Earth Surf. Dynam. Discuss., 3, 1005–1059, 2015
www.earth-surf-dynam-discuss.net/3/1005/2015/
doi:10.5194/esurfd-3-1005-2015

Earth Surface Dynamics Discussions
Open Access

Designing a suite of measurements to understand the critical zone

S. L. Brantley[1], R. DiBiase[1], T. Russo[1], Y. Shi[2], H. Lin[2], K. J. Davis[3], M. Kaye[2], L. Hill[2], J. Kaye[2], A. L. Neal[4], D. Eissenstat[2], B. Hoagland[1], and A. L. Dere[1,a]

[1]Earth and Environmental Systems Institute, Department of Geosciences, Pennsylvania State University, Pennsylvania, PA, USA
[2]Department of Ecosystem Science and Management, Pennsylvania State University, Pennsylvania, PA, USA
[3]Earth and Environmental Systems Institute, Department of Meteorology, Pennsylvania State University, Pennsylvania, PA, USA
[4]Earth and Environmental Systems Institute, Pennsylvania State University, Pennsylvania, PA, USA
[a]now at: Department of Geography & Geology, University of Nebraska Omaha, Omaha, NE, USA

Received: 21 July 2015 – Accepted: 31 July 2015 – Published: 17 September 2015

Correspondence to: S. L. Brantley (sxb7@psu.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

Abstract

Many scientists have begun to refer to the earth surface environment from the upper canopy to the depths of bedrock as the critical zone (CZ). Identification of the CZ as a worthy object of study implicitly posits that the study of the whole earth surface will provide benefits that do not arise when studying the individual parts. To study the CZ, however, requires prioritizing among the measurements that can be made – and we do not generally agree on the priorities. Currently, the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO) is expanding from a small original study area (0.08 km^2, Shale Hills catchment), to a much larger watershed (164 km^2, Shavers Creek watershed) and is grappling with the necessity of prioritization. This effort is an expansion from a monolithologic first-order forested catchment to a watershed that encompasses several lithologies (shale, sandstone, limestone) and land use types (forest, agriculture). The goal of the project remains the same: to understand water, energy, gas, solute and sediment (WEGSS) fluxes that are occurring today in the context of the record of those fluxes over geologic time as recorded in soil profiles, the sedimentary record, and landscape morphology.

Given the small size of the original Shale Hills catchment, the original measurement design resulted in measurement of as many parameters as possible at high temporal and spatial density. In the larger Shavers Creek watershed, however, we must focus the measurements. We describe a strategy of data collection and modelling based on a geomorphological framework that builds on the hillslope as the basic unit. Interpolation and extrapolation beyond specific sites relies on geophysical surveying, remote sensing, geomorphic analysis, the study of natural integrators such as streams, ground waters or air, and application of a suite of CZ models. In essence, we are hypothesizing that pinpointed measurements of a few important variables at strategic locations will allow development of predictive models of CZ behavior. In turn, the measurements and models will reveal how the larger watershed will respond to perturbations both now and into the future.

ESURFD

3, 1005–1059, 2015

1 Introduction

The critical zone (CZ) is changing due to human impacts over large regions of the globe at rates that are geologically significant (Crutzen, 2002; Vitousek et al., 1997a, 1997b; Wilkinson and McElroy, 2007). To maintain a sustainable environment requires that we learn to project the future of the CZ. Models are therefore needed that accurately describe CZ processes and that can be used to project, or “earthcast,” the future. At present we generally cannot earthcast all the properties of the CZ but we can run models to project certain processes based on scenarios of human behavior (Godderis and Brantley, 2014). However, many of our models are inadequate to make successful projections. For example, we cannot a priori predict the streamflow in a catchment even if we know the average climate conditions, current soil textures, and current vegetation, because we are often uncertain how much water is lost to evapotranspiration and to groundwater (Beven, 2011). Likewise, we cannot a priori predict the depth or chemistry of regolith on a hillslope even if we know its lithology and tectonic and climatic history, because we do not fully understand what controls the rates of regolith formation and transport (Amundson, 2004; Brantley and Lebedeva, 2011; Dietrich et al., 2003; Minasny et al., 2008). Perhaps even more unexpectedly, we often do not even agree upon which minimum measurements are needed to answer these questions at any location.

Such difficulties are largely due to two factors: (i) we cannot adequately quantify spatial heterogeneities and temporal variations in the reservoirs and fluxes of water, energy, gas, solutes, and sediment (WEGSS); and (ii) we do not adequately understand the interactions and feedbacks among chemical, physical, and biological processes in the CZ that control these fluxes. This latter problem means that the CZ (Fig. 1) is characterized by tight coupling between chemical, physical, and biological processes which exert both positive and negative feedbacks on surface processes. Modelling the CZ is fraught with problems precisely because of these feedbacks and because the

ESURFD

3, 1005–1059, 2015

presence of thresholds means that extrapolation from sparse measurements can be challenging (Chadwick and Chorover, 2001; Ewing et al., 2006).

The result of these couplings and feedbacks is that properties emerge during evolution of the system – properties such as the distribution of permeability in regolith or the distribution of soil gas vs. depth. Many of these properties are considered to move toward average steady-state values. Perturbations that occur over timescales shorter than the characteristic time needed to reach steady state for a given gradient result in short-term changes, but in general the gradients are thought to move toward steady-state average values determined by the operative feedbacks. In Fig. 1, some examples of these emergent properties are shown as depth or spatial gradients that are identified in the brown boxes. Scientists from different disciplines generally focus on different emergent properties as shown in Fig. 1, and thus tend to think about processes operating at disparate timescales. However, CZ science is built upon the hypothesis that an investigation of the entire object – the CZ – across all timescales (Fig. 1) will yield insights that disciplinary-specific investigations cannot.

This is a challenging task, given that the driving mechanisms for landscape change also span disparate timescales, from tectonic forcing over millions of years to glacial–interglacial climate change, to the recent influence of humans on the landscape. Each setting or observatory for analysis of the CZ must grapple with processes at different timescales to understand the dynamics and evolution of the system. At the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO), we have been investigating this challenge by studying the CZ in a 0.08 km^2 watershed located in central Pennsylvania (the Shale Hills catchment, Fig. 2). We measure all of the properties that are indicated in the boxes in the center of the diagram. Then, to explore the evolution and dynamics of Shale Hills, we use a suite of simulation models as shown in Table 1 (Duffy et al., 2014).

The small Shale Hills catchment, established for research in the 1970s (Lynch, 1976) and expanded as a CZO in 2007, has been a successful location for CZ research. The CZO's small scale has allowed development of a diverse but dense monitoring network

ESURFD
3, 1005–1059, 2015

that spans disciplines from meteorology to groundwater chemistry to landscape evolution. Given the small size, we referred to our measurement paradigm as “measure everything, everywhere”. For example, we inventoried all of the ~2000 trees with diameter greater than 20 cm at breast height, drilled 28 wells, sampled soil porewaters at 13 locations at multiple depths approximately every other week during the non-snow covered seasons for more than a year, and measured soil moisture at 105 locations (Fig. 2).

The Shale Hills catchment is also situated on a single lithology (shale), and this simplified the complexity of the system by simplifying the boundary conditions for models. To aid in use of models for interpretation, we used an approach of monitoring at ridgetops (i.e. 1-D sites), along catenas (2-D transects), and the full catchment (full 3-D) and we have similarly targeted models for 1-D, 2-D and 3-D simulations. The work led us to measure two types of hillslopes (2-D sites) which dominate the catchment: planar hillslopes that experience downslope but nonconvergent flow of water and soil, and swales that experience downslope convergent flow of water and soil. Much of our effort focused on understanding soils and waters in these two hydrologic units (Fig. 2).

The goal of the SSHCZO project now is to upscale from Shale Hills to the entire 164 km^2 Shavers Creek watershed (Fig. 3). The expansion from 0.08 to 164 km^2 is an expansion from a zeroth-order catchment to a watershed with three HUC-12 watersheds (this terminology refers to hydrologic unit codes as defined by the US Geological Survey, Seaber et al., 1987). While the larger watershed still lies within the Valley and Ridge physiographic province, it contains additional lithologies (sandstone, calcareous shale, minor limestone), and impacts from multiple land uses (agriculture, forest management, minor development). By necessity, to understand the interaction of WEGSS fluxes in Shavers Creek, we must move beyond the paradigm of measuring “everything, everywhere” (Fig. 2) to an approach of measuring “only what is needed”.

This phrasing, although simplistic, should resonate with any field scientist: the choice of measurement design is at the heart of any field project and resources always require choices. But when we study the CZ as a whole, we are asking, how does one allocate resources to measure and model the dynamics and evolution of the entire CZ system? This paper describes our philosophy of measurement and our previous paper describes the modelling approach (Duffy et al., 2014). Obviously, due to the wide range of CZ processes across environmental gradients, the specifics of an ideal sampling design will vary from site to site. Nonetheless we describe the philosophy behind our approach as a way to hypothesize an answer to the question, how can we adequately and efficiently measure the entire CZ? We then present specific examples for the first part of our expansion from Shale Hills to a sandstone subcatchment within Shavers Creek.

2 Rationale for the measurement plan

The choices of measurements to be made during the expansion are driven by data needs for the models under development (Table 1), at the same time that the models are driven by observations in the field. The suite of models shown in Table 1 is one way to understand the entire CZ as an object of study, rather than as a set of disparate systems. In coordination with this modelling approach, a stratified sampling plan is being implemented in the CZO and paired with geophysical, catchment-scale stream and remote sensing measurements. The models will then be used to upscale from limited point or subregion measurements to the whole watershed and from limited temporal measurements over longer timescales.

Perhaps the largest difficulty in *spatially* characterizing the CZ in any observatory is the assessment of the extremely heterogeneous land surface, including regolith and pore fluids down to bedrock. In other words, while assessment of atmospheric and surface water pools can be technically challenging, mixing of these pools is much faster than mixing of the biotic pool, the regolith and rock reservoir, or the pool of soil porewater, making assessment of the spatial distribution of these latter reservoirs exceedingly difficult (Niu et al., 2014). On the other hand, the rates of changes in the land surface pools is generally much slower than the changing atmospheric reservoir, making atmospheric sensor requirements technically difficult. Thus, the largest difficulty in *tem-*

ESURFD
3, 1005–1059, 2015

porally characterizing today's fluxes in the CZ in any observatory may be measuring the fast-changing fluxes of the atmospheric pools and fluxes.

In recognition of these difficulties, the project started at Shale Hills precisely because it is a catchment almost 100 % underlain by Rose Hill formation shale with land use strictly as managed forest. Surface heterogeneities at Shale Hills were largely related to hillslope position, colluvium related to the Last Glacial Maximum (LGM), fracturing, and relatively limited spatial variations in vegetation. To understand the CZ at the Shavers creek watershed, on the other hand, we must grapple with a more complex set of variations related to differences in lithology, land use, climate change, and landscape adjustment to changes in base level due to tectonics, eustasy or stream capture (Fig. 1). Here, base level is used to refer to the reference level or elevation down to which the watershed is currently being graded.

In recognition of the new complexities within Shavers creek, the sampling strategy was designed using a stratified sampling plan based on geological and geomorphological knowlege rather than random sampling. An implicit hypothesis underlying this approach is the idea that sampling can be more limited when it is designed as a stratified approach based on geological (geomorphological knowledge). For example, where many many randomly chosen soil pits might be necessary if the delineation of swales vs. planar hillslopes was not recognized, if these two features are recognized and representative pits are dug to investigate these features, the number of pits can be minimized. Furthermore, one of the models under development for the CZO is a regolith formation model: by using this model to understand regolith formation, the number of pits in regolith can similarly be minimized.

Measurements at Shale Hills will soon be supplemented with targeted instrumentation in the two new subcatchments of Shavers creek watershed. The subcatchments were chosen to represent two of the new lithologies in the watershed, as bedrock geology is known to exert a first-order control on WEGSS fluxes in the CZ (e.g.Duvall et al., 2004; Williard et al., 2005). The first is a new forested subcatchment underlain only by sandstone. The second subcatchment for targetted measurements is currently being

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

identified on calcareous shale. This second subcatchment will also host several farms and will allow assessment of the effects of this land use on WEGSS fluxes.

The targetted subcatchment data will be amplified by measurements of chemistry and streamflow along the mainstem of Shavers Creek as well as catchment-wide meteorological measurements to upscale from Shale Hills to Shavers Creek (Fig. 3). The upscaling will rely on only a small number of sites for soil, vegetation, pore fluid, and soil gas measurements in each subcatchment. To extrapolate from and interpolate between these limited land surface measurements, models of landscape evolution (LE-PIHM), soil development (Regolith-RT-PIHM, WITCH), distribution of biota (BIOME4, CARAIB), C and N cycling (Flux-PIHM-BGC), sediment fluxes (PIHM-SED), solute fluxes (RT-Flux-PIHM, WITCH), soil gases (CARAIB), and energy and hydrologic fluxes (PIHM, Flux-PIHM) will be used. In effect, the plan is to substitute "everything everywhere" with measurements of "only what is needed" by using models of the CZ. As a simple example, a regolith formation model is under development that will predict distributions of soil thickness on a given lithology under a set of boundary conditions. Since much of the water flowing through these small catchments flows as interflow through the soil and upper fractured zone (Sullivan et al., 2015), use of the regolith formation model is necessary to predict the distribution of permeability in the catchment. The model will be groundtruthed with pinpointed field measurements. With this approach, water fluxes in the subcatchments and in Shavers creek watershed itself will eventually be estimated.

