https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

Understanding hydrologic controls of slope response to precipitations through Machine Learning analysis applied to synthetic data

Daniel Camilo Roman Quintero[1], Pasquale Marino[1], Giovanni Francesco Santonastaso[1], Roberto Greco[1]

[1]Dipartimento di ingegneria, Università degli Studi della Campania 'Luigi Vanvitelli', via Roma 9, 81031 Aversa (CE), Italy;

Correspondence to: Daniel Camilo Roman Quintero (danielcamilo.romanquintero@unicampania.it)

Abstract:

The assessment of the response of slopes to precipitation is important for various applications, from water supply management to hazard assessment due to extreme rainfall events. It is well known that the underground conditions prior to the initiation of rainfall events control the hydrological processes that occur in slopes, affecting the water exchange through their boundaries. The present study aims at identifying suitable variables to be monitored and modelled to predict the response of the slope to precipitations. A case study consisting of a loose pyroclastic coarse grained soil cover overlaying a karstic bedrock located in the southern Apennines (Italy) is described, where field monitoring has been carried out, comprising stream level recordings, meteorological recordings, and soil water content among others. Nevertheless, to enhance the field dataset, the slope hydraulic behaviour of the case study has been simulated with a physically based model linked to a synthetic rainfall time series, getting a consistent hourly timeseries dataset of 1000 years, containing information on rainfall, aquifer water level and soil volumetric water content at different depths. Machine Learning techniques have been used to unwrap the relationships amongst the studied variables, which relations are commonly non-linear. The Random Forest

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

technique has been used to assess the way the slope response could be addressed and the importance of each variable on the slope response and the k-means clustering technique has been used to explore the geometrical disposition of data, and so the identification of seasonally recurrent different scenarios linked to the slope response. It has been shown that the slope response in terms of the rainwater being stored in the soil cover is naturally highly dependent on the rainfall amount, but water drainage and storage processes can be identified by normalizing the change in water storage with the rainfall depth. Indeed, with the methodology presented here, different hydrometeorological scenarios controlling major hydrological processes have been identified not only from the meteorological and seasonal behaviour but also from the underground conditions prior to the rainfall initiation, weighting the role, on one hand, of the field capacity value on the ease of the water to flow in and out of the soil cover and, on the other hand, of the ground water level, the increase of which gives evidence of the activation of slope drainage even during relatively intense rainfall events.

Keywords: Water storage, slope response, underground antecedent conditions, hydrological controls

1. Introduction

Slope response to precipitations is highly non-linear, in terms of runoff generation, rainwater infiltration and subsurface drainage processes, which are mostly depending on the initial soil moisture state at the onset of the rainfall event (Tromp-Van Meerveld and McDonnell, 2006; Nieber and Sidle, 2010; Damiano et al., 2017). The initial or antecedent conditions are related to hydrological processes that occur in the slope, which control water exchanges between the slope and surrounding systems (i.e., atmosphere, surface water, deep groundwater). These processes occur through the boundaries of the slope, and often evolve over long timescales.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

While the importance of antecedent conditions on overland runoff and drainage has been early identified, and their role as predisposing conditions has already been recognized (Ponce & Hawkins, 1996; Tromp-Van Meerveld & McDonnell, 2006), only recently the scientific community started providing new perspectives to better understand slope predisposing conditions to landslides (Bogaard and Greco, 2018), and so to develop physically-based models capable of integrating this hydrological knowledge for predicting their occurrence (Greco et al., 2018; Bordoni et al., 2015).

The storage of infiltrating water within the soil also requires drainage mechanisms developing in the slopes in response to precipitations to be not so effective to drain out much rainwater (Marino et al., 2020b; Greco et al., 2021). Consequently, especially for nowcasting and early warning purposes, the identification of hydrological variables suitable to identify slope predisposing conditions is extremely useful. Thus, to better understand how the complex hydrological predisposing conditions may control the processes involving the slope response in terms of water storage, a field monitoring campaign allowing the assessment of the slope water balance is highly recommended (Marino et al., 2020a).

The identification of suitable variables to be monitored in the field is indeed useful to achieve an insight of the behaviour of the interconnected hydrological systems (i.e., groundwater, surface water, soil water). The proper evaluation of the hydrological scenarios in a region of interest could impact from flood hazard assessment and the study of rainfall-induced landslides (Bogaard and Greco 2016), to the prediction of possible crop water stress conditions in relation to defoliation (Capretti and Battisti, 2007), pathogen expansions in chestnut grove (Gao and Shain, 1995), and plant mortality in a climate change context as well (McDowell et al., 2008).

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

This research focuses on a case study in an area sensitive to problems associated to both dry and wet seasons, but where particularly destructive rainfall triggered landslides occurred. The triggering of rainfall-induced shallow landslides is controlled by the increase of water storage within the soil cover after intense and persistent precipitations, leading to pore pressure build up (Bogaard and Greco 2016). Such geo-hazards are recurrent along the slopes covered with unsaturated pyroclastic deposits typical of wide areas of Campania, southern Italy (Fiorillo et al., 2001; Revellino et al., 2013). Slope equilibrium is guaranteed by the additional shear strength promoted by soil suction (Lu and Likos 2006; Greco and Gargano 2015), which reduction often leads to slope failure due to shear strength loss by soil wetting during rainwater infiltration (Olivares and Picarelli, 2003; Damiano and Olivares, 2010; Pagano et al., 2010; Pirone et al., 2015).

Not only the rainfall characteristics and the soil saturation play a role in the hydrological slope response to precipitations. Recent studies show that the response of soil covers to precipitations is affected by the wetness of the soil-bedrock interface, which controls the leakage of water into the underlying fractured limestone (Marino et al., 2020a, b). At the contact between soil cover and bedrock, intense weathering modifies the physical properties of the soil as well as of the fractured bedrock, which seem to be hydraulically interconnected. The changing hydraulic behaviour of the soil-bedrock interface can be related to the storage of water in the uppermost part of the fractured bedrock (Greco et al., 2018, 2014). Therefore, the aquifer water level arises as another possible monitoring variable to be considered within slope controls.

The aim of this study is to identify the major hydrological processes controlling the slope response to precipitations, and the most suitable variables to quantify them. Therefore, a methodological approach gathering wide knowledge on field process simulations and data analysis is presented. In that way, a 1-D physically based mathematical model, that couples the interaction between the unsaturated

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

pyroclastic soil cover and the underlying perched aquifer (Greco et al., 2018), was applied to describe the behaviour of a slope located in Campania (southern Italy), under a 1000-years long hourly rainfall synthetic time series. The model had been previously calibrated and validated on data from field monitoring carried out since 2001 (Greco et al., 2013; Comegna et al., 2016; Damiano et al., 2012). The results of the simulations provide the time series of soil suction, water content and aquifer water level, assumed as representative of real field conditions. Once identified the rainfall events within the 1000-years hourly timeseries, a specific dataset is built with the antecedent conditions one hour before the beginning of any rainfall event, comprising the previously named variables plus the total rainfall amount and the change in water stored in the soil cover at the end of the rainfall events. The dataset was analysed with Machine Learning techniques, i.e., clustering, and random forest, to disentangle the nonlinear processes controlling the slope hydraulic behaviour due to rainfall events, and thus to identify the role played by the variables, suitable to be monitored in the field, on the slope response to precipitations.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

2. Materials and methods

The studied slope belongs to the Partenio Massif, and it has the typical characteristics of many pyroclastic slopes of Campania (southern Italy) (Greco et al., 2018). Indeed, three major zones characterized by unsaturated pyroclastic deposits can be identified in Campania (Cascini et al., 2008): Campanian Apennine chain, composed by carbonate rock covered by a variable layer of pyroclastic soil (from 0.1 to 5 m); Phlegraean district, formed by underlying densely fractured volcanic tuff bedrock, placed under several meters of pyroclastic soils; and Sarno and Picentini Mountains, where a thin layer of pyroclastic material is over a terrigenous bedrock. In these three areas, the thickness of the soil covers results quite variable, according to the slope inclination and to the distance from the eruptive centre (De Vita et al., 2006; Tufano et al., 2021).

