https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access

Characterization of Hillslope Hydrologic Events through a Self-Organizing Map

Eunhyung Lee and Sanghyun Kim

Department of Environmental Engineering, College of Engineering, Pusan National University, Busan, South Korea

Corresponding author: Sanghyun Kim (kimsangh@pusan.ac.kr)

Key Points:

Hydrologic events in a hillslope were analyzed using self-organizing map.

The maximum variation and response time of soil moisture are useful in process-based event identification.

Combinations of hillslope hydrological processes were responsible for five delineated event clusters.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

Abstract

Hydrologic events can be characterized as particular combinations of hydrological processes on a hillslope scale. To configure hydrological mechanisms, we analyzed a dataset using an unsupervised machine learning algorithm to cluster the hydrologic events based on the dissimilarity distances between the weighting components of a self-organizing map (SOM). The time series of soil moisture was measured at 30 points (in 10 locations with 3 varying depths) for 356 rainfall events on a steep, forested hillslope between 2007 and 2016. Soil moisture features for hydrologic events can be effectively represented by the antecedent soil moisture, maximum variation, and standard deviation of peak-to-peak time between rainfall and soil moisture response. Five clusters were delineated for hydrologically meaningful event classification in the SOM representation. The two-dimensional spatial weighting patterns in the SOM provided greater insight on relationships between rainfall characteristics, antecedent wetness, and soil moisture response at different locations and depths. The distinction of the classified events can be explained by several rainfall features and antecedent soil moisture conditions that resulted in different patterns made by combinations of hillslope hydrological processes, vertical flow, and lateral flow along either surface or subsurface boundaries for the upslope and downslope areas.

Keywords: rainfall, soil moisture, hillslope hydrology, self-organizing map, process-based characterization

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

1 Introduction

Soil moisture information is critical for assessing water storage, estimating the quantity of runoff generated, and determining slope stability for hillslopes during rainfall (Tromp Van Meerveld and McDonnel, 2005; Lu and Godt, 2008; Penna et al., 2011; Angermann et al., 2017). Hillslope hydrological processes are affected by many factors, including topography, soil texture, and eco-hydrological parameters (Western et al., 1999; Rodriguez-Iturbe et al., 2006; Liang et al., 2011; Rosenbaum et al., 2012; Baroni et al., 2013), which results in highly nonstationary and heterogeneous spatiotemporal distributions of soil moisture (Wilson et al., 2004; Penna et al., 2009). The relationship between precipitation and runoff is highly nonlinear, and the spatiotemporal variations in soil moisture, groundwater, and surface runoff are extremely difficult to predict (Ali et al., 2013; Curtu et al., 2014).

Rainfall is the primary driver of rapid variations in soil moisture and subsurface flow generation (Penna et al., 2011). Soil moisture response to rainfall events has been investigated for various topographic positions, depth profiles, and land cover conditions (He et al., 2012; Wang et al., 2013; Zhu et al., 2014; Feng and Liu, 2015). The functional relationship between rainfall events and soil moisture, depends on various factors, such as soil texture, depth, topography, and vegetation cover (Liang et al., 2011; Bachmair et al., 2012; Gwak and Kim, 2016). Rainfall characteristics, including the total quantity, duration, intensity, and dry period duration, have also been used to understand the soil moisture response (Heisler-White et al., 2008; Albertson and Kiely, 2001). Other studies of rainfall features have categorized rainfall events to analyze soil moisture variation (Lai et al., 2016; Wang et al., 2008).

Antecedent soil moisture (ASM) plays an essential role in the hydrological response at the hillslope scale (Hardie et al., 2011; Uber et al., 2018; Lee and Kim, 2020). The interaction between

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

the spatial distribution of ASM and rainfall events, determines various hydrological processes, such as the occurrence of preferential flow, soil moisture variation patterns, subsurface stormflow, and runoff generation (Bachmair et al., 2012; Zhang et al., 2011; Saffarpour et al., 2016; Wiekenkamp et al., 2016). The wetter ASM and the greater rainfall events resulted in a greater variation in soil moisture and deeper rainwater percolation (Zhu et al., 2014; Lai et al., 2016; Lee and Kim 2020).

Due to the generation of distinct hillslope flow paths, vertical flows such as matrix, bypass, and lateral flows along different boundaries (e.g., subsurface stormflow over bedrock and surface overland flow) can vary along the transect of the hillslope (Wienhöfer and Zehe, 2014). Previous studies have investigated the functional relationship between rainfall and soil water storage (Castillo et al., 2003; Crow and Ryu, 2009; Tramblay et al., 2012). However, the influence of rainfall features such as rainfall amount, intensity, duration, and ASM conditions on the generation of hillslope flow paths and their distributions at the hillslope scale have not been sufficiently explored. Other studies on hillslope hydrology have focused on several events to identify specific flow paths (e.g., subsurface lateral flow) using intensively collected field measurements over relatively short periods (Freer et al., 2004; Kim 2009; Penna et al., 2011; Wienhöfer and Zehe, 2014).

A comprehensive approach can be useful for addressing the holistic behavior of hydrological processes using a dataset of substantial number of events collected over many years. Identification of specific hydrological processes through visual inspection of field data can be labor-intensive, and the accuracy of analysis can be marginal and subjective if the size of the dataset is large.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

Machine learning techniques have been applied to soil moisture data from in-situ measurements (Ley et al., 2011; van Arkel and Kaleita, 2014; Carranza et al., 2021), remote sensing applications (Ahmad et al., 2010; Srivastava et al., 2013), and the analysis of hydrological model performance (Shrestha et al., 2009; Herbst et al., 2009). Supervised learning algorithms have been used to improve predictions of subsurface flow in a hillslope (Bachmair and Weiler, 2012), downscale satellite soil moisture data (Srivastava et al., 2013), and estimate soil moisture obtained through regression analysis (Ahmad et al., 2010). Critical soil moisture sampling points have also been identified using unsupervised learning algorithms (Van Arkel and Kaleita, 2014; ; Liao et al., 2017). Most studies involving machine learning algorithms for the analysis of soil moisture have focused on estimating and determining the appropriate measurement locations for assessing variations in mean soil moisture. However, soil moisture response can be further explored in the context of hydro-meteorological (rainfall), hydro-historic (ASM), and topographic (location and depth) controllers at the hillslope scale.

A self-organizing map (SOM), which is an unsupervised neural network method, has been used to investigate datasets representing ecosystems, animals, catchment classification, and crop evapotranspiration (Ley et al., 2011; Casper et al., 2012; Ismail et al., 2012; Farsadnia et al., 2014). An SOM can be an effective tool for understanding large hydrologic data by reducing the dimensionality of a dataset, which can provide hydrologic interpretation. Furthermore, and SOM can be used to successfully address the nonlinear relationship between hydrologic variables (di Prinzio et al., 2011; Ley et al., 2011; Toth, 2013; Chen et al., 2018). The highly heterogeneous and extremely nonstationary variation in soil moisture between the upslope and downslope areas alongside the upper, middle, and lower soil layers of a hillslope can be analyzed using an SOM.

In this study, we aimed to answer the following research questions:

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences Discussions

1. How can machine learning algorithms be used to understand the soil moisture response patterns at the hillslope scale?
2. Can delineated clusters of hydrologic events be explained by different hillslope hydrological processes?

In this study, an alternative method for understanding hillslope hydrologic behavior was explored through long-term data analysis using SOM. Hydrologic events for the hillslope scale can be characterized through the rigorous classification of a large hydrologic dataset. Particularly, machine learning algorithms provide several opportunities for understanding hydrologic events by transforming a substantial dataset into compact clusters and delineating the hierarchical relationship between clusters, which can be useful for exploring process-based interpretations and obtaining an efficient monitoring network. We used hydrologic data (rainfall and soil moisture) to analyze and characterize the highly complex relationships between ASM, rainfall characteristics, and soil moisture responses, which include variation in soil moisture and the time to peak. The SOM was employed to investigate the nonlinear interactions between various rainfall characteristics and their effects on temporal changes in soil moisture and classify the multivariate datasets regarding the likely flow paths in the hillslope.

We employed the following approaches to address these research topics: First, we applied an SOM algorithm to datasets composed of rainfall features, ASM, and soil moisture status from upslope to downslope locations in the study area. The dataset was reclassified based on the weighting vectors of each neuron in the SOM map using the Euclidean distances between distinct hydrological variables from individual hydrologic events. Second, the nonlinear relationship between rainfall and soil moisture was evaluated by comparing the spatially weighted patterns of rainfall characteristics and soil wetness variables. The relationships between rainfall characteristics

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

and soil moisture at varying depths and locations were investigated, and these data were used to interpret the hydrological processes.

