{x} \left(\frac{U}{KQ^{*n} A_0^m}\right)^{1/n} \left(\frac{A_0}{A(x)}\right)^{m/n} dx \tag{18}$$

where $z(x_b)$ is the elevation at a specified baselevel location x_b. In general, Q^* varies with position, so we cannot remove it from the integral. However, in our model Q^* generally approaches a constant value with distance downstream equal to one minus the actual evapotranspiration relative to precipitation $1-\langle AET \rangle / \langle P \rangle$, which is approximately the mean runoff ratio $\langle Q \rangle / \langle P \rangle$. We will call this value Q^*_{max}. Then we can write:

$$z(x) = z(x_b) + \left(\frac{U}{KQ^{*n}_{max} A_0^m}\right)^{1/n} \chi, \tag{19}$$

where

$$\chi = \int_{x_b}^{x} \left(\frac{A(x)}{A_0}\right)^{m/n} dx. \tag{20}$$

These equations show that the elevation of a stream channel in dynamic equilibrium should be linear with respect to χ if U, K, and Q^*_{max} are uniform, and that the slope of that relationship should be:

$$k_{sn} = \left(\frac{U}{K Q^{*n}_{max} A^m_0} \right)^{1/n}, \tag{21}$$

which is often called the normalized channel steepness index. Note that this is related to but distinct from our use of "steepness" in Litwin et al. (2022).

We calculated the slopes of channel segments in χ-elevation space for the channel networks we extracted previously. Because the reference drainage area A_0 is introduced for dimensional purposes only, we can set it equal to unity, and solve for the streampower incision coefficient K:

$$K = \frac{U}{(k_{sn} Q^*_{max})^n}. \tag{22}$$

3 Results

3.1 Hydrologic and geomorphic observations

3.1.1 Discharge, baseflow, and runoff ratio

Figure 2 shows the timeseries of discharge and precipitation for both sites. Baseflow (in dark blue) at Baisman Run shows a strong annual signal, with drydown from early summer continuing until October, when a small persistent increase is combined with episodic increases in response to large storms. Unfortunately the discharge timeseries available to us at Druids Run is too short to look at annual trends, though there does appear to be a significant drydown from spring into summer, leading to low flows by late June. We did not observe no-flow conditions at the gage location, but we do know that flows were often close to or below the pressure transducer detection limit during the summer.

The storm runoff ratio is substantially more variable at Druids Run than Baisman Run. We identified 21 storm events at Druids Run and 43 storm events at Baisman Run, and found that the total event precipitation explained most of the variation in total event quickflow $Q_{f,event}$ (Figure 3). Events are colored by the antecedent baseflow, which shows that some of the variation in event runoff that cannot be explained by event precipitation may be explained by antecedent conditions. To quantify the sensitivity of event runoff to event precipitation, we fit the curve $Q_{f,event} = a_2 P^{a_1}_{event}$, where the log-space slope corresponds to the fitted exponent a_1. The exponent and standard error are 3.17±0.40 and 1.89±0.13 at Druid Run and Baisman Run, respectively. An exponent $a_1 = 1$ would indicate that the storm runoff is a constant proportion of the event precipitation. When the event runoff ratio is interpreted as the effective proportion of the watershed contributing runoff (O'Loughlin, 1986), an exponent closer to one indicates that the contributing area does not vary with storm size. This interpretation suggests that contributing areas vary with precipitation at both sites, but they are more variable at Druids Run than Baisman Run. This interpretation also suggests that as storm events approach 100 mm, nearly all of Druids Run contributes storm runoff (3A). These events are fairly frequent; the annual maximum recurrence interval of 100 mm of precipitation in 24 hours is approximately two years at our sites (NOAA, 2024).

3.1.2 Saturated areas

At Druids Run, observed saturation was highly variable in time and correlated with discharge. We measured saturation five times along nine transects, seven of which run along first order drainages or the interfluves between them, and two of which run parallel to the valley bottom (Figure 4A–E). The surveys conducted under the two highest flow conditions (C, E) had the greatest number of saturated points. Saturation was

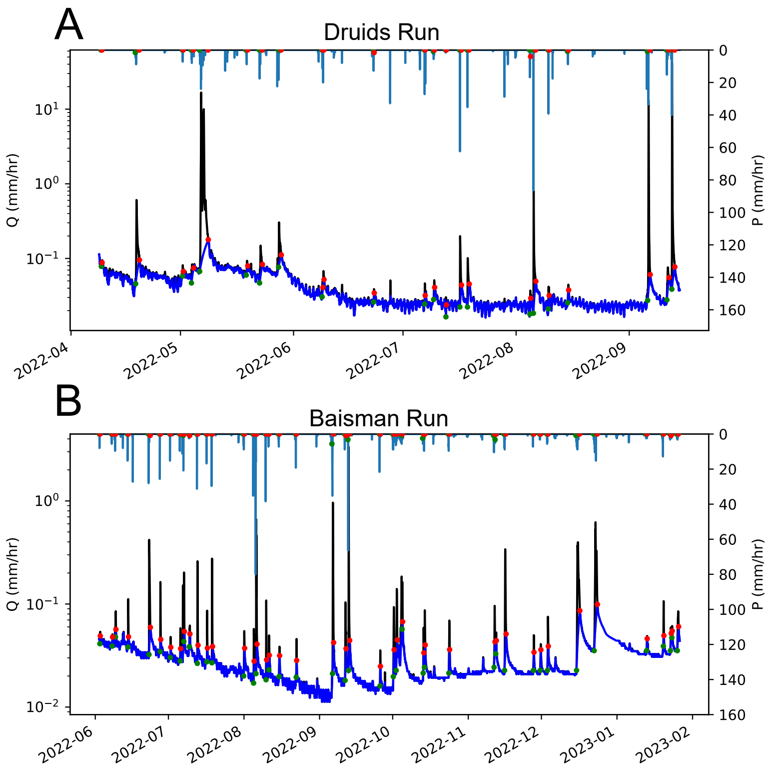

Figure 2. Timeseries of discharge Q (black), baseflow Q_b (dark blue), and precipitation P (light blue) at Druids Run (A) and Baisman Run (B). Storm events that we identified based upon the baseflow separation and precipitation begin with green dots and end with red dots, which are placed at the corresponding times on both the precipitation and discharge timeseries. Note that the timeseries for Baisman Run and Druids Run are not aligned in time.

often discontinuous with distance downstream in first order channels. Upslope areas sometimes saturated and flowed first, while downslope reaches remained dry, as flow passed through the subsurface. First order channels tend to have exposed bedrock or thin alluvial cover near their headwaters, while closer to the valley bottom they become submerged in alluvium that has sufficient capacity to move the water from upslope through the subsurface. Bedrock fractures may also play a role in redistributing surface flow to subsurface pathways.

In contrast, saturated areas were more static at Baisman Run. We measured saturation four times along six transects, four of which run perpendicular to the valley bottom, and two run parallel to it (Figure 4F–I). Regardless of discharge, we found that saturation was confined to locations at or below the distinct break in slope where the hillslopes meet the valley bottom. Within the valley bottom, saturation was not present everywhere, as the stream channel is incised into the valley bottom alluvium in some places. Flow emerges at distinct springs and seeps at the break in slope (Putnam, 2018). The springs are further evidence that subsurface pathways support baseflow, while the relatively static nature of saturated areas support our observation that event quickflow is less sensitive to event precipitation at Baisman Run than it is at Druids Run.

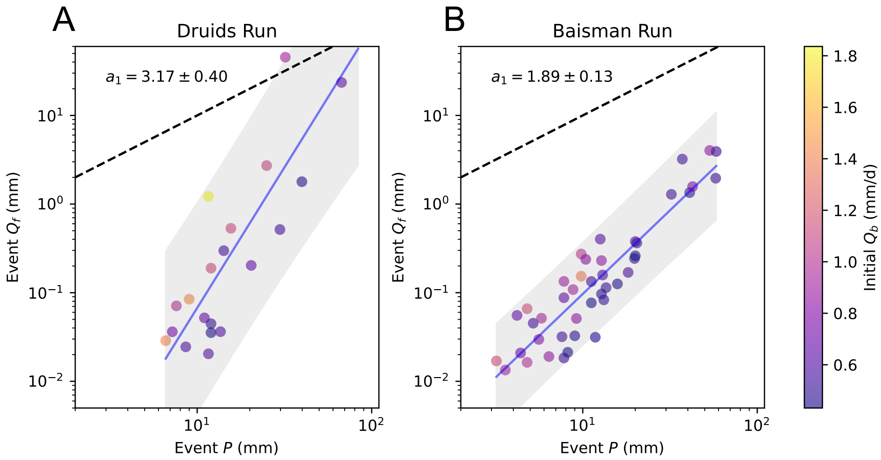

Figure 3. Event runoff characteristics for Druids Run (A) and Baisman Run (B). Event totals are calculated by summing the 15-minute precipitation and quickflow timeseries over the event durations. The points are colored by the initial baseflow Q_b. The dotted line is a 1:1 line, which represents the case where event runoff is equal to event precipitation. The blue line is a power law regression with the form $Q_{f,event} = a_2 P_{event}^{a_1}$, and the shaded area is the 95% confidence interval on the regression. The range on the coefficient a_1 is given as the standard error.

We generalized our point observations to whole-watershed predictions of saturation frequency using logistic regression. Specifically, we predicted the presence of saturation (flowing water, ponded water, or soil saturation) using topographic index and discharge using Equation 1. The parameters of the fitted model are shown in Table 1. To calculate topographic index, we first resampled the DEM to 5 m resolution to smooth over roughness in the high resolution DEM and to reflect the uncertainty in the positioning data of our saturation surveys. The resampling approach is also consistent with our measurement scheme, in which we labeled locations based upon the highest saturation class observed in a small vicinity. We calculated upslope area using the D_∞ algorithm, and slope using the same 10 m footprint used to calculate hilltop curvature. While our regression model calls for the use of baseflow discharge, we used the total discharge, as all of our samples were taken during baseflow or recession periods. This was also necessary because the timeseries of discharge at Druids Run does not overlap all the saturation surveys. For consistency, we used instantaneous discharge measurements from immediately before the surveys began. At Druids Run, we made these measurements using dilution gaging; Baisman Run, we used instantaneous discharge from the USGS gage.

	α_0	α_1	α_2
Druids Run	4.609 ± 0.637	0.174 ± 0.040	1.000 ± 0.097
Baisman Run	-7.487 ± 3.922	0.703 ± 0.103	0.081 ± 0.540

Table 1. Estimated parameter values of the logistic regression models for saturation (Equation 1), where α_0 is the intercept, α_1 is the coefficient on topographic index, and α_2 is the coefficient on the area-normalized discharge. Parameter ranges are given as standard errors.

