Hydrol. Earth Syst. Sci., 26, 429–444, 2022
https://doi.org/10.5194/hess-26-429-2022

Hydrology and Earth System Sciences
Open Access
EGU

Watershed zonation through hillslope clustering for tractably quantifying above- and below-ground watershed heterogeneity and functions

Haruko M. Wainwright[1,2], Sebastian Uhlemann[1], Maya Franklin[1], Nicola Falco[1], Nicholas J. Bouskill[1], Michelle E. Newcomer[1], Baptiste Dafflon[1], Erica R. Siirila-Woodburn[1], Burke J. Minsley[3], Kenneth H. Williams[1,4], and Susan S. Hubbard[1]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[2]Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]U.S. Geological Survey, Denver, CO 80225, USA
[4]Rocky Mountain Biological Laboratory, Crested Butte, CO 81224, USA

Correspondence: Haruko M. Wainwright (hmwainw@mit.edu)

Received: 26 April 2021 – Discussion started: 10 May 2021
Revised: 8 December 2021 – Accepted: 10 December 2021 – Published: 31 January 2022

Abstract. In this study, we develop a watershed zonation approach for characterizing watershed organization and functions in a tractable manner by integrating multiple spatial data layers. We hypothesize that (1) a hillslope is an appropriate unit for capturing the watershed-scale heterogeneity of key bedrock-through-canopy properties and for quantifying the co-variability of these properties representing coupled ecohydrological and biogeochemical interactions, (2) remote sensing data layers and clustering methods can be used to identify watershed hillslope zones having the unique distributions of these properties relative to neighboring parcels, and (3) property suites associated with the identified zones can be used to understand zone-based functions, such as response to early snowmelt or drought and solute exports to the river. We demonstrate this concept using unsupervised clustering methods that synthesize airborne remote sensing data (lidar, hyperspectral, and electromagnetic surveys) along with satellite and streamflow data collected in the East River Watershed, Crested Butte, Colorado, USA. Results show that (1) we can define the scale of hillslopes at which the hillslope-averaged metrics can capture the majority of the overall variability in key properties (such as elevation, net potential annual radiation, and peak snow-water equivalent – SWE), (2) elevation and aspect are independent controls on plant and snow signatures, (3) near-surface bedrock electrical resistivity (top 20 m) and geological structures are significantly correlated with surface topography and plan species distribution, and (4) K-means, hierarchical clustering, and Gaussian mixture clustering methods generate similar zonation patterns across the watershed. Using independently collected data, we show that the identified zones provide information about zone-based watershed functions, including foresummer drought sensitivity and river nitrogen exports. The approach is expected to be applicable to other sites and generally useful for guiding the selection of hillslope-experiment locations and informing model parameterization.

1 Introduction

Predictive understanding of watershed functions is often hindered by the heterogeneous and multiscale fabric of watersheds (e.g., Peters-Lidard et al., 2017). Heterogeneity exists within each of the watershed compartments, including above-ground compartments (i.e., plant species distribution and plant dynamics, topography) and below-ground compartments (i.e., soil and bedrock structures/properties). Such watershed *patterns* influence ecohydrological and biogeochemical *processes*, which in turn affect watershed *functions* and create emerging *patterns* as feedback (Sivapalan, 2006). Since watersheds consist of diverse biotic and abiotic compartments, watershed functions include diverse signatures, including hydrological (i.e., partition, storage, and

release of water), ecological (e.g., species adaptation, productivity), and geochemical (e.g., nutrient cycling, solute export) signatures (Sivapalan, 2006; Wagener et al., 2007). There is a multiscale nature of heterogeneity such that different processes have different characteristic scales. Watershed hydrology modeling studies typically use the grid size from 100 m to 1 km (e.g., Foster et al., 2020; Maina and Siirila-Woodburn, 2020), whereas soil moisture is known to vary on the order of several or several tens of meters (e.g., Engstrom et al., 2005; Wainwright et al., 2015), and biogeochemical dynamics can vary within 1 m or less (e.g., Burt and Pinay, 2005; Groffman et al., 2009).

There have been extensive studies investigating how the heterogeneous watershed organization influences water, energy, and nutrient cycling and their fluxes (e.g., Peters-Lidard et al., 2017). There are two directions for tackling this problem: a bottom-up Newtonian approach or a top-down Darwinian approach. The Newtonian approach is a reductionist approach that describes a system by a set of mass/energy/momentum conservation equations with spatially variable parameters. Recently, integrated hydrological and reactive transport models have been successfully implemented to describe and predict watershed behaviors from hillslope to watershed scales (e.g., Maxwell and Kollet, 2008; Li et al., 2017). In addition, the hydrological response unit concept has been used to classify the landscapes based on spatial datasets (e.g., land-cover types, elevation, and soil maps) and to parameterize hydrological models (Flugel, 1997; Aytaç, 2020). On the other hand, the Darwinian approach identifies rules or organizing principles governing spatial patterns of complex datasets and defines watersheds as self-organized and co-evolved units by watershed functional traits (McDonnell et al., 2007). Catchment scaling and similarity concepts have been used to synthesize the catchment datasets across scales and to classify catchments (Wagener et al., 2007; Thompson et al., 2011; Sawicz et al., 2011; Krause et al., 2014).

In parallel, there have been recently significant advances in understanding and quantifying the watershed-scale heterogeneity of bedrock-to-canopy terrestrial compartments, which regulate water and nutrient cycling and their exports, particularly through the critical-zone observatory (CZO) network (Brantley et al., 2017). In particular, Pelletier et al. (2018) have highlighted the control of slope aspects on ecosystem and critical-zone systems, finding that, for example, in water-limited systems, the north-facing slopes have less evapotranspiration and hence higher soil moisture, deeper weathering, and larger nutrient retention in soil (e.g., Hinckley et al., 2014). Such advances are largely attributed to a variety of spatially extensive characterization technologies across bedrock-to-canopy compartments, which provide various *patterns* (Hubbard et al., 2020). High-resolution digital elevation models (DEMs) from lidar have been applied to better understand the relationship between geomorphology and hydrology (Prancevic and Kirchner, 2019) as well as to measure snow depths and snow-water equivalent (SWE) over a basin scale (Painter et al., 2016). Lidar data were also able to inform near-surface soil properties (Patton et al., 2018; Gillin et al., 2015), hydrological connectivity (Jencso et al., 2009), and biogeochemical hotspots (Duncan et al., 2013). In addition, hyperspectral remote sensing can map plant traits (e.g., Asner et al., 2015), leaf water content (e.g., Colombo et al., 2008), leaf chemistry (e.g., Feilhauer et al., 2015), and other properties, which are also proxies for soil biogeochemistry (Madritch et al., 2014). At the same time, geophysics has been extensively used to characterize the subsurface structure and to estimate soil and bedrock properties. Surface geophysics has been used to measure bedrock depth, weathering zone thickness, and other properties (e.g., de Pasquale et al., 2019), contributing to hillslope-scale hydrological characterization and modeling. Surface electrical resistivity tomography (ERT) and seismic data have also revealed the influence of tectonic stresses and hydrological processes on bedrock fracturing and weathering (Rempe and Dietritch, 2018; St. Clair et al., 2015). Airborne geophysics – particularly airborne electromagnetic (EM) surveys – was originally developed for mineral exploration but is now increasingly used for water-resource applications (e.g., Barfod et al., 2018; Ball et al., 2020).

Despite these advancements, there are still challenges in associating watershed *functions* with heterogenous watershed *patterns*. It remains a challenge to integrate multitype and multiscale datasets, including ground-based point measurements and airborne or satellite remote sensing datasets. Although the above-ground properties, such as topography and plant characteristics, can be mapped over the watershed scale, the subsurface variability is still difficult to map over that scale, which is one of the biggest uncertainties in hydrology (Fan et al., 2019). In addition, even though hillslope-scale characterization and experiments (such as tracer tests) can be extremely useful for providing detailed information about watershed functions (e.g., Hinckley et al., 2014), it is difficult to select several hillslopes for such intensive characterization or to gauge the representativeness of one hillslope for an entire watershed.

In this study, we develop and test the concept of a zonation approach for tractably characterizing the organization of a watershed based on multiple spatial data layers and how these characteristic *patterns* aggregate to predict watershed *functions*. The clustering-based zonation approaches have emerged recently as effective spatial data integration methods that use spatial clustering to identify regions or zones that have unique distributions of heterogeneous properties and key functions relative to neighboring regions (Hubbard et al., 2013; Wainwright et al., 2015; Devadoss et al., 2020; Hermes et al., 2020). For watershed zonation, we consider a hillslope to be a fundamental unit for watershed hydrology and element cycling, funneling water and elements from the ridge to the river (Fan et al., 2019) and also representing aspect controls on critical zones (Pelletier et al., 2018). We

follow Band (1989) and Band et al. (1991, 1993) to investigate the appropriate scales of hillslopes for capturing the watershed heterogeneity while limiting the internal variance within hillslopes.

We then hypothesize that (1) a hillslope is an appropriate unit for capturing the watershed-scale heterogeneity of key bedrock-through-canopy properties and for quantifying the co-variability of these properties representing coupled ecohydrological and biogeochemical interactions, (2) we can identify a group of hillslopes or watershed-scale zones that have unique distributions of these properties relative to neighboring parcels, and (3) the identified zones can capture the variability of key watershed functions. We demonstrate our approach using the airborne and spatial datasets collected in the East River watershed region near Crested Butte, Colorado, USA (Hubbard et al., 2018). We apply and compare multiple clustering methods to understand the characteristics, commonality, and differences among each method. Finally, we validate the zonation hypothesis based on the datasets that define key watershed functions, including drought sensitivity of plant productivity and water/nitrogen export.

2 Site and data

We consider the domain of approximately 15 km by 15 km (Fig. 1) near Gothic, Colorado, USA, which is the same area used in a recent study (Wainwright et al., 2020). As described in Hubbard et al. (2018), the domain includes four catchments, the East River, Washington Gulch, Slate River, and Coal Creek. It is a part of the Elk Range in the Rocky Mountains, with elevation from 2800 to 4000 m (Fig. 1a). The major land-cover types (NLCD 2011 map; Homer et al., 2012) are rock outcrop (12 %), evergreen forest (29 %), deciduous forest (18 %), grassland (30 %), and woody wetland (6 %). Geology within the domain is diverse, including Paleozoic, Mesozoic, and Cenozoic sedimentary rocks (siltstones and sandstones of the late Permian Maroon Formation; shales and sandstones of the upper Cretaceous Mancos Shale and Mesaverde group, respectively; siltstones and sandstones of the Eocene Wasatch Formation) and Miocene igneous intrusive rocks of predominantly granodioritic composition (Gaskill et al., 1991).