2 Materials and Methods

2.1 Study Area and Data Acquisition

The hillslope (4000 m^2) selected for the study, is in the Sulmachun watershed (8.5 km^2), which is a headwater of the Imjin River in northwestern South Korea (Figure 1). The study area is primarily covered by a mixture of *Polemoniales*, shrubby *Quercus*, and a coniferous canopy of *Pinus densiflora*, with slopes varying between 30° and 45°. Rainfall, streamflow, and other hydrometeorological records (e.g., temperature and relative humidity) have been collected over the last 25 years from seven hydrologic monitoring stations in this watershed (Figure 1). The mean annual rainfall for the last two decades was approximately 1,500 mm; 70% of the total rainfall occured during the Asian monsoon season between June and August. Precipitation in the form of snowfall occured between December and March. The mean annual evaporation was approximately 420 mm and estimated via eddy-covariance method using data obtained from a flux tower (adjacent hydrologic monitoring station) located 50 m away from the study area. The average daily temperature varied between −15°C and 35°C. The hillslope bedrock consists of granite with extensively weathered areas. Elevations range between 200 and 260 m above sea level, and the surface slope varies between 20° and 35°. Leptosol and Cambisol (classifications from the Food and Agricultural Organization of the United Nations) are the dominant soils in the upslope and downslope areas, respectively. Analysis of 15 soil samples (from 5 points each for the upslope and downslope areas at depths of 30 cm) indicated that the predominant soil textures were sandy-loam and loamy-sand. The average porosities for the upslope and downslope areas were 49% and 48%, respectively. Multiple insertions of an iron pole to each grid cell (0.5 × 0.5 m) indicated that the soil depth along the hillslope varied between 25 and 95 cm. The depth of the root zone was approximately 20–30 cm.

Rainfall (used to describe rainfall characteristics) was recorded at hourly intervals using a rainfall gauge (Automatic Rain Gauge System, Eijkelkamp) placed under the canopy. The soil moisture time series were measured using a multiplex-based time domain reflectometer (TDR; MiniTRASE, SoilMoisture, 2004) at five locations each for upslope (UP1–UP5) and downslope (DO1–DO5) (Figure 1). At each location, three TDR sensors (waveguides) were inserted parallel to the surface at depths of 10, 30, and 60 cm into the upslope side of the installation trench, that filled with soil. Soil moisture measurements were collected hourly between 2007 and 2016. There were 356 rainfall events during the study period. A rainfall event was defined by a minimum dry period of 1 d and at least 1 mm of rainfall.

2.2 Data analysis for soil moisture response

For a given rainfall event, the soil moisture variation at a particular point in the hillslope depends not only on the rainfall but also on other environmental factors such as the location, depth, and soil texture. To consider the relative variation (%) of water storage normalized by the ASM condition, we used the soil moisture difference index, which is the percentage of maximum soil moisture difference (Zhu et al., 2014), to represent the soil moisture variation (VAR):

$$\Delta\theta(\%) = \frac{\theta_{max} - \theta_{ant}}{\theta_{ant}} \cdot 100 \quad (4)$$

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

where θ_{max} is the maximum soil moisture during a rainfall event and the subsequent period ($\leq$ 4 h), and θ_{ant} is the soil moisture measurement before the rainfall event (2 h).

We also calculated the time to peak to peak (P2P in h), which is the time difference between the peak of rainfall and the maximum soil moisture variation. The standard deviation of P2P (SDP2P) for the measuring points was used to represent the homogeneity of the soil moisture responses (Kim, 2009). The time series of the soil moisture was converted to address distinct response features for rainfall events. Depending on soil moisture responses in the transect, location, and depth, 12 different soil moisture response features were delineated as follows: behavior of all measurements (total); measurements at upslope points (upslope); and those for downslope (downslope); measurements for depths of 10 cm (10 cm), 30 cm (30 cm), and 60 cm (60 cm); measurements for upslope at depths of 10 cm (UP10 cm), 30 cm (UP30 cm), and 60 cm (UP60 cm); measurements for downslope at depths of 10 cm (DO10 cm), 30 cm (DO30 cm), and 60 cm (DO60 cm).

2.3 Unsupervised Machine Learning Algorithm

The SOM utilizes an unsupervised learning algorithm that can be useful for pattern recognition of multivariate datasets from different observations. The SOM is typically a two-dimensional (2D) grid composed of either hexagonal or rectangular elements. In this study, we used a hexagonal lattice as the output layer because it resulted in better information propagation when updating more neighborhood neurons than that of the rectangular lattice (Kohonen, 2001). Based on the recommended output dimension of $5\sqrt{r}$ (Kohonen, 2001), where r is the number of events, and the 356 total rainfall events used in this study, the array structure of the SOM was specified as a 16×6 matrix, which corresponded to 96 neurons, namely, the grid cells in the SOM. Each neuron

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
Discussions

had a different weighting vector (w_{ab}), where the subscripts a and b represent the address codes for the variable and node, respectively. A random number was used to initialize the weighting vectors in the neurons. On populating the dataset with rainfall characteristics and soil moisture data, we can obtain the spatial pattern of SOM (Kohonen, 2001).

Input variables for SOM computation were obtained from rainfall features such as rainfall duration (DUR), rainfall amount (AMO), rainfall intensity (INT), ASM, and maximum soil moisture difference index to represent VAR and SDP2P for upslope areas at depths of 10, 30, and 60 cm, and those for the downslope area at depths of 10, 30, and 60 cm, respectively. A min-max transformation was applied to all input variables to fit the bounds of data between zero and one, except SDP2P, which was <1 in most of the data.

SOM maps were established for each variable, and the distance between the input vector and weighting vector can be calculated as follows:

$$d_b = \sqrt{\sum_{a=1}^{v}(w_{a,b} - x_a)^2}, \tag{5}$$

where v is the number of variables.

The best neuron can be identified as the neuron with the minimum value of d_b indicating the best fitness to the characteristics of each rainfall event among every neuron in the SOM. Once the neuron is chosen, the weighting vector should be re-evaluated using Eq. 6 for the renewal weighting vector as follows:

$$\Delta w_{a,b} = \begin{Bmatrix} \alpha(x_a - w_{a,b}) & b = b^* \\ 0 & b \neq b^* \end{Bmatrix}$$

$$w_{a,b}^{new} = w_{a,b}^{old} + \Delta w_{a,b}, \tag{6}$$

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

where α (= 0.5) is the acceleration coefficient, and b^* is the winner neuron. Neurons adjacent to the winner neuron are also updated by applying Eq. (6). The radius of neighboring neurons and acceleration coefficient decreased linearly (from 16 to 1) as the number of iterations increased.

After updating the algorithm, all neurons in the SOM maps fit weighting vectors to the multiple datasets used in this study. The probability density function of each input variable leading to selection of specific SOM nodes can then be inferred from the weighting vector. The input variables in each neuron can be displayed in component planes, which are a spatial pattern in SOM maps. The nonlinear relationship between variables was identified through visual comparison between the spatially distributed weightings in each component plane (Adeloye et al., 2011; Farsadnia et al., 2014; López García and Machón González, 2004; Park et al., 2003).

2.4 Clustering of Hydrologic events

Clusters within the dataset can be delineated by applying the dendrogram classification method and by evaluating the dissimilarity between the weighting vectors (Montero and Vilar, 2014). We used a hierarchical method as the resulting dendrogram structure provided a better representation of the relationships between clusters than the results obtained using non-hierarchical methods. The hierarchical method forms clusters by binding datasets with shorter distances between them. The Euclidean distance function was employed to evaluate the dissimilarity as it is suitable for shape-based comparisons between soil moisture series collected simultaneously (Iglesias and Kastner, 2013). This method has also been used to identify clusters of soil moisture data (Van Arkel and Kaleita, 2014). The Euclidean distance between two weighting vectors in neurons (b_1 and b_2) can be expressed as follows:

$$d_{b_1 b_2} = \left[\sum_{a=1}^{v} (w_{a,b_1} - w_{a,b_2})^2\right]^{0.5}, \quad (7)$$

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

The relationship that has the shortest distance between neurons is assigned to the first cluster, and the weighting vectors of the first cluster can be expressed as:

$$\mu_{c_1,a} = \frac{n_{b_1}\mu_{b_1} + n_{b_2}\mu_{b_2}}{n_{b_1} + n_{b_2}} \tag{8}$$

where μ_{b_1} and μ_{b_2} are the variable weighting vectors in the neurons (b_1 and b_2), respectively; n_{b_1} and n_{b_2} are set to 1 in this relationship, but these values are set to the number of components during the comparison of clusters. Additionally, we used Ward's method to evaluate the dissimilarity between two weighting vectors of each neuron, and between each cluster, namely, this was the chosen algorithm in our hierarchical clustering method (Ward, 1963). When the dissimilarity between two clusters (c_1 and c_2) is calculated, the distance between clusters can be expressed as:

$$d_{cluster} = \sum_{a=1}^{v} \frac{\left\|\mu_{a,c_1} - \mu_{a,c_2}\right\|^2}{\frac{1}{n_{c_1}} + \frac{1}{n_{c_2}}}, \tag{9}$$

where μ_{a,c_1} and μ_{a,c_2} are the averages of clusters c_1 and c_2, respectively, and n_{c_1} and n_{c_2} are the numbers of components for clusters c_1 and c_2, respectively. A dendrogram can be constructed based on the resulting $d_{cluster}$, and the upper part from a designated horizontal line can be recognized as the structure of the final clusters.