A synthetic dataset of slope hydrologic response to precipitations, based on soil cover interaction with rainfall (in the uppermost boundary) and underlying shallow aquifer (in the lowermost boundary), has been built. The synthetic dataset has been considered reliable to reproduce slope response to climate forcing in terms of soil volumetric water content, perched aquifer water level, and rainfall.

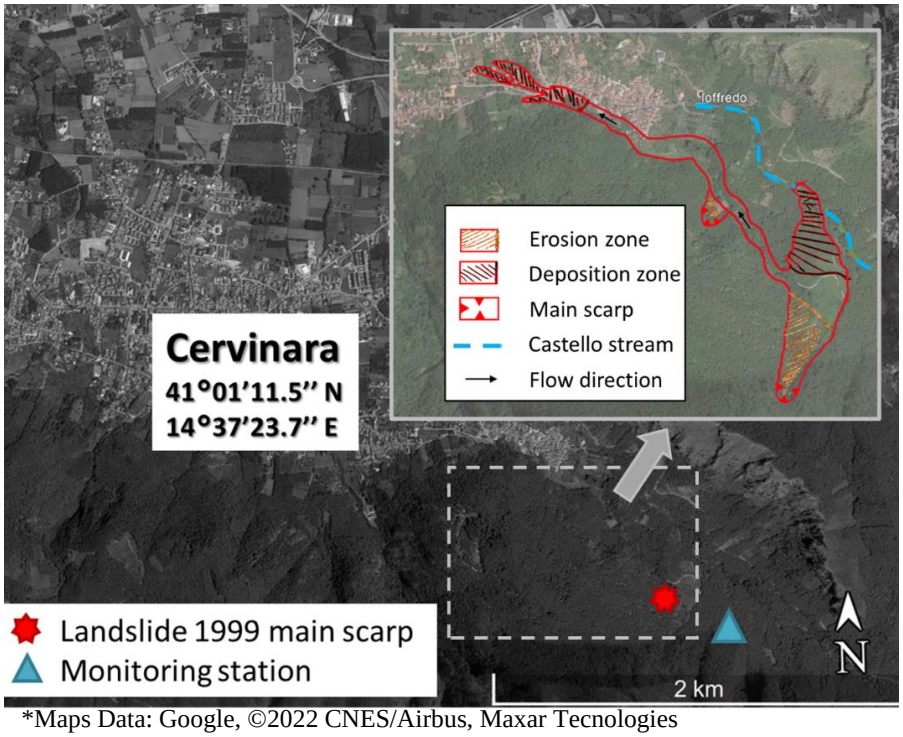
Moreover, the obtained data set has been analysed with Machine Learning techniques, as they result quite powerful to identify non-linear relationships between variables. Besides, the results which come from these data analyses have been compared a priori to the limited dataset from field monitoring, showing a reasonable agreement.

2.1. Case study

The study area refers to the north-east slope of Monte Cornito, part of the Partenio Massif (Campania, southern Italy), 2 km from the town of Cervinara, about 40

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

km northeast of the city of Naples. The slope was involved in a series of rapid shallow landslides after a rainfall event of 325 mm in 48 hours during the night between 15–16 December 1999, causing casualties and heavy damages (Fiorillo et al., 2001). A field monitoring station was installed nearby the big landslide scarp since 2001. Further details of the investigated zone, with indications of the area affected by the largest of the landslides triggered in 1999, are shown in Figure 1.



*Maps Data: Google, ©2022 CNES/Airbus, Maxar Tecnologies

Figure 1. Location of the study area and indication of the zone affected by a large landslide in 1999[1]

Partenio Massif is part of the southern Apennines area. The bedrock mainly consists of Mesozoic-Cenozoic fractured limestones, mantled by loose

[1] Google (2022) *Cervinara, Italy*. Available at: https://www.google.com/maps/@41.0114559,14.6411297,2097m/data=!3m1!1e3 (Accessed: 7 March 2022)

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

pyroclastic deposits, resulting from the explosive volcanic activity of Somma-Vesuvius and Phlegrean Fields, which occurred over the last 40.000 years (Rolandi et al., 2003).

The fractured limestone formations of the southern Apennines often host large karst aquifers, through which a basal groundwater circulation occurs, for which regional groundwater recharge between 100 and 500 mm/year has been estimated, with 200 mm/year regarding the area of Cervinara (Allocca et al., 2014). Moreover, recent studies showed that, in the upper part of the karst system, denoted as epikarst (Hartmann et al., 2014), more permeable and porous than the underlying rock, a perched aquifer often develops (Williams, 2008; Celico et al., 2010). It temporally stores water and favors the recharge of the deep aquifer through the larger fracture system. The water, which is accumulated temporally in the epikarst, also reappears at the surface in small ephemeral streams.

Specifically, the slope of Cervinara has an inclination between 35° and 50°, at an elevation between 500 m and 1200 m above sea level. The soil cover, usually in unsaturated conditions, is the result of the air-fall deposition of the materials from several eruptions, so it is generally layered. It mainly consists of layers of volcanic ashes (with particle size in the range of sands to loamy sands) alternating with pumices (sandy gravels), laying upon the densely fractured limestone bedrock. Near the soil-bedrock interface, a layer of weathered ashes, characterized by finer texture (silty sand), with lower hydraulic conductivity, moderate plasticity and low cohesion, is often observed (Damiano et al., 2012).

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

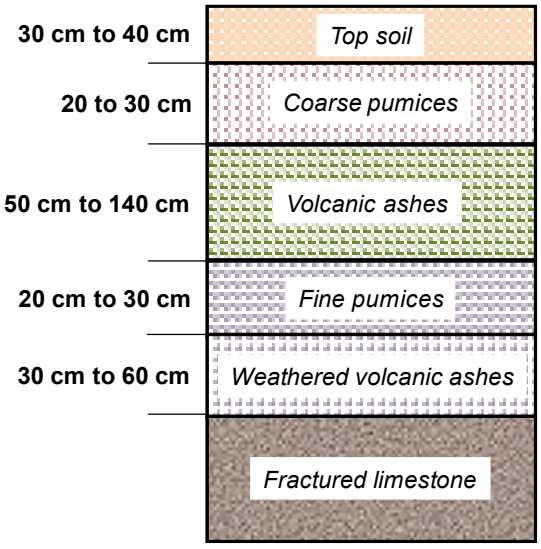


Figure 2. Characteristic soil profile for the slope near the scarp of the landslide occurred in 1999 in Cervinara

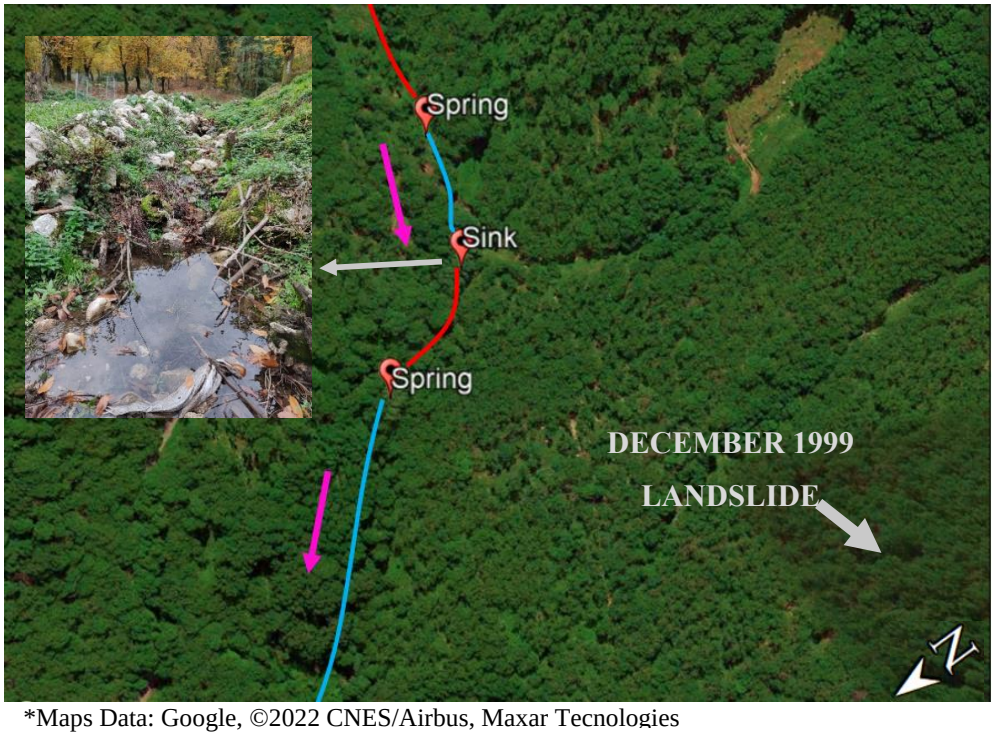
The soil cover thickness varies spatially from a minimum of 1.0 m, in the steepest part of the slope, to larger values at its foot (up to 4-5 meters). Figure 2 shows the soil layers constituting the cover, as found throughout the slope near the main scarp of 1999 landslide.