The spatial data layers include the USGS land-cover map (NLCD, 2011), the digital geological map of Colorado (Green, 1992), and the soil texture maps (percent clay and percent sand) from the POLARIS database (Chaney et al., 2016, 2019). In addition, we used four airborne datasets (Sect. S1 in the Supplement): an airborne electromagnetic (AEM) survey acquired in fall 2017 (Minsley and Ball, 2018; Uhlemann et al., 2022; Zamudio et al., 2020), lidar and hyperspectral data collected by the National Ecological Observation Network (NEON) team in June 2018 (Chadwick et al., 2020), and NASA Airborne Snow Observatory (ASO) data collected in April 2018 (Painter, 2018).

To test the zonation hypothesis, we used datasets representing two key functions: foresummer drought sensitivity of plant productivity and river nitrogen export. The foresummer drought sensitivity map based on the Landsat normalized difference vegetation index (NDVI) was developed by Wainwright et al. (2020) to represent the plant productivity responses to early snowmelt and subsequent drought conditions in the primary growing season. To create this map, Wainwright et al. (2020) performed the linear regression of the historical peak Landsat NDVI as a function of the Palmer drought severity index in June. They then defined the slope of the linear fit as the foresummer drought sensitivity, which represented the magnitude of peak plant productivity changes with respect to the drought condition in the growing season. Since the satellite images were not used in clustering, they were considered independent datasets. In addition, the annual discharge and nitrogen export were computed from the streamflow and chemistry data at the subcatchments (Fig. 1) within the domain used in Carroll et al. (2018). The nitrogen export was computed in the same manner as Newcomer et al. (2021).

3 Methodology

3.1 Watershed and hillslope characteristics

We followed Wainwright et al. (2020) to compute pixel-by-pixel topographic metrics based on the lidar DEM, including slope, topographic wetness index (TWI), and annual net potential solar radiation. We selected metrics based on the previous studies that reported their relevance to soil moisture, soil thickness, water quality, and others (Mohanty et al., 2000; Gillin et al., 2015; Lintern et al., 2018). The annual net potential solar radiation – a function of the aspect, slope, and solar angle – is considered a better metric to represent the intensity of solar radiation than the aspect itself, which is circular (0 and 360° are the same; Wainwright et al., 2020). In parallel, we delineated each hillslope based on the DEM, following Noël et al. (2014) and using Topotoolbox (Schwanghart and Scherler, 2014). The hillslope delineation algorithm first identifies the stream segments that are a collection of pixels that have larger flow accumulation (i.e., a larger drainage area) than a certain threshold value, using the flow routing algorithms in Topotoolbox. It then finds the pixels in both sides of each stream segment and also the drainage area leading to each of these pixels. This process yields two lateral hillslopes in both sides of each stream segment. For the first-order stream, the algorithm also defines the headwater hillslope, which is the drainage area leading to the pixel at the origin of the stream. Although these hillslopes are based on surficial water routing, we assume that the DEM captures near-surface hydrological connectivity as documented by Jencso et al. (2009).

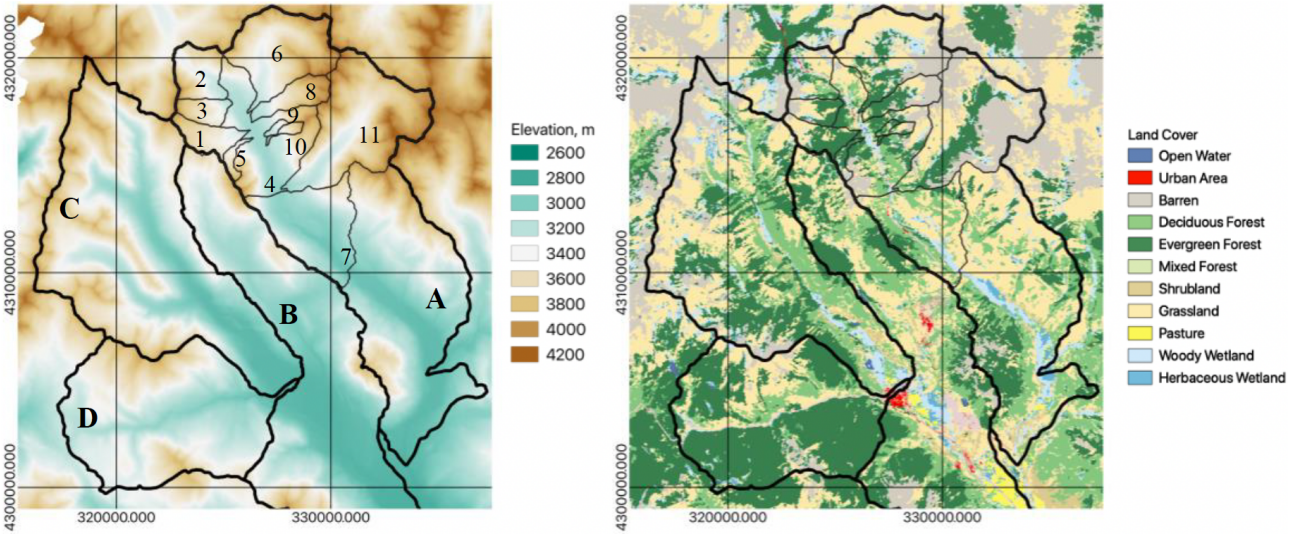

Figure 1. Study domain with **(a)** elevation and **(b)** land-cover map (USGS NLCD 2011 map; Homer et al., 2012). The black lines are the boundaries of the four watersheds: (A) East River, (B) Washington Gulch, (C) Slate River, and (D) Coal Creek. The thin black boundaries are subcatchments of the East River watershed from Carroll et al. (2018): (1) Rock, (2) East Above Quigley (EAQ), (3) Quigley, (4) East Below Copper (EBC), (5) Gothic, (6) Rustlers, (7) Pumphouse, (8) Bradley, (9) Marmot, (10) Avery, and (11) Copper.

We then defined the characteristics of each hillslope, based on the spatial data layers (Fig. S1). We computed the average values of the DEM topographic metrics, AEM-based bedrock resistivity, NEON products (NDVI, Normalized Difference Water Index, NDWI, and biomass) at the peak growing season (2018), peak SWE (2018), and soil texture in each hillslope. In addition, we computed the relief of each hillslope, which was the difference between the minimum and maximum elevations. For the categorical variables such as land-cover types and geology, we computed the percent coverage of each plant type and surface geology in each hillslope. These 17 hillslope features are defined for each hillslope (Table S1).

As Band (1989) and Band et al. (1991) noted, the hillslope delineation can create different sets of hillslopes, depending on the threshold drainage areas that define the stream segments. For the hillslope metrics defined by the average (e.g., elevation, slope, and others), we evaluated how these averaged metrics can represent the overall variability of the watershed properties and how this representation changes depending on the different threshold drainage areas. We consider that the variance of pixel-by-pixel properties represents the overall variability over this domain, while the variance of hillslope-averaged metrics represents the variability across the hillslopes. We computed the ratio between the across-hillslope variance and the overall variance, representing how much watershed-scale variability the hillslopes can capture. In addition, as a contrast, we computed the variance of each property within the upscaled pixels by taking the averaged values in larger pixels compared to the original 9 m pixel. We computed the ratio between this across-upscaled-pixel variance and the overall variance. In this way, we can investigate the difference between hillslope-based spatial aggregation compared to standard pixel-based upscaling for each of the key watershed properties.

3.2 Cluster-based approach to identify watershed zones

Based on the hillslope features, we first evaluated the correlations among multivariate above-/below-ground properties. Although such multivariate co-variability has been analyzed using principal component analysis (Devadoss et al., 2020), we used scatter plots and correlation coefficients in this study because of nonlinearity. We then applied three commonly used unsupervised clustering methods: K-means (KM), hierarchical clustering (HC), and Gaussian mixture models (GMMs; Hastie et al., 2001; Kassambara, 2017). We scaled each feature by the mean and standard deviation and defined the dissimilarity between two data points based on the Euclidean distance. The characteristics of each method are described in Sect. S2. Multiple methods are often evaluated based on true classes or labeled datasets (Rodriguez et al., 2019), which are not available here. We used a silhouette score, which represented how similar a given data point was to its own cluster compared to other clusters. For GMMs, we used the Bayesian information criterion (BIC) to select the number of clusters.

After the clusters are defined, we transfer the clusters – the group of hillslopes that have similar features – to the spatial map as zones. We identify the common zones across the three methods as well as the zones that differ. We then evaluate the distribution of hillslope features and functions in each zone using box plots to define the characteristics of each zone. The

foresummer drought sensitivity (Wainwright et al., 2020) is a watershed function available throughout the domain that allows us to quantify the hillslope-average values in the same manner as other spatial data layers. In addition, we computed the spatial coverage of each zone in each subcatchment and compared them to the ratio between the annual nitrogen export and total discharge, which is considered a key metric indicating how watersheds retain and lose nutrients (Newcomer et al., 2021).

4 Results

4.1 Hillslope scales

The ratio between the across-hillslope variance and the overall variance (Fig. 2a) is generally high for the elevation, net annual potential solar radiation (radiation), peak SWE, and bedrock resistivity up to 0.75, which means that the hillslope-averaged metrics capture the watershed-scale variability of these variables and that the within-hillslope variability is small compared to the across-hillslope variability. TWI and NDVI, on the other hand, have a low ratio, which means that the within-hillslope variability is significant. The ratio increases as the drainage area decreases, since the smaller the hillslopes, the better they capture small-scale variability. However, the variance ratio of the elevation and radiation reaches a plateau with the drainage area around 10^6 m^2. This means that the internal variability within hillslopes is limited up to this threshold drainage area and that the hillslope metrics are representative of the overall variability up to this threshold. Based on this result, we selected 810 000 m^2 as the threshold drainage area in the subsequent analysis.

We can compare this hillslope-based averaging (Fig. 2a) with pixel-based averaging/upscaling (Fig. 2b). Similarly to hillslope-based averaging, the ratio between the across-upscaled-pixel variance and the overall variance increases as the pixel size decreases. The magnitude is similar such that the elevation, peak SWE, and bedrock resistivity have a higher ratio, meaning that the upscaled properties can capture the overall variability of these properties. The exception is the radiation, since the pixel-averaged radiation captures only up to 60 % of the overall variance, while the hillslope-averaged radiation captures up to 70 %. In addition, differently from the hillslope average, the ratio for radiation in Fig. 2b keeps increasing without reaching a plateau. This means that a representative size or scale does not exist when we use pixel-based upscaling.

4.2 Watershed zonation

Figure 3 shows that the above- and below-ground hillslope features are significantly correlated with each other. In particular, elevation is correlated with all other features except for net annual potential radiation. The hillslopes at higher elevation have steeper slopes and lower TWI, higher bedrock resistivity, and higher peak SWE. There are some nonlinearities: TWI increases beyond the linear function at lower elevations. The relationship between peak NDVI and elevation is quadratic, having peaks at middle elevation (corresponding to around 3300 m). The net annual potential radiation is weakly correlated with slope, TWI, and peak SWE. The correlation coefficients among the hillslope-averaged features are significantly higher than the pixel-by-pixel ones (Fig. S2).

