Testing the Fill-and-Spill Model of Subsurface Lateral Flow Using Ground-Penetrating Radar and Dye Tracing

Jonathan E. Nyquist,* Laura Toran, Lacey Pitman, Li Guo, and Henry Lin

Preferential flow (PF), which bypasses large portions of the soil or subsurface matrix, is critical in the transport of water and dissolved constituents in the unsaturated zone. To test the “fill-and-spill” model of hillslope hydrology that describes the generation and pattern of downslope lateral PF after storms, we used dye tracer and time-lapse, ground-penetrating radar (GPR) on a forested hillslope in the Susquehanna–Shale Hills Critical Zone Observatory. We injected 50 L of water mixed with Brilliant Blue dye (4 g L$^{-1}$) into a shallow trench cut perpendicular to the slope and used GPR to monitor the tracer downslope across a 1.0- by 2.0-m grid. The site was then excavated to the soil–saprock interface and photographed to document the dye pathways. We observed vertical dye fingering near the infiltration trench. Downslope lateral PF at the soil–saprock boundary was limited to $\sim 0.40$ m, which is evidence that the soil–saprock interface did not fill-and-spill. The extent, depth, and direction of the downslope PF indicated by GPR generally matched the dye staining patterns in the excavation, but the resolution of the 800-MHz GPR antenna was insufficient to distinguish small fingers of dye. A revised fill-and-spill model was proposed for this site that incorporates the PF through fractured saprock before water encounters fresh bedrock surface. This study demonstrates that GPR integrated with dye tracer infiltration can provide a useful means of testing hillslope hydrological hypotheses and unraveling the complexity of PF at the hillslope scale in a field setting.

Hillslope hydrology comprises complex processes that remain poorly understood despite decades of study. Unsaturated porous media flow models are commonly used to model flow and transport despite a general realization that Richards’ equation does not capture preferential flow (PF), which is typically dominant. Beven and Germann (2013) attributed the continuing reliance on porous media flow models to the increasing availability of unsaturated, porous media, modeling software, and insufficient data to parameterize PF models on the macropore scale. Burt and McDonnell (2015) suggested that the conceptual model of “fill-and-spill” (Graham et al., 2010; Tromp-van Meerveld and McDonnell, 2006) may provide a better working hypothesis for hillslope runoff phenomena. The proposed mechanism is rapid infiltration of precipitation through the soil layer down to the fresh bedrock surface or other less permeable horizon followed by lateral flow along the interface, with water filling some small depressions and then overflowing down along the hillslope. Fill-and-spill behavior may also result from excess saturation causing the water table to rise into shallower transmissive layers (McDonnell, 2013). After a certain amount of precipitation reaches the ground, the fill-and-spill model predicts rapid lateral flow downslope, following PF pathways controlled by the topography of the low-permeability horizon and microchannels above the bedrock (Graham et al., 2010). The fill-and-spill model has been applied to describe and predict the generation and pathways of downslope flow above impermeable layers (e.g., the bedrock surface), flow dynamics at less permeable layers (e.g., the allocation of vertical percolation and downslope flow), and the threshold response of discharge to total precipitation of a hillslope (Graham et al., 2010; Tromp-van Meerveld and McDonnell, 2006).
Given the potentially broad applicability of the fill-and-spill model (McDonnell, 2013), field verification is needed. However, determining whether fill-and-spill is an important mechanism for lateral flow in different natural settings is challenging because instrumenting a hillslope with a sufficient density of sensors to map subsurface flow networks is often impractical, and doing so would probably alter the system under study. Salve et al. (2012) instrumented a hillslope with wells for neutron probes, time domain reflectometry, and electrical resistivity. They found that bedrock wetting occurred before the saprolite was saturated and that the responses were highly heterogeneous. They also observed that the instrumentation was better at detecting wetting fronts than soil moisture values. In addition to detailed soil moisture instrumentation, predicting flow paths for fill-and-spill requires detailed knowledge of bedrock microtopography, which is buried beneath the subsurface.

Dye tracer tests remain the most direct technique for mapping PF networks, and numerous studies have demonstrated the variety and complexity of PF in natural systems (e.g., Anderson et al., 2009; Beven and Germann, 2013; Flury and Flühler, 1995; Flury et al., 1994; Wang and Zhang, 2011). Unfortunately, dye tracer tests have a limited ability to show the temporal evolution of the PF process, and they require excavation, which is time consuming. Anderson et al. (2009) used dye staining and excavation to trace PF pathways through a 30-m section of the Russell Creek research watershed in British Columbia, Canada, but it would be difficult to conduct a similar experiment throughout an entire hillside or watershed. Furthermore, large-scale excavation irreparably destroys the natural soil fabric of the system under study, which negates the possibility of repeated studies. Consequently, there is interest in nonintrusive methods for characterizing subsurface water movement.