The pyroclastic ashes usually exhibit high porosity (up to 75%) and quite high values of saturated hydraulic conductivity (ranging up to $1x10^{-5}$ m/s). Thus, this kind of soil lets rainwater infiltrate during the most intense rainfall events, with little runoff generation, and it can store a large amount of water.

The climate is Mediterranean, which is characterized by dry and warm summer and rainy autumn and winter, with mean annual precipitation of about 1600 mm, mostly occurring between October and April. The total potential evapotranspiration ET_0, estimated with the Thornthwaite formula (Shuttleworth, 1993), is between 700 mm and 800 mm in the altitude range between 750 m and 400 m (Greco et al., 2018). The vegetation mainly consists of widespread deciduous chestnuts, with a dense understory of brushes and ferns, growing

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

206 during the flourishing period (between May and September). In fact, visual
207 inspections of the soil profile showed a large amount of organic matter and roots.
208 In most cases, roots are denser in the uppermost part of the soil cover and become
209 sparse between the depth of 1.50 m and 2.00 m below the ground surface,
210 reaching the basal limestones and penetrating the fractures.



211

212 **Figure 3. Identification of surface water flow in the Castello stream at the beginning**
213 **of the rainy season in November 2021 by visual recognition of springs and sinks in**
214 **the watercourse**[2]

215 Moreover, in the surrounding area, several ephemeral and perennial springs are
216 present, mostly located at the foot of the slopes, which supply a network of small
217 creeks and streams, allowing to show the activity of the aquifer discharge to the

[2] Google (2022) *Cervinara, Italy*. Available at: https://www.google.com/maps/@41.0088511,14.65137,786m/data=!3m1!1e3 (Accessed: 7 March 2022)

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

surface water. An indication regarding the Castello stream (the main stream for this side of the basin), with springs, is shown in Figure 3, where, during a field recognition in November the 11th 2021, the surface water flow appeared (springs) and disappeared (sinks) in some points along the stream course. Normally the stream exhibits its lowest water depth values up to the beginning of the late autumn (Marino et al., 2020a, p.3.3), but it is interesting to note that the surface water in the stream emerging from the epikarstic springs is feasible to be monitored as an indicator of the slope drainage status.

2.1.1. Field monitoring data

Several hydrological monitoring activities have been carried out at the slope of Cervinara since 2001, initially consisting of measurements of precipitations and manual readings (every two weeks) of soil suction by "Jet-fill" tensiometers, equipped with a Bourdon manometer (Damiano et al., 2012). Afterwards, since November 2009, an automatic monitoring station has been set at an elevation of 585 m a.s.l., near a narrow track close to the landslide scarp of December 1999. The installed instrumentation consisted of tensiometers, time domain reflectometry (TDR) probes for water content measurements, and a rain gauge (Greco et al., 2013; Comegna et al., 2016).

Since 2017, the hydro-meteorological monitoring was enriched (Marino et al., 2020a), aiming at understanding the seasonal behaviour of the slope and the interactions between the hydrological systems, i.e., the unsaturated soil cover, the epikarst, and the underlying fractured bedrock.

Specifically, the data collected by tensiometers and TDR probes were supplemented with those from a meteorological station (composed by a thermo-hygrometer, a pyranometer, an anemometer, a thermocouple for soil temperature measurement, and a rain gauge), and with the water level in two streams at slope

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

foot, so to gain useful information for the assessment of the water balance of the studied slope.

The data collected during these field monitoring activities have been useful to highlight seasonally recurrent soil moisture distributions. More details about the recurrent seasonal behaviour related to the moisture conditions of the soil profile of the area of Cervinara can be found in Marino et al. 2020a.

2.2. Synthetic dataset

Aiming at identifying suitable variables to be monitored in the field for the identification of the conditions controlling different slope responses to the precipitations, a rich dataset of rainfall and underground monitored variables, such as soil moisture and groundwater level, is needed. However, a complete field monitored dataset is not always possible to be analyzed and, when it exists, it is commonly available for short periods, granting a relatively low measurement density. Hence, a synthetic dataset, aiming at improving the information obtained from field monitoring, has been generated. This dataset has been obtained by means of the physically based mathematical model described hereinafter (section 2.2.2). The model has been run with a 1000-years synthetic hourly rainfall series, obtained with a stochastic rainfall generator, for which further details are given in section 2.2.1.

2.2.1. Definition of synthetic rainfall events

The Neyman-Scott rectangular pulse model (NSRP) has been used to obtain a 1000-years long synthetic hourly series of precipitations. The NSRP model reproduces the precipitation process as a set of rain clusters, composed by possibly overlapping rain cells embodied by rectangular pulses, each one with random origin. The storm duration is represented by the cell width and its height represents the associated rainfall intensity, so that when multiple cells overlap,

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

the total intensity is the sum of the intensities of the overlapping cells (Rodriguez-Iturbe et al. 1987; Cowpertwait et al. 1996).

NSRP model calibration requires the identification of five parameters, using the method of moments (Peres and Cancelliere, 2014), based on available rainfall data for the investigated site. Specifically, the data from the rain gauge station of Cervinara, situated near the Loffredo village, belonging to the Civil Protection Agency of Campania Region available from January 2001 to December 2017 with a time resolution of 10 min, were used.

The aim of this study is the identification of variables expressing the slope conditions responsible of different responses to precipitations. In that sense, it is important to define the events within the rainfall time series to clearly distinguish antecedent conditions from the effects of the current rainfall event.

In other words, within the 1000-years long time series, a criterion should be identified to separate rainfall events, so that a new event begins only when the effects of the previous one disappeared. For this study, the events were defined as periods with at least 2mm of rainfall, preceded and followed by at least 24h with less than 2mm (i.e., smaller than the mean daily potential evapotranspiration estimated for the case study).

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

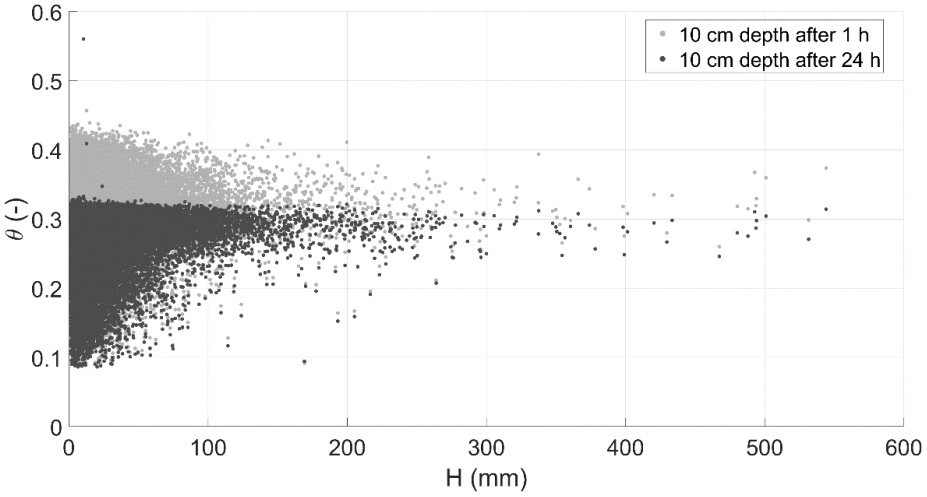


Figure 4. Scatter plot of event rainfall depth and mean volumetric water content of the top 10 cm soil depth 1 hour (grey dots) and 24 hours (black dots) after the end of each rainfall event

In fact, the mean volumetric water content (θ) at 10 cm depth drops below soil field capacity ($\theta \cong 0.35$) 24 hours after the end of each event (Figure 4) in all the cases in which such value was overcome at the end of the event. This shows that a dry interval of 24 hours after a rainfall event is long enough for drainage processes to remove from the topsoil most of the water infiltrated from the previous event. As topsoil moisture controls the infiltration capacity at ground surface, after such interval the infiltration of new rainfall is only little affected by the remnants of the previous rainfall event.