3 Results

3.1 Soil moisture responses of all measuring points during rainfall events

The statistics of soil moisture response from 30 points are summarized regarding P2P and maximum variation, as displayed in Fig. 2(a) and 2(b). The P2P ranged from –2 d to + 4 d, indicating that the maximum soil moisture can be reached even before the rainfall peak. SDP2P tends to increase at deeper depths except for locations DO2 and DO5.

While the mean P2P for the upslope area was 0.24 d, the downslope area was 0.02 d. Furthermore, SDP2Ps for upslope and downslope were 0.75 and 0.66 d, respectively. The means of P2P at depths of 10, 30, and 60 cm were -0.08, 0.04, and 0.011 for the downslope and 0.1, 0.24, and 0.38, respectively. The difference in P2P between other points at an identical depth for the downslope was smaller than that for the upslope. This suggests that the soil moisture response in the downslope area is faster and more uniform than the upslope area. The accumulated soil water flow from the upslope area to the downslope area appears responsible for quicker and less spatially variable soil moisture responses in the downslope area. The maximum variation did not display any notable pattern for the transect to the downslope and the depth profile. This is partially because the spatial distribution of antecedent moisture is difficult to characterize because of the temporally varied rainfall event feature and its interaction with the nonlinear soil water process (e.g., hysteresis between soil moisture and soil tension).

3.2 Soil moisture responses feature in measuring locations and depths

The soil moisture response features (e.g., ASM, maximum variation, and SDP2P) were expressed in 12 different spatially averaged responses (Fig. 3) depending on the depth and location. As displayed in Fig. 3(a), the ASM in the downslope area was higher than that in the upslope area. It is apparent that the deeper the depth, the higher the ASM in the downslope area, but those for the upslope area did not display any notable trend in the depth profile. This means that soil water infiltration upslope did not necessarily always occur in all depth profiles.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

The maximum variation in the downslope area was higher than that of the upslope area, as displayed in Fig. 3(b). The mean maximum variation in the downslope area (50.67%) was higher than that of the upslope area (38.73%), and the mean maximum variations at depths of 10, 30, and 60 cm for the upslope area were 44.51%, 34.27%, and 37.39%, while those for the downslope area were 64.49%, 40.83%, and 46.69%, respectively. This indicates higher wetness along both surface and subsurface boundaries, and this trend is pronounced in the downslope direction.

The SDP2Ps for the 12 soil moisture datasets represent the degree of spatial heterogeneity in the temporal soil moisture response. The statistics of the SDP2P (Fig. 3(c)) revealed that the downslope response varied less than that of upslope. While the SDP2P of downslope displayed an apparent increasing trend at deeper depths, those for upslope were similar in-depth profile. The difference in the SDP2P profile between the upslope and downslope indicates that the impact of rainfall on soil moisture response timing can be completely different between upslope and downslope.

The relationships of each response feature (e.g., AMS, VAR, and SDP2P) among different soil moisture datasets can be visualized through the heat map presented in Fig. 4. The heat map consisted of coefficients of determination between different soil moisture datasets representing spatial correlations for different locations, depths, and their combinations. As displayed in Fig. 4, the heat maps for ASM ranged from 0.88 to 0.99, and those for VAR and SDP2P were from 0.78 to 0.98 and from 0.40 to 0.90, respectively. The relationship between upslope and downslope (${}_2C_2$; namely first combination), those between identical depths (${}_3C_2$; namely second combination), and those for different depths and locations (${}_6C_2$; namely third combination) indicate the heterogeneity of different soil moisture features in the spatial context. The first combinations for ASM, VAR, and SDP2P were 0.81, 0.72, and 0.53; the means of second and third combinations were 0.95, 0.84,

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

and 0.62, and 0.83, 0.69, and 0.35 for ASM, VAR, and SDP2P, respectively. This suggests that the spatial distribution of ASM did not demonstrate meaningful spatial variability, but those for VAR and SDP2P were substantial. Namely, VAR and SDP2P can be useful variables to characterize the spatial variation of the soil moisture response for the application of SOM.

3.3 Composition and clustering of SOM

The dataset of hydrologic measurements (356 × 15) was transformed through 96 neurons and output regarding a matrix (16 × 6) through the iterative application of Eqs. (5) and (6), respectively: Namely, 15 hydrologic variables from 356 events were expressed compactly in the SOM. Dissimilarity regarding the Euclidean distance between the output neurons was then used to construct the dendrogram. Many alternatives exist in the number of clusters, depending on the complexity of the dendrogram structure. In this study, five clusters were selected based on a heuristic approach to achieve a hydrologically meaningful classification of events and parsimonious clustering. The relation to notable hydrological processes such as lateral flow or vertical preferential flow and the redundancy check in cluster number were essential factors in the heuristic approach. Figure 5(a) illustrates the resulting dendrogram for the five clusters. The structure of the dendrogram demonstrates the relationships between groups of clusters and between individual clusters. For example, the relationship between clusters 4 and 5 had a lower hierarchy than clusters 1 and 2. Figure 5(b) presents the output SOM (16 × 6) delineated from the dendrogram analysis, which is a structural array identical to the delineated dendrogram with neurons for each cluster. The spatial distributions between other clusters and the corresponding numbers of neurons indicate the areal portion of each cluster from all clusters and their connections with adjacent clusters.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions

Table 1 presents the average of vector components, such as the AMO, DUR, INT, and average ASM among all measuring points (ASMTOT) in volumetric %, alongside an average of the soil moisture difference indices (Δθ) in five upslope locations and five downslope locations at depths of 10, 30, and 60 cm, as VUP10, VUP30, VUP60, VDO10, VDO30, and VDO60. Additionally, it presents the SDP2P in five upslope and five downslope locations at depths of 10, 30, and 60 cm, as SUP10, SUP30, SUP60, SDO10, SDO30, and SDO60, respectively, for the five clusters displayed in Fig. 5(b).

As displayed in Figure 5(b), clusters 1 and 2 were in the upper part of the SOM. Table 1 indicates that the rainfall characteristics of clusters 1 and 2, such as DUR, AMO, and INT, were relatively low, but those for the ASM were similar to the mean ASM for all clusters (Table 1). The average soil moisture difference indices were less than 5% for cluster 1 because the low AMO and intensity resulted in a limited increase in soil water storage, and the loss due to evaporation offset a substantial proportion of the precipitation (Albertson and Kiely, 2001; Ramirez et al., 2007). Cluster 2 had higher AMO and intensities and more significant average soil moisture differences than Cluster 1 (Table 1). The intermediate part of the SOM (Figure 5(b)) is associated with Cluster 3, which revealed higher rainfall durations, quantities, and intensities than those for clusters 1 and 2, which resulted in higher soil moisture difference indices for Cluster 3 than for clusters 1 and 2 (Table 1). One notable feature of cluster 3 was the increasing trend of soil moisture difference indices with depth (DO60 > DO30) for the downslope area, whereas those of clusters 1 and 2 displayed decreased soil moisture difference indices with depth (DO30 > DO60) (Table 1). The pattern of soil moisture difference indices for Cluster 3 can provide evidence for vertical infiltration in all depth profiles for upslope and apparent lateral flow downslope (Table 1 and Figure 4), which appears completely different from those for clusters 1 and 2. Clusters 4 and 5

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

were events demonstrating a larger soil moisture difference index, namely, significant events, in the SOM classification (Table 1). Cluster 4 displayed two distinctive features compared to the other clusters. One is that the ASM of cluster 4 was the lowest among all the clusters. However, the soil moisture difference indices of 30 and 60 cm in the downslope area for Cluster 4 were significantly higher than those in clusters 1, 2, and 3. The other is that the difference in VAR between the upslope and downslope is most pronounced in Cluster 4. This means that the hydrological processes between the upslope and downslope can be substantially distinct from each other. Both rainfall characteristics and soil moisture difference indices for Cluster 5 were significantly higher than those for all other clusters. Many of the measurement points in Cluster 5 were saturated during rainfall events, and the soil moisture at a depth of 60 cm displayed higher variation than that at 30 cm, which indicates that subsurface stormflow was generated along the bedrock in both the upslope and downslope areas.

3.4 Component planes for variables

The component planes of 16 variables and their visual comparisons can provide insight into the nonlinear relationships between the 16 hydrological variables. Figure 6 illustrates the SOM distributions for the component weightings of the 16 variables. Both the spatial distributions and the scales of weightings (scale bar) in Fig. 6 represent the characteristics of impacts (rainfall features and ASM) and consequences (average of soil moisture difference and SDP2P).

The component planes for AMO, DUR, and INT demonstrate higher weightings in the lower right part, as displayed in Figs. 6(a)–6(c). The component plane (Fig. 6(d)) for the ASM displayed lower and higher weightings in the left and right parts. The visual comparison of Figs. 6(a)–6(d) indicates a negligible relationship between rainfall features and ASM. The component

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

planes for upslope soil moisture difference at depths of 10, 30, and 60 cm (Fig. 6(e)–6(g)) displayed similar spatial weightings to those for rainfall features. The high weightings for the soil moisture difference index at 10 cm depth were mainly distributed to clusters 4 and 5, and the weightings tended to concentrate in Cluster 5 at deeper depths (Fig. 5). The comparison between ASM and maximum soil moisture differences indicated that ASM did not influence the VAR index.