We used the logistic model to predict the odds of saturation for the range of topographic index values in each watershed and the range of discharge values at which sat-

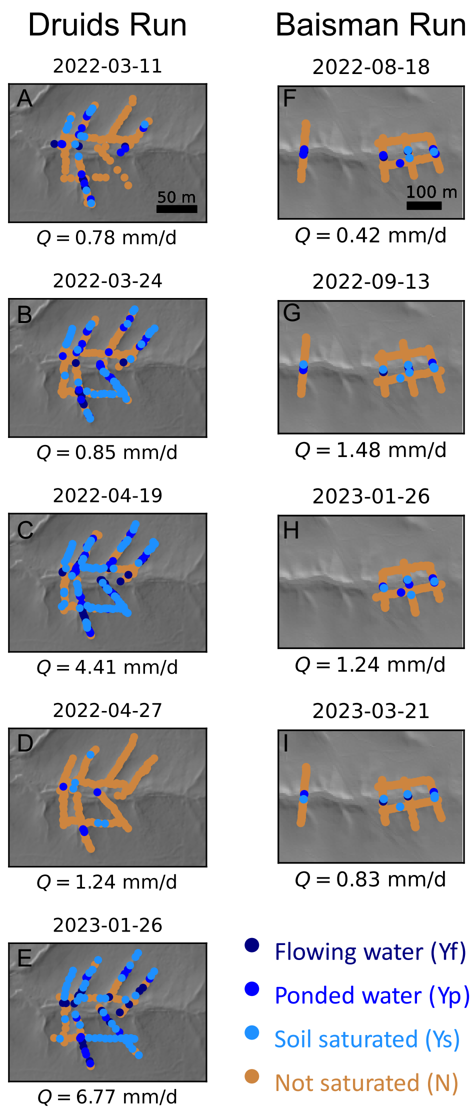

Figure 4. Observations of saturation made on transects at Druids Run (A–E) and Baisman Run (F–I). The latter plots have been rotated 90 degrees such that north is in the direction of the positive x-axis. In both figures, flow in the valley bottom is from right to left. The classification and sampling approaches are described in Section 2.2.

uration surveys were conducted (Figure 5). In Figure 5, the topographic index value at which the black dashed line intersects the odds ratio curves is the critical value of TI where saturation becomes more likely than not for a given value of discharge. We plotted this together with the probability density of watershed topographic index (orange) to show how the critical TI relates to the distribution of TI for the watershed.

The regression model for Druids Run in Figure 5A shows that the predicted odds of saturation varies substantially with discharge. When discharge is small, the critical TI value confines likely saturation to a very small portion of the total watershed area, while for large discharge the critical value of TI is low enough that most of the watershed is likely to be saturated. This supports the high variability of saturation in space and time that we inferred from the pointwise measurements.

The logistic regression model predicts very different behavior for Baisman Run (Figure 5B). First, we notice that the saturation odds curve does not vary with discharge, such that all curves overlap. This is reflected in the regression parameter $alpha_2$ on discharge (Table 1), which is much smaller and more uncertain for Baisman Run than Druids Run. As a result, the critical value of topographic index is nearly constant with time. Second, we notice that the curves are narrower and steeper than those estimated for Druids Run, such that the odds of saturation increases more abruptly around the critical value of TI. This is reflected in the regression parameter α_1 on topographic index, which is much larger at Baisman Run than Druids Run. This supports our observation that saturation emerges abruptly at the transition from hillslopes to valley bottoms.

The logistic regression models also allowed us to generalize the saturation predictions to the entire watersheds. We predicted saturation through time for the discharge timeseries in Figure 2 and for all raster points based upon their topographic index. We then classified whether each point was "wet" (exceeded criteria for saturation greater than 95% of the time), "dry" (exceeded criteria for saturation less than 5% of the time), or variably saturated if it met neither of those criteria.

Figure 6 shows a dramatic difference in the hydrological function of the two sites based on the logistic regression model predictions. The predicted channel network at Druids Run was ephemeral until close to the watershed outlet. Saturation occurred occasionally in zero-order basins and up onto the hillslopes. Some of the hillslopes we sampled that appear as "dry" may in fact saturate occasionally, but less than 5% of the time. In contrast, the regression model predicted that Baisman Run had a continually wet stream channel over the course of our observation period, and did not experience saturation on the hillslopes.

Analysis of rainfall-runoff and saturation data reveal the dramatic difference between hydrological function of the two sites. When the permeable subsurface is thin, as at Druids Run, much of the landscape saturates and desaturates relatively easily in response to precipitation, and the effective proportion of the watershed contributing runoff varies substantially. In contrast, when the permeable subsurface is thick, as at Baisman Run, the same precipitation causes modest or no change in saturated areas, though new subsurface flow paths may still be activated with increasing storm size, such that the effective contributing area increases with increasing wetness.

3.1.3 Hillslope length and relief

Both hillslope length and relief are greater at Baisman Run than Druids Run. The channel networks and hilltop points from which hillslope length and relief were defined are shown in Figure 1. Totals of 5.3×10^4 and 7.0×10^4 hilltop points with unique length and relief were identified at Druids Run and Baisman Run, respectively. The median hillslope length is 88.3 m at Druids Run and 177.3 m at Baisman Run, while median relief was 2.9 m at Druids Run and 6.7 m at Baisman Run. Figure 7A shows that there is no

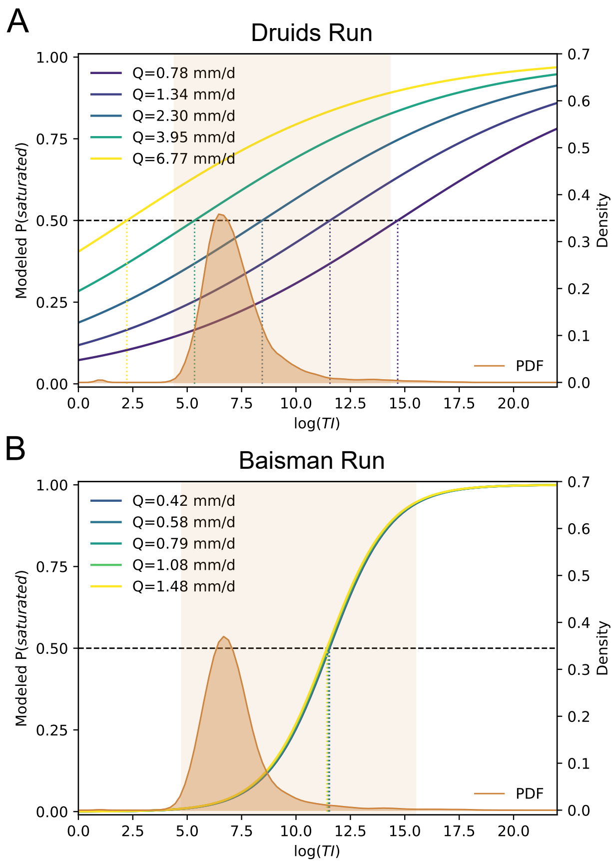

Figure 5. Regression results for Druids Run (A) and Baisman Run (B). The regression model has the form given in Equation 1. The modeled probability of saturation is given in terms of topographic index and discharge, where discharge varies logarithmically across the range of saturation survey discharge values. There is a dashed line at the 50% probability mark, and where this intersects each one of the probability curves, there is a dotted line dropped to the x-axis. This indicates the critical value of topographic index at which saturation is more likely than not to occur given that value of discharge. On the opposing axis is the probability density of topographic index, estimated with a kernel density approach. The lighter shaded region indicates the range of TI values sampled in our surveys, which indicates good topographic index coverage of our samples.

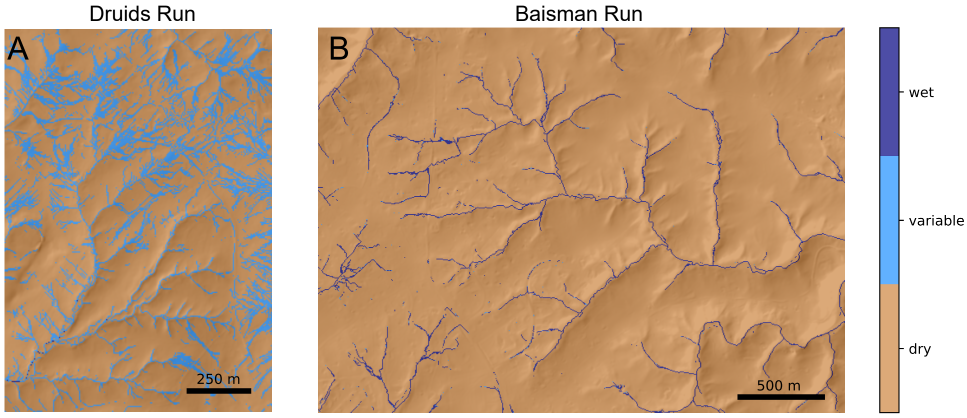

Figure 6. Classified saturated areas for Druids Run (A) and Baisman Run (B), based on the logistic regression model in Equation 1 and shown in Figure 5, and the runoff timeseries shown in Figure 2. The modeled probability necessary for saturation was set at 50%. A location was classified as "wet" if it exceeded criteria for saturation greater than 95% of the time, "dry" if it exceeded criteria for saturation less than 5% of the time, or variably saturated if it was in between. At Druids Run we predicted persistent saturation near the watershed outlet, an ephemeral channel network above that, and occasional saturation on some flat and concave hillslopes, and generally dry convex hillslopes. At Baisman Run, we predicted persistent saturation in the channel network, and dry conditions everywhere else.

overlap in the interquartile range (IQR) of hillslope length or relief for the two sites. The strength and sign of this difference supports our hypothesis that the site with a thick permeable subsurface will have greater hillslope length and relief than that with a thin permeable subsurface.

3.2 Landscape evolution parameterization

While both the hydrological and geomorphic differences between Druids Run and Baisman Run support our hypotheses, we have not yet established that the subsurface is the link between the emergent hydrological function and morphology. To do so, we estimated the parameters for DupuitLEM, and ran the model under conditions that approximate those found at our sites. Using the approaches described in Sections 2.6 and 2.7, we estimated all the parameters needed to run the model without calibration.