With the assumed separation criterion, a total of 53061 rainfall events within 1000 years are obtained, with durations ranging between 1 and 570 hours, and total rainfall depth between 2 and 710 mm.

2.2.2. Slope hydrological model

A simplified 1-D model has been built, previously validated according to the data collected during the hydrological monitoring activities (Greco et al., 2013), to investigate the hydrological response of the slope to synthetic hourly precipitation data. The unsaturated flow through the soil cover is modelled with

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

1-D head-based Richards' equation (Richards, 1931), assuming for simplicity a single homogeneous soil layer, and it is coupled with a model of the saturated water accumulated in the perched aquifer. The adoption of a 1-D model is allowed thanks to the geometry of the considered soil cover, as well as to the prevailing water potential gradients orthogonal to the ground surface in unsaturated conditions.

The root water uptake has been accounted in the source term of the model, according to the expressions by Feddes et al. (1976), based on estimated potential evapotranspiration, with maximum root penetration depth equal to the soil cover thickness and triangular root density shape.

Two boundary conditions are considered for the unsaturated soil cover. At ground surface (i.e., the upper boundary condition), if the rainfall intensity is greater than the current infiltration capacity, the excess rainfall forms overland runoff. Otherwise, all rainfall intensity is set as infiltration. The bottom boundary condition links the soil cover to a perched aquifer developing in the fractures and hydraulically connected to the unsaturated cover through the weathered soil layer (less conductive and capable of retaining much water), located at the contact between the cover and the bedrock. This soil layer penetrates the vertical conduits and fractures (Greco et al., 2013). In this context, the perched aquifer is modelled as a linear reservoir model, that receives water from the gravitational leakage of the overlying unsaturated soil cover and releases it as deep groundwater recharge and spring discharge (Greco et al., 2018). The pressure head at the soil–bedrock interface is assumed to follow the fluctuations of the water table of the underlying aquifer.

The hydraulic parameters of the soil cover have been obtained from previous laboratory tests (Damiano and Olivares, 2010) and field monitoring data analysis (Greco et al., 2013), considering the van Genuchten-Mualem model for the

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

hydraulic characteristic curves (van Genuchten, 1980). The parameters describing the hydraulic behaviour of the perched aquifer hosted in the upper part of the limestone bedrock have been derived from previous studies, which showed that the model satisfactorily reproduced the fluctuations of water potential and moisture, observed at various depths in the unsaturated soil cover, both during rainy and dry seasons (Greco et al., 2013; 2018). Model parameters are summarized in Table 1.

Table 1. Hydraulic parameters of the coupled model of the unsaturated soil cover and of the aquifer hosted in the Epikarst (Greco et al. 2021)**.**

Soil cover	Soil cover thickness (m)	2
	Saturated water content (-)	0.75
	Residual water content (-)	0.01
	Air entry value (m^{-1})	6
	Shape parameter (-)	1.3
	Saturated hydraulic conductivity (m/s)	$3x10^{-5}$
Epikarst	Epikarst thickness (m)	14
	Effective porosity (-)	0.005
	Time constant of linear reservoir (days)	871 days

The equations have been numerically integrated with the finite difference technique, with a time step of 1 hour over a spatial grid with vertical spacing of 0.02 m.

It is important to note that, even if the model simplifies the reality assuming a homogeneous soil profile, a more complex approach considering a layered profile would lead to difficult application of the model at less detailed scales such as regional and catchment scales. Consequently, the hydraulic properties of the homogeneous soil layer should be considered as effective properties, useful to reliably reproduce the observed phenomena.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

2.2.3. Synthetic hydrometeorological data

As it has been stated from previous sections, the dataset comes from the simulation of the hydrologic response of a slope to 1000-years long hourly rainfall time series, carried out with a physically based model, calibrated for the case study. The output contains the time series of soil water content and suction at all depths throughout the soil cover, of the water exchanged between the soil and the atmosphere, of the leakage through the soil-bedrock interface, and of the predicted water level of the underlying aquifer.

One hour before the onset of each rainfall event, the following variables have been extracted, as they would be measurable in the field and are representative of antecedent conditions: the aquifer water level (h_a), the mean volumetric water content in the uppermost 6 cm of soil cover (θ_6) and the mean volumetric water content in the uppermost 100 cm of soil cover (θ_{100}). To quantify the effects of rainfall on the slope response, the change of the water stored in the soil cover at the end of each rainfall event (ΔS) has been computed and compared with the total rainfall depth of the event (H).

Specifically, the inclusion of soil water content information has been chosen, as it can be obtained from available satellite-derived remote sensing products (Pan et al., 2020; Paulik et al., 2014) or from field sensor networks (Wicki et al., 2020). Besides, as the model introduces a linear relationship to estimate the outflow from the groundwater system, the monitored stream water level has been considered comparable to the simulated groundwater level, as the two variables are assumed directly proportional in the model.

2.3. Data analysis techniques

The resulting dataset has been analyzed with Machine Learning techniques, aiming at capturing the complex interactions between the hydrological subsystems (i.e., soil cover, fractured bedrock, surface water). Indeed, the

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

analysis of the data is not only constrained to classical statistical analyses, such as data frequency distributions, but also to data classification based on their geometrical distribution, and on quantifying the importance of the considered antecedent variables on the simulated response as well.

2.3.1. Variable importance assessment by Random Forest

The Random Forest is a Machine Learning method that sets its basis on the theory of regression/classification trees, bagging data and capturing even the complex or non-linear interactions in-between the data of a set with relatively low bias (Breiman, 2001). This method is often used to forecast a desired variable based on predictor variables in terms of regression or classification set of randomly constructed trees. In this case, a regression based Random Forest technique is applied to predict the slope response (ΔS) at the end of each rainfall event, using as predictors all possible triplets of variables described in the section 2.2.3 (total rainfall depth, aquifer water level, mean volumetric water content in the uppermost 6 cm and mean volumetric water content in the uppermost 100 cm): A total of 100 trees with a maximum leaf split of 20 nodes have been built for all variable combinations and trained with the 80% of the dataset randomly selected to obtain different regressor models. Then, the best triplet of predictor variables is selected according to the lowest value of the Root Mean Squared Error (RMSE) calculated using the test data set consisting of the 20% of the remaining data.

Furthermore, to understand how a single variable affects the regression model, the predictor importance is measured by the sensitivity of Random Forest model to the predicted variable (i.e., soil cover response), which is proportional to the RMSE, by permuting on purpose the variables between the levels of the model and estimating the change in the RMSE. Hence the most important variable is the one that exhibits the greatest change in RMSE after the permutations. Hence, the importance of predictor variables follows the magnitude of the change in RMSE.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

This feature is usually referred to in relative terms to the most important variable, called variable relative importance (Hastie et al., 2008).

2.3.2. Data classification by clustering analysis

The exploratory analysis of spatial large datasets is often performed by means of clustering techniques, aiming at identifying different classes in the data, accounting on the distribution of the variables under study. There are two types of clustering algorithms used for class identification purposes: algorithms based on the density of points and algorithms based on the distance between points. The algorithm used here is named k-means, and it is a distance-based procedure to cluster data, based on the number of desired clusters and their centroids. The algorithm proceeds to iteratively minimize the variance of the Euclidean distance of cluster elements from their centroids, by accordingly moving these latter. Consequently, the data labelling is done based on their geometrical disposition in the studied plane or space, depending on the target number of clusters to be identified (Lloyd, 1982; Arthur and Vassilvitskii, 2007). When variables with very different magnitudes are being related for clustering purposes, it is convenient to normalize the data keeping the relative distances between observations. Therefore, the clustering here is applied to the standardized data to exploit the variance of each variable and keeping the geometrical disposition between observations stable.