The clustering results are then mapped spatially as the watershed zones (Fig. 4). The number of clusters is six, since the highest silhouette score is high at six clusters across the three methods (Fig. S3a), and BIC for GMM is also the highest at six clusters (above four; Fig. S3b). To compare the results from these three methods, we first identified six common zones across the three methods that have the overlapping coverage, starting from the GMM-based map as a basis. Zone 7 appears only in the HC result (Fig. 4c) and is designated as a separate zone; the hillslopes in Zone 7 are included either in Zone 3 or Zone 4 in the GMM and KM results. The frequency maps (Fig. S4) show that most hillslopes are consistently categorized in the same zones among the three methods. The zonation maps from GMM (Fig. 4a) and KM (Fig. 4b) are quite similar, although small fractions of the hillslopes have different designations between the two maps. The zonation map from HC (Fig. 4c) is also similar to the ones from GMM and KM, except for Zone 7, which is a part of Zone 3 or Zone 4 in the GMM and KM result.

Box plots shows the characteristics of each identified zone (Figs. 5 and 6). Although each map has six zones, we use seven zones that are found across the three clustering methods. The statistics are computed by aggregating the zonation from the three methods (Fig. S4) such that the zonation of each method has a weight of 1/3. Elevation (Fig. 5a) increases from Zones 1 to 6, while Zone 7 is similar to Zone 4. Slope (Fig. 5b) shows a similar progression to elevation, which is consistent with the correlation plot (Fig. 3), although Zone 7 has steeper hillslopes than expected for its elevation range. Relief (Fig. 5c) also has a similar order to the elevation, except that Zones 4 to 7 are similar, because the higher-elevation hillslopes are smaller with lower relief, even though the slope is high. TWI (Fig. 5d) has an opposite trend to elevation, consistent with the correlation plot (Fig. 3). In terms of the net annual potential radiation (Fig. 5e), the grassland-dominated hillslopes (Zones 1, 2, and 5) tend to be higher, while the conifer-dominated hillslopes (Zones 3, 4, and 7) are lower except for Zone 4 being around the average. The bedrock resistivity (Fig. 5f) is higher at the higher-elevation zones (Zones 4–7) or an intrusive-rock-dominated zone (Zone 4). NDVI (Fig. 5g) is higher in the forest-dominated zones (Zone 2, 3, 4, and 7). Zone 6 has a significantly lower NDVI, since it is predominantly in the barren region. NDWI (Fig. 5h) is lower in the low-elevation zones (Zone 1). Peak SWE (Fig. 5i) is associated with the

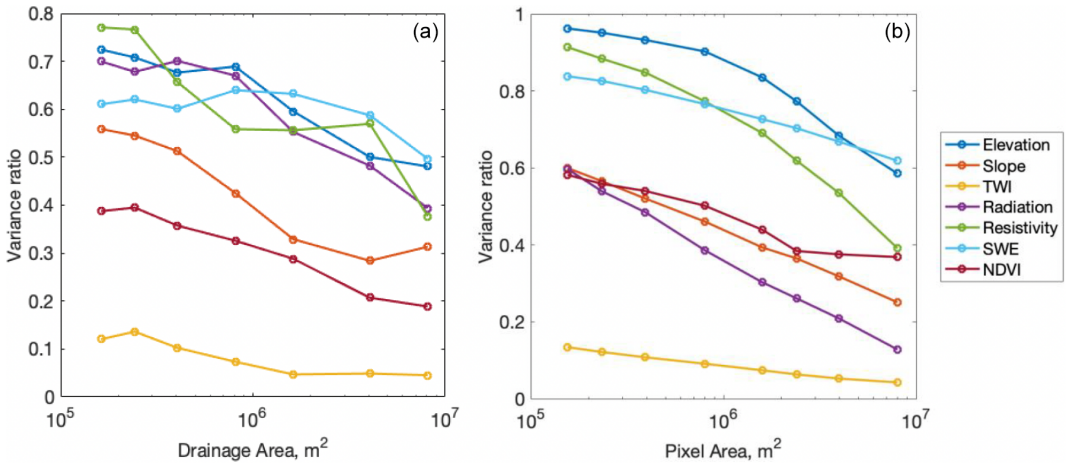

Figure 2. Variance ratio **(a)** between the across-hillslope variance and the overall variance as a function of the threshold drainage area and **(b)** between the across-upscaled-pixel variance and the overall variance as a function of the pixel area.

elevation such that it is the highest in Zone 6, which has the highest elevation and the lower net potential radiation.

With respect to geology (Fig. 6a–c), the low-elevation zones (Zones 1 and 2) are predominantly shale, while the mid- to high-elevation zones (Zones 3 and 5) have both shale and sandstone. Zones 4, 6, and 7 have a high percentage of intrusive rock coverage. Among the conifer-dominated areas (Zones 3, 4, and 7), Zone 4 is associated with intrusive rock, while Zone 3 is associated with shale and sandstone, and Zone 7 has a mixture of these three bedrock types. In terms of land-cover types (Fig. 6d–g), the barren area coverage (Fig. 6d) increases with elevation from Zones 1 to 6. The grassland region (Fig. 6e) exists across all the zones, but it is particularly high in the low-elevation zones (Zones 1 and 2) and the higher-elevation south-facing zone (Zone 5). The coverage of deciduous forests (Fig. 6f) is higher at lower elevation (i.e., Zones 1 and 2); Zone 2, in particular, has high aspen coverage. The evergreen forest regions (Fig. 6g) appear in Zones 3, 4, and 7, which are associated with different bedrock types and slopes as mentioned above. Finally, the wetland (woody and herbaceous combined) coverage is higher in the low-elevation Zone 1. The characteristics of each zone are summarized in Table 1.

4.3 Watershed functions

Foresummer drought sensitivity (Fig. 7a) has an opposite trend to elevation (Fig. 5a), with higher sensitivity in lower-elevation zones (Zones 1 and 2) and the lowest in Zone 6. Zone 5 is an exception such that this zone – high-elevation south-facing grassland-dominated hillslopes – has higher sensitivity than the zones that are in a similar elevation range. The conifer-dominated zones (Zones 3, 4, and 7) have lower-than-average sensitivity. Tukey's pairwise comparison metrics suggest that the differences are significant except in the case of conifer-dominated zones (3, 4, and 7).

We found that the magnitude of annual nitrogen exports scaled log-linearly with discharge, albeit with bifurcation among some of the smaller catchments (Fig. 7c). The Rock and Gothic subcatchments, which are predominantly within conifer-dominated Zones 3, 4, and 7 (Fig. 7b), have lower nitrogen expected from this simple scaling relationship. By contrast, Zone-2-dominated subcatchments (Marmot, Avery) have a larger nitrogen export than would be expected from this relationship. In addition, we evaluated the association of the N mass export and percent coverage of each zone quantitatively. Since the number of points is not large, we divided the N export data into the two groups of the subcatchments that have the coverage of each zone larger than 50 % or less (Fig. 7d and e). The association is statistically significant for Zone 2 and for the conifer-dominated zones combined (Zones 3, 4, and 7), with a p value of less than 0.01. The larger coverage of Zone 2 is associated with a higher N export, while the conifer-dominated zones are associated with a lower export. The individual data points are also shown in Fig. S5.

5 Discussion

We first identified the co-variability between topographic metrics, elevation, bedrock resistivity, and plant signatures at the watershed scale. In particular, the near-surface bedrock electrical resistivity is significantly correlated with slope and elevation. The detailed analysis of AEM by Uhlemann et al. (2022) has found that (1) bedrock resistivity is primarily controlled by bedrock type and that (2) within the shale and sandstone, the bedrock resistivity is affected by the extent of hydrothermal alteration. The relationship between topography, elevation, and bedrock strength has been documented extensively (e.g., Selby, 1987; Clarke and Burbank, 2010). Montgomery (2001) showed that, in low-precipitation and

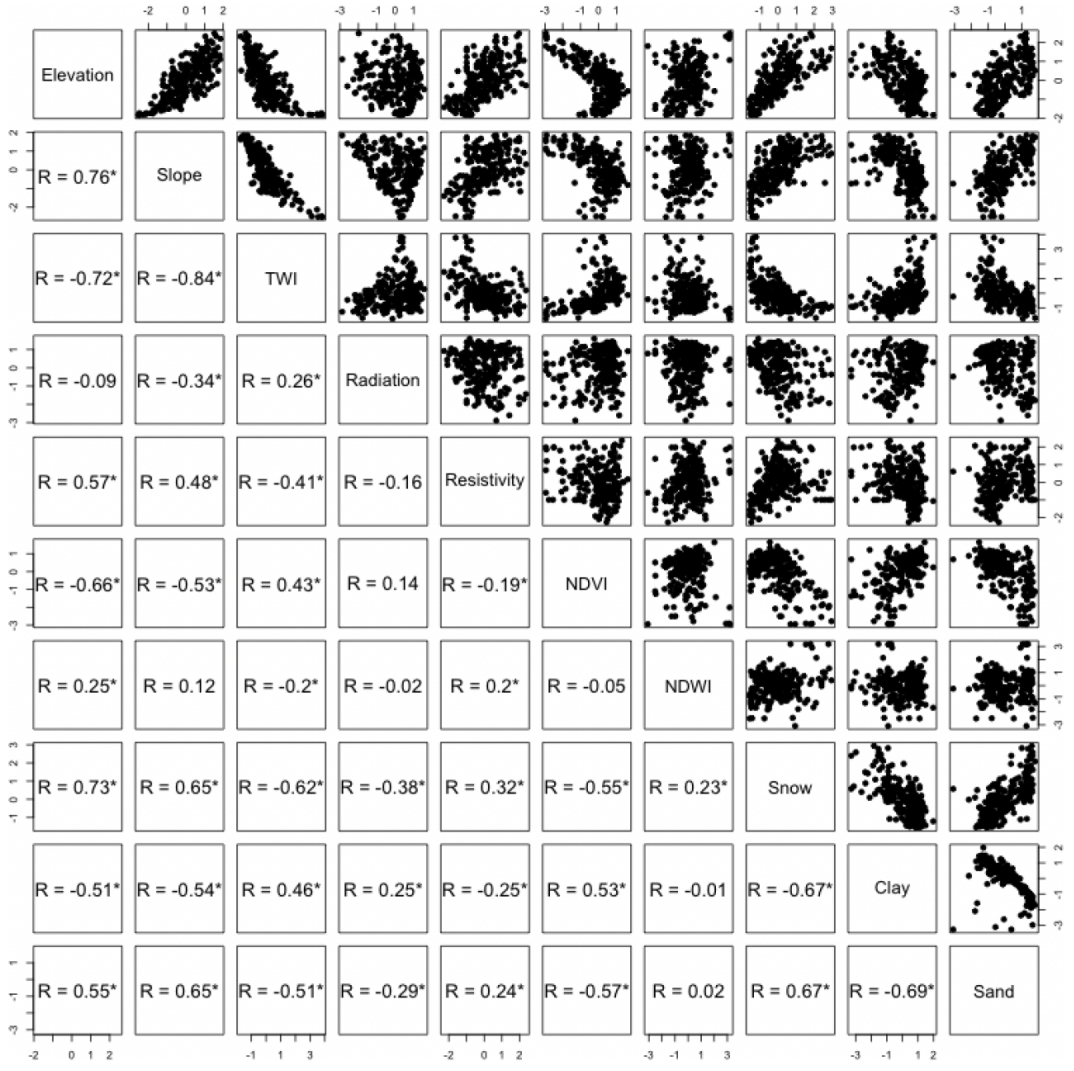

Figure 3. Correlation and scatter plots (Pearson's correlation coefficients) among selected hillslope features (Table S1). The * sign represents p values < 0.01.