In each subcatchment, we have given names to arrays of instruments for clarity of description. The array of instruments in soil pits (1 m × 1 m× ∼ 2 m deep) and in trees near the pits along a catena is referred to as "ground hydrological observation gear" (Ground HOG). Vegetation is being assessed at transects located coincident with the Ground HOG. Geophysical surveys and geomorphic analysis using lidar are being conducted to interpolate between or extrapolate beyond the catenas.

In addition to Ground HOG, the energy, water, and carbon fluxes are being measured using "tower hydrologic observation gear" (Tower HOG). Ground and Tower HOGs are in turn accompanied by measurements of stream flow, chemistry and temperature,

ESURFD

3, 1005–1059, 2015

groundwater levels and chemistry. These streams and ground waters are natural spatial and temporal integrators over the watershed and therefore provide constraints on the 3-D-upscaled models. Stream and ground water data will parameterize and constrain model-data comparison and data assimilation, as described below.

3 Data assimilation

The choice of targeted measurements are derived at least in part from an observational system simulation experiment (OSSE) completed for the Shale Hills catchment using the Flux-PIHM model (Table 1) (Shi et al., 2014b). The OSSE evaluates how well a given observational array describes the state variables that are targeted by Flux-PIHM. Specifically, this OSSE (Shi et al., 2014b) emphasized water and energy fluxes for the catchment.

Prior to the OSSE, a sensitivity analysis was performed (Shi et al., 2014a) to determine the six most influential model parameters that were needed to constrain and produce a successful simulation. We defined "successful simulation" as one that reproduced the temporal variations of the four land surface-hydrologic fluxes (stream discharge, sensible heat flux, latent heat flux, and canopy transpiration), and the three state variables (soil moisture, water table depth, and surface brightness temperature) (Table 1) with high correlation coefficients and small root mean square errors. Once the six most influential model parameters were determined – porosity, van Genuchten alpha and beta, Zilitinkevich parameters, minimum stomatal resistance, and canopy water storage – the OSSE was then performed.

The OSSE evaluated which of the fluxes and state variables were most important in constraining those model parameters. Shi et al. (2014b) found that the calibration coefficients for the most important model parameters were most sensitive to observations of (i) stream discharge, (ii) soil moisture, and (iii) surface brightness temperature. (Alternately, instead of brightness temperature, measurements could focus on sensible and latent heat fluxes.) The OSSE has also been validated with assimilation of field observations at Shale Hills (Shi et al., 2015b).

On the basis of this OSSE, we are targeting measurement of stream discharge, soil moisture, and surface brightness temperature for each of the SSHCZO subcatchments on shale, sandstone, and calcareous shale. These measurements should allow us to reproduce subcatchment-averaged land–atmosphere fluxes and subsurface hydrology adequately. Once the three subcatchments are parameterized, the models will then be upscaled to the entire Shavers Creek watershed using information from lidar, SSURGO, geological maps, geophysical surveying, and land use.

Currently, the OSSE has only been used for assimilation of water and energy data but is being expanded to include biogeochemical variables. In other words, our ultimate aim is to complete an OSSE for C and N fluxes in each subcatchment. In the long run, we could also extend the OSSE to assimilate data for other solutes and for sediments.

Modeling results from Shale Hills indicated that an accurate simulation of the subcatchment spatial patterns in soil moisture were achieved using a relatively limited set of hydrologic measurements made at a few points (Shi et al., 2015a). Specifically, we had to measure (i) stream discharge at the outlet, (ii) soil moisture at a few locations, and (iii) groundwater levels at a few locations. The soil moisture (ii) and groundwater (ii) data used to calibrate the model were from 3 nearly co-located sites in the valley floor. These sites (referred to as RTHnet on Fig. 2) were the only sites with continuous data at the time of model calibration. Notably, COSMOS data were not yet available. The measurements were averaged across the three RTHnet sites (see data posted at http://criticalzone.org/shale-hills/data/dataset/3615/) to provide one calibration point in the model. Extending from this calibration point to the entire catchment was attempted using data from the SSURGO database (http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/). However, because of the coarseness of SSURGO, this was not successful for the very small Shale Hills catchment. Therefore, porosity, horizontal and vertical saturated hydraulic conductivity, and the van Genuchten parameters α and β were separately measured for each soil series and then were averaged

ESURFD
3, 1005–1059, 2015

for the whole soil column for each soil series (Supplement Table S2). These soil core measurements for each soil series were used to constrain the shape of the soil water retention curve for each soil series in the catchment in the model.

The result of this effort was that for the monolithologic 0.08 km^2 catchment of Shale Hills, five soil series were identified and soil properties measured (Lin et al., 2006). As we proceed with work on the new subcatchments, one of two approaches will be used. First, it is possible that relatively few soil moisture measurement locations are required in any given catchment, as long as we can obtain soil hydraulic properties for each soil series. Using the SSURGO soils database, such measurements could be made to parameterize the model. Alternately, spatially extensive soil moisture measurements based on COSMOS may be adequate to infer the variations in soil hydraulic properties on a series-by-series basis or based on geomorphological criteria. The overall plan is to use (i) SSURGO, (ii) geomorphological constraints, (iii) COSMOS, and (iv) soil moisture measurements along the catenas to parameterize Flux-PIHM.

4 Implementation in the Garner Run subcatchment

In this section we introduce the Garner Run subcatchment, one of the two new focus areas planned within the Shavers Creek watershed. In addition to describing the geologic and geomorphologic setting, we detail the sampling strategy. Preliminary observations and measurements from soil pits, vegetation surveys, and surface water monitoring are also presented.

4.1 Geologic and geomorphic context of Garner Run

A central underlying hypothesis of the CZO work is that the use of geomorphological analysis can inform the sampling strategy so that measurements can be limited in number. Therefore, we include a description of current knowledge of the geomorphological setting of the new subcatchment at Garner Run. The subcatchment drains a synclinal valley underlain by the Silurian Tuscarora Formation between the NW–SE trending ridges of Tussey Mountain and Leading Ridge (Figs. 3–5). The Tuscarora Formation, which locally consists of nearly pure quartzite with minor interbedded shales, is the ridge-forming unit that caps the highest topography in the Shavers Creek watershed. The hillslopes of both Tussey Mountain and Leading Ridge are nearly dip slopes, i.e, the hillslopes parallel bedding in the sandstone (Fig. 5). Indeed, subtle bedding planes can be observed in lidar-derived elevation data (Fig. 6b). The strong lithologic control on landscape form is manifested clearly in the high-resolution (1 m) bare-earth lidar topography.

One question that must be addressed is whether the Rose Hill Shale of the Clinton Group, which underlies Shale Hills and lies stratigraphically above the Tuscarora, may be present in the Garner Run subcatchment. Down-valley of the Garner Run study area, the Rose Hill Shale has been mapped in a low-sloping bench at the foot of Tussey Mountain (Figs. 3 and 4, Flueckinger, 1969). Although the entirety of the Garner Run study area is mapped as the Tuscarora Formation, the continuation of a low-sloping bench along the entire valley (Fig. 7) could be consistent with the presence of Rose Hill Shale throughout the catchment. In general, bedrock exposure is poor in the Shavers Creek watershed, but lidar topographic analysis, field mapping, and targeted geophysical surveys will aid in resolving uncertainties in subsurface composition necessary for modeling water, solute, and sediment fluxes.

The hillslope morphology of the Garner Run subcatchment shows a number of striking contrasts to that of the Shale Hills catchment. Most notably, the hillslopes of Tussey Mountain and Leading Ridge are nearly planar in map-view: they have not been dissected with the streams and swales common in the shale topography of much of Shavers Creek (including Shale Hills). Hillslopes underlain by the Tuscarora Formation are also nearly 10× longer (300–600 m) than those underlain by other geologic units within Shavers Creek (Fig. 6), including shales. In Shale Hills, for example, hillslopes are 50–100 m in length (Fig. 7). Furthermore, the hillslopes at Garner Run are less

ESURFD
3, 1005–1059, 2015

steep (mean slope = 12–17°) compared to those at Shale Hills (mean slope = 14–21°), despite having stronger underlying bedrock (quartzite vs. shale).

This observation of steeper hillslopes in Shale Hills vs. Garner Run is particularly curious given their presumably similar histories of climate and tectonism. If the two landscapes were in a topographic steady-state with erosion rate equal to rock uplift rate, we would expect Garner Run to have evolved with time to have steeper slopes in order to erode and transport its more resistant bedrock and coarser sediment. This hillslope conundrum could be related to the role of local base level and transient landscape adjustment (Whipple et al., 2013).

Specifically, analysis of stream longitudinal profiles on Garner Run and the mainstem of Shavers Creek reveals prominent knickpoints at elevations of ~ 320 m and 380 m, respectively (Fig. 7). Such breaks in channel slope geomorphically insulate the upper stream reaches from the mainstem of Shavers Creek and could be consistent with different rates of base level fall upstream and downstream of the knickpoint. Equivalently, the knickpoints could delineate different rates of local river incision into bedrock in the upper and lower reaches. Published cosmogenic nuclide-derived bedrock lowering rates ranging from 5–10 m Myr^{-1} from similar nearby watersheds (Miller et al., 2013; Portenga et al., 2013) may be a good estimate for rates in Garner Run upstream of the knickpoint (Fig. 7). These rates are indeed 3–4 times lower than bedrock lowering rates inferred for the Shale Hills catchment (20–40 m Myr^{-1}), which lies downstream of the knickpoint on Shavers Creek (Ma et al., 2013; West et al., 2014, 2013).

The origin and genesis of these knickpoints is likely due to some combination of the following: regional baselevel adjustment on the Susquehanna River since the Neogene (3.5–15 Ma) due to epierogenic uplift (Miller et al., 2013), stream capture and drainage reorganization (e.g. Willett et al., 2014), or temporal and spatial variations in bedrock exposure at the surface (e.g. Cook et al., 2009). Testing these competing controls will require additional direct measurements of bedrock lowering rates with cosmogenic nuclides at Garner Run, in addition to bedrock river incision models. Such models can account for both variations in rock strength and temporal changes in relative base level.

ESURFD
3, 1005–1059, 2015

In addition to variations in structure, lithology, and base level, Quaternary climate variations have left a strong imprint on the landscape of Shavers Creek. While the relict of the periglacial processes at Shale Hills are mostly observed in the subsurface colluvial stratigraphy (West et al., 2013), at Garner Run these processes have left behind boulder fields, solifluction lobes, and landslides observed at the land surface (Fig. 6). Such features are found throughout central Pennsylvania south of the LGM (last glacial maximum) limit (Gardner et al., 1991). These features document a major reorganization of the uppermost CZ by processes such as permafrost thaw. For example, the Leading Ridge hillslope (the southern hillslope defining the Garner Run subcatchment) is characterized by a hummocky topography at the 5–10 m scale, with partially vegetated boulder fields observed to be common. The other side of the catchment – Tussey Mountain hillslope – is steeper at the top, has greater relief, retains evidence of past translational slides, and contains open, unvegetated boulder fields. At the foot of the Tussey Mountain hillslope is a strong slope break that demarcates a low-sloping region characterized by abundant solifluction lobes (Figs. 6 and 7). Such features were either not as active, or their evidence has been erased or buried, at the Shale Hills subcatchment.

4.2 Water and energy flux measurements at Garner Run: Tower HOG

One of our major focuses is measuring precipitation and evapotranspiration (ET). These fluxes are drivers for landscape evolution as they are manifested today (Fig. 1). These measurements also are needed for the land surface water balance to constrain today's WEGSS fluxes. First, we are installing a laser disdrometer (LPM, Theis Clima GmbH) to measure precipitation amount and type in Garner Run. Another disdrometer has been in use at Shale Hills since 2008. The disdrometer will be deployed as part of the "tower hydrological observation gear" – referred to here as Tower HOG (Table 2). Tower HOG will be placed outside the watershed on Tussey Mountain ridgeline (Fig. 3). The remote, rocky terrain in Garner Run made constructing a new tower in the center of the watershed challenging. In contrast, a communications tower that is surrounded by representative forests already exists on the ridge top above the watershed, and we have therefore chosen this to host the eddy covariance flux instrumentation. Although the measurement footprint (i.e. fetch) for the tower measurements will include other areas, the tower instrumentation will be sensitive to fluxes from the forest in Garner Run. The tower measurements can also be compared to regional measurements such as the National Atmospheric Deposition Program measurements and samples of rainwater. For example, according to the nearest NADP site, Garner Run receives 1006 $mm\,yr^{-1}$ precipitation with an average pH of 5.0 (Thomas et al., 2013).

In addition to precipitation, sensible and latent heat fluxes (i.e. using eddy covariance), or skin temperature (upwelling terrestrial radiation) must also be measured to constrain Flux-PIHM (Shi et al., 2013). A small clearing below the tower site on the Tussey ridgeline makes the site unsuitable for skin temperature measurements representative of the forest, so we are only collecting eddy covariance measurements at the Tussey ridgeline. Of course, the complex terrain at both Shale Hills and Garner Run make eddy covariance measurements difficult to interpret in stable micrometeorological conditions. Since the primary energy partitioning happens during the day, however, daytime flux measurements are sufficient to constrain the modeling. For the Garner Run subcatchment, in addition, we also may be able to use upwelling infrared radiation measurements currently being made at the nearby Shale Hills. These radiative energy fluxes are measured using a four component radiometer, i.e., one that measures upwelling and downwelling terrestrial and solar radiation (Table 2). With both the EC measurements at Garner Run and radiative flux measurements at Shale Hills, Flux-PIHM should be well constrained.