The exclusive vertical flow impact can be one possible explanation for the relationship between the component plane for VUP10 and those for VUP30 or VUP60 (Figs. 6(e), 6(f), and 6(g)) because there were negligible upslope contributing areas or small topographic wetness indices (Fig. 1) in upslope locations. The high weightings of 10 cm for the upslope area were distributed in two parts of SOM (the lower left and lower right) (Fig. 6(e)), but those of deeper depths for the upslope area were shifted toward the lower-right part of SOM (Figs. 6(f) and 6(g)). Weightings in VUP10 were associated with AMO and INT, but those for VUP60 correlated only with AMO. This pattern of weighting shift was found between VUP30 and VUP60, which can be attributed to the effect of vertical infiltration (Li et al., 2013). This relationship along the vertical profile is different between the upslope and downslope. The development of the vertical gradient in weightings (Figs. 6(e)–6(g)) from VUP10 to VUP60 can hardly be found in weightings from VDO10 to VDO60 (Figs. 6(h)–6(j)). This means that the flow path in the downslope area cannot be completely explained by the vertical flow.

Figures 6(k)–6(m) displays the component planes of SDP2P at a depth of 10, 30, and 60 cm on the upslope area. The weighting distributions between upslope SDP2P (Figs. 6(k)–6(m)) and ASM (Fig. 6(d)) were completely reverse patterns. The spatial distribution of SDP2P in the downslope did not reveal a notable difference in-depth profile (Figs. 6(n)–6(p)), which can be explained by the time to peak in the downslope appearing not only to be determined by the rainfall driver but is more affected by other drivers such as topography. A wider portion of component planes in the downslope was covered by lower weightings than those for upslope, indicating more uniform and quicker peak time in the downslope than upslope. Relatively uniform distribution in peak time matches wetter ASM and vice versa for dry soil moisture conditions.

4. Discussion

4.1 Characterization of classified hydrologic event

The hydrologic events classified by the SOM can be characterized through comparative feature presentation for all clusters. Lower ASM matched with a higher mean and wider bound in SDP2P, which can also be confirmed by the component planes of ASM and SDP2P. The deeper the depth, the higher the heterogeneity in response time (greater SDP2P) in most locations. This can be explained by the rainfall control to the soil moisture response time decreasing at deeper depths. Depending on the cluster, the SDP2P response between the upslope and downslope can be distinctly expressed. Clusters 1 and 2 exhibited negligible differences in hillslope transects, but those for clusters 3, 4, and 5 were substantially different. This is because the generation of lateral flow can be more significant under larger rainfall events at downslope than those for upslope. The soil moisture peak formations matched well with the maximum variation in soil moisture at the downslope. Events in cluster 1 demonstrated less variation in SDP2P for both depth profile and hillslope transact location due to the lowest AMO and INT. The impact of depth on the variation of SDP2P can be observed in Clusters 2, 3, and 5, and the deeper the depth, the higher the bound, both upslope and downslope. However, this pattern was different between the upslope and downslope in Cluster 4, which had the lowest ASM. The lowest ASM leads to substantially less response variation at 60 cm depth in the upslope, while those for the downslope revealed higher

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

variation at 60 cm depth than those for shallower depths. This means that the dominant flow path between the upslope and downslope was different in cluster 4.

The increasing pattern of the soil moisture difference index corresponds to increasing rainfall features such as DUR and INT from clusters 1 to 5. However, the depth profile of maximum VAR was different between clusters 4 and 5. While the scale of soil moisture recharge demonstrated an apparent decrease in the depth profile for Cluster 4, those for Cluster 5 had different surface and subsurface boundaries (at depths of 10 and 60 cm). This indicates that the dominant hydrological processes for Cluster 4 appear restricted to the surface as the vertical flow, but those for Cluster 5 existed at both the surface and subsurface boundaries regarding both vertical and lateral flows.

The impact of rainfall events on water storage can be useful for understanding the change in various hydrological statuses for each cluster. The storage changes (Table 2) were estimated by multiplying the soil moisture change by the corresponding depth for each waveguide (e.g., 200 mm for 10 and 30 cm depths and 300 mm for 60 cm depth). Water storage analysis for Cluster 1 demonstrated negligible changes under 2% (the measurement accuracy of TDR) in soil moisture that occurred for both the upslope and downslope areas. Rainfall impacts to Cluster 2 can be classified as an intermediate category because both clusters introduced meaningful storage changes (mm) in the downslope area. Significant changes in water storage were found for clusters 3, 4, and 5, regardless of the quantity of rainfall. Substantial increases in storage change at a depth of 60 cm in the downslope area indicated the generation of subsurface stormflow for clusters 3, 4, and 5. The main difference between clusters 4, 3, and 5 was whether the subsurface lateral flow was generated in the upslope area. Clusters 3 and 5 can be characterized as high rainfall and high ASM, which resulted in subsurface lateral flow in both the upslope and downslope areas. The soil

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

moisture changes and storage for cluster 4 indicated an apparent decreasing trend in the depth profile in the upslope area. The storage changes and soil moisture difference indices at depths of 10 and 30 cm in the upslope area for Cluster 4 were greater than those for Cluster 3 due to higher AMO, DUR, and INT. However, the storage change at a depth of 60 cm in the upslope for Cluster 4 was smaller than that of Cluster 3, which can be explained by the lower infiltration under drier ASM conditions (Zhu et al., 2014; Mei et al., 2018; He et al., 2020).

4.2 Configuration of hydrological processes

The application of SOM, an unsupervised machine learning algorithm, to the dataset provided an integrated assessment to evaluate and characterize hydrologic events. The recharge patterns of water storage for the soil layers of the hillslope were characterized by several distinct clusters. The distinct distribution of characteristics of soil moisture responses can be explained by the different combinations of drivers (rainfall and ASM) and hydrological processes (vertical flow, surface, and subsurface lateral flows) for each cluster. The hillslope hydrological flow path was characterized by comparing the component planes between UP10 and UP30 or UP60, and other combinations of soil moisture component planes, such as those of DO10 and DO30 or DO60 regarding SDP2P and VAR.

The rainfall events can be classified into three distinct categories, which depend on the rainfall characteristics, and five further refined clusters as follows: insignificant events for Cluster 1, intermediate events for Cluster 2, and significant events for clusters 3, 4, and 5 (Table 3). Further classification of significant events indicated that the effects of antecedent moisture conditions and AMO were critical for delineating clusters 3, 5, and 4. The generation of hydrological processes based on the significant soil moisture changes over 2% and increasing patterns of SDP2P (0.11 for

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

10 cm, 0.18 for 30 cm, and 0.22 for 60 cm depths) at greater depths was the threshold feature between the insignificant and intermediate events. The primary difference between the intermediate and significant events was the significant response in both the upslope and downslope areas and the substantial development of interface flow between the bedrock and soil layer in the downslope area. This indicates that the lateral flow along boundaries (subsurface and surface) was stronger than that at intermediate depths, and the downslope lateral flow tended to be generated through boundaries either along the surfaces or bedrock. Furthermore, ASM was substantially higher for clusters 3 and 5 than for Cluster 4, and the SDP2D in clusters 3 and 5 were lower for all points than those for Cluster 4. This can be explained by the development of preferential pipe flow, which is more common at greater depths under wetter conditions (Lai et al., 2016; Uber et al., 2018; Uchida et al., 2001; Wienhöfer and Zehe, 2014). Low variation and soil moisture changes in UP60 for Cluster 4 indicated that low antecedent moisture conditions limit the generation of lateral flow into the upslope area, and that of Cluster 3 can be explained by even the fewer rainfall events in Cluster 3 than those for Cluster 4 being sufficient to activate subsurface lateral flow in the upslope. Extreme rainfall events were mainly associated with Cluster 5. Lateral storm flow likely occurred in both the upslope and downslope areas of Cluster 5. Effective drainage during extreme events appears to be strongly associated with lateral flow generation along the two boundaries in the soil media (i.e., surface and bedrock) (Uchida et al., 2001; Freer et al., 2004; Haga et al., 2005; Kim, 2009; Wienhöfer and Zehe., 2014; Angermann et al., 2017). The impact of extreme rainfall conditions dominates other controls (e.g., land cover and topography) regarding hillslope runoff generation (Feng and Liu, 2015).

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

5. Conclusion

Rainfall characteristics and responses of soil moisture at the hillslope scale were explored by applying SOM to a dataset with a large number of hydrologic events. Hydrologic events were characterized for rainfall and soil moisture data collected over ten years from a steep hillside. Based on a delineated dendrogram, the classification of neurons into five clusters provided meaningful interpretations to understand hydrologic events.

The nonlinear relations between hydrologic variables were effectively expressed in the 2D SOM presentations of variables. The apparent relationship between ASM and peak time variation indicates that the hydrologic response is more feasible under wetter conditions. Water storage analysis for each event from different clusters suggests that spatially different combinations of VAR can be attributed to the identified hydrologic response for each cluster. Combinations of upslope and downslope spatial patterns of hillslope hydrological processes, vertical flow, and lateral flow along surface or subsurface boundaries were responsible for the distinctions between the event clusters. Depending on rainfall and ASM conditions delineated from each cluster, the spatial distribution of hydrological processes can be predicted to be useful for obtaining systematic insight into the hillslope hydrological response. The meta-heuristic classification of hydrologic events provides intuition for hydrologic conditions and their drivers, which is vital for designing a process-based hillslope hydrology model.