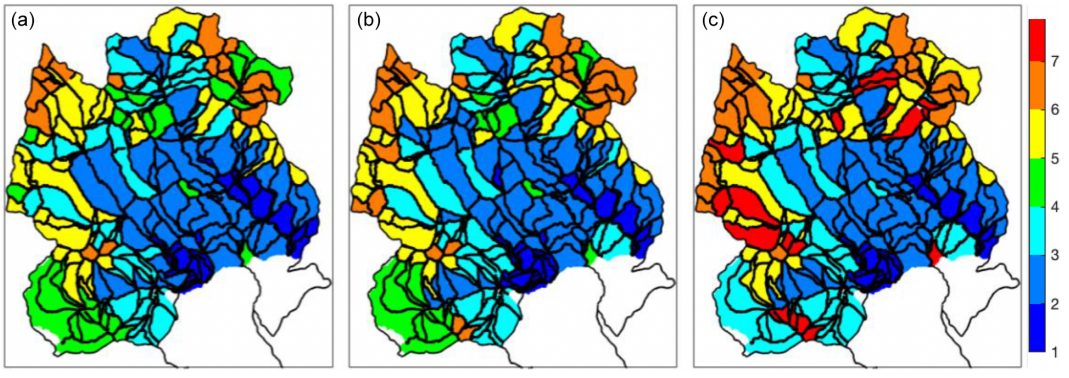

Figure 4. Hillslope zonation based on **(a)** Gaussian mixture model (GMM), **(b)** K-means (KM), and **(c)** hierarchical clustering (HC). Different zones are represented by the cluster index from 1 to 7. There are six common zones among the three methods, and one additional zone (Zone 7) appears in the HC result **(c)**. The black lines are the hillslope boundaries.

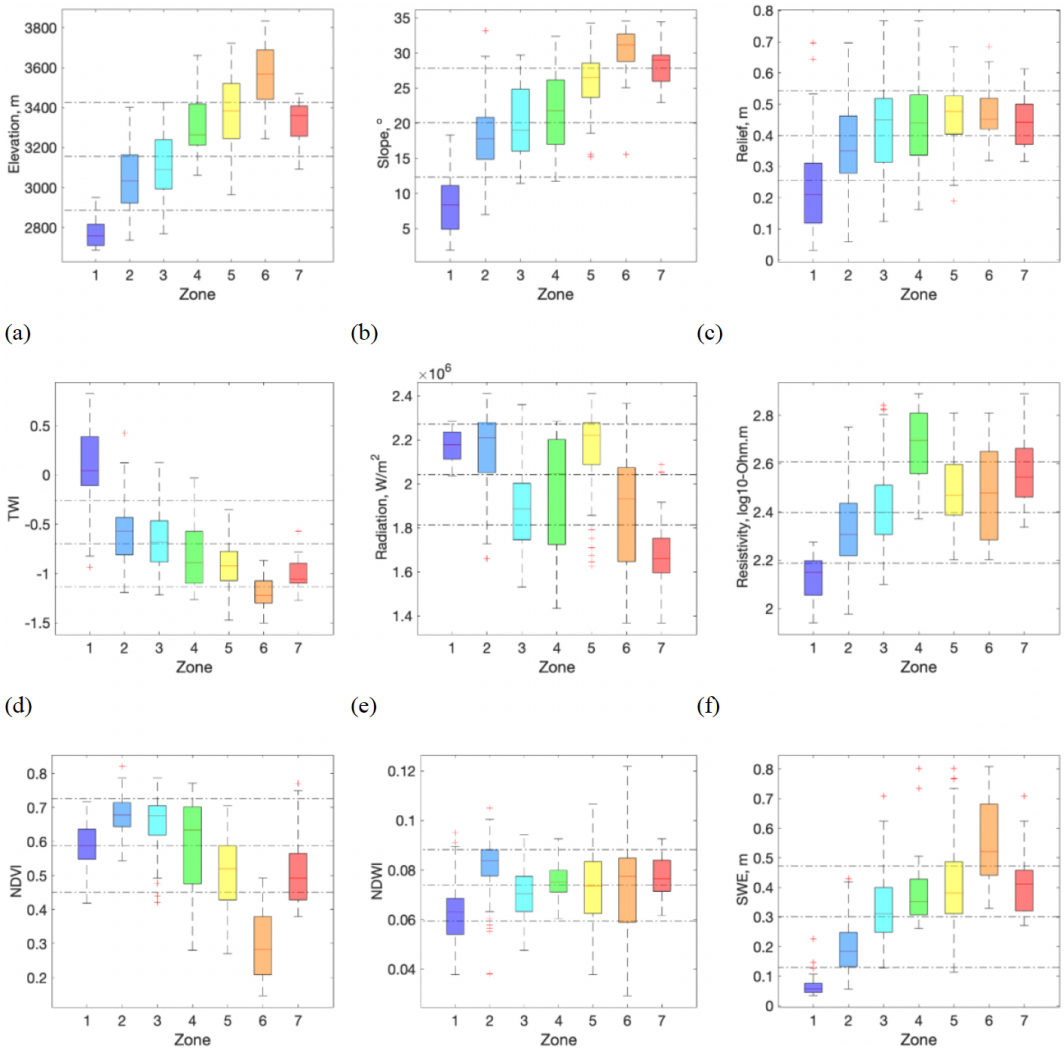

Figure 5. Box plots of hillslope pattern features: **(a)** elevation, **(b)** slope, **(c)** relief, **(d)** TWI, **(e)** annual net potential radiation, **(f)** bedrock electrical resistivity (top 20 m), **(g)** peak NDVI in 2018, **(h)** NDWI (at peak NDVI) in 2018, and **(i)** peak SWE in 2018. The horizontal lines are the overall mean and plus/minus 1 standard deviation.

Table 1. Summary of the features in each zone based on Figs. 5 and 6.

Zone	Elevation	Slope	TWI	Resistivity	Radiation	Peak NDVI	Peak SWE	Plant types	Geology
1	Low	Low	High	Low	Mid–high	Mid	Low	Grassland	Shale
2	Low–mid	Low–mid	Mid–high	Low–mid	Mid–high	Mid–high	Low–mid	Grassland–deciduous	Shale
3	Mid	Mid	Mid–high	Mid	Low	Mid–high	Mid	Evergreen	Shale–sandstone
4	Mid	Mid–high	Low–mid	High	Mid	Mid–high	Mid–high	Evergreen	Intrusive
5	Mid–high	Mid–high	Low–mid	Mid–high	Mid–high	Low–mid	Mid–high	Grassland	Shale–sandstone
6	High	High	Low	High	Mid–low	Low	High	Barren	Shale–intrusive
7	Mid–high	High	Low–mid	Mid–high	Low	Low–mid	Mid–high	Evergreen	Mixed

low soil production environments, such as the East River, the erosion rate is dependent on rock strength and slope steepness, whereby erosion-resistant rocks (i.e., hydrothermally altered sandstones and shales and/or igneous rocks) are on steeper slopes than more erodible rocks. Erosion-resistant rocks also remain at higher elevations, while erodible rock erodes away. This work provides information on how surface topographic features are linked to bedrock variability and could be potentially used to inform bedrock variability,

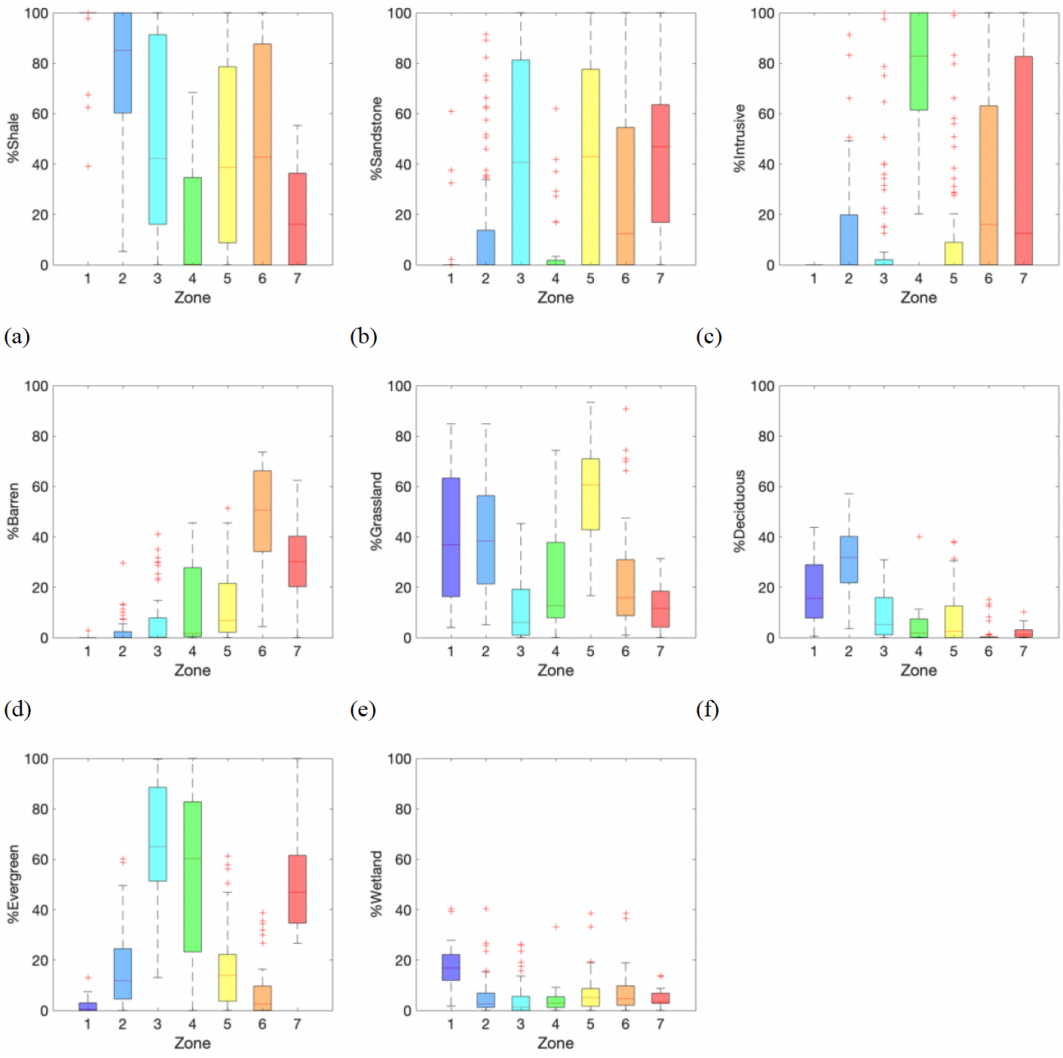

Figure 6. Box plots of hillslope pattern features with respect to geology and land-cover types: the percent coverage of **(a)** shale, **(b)** sandstone, **(c)** igneous (intrusive) rock, **(d)** barren, **(e)** grassland, **(f)** deciduous forest, **(g)** evergreen forest, and **(h)** riparian zone.

which is considered to be one of the largest uncertainties in watershed science (Fan et al., 2019).