4.3 Vegetation mapping

Vegetation has important impacts on the WEGSS fluxes and has important but poorly understood impacts on regolith formation and sediment transport. As we study individual subcatchments to understand WEGSS budgets, we seek to learn enough about the fluxes to extrapolate to the entire Shavers creek watershed: we therefore seek to

ESURFD
3, 1005–1059, 2015

understand some of the biogeochemical controls on WEGSS fluxes. For example, one of our goals is to assess nitrate fluxes out of Shavers creek. To do this, it may be necessary to determine tree species (Williard et al., 2005). Ultimately, we plan to run an OSSE to compare model predictions to measurements as a way to determine the important parameters for predicting carbon and nitrogen fluxes.

The vegetation has already been mapped in Garner Run subcatchment. The objective of the ground-based vegetation sampling design for the subcatchment was to measure spatial variability in vegetation across the catena (ridge top, midslope, and valley floor positions) defined for GroundHOG. These measurements set the stage for later re-measurements to understand temporal variability. For example, future assessments will quantify above-ground biomass, an important carbon pool. Variability in forest composition, standing biomass, and productivity across a watershed is generally related to gradients in biotic and abiotic resources such as soil chemistry or structure, water flux, and incoming solar energy. Therefore, the relatively restricted vegetation analysis design (Fig. 5) will be upscaled based on the team's developing knowledge of the distribution of soils across the watershed as well as lidar-based estimates of tree biomass and seasonal patterns of leaf area index and tree diameter growth. Given that we have not yet run an OSSE for carbon or nitrogen fluxes, our measurements of vegetation are relatively broad to enable future such analysis.

The vegetation measurements are important not only for C and N fluxes, but also for water flux. At Shale Hills, seasonal variation in tree transpiration has been estimated using tree sap flux sensors (Meinzer et al., 2013). While we sampled many different tree species in multiple locations at Shale Hills (Fig. 2), a more restricted number will be sampled at Garner Run. For example, sapflux sensors are being deployed at only the midslope positions of Ground HOG (Fig. 5). While eddy flux and soil moisture dynamics provide estimates of total transpiration and evaporation, sap flux provides direct estimates of tree transpiration that can constrain model predictions of transpiration. Collectively, these measures will help evaluate Flux-PIHM model processes. In

ESURFD
3, 1005–1059, 2015

addition, all approaches to measuring water fluxes are imperfect; errors can best be constrained when multiple approaches are used.

In addition to these sapflux measurements limited to midslope pits, vegetation has been sampled in linear transects parallel to the slope contour at each of the four soil pits (Fig. 5, Sect. 4.4), i.e., at each of the following pits: Leading Ridge ridge top (LRRT), Leading Ridge midslope (LRMS), Leading Ridge valley floor (LRVF), Tussey Mountain midslope (TMMS). Each vegetation transect was 10 m along the direction perpendicular to the valley axis and ~ 700–1400 m parallel to the valley axis.

Measurements along the transects yielded vegetation and forest floor cover data for 4.1 ha in the subcatchment (Table 3). They provide vegetation input data for land surface hydrologic models, and also evaluation data for a spatially-distributed biogeochemistry model (Flux-PIHM-BGC, Table 1). In the transected area 2241 trees > 10 cm diameter at breast height were measured, mapped, and permanently tagged. Understory vegetation composition was measured at 5 m intervals along transects and coarse woody debris was measured in 25 m planar transects parallel to the main transect, spaced every 100 m. Forest floor cover was classified as rock (typically boulder clasts from periglacial block fall), bare soil, or leaf litter every 1 m along each transect, and the dimensions (*a*, *b*, *c* axes) of the five largest exposed rocks was recorded every 25 m. Forest floor biomass was measured every 25 m along transects by removing the organic horizon from a 0.03 m^2 area for laboratory analysis: samples were dried, weighed and measured for carbon loss on ignition.

The results from these linear transects document variations in vegetation across catena positions (Table 3), as well as spatial variation in vegetation within a position. For example, mean tree basal area (BA; the ratio of the total cross-sectional area of tree stems ratioed to the total land surface area) in the LRRT transect is 25.3 $m^2\,ha^{-1}$; however, BA measurements ranged from 0 to 79 $m^2\,ha^{-1}$. Similarly, 16 % of points sampled every meter in LRRT fell on rock, yet at certain points along the transect rock cover was as high as 100 % or as low as 0 %. Vegetation measurements will be combined with data on surface rockiness (from transects) and a suite of ground and remotely sensed

ESURFD
3, 1005–1059, 2015

measurements from the watershed such as slope, curvature, aspect, solar radiation, and soil depth to model vegetation dynamics from environmental conditions and interpolate vegetation structure in areas of the watershed not directly sampled. Future resampling of linear transects will allow assessment of carbon uptake in vegetation, as well as changes in forest composition and structure.

Additional key vegetation parameters will be assessed at the soil pits described in Sect. 3.4. These additional measurements include root distributions, leaf area index (LAI, described in the next paragraph), litter fall, tree diameter growth and tree sap flux. Root distributions are being measured at all four soil pits in Garner Run using a combination of soil cores to accurately assess the high length densities near the surface. Root distributions, combined with soil water depletion patterns, can inform depth of tree water use over the season, which is an input parameter in the PIHM suite of models. Currently, a look-up table (http://www.ral.ucar.edu/research/land/technology/lsm/parameters/VEGPARM.TBL) is used to determine the rooting depth of each landcover type in the PIHM suite of models. Using field measured rooting depth as model input may improve the modeling of water uptake. In addition, profile wall mapping is being used to analyze the architecture, mycorrhizal colonization, and anatomy of deep roots. By characterizing and understanding the controls on root traits along a hillslope, we will eventually be able to use such observations to inform both models of water cycling (Flux-PIHM) and regolith formation (RT-Flux-PIHM, see Table 1).

At weekly intervals in the spring and fall and monthly during the summer, LAI will be assessed with a Li-2200 plant canopy analyzer (LI-COR Inc., Lincoln, Nebraska USA). The Moderate Resolution Imaging Spectroradiometer (MODIS) also provides remotely-sensed 8 day composite LAI (Knyazikhin et al., 1999; Myneni et al., 2002). The MODIS LAI product, however, has a spatial resolution of 1 km^2, which cannot resolve the spatial structure in LAI within small watersheds. The product also has a notable bias compared to field measurements (e.g. Shi et al., 2013). The LAI field measurements will be used for detailed information on leaf phenology, which is an important driver for the modeling of water and carbon fluxes for land surface and hydrologic mod-

ESURFD
3, 1005–1059, 2015

els (e.g., PIHM, Flux-PIHM (Table 1)), and provides calibration or evaluation data for biogeochemistry models like Flux-PIHM-BGC (Naithani et al., 2013; Shi et al., 2013).

Another important value we must estimate is net primary productivity (NPP). With NPP it is possible to constrain carbon and nutrient fluxes in vegetation stocks, which can be large components of the overall budgets. To estimate aboveground NPP, we will measure annual variation in trunk growth with dendrobands emplaced on examples of each of the six dominant tree species near each soil pit site. In addition, traps at each soil pit are also being used to assess litter fall. One of the key model outputs of Flux-PIHM-BGC is NPP, which can be evaluated using these measured data.

4.4 Soil pit measurements and Ground HOG instrumentation

4.4.1 Soil observations

To first order, the Garner Run subcatchment land surface falls into one of three categories: (i) fully soil mantled with few boulders emerging at the ground surface, (ii) boulder-covered with tree canopy, and (iii) boulder-covered without tree canopy. To assess the spatial heterogeneity of soils in the Garner Run subcatchment, we focused efforts on four soil pits: three on the north-facing planar slope of Leading Ridge (LRRT, LRMS, LRVF) and one mid-slope pit on the south-facing slope of Tussey Mountain (TMMS) (Fig. 5). This deployment of observations in soil pits along a catena, with an additional pit on the opposite valley wall, is here referred to as "Ground HOG" (ground hydrological observation gear) (Fig. 5, Supplement Fig. S1) and is the result of our focus on a minimalist sampling design.

In addition, the surface cover at Garner Run consists of coarse blocks of the Tuscarora sandstone ranging in diameter from ~ 10–200 cm, making it challenging to excavate large soil pits, limiting the number of such installations (Table 3). Three pits were dug entirely by hand (LRRT, LRMS, and TMMS). The Leading Ridge Valley Floor (LRVF) pit was dug by hand and was deepened using a jackhammer until the inferred contact with intact bedrock was reached. All four pits locations were selected on slopes that were planar in planview to avoid areas of convergent flow. The midslope pits were located on convex-up hillslopes for reasons discussed below. Given our catena design, we excavated pits in the following soil series: TMMS, LRRT and LRMS (Hazleton-Dekalb association, very steep), and LRVF (Andover extremely stony loam, 0–8 % slopes).

The rationale for the positions of the pits in Ground HOG are as follows. First, regolith formation at a ridge top is the simplest to understand and model (see, for example, Lebedeva et al., 2007, 2010) because net flux of water and earth materials is largely 1-D: i.e., net water flux is downward and net earth material flux is upward over geological time. Regolith-RT-PIHM is a model under development to simulate regolith development quantitatively for such 1-D systems, using constraints from cosmogenic isotope analysis (Table 1). Second, Regolith-RT-PIHM will also be able to model convex-upward hillslopes by assessment of the hillslope as a 2-D system that incorporates downslope transport of water and soil (e.g. Lebedeva and Brantley, 2013). By analyzing soil pits along a planar hillslope as we did for Shale Hills (Jin et al., 2010b), both 1-D and 2-D models of regolith formation will be enabled. With such conceptual and numerical models, we will extrapolate to other hillslopes within Shavers creek watershed. Third, at Shale Hills we discovered that both planar hillslopes and swales were important, requiring measurements at both (Graham and Lin, 2010; Jin et al., 2011; Thomas et al., 2013). No such swales have been observed at Garner Run, allowing focus on just one catena in the minimalist design. Finally, the importance of aspect on soil development and WEGSS fluxes at Shale Hills has been noted (Graham and Lin, 2011, 2010; Ma et al., 2011; West et al., 2014) on shale, as well as on sandstones in Pennsylvania (Carter and Ciolkosz, 1991). For that reason, one additional pit was sited on the northern side of the catchment to make observations to constrain the effect of aspect (Fig. 5).

At each pit location, we described the soil profile, which typically had the following structure: an upper rocky layer with a thin organic soil, a leached layer with large clasts mostly absent, a sandy mineral soil with a thin layer of accumulated organic and sesquioxide material, and a deeper clay-rich layer with larger rock fragments interspersed (Fig. S3, Table S2). Additionally, for each pit we sampled soils at 10 cm intervals for chemistry, grain size, organic matter, and composition analysis (Table S4).

Most of the Garner Run subcatchment has been mapped to lie on Tuscarora sandstone (Flueckinger, 1969). This sandstone, deposited in the Lower Silurian, has been interpreted as reworked beach sediments during original deposition (Cotter, 1982). The unit has been mildly metamorphosed so that pressure solution has cemented the fabric of the rock: as such, the unit is often referred to as a quartzite. Cotter reported the unit to be close to 98 % SiO_2. Weathering of sandstone is largely controlled by the porosity, the fraction of non-quartz grains, the composition of the cement (Turkington and Paradise, 2005), and the pH of soil porewaters (Certini et al., 2003). The porosity is important because it dictates how much water enters the weathering rock; in addition, during seasonal drying, salts deposited inside a sandstone can crystallize and disintegrate the rock (Labus and Bochen, 2012). Thermal cycling can also crack sandstones (Turkington and Paradise, 2005) as can tree roots (Amundson, 2004).

The average of the bulk compositions of four rock samples collected from the bottoms of the GroundHOG soil pits were used to estimate an average composition of the quartzite for comparison to similar analyses of bulk regolith samples (all measured using Li metaborate fusion followed by analysis by inductively coupled plasma atomic emission spectroscopy, Table S3). In Garner run samples, the Tuscarora was observed to be close to 98 % SiO_2. A small amount of titanium (Ti), generally present in sandstones in highly insoluble minerals, was observed to be present (Table S3). By calculating the normalized concentrations for elements assuming Ti is insoluble, we assessed the loss or gain of elements from the regolith as compared to Ti in the underlying Tuscarora sandstone. These normalized concentrations are referred to as mass transfer coefficients, τ_{ij}, where i is the immobile element and j is the mobile element (Anderson et al., 2002; Brimhall and Dietrich, 1987). From this assessment of regolith mass balance, it was observed that Al, Ca, Na, Si, and P were either largely unchanged ($\tau \approx 0$) or highly depleted ($\tau < 0$) compared to the underlying rock. In contrast, Mg, K, and

Fe were all significantly enriched in the soils ($\tau > 0$) compared to the protolith (Fig. S3). On each plot, the star represents the parent composition ($\tau \approx 0$), plotted at an arbitrary depth.