3.2.1 Hydrologic parameters

We first estimated the transmissivity using Equation 9. We estimated the parameters β_0 and β_1 by fitting Equation 7 using topographic index, discharge, and saturation survey data. With the fitted model, we determined the optimal threshold probability p^* at which saturation was likely to occur. While we could have chosen 50% as we did in the regression model for saturated area, we found that this performed poorly on the simpler two-parameter formulation used to calculate transmissivity. The selected value of p^* should balance correctly classifying points as saturated (high true positive ratio (TPR)) and minimizing the number of points that are misclassified as saturated (low false positive ratio (FPR)). Plotting TPR against FPR gives the receiver operating character-

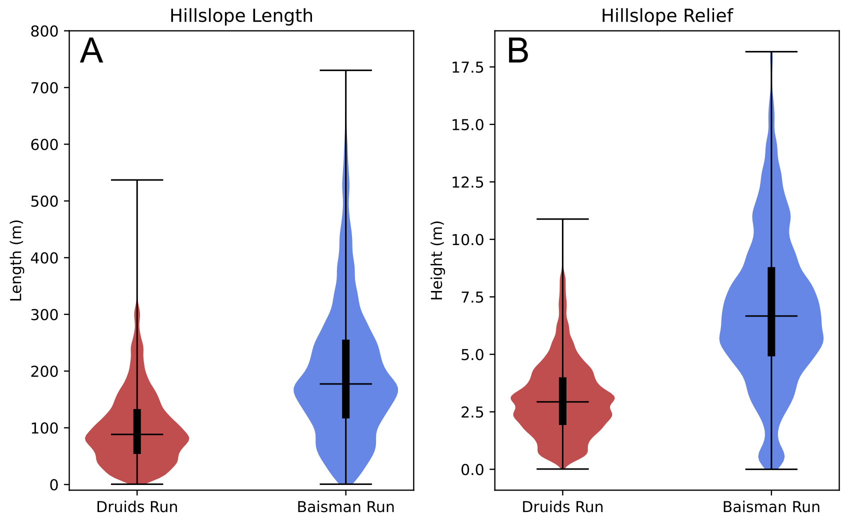

Figure 7. Violin plots of hillslope length and relief for Druids Run (A) and Baisman Run (B). Hillslope length is the length along a flow path from a hilltop point to the nearest channel point along a flowpath. Hillslope relief is the drop in elevation over that distance. Violin plots show the median, minimum, and maximum (horizontal lines) values and the interquartile range (wider vertical bar).

istic curve, from which we selected the optimal threshold probability by maximizing the difference TPR-FPR. The results of this process are shown in Figure 8. Using the optimal p^*, we estimated the transmissivity from Equation 9 10,000 times using Monte Carlo simulations to determine the uncertainty due to the variance and covariance of the logistic regression parameters. The median and quartiles of transmissivity are reported in Table 2. This approach predicts that the transmissivity at Baisman Run is nearly 8 times higher than at Druids Run. There is no overlap between the IQRs of the estimated transmissivities, which suggests a robust difference between the two sites. While the true uncertainty is likely much larger as a result of methodological choices (raster resolution, flow routing method, threshold selection method), experimentation suggested that the median transmissivity is always larger at Baisman Run than Druids Run when the same methodology is applied to both sites.

	Transmissivity (m^2/d)			Regression Parameters			
	Med	LQ	UQ	$\overline{\beta_0}$	$\overline{\beta_1}$	ρ^*	p^*
Druids Run	1.12	0.88	1.40	-0.691	0.268	-0.660	0.341
Baisman Run	8.46	7.07	10.23	-3.113	0.676	-1.668	0.159

Table 2. Median (Med), lower and upper quartiles (LQ, UQ) of transmissivity estimated from the logistic regression model, and the associated regression model parameters. The bar over a variable indicates the mean value.

To estimate the effective hydraulic conductivity from transmissivity, we first estimated the permeable thickness. At Druids Run, data from the USDA Soil Survey sug-

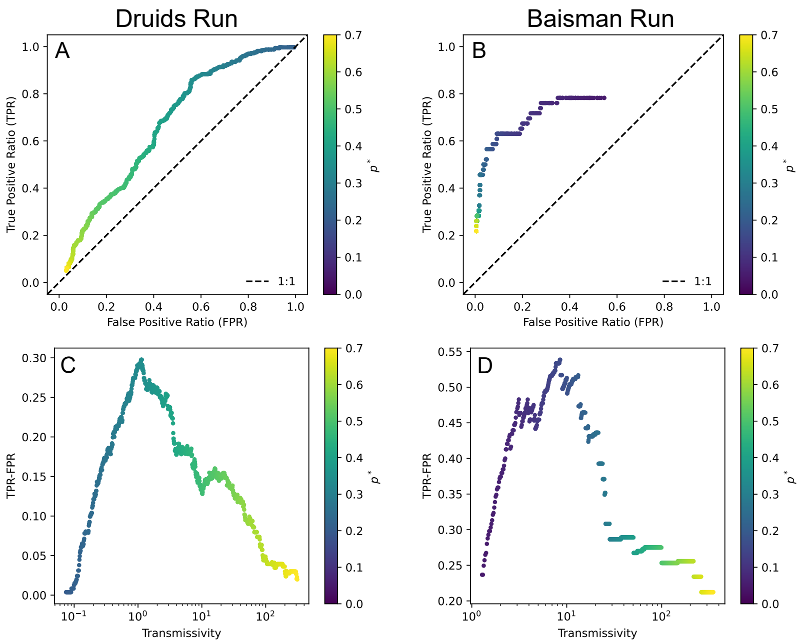

Figure 8. Results of the TPR-FPR analysis. (A–B) The receiver operating characteristic curve for Druids Run and Baisman Run, respectively, colored by the threshold value p^* used to obtain each combination of quantities. (C–D) The difference TPR-FPR, which we seek to maximize, plotted against the transmissivity value associated with each threshold p^*. We selected the transmissivity associated with the largest value of TPR-FPR.

gested a strong permeability contrast at the base of the A horizon, so we used the characteristic A horizon thickness as our permeable thickness (Staff & Natural Resources Conservation Service, United States Department of Agriculture., 2023). At Baisman Run, there is no strong permeability contrast within the soil profile, so we used the entire soil profile thickness, weighted for the different soil types found in the watershed. We added 2 m to this value to account for the importance of flow through the shallow saprolite (Cosans, 2022), which is below the maximum depth considered by the USDA Soil Survey. We divided transmissivity by the permeable thickness, and found that the effective hydraulic conductivity is similar between the two sites (2.83×10^{-5} and $2.43 \times 10{-}5$ for Druids Run and Baisman Run, respectively); the majority of the difference in transmissivity is due to the difference in permeable thickness. The values are shown in Table 3.

We estimated drainable porosity and plant-available water content from literature values. We assumed drainable porosity was constant and equal to 0.25 at both sites, which is typical for materials with medium sand to medium gravel texture (Johnson, 1967). While drainable porosity is an important variable for regulating the degree to which the water table rises and falls in response to recharge, it has a relatively narrow range of possible values in comparison to other parameters, so a possible difference between the sites should not have a strong effect on our results. We estimated plant-available water content as 0.19 and 0.14 for Druids Run and Baisman Run respectively using characteristic values for our sites from the USDA Soil Survey.

Lastly, climatological variables were estimated using the approaches described in the methods section with weather station data and literature values. The relevant values are shown in Table 3.

Overall, our characterization of hydrological properties is a rather coarse simplification of reality. Exponential distributions for precipitation do not capture the importance of extreme events (Rossi et al., 2016), while subsurface properties are often dynamic in time and vary both with depth and landscape position (e.g., St. Clair et al., 2015; Pedrazas et al., 2021). Still, our prior modeling work showed that even with these simplifications the model can produce rich and complex emergent hydrologic behavior (Litwin, Tucker, et al., 2023). Our approach here can serve as a starting place for future work that accounts for higher-order controls on runoff generation.

Name	Symbol	Units	Druids Run	Baisman Run
Hydraulic conductivity	k_s	m/s	2.84e−5	2.43e−5
Permeable thickness	b	m	0.46	4.03
Plant-available water content	n_a	-	0.19	0.14
Drainable porosity	n_e	-	0.25	0.25
Mean storm duration	$\langle t_r \rangle$	s	1.02e4	1.02e4
Mean interstorm duration	$\langle t_b \rangle$	s	1.11e5	1.11e5
Mean storm depth	$\langle d_s \rangle$	m	4.50e−3	4.50e−3
Interstorm potential ET rate	pet	m/s	2.58e−8	2.58e−8

Table 3. All hydrological parameters needed to run DupuitLEM. The values for n_e, $\langle t_r \rangle$, $\langle t_b \rangle$, $\langle d_s \rangle$, and pet are identical at the two sites.

3.2.2 Geomorphic parameters

The uplift or baselevel change rate U is an important model parameter and is needed to obtain estimates of both the hillslope diffusivity D and the streampower incision coefficient K. We equated U with the denudation rate estimated from in-situ ^{10}Be,, assuming that the Piedmont physiographic province is near dynamic equilibrium between baselevel change and denudation. Portenga et al. (2019) estimated the mean denudation rate of the Piedmont in the nearby Potomac River basin as 11.4 m/Myr (IQR 7.6 – 15.0) assuming an average rock density of 2700 kg/m^3. To quantify the uncertainty in U, and its contribution to the uncertainty in D and K, we estimated a probability distribution for U based on the box plot in Figure 4 of Portenga et al. (2019). The data did not appear particularly skewed, so we modeled denudation with a normal distribution, which we truncated to permit only positive values.

We estimated the diffusivity based on hilltop curvature, as presented in Equation 14. All the parameter values needed are shown in Table 4, and the distributions of the log of hilltop curvature are shown in Figure 9A. Hilltop curvature is quite similar at both sites. This is surprising since different processes likely contribute to diffusive transport at Druids Run versus Baisman Run. For example, freeze-thaw effects may be more important in the exposed, rocky soils at Druids Run, while treethrow may be more important in the forest-covered soils at Baisman Run. We estimated the diffusivity and its uncertainty by Monte Carlo simulation, sampling the distribution of U 10,000 times, and selecting 10,000 values from the hilltop curvature dataset independently with replacement. The distributions of diffusivity from the Monte Carlo simulation are shown in Figure 9B. The median diffusivity is 8.6e−3 m^2/yr (IQR 4.4e−3 – 1.7e−2) at Druids Run, and 9.3e−3 m^2/yr (IQR 4.3e−3 – 1.9e−2) at Baisman Run.

	C_{HT} (m^{-1})			U (m/yr)		
	Med	LQ	UQ	Med	LQ	UQ
Druids Run	−1.272e−3	−2.084e−3	−7.053e−4	1.193e−5	7.561e−6	1.495e−5
Baisman Run	−1.125e−3	−2.123e−3	−6.571e−4	1.193e−5	7.561e−6	1.495e−5

Table 4. Hilltop curvature C_{HT} and uplift U for Baisman Run and Druids Run. Negative curvature indicates convexity. Uplift values are the same for both sites.

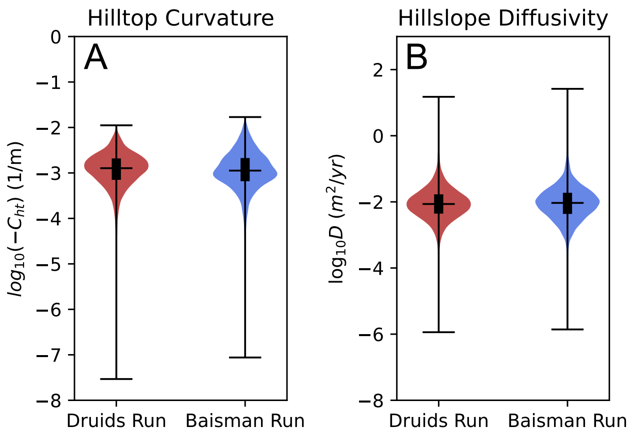

Figure 9. Violin plots of the log of hilltop curvature and log of hillslope diffusivity for Druids Run (A) and Baisman Run (B). Violin plots show the median, minimum, and maximum (horizontal lines) values and the interquartile range (wider vertical bar). Both distributions are similar, though Druids Run has slightly higher curvature, and therefore slightly lower diffusivity.