As the k-means algorithm does not automatically estimate the optimal number of clusters to be identified within the dataset, the Silhouette metric has been used here to evaluate the preferred number of clusters (Rousseeuw, 1987; de Amorim and Hennig, 2015). In fact, this metric quantifies the quality of cluster identification by scoring the difference between the overall average intra-cluster distances and the average inter-cluster distances related to the maximum between the latter two. In that way the metric would always be a value ranging from -1 and 1, where typically 1 means that clearly distinguished clusters have been

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

identified, 0 means that the identified clusters are indifferent, and -1 means that data are mixed in the identified clusters.

3. Results and discussion

The analysis is developed on both field monitored and synthetic datasets, to quantify the information provided by the defined antecedent variables useful to predict the seasonal changes of the slope response to precipitations.

3.1. Role of measurable variables on the slope response

To select the most appropriate triplets of predictor variables, four Random Forest models for predicting the change in water storage (ΔS) in the soil cover, associated to rainfall events, are trained on the dataset consisting of all possible combinations of the synthetic variables: $\langle H, \theta_6, \mathrm{h_a} \rangle$, $\langle H, \theta_{100}, \mathrm{h_a} \rangle$, $\langle H, \theta_6, \theta_{100} \rangle$ and $\langle \theta_6, \theta_{100}, \mathrm{h_a} \rangle$. For each Random Forest model, the values of the Root Mean Square Error (RMSE) are calculated, and the importance of each predictor variable is evaluated according to the procedure described in Section 2.3.1. The obtained results are reported in Table 2.

Table 2. RMSE and variable importance for H, θ_6, θ_{100} and h_a in the prediction of soil response described as ΔS

		Importance			
Dataset	RMSE	H	θ_6	θ_{100}	$\mathrm{h_a}$
$\langle H, \theta_6, \mathrm{h_a} \rangle$	5.353	0.963	0.024	-	0.012
$\langle H, \theta_{100}, \mathrm{h_a} \rangle$	4.336	0.964	-	0.024	0.010
$\langle H, \theta_6, \theta_{100} \rangle$	4.706	0.962	0.014	0.022	-
$\langle \theta_6, \theta_{100}, \mathrm{h_a} \rangle$	24.665	-	0.313	0.340	0.345

As it could be seen from Table 2, an expected behavior on the slope response, quantified as ΔS, answers back highly depending on the rainfall amount as the most important variable in the triplet, while the antecedent conditions show an

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

almost negligible importance with less than 3%. Additionally, the absence of total rainfall among the predictors leads to a substantial increase in the RMSE (24.665). This result is physically expected; the more it rains the more the water storage in the slope increases. Nevertheless, such strong relationship between H and ΔS could conceal important hydrological processes going on the slope. For this reason, normalization schemes could be introduced to account different processes in the slope response. In this case the ratio $\Delta S/H$ is used as a measure of the association between the amount of rain and how much it was stored in the soil cover, pursuing to detach the water drainage processes from the water accumulation processes. To this end, additional Random Forest models are trained on the same four datasets for predicting the normalized water storage in slope $\Delta S/H$.

Table 3. RMSE and variable importance for H, θ_6, θ_{100} and h_a in the prediction of soil response described as $\Delta S/H$

		Importance			
Dataset	RMSE	H	θ_6	θ_{100}	h_a
$\langle H, \theta_6, h_a \rangle$	0.213	0.352	0.329	-	0.319
$\langle H, \theta_{100}, h_a \rangle$	0.197	0.293	-	0.405	0.302
$\langle H, \theta_6, \theta_{100} \rangle$	0.203	0.340	0.261	0.399	-
$\langle \theta_6, \theta_{100}, h_a \rangle$	0.210	-	0.292	0.414	0.293

Normalizing the slope response as $\Delta S/H$ (Table 3) highlights the important role played by the other variables. It is worth to note the importance of the ground water level, which can be compared with the importance of the two water contents and the total rainfall depth. The importance of h_a on the slope response suggests the presence of conditions of change in soil storage less connected to

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

rainfall values and more linked to the capability of water exchange between the aquifer and the soil cover, as it will be discussed in the following sections. Moreover, in Table 3 the variables with the lowest RMSE values are total rainfall depth, aquifer water level and mean volumetric water content in the uppermost 100 cm. According to the Random Forest model, they are the most informative for predicting the slope response. Therefore, the triplet $\langle H, \theta_{100}, \mathrm{h_a} \rangle$ is used for further analysis.

3.2. Underground antecedent conditions

The field monitoring activities allow to get a complete dataset that traces the rainfall values coupled with the soil mean volumetric water content in the uppermost meter of the soil profile (θ_{100}) and the water depth of the Castello stream ($\mathrm{h_s}$), both measured hourly for three years. The field monitored data, composed by 57 rainfall events, include the water level of the Castello stream rather than the direct measurement of the aquifer water level ($\mathrm{h_a}$). Nevertheless, a direct relationship links the water level in the aquifer and the water level in the stream, as assumed for the mathematical modelling. This dataset has been enriched synthetically, as it has been described in section 2.2.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

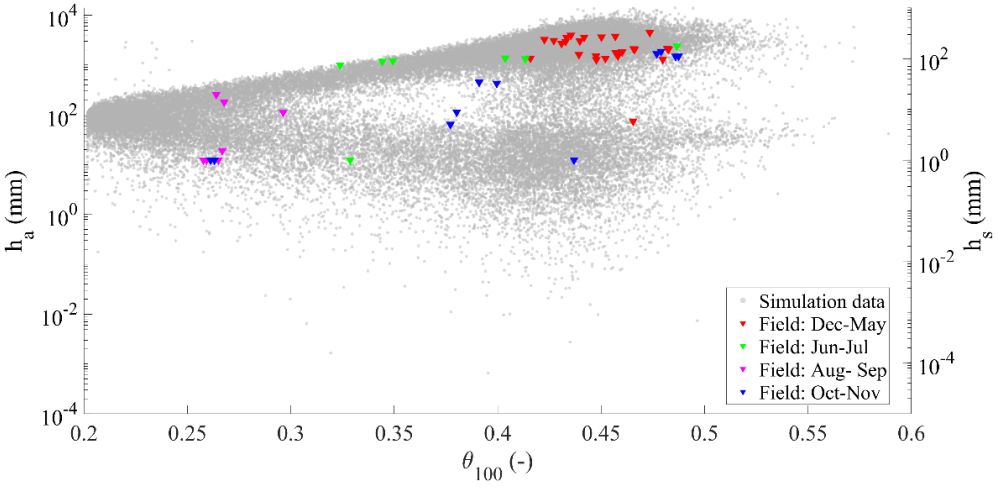


Figure 5. Field monitored mean volumetric water content in the upper meter of the soil profile (θ_{100}) and water depth in the Castello stream (h_s) compared with simulated data (the vertical axis is plotted in logarithmic scale to help visualizing of small water levels).

Therefore, to analyze the effects of the underground conditions on the slope response, Figure 5 shows the simulated data (circular dots in the background) and the field monitored data (triangular colored dots). Logarithmic axes are used to distinguish the very low aquifer water level from the high values.