These observations are consistent with arguments in the literature that ridgetop soils are residual, poorly developed, and thin (Ciolkosz et al., 1990). In contrast, downslope soils generally developed not only from rock in place but also from colluvium (Fig. 5). Furthermore, soils in Pennsylvania commonly show a brown over red color layering that has been attributed to exposure of earlier regolith to weathering (producing the red layer) followed by emplacement of colluvium that experienced additional weathering (the brown layer) (Hoover and Ciolkosz, 1988). Such polygenetic histories will make regolith formation modelling more complex. The addition of Mg, K, and Fe to the soils, even at the ridgetop where downslope transport is unlikely to have been significant (Fig. S3), could either be explained by exogenous additions to the soil or by protolith compositional variation which was not assessed in the small set of 5 rock samples. For example, some interfingered shales are known to occur within the Tuscarora formation (Flueckinger, 1969) and could have provided the excess Mg, K, and Fe. Alternately, addition of these elements could have been caused by (i) dust inputs (Ciolkosz et al., 1990) which were likely to be important especially during the glacial period and just after, or (ii) fines percolating downward from weathering of the overlying Rose Hill shale before it was eroded away (Fig. 4). Movement of fines out of the Rose Hill shale is known to be happening today from our work at Shale Hills (Jin et al., 2010a).

4.4.2 Ground HOG

The Ground HOG instrumentation enables the in situ measurement of soil moisture and temperature, as well as gas and pore-fluid compositions, all at multiple depths (Fig. 5, Fig. S2). Ground HOG complements the atmospheric measurements taken by Tower HOG instrumentation (Sect. 3.2). Because the sites are difficult to access, measurements were automated to the extent possible. However, the lack of access to electricity and the cost of automated sensors (for CO_2 for example) meant that a completely

ESURFD
3, 1005–1059, 2015

automated monitoring system was unfeasible as well. Therefore, our final approach (Fig. S2) included a few automated components recording a continuous time series of data, coupled with additional components to be monitored manually, but with lower temporal resolution.

In selecting depths for soil sampling we wanted to instrument the site so that results could be compared across all watersheds, which meant we focused on a depth-based (as opposed to horizon-based) sampling scheme. In addition, we wanted to emphasize surface soils that have the highest water and biogeochemical flux rates. These layers also have the strongest influence on the atmospheric boundary layer. At the same time, we wanted to also document deep soil processes critical to understanding weathering and subsurface flowpaths. Thus, our final depth distribution included samples at 10, 20, and 40 cm from the top of the mineral soil (we used the top of the mineral soil as the depth reference because the O horizon depth varies greatly across the sites and among land-use types) and 20 cm above the bottom of the soil pit (coded "D-20"). At these four depths we installed from 1 to 4 component devices of the Ground HOG in each pit.

Automated soil moisture and temperature sensors (Hydra Probe, Stevens Water Monitoring Systems, Inc. Portland, OR) were emplaced to monitor at 10, 20 and 40 cm depths on the uphill face of each pit. In addition, TDR waveguides (Jackson et al., 2000) for manual point estimates of soil moisture were installed at the same depths plus D-20 on the uphill pit face, and the left and right pit faces (facing uphill). Wave guides are paired metal rods on a single cable that conduct a signal for time-domain reflectometry. The rods are 20 cm long and hand-made (Hoekstra and Delaney, 1974; Topp et al., 1980; Topp and Ferre, 2002). We placed 12 (4 depths × 3 pit faces) in each pit. The automated sensors were emplaced at depths expected to have the most dynamic soil moisture. In contrast, the waveguides measure deep soil moisture where temporal variability is expected to be low. The use of waveguides added spatial replication at all depths (Fig. 5, Fig. S2).

ESURFD
3, 1005–1059, 2015

Co-located with every soil moisture waveguide is a soil gas access tube to sample soil gas for measurements of the depth distribution of CO_2 and O_2 at a low temporal frequency. At 20 cm below the soil surface and 20 cm above the bottom of the uphill face of the pit, sensors are continuously measuring soil CO_2 (GP001 CO_2 probe, Forerunner Research, Canada) and O_2 (SO-110 Sensor, Apogee Intruments, Utah, USA) at the two midslope catena positions. We selected the midslope catenas for these sensors because they provide the best locations for contrasting north and south aspects. We placed one sensor at the D-20 location to document controls on acid and oxidative weathering near the bedrock interface. The second sensor is near the surface to monitor a zone of high biological CO_2 and O_2 processing. We did not install the sensors at the shallowest depth (10 cm) because we found that high diffusion and advection at shallower depths causes the gas concentrations at 10 cm to reflect atmospheric conditions, providing less information on soil biology (Jin et al., 2014) (Hasenmueller et al., 2015).

Lysimeters (Super Quartz, Prenart Equipment ApS, Denmark) have been emplaced to allow periodic manual sampling of soil pore water for chemical analysis at 20 cm and D-20 cm depths in all catena locations. The rationale for these depths is the same as described above for the automated CO_2 and O_2 sensors (they are co-located in the midslope pits). Overall, these Ground HOG measurements will parameterize the regolith formation models (Table 1) and will be used to test hypotheses linking hydrology, biotic production/consumption of soil gases, and weathering rates.

4.5 Upscaling from the pits to the catena using geophysics

To supplement the Ground HOG observations, we use geophysical and large-footprint methods to interpolate between and extrapolate beyond soil pits. For example, a cosmic-ray neutron detector (CR-1000B, Hydroinnova Inc.) has been emplaced to measure large-scale (~ 0.5 km radius) average soil moisture every 30 min. This COSMOS unit, already used in a variety of ecosystems (Zreda et al., 2013), will measure spatially averaged (3-D) soil moisture content within the watershed. Data processing methods have been developed that can account for various types of moisture storage (e.g. canopy storage, snow, water vapor, Franz et al., 2013; Zweck et al., 2013). The sensor has been installed near the LRVF (Leading Ridge valley floor) pit to provide spatially averaged moisture estimates across the valley.

The COSMOS fills in the gap between small-scale point measurements (Fig. 5) and large-scale satellite remote sensing. The footprint of COSMOS is optimal for hydrometeorological model calibration and validation at small watersheds. One sensor was installed at Shale Hills in 2011 and we are currently testing the COSMOS data with PIHM. We anticipate the results from both catchments will yield insights into the capabilities of cosmic-ray moisture sensing technology in steep terrain and will offer valuable insights into the problem of upscaling soil moisture measurements.

Ground HOG measurements will be further complemented by geophysical mapping along the catenas, including ground penetrating radar (GPR) transects of subsurface structure. Electromagnetic induction (EMI) mapping of soil electric conductivity will similarly be used to measure soil spatial variations between pits. We plan repeated GPR and EMI surveys, in combination with terrain analysis using lidar topography, to identify subsurface hydrological features and soil distribution using published procedures (Zhu et al., 2010a, b). We will also field check regolith depths using augers, drills, etc. With repeated geophysical surveys over time (e.g., different seasons and/or before and after storm events), we will explore temporal changes in heterogeneous soilscapes and subsurface hydrologic dynamics, as demonstrated in the previous studies at Shale Hills (Guo et al., 2014; Zhang et al., 2014).

Such geophysical mapping is necessary to link between soil-pit point measurements. Depth to bedrock along the catenas will also be mapped using the geophysical surveys and compared to pit measurements. These data can be used for upscaling biogeochemical patterns and processes. For example, we expect that soil depth and soil moisture exert the strongest controls on variation in soil gas concentrations, as observed in many places, including Shale Hills (Hasenmueller et al., 2015; Jin et al., 2014). Empirical relationships among these variables developed at Ground HOG points can be

ESURFD
3, 1005–1059, 2015

coupled with catchment scale soil moisture (from COSMOS) and soil depth (from GPR) data to upscale soil gas characteristics to the whole catchment.

An example of the utility of this approach is shown here from an investigation completed using a ground penetrating radar unit (TerraSIRch Subsurface Interface Radar System-3000). The unit was used to map the depth to bedrock in the Garner Run hillslope near the three major monitoring sites (LRVF, LRMS, LRRT) (Fig. 5). Multiple GPR traverses were completed by pulling the antennae along the ground surface. A distance-calibrated survey wheel with encoder was bolted onto these antennae to provide greater control of signal pulse transmission and data collection. The survey wheel occasionally slipped in the challenging terrain, resulting in some line lengths recorded by the survey wheel which were slightly different than the actual lengths. In order to surface normalize the radar records collected, relative elevation data were collected at major slope breaks along the traverse line with an engineering level and stadia rod.

A traverse line was established that ascends Leading Ridge in essentially a west to east direction from near Garner Run to the summit, running from about 494 to 588 m a.s.l. (Fig. 5). The dominant soils mapped along this traverse line (Table S2) include: Andover, Albrights, Hazleton, and Dekalb. The very deep, poorly drained Andover and moderately well to somewhat poorly drained Albrights soils have been reported in general to have formed in colluvium derived from acid sandstone and shale on upland toe-slope and foot-slope positions. The moderately deep, excessively drained Dekalb and the deep and very deep, well-drained Hazleton soils formed on higher-lying slope positions in residuum weathered from acid sandstone. These soils have moderate potential for penetration with GPR.

The traverse line was cleared of debris but the ground surface remained highly irregular with numerous rock fragments and exposed tree roots. These obstacles often halted the movement and caused poor coupling of the antennas with the ground. In this study, flags were inserted in the ground at noticeable breaks in the topography along the traverse line. User marks were inserted on the radar records as the antenna

ESURFD

3, 1005–1059, 2015

passed by these survey flags. Later, the elevations of these points were determined using an engineering level and stadia rod. The elevation data were entered into the radar data files and used to “surface normalize” or “terrain correct” the radar records.

In this preliminary investigation, the soil–bedrock interface was not easy to identify. This was attributed to poor antenna coupling with the ground surface in the challenging rocky terrain, noise in the radar records caused by rock fragments in the overlying soil, irregular and fractured bedrock surfaces, and varying degrees of hardness in both rock fragments and the underlying bedrock. These factors weakened the amplitude, consistency and continuity of reflections from the soil–bedrock interface. Nevertheless, we describe the preliminary results below.

Figure 9 shows two surface-normalized plots of the data that were collected with the 400 MHz antenna as it was pulled down Leading Ridge from the summit area to near the Garner Run (the stream). In these plots, the distance scale is measured from the summit area to near Garner Run. While differences in gross reflection patterns can be used to differentiate rock from soil, on these images the soil–bedrock interface is diffuse. However, we collected four repeated GPR transects using both 400 and 270 MHz antenna. Compared with the 400 MHz antenna, the lower resolution of the 270 MHz antenna has smoothed-out irregularities in the bedrock surface and reduced the noise from smaller, less extensive subsurface features, thus improving the interpretability of the soil–bedrock interface. Based on a total of 14 748 soil-depth measurements from $\sim$ 400 m long GPR images along this traverse line, the interpreted depth to bedrock averaged 1.37 m, with a range of 0.58 to 2.42 m. Table 4 summarizes the depth to bedrock values estimated from the two radar traverses shown in Fig. 9. Each entry in the table indicates the frequency of depth to bedrock data collected with the 400 MHz antenna along a traverse line that descended from the Leading Ridge. Data are grouped into four soil depth classes. The GPR-derived soil depths are reasonable compared to the values we estimated in the soil pits.

ESURFD
3, 1005–1059, 2015

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

4.6 Hydrology: groundwater measurements

Several methods are needed in a catchment to characterize physical and chemical interactions of water with regolith and rock. First, physical inputs and outputs to a catchment, including precipitation, interception, ET, soil infiltration, and groundwater discharge, must be understood. In fact, however, groundwater flows are often omitted from comprehensive hydrology–meteorology–vegetation models (e.g. the Variable Infiltration Capacity (VIC) hydrologic model, or the Noah Land-Surface Model (LSM)); however, at Shale Hills, we have estimated that 5 % of the nonevapotranspired water that enters the catchment reaches the regional groundwater table and flows to the stream as a deep flow component (Sullivan et al., 2015). At Garner Run, we also expect groundwater to play a significant role in streamflow and geochemical dynamics. For example, some researchers have found that drainage and runoff on sandstone catchments is controlled to great extent by bedrock (Hattanji and Onda, 2004), and specifically by flow through fractures in the upper meters of sandstone directly beneath the soil (Williams et al., 2010). In this section and the next section we focus on quantifying flows through and between surface water and groundwater. We aim to measure the relative magnitudes, timing, and spatial variability of these fluxes. We emphasize methodologies for measuring and characterizing groundwater and streamwater to characterize groundwater residence times, identify subsurface flow paths, and the drivers and controls on water-rock interactions.

Our plans for well installation and solid earth sampling by coring are reduced compared to sampling at Shale Hills. At Shale Hills, 28 wells were emplaced and then intermittently monitored (Fig. 2). In Garner run, two deep cores (> 50 m) will be extracted at two locations near the Garner Run catchment, one (∼ 100 m) on Tussey ridge, i.e. the ridge that divides Shavers Creek from the watersheds to the northwest and one (50–75 m) on the smaller divide within Shavers Creek between Garner Run and Roaring Run (see Fig. 3). Three shallow wells will be installed and cores (∼ 10 m) will be collected at the catena sites (Fig. 5). Two to four additional monitoring wells will be in-

ESURFD
3, 1005–1059, 2015

stalled along the stream reach on the valley floor. In drilling boreholes for assessment of groundwater, we also sample borehole solid-phase chemistry and mineralogy.