We calculated the streampower incision coefficient K using Equation 22 by estimating n, k_{sn}, and Q^*_{max}. We first conducted a χ-analysis of the channel networks of both sites to determine the streampower exponent n and then the appropriate steepness index k_{sn}. Lastly, we estimated the maximum dimensionless discharge Q^*_{max} based on available hydrologic data.

To calculate the optimal coordinate χ, we need to estimate the concavity index m/n (see Equation 20) for which the channel network collapses to a single line in χ-elevation space (Perron & Royden, 2013). We tried a range of values for the concavity index and determined that $m/n = 1/2$ produced a satisfactory collinearity of channels for both of the sites. Independently estimating the exponents m and n is challenging (Harel et al., 2016), so we chose the combination $m = 1/2$ and $n = 1$ for consistency with our prior modeling studies.

We determined k_{sn} from the slope of the relationship between χ and elevation for individual channel segments using the method described by S. M. Mudd et al. (2014). We estimated K using the segments that are above the 40^{th} percentile of channel network drainage area, which are colored by k_{sn} in Figure 10A–B. We selected this drainage area cutoff to isolate channel segments where Q^* is less likely to vary with distance downstream. We found that channel segments with smaller upslope areas were often less linear in χ-elevation space, which may indicate a change in Q^* with area. Figure 10C shows the distribution of k_{sn} values that meet these criteria. We found that k_{sn} was nearly twice as high at Druids Run, with a median of 2.774 (IQR 2.163 – 3.284), as Baisman Run, with a median of 5.23 (IQR 4.747 – 7.017).

	k_{sn} (m)			Exponents (-)		Runoff (-)
	Median	LQ	UQ	m	n	Q^*_{max}
Druids Run	2.774	2.163	3.284	0.5	1	0.3
Baisman Run	5.230	4.747	7.017	0.5	1	0.3

Table 5. Channel steepness index k_{sn}, streampower exponents, and maximum runoff rate Q^*_{max} for Baisman Run and Druids Run.

We estimated the maximum dimensionless discharge Q^*_{max} at Baisman Run as the long-term average runoff ratio $\langle Q\rangle/\langle P\rangle = 0.3$ (Cosans, 2022). From our short timeseries at Druids Run, we calculated a runoff ratio of 0.57. Because k_{sn} depends on the product of K and Q^*_{max} (Equation 21) in our model, these data suggest that the factor of two difference in k_{sn} between our sites could be due to the difference in the hydrology, expressed in Q^*_{max}, rather than a difference in material and geomorphic properties, expressed in K. While that would support our hypothesis, we will conservatively set $Q^*_{max} = 0.3$ for Druids Run as a first estimate, matching Baisman Run.

With all components of Equation 22 estimated, we used the same Monte Carlo procedure to calculate K and its uncertainty. Figure 10D shows that K is substantially higher at Druids Run than at Baisman Run when Q^*_{max} is set equal. The median at Druids Run is 1.34e−5 yr^{-1} (IQR 8.24e−6–1.98e−5), while at Baisman Run it is 6.49e−6 yr^{-1} (IQR 3.83e−6 – 9.66e−6). The full table of geomorphic parameters are shown in Table 6.

Name	Symbol	Units	Druids Run	Baisman Run
Uplift rate	U	m/yr	1.143e−5	1.143e−5
Hillslope diffusivity	D	m^2/yr	8.611e−3	9.285e−3
Streampower incision coefficient	K	1/yr	1.334e−5	6.546e−6
Contour length	v_0	m	30	30

Table 6. Geomorphic parameters needed to run DupuitLEM. We used the median value from the estimated parameter distributions for U, D, and K. The values for U and the characteristic contour length v_0 are identical at the two sites.

The difference in streampower incision coefficient between the two sites potentially confounds our interpretation of subsurface hydrologic controls on emergent hillslope length and hydrological function, assuming the difference is due to a contrast in material properties rather than hydrology. Our estimated subsurface hydrological variables support our perceptual model of how the sites should be different if they have coevolved with their hydrology; lower transmissivity at Druids Run should lead to more surface runoff and channel incision, and greater extent of variably saturated areas than the high transmissivity conditions at Baisman Run. However, a higher streampower incision coefficient may indicate that runoff is more effective at detaching and transporting sediment out of the watershed at Druids Run, which could also lead to closer spacing of channels and shorter hillslopes (Perron et al., 2008).

To test whether subsurface hydrology is necessary and sufficient for explaining the difference in variable source areas and hillslope length at the two sites, we ran four simulations, shown in Figure 11: two that represent our best estimates of hydrological and

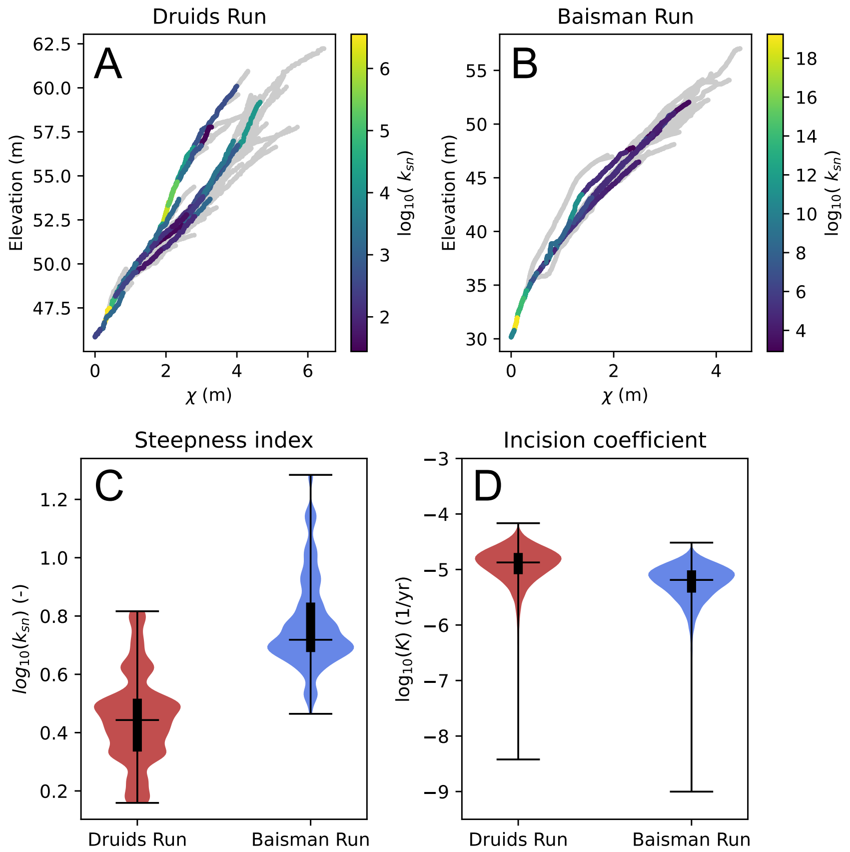

Figure 10. χ-elevation plots for Druids Run (A) and Baisman Run (B) for a concavity index $m/n = 0.5$. Channel segments are colored by their steepness index k_{sn} where the upslope area is greater than the 40$^{\text{th}}$ watershed area percentile, and are otherwise gray. (C) the distributions of k_{sn} for the segments colored in (A) and (B), showing generally higher channel steepness at Baisman Run than Druids Run. (D) distributions of the streampower incision coefficient K from Monte Carlo simulations. k_{sn} scales inversely with the erodibility, such that the streampower incision coefficient is lower at Baisman Run than Druids Run.

geomorphic parameters as described above (DR-DR, BR-BR), and two where we swapped the geomorphic parameters (DR-BR, BR-DR). Our best estimate cases helped discriminate how well DupuitLEM can capture landscape geomorphic and hydrologic dynamics at our sites. By comparing the best estimate simulations with simulations that have the same hydrological parameters but swapped geomorphic parameters, we determined whether geomorphic process rates alone explained the differences in morphology when the landscape coevolves with hydrology. Because we started with a randomized rough surface as an initial condition, we did not expect the simulation results to look exactly like Druids Run or Baisman Run. Instead, we compared them on the basis of aggregate properties including the hillslope length and relief, and saturation behavior.

		Hydrologic Variables (k_s, b, n_a)	
		Druids Run (DR)	**Baisman Run (BR)**
Geomorphic Variables (K, D)	**DR**	DR-DR	BR-DR
	BR	DR-BR	BR-BR

Figure 11. Four boxes indicating the four simulations we conducted. Colored boxes indicate the correctly matched hydrologic and geomorphic parameters, while white boxes indicate the ones in which the geomorphic variables are swapped. The listed hydrological and geomorphic variables are those that are varied, while all others are kept the same.

Lastly, we considered what happens when the differences in observed channel steepness were due to differences in runoff ratio (Q^*_{max}) rather than material properties (K). In our model formulation, determining the right value of Q^*_{max} should be an iterative process, in which the value of Q^*_{max} is estimated in order to determine erodibility, the model is run forward, the discharge and precipitation from the simulated landscape are used to recalculate Q^*_{max}, and then the streampower incision coefficient is adjusted accordingly. This would be repeated until the estimated Q^*_{max} value matches the value produced by the simulation. If there is a mismatch, the channel steepness of the modeled topography will be offset from that measured at the site. While we did not do a complete iterative solution, we did adjust Q^*_{max} and K according to the results of our first simulation.

3.3 Landscape evolution results

The landscape evolution model results showed the important effect of subsurface hydrology on the emergent landscapes, and revealed the complexity of interactions between hydrologic and geomorphic processes. We first simulated topography for the four cases presented in Figure 11, and analyzed the hillslope properties and persistence of saturated areas using the same criteria as we used for the field sites. The only necessary difference was that we identified channel heads using a threshold on topographic curvature ($\nabla^2 z > 0.001$), because the DrEICH algorithm performed poorly on our model simulations, which are much lower resolution than the lidar-derived DEMs. Because the transmissivity is the primary difference in hydrological variables, we call the cases with hydrology like Druids Run (DR-DR and DR-BR) the low transmissivity cases, and cases with hydrology like Baisman Run (BR-BR and BR-DR) the high transmissivity cases.

The most striking pattern in the hillshades shown in Figure 12A is that the low transmissivity cases were substantially more dissected than the high transmissivity cases. DR-DR and DR-BR have extensive fluvial dissection that extends onto hillslopes, which appears more extensive than we observed at Druids Run. However, the broad undissected hillslopes in BR-BR and BR-DR are similar to what we observed at Baisman Run. Despite some visual similarities, Figure 12B–C shows that BR-BR and BR-DR cases tended to overpredict hillslope length and relief. Also, contrary to our expectations, in the low transmissivity cases where the geomorphic properties have been swapped (DR-DR versus DR-BR), the difference in hillslope length and relief appeared to be comparable to the difference between Baisman Run and Druids Run (for a better view of length and relief at the field sites, see Figure 7). However, the presence of fluvial dissection broadly across these modeled topographies makes direct comparison with our field sites more difficult. When the transmissivity is large, the channel network is very well defined, and we found less apparent effect of the difference in geomorphic parameters. While the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of hillslope length at BR-DR are smaller than those at BR-BR, their medians are approximately the same (Figure 12B).