Four major seasonally recurrent conditions could be identified for the water in the subsurface system from field monitored data: first, a condition usually occurring between December and May is characterized by the highest water content in the soil and the highest measured water level in the stream. Second, the period from June to July is characterized by intermediate water content values, with still high level in the stream. Third, the period from August to September is characterized by the lowest values of water content in the soil, but also the lowest water depth h_s measured in the stream (few centimeters, in some cases nearly zero). Finally, the period from October to November is characterized by a wide range of values in soil water content and a relatively low range of stream water depth.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

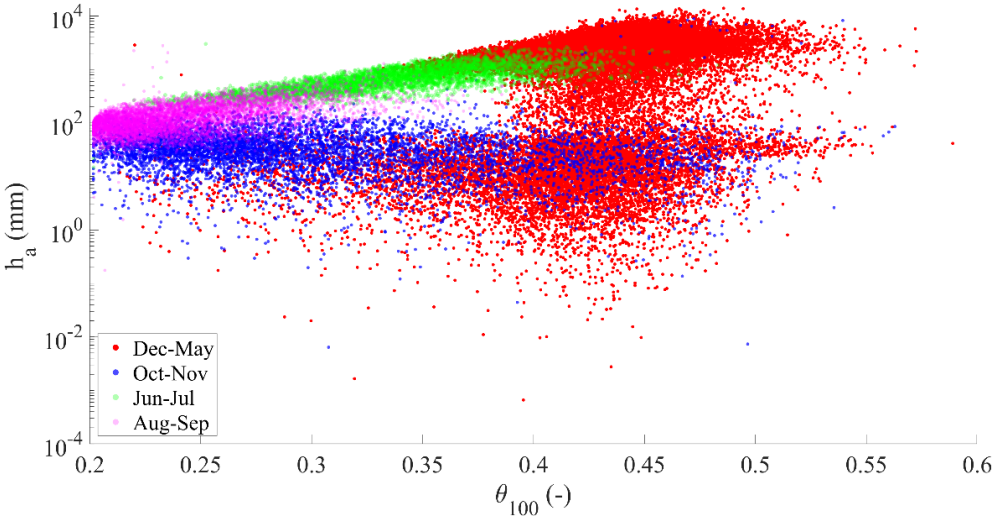


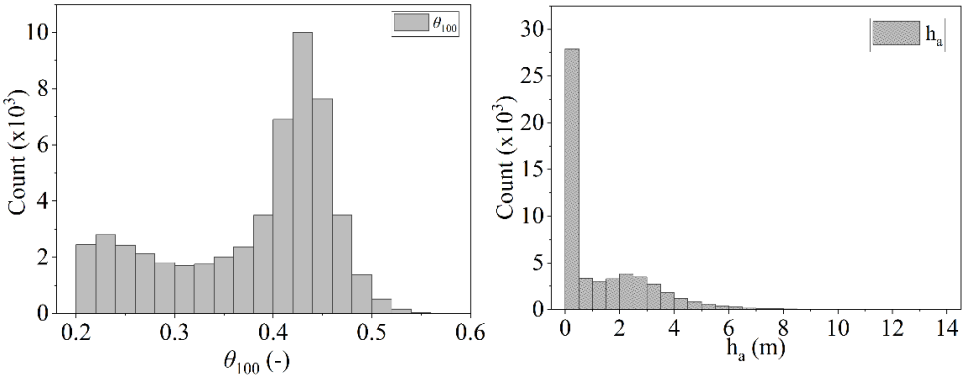
Figure 6. Seasonal behavior of the aquifer water level (h_a) and the mean volumetric water content of the upper meter of the soil profile (θ_{100}) for the synthetic dataset (the vertical axis is plotted in logarithmic scale to help visualizing small water levels).

The underground antecedent conditions are naturally linked to a seasonal behavior dominated by the hydrological conditions which can be traced in time as it can be seen from the synthetic data (Figure 6). The months from December to April follow a winter and spring behavior, characterized by wet soil conditions and aquifer water levels ranging from low to high. From June to July, a late spring behavior is visible, characterized by relatively dry soil (i.e., most of the data falling below soil field capacity), in combination with relatively high groundwater levels (indicating a still active slope drainage). In August and September, a summer like behavior is shown, with the driest soil water content and generally low aquifer water level. Finally, in October and November, the end of the dry season is shown: a wide range of soil wetness coupled with a still low aquifer water level.

For both the field monitored and synthetically obtained datasets, the observed conditions are the result of the time lag between the beginning of the rainy season and the slope response. The recurrent seasonal behavior observed for the

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

535 synthetic dataset, although delayed or anticipated owing to the year-by-year
536 variability of rainfall, is close to that observed in the field.



537

538 **Figure 7. Histograms for data distributions of θ_{100} (left) and h_a (right) for the**
539 **synthetic dataset**

540 The overall situation for the synthetic dataset of antecedent conditions (i.e.,
541 duplets $\langle \theta_{100}, h_a \rangle$) can be described by the distribution of each individual
542 variable, which can be seen in the histograms shown in Figure 7. It is interesting
543 to note that, for both θ and h_a, a bimodal behaviour is observed, corresponding
544 to dry and wet field conditions.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

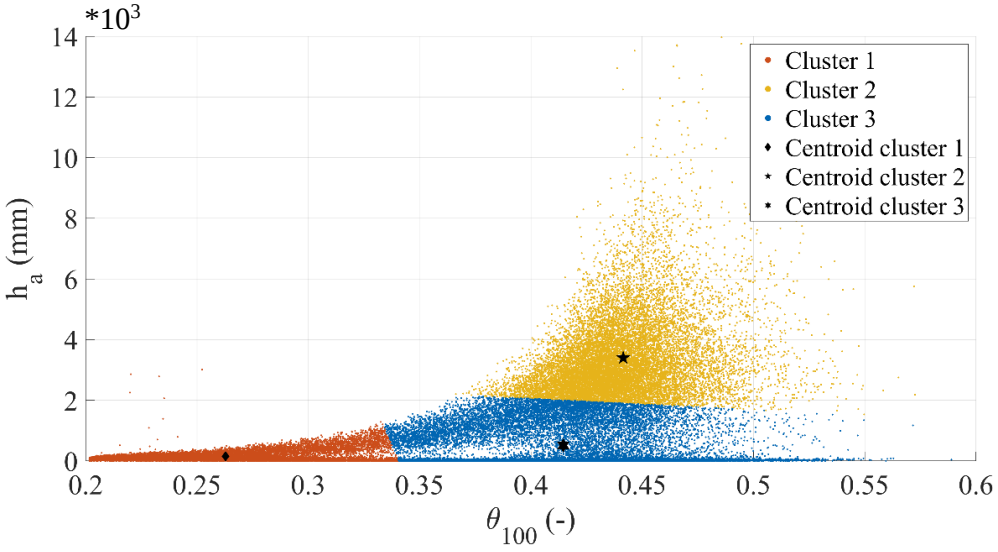
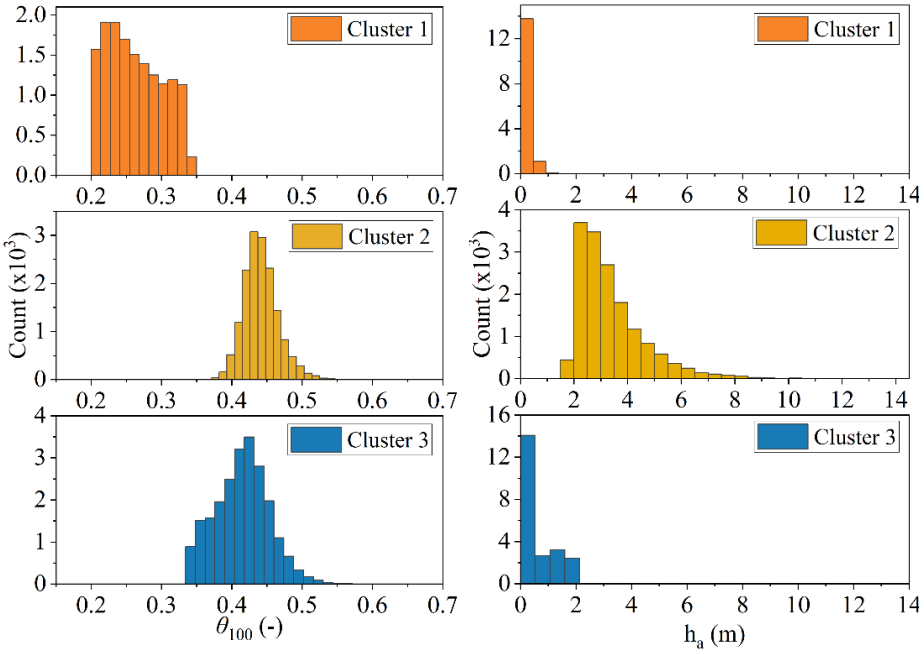


Figure 8. Identified clusters for the duplets $\langle \theta_{100}, h_a \rangle$ representing underground antecedent conditions of the synthetic dataset. For each cluster, the centroids are shown.