Geophysical imaging provides continuous, nondestructive characterization of large subsurface areas. Ground-penetrating radar (GPR) has the highest spatial resolution of all current geophysical methods and is especially sensitive to the distribution and dynamics of soil moisture. Recently, Guo et al. (2014) imaged lateral PF networks on a hillslope in the Susquehanna–Shale Hills Critical Zone Observatory (SSHCZO) using time-lapse, common-offset GPR. In their experiment, water was introduced into a trench cut perpendicular to the hillslope, and GPR data were collected at regular time intervals before, during, and after infiltration along a series of transects downslope that were parallel to the trench. Their analysis of changes in radar data showed only localized changes in reflection amplitude. They did not see any large-scale changes in the radar velocities affecting the time–depth alignment of the radar traces; therefore, they were able to compare time-lapse radargrams using direct subtraction. They attributed the small radargram amplitude changes to PF confined to a series of macropore conduits above a dense layer, which they could trace downslope by connecting the changed areas on successive radargrams. Data from moisture sensors at the soil pit face that was 0.5 m downslope of the radar grid were consistent with their proposed explanation for the temporal evolution of the changes in the radar data, but, in the absence of direct ground truth data within the radar grid, their interpretation remains somewhat speculative. In fact, it remains challenging to map and monitor subsurface flow at the field scale (Allaire et al., 2009; Angermann et al., 2017), making the verification of the flow patterns derived from the GPR data difficult. Consequently, in most of the previous applications of GPR to investigate subsurface flow, ground truth of the flow pathways, such as by dye staining, is generally lacking (e.g., Doolittle et al., 2012; Gish et al., 2002; Truss et al., 2007; Zhang et al., 2014).

Therefore, to further demonstrate and validate the application of GPR technology in subsurface flow investigation, a direct comparison between the GPR-derived and actual flow patterns is required. Our objective was to partially replicate the experiment performed by Guo et al. (2014) with the addition of brilliant blue dye to the water release and subsequent excavation to compare changes detected using three-dimensional, time-lapse GPR with PF patterns mapped by dye tracing and excavation. The effectiveness of time-lapse GPR to detect and locate flow pathways was then evaluated to distinguish closely located flow pathways and to map the direction of the flow pathways. To our best knowledge, this is the first time that time-lapse GPR has been compared with dye staining to characterize subsurface flow dynamics in the field. Moreover, we hypothesized that the fill-and-spill model may be in effect for portions of the SSHCZO, mainly along planar hillslopes and near ridgetop areas where the depth to bedrock is shallow. Previous studies conducted in the same catchment have shown that for some storm events, water arrives more quickly at depth via lateral PF routes (Graham and Lin, 2011; Lin and Zhou, 2007). We further hypothesized that three-dimensional GPR could be used to map the shallow bedrock microtopography in sufficient detail to delineate fill-and-spill pathways.

Materials and Methods

Site Description

Our study site was located on the south-facing backslope side of the SSHCZO (Fig. 1), which is part of the forested Shaver Creek watershed located 24 km south-southwest of State College, PA. The SSHCZO is one of 10 critical zone observatories funded by the National Science Foundation. A humid, temperate forest, roughly 110 yr old and comprised of largely deciduous trees with some conifers, covers the watershed (Naithani et al., 2013). The region is underlain by the Silurian-age Rose Hill Formation, which consists of sequences of olive-colored shale, gray siltstones, and red hematite sandstone that were deposited in marine or brackish-water environments (Folk, 1960). At our site, the Rose Hill is present as a shale composed primarily of illite (58% w/w),
quartz (30% w/w), and vermiculated chlorite (11% w/w), plus trace amounts of feldspar, anatase, Fe oxides, and zircon (Brantley et al., 2016). The bedrock is highly fractured, with folded and faulted layers oriented northeast to southwest (N54° W), dipping 25 to 76° to the northwest (Jin et al., 2010).

Jin et al. (2011) made the distinction between bedrock, saprock, and regolith, using bedrock to refer to chemically unaltered parent rock. Saprock is fractured and chemically altered but is competent to the extent that it must be drilled to obtain samples. Regolith has been disaggregated and highly altered and can be sampled using a hand auger. In our study, we excavated using hand tools; thus, we removed soil and regolith to expose the saprock surface but not true bedrock.

Previous work at the study site has revealed lateral flow of ≥1 m based on geophysical monitoring during injection tests. Two injection experiments were conducted at a site 10 m upslope from the present study area. One injection was monitored with soil moisture sensors and GPR (Guo et al., 2014), which suggested lateral flow as far as 1.3 to 1.8 m downslope. The second experiment was monitored with electrical resistivity tomography and showed an increase in soil moisture 1.3 m downslope (Lichtner et al., 2012). These experiments were in part motivation for the present, more detailed survey, where we have added dye injection followed by excavation. A site ~10 m downslope from the earlier studies was selected for excavation.