In addition, our results show that plant signatures and peak SWE are primarily correlated with elevation. The quadratic dependency of NDVI and above-ground biomass on elevation is consistent with previous studies (e.g., Wainwright et al., 2020); productivity and biomass increase with elevation up to 3300–3500 m because plants have more access to snow-derived water at higher elevation. Above 3500 m, plant productivity is limited by temperature. Net annual potential radiation is not correlated with elevation but is associated with conifer coverage and peak SWE. The conifer forest and high vegetative biomass are associated with north-facing slopes due to higher water availability, which is consistent with previous synthesis studies (e.g., Pelletier et al., 2018). The spatial variability of peak SWE is also consistent with previous studies being associated with elevation and radiation (Anderson et al., 2014). In addition, the conifer-dominated zones are associated with high-resistivity bedrock, while aspen forests are located within the less resistive bedrock. This is consistent with past studies showing that conifer trees grow in the shallow bedrock regions, while aspens require the establishment of deeper root systems (e.g., Burke and Kasahara, 2011). Although there have been studies that have reported the co-variability of plants, topographic metrics, and aspect-controlled bedrock weathering (e.g., Pelletier et al., 2018), our analysis suggested an additional control of geological variability.

To capture such co-variabilities, we hypothesized that hillslopes would be appropriate spatial units to aggregate key properties and to classify into zones. However, the hillslopes could be defined at different scales based on the threshold drainage area, as documented by Band (1989) and Band et al. (1991). We computed the variance in hillslope-averaged

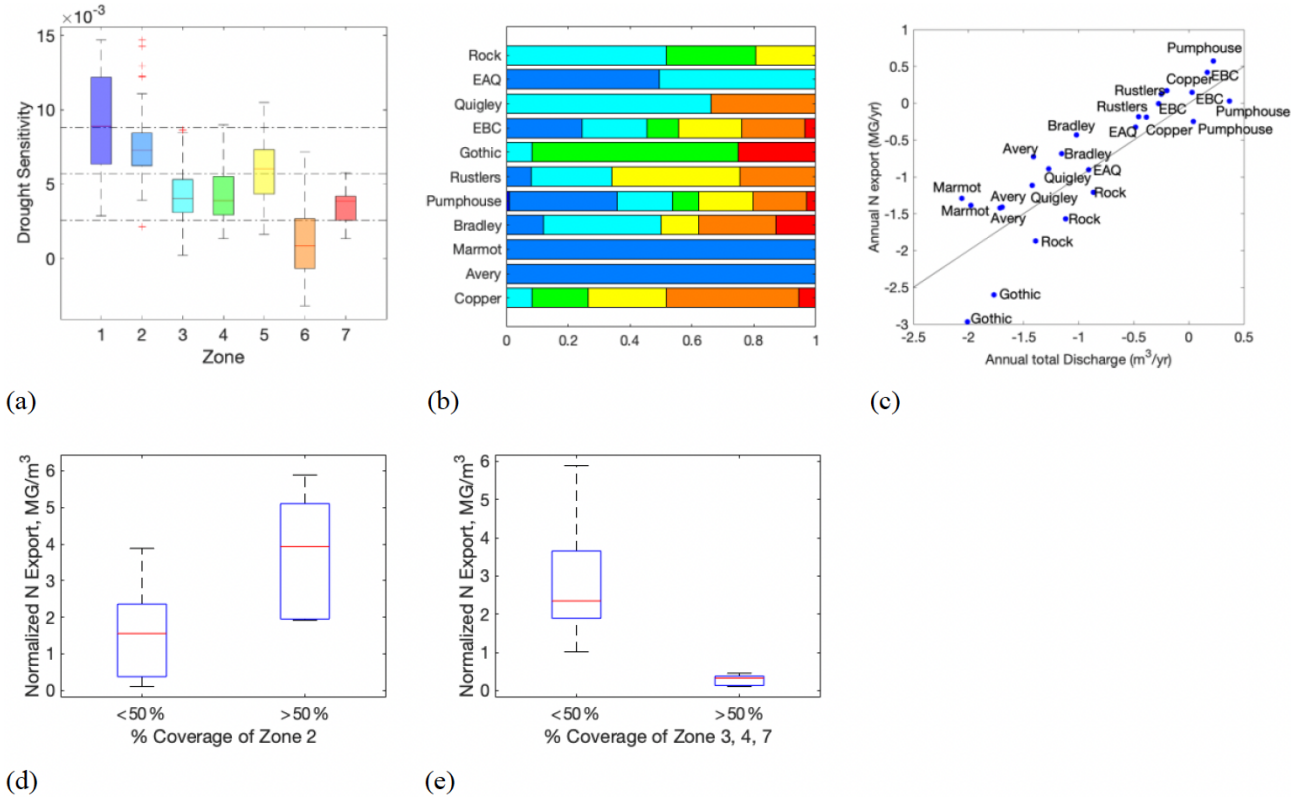

Figure 7. **(a)** Box plot of the foresummer drought sensitivity as a function of zones, **(b)** zonation coverage of each subcatchment, **(c)** annual discharge vs. N export in 2015, 2016, and 2017, and **(d, e)** normalized N export (N export divided by annual discharge) as a function of the percent coverage of **(d)** Zone 2 and **(e)** Zones 3, 4, and 7 as the conifer-dominated zone, respectively. In **(b)**, the color code is the same as Figs. 3–5.

metrics compared to the overall variance as a measure for quantifying how well hillslopes could capture the overall heterogeneity. We then investigated the effect of the threshold drainage area on this variance ratio. Our results show that, although this ratio decreases as the drainage area increases, it stays flat up to a certain threshold drainage area, particularly for elevation and annual net potential radiation. This suggests that there is a representative scale of hillslopes, up to which the within-hillslope variability is limited, and the hillslope-averaged metrics can capture most of the overall variability across the domain. Such a representative scale has been known to exist as a representative elementary area (Wood et al., 1988). This representative scale is important, since Pelletier et al. (2018) showed the significant impact of hillslope aspects and slopes on evapotranspiration, weathering, nutrient cycling, and others.

On the other hand, standard pixel-based upscaling (i.e., taking an average in larger pixels) does not show such a plateau, since larger pixels include a mixture of hillslopes with different aspects. This is consistent with Band et al. (1991), which documented that the hillslope-based partition had much greater efficiency for upscaling than pixel-based scaling by minimizing the within-unit variability and maximizing the ability to capture the watershed-scale variability. Although pixel-based (or grid-based) spatial scaling is still the mainstream for remote sensing and hydrological modeling (e.g., Peng et al., 2017; Foster et al., 2020), our results suggest that the hillslope-based partitioning would be a more effective template for scaling, particularly for mountainous hydrology, where aspect and solar radiation play a significant role (Pelletier et al., 2018; Fan et al., 2019).

After the appropriate hillslope scale is defined, we computed the hillslope features (including the hillslope average of key properties as well as the percent coverage of categorical variables) and applied unsupervised clustering to these features. This clustering process takes advantage of the co-variability between above-ground and below-ground properties and reduces the multidimensional parameter space to one-dimensional classes (Hastie et al., 2001). In the East River watershed, the zones are primarily dictated by elevation, while aspect or radiation poses orthogonal controls to the elevation. Geology and bedrock are also correlated with elevation, such that hillslopes dominated by crystalline rocks and hydrothermally altered shales and sandstones are primarily found at higher elevations. The plant types are associated with both elevation and radiation, such that the dominant

type changes from grassland, aspen, conifer, and grassland to barren from low to high elevations, and conifers are dominant in the north-facing hillslopes.

Instead of relying on a single clustering method, we compared three commonly used clustering methods: hierarchical clustering, *k*-means, and Gaussian mixture models. Although clustering has been applied in many hydrology or environmental science applications, the selection of methods is often arbitrary and subjective (e.g., Sawicz et al., 2011; Wainwright et al., 2015; Devadoss et al., 2020). Such methodological or conceptual model uncertainties are important to characterize, since they are often larger than parameter uncertainty (Neuman, 2003; Ye et al., 2010). Based on the frequency map, we can also identify the hillslopes that are consistently categorized in each zone as the most representative hillslopes in that zone. In addition, our results show that different methods yield similar zones provided that the distance metrics are the same. Differences could be explained as the further division of an identified zone, such that one method divided a zone into two, due to a particular metric. For example, the conifer-dominated hillslopes were divided into Zones 3, 4, and 7, which could provide insights into the impact of different bedrock types and topographic features such as slope.

The identified zones are evaluated with the metrics associated with key watershed functions, including foresummer drought, sensitivity of plant productivity, and river nitrogen export from subcatchments. Drought sensitivity is mainly affected by the dominant plant species in the hillslopes, which is consistent with Wainwright et al. (2020). The conifer-dominated hillslopes have lower sensitivity to the droughts, since they are located mostly in north-facing hillslopes and conifer trees have deeper roots. However, there is a difference in two grassland-dominated zones, Zone 5 and Zone 1, with the high-elevation grassland less sensitive to droughts, possibly owing to increased water availability at higher elevation. In addition, in a separate study by Yan et al. (2021), the soil thickness and associated parameters such as soil diffusion coefficients are found to be distinctly different between the two zones in this domain (Zone 1 and Zone 2).