Code and Data Availability

Code will be available through repository https://www.re3data.org/ when paper is accepted. Data is uploaded as supplementary materials.

Author contribution

Enhyung Lee and Sanghyun Kim and several former graduate students had collected data for the study area. Eunhyung Lee developed model code and performed simulation. Sanghyun Kim prepared manuscript with contribution from Ennhyung Lee.

Competing interests

The authors declare that they have no conflict of interest

Acknowledgments

This study was financially supported by the Basic Research Program (2016R1D1A1B02008137) of the National Science Foundation of the Republic of Korea.

References

Adeloye, A. J., Rustum, R., and Kariyama, I.D.: Kohonen self-organizing map estimator for the reference crop evapotranspiration, Water Resour. Res., 47(8), https://doi.org/10.1029/2011WR010690, 2011.

Ahmad, S., Kalra, A., and Stephen, H.: Estimation soil moisture using remote sensing data: A machine learning approach, Adv. Water Resour., 33(1), 69-80, https://doi.org/10.1016/j.advwatres.2009.10.008, 2010.

Albertson, J. D., and Kiely, G.: On the structure of soil moisture time series in the context of land surface models, J. Hydrol., 243, 101-119, https://doi.org/10.1016/S0022-1694(00)00405-4, 2001.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences Discussions

Ali, M., Fiori, A., and Bellotti, G.: Analysis of the nonlinear storage-discharge relation for hillslopes through 2D numerical modelling, Hydrol. Process., 27, 2683-2690, https://doi.org/10.1002/hyp.9397, 2013.

Angermann, L., Jackish, C., Allroggen, N., Sprenger, M., Zehe, E., Tronicke, J., Weiler, M., and Blume, T.: Form and function in hillslope hydrology: characterization of subsurface flow based on response observations, Hydrol. Earth Syst. Sci., 21, 3727-3748, https://doi.org/10.5194/hess-21-3727-2017, 2017.

Bachmair, S., Weiler, M., and Troch, P. A.: Intercomparing hillslope hydrological dynamics: Spatio-temporal variability and vegetation cover effects, Water Resour. Res., 48, W05537, https://doi.org/ 10.1029/2011WR011196, 2012.

Baroni, G., Ortuani, B., Facchi, A., and Gandolfi, C.: The role of vegetation and soil properties on the spatio-temporal variability of the surface soil moisture in a maize-cropped field, J. Hydrol., 489, 148-159, https://doi.org/10.1016/j.jhydrol.2013.03.007, 2013.

Carranza, C., Nolet, C., Pejij, M., Ploeg, and van der Ploeg, M.: Root zone soil moisture estimation with random forest, J. Hydrol., 593, 125840, https://doi.org/10.1016/j.jhydrol.2020.125840, 2021.

Casper, M.C., Grigoryan, G., Gronz, O., Gutjahr, O., Heinemann, G., Ley R., and Rock, A.: Analysis of projected hydrological behavior of catchments based on signature indices, Hydrol. Earth Syst. Sci., 16, 409-421, https://doi.org/10.5194/hess-16-409-2012, 2012.

Castillo, V.M., Gomez-Plaza, A., and Martinez-Mena, M.: The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach, J. Hydrology, 284(1-4), 114-130, https://doi.org/10.1016/S0022-1694(03)00264-6, 2003.

Chen, I.T., Chang, L.C., and Chang, F.J.: Exploring the sptio-temporal interrelation between groundwater and surface water by using the self-organizing maps, J. Hydrol., 556, 131-142, https://doi.org/10.1016/j.jhydrol.2017.10.015, 2018.

Crow, W.T., and Ryu, D.: A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals, Hydrol. Earth Syst. Sci., 13, 1-16, https://doi.org/10.5194/hess-13-1-2009, 2009.

Curtu, R., Mantilla, R., Fonley, M., Cunha, L.K., Small, S.J., Jay, L.O., and Krajewski, W.F.: An integral-balance nonlinear model to simulate changes in soil moisture, groundwater and surface runoff dynamics at the hillslope scale, Adv. Water Resour., 71, 125-139, https://doi.org/10.1016/j.advwatres.2014.06.003, 2014.

di Prinzio, M., Cstellarin, A., and Toth, E.: Data-driven catchment classification: application to the pub problem, Hydrol. Earth Syst. Sci., 15, 1921-1935, https://doi.org/10.5194/hess-15-1921-2011, 2011.

Farsadnia, F., Kamrood, M. R., Nia, A. M., Modarres, R., Bray, M. T., Han, D., and Sadatinejad, J.: Identification of homogeneous regions for regionalization of watersheds by two-level self-organizing feature maps, J. Hydrol., 509, 387-397, https://doi.org/10.1016/j.jhydrol.2013.11.050, 2014.

Gwak, Y., Kim, S., Factors affecting soil moisture spatial variability for a humid forest hillslope, Hydrolo. Process. 31(2), 431-445, https://doi.org/10.1002/hyp.11039, 2016

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Feng, H., and Liu, Y.: Combined effects of precipitation and air temperature on soil moisture in different land covers in a humid basin, J. Hydrol., 531, 1129-1140, https://doi.org/10.1016/j.jhydrol.2015.11.016, 2015.

Freer, J., McDonnell, J., Beven, K., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., and Kendall, C.: The role of bedrock topography on subsurface storm flow, Water Resour. Res., 38(12), W1269, https://doi.org/ 10.1029/2001WR000872, 2004.

Haga, H., Matsumoto, Y., Matsutani, J., Fujita, M., Nishida, K., and Sakamoto, Y.: Flow paths, rainfall properties, and antecedent soil moisture controlling lags to peak discharge in a granitic unchanneled catchment, Water Resour. Res., 41, W12410, https://doi.org/10.1029/2005WR004236, 2005.

Hardie, M.A., Cotching, W.E., Doyle, R.B., Holz, G., Lisson, S., and Mattern, K.: Effect of antecedent soil moisture on preferential flow in a texture-contrast soil, J. Hydrol., 398(3-4), 191-201, https://doi.org/10.1016/j.jhydrol.2010.12.008, 2011.

He, Z., Jia, G., Liu, Z., Zhang, Z., Yu, X, and Xiao, P.:Field studies on the influence of rainfall intensity, vegetation cover and slope length on soil moisture infiltration on typical watersheds of the Loess Plateau, China, Hydrol. Process., 34, 4904-4919, https://doi.org/10.1002/hyp.13892, 2020.

He, Z., Zhao, W., Liu, H., and Chang, X.: The response of soil moisture to rainfall event size in subalpine grassland and meadows in a semi-arid mountain range: a case study in northwestern China's Qilian Mountains, J. Hydrol., 420-421, 183-190, https://doi.org/10.1016/j.jhydrol.2011.11.056, 2012.

Heisler-White, J. L., Knapp, A. K., and Kelly, E. F.: Increasing precipitation event size increases aboveground net primary productivity in a semi-arid grassland, Oecologia, 158, 129-140, https://doi.org/10.1007/s00442-008-1116-9, 2008.

Herbst, M., Gupta, H.V., and Casper, M.C.: Mapping model behavior using self-organizing maps, Hydrol. Earth Syst. Sci., 13, 395-409, https://doi.org/10.5194/hess-13-395-2009, 2009.

Iglesias, F., and Kastner, W.: Analysis of similarity measures in times series clustering for the discovery of building energy patterns, Energies, 6, 579-597, https://doi.org/10.3390/en6020579, 2013.

Ismail, S., Shabri, A., and Samsudin, R.: A hybrid model of self organizing maps and least square support vector machine for river flow forecasting, Hydrol. Earth Syst. Sci., 16, 4417-443, https://doi.org/10.5194/hess-16-4417-2012, 2012.

Kim, S.: Characterization of soil moisture responses on a hillslope to sequential rainfall events during late autumn and spring, Water Resour. Res. 45, W09425, https://doi.org/ 10.1029/2008WR007239, 2009.

Kohonen, T.: Self-Organizing Maps, third ed., Springer, Berlin, 2001.

Lai, X., Liao, K., Feng, H., and Zhu, Q.: Responses of soil water percolation to dynamic interactions among rainfall, antecedent moisture and season in forest site, J. Hydrol., 540, 565-573, https://doi.org/10.1016/j.jhydrol.2016.06.038, 2016.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Lee, E. and Kim, S.: Pattern similarity based soil moisture analysis for three seasons on a steep hillslope, J. Hydrol., 551, 484-494, https://doi.org/10.1016/j.jhydrol.2017.06.028, 2017.

Lee E., and Kim, S.: Characterization of runoff generation in a mountainous hillslope according to multiple threshold behavior and hysteretic loop features, J. Hydrol., 590, 125534, https://doi.org/10.1016/j.jhydrol.2020.125534, 2020.

Ley, R., Casper, M. C., Hellebrand, H., and Merz, R.: Catchment classification by runoff behavior with self-organizing maps (SOM), Hydrol. Earth Syst. Sci., 15, 2947-2962, https://doi.org/10.5194/hess-15-2947-2011, 2011.