All core samples will be analyzed for bulk chemistry and mineralogy to characterize the weathering reactions and protolith in the critical zone. All boreholes will have groundwater monitoring wells installed, with screened intervals spanning the water table and with instrumentation as shown in Fig. 5. Monitoring at the wells will include hourly water level measurements using autonomous pressure loggers, hourly temperature measurements at two depths below the water table, and monthly water samples collected and analyzed for major ion chemistry. A pumping test will be conducted at the adjacent valley floor wells to measure aquifer storativity and hydraulic conductivity.

Deep core samples and groundwater monitoring will provide a baseline understanding of the geologic/pedologic and hydrologic system on the new sandstone lithology. Subsequent hypotheses about controls on weathering and hydrologic dynamics, as well as historical flow and solute fluxes, will be constrained by these observations at the catchment boundaries.

4.7 Hydrology: streamflow and chemistry measurements

The Garner Run study reach is approximately 500 m long within the catchment (Fig. 5) and consists of a rocky, often braided, channel. We have deployed a flume at the downstream end of the reach to measure discharge, and are monitoring stage continuously using a pressure transducer (Hobo U-20, Onset Computer Corp., Hyannis, MA). Surface water – groundwater (SW-GW) exchange characteristics have been measured using a short-term deployment of a distributed temperature sensor (DTS), and will be supplemented by a series of tracer injection tests to investigate hyporheic exchange characteristics over a wider range of stream discharges. Stream chemistry, including DO, pH, TDS, NO_3^-, SO_4^{2-}, Ca, K, Mg, Mn, Na, Fe, and Si, are being measured biweekly or monthly in the field with handheld electrodes along the 500 m reach, or by grab sampling and laboratory analysis (inductively coupled plasma atomic emission spectroscopy or ion chromatography).

Stream chemistry is also being monitored intermittently using higher temporal resolution by using a s::can spectrometer and an autosampler during storm events. The s::can is an in-situ measurement instrument for several water quality parameters (pH, TDS, DOC, NO_3^-,DO, NH_4^+, K, F (s::can, GmBH, Vienna, Austria). The chemistry and tracer test data will help quantify the flux of fluid and solutes through the subcatchment. The stream chemistry and discharge data will be combined with soil moisture, soil pore water chemistry, and groundwater data to estimate relative contributions to the stream, and underlying processes related to weathering in the near surface and aquifer.

Preliminary results from Garner Run indicate lower concentrations of Ca, Mg, and K, compared to Shale Hills. In addition, as expected, an initial constant injection tracer test at Garner Run revealed significant exchange with the subsurface during low-flow conditions ($\sim 0.004\,m^3\,s^{-1}$). Tracer test and temperature results documented that the stream sometimes loses and sometimes gains water in different sections over the 500 m experimental reach. One point stood out along the reach: both DTS and stream chemistry measurements are consistent with a significant input of groundwater at ~ 100 m downstream of the catena (Fig. 5). The DTS time series data will be analyzed to identify locations and magnitudes of groundwater inputs, as well as characteristic responses to rainfall events or changes in stream discharge. In combination with the tracer tests, DTS, and chemistry results, we will use well logs and lidar topography to explain the lithological and geomorphologic controls on the SW-GW system.

To characterize the major controls and processes governing WEGSS fluxes through the entire Shavers Creek catchment, we are making strategic measurements across the watershed to represent variability: stream discharge, stream chemistry, lithology, and geomorphology. Stream discharge and chemistry are being monitored along the main stem of Shavers Creek (SCAL, SCBL, and SCO) as shown in Fig. 3. At each location we are constructing a stage-discharge rating curve, and monitoring stage continuously using pressure transducers (Hobo U-20, Onset Computer Corp., Hyannis, MA). Streamwater-groundwater exchange characteristics will be measured as the channel crosses varying lithologies using a series of tracer injection tests. Stream chemistry

ESURFD
3, 1005–1059, 2015

will be measured monthly at each sampling site along Shavers creek. Analyses from the main stem of Shavers Creek provides a spatial integration of solute behavior from upstream lithologies and land use types. Eventually, with data from the three subcatchments on shale, sandstone, and calcareous shale, we will make estimates for nonmonitored catchments and test up-scaled estimates of the processes observed in each small watershed.

Preliminary stream chemistry and discharge results indicate significant variability among the three monitoring locations along Shavers Creek (Fig. 10). We see declining concentrations with increasing discharge for Mg and Ca (not shown), and somewhat chemostatic behavior for Si, K, nitrate and others. In this context, chemostatic is used to refer to concentrations of a stream that vary little with discharge (Godsey et al., 2009). Concentrations of Si decrease downstream (a dilution trend), while concentrations increase for Mg and nitrate, possibly due to agricultural amendments in the lower half of the watershed. The variety of behaviors will be investigated with respect to land use and lithology changes through the catchment.

5 Conclusions: measuring and modelling the CZ

Many environmental scientists worldwide are embracing the concept of the critical zone – the surface environment considered over all relevant timescales from the top of the vegetation canopy to the bottom of ground water. CZ science is built upon the hypothesis that an investigation of the entire object – the CZ – will yield insights that more disciplinary-specific investigations cannot. To understand the evolution and dynamics of the CZ, we are developing a suite of simulation models as shown in Table 1 (Duffy et al., 2014). These models are being parameterized based on measurements made at the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO) which is currently expanding from less than 1 to 165 km^2.

In this paper we described an approach for assessing the CZ in the larger watershed. In effect, our measurement design is a hypothesis in answer to this question: if we

ESURFD
3, 1005–1059, 2015

want to understand the dynamics and evolution of the entire CZ, what measurements are needed and where should they be made? Our approach emphasizes upscaling from 1-D to 2-D to 3-D using a catena paradigm for ground measurements that are extended with geological, geophysical, lidar, stream and meteorological measurements. Of course, our dataset has very low or no sampling replication within each catchment and we have only designed for one catchment per parent material. Obviously, there is a tension between monitoring a core dataset over time (a geological or hydrological approach) vs. the replication that is needed for spatial characterization (a soil science or ecological approach). Our spatial design was chosen based on the implicit assumption that implementation of Ground HOG and Tower HOG in each subcatchment could be upscaled to the entire watershed by interpolation, extrapolation, and modelling as described in Table 1. For example, we are testing the hypothesis that fewer soil pits are needed because we are using a regolith formation model and geological knowledge to site the few pits that we dig.

As an example of this approach, we point to our earlier observation of loss of Al, Na, Si and P from the soils at the same time that we identified significant enrichment in Mg, K, and Fe (Fig. 8). Simple mass balance arguments can be used to show that the enrichments in these latter elements are not likely due to residual accumulation during weathering of the parent orthoquartzite: prohibitively large thickness of quartzite would have had to weather away without loss of any Mg, K or Fe to enrich the soils adequately. On the other hand, accumulation of dust during weathering over a significant time period could explain the enrichment. Alternately, downward mobilization of fine particles from weathering of the overlying Rose Hill shale or interfingered shaley units might adequately explain the enrichment in these elements. Use of Regolith-RT-PIHM (Table 1) or WITCH (Godderis et al., 2006) to model regolith formation should allow testing of the feasibility of these or other ideas. With regolith formation models we can also extrapolate point measurements of soil thickness and porosity from catena observations to the broader Garner Run subcatchment and to other similar subcatchments

ESURFD
3, 1005–1059, 2015

in the Shavers creek watershed. In other words, the numerical models in Table 1 will be used to extend beyond the limited observations.

Of course, we can also augment the sampling design described here with brief measurement campaigns inside and outside the subcatchments or Shavers creek watershed as warranted. For example, while we will only monitor soil CO_2 continuously at a few catena positions and soil depths, we can augment these high frequency data with spatially extensive, but temporally limited measurements using manual soil gas samplers. Likewise, we may characterize vegetation and surface soil properties at 3–5 additional catchments of each parent material type using the transect design that we initiated at Garner Run (Fig. 5). In general, these outside measurements will be discipline-specific excursions to understand a specific variable. Another example is a set of measurements that are ongoing in a catchment to the north of Shavers creek to investigate regolith formation and hillslope form where the erosion rate is considerably faster. At this site, we anticipate learning how to parameterize or run models of regolith formation by exploring the impact of the rate of erosion (Table 1).

As we improve our understanding of the behavior of components of the critical zone, the point is to discover system-wide patterns and processes. Throughout, upscaling will remain a challenge. There is no comprehensive mathematical model of the critical zone, partly because it would be arduous to parameterize and perhaps more importantly because we do not yet understand all the interacting governing processes (Fig. 1). The research in Shavers Creek, and the work done at other critical zone observatories around the world, is an attempt to develop a system-wide process model (or ensemble of models) and to identify the essential measurements required for parameterization. The most robust models we have are conceptual models, and the most predictive are complex numerical simulations. However, both typically include only a portion of the critical zone. We seek a model that successfully explains the dynamics between topography, groundwater levels, and regolith thickness – at present we are working mostly with conceptual relationships drawn between pairs of factors (Fig. 1).

ESURFD
3, 1005–1059, 2015

The Supplement related to this article is available online at doi:10.5194/-15-1005-2015-supplement.

Author contributions. S. L. Brantley, H. Lin, K. J. Davis, A. L. Neal, J. Kaye, and D. Eissenstat designed the study and wrote the sections on soil analysis, GPR, eddy correlation, GIS analysis, biogeochemical analysis, and tree root analysis respectively. R. DiBiase spearheaded the geomorphological treatment, and T. Russo and B. Hoagland led the ground and stream water research. Y. Shi performed the OSSE and worked on coupling models with PIHM. M. Kaye designed and led the vegetation surveys; L. Hill and J. Kaye designed and implemented the soil pit sensor research; A. Dere completed the soil descriptions.

Acknowledgements. Field help was provided by B. Forsythe, D. Pederson, and R. Davis. D. Arthur and J. Williams are acknowledged for data organization; A. Orr for photographing the soil pits. C. Duffy is acknowledged for leadership at Shale Hills. S. Macdonald is acknowledged for Garner Run soil geochemistry analysis. L. Li is acknowledged for implementation of the COSMOS at Garner Run, and E. Kirby and P. Bierman for planning the cosmogenic measurements. This work was funded by NSF Critical Zone Observatory program grants to C. Duffy (EAR 07-25019) and SLB (EAR 12-39285, EAR 13-31726).

References

Amundson, R.: Soil formation, in: Treatise in Geochemistry: Surface and Ground Water, Weathering, and Soils, edited by: Drever, J. I., Elsevier Pergamon, Amsterdam, 1–35, 2004.

Anderson, S. P., Dietrich, W. E., and Brimhall, G. H.: Weathering profiles, mass balance analysis, and rates of solute loss: linkages between weathering and erosion in a small, steep catchment, Geol. Soc. Am. Bull., 114, 1143–1158, 2002.

Bao, C., Li, L., Shi, Y., Sullivan, P. L., Duffy, C. J., and Brantley, S. L.: RT-Flux-PIHM: a hydrogeochemical model at the watershed scale, Water Resour. Res., submitted, 2015.

Berg, T. M., Edmunds, W. E., Geyer, A. R., and others: Geologic map of Pennsylvania, 2nd edn., Pennsylvania Geological Survey, 4th ser., Map 1, 3 sheets, scale 1 : 250,000, 1980.

Beven, K. J.: Rainfall–Runoff Modelling: The Primer, Wiley-Blackwell, 2011.

ESURFD

3, 1005–1059, 2015

Brantley, S. L. and Lebedeva, M.: Learning to read the chemistry of regolith to understand the Critical Zone, Annu. Rev. Earth Pl. Sc., 39, 387–416, 2011.
Brimhall, G. and Dietrich, W. E.: Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenisis, Geochim. Cosmochim. Ac., 51, 567–587, 1987.
Carter, B. and Ciolkosz, E.: Slope gradient and aspect effects on soils developed form sandstone in Pennsylvania, Geoderma, 49, 199–213, 1991.
Certini, G., Hillier, S., McMurray, E., and Edwards, A.: Weathering of sandstone clasts in a forest soil in Tuscany (Italy), Geoderma, 116, 357–372, 2003.
Chadwick, O. A. and Chorover, J.: The chemistry of pedogenic thresholds, Geoderma, 100, 321–353, 2001.
Ciolkosz, E. J., Carter, B., Hoover, M. T., Cronce, R., Waltman, W., and Dobos, R.: Genesis of soils and landscapes in the Ridge and Valley province of central Pennsylvania, Geomorphology, 3, 245–261, 1990.
Cook, K. L., Whipple, K. X., Heimsath, A. M., and Hanks, T. C.: Rapid incision of the Colorado River in Glen Canyon – insights from channel profiles, local incision rates, and modeling of lithologic controls, Earth Surf. Proc. Land., 34, 994–1010, 2009.
Cotter, E.: Tuscarora formation of Pennsylvania : guidebook, Society of Economic Paleontologists and Mineralogists, Eastern Section, 1982 field trip, Lewisburg, Pa., Dept. of Geology and Geophysics, Bucknell University, 1982.
Crutzen, P. J.: Geology of mankind, Nature, 415, 23–23, 2002.
Dietrich, W. E., Bellugi, D. G., Sklar, L. S., Stock, J. D., Heimsath, A. M., and Roering, J. J.: Geomorphic Transport Laws for Predicting Landscape Form and Dynamics, Dynamics, In Prediction, in: Geomorphology, Geophyscial Monograph, edited by: Wilcock, P. R., Iverson, R. M., 135, American Geophysical Union, 103–132, doi:10.1029/135GM09, 2003.
Duffy, C., Shi, Y., Davis, K., Slingerland, R., Li, L., Sullivan, P. L., Godderis, Y., and Brantley, S. L.: Designing a suite of models to explore critical zone function, Procedia Earth and Planetary Science, 10, 7–15, doi:10.1016/j.proeps.2014.08.003, 2014.
Duvall, A., Kirby, E., and Burbank, D.: Tectonic and lithologic controls on bedrock channel profiles and processes in coastal California. J. Geophys. Res.-Earth, 109, F03002, doi:10.1029/2003JF000086, 2004.