Nitrogen export is known to be controlled by multiple factors, such as plant retention (Aber et al., 1998), soil cover (Sickman et al., 2002), river corridor features (e.g., floodplains) (Rogers et al., 2021), and geology (Wan et al., 2019; Arora et al., 2020; Maavara et al., 2021). Our analysis showed that the conifer-dominated zone (Zones 3, 4, and 7) is associated with low nitrogen export, while Zone 2, which has the highest fraction of aspen forests, is associated with higher N export. Conifer forests have previously been observed to have high nitrogen retention (Abrahamsen and Stuanes, 1998; Sollins and McCorison, 1981), particularly through ectomycorrhizal uptake (Högberg et al., 2011). By contrast, deciduous aspen forests produce annual leaf litters, with a low C : N content, releasing nitrogen back to the soil (e.g., DeByle et al., 1985; Köchy and Wilson, 1997; Buck and St. Clair, 2012). In addition, Wan et al. (2019) found that shale was a significant geogenic source of nitrogen, based on the intensive observation site located in Zone 2 of this watershed, which is consistent with the global data synthesis (Houlten et al., 2018). We found that the conifer-dominant hillslopes, including those in Zone 3, were in sandstone or igneous rock regions or in more resistive shale with a lower fracture density. These factors may be combined to reduce the nitrogen delivery to the river in the conifer-dominated hillslopes and thus the lower observed river nitrogen exports from that subwatershed. We acknowledge that our approach does not explicitly account for microtopography or the small-scale heterogeneity in the wetland areas, which are considered biogeochemical hotspots because they are small in area but have outsized impacts on N export (Duncan et al., 2013; Rogers et al., 2021). However, we assume that zonation can still capture a large-scale pattern such that Zone 1 has the largest coverage of the wetland region.

6 Conclusion

In this paper, we have developed a watershed zonation approach for characterizing watershed organization and functions based on the bedrock-to-canopy remote sensing (lidar, hyperspectral, and electromagnetic surveys) and spatial data layers (soil, geology, and land-cover maps). We first delineated the domain into a set of hillslopes by computing flow routing based on the lidar-based DEM. To choose an appropriate scale of hillslopes, we defined the ratio between the across-hillslope variance (i.e., the variance of hillslope-averaged properties) and the overall variance and investigated the impact of different threshold drainage areas on the ability of the hillslope-averaged metrics to represent the watershed-scale heterogeneity. We then applied unsupervised clustering to the hillslope features, including the hillslope-averaged metrics of spatial data layers, as well as the percent coverage of categorical variables within each hillslope. Our major findings are that (1) we can define the scale of hillslopes at which the hillslope-averaged metrics can capture the majority of the overall heterogeneity, particularly for elevation and net annual potential radiation, (2) the identified zones (i.e., the groups of hillslopes) have representative characteristics with respect to co-varied bedrock-to-canopy properties, (3) different clustering methods generate similar zonation *patterns* given the same distance criteria, and (4) the identified zones could provide information about zone-based watershed functions, including foresummer drought sensitivity and river nitrogen exports.

Our approach is similar to catchment clustering and classification of past studies (e.g., Wagener et al., 2007; Krause et al., 2014; Kuentz et al., 2017). Those studies, however, defined the classes of watersheds based on streamflow signatures, which were sparse in space, particularly in headwater catchments. In our study, the expansion of spatial data lay-

ers from various remote sensing data layers provides alternative opportunities to apply clustering to the regions without streamflow measurements. In addition, although our approach is similar to the hydrological response unit approach (Flugel, 1997) or hillslope clustering done by Chaney et al. (2018), their main purpose was to parameterize hydrological models. Our analysis confirmed the significance of zonation using the key watershed properties and functions based on observational datasets, including novel datasets such as subsurface structures and signatures from airborne EM.

We recognize that every hillslope is unique, each with distinct topographic positions, geology, and vegetation. However, since hillslope-scale studies and experiments are an integral part of watershed science and critical zone science (Hinckley et al., 2014; Tokunaga et al., 2019; Wan et al., 2019), it is important to identify the similarities and differences among the hillslopes. Our approach provides an objective way of classifying different hillslopes, including the evaluation of hillslope scales and the variability associated with different clustering methods. Using our zonation approach, we can guide the experimental plot placement as well as evaluate the representativeness of the selected hillslopes within the domain of interest. After a particular hillslope is identified, we can use pixel-by-pixel clustering to map the heterogeneity within each hillslope associated with microtopography and hillslope positions based on high-resolution images and lidar (e.g., Park and Van De Giesen, 2004; Falco et al., 2019; Devadoss et al., 2020; Hermes et al., 2020). This hierarchical representation provides a tractable framework for watershed characterization. Furthermore, Yan et al. (2021) developed two separate parameterizations for the soil evolution model in the two hillslopes that belonged to Zones 1 and 2 and then simulated the variable soil thickness within those hillslopes. In this way, our method can improve model parameterization in large-scale hydrological models by honoring distinct boundaries and water/element export contributions and provide a new comprehensive way of linking above- and below-ground properties to watershed functions critical to maintaining water resources.

Code availability. We used Topotoolbox (Schwanghart and Scherler, 2014) for processing spatial data layers. In addition, we used statistical software R (R Core Team, 2013, http://www.R-project, last access: 26 January 2022) for statistical analysis, including the "kmeans" function for KM, the "hclust" function for HC, and the "mclust" function (Scrucca et al., 2016, https://mclust-org.github.io/mclust/, last access: 26 January 2022) for GMM. We used MATLAB R2021a (Mathworks, 2022, http://www.mathworks.com, last access: 20 January 2022) for visualization.

Data availability. All data files associated with calculating nitrogen fluxes (gap-filled nitrogen concentrations and gap-filled discharge) are freely available at https://doi.org/10.15485/1841262 (Wainwright et al., 2022).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/hess-26-429-2022-supplement.

Author contributions. HMW was responsible for conceptualizing the study, undertaking the formal analysis, developing the methodology and codes, and writing and preparing the original draft of the paper. SU conceptualized the study, processed the airborne EM data, and reviewed and edited the paper. MF was responsible for analyzing the streamflow data. NF was responsible for processing the NEON remote sensing data products. NJB conceptualized the study, and reviewed and edited the paper (particularly on the nitrogen cycle). MEN conceptualized the study, analyzed the streamflow data, and reviewed and edited the paper (particularly on nitrogen export). BD conceptualized the study, and reviewed and edited the paper (particularly on geological heterogeneity). ERSW conceptualized the study, and reviewed and edited the paper (particularly on ecohydrology). BJM conceptualized the study, processed the airborne EM data, and reviewed and edited the paper. KHW conceptualized the study, and reviewed and edited the paper. SSH conceptualized the study, developed the methodology, acquired funding, and reviewed and edited the paper.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. This material is based upon work supported as part of the Watershed Function Scientific Focus Area funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-AC02-05CH11231. This work was also supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program and Workforce Development and Education at Lawrence Berkeley National Laboratory. The airborne EM data were acquired with funding from the USGS Mineral Resources Program. We thank Lawrence Band at the University of Virginia, Neal Pastick at USGS, and one anonymous reviewer for careful review and constructive comments.

Financial support. This research has been supported by the U.S. Department of Energy (grant no. DE-AC02-05CH11231). This work was also supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program and Workforce Development and Education at Lawrence Berkeley National Laboratory.

Review statement. This paper was edited by Genevieve Ali and reviewed by Lawrence Band and one anonymous referee.

References

Aber, J. D., McDowell, W., Nadelhoffer, K., Magill, A., Berntson, G., Kamakea, M., McNulty, S., Currie, W., Rustad, L., and Fernandez, I.: Nitrogen saturation in temperate forest ecosystems, BioScience, 48, 921–934, https://doi.org/10.2307/1313296, 1998.

Abrahamsen, G. and Stuanes, A. O.: Retention and leaching of N in Norwegian coniferous forests, Nutr. Cycl. Agroecosyst., 52, 171–178, 1998.

Anderson, B. T., McNamara, J. P., Marshall, H. P., and Flores, A. N.: Insights into the physical processes controlling correlations between snow distribution and terrain properties, Water Resour. Res., 50, 4545–4563, 2014.

Arora, B., Burrus, M., Newcomer, M. E., Steefel, C. I., Carroll, R. W. H., Dwivedi, D., Dong, W., Williams., K. H., and Hubbard, S. S.: Differential C-Q Analysis: A New Approach to Inferring Lateral Transport and Hydrologic Transients Within Multiple Reaches of a Mountainous Headwater Catchment, Frontiers in Water, 2, 24, https://doi.org/10.3389/frwa.2020.00024, 2020.

Asner, G. P., Martin, R. E., Anderson, C. B., and Knapp, D. E.: Quantifying forest canopy traits: Imaging spectroscopy versus field survey, Remote Sens. Environ., 158, 15–27, https://doi.org/10.1016/j.rse.2014.11.011, 2015.

Aytaç, E.: Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based *k*-means clustering, a demonstration on Western Black Sea Region of Turkey, Int. Soil Water Conserv. Res., 8, 321–331, 2020

Ball, L. B., Davis, T. A., Minsley, B. J., Gillespie, J. M., and Landon, M. K.: Probabilistic categorical groundwater salinity mapping from airborne electromagnetic data adjacent to California's Lost Hills and Belridge oil fields, Water Resour. Res., 56, e2019WR026273, https://doi.org/10.1029/2019WR026273, 2020.

Band, L. E.: Spatial aggregation of complex terrain, Geogr. Anal., 21, 279–293, 1989.

Band, L. E., Peterson, D. L., Running, S. W., Coughlan, J., Lammers, R., Dungan, J., and Nemani, R.: Forest ecosystem processes at the watershed scale: basis for distributed simulation, Ecol. Model., 56, 171–196, 1991.

Band, L. E., Patterson, P., Nemani, R., and Running, S. W.: Forest ecosystem processes at the watershed scale: incorporating hillslope hydrology, Agric. Forest Meteorol., 63, 93–126, 1993.

Barfod, A. A. S., Møller, I., Christiansen, A. V., Høyer, A.-S., Hoffimann, J., Straubhaar, J., and Caers, J.: Hydrostratigraphic modeling using multiple-point statistics and airborne transient electromagnetic methods, Hydrol. Earth Syst. Sci., 22, 3351–3373, https://doi.org/10.5194/hess-22-3351-2018, 2018.

Brantley, S. L., McDowell, W. H., Dietrich, W. E., White, T. S., Kumar, P., Anderson, S. P., Chorover, J., Lohse, K. A., Bales, R. C., Richter, D. D., Grant, G., and Gaillardet, J.: Designing a network of critical zone observatories to explore the living skin of the terrestrial Earth, Earth Surf. Dynam., 5, 841–860, https://doi.org/10.5194/esurf-5-841-2017, 2017.

Buck, J. R. and Clair, S. B. S.: Aspen increase soil moisture, nutrients, organic matter and respiration in Rocky Mountain forest communities, PLoS One, 7, e52369, https://doi.org/10.1371/journal.pone.0052369, 2012.