**Dye Injection**

The dye tracer was released into a trench dug perpendicular to the hillslope and 0.2 m upslope of the first radar line. The trench was 1.15 m long, 0.20 m wide, and 0.18 m deep. Water and dye were added to the trench by pumping into a 1.15-m-long polyvinyl chloride pipe with a horizontal slot running its length to ensure even dispersal along the trench. The pipe was maintained at a constant ponding depth of 10 cm in the infiltration trench during the infiltration experiments (Fig. 2). The bottom of the infiltration trench was located within the Bw horizon. Two injections were conducted on 18 July 2013, which was in the dry period of the Shale Hills catchment, with a soil moisture content of ~15% (v/v) on the planar hillslope (Naithani et al., 2013). No rainfall events were recorded during the week before the experiment. A total of 53 L of water was first injected into wet the soil, and a second injection was performed using 53 L of water mixed with Brilliant Blue dye at a concentration of 4 g L⁻¹ (Petersen et al., 2001; Wang and Zhang, 2011). Addition of the dye increased the conductivity of the solution, which measured 3 mS cm⁻¹. The injection volume was based on the previous experiment at the site, which resulted in downslope travel of ~2 m along preferential flow paths. The first infiltration took 10 min, and the second infiltration took 27 min.

**Monitoring Using Time-Lapse Ground-Penetrating Radar**

Numerous studies have investigated GPR as a tool to map moisture (see Huisman et al. [2003] and Grote et al. [2010] for reviews). For typical GPR systems with a fixed transmitter–receiver offset, the goal is to map reflections. Ground-penetrating radar energy is reflected at boundaries between layers with contrasting dielectric properties:

\[
R = \frac{\sqrt{\varepsilon_2} - \sqrt{\varepsilon_1}}{\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}} \tag{1}
\]

where \( R \) is the reflection coefficient, and \( \varepsilon_1 \) and \( \varepsilon_2 \) are the dielectric constants in Media 1 and 2, respectively. Topp et al. (1980) developed an empirical relationship between soil water content and dielectric constant:

\[
\theta = 4.3 \times 10^{-6} \varepsilon^{-1} - 5.5 \times 10^{-4} \varepsilon^{-2} + 2.92 \times 10^{-2} \varepsilon - 5.3 \times 10^{-2} \tag{2}
\]

where \( \theta \) is the volumetric soil water content (m³ m⁻³). Because most geologic materials have relative dielectric constants in the

---

**Fig. 1.** The dye tracer test site was located on the south-facing backslope of the Susquehanna–Shale Hills Critical Zone Observatory. The size of the symbol marking the location of the ground-penetrating radar survey grid (actual size: 1 by 2 m) is exaggerated for visibility. Left panel shows the relative locations of the dye tracer test site and the soil borrow pit used for the surface water release experiment.
range of 3 to 30, compared with a dielectric constant of 81 for water (Reynolds, 1997), even small changes in saturation with time will alter radar reflection patterns significantly. Using GPR to detect PF pathways by mapping soil moisture changes requires comparing GPR images from before and after events. However, detection of moisture changes by direct-subtraction time-lapse radargrams is an inherently unstable process (Versteeg and Birken, 2000) because changes in soil moisture alter the propagation velocity of the radar waves in addition to the amplitude of the radar reflections. Velocity changes shift locations of the reflections on the radargram, and these shifts accumulate with depth, causing artifacts on the differenced radargram below the area of actual moisture change because of misalignment of the before and after radargrams (Truss et al., 2007). Because our experiment focused on the top 0.5 m of the subsurface, travel times were short and this misalignment was minor.

Our GPR survey grid covered an area 2 m wide perpendicular to the hillslope by 1 m parallel to the hillslope. For each radar survey, we collected data over a three-dimensional grid comprised of 21 lines (each 2 m long) spaced 0.05 m apart. For time-lapse GPR, the data collection should be as reproducible as possible to avoid introducing artifacts on the differenced radargrams between successive GPR scans. Following the example of Haarder et al. (2011), we constructed a wooden frame with an adjustable bar and pegs along the sides to guide the antenna along each of the lines to ensure that the lines were parallel, evenly spaced, and reproducible (Fig. 2). A tow rope was used to draw the antenna along the line smoothly and to reduce the amount of foot traffic on the grid. A survey wheel attached to the antenna recorded the distance along the lines. To assess reproducibility, two complete radar surveys of the grid were completed before the injection (background), after the initial release of water (post-water), and again after the release of the water and dye solution (post-dye).

We used a Mala radar system with a shielded antenna operating at a center frequency of 800 MHz and with a 0.14-m separation between the transmitter and receiver antennae. We collected 400 samples per trace at a time step of 0.1164 ns for a total record length of 46.434 ns. The traces were triggered every centimeter along the 2-m lines. Data collection for all 21 lines took ~15 min.