Li, X. Y., Zhang, S. Y., Peng, H. Y., Hu, X., and Ma, Y. J.: Soil water and temperature dynamics in shrub-encroached grasslands and climatic implications: Results from inner Mongolia steppe ecosystem of north China, Agric.For.Meteorol., 171, 20-30, https://doi.org/10.1016/j.agriformet.2012.11.001, 2013.

Liang, W. L., Kosugi, K., and Mizuyama, T.: Soil water dynamics around a tree on a hillslope with or without rainwater supplied by stemflow, Water Resour. Res., 47, W02541, https://doi.org/10.1029/2010WR009856, 2011.

Liao, K., Zhou, Z., Lai, X., Zhu, Q., and Feng, H.: Evaluation of different approaches for identifying optimal sites to predict mean hillslope soil moisture content, J. Hydrol., 547, 10-20, https://doi.org/10.1016/j.jhydrol.2017.01.043, 2017.

López García, H., and Machón González, I.: Self-organizing map and clustering for wastewater treatment monitoring, Engineering Application of Artificial Intelligence, 17, 215-225, https://doi.org/10.1016/j.engappai.2004.03.004, 2004.

Lu, N., and Godt, J.: Infinite slope stability under steady unsaturated seepage conditions, Water Resour. Res., W11404, https://doi.org/10.1029/2008WR006976, 2008.

Mei, X., Zhu, Q., Ma, L., Zhang, D., Wang, Y., and Hao, W.: Effect of stand origin and slope position on infiltration pattern and preferential flow on a Loess hillslope, Land Degrad. Dev.., 29, 1353-1365, https://doi.org/10.1002/ldr.2928, 2018.

Montero, P., and Vilar, J.A.: TSclust: An R package for time series clustering, J. Stat. Softw., 62(1), 1-43, https://doi.org/10.18637/jss.v062.i01,2014

Park, Y. S., Cereghino, R., Compin, A., and Lek, S.: Applications of artificial neural networks for pattering and predicting aquatic insect species richness in running waters, Ecol Model., 160, 265-280, https://doi.org/10.1016/S0304-3800(02)00258-2, 2003.

Penna, D., Borga, M., Norbiato, D., and Fontana, G. D.: Hillslope scale soil moisture variability in a steep alpine terrain, J. Hydrol., 364, 311-327, https://doi.org/10.1016/j.jhydrol.2008.11.009, 2009.

Penna, D., Tromp van Meerveld, H. J., Gobbi, A., Borga, M., and Fontana, G. D.: The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment, Hydrol. Earth Syst. Sci., 15, 689-702, https://doi.org/10.5194/hess-15-689-2011, 2011.

Ramirez, D. A., Bellot, J., Domingo, F., and Blasco, A.: Can water responses in stipa tenacissima L. during the summer season be promoted by non-rainfall water gains in soil?, Plant and Soil, 291, 67-79, https://doi.org/10.1007/s11104-006-9175-3, 2007.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Rodriguez-Iturbe, I., Isham, V., Cox, D. R., Manfreda, S., and Porporato, A.: Space-time modeling of soil moisture: Stochastic rainfall forcing with heterogeneous vegetations, Water Resour. Res., 42, W06D05, https://doi.org/10.1029/2005WR004497, 2006.

Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J. A., Peterson, T. J., Weuthen, A., Western, A. W., and Vereecken, H.: Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale, Water Resour. Res., 48, W10544, https://doi.org/10.1029/2011WR011518, 2012.

Saffarpour, S., Western, A.W., Adams, R., and McDonnell, J.J.: Multiple runoff processes and multiple thresholds control agricultural runoff generation, Hydrol. Earth Syst. Sci., 20, 4525-4545, https://doi.org/10.5194/hess-20-4525-2016, 2016.

Shrestha, D.L., Kayastha, N., and Solomatine, D.P.: A novel approach to parameter uncertainty analysis of hydrological models using neural networks, Hydrol. Earth Syst. Sci., 13, 1235-1248, https://doi.org/10.5194/hess-13-1235-2009, 2009.

Srivastava, P.K., Han, D., Ramirez, M.R., and Islam, T.: Machine learning techniques for downscaling smos satellite soil moisture using modis land surface temperature for hydrological application, Water Resour. Manag., 27, 3127-3144, https://doi.org/10.1007/s11269-013-0337-9, 2013.

Toth, E.: Catchment classification based on characterization of streamflow and precipitation time series, Hydrol. Earth Syst. Sci., 17, 1149-1159, https://doi.org/10.5194/hess-17-1149-2013, 2013.

Tromp van Meerveld, I., and McDonnell, J.J.: Comment to “Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, J.Hydrol., 286, 113-134”, J.Hydrol., 303, 307-312, https://doi.org/10.1016/j.jhydrol.2004.09.002, 2005.

Tramblay, Y., Bouaicha, R., Brocca, L., Dorigo, W., Bouvier, C., Camici, S., and Servat, E.: Estimation of antecedent wetness conditions for flood modelling in northern morocco, Hydrol. Earth Syst. Sci., 16, 4375-4386, https://doi.org/10.5194/hess-16-4375-2012, 2012.

Uber, M., Vandervaere, J.P., Zin, I., Braud, I., Heistermann, M., Legout, C., Molinie, G., and Nord, G.: How does initial soil moisture influence the hydrological response? A case study from southern france, Hydrol. Earth Syst. Sci., 22, 6127-6146, https://doi.org/10.5194/hess-22-6127-2018, 2018.

Uchida, T., Kosugi, K., and Mizuyama, T.: Effects of pipeflow on hydrological process and its relations to landslide, a review of pipeflow studies in forested headwater catchments, Hydrol. Process. 15, 2151-2174, https://doi.org/10.1002/hyp.281, 2001.

Van Arkel, Z., and Kaleita, A. L.: Identifying sampling locations for field-scale soil moisture estimation using K-means clustering. Water Resour. Res. 50 (8), 7050-7057, https:/doi.org/10.1002/2013WR015015, 2014.

Wang, S., Fu, B., Gao, G., Liu, Y., and Zhou, J.: Responses of soil moisture in different land cover types to rainfall events in a re-vegetation catchment area of the Loess Plateau, China, Catena, 101,122-128, https://doi.org/10.1016/j,catena.2012.10.006, 2013.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Wang, X. P., Cui, Y., Pan, Y. X., Li, X. R., Yu, Z., and Young, M. H.: Effects of rainfall characteristics on infiltration and redistribution patterns in revegetation-stabilized desert ecosystems, J. Hydrol., 358, 134-143, https://doi.org/10.1016/j,jhydrol.2008.06.002, 2008.

Western, A. W., Grayson, R. B., Blöschl, G. Willgoose, G. R., McMahon, T. A., Observed spatial organization of soil moisture and its relation to terrain indices, Water Resour. Res., 35(3), 797-8110, https://doi.org/10.1029/1998WR900065, 1999.

Wiekenkamp, I., Huisman, J.A., Bogena, H.R., Lin, H.S., and Vereecken, H.: Spatial and temporal occurrence of preferential flow in a forested headwater catchment, J. Hydrol., 534, 139-149, https://doi.org/10.1016/j.jhydrol.2015.12.050, 2016.

Wienhöfer, J., and Zehe, E.: Predicting subsurface stormflow response of a forested hillslope: the role of connected flow paths, Hydrol. Earth Syst. Sci., 18, 121-138, http://doi.org/10.5194/hess-18-121-2014, 2014.

Wilson, D. J., Western, A. W., and Grayson, R. B.: Identifying and quantifying sources of variability in temporal and spatial soil moisture observations, Water Resour. Res., 40, W02507, https://doi.org/10.1029/2003WR002306, 2004.

Zhang, Y., Wei, H., and Nearing M.A.: Effects of antecedent soil moisture on runoff medeling in small semiarid watersheds of southeastern Arizona, Hydrol. Earth Syst. Sci., 15, 3171-3179, https://doi.org/10.5194/hess-3171-2011, 2011.