Burke, A. R. and Kasahara, T.: Subsurface lateral flow generation in aspen and conifer-dominated hillslopes of a first order catchment in northern Utah, Hydrol. Process., 25, 1407–1417, 2011.

Burt, T. P. and Pinay, G.: Linking hydrology and biogeochemistry in complex landscapes, Prog. Phys. Geogr., 29, 297–316, 2005.

Carroll, R. W., Bearup, L. A., Brown, W., Dong, W., Bill, M., and Williams, K. H.: Factors Controlling Seasonal Groundwater and Solute Flux from Snow-Dominated Basins, Hydrol. Process., 32, 1–16, https://doi.org/10.1002/hyp.13151, 2018.

Chadwick, K. D., Brodrick, P. G., Grant, K., Goulden, T., Henderson, A., Falco, N., Brodie, E. L., Steltzer, H., Rick Williams, C. F., Blonder, B., Chen, J., Dafflon, D., Damerow, J., Hancher, M., Khurram,A., Lamb, L., Lawrence, C. R., McCormick, M., Musinsky, J., Pierce, S., Polussa, A., Porro, M. H., Scott, A., Singh, H. W., Sorensen, P. O., Varadharajan, C., Whitney, B., and Maher, K.: Integrating airborne remote sensing and field campaigns for ecology and Earth system science, Methods Ecol. Evol., 11, 1492–1508, 2020.

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., and Odgers, N. P.: POLARIS: A 30-meter probabilistic soil series map of the contiguous United States, Geoderma, 274, 54–67, 2016.

Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, Hydrol. Earth Syst. Sci., 22, 3311–3330, https://doi.org/10.5194/hess-22-3311-2018, 2018.

Chaney, N. W., Minasny, B., Herman, J. D., Nauman, T. W., Brungard, C. W., Morgan, C. L., McBratney, A. B., Wood, E. F., and Yimam, Y.: POLARIS soil properties: 30-m probabilistic maps of soil properties over the contiguous United States, Water Resour. Res., 55, 2916–2938, https://doi.org/10.1029/2018WR022797, 2019.

Clarke, B. A. and Burbank, D. W.: Bedrock fracturing, threshold hillslopes, and limits to the magnitude of bedrock landslides, Earth Planet. Sci. Lett., 297, 577–586, 2010.

Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., and Panigada, C.: Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling, Remote Sens. Environ., 112, 1820–1834, 2008.

DeByle, N. V.: Water and watershed, in: Aspen: Ecology and management in the western United States, edited by: DeByle, N. V. and Winokur, R. P., USDA Forest Service General Technical Report RM-119, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 119, 153–160, 1985.

de Pasquale, G., Linde, N., and Greenwood, A.: Joint probabilistic inversion of DC resistivity and seismic refraction data applied to bedrock/regolith interface delineation, J. Appl. Geophys., 170, 103839, https://doi.org/10.1016/j.jappgeo.2019.103839, 2019.

Devadoss, J., Falco, N., Dafflon, B., Wu, Y., Franklin, M., Hermes, A., Hinckley, E. L., and Wainwright, H.: Remote Sensing-Informed Zonation for Understanding Snow, Plant and Soil Moisture Dynamics within a Mountain Ecosystem, Remote Sens., 112, 2733, https://doi.org/10.3390/rs12172733, 2020.

Duncan, J. M., Groffman, P. M., and Band, L. E.: Towards closing the watershed nitrogen budget: Spatial and temporal scaling of denitrification, J. Geophys. Res.-Biogeo., 118, 1105–1119, 2013.

Engstrom, R., Hope, A., Kwon, H., Stow, D., and Zamolodchikov, D.: Spatial distribution of near surface soil moisture and its relationship to microtopography in the Alaskan Arctic coastal plain, Hydrol. Res., 36, 219–234, 2015.

Falco, N., Wainwright, H., Dafflon, B., Léger, E., Peterson, J., Stelzer, H., Wilmer, C., Rowland, J. C., Williams, K. H., and Hubbard, S. S.: Remote sensing and geophysical characterization of a floodplain-hillslope system in the East River watershed, Colorado. Watershed Functionality Scientific Focus Area, 2019.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., van Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope hydrology in global change research and Earth system modeling, Water Resour. Res., 55, 1737–1772, https://doi.org/10.1029/2018WR023903, 2019.

Feilhauer, H., Asner, G. P., and Martin, R. E.: Multimethod ensemble selection of spectral bands related to leaf biochemistry, Remote Sens. Environ., 164, 57–65, https://doi.org/10.1016/j.rse.2015.03.033, 2015.

Flugel, W. A.: Combining GIS with regional hydrological modelling usinghydrological response units (HRUs): An application from Germany, Math. Comput. Simulat., 43, 297e304, https://doi.org/10.1016/S0378-4754(97)00013-X, 1997.

Foster, L. M., Williams, K. H., and Maxwell, R. M.: Resolution matters when modeling climate change in headwaters of the Colorado River, Environ. Res. Lett., 15, 104031, https://doi.org/10.1088/1748-9326/aba77f, 2020.

Gaskill, D. L., Mutschler, F. E., Kramer, J. H., Thomas, J. A., and Zahony, S. G.: USGS Geologic map of the Gothic Quadrangle, Gunnison County, Colorado, https://doi.org/10.3133/gq1689, 1991.

Gillin, C. P., Bailey, S. W., McGuire, K. J., and Gannon, J. P.: Mapping of hydropedologic spatial patterns in a steep headwater catchment, Soil Sci. Soc. Am. J., 79, 440, https://doi.org/10.2136/sssaj2014.05.0189, 2015.

Green, G. N.: The Digital Geologic Map of Colorado in ARC/INFO Format: U.S. Geological Survey Open-File Report 92-0507, 9 pp., available at: http://pubs.usgs.gov/of/1992/ofr-92-0507 (last access: 20 January 2022), 1992.

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., Tague, C., Tonitto, C. and Vidon, P.: Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models, Biogeochemistry, 93, 49–77, 2009.

Hastie, T., Tibshirani, R., and Friedman, J. H.: The elements of statistical learning: data mining, inference, and prediction, Springer, New York, USA, ISBN-13 978-0387848570, 2001.

Hermes, A. L., Wainwright, H. M., Wigmore, O. H., Falco, N., Molotch, N., and Hinckley, E. L. S.: From patch to catchment: A statistical framework to identify and map soil moisture patterns across complex alpine terrain, Front. Water, 2, 48, https://doi.org/10.3389/frwa.2020.578602, 2020.

Hinckley, E.-L. S., Barnes, R. T., Anderson, S. P., Williams, M. W., and Bernasconi, S. M.: Nitrogen retention and transport differ by hillslope aspect at the rain-snow transition of the Colorado Front Range, J. Geophys. Res.-Biogeo., 119, 1281–1296, https://doi.org/10.1002/2013JG002588, 2014.

Högberg, P., Johannisson, C., Yarwood, S., Callesen, I., Näsholm, T., Myrold, D. D., and Högberg, M. N.: Recovery of ectomycorrhiza after "nitrogen saturation" of a conifer forest, New Phytologist, 189, 515–525, 2011.

Homer, C. H., Fry, J. A., and Barnes, C. A.: The national land cover database, US Geological Survey Fact Sheet, 3020, 1–4, 2012.

Houlton, B. Z., Morford, S. L., and Dahlgren, R. A.: Convergent evidence for widespread rock nitrogen sources in Earth's surface environment, Science, 360, 58–62, 2018.

Hubbard, S. S., Gangodagamage, C., Dafflon, B., Wainwright, H., Peterson, J., Gusmeroli, A., Ulrich, C., Wu, Y., Wilson, C., Rowland, J. and Tweedie, C.: Quantifying and relating land-surface and subsurface variability in permafrost environments using LiDAR and surface geophysical datasets, Hydrogeol. J., 21, 149–169, 2013.

Hubbard, S. S., Williams, K. H., Agarwal, D., Banfield, J., Beller, H., Bouskill, N., Dafflon, B., Dwivedi, D., Falco, N., Faybishenko, F., Maxwell, R., Nico, P., Steefel, C., Steltzer, H., Tokunaga, T., Tran, P. A., Wainwright, H., and Varadharajan, C.: The East River, Colorado, Watershed: A mountainous community testbed for improving predictive understanding of multiscale hydrological–biogeochemical dynamics, Vadose Zone J., 17, 1–25, 2018.

Hubbard, S. S., Varadharajan, C., Wu, Y., Wainwright, H., and Dwivedi, D.: Emerging technologies and radical collaboration to advance predictive understanding of watershed hydrobiogeochemistry, Hydrol. Process., 34, 3175–3182, 2020.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: Transferring reach-and plot-scale understanding to the catchment scale, Water Resour. Res., 45, https://doi.org/10.1029/2008WR007225, 2009.

Kassambara, A.: Practical guide to cluster analysis in R: Unsupervised machine learning, vol. 1, Sthda, ISBN-13 978-1542462709, 2017.

Krause, S., Freer, J., Hannah, D., Howden, N., Wagener, T., and Worrall, F.: Catchment similarity concepts for understanding dynamic biogeochemical behavior of river basins, Hydrol. Process., 28, 1554–1560, https://doi.org/10.1002/hyp.10093, 2014.

Köchy, M. and Wilson, S. D.: Litter decomposition and nitrogen dynamics in aspen forest and mixed-grass prairie, Ecology, 78, 732–739, 1997.

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrol. Earth Syst. Sci., 21, 2863–2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

Li, L., Maher, K., Navarre-Sitchler, A., Druhan, J., Meile, C., Lawrence, C., Moore, J., Perdrial, J., Sullivan, P., Thompson, A., and Jin, L.: Expanding the role of reactive transport models in critical zone processes, Earth-Sci. Rev., 165, 280–301, 2017.

Lintern, A., Webb, J. A., Ryu, D., Liu, S., Waters, D., Leahy, P., Bende-Michl, U., and Western, A. W.: What are the key catch-

ment characteristics affecting spatial differences in riverine water quality?, Water Resour. Res., 54, 7252–7272, 2018.

Maavara, T., Siirila-Woodburn, E. R., Maina, F., Maxwell, R. M., Sample, J. E., Chadwick, K. D., Carroll, R., Newcomer, M. E., Dong, W., Williams, K. H., and Steefel, C. I.: Modeling geogenic and atmospheric nitrogen through the East River Watershed, Colorado Rocky Mountains, PLOS ONE, 16, e0247907, https://doi.org/10.1371/journal.pone.0247907, 2021.

Madritch, M. D., Kingdon, C. C., Singh, A., Mock, K. E., Lindroth, R. L., and Townsend, P. A.: Imaging spectroscopy links aspen genotype with below-ground processes at landscape scales, Philos. T. Roy. Soc. B, 369, 20130194, https://doi.org/10.1098/rstb.2013.0194, 2014.