Fig. 2. Water and Brilliant Blue dye were pumped into a shallow trench upslope of the ground-penetrating radar (GPR) survey grid. A slotted, 1-m-long polyvinyl chloride pipe (inset) was used to ensure that the tracer solution was dispersed uniformly and to maintain a constant head during the injection. An adjustable wooden frame acted as a GPR antenna guide to ensure reproducible data collection.
The radar velocity was calibrated after the tracer experiments by hammering a steel rod parallel to the surface into the excavated wall along the soil–saprock interface at a known depth, running a radar line over the rod, and picking the arrival time in the resulting radargram. This velocity was used in subsequent time-to-depth conversion, but the assumption of a uniform, constant velocity must be considered approximate because of the heterogeneity of the soil and changes in moisture content during the experiments.

**Ground-Penetrating Radar Processing**

Post-processing of the radar data comprised the following steps: (i) removal of the direct current component by detrending and dewow filtering to remove low-frequency noise; (ii) first-arrival time zero adjustment; (iii) low-pass filtering to remove frequencies >1600 MHz; (iv) background removal by subtracting the average of data below 15 ns, which corresponded to signal degraded by attenuation in the shale saprock beneath the regolith; (vi) inverse amplitude gain; and (vii) time-to-depth conversion using a constant velocity of 0.077 m ns⁻¹. Because of velocity heterogeneity and perturbations created by the water release, we chose not to migrate the data. Guo et al. (2014) refined the algorithm of Boschetti et al. (1996) to align the first breaks of the traces. That proved unnecessary for our data because the slow, gentle movement of the antenna during data collection and the use of a frame to guide the radar antenna yielded data with no noticeable trace-to-trace time shifts, and repeat traces along the line matched within 1 cm. The gain function we used was inverse amplitude decay, which computes the best fitting exponential of the form to the mean trace (Tzanis, 2010). We calculated the decay curve using data for the first line of the background survey (closest to the trench) and then applied the same gain to all lines for consistency when calculating temporal changes. All the standard processing steps were conducted using matGPR (Tzanis, 2010), a freeware package written as a toolbox for MATLAB (The MathWorks, 2015), which was also used to difference the radargrams to highlight temporal changes. The choice of differencing schemes and alternatives is discussed below because it is dataset dependent.

The processed and differenced two-dimensional time-lapse radargrams were then loaded into GPR-Slice v7.0 (Goodman, 2015), a processing package designed for three-dimensional visualization of GPR data. The data were first aggregated into 20 horizontal time slices spanning 4.3 ns, with 50% overlap between slices, to cover the full 15-ns region of interest. The data within each time slice were interpolated to fill the gaps between the radar lines on a grid cell size of 1 cm using inverse distance squared weighting. Then the time slices were interpolated vertically to complete the three-dimensional grid used for visualization.

**Excavation**

After the tracer experiments and GPR surveys were complete, the soil and regolith were excavated to expose and document PF pathways and to compare with the interpretation of the time-lapse GPR data. The excavations were performed manually using shovels and trowels and a pickax once the saprock was reached. The excavation started 0.20 cm downslope of the injection trench to preserve the stability of the excavation wall and to prevent accidental contamination by the heavily dyed soil in the trench. To document the vertical and lateral change in PF pathways, the soil was excavated in a series of horizontal layers at depths of 0.05, 0.08, 0.25, and 0.38 m and photographed. Excavation ceased in the downslope direction once the dye no longer stained the soil and bedrock.

**Identification of Dye-Stained Pixels**

The procedure to identify dye-stained areas within an excavation can range from purely qualitative (i.e., separating stained from unstained soil based on visual inspection and investigator judgment) to quantification of dye concentration from the soil color (Persson, 2005). The latter requires high-quality camera equipment and careful color correction of the photographic images for white balance, exposure settings, and uneven illumination in the field. In addition, laboratory analysis of soil samples must be performed to develop a soil-specific calibration curve for color vs. concentration. We opted for a qualitative approach but chose one that is independent of the observer by using a k-means unsupervised clustering algorithm to classify the pixels as either stained or unstained.

The RGB color values were first converted into the L*a*b* color space (Baldevbhai and Anand, 2012), also known as the “CIELAB, 1976” color space. In this color space, the L axis ranges from light to dark, the a axis ranges from red to green, and the b axis ranges from blue to yellow. Thus, all of the chromicity information is contained within the a–b plane, independent of pixel brightness. The L*a*b* color space has the further advantage of perceptual uniformity. A change of the same amount at any point along any axis should produce an equivalent change in human perception, an approach pioneered in the development of the Munsell color system.