ESURFD
3, 1005–1059, 2015

Ewing, S. A., Sutter, B., Owen, J., Nishiizumi, K., Sharp, W., Cliff, S. S., Perry, K., Dietrich, W., McKay, C. P., and Amundson, R.: A threshold in soil formation at Earth's arid-hyperarid transition, Geochim. Cosmochim. Ac., 70, 5293–5322, 2006.

Flueckinger, L. A.: Report 176, Geology of a portion of the Allensville Quadrangle, Centre and Huntingdon Counties, Pennsylvania, Bureau of Topographic and Geologic Survey, Commonwealth of Pennsylvania State Planning Board, 1969.

Franz, T. E., Zreda, M., Rosolem, R., and Ferre, T. P. A.: A universal calibration function for determination of soil moisture with cosmic-ray neutrons, Hydrol. Earth Syst. Sci., 17, 453–460, doi:10.5194/hess-17-453-2013, 2013.

Gardner, T. W., Ritter, J. B., Shuman, C. A., Bell, J. C., Sasowsky, K. C., and Pinter, N.: A periglacial stratified slope deposit in the valley and ridge province of central Pennsylvania, USA: Sedimentology, stratigraphy, and geomorphic evolution, Permafrost Periglac., 2, 141–162, 1991.

Godderis, Y. and Brantley, S. L.: Earthcasting the future Critical Zone, Elementa, 1, 1–10, doi:10.12952/journal.elementa.000019, 2014.

Godderis, Y., Francois, L., Probst, A., Schott, J., Moncoulon, D., Labat, D., and Viville, D.: Modelling weathering processes at the catchment scale: the WITCH numerical model, Geochim. Cosmochim. Ac., 70, 1128–1147, 2006.

Godsey, S. E., Kirchner, J. W., and Clow, D. W.: Concentration-discharge relationships reflect chemostatic characteristics of US catchments, Hydrol. Process., 23, 1844–1864, doi:10.1002/hyp.7315, 2009.

Graham, C. and Lin, H. S.: Controls and frequency of preferential flow occurrence: a 175-event analysis, Vadose Zone J., 10, 816–831, 2011.

Graham, C. G. and Lin, H. S.: Spatial–temporal patterns of preferential flow occurrence in the Shale Hills CZO based on real-time soil moisture monitoring, Vadose Zone J., special issue, 10, 816–831, doi:10.2136/vzj2010.0119, 2010.

Guo, L., Chen, J., and Lin, H.: Subsurface lateral flow network on a hillslope revealed by time-lapse ground penetrating radar, Water Resour. Res., 50, 21, 2014.

Hasenmueller, E., Jin, L., Stinchcomb, G., Lin, H., Brantley, S. L., and Kaye, J. P.: Topographic controls on the depth distribution of soil CO_2 in a small temperate watershed, Appl. Geochem., 63, 58–69, doi:10.1016/j.apgeochem.2015.07.005, 2015.

Hattanji, T. and Onda, Y.: Coupling of runoff processes and sediment transport in mountainous watersheds underlain by different sedimentary rocks, Hydrol. Process., 18, 623–636, doi:10.1002/hyp.1262, 2004.

Hoover, M. T. and Ciolkosz, E. J.: Colluvial soil parent material relationships in the Ridge and Valley physiographic province of Pennsylvania, Soil Sci., 145, 163–172, 1988.

Jin, L., Ravella, R., Ketchum, B., Bierman, P. R., Heaney, P., White, T., and Brantley, S. L.: Mineral weathering and elemental transport during hillslope evolution at the Susquehanna/Shale Hills Critical Zone Observatory, Geochim. Cosmochim. Ac., 74, 3669–3691, 2010a.

Jin, L., Ravella, R., Ketchum, B., Bierman, P. R., Heaney, P., White, T. S., and Brantley, S. L.: Mineral weathering and elemental transport during hillslope evolution at the Susquehanna/Shale Hills Critical Zone Observatory, Geochim. Cosmochim. Ac., 74, 3669–3691, 2010b.

Jin, L., Andrews, D. M., Holmes, G. H., Lin, H., and Brantley, S. L.: Opening the "black box": water chemistry reveals hydrological controls on weathering in the Susquehanna Shale Hills Critical Zone Observatory, Vadose Zone J., 10, 928–942, doi:10.2136/vzj2010.0133, 2011.

Jin, L., Ogrinc, N., Yesavage, T., Hasenmueller, E. A., Ma, L., Sullivan, P. L., Kaye, J., Duffy, C., and Brantley, S. L.: The CO_2 consumption potential during gray shale weathering: insights from the evolution of carbon isotopes in the Susquehanna Shale Hills critical zone observatory, Geochim. Cosmochim. Ac., 142, 260–280, doi:10.1016/j.gca.2014.07.006, 2014.

Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P., Christensen, T. R., Cramer, W., Matveyeva, N. V., McGuire, A. D., Murray, D. F., Razzhivin, V. Y., Smith, B., Walker, A. D., Anderson, P. M., Andreev, A. A., Brubaker, L. B., Edwards, M. E., and Lozhkin, A. V.: Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, amid future projections, J. Geophys. Res., 108, 8171, doi:10.1029/2002JD002559, 2003.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, P., Myneni, R. B., Nemani, R. R., and Running, S. W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm, Theoretical Basis Document, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA, 1999.

Labus, M. and Bochen, J.: Sandstone degradation: an experimental study of accelerated weathering, J. Soil Sci., 32, 453–459, 2012.

ESURFD

3, 1005–1059, 2015

Lebedeva, M. and Brantley, S. L.: Exploring geochemical controls on weathering and erosion of convex hillslopes: beyond the empirical regolith production function, Earth Surf. Proc. Land., 38, 1793–1807, doi:10.1002/esp.3424, 2013.
Lebedeva, M. I., Fletcher, R. C., Balashov, V. N., and Brantley, S. L.: A reactive diffusion model describing transformation of bedrock to saprolite, Chem. Geol., 244, 624–645, 2007.
Lebedeva, M. I., Fletcher, R. C., and Brantley, S. L.: A mathematical model for steady-state regolith production at constant erosion rate, Earth Surf. Proc. Land., 35, 508–524, 2010.
Lynch, J. A.: Effects of Antecedent Soil Moisture on Storm Hydrographs, Doctor of Philosophy, Forestry, The Pennsylvania State University, University Park, PA, 1976.
Ma, L., Jin, L., and Brantley, S. L.: How mineralogy and slope aspect affect REE release and fractionation during shale weathering in the Susquehanna/Shale Hills Critical Zone Observatory, Chem. Geol., 290, 31–49, 2011.
Ma, L., Chabaux, F., West, N., Kirby, E., Jin, L., and Brantley, S. L.: Regolith production and transport in the Susquehanna Shale Hills Critical Zone Observatory, Part 1: Insights from U-series isotopes, J. Geophys. Res.-Earth, 118, 722–740, doi:10.1002/jgrf.20037, 2013.
Meinzer, F. C., Woodruff, D. R., Eissenstat, D. M., Lin, H. S., Adams, T. S., and McCulloh, K. A.: Above- and belowground controls on water use by trees of different wood types in an eastern US deciduous forest, Tree Physiol., 33, 345–356, 2013.
Miller, S. R., Sak, P. B., Kirby, E., and Bierman, P. R.: Neogene rejuvenation of central Appalachian topography: evidence for differential rock uplift from stream profiles and erosion rates, Earth Planet. Sc. Lett., 369–370, 13, 2013.
Minasny, B., McBratney, A. B., and Salvador-Blanes, S.: Quantitative models for pedogenesis – a review, Geoderma, 144, 140–157, 2008.
Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., and Running, S. W.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83, 214–231, 2002.
Naithani, K. J., Baldwin, D., Gaines, K., Lin, H. S., and Eissenstat, D. M.: Spatial distribution of tree species governs the spatio-temporal interaction of leaf area index and soil moisture across a landscape, PloS ONE, 8, 12, doi:10.1371/journal.pone.0058704, 2013.
Niu, X., Williams, J. Z., Miller, D., Lehnert, K. A., Bills, B., and Brantley, S. L.: An ontology driven relational geochemical database for the Earth's Critical Zone: CZchemDB, J. Environ. Inf., 23, 10–23, doi:10.3808/jei.201400266, 2014.

Portenga, E. W., Bierman, P. R., Rizzo, D. M., and Rood, D. H.: Low rates of bedrock outcrop erosion in the central Appalachian Mountains inferred from in situ Be-10, Geol. Soc. Am. Bull., 125, 201–215, 2013.
Qu, Y. and Duffy, C. J.: A semi-discrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43, 1–18, 2007.
Seaber, P., Kapinos, F. P., and Knapp, G. L.: Hydrologic Unit Maps, US Geological Survey, Water-Supply Paper, 2294, , Denver, CO, 1987.
Shi, Y., Davis, K. J., Duffy, C. J., and Yu, X.: Development of a coupled land surface hydrologic model and evaluation at a critical zone observatory, J. Hydrometeorol., 14, 1401–1420, 2013.
Shi, Y., Baldwin, B. C., Davis, K. J., Yu, X., Duffy, C. J., and Lin, H.: Simulating high resolution soil moisture patterns in the Shale Hills catchment using a land surface hydrologic model, Hydrol. Process., doi:10.1002/hyp.10593, 2015a.
Shi, Y., Davis, K. J., Zhang, F., Duffy, C., J., and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: a multivariate real-data experiment, Adv. Water Resour., 83, 421–427, doi:10.1016/j.advwatres.2015.06.009, 2015b.
Sullivan, P. L., Hynek, S., Gu, X., Singha, K., White, T. S., West, N., Kim, H., Clarke, B., Kirby, E., Duffy, C., and Brantley, S. L.: Oxidative dissolution under the channel leads geomorphological evolution at the Shale Hills catchment, Am. J. Sci., submitted, 2015.
Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resour. Res., 33, 309–319, doi:10.1029/96WR03137, 1997.
Thomas, E., Lin, H., Duffy, C., Sullivan, P., Holmes, G. H., Brantley, S. L., and Jin, L.: Spatiotemporal patterns of water stable isotope compositions at the Shale Hills Critical Zone: linkages to subsurface hydrologic processes, Vadose Zone J., 12, doi:10.2136/vzj2013.01.0029, 2013.
Turkington, A. and Paradise, T.: Sandstone weathering: a century of research and innovation, Geomorphology, 67, 229–253, doi:10/1016/j.geomorph.2004.09.028, 2005.
Vitousek, P. M., Aber, J. D., Horwath, R. W., Lubchenco, J., and Melillo, J. M.: Human alteration of the global nitrogen cycle, Science, 277, 494–499, 1997a.
Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M.: Human domination of Earth's ecosystems, Science, 277, 494–499, 1997b.

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

Warnant, P., Francois, L., Strivay, D., and Gerard, J.-C.: CARAIB: a global model of terrestrial biological productivity, Global Biogeochem. Cy., 8, 255–270, 1994.
West, N., Kirby, E., Bierman, P. R., Slingerland, R., Ma, L., Rood, D., and Brantley, S. L.: Regolith production and transport at the Susquehanna Shale Hills Critical Zone Observatory: Part 2 – Insights from meteoric 10Be. J. Geophys. Res.-Earth, 118, 1–20, doi:10.1002/jgrf.20121, 2013.
West, N., Kirby, E., Bierman, P. R., and Clarke, B. A.: Aspect-dependent variations in regolith creep revealed by meteoric 10Be, Geology, 42, 507–510, doi:10.1130/G35357.1, 2014.
Whipple, K. X., Dibiase, R. A., and Crosby, B. T.: Bedrock rivers, in: Treatise in Geomorphology, Fluvial Geomorphology, edited by: Wohl, E., Academic Press, San Diego, CA, 550–573, 2013.
Wilkinson, B. H. and McElroy, B. J.: The impact of humans on continental erosion and sedimentation, Geol. Soc. Am. Bull., 119, 140–156, 2007.
Willett, S. D., McCoy, S. W., Perron, J. T., Goren, L., and Chen, C. Y.: Dynamic reorganization of river basins, Science, 343, 1–9, doi:10.1126/science.1248765, 2014.
Williams, J., Reynolds, R., Franzi, D., Romanowics, E., and Paillet, F.: Hydrogeology of the Potsdam sandstone in northern New York, Can. Water Resour. J., 35, 399–415, 2010.
Williard, K., Dewalle, D., and Edwards, P.: Influence of bedrock geology and tree species composition on stream nitrate concentrations in mid-Appalachian forested watersheds, Water Air Soil Poll., 160, 55–76, 2005.
Zhang, J., Lin, H. S., and Doolittle, J.: Soil layering and preferential flow impacts on seasonal changes of GPR signals in two contrasting soils, Geoderma, 213, 560–569, doi:10.1016/j.geoderma.2013.08.035, 2014.
Zhu, Q., Lin, H., and Doolittle, J.: Repeated electromagnetic induction surveys for determining subsurface hydrologic dynamics in an agricultural landscape, Soil Sci. Soc. Am. J., 74, 12, 2010a.
Zhu, Q., Lin, H., and Doolittle, J.: Repeated electromagnetic induction surveys for improved soil mapping in an agricultural landscape, Soil Sci. Soc. Am. J., 74, 1763–1774, doi:10.2136/sssaj2010.0056, 2010b.
Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R.: Corrigendum to "COSMOS: the COsmic-ray Soil Moisture Observing System" published in Hydrol. Earth Syst. Sci., 16, 4079–4099, 2012, Hydrol. Earth Syst. Sci., 17, 1065–1066, doi:10.5194/hess-17-1065-2013, 2013.