Zhu, Q., Nie, X. F., Zhou, X. B., Liao, K. H., and Li, H. P.: Soil moisture response to rainfall at different topographic positions along a mixed land-use hillslope, Catena, 119, 61-70, https://doi.org/10.1016/j.catena.2014.03.010, 2014.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

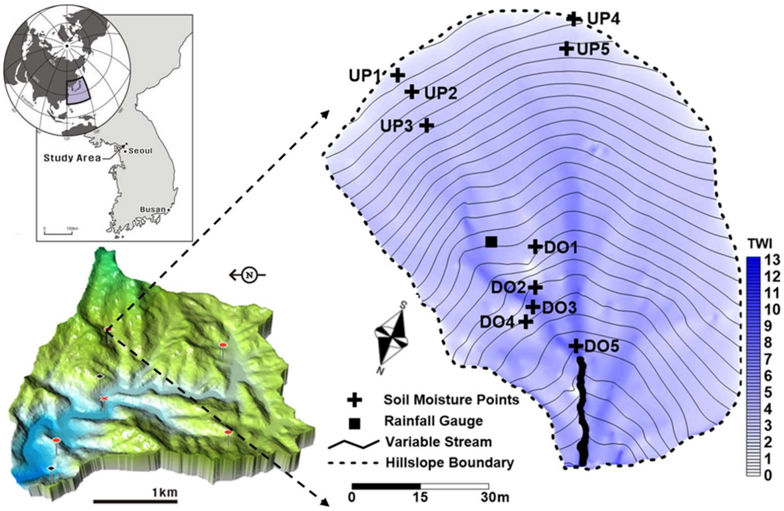
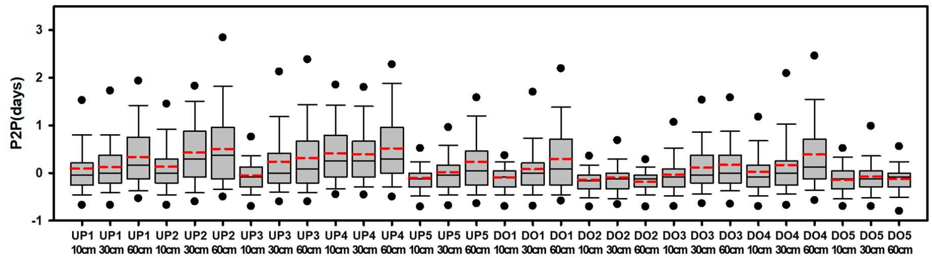


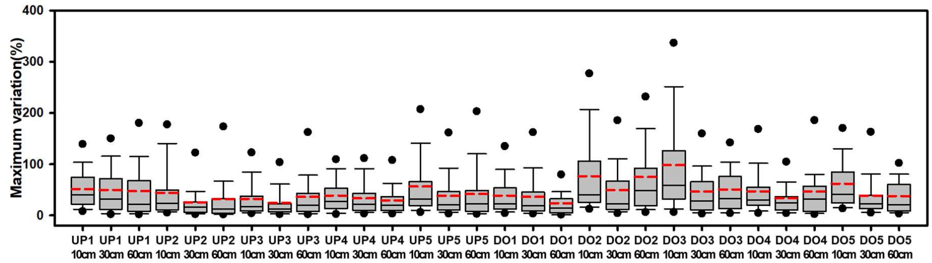
Figure 1. Location of Sulmachun watershed in South Korea with hydrologic monitoring (rainfall and evapotranspiration) stations (lower left) and study area with terrain contours, topographic wetness index (TWI), and soil moisture monitoring points (right). (We created this map)

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
Discussions



(a)



(b)

Figure 2 Boxplots of soil moisture responses of P2P (a) and maximum variation (b) for 30 points.

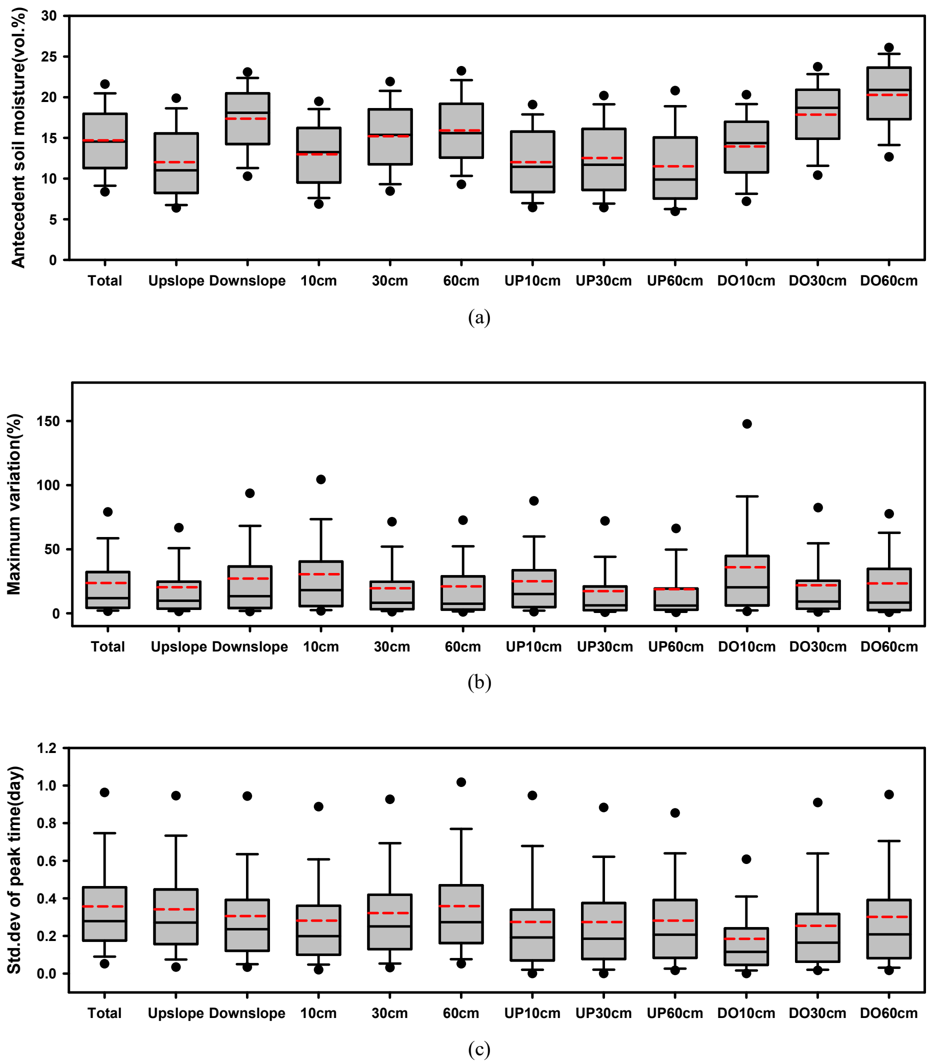
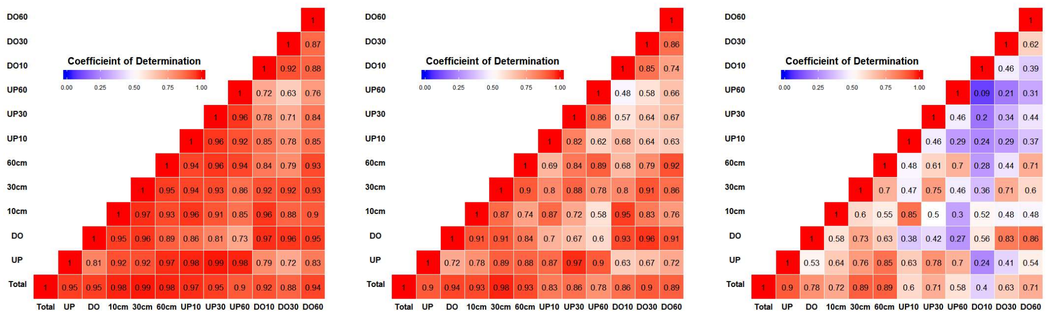


Figure 3 Box plots of antecedent soil moisture (a), maximum variation (b), and standard deviation of peak time (SDP2P) (c) of 12 time series of soil moistures.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

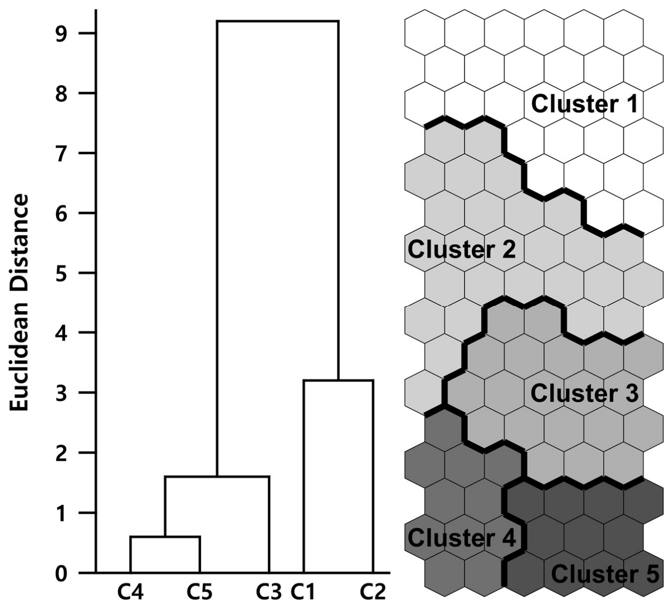


(a) (b) (c)

Figure 4. Heat maps of coefficient of determination among combinations of (a) antecedent soil moisture, (b) maximum variation, (c) standard deviation of peak time.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