The k-means clustering (Jain, 2010) was then applied to the pixel values in the a–b plane so they are grouped by color independent of illumination. The k-means algorithm is an unsupervised classification algorithm; the operator does not provide a priori information about which pixels are stained. The input parameters are (i) a distance measure, (ii) the number of clusters, k, and (iii) initial seed points for each cluster. The algorithm then seeks to divide the dataset into the k clusters comprised of pixels that are closer to the mean of each cluster than the means of neighboring clusters. As a distance measure, we chose the commonly used Euclidean distance as measured in the a–b color plane. The number of clusters is chosen to be the minimum where dye-stained pixels are not lumped with other color changes, such as soil horizon boundaries. This required minimum number of clusters is site dependent. The initial cluster seeds are chosen randomly, but running the clustering algorithm repeatedly with different seed points (we chose...
n = 6) mitigates the danger that the solution will converge to a local minimum. The color conversion and $k$-means analysis were performed using MATLAB, and examples of the procedure are provided in the software’s documentation (The MathWorks, 2015).

**Infiltration Experiment at the Secondary Site**

An additional infiltration experiment was performed on a nearby hillside (Fig. 1) in the same formation where the soil had been removed for road aggregate, exposing the saprock (Nyquist et al., 2015). Water was poured on the exposed saprock surface through a perforated pipe to create a line source. The wetted surface was photographed to show where lateral flow and infiltration occurred relative to the exposed fractures on the saprock surface. Water was also directly poured on exposed fractures and infiltration rates were timed. This simple experiment was used to provide direct observation of the relationship between fracture patterns and lateral flow and additional ground-truthing for the infiltration experiment.

**Results and Discussion**

**Excavation and Saprock Infiltration**

Dye was not visible in the O or A horizons (a total of ~5 cm thick). At a depth of 7 to 8 cm, a finger of dye was located in the center part of the excavated area, centered on a root that had a diameter of 2 cm, which indicated a PF pathway along a root channel. Deeper excavation showed that the dye was no longer centered on this root. However, near the center of the excavation pit, two large fingers of dye (~15 by 12 cm) were exposed, roughly 15 cm apart, that started at the base of several roots. Furthermore, the two largest roots were located above the two largest dye plumes. Relic bedding planes were observed below the roots, and the dye stained the surface of the bedding planes. The bedding planes were oriented 33° off strike from the injection trench, which had a strike of 269°.

By the time the excavation had reached a depth of 0.25 m, distinct fingers of dye were apparent, two of which are visible in Fig. 3. It was clear that the dye infiltration was at least partially controlled by the relict bedding. At the base of the excavation pit, the widths of the dye fingers increased. From the injection trench to the end of the dye-stained area, the extent of lateral flow was ~0.40 m. The dye traveled more than 0.38 m in the vertical direction and faded rapidly beyond this depth across most of the horizontal plane of excavation (Fig. 3). Some dye was observed on the left side (looking upslope) below this plane, but the total depth of penetration was not measured because manual excavation of the saprock was not feasible. The total excavated volume stained with dye was 0.4 m downslope by 0.65 m wide by 0.35 m deep. From a simple mass balance calculation (multiplying the observed volume stained by dye times porosity), we can deduce that, for a porosity of 45%, one-third of the dye solution infiltrated the saprock instead of flowing laterally along the soil–saprock boundary. Although the estimation of the visible dye area is rough, the selected porosity is overly conservative. Lower porosities would require even more infiltration to explain the limited extent of lateral flow.

The heterogeneous nature of the saprock was revealed in the infiltration experiment at the nearby hillside where saprock was exposed. Water poured on the saprock surface followed the surfaceto-topography for several meters, crossing bedding planes, but infiltrated once it encountered a permeable fracture. Later, some isolated wet areas appeared several meters downslope from where the water had infiltrated, forming return flow (Fig. 4). The saprock fractures are in part the result of repeated freeze–thaw cycles (Jin et al., 2010). Fractures of this origin close at depth, and when they filled with water the decrease in permeability with depth could create return flow. Water poured directly on fractures showed that a fracture set with 10-m spacing had higher infiltration rates than other fractures, including fracture intersections. The experiment also revealed that not all fractures lead to infiltration, which helps explain previous experiments nearby in the Shale Hills study site that showed more lateral flow than the trench excavated here. This experiment showed PF pathways along horizontal fissures in...
saprock that conducted water downslope in addition to vertical infiltration through the fractured saprock.

**Ground-Penetrating Radar Differencing**

Determining temporal changes by simple radargram differencing is an inherently unstable process (Versteeg and Birken, 2000) because changes in soil moisture change the velocity of radar waves and create a phase misalignment between the before and after radargrams. We investigated several techniques to compensate the phase shift between the radargrams collected before and after the tracer release, such as dynamic time warping (Hale, 2013; Nyquist et al., 2014) and complexity-invariant distance measures (Batista et al., 2014). We did not see a significant improvement over direct differencing of the radargrams’ imaging of the areas with soil moisture changes, so we elected to use the simplest approach (i.e., direct subtraction of the corresponding background survey line). In our experiment, the addition of the tracer solution principally changed the radar reflection amplitudes. Although velocity changes associated with water retention in the upper 0.5 m created some cumulative delay, they did not cause enough phase shift to obscure the changes apparent in the differenced radargrams.