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

Zweck, C., Zreda, M., and Desilets, D.: Snow shielding factors for cosmogenic nuclide dating inferred from Monte Carlo neutron transport simulations, Earth Planet. Sc. Lett., 379, 64–71, 2013.

ESURFD

3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

Table 1. Designing a suite of CZ models.

	Modeling purpose	Model	Timescale of interest
Numerical models in use at SSHCZO	Topography (land scape evolution)	LE-PIHM	Days–millions of years
	Regolith composition and structure	Regolith-RT-PIHM, WITCH[a]	Hours–millions of years
	Distribution of biota	BIOME4[b], CARAIB[c], ED2	Day –centuries
	C and N pools and fluxes	Flux-PIHM-BGC	Days–decades
	Sediment fluxes	PIHM-SED	Hours – decades
	Solute chemistry and fluxes	RT-Flux-PIHM[d], WITCH	Hours–decades
	Soil CO_2 concentration and fluxes	CARAIB	Hours–decades
	Energy and hydrologic fluxes	PIHM[d], Flux-PIHM[f]	Hours–decades
Geological factor	Uplift rate, bedrock composition, bedrock physical properties, pre-existing geological factors such as glaciation		
External driver	Energy inputs, chemistry of wet and dry deposition, atmospheric composition, climate conditions, anthropogenic activities		

[a] Godderis et al. (2006)
[b] Kaplan et al. (2003)
[c] Warnant et al. (1994)
[d] Bao et al. (2015)
[e] Qu and Duffy, 2007)
[f] Shi et al. (2013)

ESURFD
3, 1005–1059, 2015
Designing a suite of measurements to understand the critical zone
S. L. Brantley et al.

Table 2. Measurements and instrumentation for Tower HOG system.

Measurement	Manufacturer	Model	Collection frequency
$[CO_2]$, $[H_2O]$	Li-cor	LI-7500A CO_2/H_2O analyzer	10 Hz[c]
3-D wind velocity, virtual temperature	Campbell Scientific	CSAT3 sonic anemometer	10 Hz[c]
Precipitation	Thies Clima	LPM disdrometer	Every 30 min
Precipitation type	Thies Clima	LPM disdrometer	Every 30 min
T_{air}	Vaisala	HMP60 humidity and temperature probe	Every 30 min
Relative Humidity	Vaisala	HMP60 humidity and temperature probe	Every 30 min
Longwave Radiation[a]	Kipp and Zonen	CGR3 pyrgeometer	Every 30 min
Shortwave Radiation[a]	Kipp and Zonen	CMP3 pyranometer	Every 30 min
Snow depth[b]	Campbell Scientific	SR50A sonic ranging sensor	Every 30 min
Digital Imagery	Campbell Scientific	CC5MPX digital camera	Every 24 h

[a] All four components of radiation (upwelling and downwelling (longwave and shortwave)) will only be measured at Shale Hills Tower HOG due to the location of the Garner Run Tower HOG. To model Garner Run we will use the Shale Hills data.
[b] originally designed as part of tower system but will be deployed at LRVF Ground HOG location because the Garner Run tower will be located outside of the catchment.
[c] The turbulent fluxes (sensible and latent heat) and the momentum flux are computed at 30 min intervals via eddy covariance using these data collected at 10 Hz.

ESURFD
3, 1005–1059, 2015

Table 3. Vegetation sampling in the Garner Run subcatchment.

Site[1]	Sample area (ha)	Tree basal area ($m^2 ha^{-1}$)	Tree density ($trees ha^{-1}$)	Tree species richness (# species)	Dominant tree species (% basal area)	Forest floor cover (% rock)	Mean rock diameter (cm)	Organic horizon C ($g m^{-2}$)
LRRT	1	25.3	607	9	*Quercus prinus* (44 %) *Acer rubrum* (19 %) *Pinus strobus* (19 %) *Nyssa sylvatica* (12 %)	16	29	1775
LRMS	1.4	25.1	610	12	*Betula lenta* (37 %) *Quercus prinus* (21 %) *Nyssa sylvatica* (15 %) *Quercus rubra* (10 %)	28	45	2208
LRVF	0.7	24.6	371	14	*Quercus rubra* (26 %) *Betula lenta* (23 %) *Quercus prinus* (20 %) *Acer rubrum* (14 %)	36	43	1122
TMMS	1	18.5	519	9	*Acer rubrum* (32 %) *Betula lenta* (29 %) *Nyssa sylvatica* (25 %)	34	60	n/a

[1] LRRT: Leading Ridge ridge top, LRMS: leading Ridge midslope, LRVF: leading Ridge valley floor, TMMS: tussey Mountain midslope. Measurements were made in linear belt transects 700 to 1400 m long and 10 m wide centered at each soil pit position.

ESURFD
3, 1005–1059, 2015

Table 4. Frequency distribution of depth to bedrock along the transect (Fig. 9).

Depth to bedrock	Upper section	Lower section
Shallow (< 0.5 m)	0.00	0.00
Moderately Deep (0.5 to 1 m)	0.26	0.04
Deep (1 to 1.5 m)	0.51	0.48
Very Deep (> 1.5 m)	0.24	0.48

ESURFD

3, 1005–1059, 2015

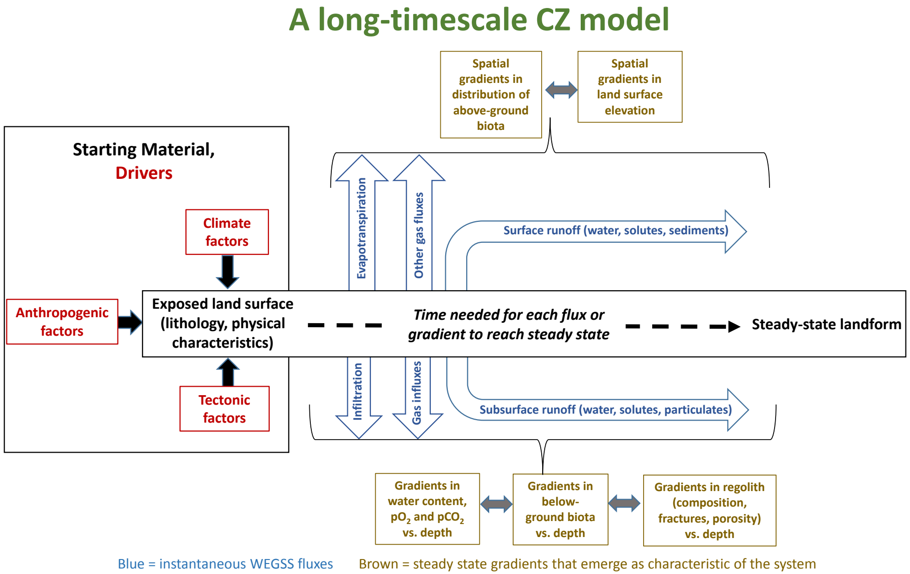

Figure 1. Critical zone science is aimed at understanding the architecture, character, and dynamics of the earth surface system at all different time scales. As rock of a certain lithology and set of structural characteristics is exposed at earth's surface due to uplift or erosion, climate-driven inputs transform rock to regolith. All of the properties in boxes to the right of the diagram can be considered properties which may sometimes reach a steady state after increasingly long exposure times. In other words, after an initial transient period, these characteristics can reach dynamic equilibrium. For example, regolith thickness can become constant when rate of erosion equals the rate of weathering advance. Likewise, the nature and distribution of biota may become constant for some period. As emphasized by the figure, ecosystems are established quickly compared to some geological changes, and can therefore often be studied as if some of the other characteristics in the diagram (e.g. regolith thickness and character, uplift rate, landscape curvature) are constant boundary conditions. However, over the longest timescales, all properties vary and can affect one another. Red boxes indicate drivers, black indicates the system under study, blue indicates the WEGSS fluxes, and brown boxes indicate gradients.

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

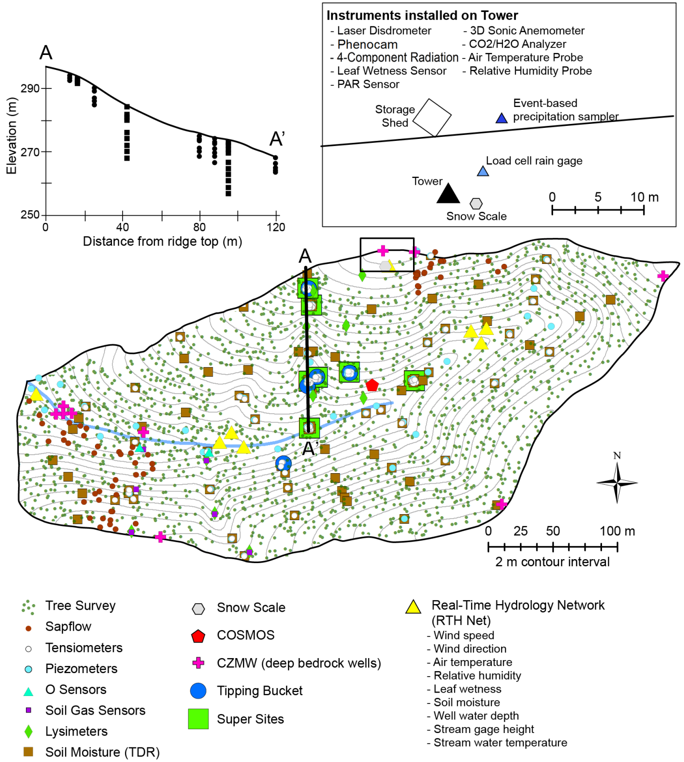

Figure 2. Mapped summary of the "everything, everywhere" sampling strategy at the Shale Hills subcatchment. Insets show soil moisture sensors (circles) and lysimeters (squares) along the transect shown on the map. Sensor and lysimeter depths are exaggerated five times compared to the land surface elevation. Second inset shows instrumentation deployed at the meteorological station on the northern ridge. Small black dots on the map are the trees that were surveyed and numbered. As we upscale the CZO to all of Shavers creek, many measurements will be eliminated as we emphasize only a Ground HOG and Tower HOG deployment as described for the Garner Run subcatchment.

ESURFD

3, 1005–1059, 2015

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

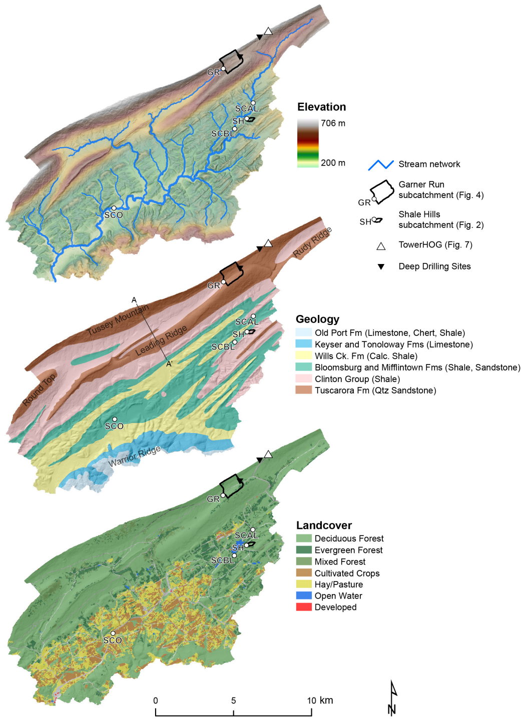

Figure 3. Map of Shaver Creek Watershed, highlighting **(a)** topography derived from airborne lidar, **(b)** geology (Berg et al., 1980), and **(c)** landuse (NLCD, 2011). In moving from measure-everything-everywhere (our paradigm in the 8 ha Shale Hills catchment (SH) to measure-only-what-is-needed in the Shavers Creek Watershed (164 km^2)), we chose to investigate two new first-order sub-catchments: a forested sandstone site (along Garner Run, marked GR) and an agricultural calcareous shale site (to be determined). In addition, three sites on Shavers Creek have been chosen as stream discharge and chemistry monitoring sites (marked SCAL – Shavers Creek Above Lake, SCBL – Shavers Creek Below Lake, and SCO – Shavers Creek Outlet).