Swapping geomorphic parameters had a relatively minor effect on hydrological function. Figure 13A shows that simulations with swapped geomorphic parameters but the same hydrologic parameters have very similar saturated area patterns, whereas there is a substantial difference between simulations that have different hydrologic parameters. The low transmissivity cases have large variably saturated areas that extend onto hilltops, as at Druids Run, though there are no hilltops that are classified as dry in the low transmissivity cases. They also show more persistent saturation in valley bottoms and zero-order basins than observed in Druids Run (Figure 13A–B). The saturated areas modeled in the high transmissivity cases look very similar to those observed at Baisman Run, where there is persistent saturation in valley bottoms and dry hilltops. The fractional saturated areas are similar to those observed at the sites as well (Figure 13B).

Next we examined the emergent runoff ratio and adjusted the fluvial parameters to account for the difference between the runoff ratio and the initial estimate of Q^*_{max}. The emergent runoff ratio for the high transmissivity cases were 0.33 and 0.32 for BR-BR and BR-DR respectively, which were very close to our initial estimate of 0.3, which was the observed runoff ratio at Baisman Run. The difference in geomorphic parameters had little effect on emergent runoff ratio in these cases. In the low transmissivity cases, the runoff ratio was significantly higher than our initial estimate of 0.3. We found runoff ratios of 0.86 and 0.81 for DR-DR and DR-BR respectively. These values are again not highly sensitive to the difference in geomorphic parameters, but both are substantially higher than our initial estimate, and higher than our field estimate of 0.57 for Druid Run. However, this is consistent with our observation that DR-DR and DR-BR have much more extensive saturated areas than Druids Run. These higher runoff ratios suggest that we should increase estimated Q^*_{max}, and therefore decrease the estimated K at Druids Run. If we increase Q^*_{max} to 0.6, the corresponding K values is 6.68e$-$6 yr^{-1}, which is within 3% of the K value we estimated for Baisman Run. The geomorphic results of this increase are shown in Figure 14. The hydrologic effect of this increase is minimal, as shown in Figure S2.

Adjusting the streampower incision coefficient for differences in Q^*_{max} nearly eliminates the difference in emergent morphology and hydrology between cases with swapped geomorphic parameters. The hydrological function of the landscapes is very similar when geomorphic parameters are swapped, which is expected given that there was little difference in hydrological function between the original cases with swapped geomorphic parameters. The emergent runoff ratio for DR-DR is now 0.78, which is slightly lower than we calculated previously. The emergent topography looks very similar when geomorphic parameters are swapped, and distributions of hillslope length and relief are nearly identical (Figure 14). This suggests that the differences in the geomorphic parameters, and

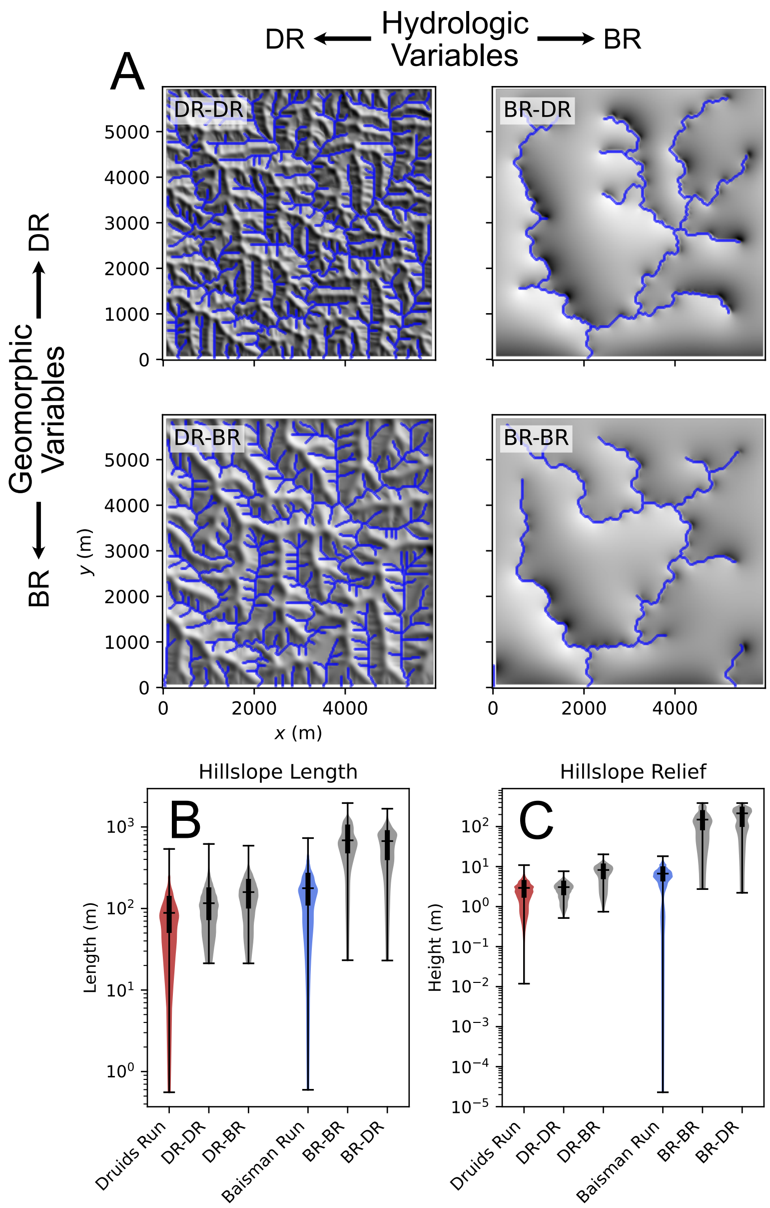

Figure 12. (A) Hillshades of model results in the same configuration as shown in Figure 11. Dissection is substantially higher in cases with Druids Run hydrological variables than Baisman Run hydrological variables. (B, C) Log-scaled violin plots of hillslope length and relief, comparing the field data (labelled "Druids Run" and "Baisman Run") to the four modeled cases. Horizontal lines represent the maximum and minim values, while the vertical bar represents the interquartile range.

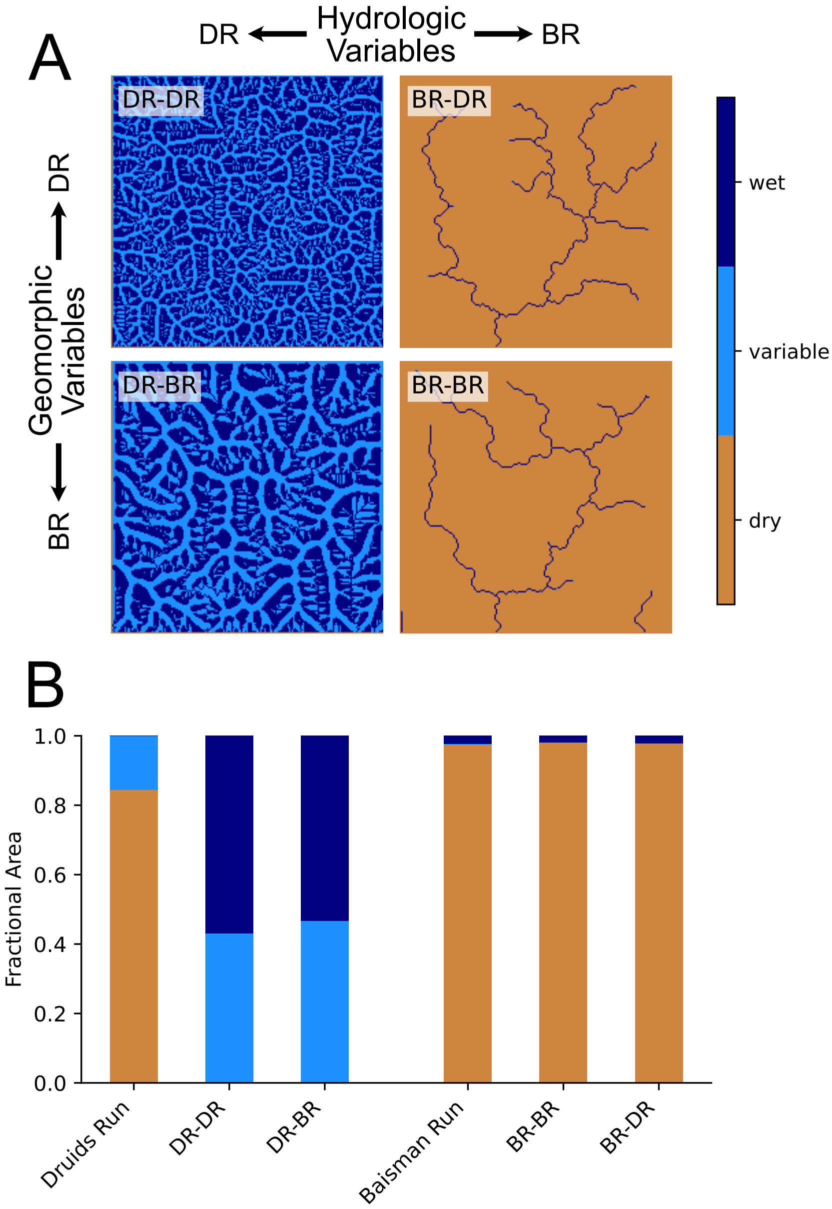

Figure 13. (A) Map view of saturated area classes for model results in the same configuration as shown in Figure 11 and Figure 12A. Saturated area behavior is not highly sensitive to swapping geomorphic variables, while it is sensitive to swapping hydrological variables. (B) Fractional area that is classified as wet, variable, and saturated based on field data (labelled “Druids Run” and “Baisman Run”) and the four modeled cases. Cases that have the hydrological variables associated with Baisman Run appear similar to the field characteristics of Baisman Run. Cases that have the hydrological variables associated with Druids Run show more persistent saturation than the field characteristics of Druids Run.

in particular the intrinsic erodibility of the rock and regolith, are not responsible for the differences in emergent morphology. Instead, what we see in the difference in morphology between Druids Run and Baisman Run is more likely a combination of effects driven by the difference in their subsurface hydrology, as (1) the difference in transmissivity changes the extent of saturated areas and surface water on the landscape, which changes the proportion of the landscape that experiences fluvial erosion, and (2) higher runoff ratios increase the efficiency of water-driven sediment transport in areas where there is saturation, which further incises the landscape.

Our results also showed that there is more work to do to understand the controls on the geomorphic evolution of our sites. For instance, adjusting Q^*_{max} did not bring us closer to the true hillslope length and relief. Figure 15 shows how the true cases DR-DR and BR-BR compare to the hillslope length and relief of Druids Run and Baisman Run, respectively. The number in parentheses following the model label is the estimated value of Q^*_{max}. The values of hillslope length and relief from simulation DR-DR (0.6) were farther from the true values at Druids Run than those from simulation DR-DR (0.3). At the same time, we know that the channel steepness k_{sn} from the simulation DR-DR (0.3) will not match k_{sn} of Druids Run, because we overestimated the streampower incision coefficient K relative to the emergent value of Q^*_{max}. More work is needed to understand both the possible difference in other parameters (e.g., the denudation rate) and limitations of model structure for capturing our sites, but it is clear that the difference in the hydrology of the sites is an important component of their geomorphic evolution.