Comparison of the same radar line extracted from the two separate background surveys showed that the heterogeneous structure of the soil was closely reproduced. There were slight changes in the reflection energies between the two surveys, so differencing did not result in complete cancelation, yielding instead a fainter ghost image of the original structure. The degree of cancelation varied from line to line and along a given line, probably the result of small changes in antenna angle and coupling with the ground surface during data collection. A sample trace selected from the center of the first line from both background surveys further illustrates the reproducibility of the data (Fig. 5).

---

**Fig. 4.** Photograph of soil moisture revealed in the infiltration experiment along the excavated dip slope (secondary site). Water moved laterally ~4.5 m until preferential flow was encountered. Further lateral flow may take place through the horizontal fissures in the saprock and form the observed return flow at a downslope location (wet surfaces marked at 2.8 m).

**Fig. 5.** Comparison of the line closest to the injection trench from the two background surveys collected before injection, the differenced section, and center traces from both lines showing the reproducibility of the data. The first 2 ns (8 cm) have been muted to eliminate the air wave. Depth was calculated assuming a homogeneous velocity of 0.077 m ns⁻¹.
The before, after, and differenced radargrams for the line closest to the trench (20 cm away) for the water release with no dye added shows changes in the reflection amplitudes (Fig. 6). The GPR survey line is centered on the trench, which extends from 0.5 to 1.5 m, and the reflection amplitudes have changed in a roughly corresponding zone, with a disturbance width that increases with depth. There is an indication of asymmetry, with more of the change occurring between 0.4 and 1.0 m than between 1.0 and 1.4 m, which is consistent with the dip of the relict bedding and dye staining pattern seen in Fig. 3.

The same line after the dye release (Fig. 7) shows a similar pattern with a slight increase in the magnitude of the change in reflection

---

**Fig. 6.** Comparison of the line closest to the injection trench from before and after the water injection. The differenced section shows the change starting at about 15 cm broadening with depth and trending slightly to the left, which is consistent with infiltration path seen in the excavation (Fig. 3). Note that the vertical blue lines in the differenced radargram align with the ends of the injection trench with ran from 50 to 150 cm and was upslope 20 cm from this line. Change in amplitude and a slight phase shift can be seen in the sample trace from the center of the section. The first 2 ns (8 cm) have been muted to eliminate the air wave. Depth was calculated assuming a homogeneous velocity of 0.077 m ns⁻¹.

**Fig. 7.** Comparison of the line closest to the injection trench before and after the dye solution injection. Note that the vertical blue lines in the differenced radargram align with the ends of the injection trench, which ran from 50 to 150 cm and was upslope 20 cm from this line. The change in the differenced section has slightly increased in size (compared with Fig. 6). The phase shift seen in the sample trace has also increased. The first 2 ns (8 cm) have been muted to eliminate the air wave. Depth was calculated assuming a homogeneous velocity of 0.077 m ns⁻¹.
amplitude. This reflectivity increase is the opposite of what was reported by Haarder et al. (2011), who found a radar response weakened by signal attenuation in the zone of dye infiltration, which they attributed to the elevated conductivity of the dye solution. There are two competing factors at work here: (i) signal enhancement by the increased contrast between wet and dry layers; and (ii) attenuation by conductive pore fluids. In a study of vertical infiltration, Haarder et al. (2011) saturated a portion of the sediments below their grid, and thus attenuation dominated. However, they reported an increase in radar reflection amplitude along the edges of the infiltration area, which they attributed to lateral flow at the edges of the wetting front. This reflectivity increase just outside the infiltration area would be where the tracer-induced moisture changes remained below saturation. We suggest that the combination of well-drained soils and slow lateral migration kept our study plot unsaturated and that the reflectivity increases were more significant than the losses due to signal attenuation; thus, the area influenced by the infiltration appeared as a zone with enhanced reflectivities.

It is interesting to compare the three-dimensional visualization of the background GPR data (no differencing) with the excavation. In general, the data simply reflect the highly heterogeneous soil layer, with diffractions from rock fragments and no continuous reflection horizons. In the two-dimensional radargrams, it is difficult to determine the soil thickness from the radar because there is no clear reflection from the saprock (Fig. 5), probably because of the gradual change in dielectric constant across this gradational boundary. Even in soil profiles, it is difficult to delineate the soil–saprock interface. However, structure does emerge in the three-dimensional radar images near the soil–saprock interface (Fig. 8). A horizontal slice through the GPR data hints at diagonal banding consistent with the strike direction of the saprock seen in the excavation. The horizontal resolution for the GPR at this depth can be estimated from the radius of the first Fresnel zone (Reynolds, 1997):

\[
r = \left( \frac{\lambda^2}{16} + \frac{\lambda z}{2} \right)^{1/2}
\]

where \( \lambda \) is the wavelength and \( z \) is the depth. At our site, this equation predicts a radius of \(~8\) cm at a depth of \(35\) cm. Thus, the horizontal resolution is insufficient to image the bedding clearly, but larger irregularities in the saprock with the same strike are visible. However, these are only visible in the horizontal slices through the three-dimensional data and are not visible in the individual two-dimensional radargrams.