The k-means clustering technique has been used to investigate the geometrical distribution of the duplets $\langle \theta_{100}, h_a \rangle$, with number of clusters ranging from 2 to 7. According to the Silhouette metric, the optimal number of clusters is 3, with a metric value of 0.7, allocating the 28%, 30% and 42% of the data in clusters 1, 2 and 3 respectively. Figure 8 shows the 3 clusters obtained within the synthetic dataset. Centroid positions are also displayed, showing the zones of the clouds where most of the dots are gathered. This representation of the data use both vertical and horizontal axes in linear scale to let visualize distance magnitudes between the different clusters, but it corresponds to the same dataset shown in Figure 6.

The distribution of the data after clustering is also analyzed for each cluster and the histograms are shown in Figure 9. It looks clear that the clusters capture different couplings of dry and wet underground antecedent conditions.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023



562

563 **Figure 9. Histograms for data distributions of θ_{100} (left) and h_a (right) according**
564 **to each identified cluster in the duplets $\langle \theta_{100}, h_a \rangle$**

565 In fact, cluster 1 captures dry conditions, with a volumetric water content below
566 the field capacity θ_{fc} (it was estimated as 0.35 with the empirical relationship
567 proposed by Twarakavi et al. (2009) according to the van Genuchten model
568 parameters) and low values of h_a. Differently, clusters 2 and 3 capture scenarios
569 related to relatively wet soil cover conditions (i.e., $\theta_{100} > \theta_{fc}$), coupled to low h_a
570 in cluster 3, gathering scenarios normally observed in late autumn, and to the
571 highest h_a conditions for cluster 2, comprising conditions normally occurring in
572 late winter and spring.

573 The two chosen variables, θ_{100} and h_a, allow identifying three different
574 antecedent slope conditions one hour before the onset of any rainfall event.
575 Hence, it is worthy to investigate how these different antecedent conditions may
576 be related to different slope responses to precipitations.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

3.3. Effects of underground antecedent conditions on the slope response to rainfall

The analysis of the data has been focused on identifying clusters within the triplets $\langle \theta_{100}, h_a, \Delta S/H \rangle$, aiming to evaluate the slope response as the amount of rainwater being stored/drained in the soil cover. The results are being plotted in the space composed by the variables that can be monitored in the field: (θ_{100}, h_a, H).

As it is not always expected to experience increased soil storage during rainfall events, the identification of draining slope conditions is an important aspect.

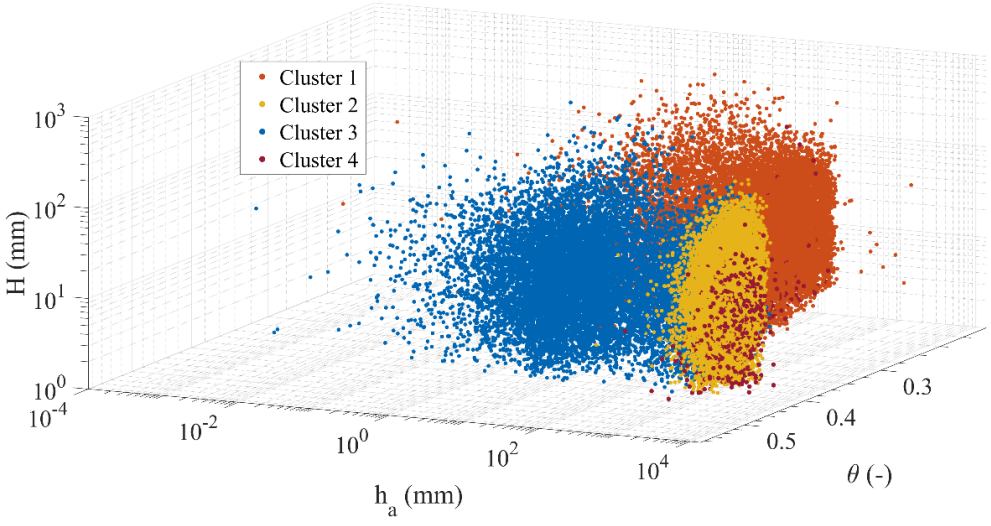


Figure 10. Clustering results of the synthetic data triplets $\langle \theta_{100}, h_a, \Delta S/H \rangle$ represented in the space (θ_{100}, h_a, H)

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

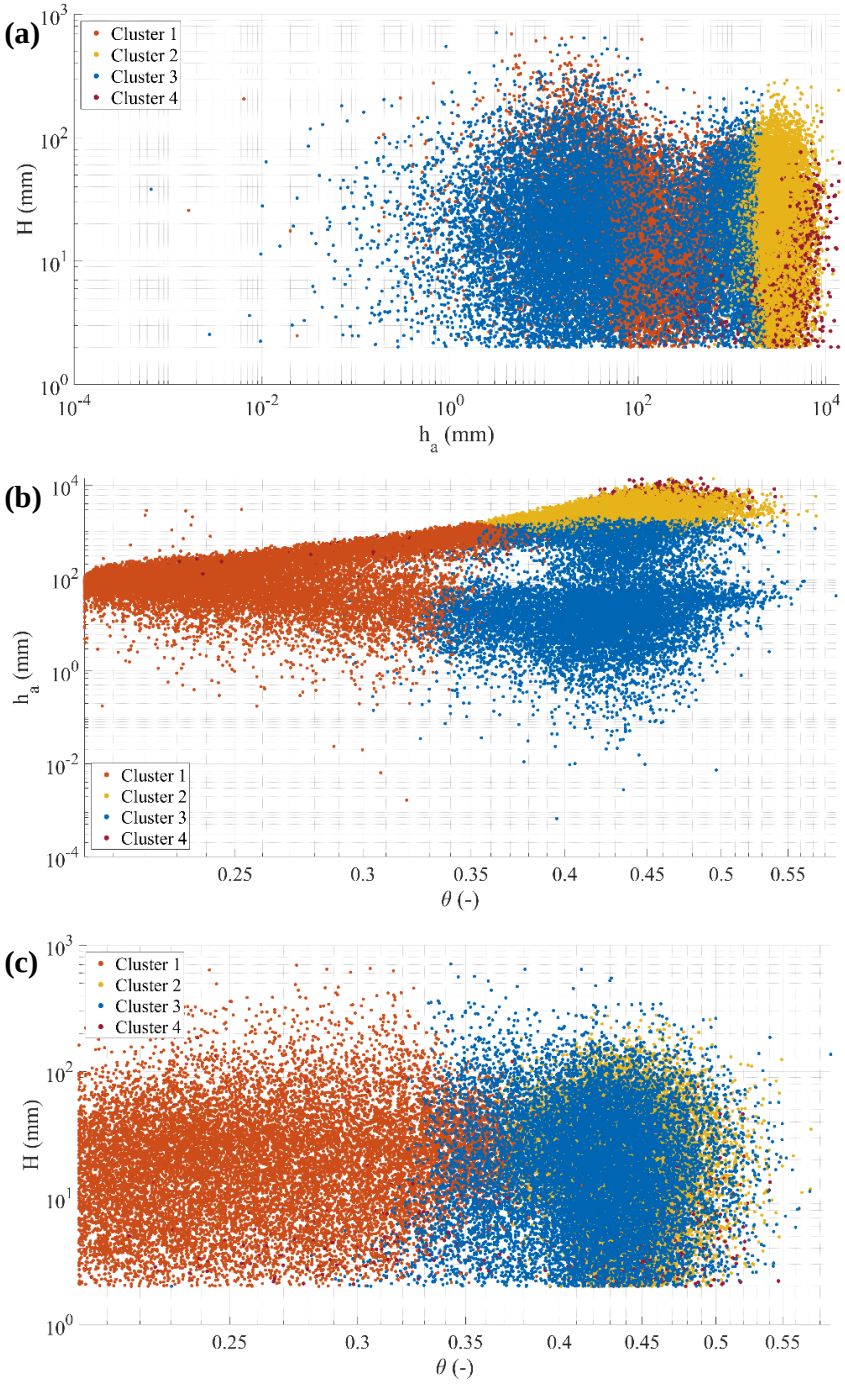


Figure 11. Clustering results of the triplets $\langle \boldsymbol{\theta_{100}}, \mathbf{h_a}, \mathbf{\Delta S/H} \rangle$ in (a) $(\boldsymbol{\theta_{100}}, \mathbf{h_a})$ plane; (b) $(\boldsymbol{\theta_{100}}, \mathbf{H})$ plane; (c) $(\mathbf{H}, \mathbf{h_a})$ plane

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

Figure 10 and Figure 11 show the data clusters for the triplets $\langle \theta_{100}, h_a, \Delta S/H \rangle$, for any identified rainfall event, represented in the (θ_{100}, h_a, H) space in a logarithmic axis representation. The Silhouette metric in this case suggests 4 as an optimal number of clusters with a metric value of 0.61. It is remarkable that three of the clusters are close to those already identified from the antecedent (seasonally recurrent) underground conditions (section 3.2).