4 Discussion

4.1 The expression of subsurface hydrology in landscape evolution

It is well known that transmissivity affects the hydrological function of landscapes. All distributed hydrological models built on Darcy's law will show a similar effect; the transmissivity, or more generally the depth-integrated hydraulic conductivity, affects the aquifer thickness and hydraulic gradient needed to convey a given water flux. This in turn determines how the water table will interact with the surface and produce overland flow (e.g., Beven & Kirkby, 1979; Li et al., 2014; Nippgen et al., 2015; Marçais et al., 2017). While there are limits to the Darcian approach for landscape scale runoff generation (e.g., Uchida et al., 2005), it has proved useful for understanding and predicting runoff, subsurface transport, and saturated areas.

Previous work toward understanding the role of transmissivity in topographic evolution (Luijendijk, 2022; Litwin et al., 2022; Litwin, Tucker, et al., 2023) is a logical extension of the hydrological study of runoff generation, as sediment transport is an important consequence of runoff generation. It has only recently received attention, in part because considering the long-term effects of this coevolution is computationally intensive, and in part because it relies on subsurface properties that are hard to estimate. Often, landscape evolution modelers select the minimally-complex model needed to explain their observations. As a result, they have often excluded subsurface hydrology, despite the widespread importance of subsurface flow for runoff generation (Wu et al., 2021). However, we have shown here that there are some cases where the subsurface hydrology is indispensable for understanding the evolution of landscapes. The importance of subsurface runoff generation for a particular application of a landscape evolution model is dependent on the geologic and climatic setting, but also on the scale of interest. Studies focusing on watershed scales of 1-10s of kilometers may find that capturing subsurface flow is essential, while these details may be less important in the evolution of entire orogens, where the length of subsurface flow paths relevant to runoff generation is shorter than the scales of geomorphic interest.

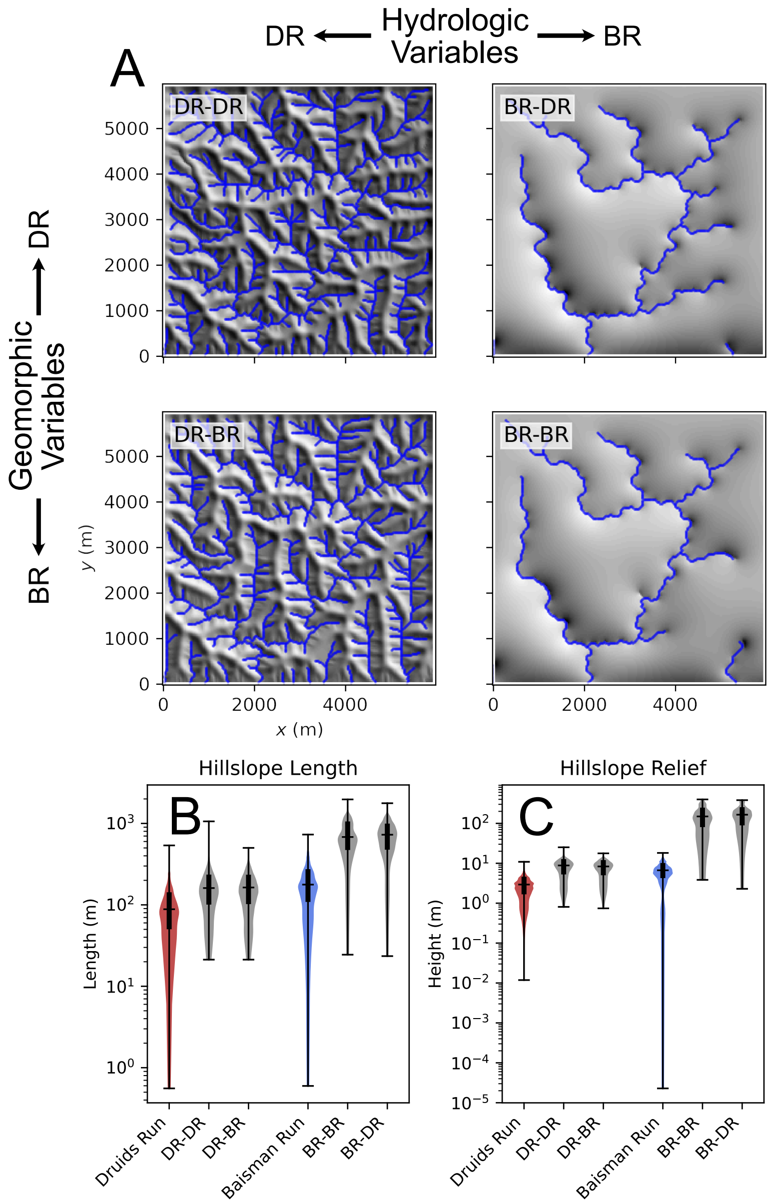

Figure 14. (A) Hillshades of model results in the same configuration as shown in Figure 11, only $Q^*_{max} = 0.6$ was used to determine the streampower incision coefficient for cases with Druids Run geomorphic variables. Visual comparison of results suggests that the difference in hydrology between the two sites is the primary control on emergent morphology. (B, C) Violin plots of hillslope length and relief, comparing the field data (labelled "Druids Run" and "Baisman Run") to the four modeled cases. There is little difference between simulations with swapped geomorphic variables (comparing down columns), while there is still substantial sensitivity to swapped hydrological variables (compare across rows). All four modeled cases still have length and relief greater than those observed in the field.

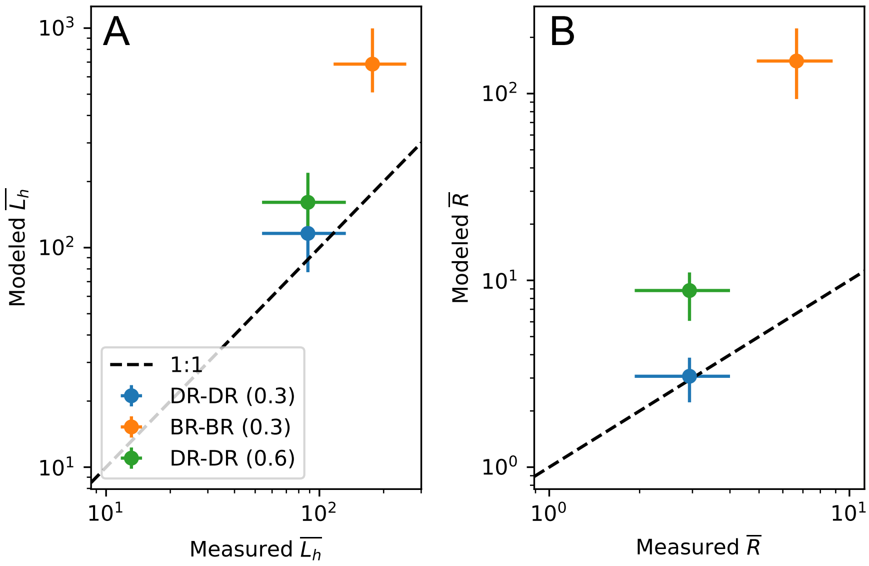

Figure 15. Modeled versus observed hillslope length (A) and hillslope relief (B). Results are only shown for the true cases (not the swapped parameter cases). The number following the simulation name is the value of Q^*_{max}. For Druids Run, simulations using both the original and updated estimates of Q^*_{max} are shown. The points are median values, and the error bars show the interquartile range.

We were able to show that hydrology was indispensable for understanding our sites, in part because we had hydrologic data for comparison, rather than just surface topography. Adding the hydrological dimension can help get the right answer for the right reasons in landscape evolution models, or discriminate when we have not gotten the right answer for the right reasons. This kind of approach could be useful beyond one-to-one site comparisons. For instance, we might be able to examine the effects of hydrological versus geomorphic processes on landscape evolution and hydrologic function across landscapes with different lithologies, by using rainfall-runoff relationships or other hydrological indicators that are more widely available than saturation. Constraining subsurface properties is still challenging, but methods like the logistic regression approach we presented here may be useful, especially as they are improved and refined.

4.2 Parameter estimation and limits of DupuitLEM

While our results provide evidence for a critical link between subsurface hydrology and landscape evolution, there are clear discrepancies between the characteristics of Baisman Run and Druids Run that we observed and those we were able to model with DupuitLEM. Some of these discrepancies could be due to our choice of model parameters, while others appear to be structural limitations of DupuitLEM.

Our results here and in prior studies (Litwin et al., 2022; Litwin, Tucker, et al., 2023) demonstrate that emergent topography and hydrology are highly sensitive to transmissivity, so the accuracy of the transmissivity estimate is likely a factor in model–data discrepancies. Our novel approach to estimate transmissivity relied on topographic index as a measure of hydrological similarity (O'Loughlin, 1986; Beven & Kirkby, 1979). However, our results showed that topographic index and discharge, when combined in Equation 7, were only modestly good predictors of saturated area (Figure 8). Furthermore, topographic index is a resolution-dependent quantity (Zhang & Montgomery, 1994), which means that the resulting transmissivity that we calculate also depends on DEM resolution. While accounting for this effect is unlikely to change the relative magnitudes of transmissivity between the sites, it may change the estimated values. This was a problem with calibrated transmissivities in TOPMODEL as well (Beven, 1997), so some of the strategies that have been devised to reduce the scale dependence in that context (e.g., Saulnier et al., 1997) may be useful for improving our transmissivity estimates as well.

Our model results also showed that hillslope length and relief were too large in the simulated landscapes regardless of transmissivity. This could suggest that the relative magnitude of hillslope diffusivity to the fluvial erosion efficiency is too large (Perron et al., 2008; Theodoratos et al., 2018). Our modelled cases are generally able to reproduce observed hilltop curvature (Figure S3A), which suggests that the diffusivity is not the primary issue. Modelled channel steepness, however, is systematically larger than the channels from which the parameters were defined (Figure S3B). One likely issue that could explain this discrepancy arises from using K estimates from 1D channel profiles in a 2D model. Hillslopes in the 2D model contribute material to valleys that rivers must remove. This decreases their erosional efficiency compared to what is expected when estimating K from a 1D profile in which the river only needs to erode at a rate U (Equation 16). This topic requires further exploration than can be accommodated here, and will be covered in future work.

While there are limitations to our ability to estimate transmissivity and other process rates, we know that some key hydrological and geomorphic processes and features are missing from DupuitLEM. For example, DupuitLEM does not have a pathway for evaporation or transpiration of water once it has reached the saturated zone. Especially in cases where the water table is close to the surface, evaporation of saturated zone water is likely a significant control on hydrologic dynamics. Including it would decrease the proportion of the watershed that stays saturated during interstorm periods and decrease antecedent wetness when storms arrive.