Horizontal depth slices through the three-dimensional volume made from interpolation of the differenced radargrams after the water and dye solution releases clearly show the change near the infiltration trench dominated by vertical infiltration with limited lateral PF (Fig. 9). The second injection doubled the amount of water released but only slightly increased the size of the plume delineated by the radar, providing further evidence that much of the fluid went into the saprock. Once infiltrated into the saprock, it could not be tracked much farther because of strong attenuation of the radar signal by the shale. Including the 20-cm offset between the infiltration trench and the radar grid, the total lateral flow was only about 40 cm. Comparing the dye staining pattern with images constructed by compositing horizontal and vertical slices through the radar difference volume (Fig. 9) shows agreement. The pattern after the water and dye releases is similar. In both cases, the extent of the changed area is limited to \(30\) to \(40\) cm from the injection trench. The excavation data are also consistent with limited lateral flow.

To compare a vertical slice through the three-dimensional GPR data with the dye staining pattern on the wall of the excavation, we used \(k\)-means clustering to classify pixels as either strained or unstained. We found four to be the minimum number of clusters required for the \(k\)-means algorithm to segment the dye-stained pixels into a distinct cluster, although the white rope and the nearly white portions of the ruler were classified as part of this cluster as well because white objects have a much stronger blue component than soil. When this cluster is superimposed on the corresponding vertical slice through the three-dimensional radar data (Fig. 10D), there is good qualitative agreement between the staining pattern and the change in GPR data. There are no stained areas that GPR did not detect, and the photograph of the excavation shows an unstained area on the wall near the center of the trench that matches an area of minimal change in the radar images. However, there is disagreement near the right edge, where the GPR shows change that was not seen in the dye staining of the excavation.
Results from dye-staining patterns and time-lapse GPR surveys suggest that there was limited fill-and-spill flow at the soil–saprock interface in the study site, where vertical infiltration dominated the flow regime in the permeable saprock (Fig. 8 and 9). Therefore, we proposed a refined fill-and-spill model to include the PF process through fractured saprock before water encounters the bedrock surface (Fig. 11). As depicted in this conceptual model, precipitation rapidly infiltrates the thin soil layer and forms limited lateral PF along the regolith–saprock interface before flow converges into the fissures in the saprock; then the bulk of the water infiltrates the regolith–saprock interface and percolates vertically through the fractured saprock down to the unweathered bedrock surface. The infiltration into the saprock layer results in heterogeneity whereby nearby locations may vary in the extent of lateral flow.

Furthermore, part of the water may flow downslope along the horizontal fissures and reemerge at the ground surface as return flow. Once water encounters the less-permeable bedrock surface, a transient water table may form in small depressions and trigger continuous lateral PF downslope above the bedrock surface; some water may percolate into the aquifer as deep drainage (Fig. 11). This two-layer fill-and-spill model is consistent with the observations of Salve et al. (2012) that saprolite can have PF paths that transfer water to the bedrock without the saprolite wetting up. Because weathered bedrock (or saprock) is commonly present in hillslopes with a thin soil cover, the refined fill-and-spill model has a potentially broader applicability to predict hillslope hydrology. Future efforts are needed to test and improve the fill-and-spill model and incorporate it into hillslope hydrology models.

Fig. 9. Horizontal slices from 8 to 53 cm made through the three-dimensional volume at composited depths were constructed from the volume of radargrams collected after dye solution injection minus the background radargrams. The time window for each slice encompasses the return energy across roughly half a wavelength of ground-penetrating radar signal, with successive slices overlapping to show the change with depth. The scale is in squared amplitude. There is some indication of fingering in the radar slices, but it is blurred by the lack of resolution and interpolation smoothing. The extent of lateral flow is clear. Most of the change is in the 20 cm of the radar grid closest to the injection trench. The trench is at 1.2 m in the y direction and extends 0.5 to 1.5 m in the x direction, as shown for the uppermost slice in the top left.
Time-lapse GPR could image bulk moisture changes in the shallow subsurface but could not resolve the finer structure of fingering visible in the dye excavation at the SSHCZO. At this site, the bulk of the tracer infiltrated vertically to the regolith–saprock boundary. Then, rather than follow the fill-and-spill model, it infiltrated fractures in the saprock. This saprock infiltration was unexpected because geophysical and soil moisture probe data from similar experiments at nearby sites with deeper soils (Guo et al., 2014; Lichtner et al., 2012) showed evidence of lateral flow for at least 1 m.