Maina, F. Z. and Siirila-Woodburn, E. R.: Watersheds dynamics following wildfires: Nonlinear feedbacks and implications on hydrologic responses, Hydrol. Process., 34, 33–50, 2020.

Mathworks: MATLAB R2021a, Mathworks [code], available at: http://www.mathworks.com, last access: 20 January 2022.

Maxwell, R. M. and Kollet, S. J.: Quantifying the effects of three-dimensional subsurface heterogeneity on Hortonian runoff processes using a coupled numerical, stochastic approach, Adv. Water Resour., 31, 807–817, https://doi.org/10.1016/j.advwatres.2008.01.020, 2008.

Minsley, B. J. and Ball, L. B.: Airborne geophysical characterization of geologic structure in a mountain headwater system, 7th International Workshop on Airborne Electromagnetics, 17–20 June 2018, Kolding, Denmark, available at: https://pubs.er.usgs.gov/publication/70217684 (last access: 22 January 2022), 2018.

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., and Selker, J.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43, W07301, https://doi.org/10.1029/2006wr005467, 2007.

Mohanty, B. P., Famiglietti, J. S., and Skaggs, T. H.: Evolution of soil moisture spatial structure in a mixed vegetation pixel during the Southern Great Plains 1997 (SGP97) Hydrology Experiment, Water Resour. Res., 36, 3675–3686, https://doi.org/10.1029/2000WR900258, 2000.

Montgomery, D. R.: Slope distributions, threshold hillslopes, and steady-state topography, Am. J. Sci., 301, 432–454, 2001.

Neuman, S. P.: Maximum likelihood Bayesian averaging of uncertain model predictions, Stoch. Env. Res. Risk A., 17, 291–305, 2003.

Newcomer, M. E., Bouskill, N. J., Wainwright, H., Maavara, T., Arora, B., Siirila-Woodburn, E. R., Dwivedi, D., Williams, K. H., Steefel, C., and Hubbard, S. S.: Hysteresis Patterns of Watershed Nitrogen Retention and Loss over the past 50 years in United States Hydrological Basins, Global Biogeochem. Cycles, 35, e2020GB006777, https://doi.org/10.1029/2020GB006777, 2021.

Noël, P., Rousseau, A. N., Paniconi, C., and Nadeau, D. F.: Algorithm for delineating and extracting hillslopes and hillslope width functions from gridded elevation data, J. Hydrol. Eng., 19, 366–374, 2014.

Painter, T.: ASO L4 Lidar Snow Water Equivalent 50m UTM Grid, Version 1, NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, https://doi.org/10.5067/M4TUH28NHL4Z, 2018.

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., and Mattmann, C.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo, Remote Sens. Environ., 184, 139–152, 2016.

Park, S. J. and Van De Giesen, N.: Soil–landscape delineation to define spatial sampling domains for hillslope hydrology, J. Hydrol., 295, 28–46, 2004.

Patton, N. R., Lohse, K. A., Godsey, S. E., Crosby, B. T., and Seyfried, M. S.: Predicting soil thickness on soil mantled hillslopes, Nat. Commun., 9, 1–10, 2018.

Pelletier, J., Barron-Gafford, G. A., Gutiérrez-Jurado, H., Hinckley, E. L. S., Istanbulluoglu, E., McGuire, L. A., Niu, G. Y., Poulos, M. J., Rasmussen, C., Richardson, P., Swetnam, T. L., and Tucker, G. E.: Which way do you lean? Using slope aspect variations to understand Critical Zone processes and feedbacks, Earth Surf. Process. Landf., 43, 1133–1154, https://doi.org/10.1002/esp.4306, 2018.

Peng, J., Loew, A., Merlin, O., and Verhoest, N. E. C.: A review ofspatial downscaling of satellite remotelysensed soil moisture, Rev. Geophys., 55, 341–366, https://doi.org/10.1002/2016RG000543, 2017.

Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E. C., van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T. E., and Woods, R.: Scaling, similarity, and the fourth paradigm for hydrology, Hydrol. Earth Syst. Sci., 21, 3701–3713, https://doi.org/10.5194/hess-21-3701-2017, 2017.

Prancevic, J. and Kirchner, J.: Topographic Controls on the Extension and Retraction of Flowing Streams, Geophys. Res. Lett., 46, 2084–2092, https://doi.org/10.1029/2018gl081799, 2019.

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [code], ISBN 3-900051-07-0, available at: http://www.R-project.org/ (last access: 26 January 2022), 2013.

Rempe, D. M. and Dietritch, W. E.: Direct observations of rock moisture, a hidden component of the hydrologic cycle, P. Natl. Acad. Sci. USA, 115, 2664–2669, 2018.

Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. D. F., and Rodrigues, F. A.: Clustering algorithms: A comparative approach, PloS One, 14, e0210236, https://doi.org/10.1371/journal.pone.0210236, 2019.

Rogers, D. B., Newcomer, M. E., Raberg, J. H., Dwivedi, D., Steefel, C., Bouskill, N., Nico, P., Faybishenko, B., Fox, P., Conrad, M., and Bill, M.: Modeling the Impact of Riparian Hollows on River Corridor Nitrogen Exports, Front. Water, 3, 590314, https://doi.org/10.3389/frwa.2021.590314, 2021.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15, 2895–2911, https://doi.org/10.5194/hess-15-2895-2011, 2011.

Schwanghart, W. and Scherler, D.: Short Communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences, Earth Surf. Dynam., 2, 1–7, https://doi.org/10.5194/esurf-2-1-2014, 2014.

Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E.: mclust 5: clustering, classification and density estimation using Gaussian finite mixture models, The R Journal, 8, 289–

317, https://doi.org/10.32614/RJ-2016-021, 2016 (code available at: https://mclust-org.github.io/mclust/, last access: 26 January 2022).

Selby, M. J.: Rock slopes, in: Slope Stability, edtied by: Anderson, M. G. and Richards, K. S., Chichester, John Wiley & Sons, 475–504, 1987.

Sickman, J. O., Melack, J. M., and Stoddard, J. L.: Regional analysis of inorganic nitrogen yield and retention in high-elevation ecosystems of the Sierra Nevada and Rocky Mountains, in: The Nitrogen Cycle at Regional to Global Scales, Springer, Dordrecht, 341–374, 2002.

Sivapalan, M.: Pattern, process and function: elements of a unified theory of hydrology at the catchment scale, Encyclopedia of hydrological sciences, https://doi.org/10.1002/0470848944.hsa012, 2006.

Sollins, P. and McCorison, F. M.: Nitrogen and carbon solution chemistry of an old growth coniferous forest watershed before and after cutting, Water Resour. Res., 17, 1409–1418, 1981.

St. Clair, J. S., Moon, S., Holbrook, W. S., Perron, J. T., Riebe, C. S., Martel, S. J., Carr, B., Harman, C., Singha, K. D., and Richter, D. D.: Geophysical imaging reveals topographic stress control of bedrock weathering, Science, 350, 534–538, 2015.

Thompson, S. E., Harman, C. J., Troch, P. A., Brooks, P. D., and Sivapalan, M.: Spatial scale dependence of ecohydrologically mediated water balance partitioning: A synthesis framework for catchment ecohydrology, Water Resour. Res., 47, W00J03, https://doi.org/10.1029/2010WR009998, 2011.

Tokunaga, T. K., Wan, J., Williams, K. H., Brown, W., Henderson, A., Kim, Y., Tran, A. P., Conrad, M. E., Bill, M., Carroll, R. W., and Dong, W.: Depth-and time-resolved distributions of snowmelt-driven hillslope subsurface flow and transport and their contributions to surface waters, Water Resour. Res., 55, 9474–9499, 2019.

Uhlemann, S., Dafflon, B., Wainwright, H. M., Williams, K. H., Minsley, B., Zamudio, K., Carr, B., Falco, N., Ulrich, C., and Hubbard, S.: Shale bedrock variability correlates with surface morphology and vegetation distribution, and controls hydraulic properties, Sci. Adv., in review, 2022.

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, Geogr. Compass., 1, 901–931, https://doi.org/10.1111/j.1749-8198.2007.00039.x, 2007.

Wainwright, H. M., Dafflon, B., Smith, L. J., Hahn, M. S., Curtis, J. B., Wu, Y., Ulrich, C., Peterson, J. E., Torn, M. S., and Hubbard, S. S.: Identifying multiscale zonation and assessing the relative importance of polygon geomorphology on carbon fluxes in an Arctic tundra ecosystem, J. Geophys. Res.-Biogeo., 120, 788–808, 2015.

Wainwright, H. M., Steefel, C., Trutner, S. D., Henderson, A. N., Nikolopoulos, E. I., Wilmer, C. F., Chadwick, K. D., Falco, N., Schaettle, K. B., Brown, J. B., Steltzer, H., Williams, K. H., and Enquist, B. J.: Satellite-derived foresummer drought sensitivity of plant productivity in Rocky Mountain headwater catchments: spatial heterogeneity and geological-geomorphological control, Environ. Res. Lett., 15, 084018, https://doi.org/10.1088/1748-9326/ab8fd0, 2020.

Wainwright, H., Uhlemann, S., Franklin, M., Falco, N., Bouskill, N., Newcomer, M., Dafflon, B., Woodburn, E., Minsley, B., Williams, K., and Hubbard, S.: Data used in Wainwright, H. M. et al. 2021, "Watershed zonation through hillslope clustering for tractably quantifying above- and belowground watershed heterogeneity and functions", Watershed Function SFA, ESS-DIVE repository [data set], https://doi.org/10.15485/1841262, last access: 20 January 2022.

Wan, J., Tokunaga, T. K., Williams, K. H., Dong, W., Brown, W., Henderson, A. N., Newman, A. W., and Hubbard, S. S.: Predicting sedimentary bedrock subsurface weathering fronts and weathering rates, Sci. Rep.-UK, 9, 1–10, 2019.

Wood, E. F., Sivapalan, M., Beven, K., and Band, L.: Effects of spatial variability and scale with implications to hydrologic modeling, J. Hydrol., 102, 29–47, 1988.

Yan, Q., Wainwright, H., Dafflon, B., Uhlemann, S., Steefel, C. I., Falco, N., Kwang, J., and Hubbard, S. S.: A hybrid data–model approach to map soil thickness in mountain hillslopes, Earth Surf. Dynam., 9, 1347–1361, https://doi.org/10.5194/esurf-9-1347-2021, 2021.

Ye, M., Meyer, P. D., Lin, Y. F., and Neuman, S. P.: Quantification of model uncertainty in environmental modeling, Stoch. Environ. Res. Risk A., 24, 807–808, 2010.

Zamudio, K. D., Minsley, B. J., and Ball, L. B: Airborne electromagnetic, magnetic, and radiometric survey, upper East River and surrounding watersheds near Crested Butte, Colorado, 2017, U.S. Geological Survey, https://doi.org/10.5066/P949ZCZ8, 2020.