Our simulations were also limited to cases where the subsurface thickness is uniform across the landscape. We know this may not generally be the case. In Baisman Run, deeply weathered zones under hillslopes delay the arrival of hillslope water to streams and support baseflow, while a relatively shallow subsurface in valley bottoms may increase the likelihood of overland flow in the channels (Cosans, 2022; St. Clair et al., 2015). This pattern could increase flow persistence and drainage dissection relative to a uniform subsurface. In contrast, very thin soils on hillslopes at Druids Run allow saturation and overland flow to occur frequently, while a more permeable valley bottom may increase the subsurface conveyance in valleys relative to the amount of water that remains after storms. Depending on how the riparian area is connected to the stream, it may also store more water that can be slowly released during interstorm periods. These patterns could increase or decrease saturated areas and drainage dissection, depending on the extent of the riparian aquifer and its stream connection. In addition to shaping subsurface structure, weathering can also result in significant chemical denudation. DupuitLEM, like many geomorphic models, assumes all denudation occurs by physical erosion. This limitation is discussed in the next section.

4.3 Chemical weathering and landscape evolution in the Eastern Piedmont

We were not the first to be interested in the contrast between landscapes on schist and serpentine bedrock in the Piedmont. Cleaves et al. (1974) also made this comparison, using Pond Branch (a subwatershed of Baisman Run), and a small watershed on the Soldiers Delight Ultramafite that is south of our site. Pond Branch had been studied previously (Cleaves et al., 1970), while this paper introduced the Soldiers Delight study site. The focus of their study was to contrast the roles of physical versus chemical denudation in the two terrains, and identify the hydrological signature of the deep saprolite that is present on the schist. Cleaves et al. (1974) identified a hydrological signature very similar to what we observed: the schist bedrock site had baseflow-dominated runoff, which was highly persistent through droughts due to the large volume of storage in the saprolite. In contrast, they observed that the site on the ultramafic bedrock generated more quickflow, and had highly variable baseflow discharge, which they attributed to the lack of saprolite.

In examining the morphologies of the two sites, they determined that there was no strong evidence for differences in overall rates of denudation between the terrains. However, on the basis of a geochemical mass balance they determined that chemical weathering was responsible for approximately 90% of denudation in Soldiers Delight at present, while it was responsible for approximately 50% of denudation at Pond Branch. This would suggest a significant difference in how we should interpret the resulting morphologies, as we assumed that all denudation was due to physical erosion. Some recent work begins to provide a framework for understanding morphologic effects of chemical denudation. Ben-Asher et al. (2019) introduced a modification of the hillslope mass balance that includes chemical denudation in the form of a chemical depletion fraction (CDF). They showed that curvature should be reduced as the ratio of chemical to total denudation increases, assuming a constant hillslope diffusivity. Marcon (2019) applied this principle to several hillslopes on contrasting lithologies across the Piedmont, including sites on schist and serpentine bedrock. They found decreasing hilltop curvature with increasing CDF, where serpentine sites had the highest CDF and lowest hilltop curvatures. Interestingly, at our sites we found virtually no difference in hilltop curvature between lithologies. If the total denudation rate at both sites is indeed very similar, but chemical denudation is dramatically different, we are left with the conclusion that the identical curvature is a coincidence that arises from higher hillslope diffusivity D at Druids Run than Baisman Run. Dissolution could also have significant effects on river profiles, with consequences for interpretations of channel steepness and streampower incision coefficient K. Further research, including updated denudation estimates specific to our sites, would be needed to draw further conclusions.

4.4 Toward surface–subsurface critical zone evolution

Here we sought to understand how subsurface properties condition runoff generation and, as a result, topographic evolution. Many other processes that build and structure the critical zone, which extends from the top of the canopy to the bottom of the circulating groundwater, are closely coupled (Brantley et al., 2007; Troch et al., 2015). Besides the hydrological link, other studies have shown that subsurface properties can influence morphology by limiting the size and effectiveness of sediment to do erosional work (Callahan et al., 2019; Brocard et al., 2016). Others have investigated critical zone evolution from fresh bedrock up to the surface, and have shown some aspects of how climate, hydrology, and geomorphology condition the evolution of the subsurface (e.g., Rempe & Dietrich, 2014; Harman & Cosans, 2019; Anderson et al., 2019). There is a lot of work to be done to understand the feedbacks between surface and subsurface evolution, and how they produce, across a range of climates and lithologies, relatively similar patterns of drainage networks and soil-mantled hillslopes. At the same time, we have an increasingly detailed picture of critical zone structure. High-resolution topographic data has given us detailed insights into geomorphic processes acting at the surface (Sofia, 2020), while near-surface geophysics has allowed us to peer into the subsurface and begin to test models of subsurface evolution (Riebe et al., 2017; Parsekian et al., 2021). Our hope is that future work will consider the importance of feedbacks between subsurface hydrology and topography as we go forward in our understanding of critical zone structure and evolution.

5 Conclusions

We framed this paper with two hypotheses about how the morphology and hydrological function of two landscapes should be different, informed by our understanding of how the subsurface affects coevolution of runoff and topography. We found that both the hydrological function and morphology aligned with our predictions. Druids Run, which has a thin permeable subsurface, had more extensive variably saturated areas, more variable effective area contributing runoff, and shorter hillslopes than Baisman Run, which has a deep permeable subsurface. An analysis of the available field data further showed that the transmissivity was substantially higher at Baisman Run than Druids Run. While these findings support our hypothesis that coevolution with subsurface hydrology is important for emergent morphology and hydrological function, they did not in themselves provide a causal link. To test that link, we used a landscape evolution model with groundwater flow to show that the differences in geomorphic process rates (the hillslope diffusivity and streampower incision coefficient) were insufficient to explain the differences in morphology and hydrological function we observed. At the same time, we found discrepancies between the calibration-free model results and field data, which we discussed in the context of both parameter estimation challenges and model structure. The methods we explored here could serve as the basis for future study to uncover the importance of subsurface hydrology for the evolution and hydrological function of landscapes.

6 Open Research

All original data, model output, and scripts needed to process data and generate figures are archived on Zenodo (Litwin & Harman, 2024). The Python package DupuitLEM v1.1 (Litwin, Barnhart, et al., 2023) contains the models and scripts used to generate and post-process the model output. Landlab v2.0 (Barnhart et al., 2020) is a core dependency of DupuitLEM.

Acknowledgments

This work was supported by National Science Foundation grants EAR-2012264 and EAR-1654194. Model simulations were carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC-1920103. We appreciate help in the field from Joseph Stanley.

References

Anderson, R. S., Rajaram, H., & Anderson, S. P. (2019). Climate driven coevolution of weathering profiles and hillslope topography generates dramatic differences in critical zone architecture. *Hydrological Processes*, *33*(1), 4–19. doi: 10.1002/hyp.13307

Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., ... Bandaragoda, C. (2020). Short communication: Landlab v2.0: A software package for Earth surface dynamics. *Earth Surface Dynamics*, *8*(2), 379–397. doi: 10.5194/esurf-8-379-2020

Bazilevskaya, E., Lebedeva, M., Pavich, M., Rother, G., Parkinson, D. Y., Cole, D., & Brantley, S. L. (2013). Where fast weathering creates thin regolith and slow weathering creates thick regolith. *Earth Surface Processes and Landforms*, *38*(8), 847–858. doi: 10.1002/esp.3369

Ben-Asher, M., Haviv, I., Roering, J. J., & Crouvi, O. (2019). The potential influence of dust flux and chemical weathering on hillslope morphology: Convex soil-mantled carbonate hillslopes in the Eastern Mediterranean. *Geomorphology*, *341*, 203–215. doi: 10.1016/j.geomorph.2019.05.021

Beven, K. (1997). TOPMODEL: A critique. *Hydrological Processes*, *11*(9), 1069–1085. doi: 10.1002/(SICI)1099-1085(199707)11:9<1069::AID-HYP545>3.0.CO;2-O

Beven, K., & Kirkby, M. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, *24*(1), 43–69. doi: 10.1080/02626667909491834

Brantley, S. L., Eissenstat, D. M., Marshall, J. A., Godsey, S. E., Balogh-Brunstad, Z., Karwan, D. L., ... Weathers, K. C. (2017). Reviews and syntheses: On the roles trees play in building and plumbing the critical zone. *Biogeosciences*, *14*(22), 5115–5142. doi: 10.5194/bg-14-5115-2017

Brantley, S. L., Goldhaber, M. B., & Ragnarsdottir, K. V. (2007). Crossing Disciplines and Scales to Understand the Critical Zone. *Elements*, *3*(5), 307–314. doi: 10.2113/gselements.3.5.307

Brocard, G. Y., Willenbring, J. K., Miller, T. E., & Scatena, F. N. (2016). Relict landscape resistance to dissection by upstream migrating knickpoints. *Journal of Geophysical Research: Earth Surface*, *121*(6), 1182–1203. Retrieved from `https://onlinelibrary.wiley.com/doi/10.1002/2015JF003678` doi: 10.1002/2015JF003678

Callahan, R. P., Ferrier, K. L., Dixon, J., Dosseto, A., Hahm, W. J., Jessup, B. S., ... Riebe, C. S. (2019). Arrested development: Erosional equilibrium in the southern sierra nevada, california, maintained by feedbacks between channel incision and hillslope sediment production. *GSA Bulletin*, *131*(7), 1179–1202. Retrieved from `https://doi.org/10.1130/B35006.1` doi: 10.1130/B35006.1

Carlston, C. W. (1963). *Drainage density and streamflow*. U.S. Govt. Print. Off.

Cleaves, E. T. (1989). Appalachian Piedmont landscapes from the Permian to the Holocene. *Geomorphology*, *2*(1), 159–179. doi: 10.1016/0169-555X(89)90010-X

Cleaves, E. T., Fisher, D. W., & Bricker, O. P. (1974). Chemical Weathering of Serpentinite in the Eastern Piedmont of Maryland. *GSA Bulletin*, *85*(3), 437–444. doi: 10.1130/0016-7606(1974)85<437:CWOSIT>2.0.CO;2

Cleaves, E. T., Godfrey, A. E., & Bricker, O. P. (1970). Geochemical Balance of a Small Watershed and Its Geomorphic Implications. *GSA Bulletin*, *81*(10), 3015–3032. doi: 10.1130/0016-7606(1970)81[3015:GBOASW]2.0.CO;2

Clubb, F. J., Mudd, S. M., Milodowski, D. T., Hurst, M. D., & Slater, L. J. (2014). Objective extraction of channel heads from high-resolution topographic data. *Water Resources Research*, *50*(5), 4283–4304. doi: 10.1002/2013WR015167

Collins, D. B. G., & Bras, R. L. (2010). Climatic and ecological controls of equilibrium drainage density, relief, and channel concavity in dry lands. *Water Resources Research*, *46*(4). doi: 10.1029/2009WR008615

Cosans, C. L. (2022). *Landscape Structure, Flow Path, and Transport Co-Evolution in the Deeply Weathered Piedmont* (Thesis). Johns Hopkins University.

Crowley, W. P., Reinhardt, J., & Cleaves, E. T. (1975). *Geologic map of the Cockeysville quadrangle* (QA-3 ed.). Baltimore, Maryland: Maryland Geological Survey.

Dunne, T. (1978). Field studies of hillslope flow processes. In M. J. Kirkby (Ed.), *Hillslope hydrology*. Chichester; New York: Wiley.