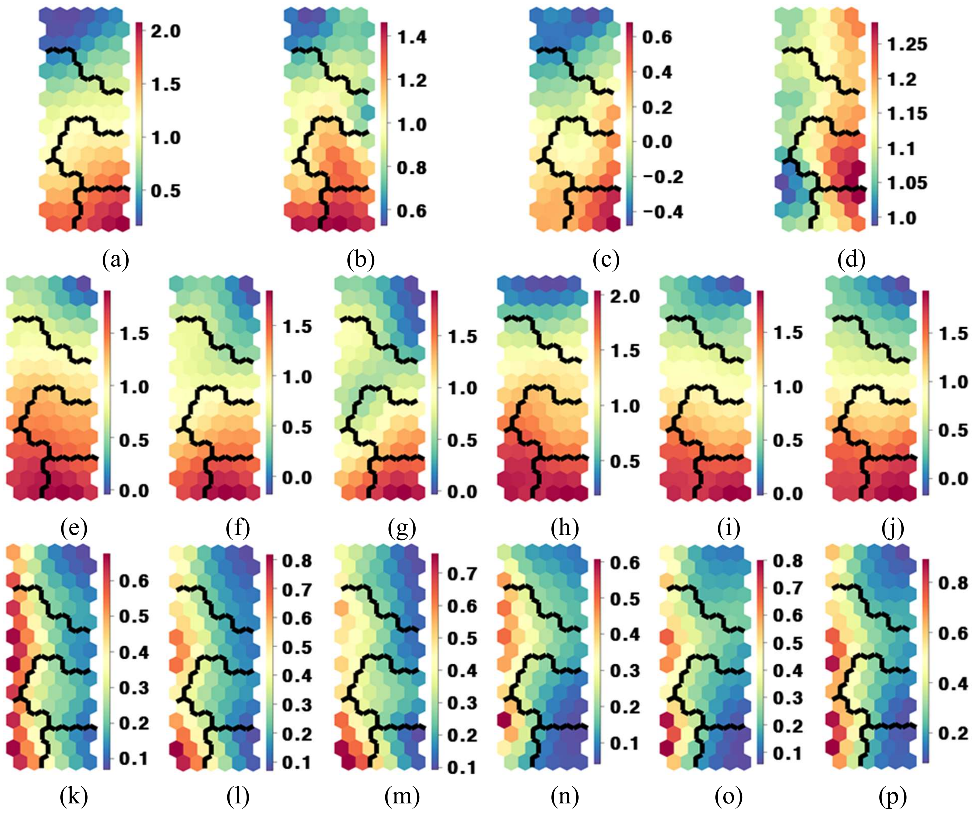


(a) (b)

Figure 5. Structure of (a) dendrogram for five clusters and (b) SOM classifications in 96 neurons through 16 × 6 matrix.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

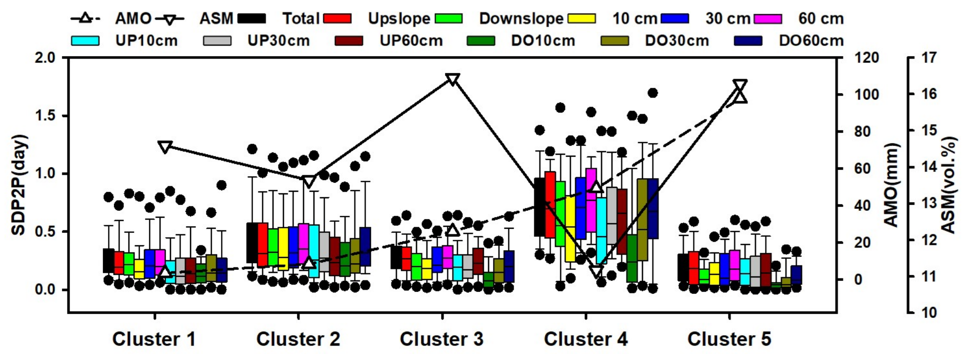
Hydrology and Earth System Sciences
Open Access
Discussions



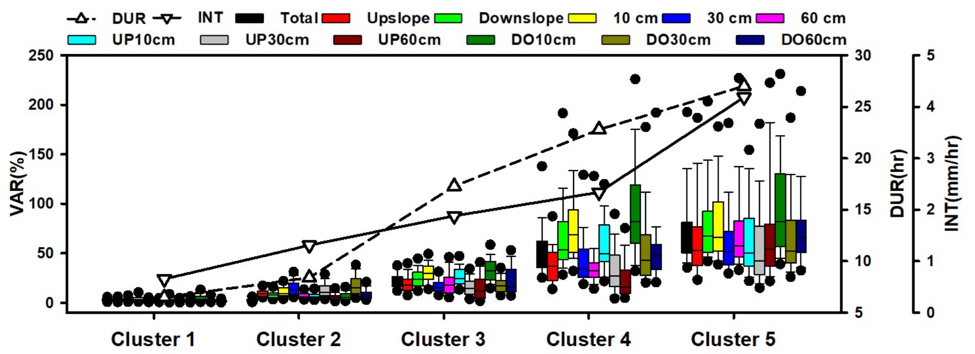
Figures 6. **(a)**–**(p)** Component planes of variable weightings for rainfall amount (AMO) (a); rainfall duration (DUR) (b); rainfall intensity (INT) (c); antecedent soil moisture (ASM) (d); soil moisture difference indices for the upslope and downslope at depths of 10, 30, and 60 cm (VUP10, VUP30, VUP60, VDO10, VDO30, VDO60) (e)-(j); standard deviation of peak time for the upslope and downslope at depths of 10, 30, and 60 cm (SUP10, SUP30, SUP60, SDO10, SDO30, and SDO60) (k)-(p)

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
Discussions



(a)



(b)

Figure 7. SDP2Ps with mean AMO and ASM for each cluster (a) soil moisture difference index with mean DUR and INT for each cluster for total, upslope, and downslope at depths of 10, 30 , and 60 cm, corresponding depths for upslope and downslope.

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Table 1. Arithmetic averages of SOM inputs for rainfall amount (AMO), rainfall duration (DUR), rainfall intensity (INT), antecedent soil moisture for all points (ASMTOT), maximum soil moisture variation (VAR), and standard deviation of peak-to-peak time (SDP2P).

Variables	numbers	AMO(mm)	DUR(hr)	INT(mm/hr)	ASMTOT(vol.%)	
cluster 1	108	3.61	6.50	0.66	14.6	
cluster 2	90	8.45	8.40	1.31	13.6	
cluster 3	75	26.08	17.28	1.88	16.4	
cluster 4	30	49.27	22.80	2.34	11.2	
cluster 5	53	97.80	27.02	4.19	16.3	
VAR(%)	VUP10	VUP30	VUP60	VDO10	VDO30	VDO60
cluster 1	3.8	2.0	2.5	4.6	2.9	1.9
cluster 2	13.2	5.7	6.8	17.5	8.6	7.2
cluster 3	26.9	16.4	16.1	33.4	18.2	22.9
cluster 4	59.1	33.0	23.4	96.1	56.6	54.8
cluster 5	66.7	60.8	73.9	100.7	68.6	77.4
SDP2P	SUP10	SUP30	SUP60	SDO10	SDO30	SDO60
cluster 1	0.21	0.20	0.21	0.16	0.22	0.22
cluster 2	0.37	0.35	0.33	0.30	0.35	0.42
cluster 3	0.22	0.22	0.26	0.11	0.18	0.22
cluster 4	0.56	0.65	0.63	0.36	0.59	0.72
cluster 5	0.17	0.17	0.20	0.06	0.09	0.12

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 2. Soil moisture changes and storage changes for all clusters at depths of 10 cm, 30 cm, and 60 cm and those for upslope and downslope.

average	cluster	10cm	30cm	60cm	upslope			downslope		
					10cm	30cm	60cm	10cm	30cm	60cm
soil moisture change (%)	1	0.5	0.4	0.3	0.4	0.3	0.3	0.6	0.5	0.4
	2	1.9	1.0	1.0	1.5	0.6	0.7	2.3	1.5	1.4
	3	4.5	2.9	3.5	3.7	2.4	2.1	5.2	3.5	5.1
	4	7.4	5.2	4.9	5.3	3.1	2.0	9.8	7.8	9.0
	5	12.0	10.8	13.3	8.9	8.7	10.0	15.4	13.3	16.8
storage change (mm)	1	1.0	0.8	0.9	0.8	0.6	0.9	1.2	1.0	1.2
	2	3.8	2.0	3.0	3.0	1.2	2.1	4.6	3.0	4.2
	3	9.0	5.8	10.5	7.4	4.8	6.3	10.4	7.0	15.3
	4	14.8	10.4	14.7	10.6	6.2	6.0	19.6	15.6	27.0
	5	24.0	21.6	39.9	17.8	17.4	30.0	30.8	26.6	50.4

https://doi.org/10.5194/hess-2021-107
Preprint. Discussion started: 13 April 2021

Hydrology and Earth System Sciences
Open Access
Discussions

Table 3. Combinations of flow paths and its hydrologic conditions for all clusters. SF: surface; SB: subsurface.

Cluster	#	Rainfall Impact	Antecedent soil moisture	Upslope		Downslope	
				Vertical flow	Lateral flow SF/SB	Vertical flow	Lateral flow SF/SB
1	108	Insignificant	Mid	No response (under 2 vol.%)		No response (under 2 vol.%)	
2	90	Intermediate	Mid	No response (under 2 vol.%)		Yes	No
3	30		High	Yes	No/Yes	Yes	No/Yes
4	53	Significant	Low	Yes	No/No	Yes	No/Yes
5	75		High	Yes	No/Yes	Yes	Yes/Yes