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

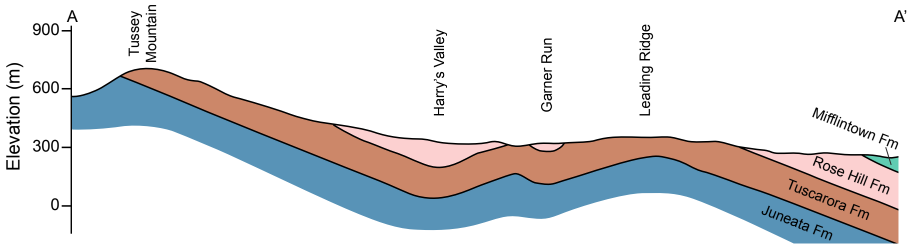

Figure 4. Geologic cross-section across Garner Run subcatchment reproduced from Flueckinger (1969). Map units are labelled from youngest to oldest: Smm (Rochester and McKenzie Members of the Mifflintown Formation), Smk (Keefer Member of the Mifflintown Formation), Srh (Rose Hill Formation), St (Tuscarora Formation), Oj (Juniata Formation). Mifflintown is Middle Silurian, Rose Hill and Tuscarora are Lower Silurian, and the Juniata is Upper Ordovician. Cross section position is downstream from the targeted subcatchment (see Fig. 3). The published map (Flueckinger, 1969) of the actual sub-catchment shows no remaining Rose Hill formation outcrop. Nonetheless, this cross-section from down valley of Garner Run subcatchment emphasizes that Rose Hill shale was originally present above the Tuscarora.

ESURFD
3, 1005–1059, 2015

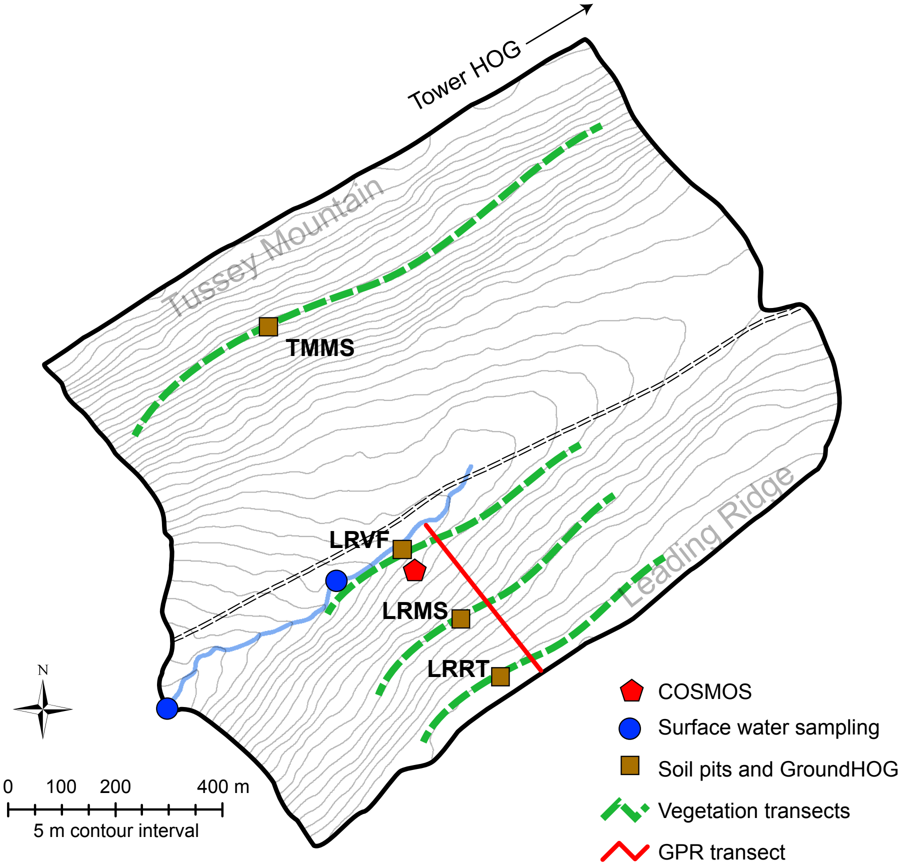

Figure 5. Map showing Garner Run subcatchment (blue line is the stream). Black dashed lines delineate Harry's Valley Road.

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

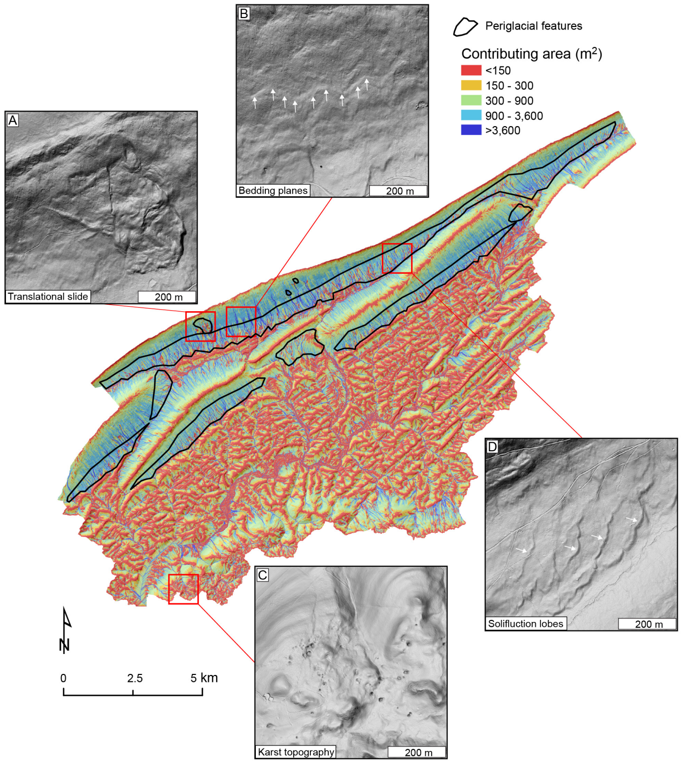

Figure 6. Map of bedrock and periglacial process controls on topography in Shavers Creek watershed. The contributing area was determined using the D-Infinity flow routing algorithm (Tarboton, 1997). The map highlights spatial variations in drainage density that correspond to sandstone (low drainage density and long hillslopes), shale (high drainage density and short hillslopes), and carbonate (intermediate drainage density and hillslope length) bedrock (see Fig. 3 for bedrock geology map). Black outlines correspond to periglacial features expressed in the 1 m lidar topography, such as landslides (inset A) and solifluction lobes (inset D). Sandstone bedding planes (inset B) and limestone karst topography (C) are also prominent.

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

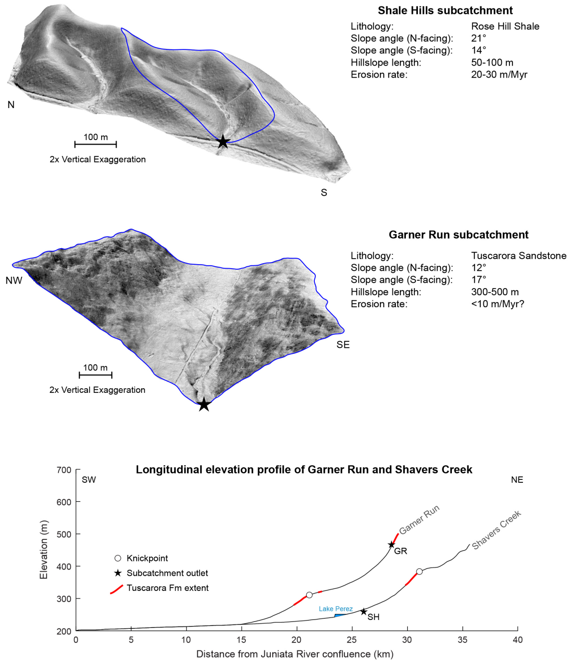

Figure 7. Perspective slopeshade maps (darker shades = steeper slopes) of Shale Hills (top panel) and Garner Run (middle panel) subcatchments, emphasizing differences in slope asymmetry and hillslope length. Soil production and erosion rates for Shale Hills subcatchment were measured based on U-series isotopes and meteoric 10Be concentrations in regolith respectively (Ma et al., 2013; West et al., 2013, 2014). Erosion rate for Garner Run subcatchment is estimated based on detrital 10Be concentrations from nearby sandstone catchments with similar relief (Miller et al., 2013). Bottom panel shows stream longitudinal profiles, highlighting the lithologic control on knickpoint locations. Note the location of the Shale Hills subcatchment (SH) downstream of the knickpoint on Shavers Creek and the location of the Garner Run subcatchment (GR) upstream of the knickpoint on Garner Run.

ESURFD
3, 1005–1059, 2015
Designing a suite of measurements to understand the critical zone
S. L. Brantley et al.

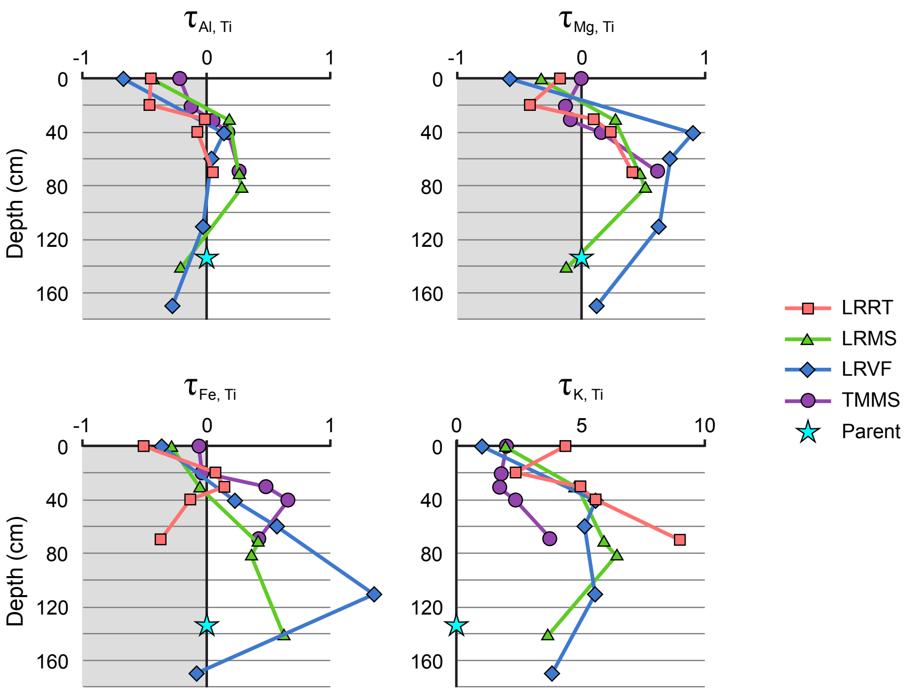

Figure 8. Four plots of normalized concentration (τ) vs. depth for soils analyzed from the four soil pits (LRVF, LRMS, LRRT, TMMS). *Y* axis indicates the depth below the organic – mineral horizon interface. The normalized concentration is the mass transfer coefficient determined using average parent composition from five rocks (Supplement Table S3) from the bottom of several of the pits and Ti as the immobile element. One explanation for these plots is that Al has largely been removed or moved downward in the profile while Mg, K, and Fe have largely been added to the profile. In these plots, $\tau = -1$ when an element is completely depleted compared to Ti in the parent material, $\tau = 0$ when no loss or gain has occurred, and is $\tau > 0$ when the element has been added to the profile.

ESURFD

3, 1005–1059, 2015

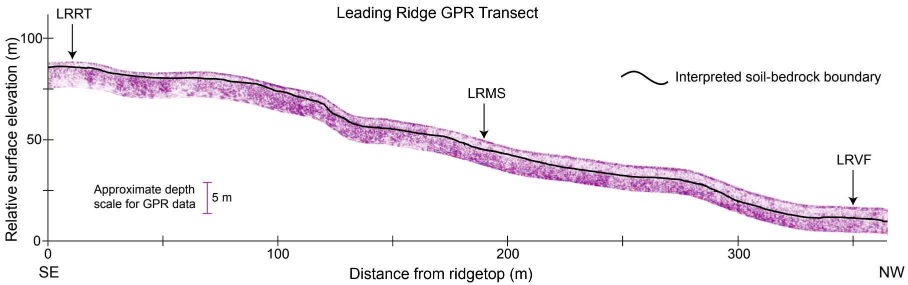

Figure 9. Ground Penetrating Radar (GPR) transect of Leading Ridge Catena, showing inferred location of bedrock–soil interface. GPR data is exaggerated by 4× compared to surface topography. Summary values are tabulated in Table 4 from these GPR measurements.

ESURFD

3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.

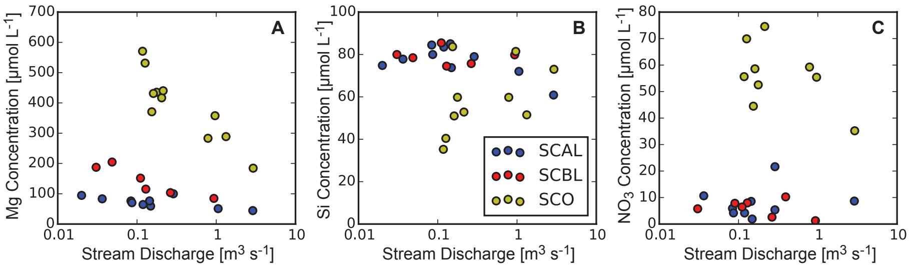

Figure 10. (a) Mg, **(b)** Si, and **(c)** Nitrate concentrations and stream discharge measured at three locations on Shavers Creek: Above Lake (SCAL, blue), Below Lake (SCBL, red), and the Outlet (SCO, yellow).

ESURFD
3, 1005–1059, 2015

Designing a suite of measurements to understand the critical zone

S. L. Brantley et al.