Eagleson, P. (1978). Introduction to Water Balance Dynamics. *Water Resources Research*, *14*(5), 705–712. doi: 10.1016/j.compchemeng.2012.11.011

Grieve, S. W., Mudd, S. M., & Hurst, M. D. (2016). How long is a hillslope? *Earth Surface Processes and Landforms*, *41*(8), 1039–1054. doi: 10.1002/esp.3884

Guice, G. L., Ackerson, M. R., Holder, R. M., George, F. R., Browning-Hanson, J. F., Burgess, J. L., ... Viete, D. R. (2021). Suprasubduction zone ophiolite fragments in the central Appalachian orogen: Evidence for mantle and Moho in the Baltimore Mafic Complex (Maryland, USA). *Geosphere*, *17*(2), 561–581.

doi: 10.1130/GES02289.1
Harel, M. A., Mudd, S. M., & Attal, M. (2016). Global analysis of the stream power law parameters based on worldwide 10Be denudation rates. *Geomorphology*, *268*, 184–196. doi: 10.1016/j.geomorph.2016.05.035
Harman, C. J., & Cosans, C. L. (2019). A low-dimensional model of bedrock weathering and lateral flow coevolution in hillslopes: 2. Controls on weathering and permeability profiles, drainage hydraulics, and solute export pathways. *Hydrological Processes*(December 2018), 1168–1190. doi: 10.1002/hyp.13385
Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. In *Int. Symp. Forest Hydrology.* doi: 10.1177/0309133309338118
Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *GSA Bulletin*, *56*(3), 275–370. doi: 10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2
Hurst, M. D., Mudd, S. M., Walcott, R., Attal, M., & Yoo, K. (2012). Using hilltop curvature to derive the spatial distribution of erosion rates. *Journal of Geophysical Research: Earth Surface*, *117*(F2), n/a-n/a. doi: 10.1029/2011jf002057
Jencso, K. G., & McGlynn, B. L. (2011). Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, *47*(11). doi: 10.1029/2011WR010666
Johnson, A. (1967). *Specific yield: Compilation of specific yields for various materials* (USGS Numbered Series No. 1662-D). Washington, D.C.: U.S. Government Printing Office. doi: 10.3133/wsp1662D
Li, H. Y., Sivapalan, M., Tian, F., & Harman, C. (2014). Functional approach to exploring climatic and landscape controls of runoff generation: 1. Behavioral constraints on runoff volume. *Water Resources Research*. doi: 10.1002/2014WR016307
Litwin, D. G., Barnhart, K. R., Tucker, G. E., & Harman, C. J. (2023). *DupuitLEM: groundwater landscape evolution with landlab [software].* Zenodo. doi: 10.5281/zenodo.7620978
Litwin, D. G., & Harman, C. J. (2024). *Data and model output for "Evidence of subsurface control on the coevolution of hillslope morphology and runoff generation" [Dataset].* Zenodo. doi: 10.5281/zenodo.10624746
Litwin, D. G., Tucker, G. E., Barnhart, K. R., & Harman, C. J. (2022). Groundwater affects the geomorphic and hydrologic properties of coevolved landscapes. *Journal of Geophysical Research: Earth Surface*, *127*(1), e2021JF006239. doi: 10.1029/2021JF006239
Litwin, D. G., Tucker, G. E., Barnhart, K. R., & Harman, C. J. (2023). *Catchment coevolution and the geomorphic origins of variable source area hydrology* [Preprint]. Preprints. doi: 10.22541/essoar.167751635.59156916/v1
Luijendijk, E. (2022). Transmissivity and groundwater flow exert a strong influence on drainage density. *Earth Surface Dynamics*, *10*(1), 1–22. doi: 10.5194/esurf-10-1-2022
Luo, W., Jasiewicz, J., Stepinski, T., Wang, J., Xu, C., & Cang, X. (2016). Spatial association between dissection density and environmental factors over the entire conterminous United States. *Geophysical Research Letters*, *43*(2), 692–700. doi: 10.1002/2015GL066941
Marçais, J., de Dreuzy, J. R., & Erhel, J. (2017). Dynamic coupling of subsurface and seepage flows solved within a regularized partition formulation. *Advances in Water Resources*, *109*, 94–105. doi: 10.1016/j.advwatres.2017.09.008
Marcon, V. (2019). *The effect of lithology on (bio) geochemical weathering: sandstone to serpentinite}* (Unpublished doctoral dissertation). THe Pennsylvania State University, State College, PA.
Mudd, S., Clubb, F., Grieve, S., Milodowski, D., Gailleton, B., Hurst, M., … Hut-

ton, E. (2022). *LSDtopotools/LSDTopoTools2: LSDTopoTools2 v0.7.* Zenodo. doi: 10.5281/ZENODO.3245040

Mudd, S. M., Attal, M., Milodowski, D. T., Grieve, S. W. D., & Valters, D. A. (2014). A statistical framework to quantify spatial variation in channel gradients using the integral method of channel profile analysis. *Journal of Geophysical Research: Earth Surface*, *119*(2), 138–152. doi: 10.1002/2013JF002981

Nippgen, F., McGlynn, B. L., & Emanuel, R. E. (2015). The spatial and temporal evolution of contributing areas. *Water Resources Research*. doi: 10.1002/2014WR016719

NOAA. (2024). *NOAA atlas 14 point precipitation frequency estimates: MD.* Retrieved 2024-01-11, from `https://hdsc.nws.noaa.gov/pfds/pfds_map_cont.html?bkmrk=md`

O'Loughlin, E. M. (1986). Prediction of Surface Saturation Zones in Natural Catchments by Topographic Analysis. *Water Resources Research*, *22*(5), 794–804.

Parsekian, A. D., Grana, D., Neves, F. d. A., Pleasants, M. S., Seyfried, M., Moravec, B. G., ... Kelleners, T. (2021). Hydrogeophysical comparison of hillslope critical zone architecture for different geologic substrates. *Geophysics*, *86*(5), WB29–WB49. Retrieved from `https://library.seg.org/doi/abs/10.1190/geo2020-0438.1` doi: 10.1190/geo2020-0438.1

Pavich, M. J. (1989). Regolith residence time and the concept of surface age of the Piedmont "Peneplain". *Geomorphology*, *2*(1), 181–196. doi: 10.1016/0169-555X(89)90011-1

Pedrazas, M. A., Hahm, W. J., Huang, M.-H., Dralle, D., Nelson, M. D., Breunig, R. E., ... Rempe, D. M. (2021). The relationship between topography, bedrock weathering, and water storage across a sequence of ridges and valleys. *Journal of Geophysical Research: Earth Surface*, *126*(4), e2020JF005848. doi: 10.1029/2020JF005848

Perron, J. T., Dietrich, W. E., & Kirchner, J. W. (2008). Controls on the spacing of first-order valleys. *Journal of Geophysical Research: Earth Surface*, *113*(4), 1–21. doi: 10.1029/2007JF000977

Perron, J. T., & Royden, L. (2013). An integral approach to bedrock river profile analysis. *Earth Surface Processes and Landforms*, *38*(6), 570–576. doi: 10.1002/esp.3302

Portenga, E. W., Bierman, P. R., Trodick, C. D., Jr., Greene, S. E., DeJong, B. D., Rood, D. H., & Pavich, M. J. (2019). Erosion rates and sediment flux within the Potomac River basin quantified over millennial timescales using beryllium isotopes. *GSA Bulletin*, *131*(7-8), 1295–1311. doi: 10.1130/B31840.1

Prancevic, J. P., & Kirchner, J. W. (2019). Topographic controls on the extension and retraction of flowing streams. *Geophysical Research Letters*, *0*(0). doi: 10.1029/2018GL081799

Putnam, S. M. (2018). *The influence of landscape structure on storage and streamflow generation in a Piedmont catchment* (Thesis). Johns Hopkins University.

Rempe, D. M., & Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy of Sciences*, *111*(18), 6576–6581. doi: 10.1073/pnas.1404763111

Riebe, C. S., Hahm, W. J., & Brantley, S. L. (2017). Controls on deep critical zone architecture: A historical review and four testable hypotheses. *Earth Surface Processes and Landforms*, *42*(1), 128–156. doi: 10.1002/esp.4052

Roering, J. J., Perron, J. T., & Kirchner, J. W. (2007). Functional relationships between denudation and hillslope form and relief. *Earth and Planetary Science Letters*, *264*(1), 245–258. doi: 10.1016/j.epsl.2007.09.035

Rossi, M. W., Whipple, K. X., & Vivoni, E. R. (2016). Precipitation and evapotranspiration controls on daily runoff variability in the contiguous United States and Puerto Rico. *Journal of Geophysical Research: Earth Surface*, *121*(1), 128–145. doi: 10.1002/2015JF003446

Sangireddy, H., Carothers, R. A., Stark, C. P., & Passalacqua, P. (2016). Controls of climate, topography, vegetation, and lithology on drainage density extracted from high resolution topography data. *Journal of Hydrology*, *537*, 271–282. doi: 10.1016/j.jhydrol.2016.02.051

Saulnier, G.-M., Obled, C., & Beven, K. (1997). Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the TOPMODEL framework. *Hydrological Processes*, *11*(9), 1331–1346. doi: 10.1002/(SICI)1099-1085(199707)11:9<1331::AID-HYP563>3.0.CO;2-9

Sofia, G. (2020). Combining geomorphometry, feature extraction techniques and earth-surface processes research: The way forward. *Geomorphology*, *355*, 107055. Retrieved 2020-11-17, from http://www.sciencedirect.com/science/article/pii/S0169555X20300258 doi: 10.1016/j.geomorph.2020.107055

Staff, S. S., & Natural Resources Conservation Service, United States Department of Agriculture. (2023). *Web Soil Survey.* https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx.

St. Clair, J., Moon, S., Holbrook, W. S., Perron, J. T., Riebe, C. S., Martel, S. J., … De Richter, D. B. (2015). Geophysical imaging reveals topographic stress control of bedrock weathering. *Science*, *350*(6260), 534–538. doi: 10.1126/science.aab2210

Theodoratos, N., Seybold, H., & Kirchner, J. W. (2018). Scaling and similarity of a stream-power incision and linear diffusion landscape evolution model. *Earth Surface Dynamics*, *6*(3), 779–808. doi: 10.5194/esurf-6-779-2018

Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T., & Valdes-Pineda, R. (2015). Catchment coevolution: A useful framework for improving predictions of hydrological change? *Water Resources Research*, *51*(7), 4903–4922. doi: 10.1002/2015WR017032

Uchida, T., Tromp-van Meerveld, I., & 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*(1), 117–133. doi: 10.1016/j.jhydrol.2005.01.012

Wu, S., Zhao, J., Wang, H., & Sivapalan, M. (2021). Regional Patterns and Physical Controls of Streamflow Generation Across the Conterminous United States. *Water Resources Research*, *57*(6), e2020WR028086. doi: 10.1029/2020WR028086

Yoshida, T., & Troch, P. A. (2016). Coevolution of volcanic catchments in Japan. *Hydrology and Earth System Sciences*, *20*(3), 1133–1150. doi: 10.5194/hess-20-1133-2016

Zhang, W., & Montgomery, D. R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, *30*(4), 1019–1028. doi: 10.1029/93WR03553