Fig. 10. The radar images are a composite of a vertical $x$–$z$ slice closest to the trench (0.20 m), a horizontal $x$–$y$ slice at the depth equal to the base of the excavation (0.38 m), and two vertical $y$–$z$ slices at the $x$ limits of the grid. The patterns of change after (A) the water release and (B) the dye solution release are similar. The scale is normalized to the brightest reflection. (C) The vertical excavation face closest to the injection trench is shown with and without the $k$-means identification of the dye-stained pixels for comparison. (D) The pixels classified as dye stained are shown superimposed on the corresponding three-dimensional radar collected after the dye release for comparison.

Fig. 11. A refined fill-and-spill model with preferential flow (PF) processes occurring at the top of two layers (saprock and bedrock). In a hillslope with a thin soil cover underlain by a weathered bedrock (or saprock) layer, the bulk of the precipitation rapidly infiltrates soils and saprock and forms limited lateral PF on the soil–saprock interface. Lateral PF may take place through the horizontal fissures in the saprock and form return flow downslope. Once water encounters a fresh bedrock surface, a transient water table may perch at small depressions and then trigger lateral PF downslope.
There is an indication that the radar data could detect large heterogeneities in the flow pattern. Near the center of the survey area, dye solution bypassed a zone roughly 10 cm wide, and there is a corresponding unchanged zone in the three-dimensional radar image, but the resolution of the radar was too low to detect finer structure. Grasmueck et al. (2005) showed that maximum detail can be resolved using “full-resolution” three-dimensional radar collection, where the spacing between the radar lines is less than a quarter wavelength. For our site, this would have dictated a maximum line spacing of a little over 1 cm, increasing the acquisition time fivefold and requiring roughly 1.24 h to collect the data even for our small 1- by 2-m grid. Thus, increased spatial resolution would come at the cost of temporal resolution, a problem when monitoring highly dynamic processes. In the future, data acquisition time may decrease because multi-antenna radar systems are becoming increasingly available, although to our knowledge none has yet been fabricated with an inter-antenna spacing as small as 1 cm. Exploring the potential benefits of full-resolution three-dimensional GPR for imaging shallow, heterogeneous soils is a promising avenue for research.

Because the contrast created by dye staining depends on soil type and illumination and because GPR change detection depends on background noise and soil dielectric proprieties, a site-specific empirical relationship between dye concentration and GPR response would be required to choose appropriate thresholds for a statistical comparison of the size and shape of the changed regions mapped by each method. This is further complicated the different resolutions of each method. Dye pattern changes can be seen on a finer scale than can be mapped by GPR. Quantitative comparison is an area for further research but was beyond the exploratory scope of this study.

Conclusions

The model that is emerging for storm flow at the SSHCZO is more complicated than simple lateral flow along a regolith–saprock interface because of unevenly spaced, permeable fractures and anisotropy of the underlying saprock. A refined fill-and-spill model is proposed to predict the complex hillslope hydrology of the study site, which is a typical forested hillslope with a thin soil cover and a fractured bedrock layer. Such an improved empirical understanding of subsurface hydrology can help model stormwater patterns in hillslopes.

In the shallow vadose zone, dye tracer tests offer the most compelling evidence of flow pathways but cannot be reproduced because excavation destroys the soil fabric. Geophysical techniques, such as GPR, provide a promising opportunity to map and monitor subsurface flow pathways noninvasively and can be repeated in the field. However, due to the lack of appropriate approaches, ground-truthing of subsurface flow patterns derived from geophysical data is often lacking. In this study, we demonstrated the potential of combining geophysical investigation and dye tracer tests to reveal subsurface PF in a natural hillslope. Given that the performance of geophysical techniques is highly dependent on local conditions, such as soil textures, soil wetness, subsurface anisotropy, and surface topography, we suggest conducting small-scale experiments that compare geophysical data with dye staining patterns to evaluate the accuracy of geophysical results and guide the selection of proper settings of the geophysical techniques (e.g., the antenna frequency of GPR). Then the optimized geophysical surveys can be applied to a greater spatial scale. We believe that geophysical techniques that are validated by direct ground truth can improve the field investigation of subsurface hydrology in future studies.

Acknowledgments

This work was supported by a seed grant from the National Science Foundation Critical Zone Observatory program, which funded the Susquehanna–Shale Hills Observatory (SSHCZO). The SSHCZO is administered by the Pennsylvania State University under grants to Christopher J. Duffy (EAR 07-25019) and Susan L. Brandt (EAR 12-39285, EAR 13-31726). The research was conducted in Penn State’s Stone Valley Forest, which is supported and managed by Penn State’s Forestland Management Office in the College of Agricultural Sciences. Li Guo and Henry Lin are supported in part by the US National Science Foundation Hydrologic Sciences Program Grant EAR-1416881 (PI: H. Lin).

References