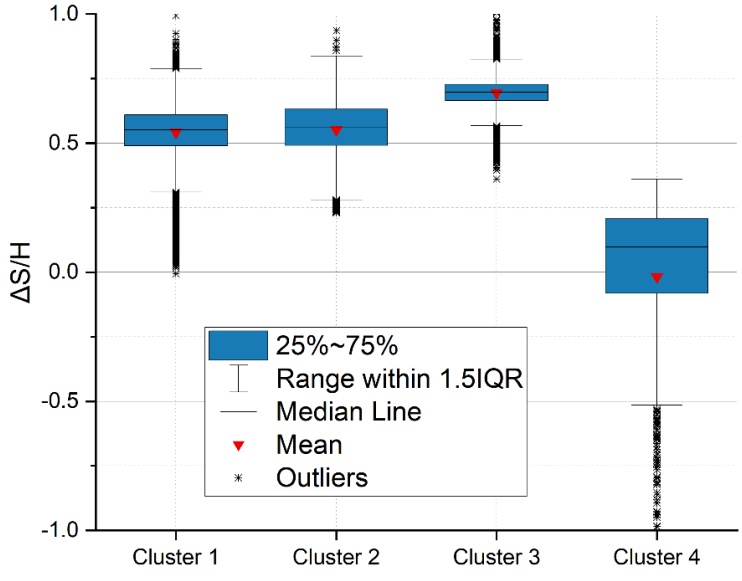


Figure 12. Distribution of the slope response $\Delta S/H$ for the data in each cluster

Specifically, cluster 1, 2 and 3 correspond to different slope processes according to $\Delta S/H$ (Figure 12). Even if cluster 1 and cluster 2 show similar responses, with slightly smaller $\Delta S/H$ for cluster 1, the controlling processes are indeed different; the conditions of cluster 1 are typically occurring in dry seasons with long dry periods between short rainfall events, leading to dry antecedent conditions, so that accumulation of water in the soil cover (increase in water storage) is expected at each event. The data in cluster 2 are typically related to wet seasons, especially in late winter and spring, where rainfall events are more frequent, leading to antecedent wet soil ($\theta_{100} \geq \theta_{fc}$) and antecedent high ground water level.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

However, these conditions do not seem to correspond to effective slope drainage, so that the slope response in cluster 2 results comparable to that observed in cluster 1 in terms of $\Delta S/H$. Instead, the conditions gathered in cluster 3 differ from those in cluster 2 for the lower aquifer water level h_a, and the highest $\Delta S/H$ indicates the lowest slope drainage.

The additional cluster 4 identified here highlights a particular slope response, as it catches all the conditions where nearly zero and negative ΔS take place, meaning an effective slope drainage during rainfall events. It is interesting to note that, even for relatively high rainfall events (above 100 mm), this slope response occurs when soil moisture is above the field capacity and when this condition is coupled with very high groundwater level, probably due to the high permeability all along the soil cover and to the hydraulic connection with the underlying aquifer.

4. Conclusions

This study aims at identifying and analysing the major hydrological controls of the slope response to precipitations and, in that way, defining suitable variables to be monitored in the field to predict such response. The studied case refers to the hydrological processes in a slope system consisting of a pyroclastic soil cover overlaying a fractured karstic bedrock, where a perched aquifer develops during the rainy season. A synthetic time series of slope response to precipitations has been built, thanks to a physically based model, previously calibrated with field monitoring data, coupled with a stochastic rainfall generator. The seasonal behaviour of the slope shows substantial agreement between synthetic and experimental data. The synthetic dataset has been explored with Random Forest and k-means clustering, to evaluate the slope response characterized as the change in water stored in the soil cover (ΔS) during precipitation events with rainfall depth H, starting from different underground antecedent conditions. These were quantified through the mean volumetric water content in the

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

uppermost meter of soil cover (θ_{100}) and the aquifer water level (h_a), one hour before the onset of rainfall.

The ratio $\Delta S/H$, which allows identifying slope response regardless the amount of event precipitation, is sensitive to both h_a and θ_{100}, with the groundwater level being the most influential antecedent variable. The underground antecedent conditions, characterized by θ_{100} and $h_a$and linked to the seasonal meteorological forcing, allow identifying different slope responses, related to the seasonally active hydrological processes.

High perched groundwater level, typical of winter and spring, indicates active slope drainage, which compensates rainwater infiltration, so that the soil storage remains stable, or even reduces, even after large rainfall events.

Differently, low perched groundwater level corresponds to impeded slope drainage. When it occurs with initially dry soil cover (typically in summer and early autumn), it tends to retain all the infiltrated rainwater as increased soil storage. When the soil cover is already wet (i.e., above the field capacity) at the onset of rainfall events, as it usually happens in late autumn and early winter, the increase of soil storage is smaller, as the soil approaches saturation.

The presented results suggest that monitoring antecedent conditions, by measuring suitable variables to identify the major hydrological processes occurring in the slope in response to precipitations, can be useful to understand such processes and to develop effective predictive models of slope response. Therefore, the proposed methodology can be replicated also in other contexts and be useful for several hydrologic applications: from the water supply towards natural streams due to infiltrated water, to the hydric stress estimation in crops (e.g., the centenary chestnut forests in the area of the case study) especially in very dry seasons, but also for the design of effective monitoring networks

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

664 exploiting geohydrological information for geohazard prevention (and early
665 waring).

666

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

Author contributions

RG and DR formulated the research aim; PM provided the field measurements; PM and GS supplied the model simulations; DR and GS curated and analyzed the data; RG oversighted the research activities; DR worked on the preparation and the data visualization.

Acknowledgements

This research is part of the Ph.D. project entitled "Hydrological controls and geotechnical features affecting the triggering of shallow landslides in pyroclastic soil deposits" within the Doctoral Course "A.D.I." of Università degli Studi della Campania "L. Vanvitelli".

The research has been also funded by Università degli Studi della Campania 'L. Vanvitelli' through the programme "VALERE: VAnviteLli pEr la RicErca".

References

Allocca, V., Manna, F., and De Vita, P.: Estimating annual groundwater recharge coefficient for karst aquifers of the southern Apennines (Italy), Hydrol Earth Syst Sci, 18, 803–817, https://doi.org/10.5194/hess-18-803-2014, 2014.

Arthur, D. and Vassilvitskii, S.: k-means++: The Advantages of Careful Seeding, in: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 1027–1035, 2007.

Bogaard, T. A. and Greco, R.: Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds, Natural Hazards and Earth System Sciences, 18, 31–39, https://doi.org/10.5194/nhess-18-31-2018, 2018.

Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdisciplinary Reviews: Water, 3, 439–459, https://doi.org/10.1002/wat2.1126, 2016.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

693

694 Bordoni, M., Meisina, C., Valentino, R., Lu, N., Bittelli, M., and Chersich, S.:
695 Hydrological factors affecting rainfall-induced shallow landslides: From the field
696 monitoring to a simplified slope stability analysis, Eng Geol,
697 https://doi.org/10.1016/j.enggeo.2015.04.006, 2015.

698 Breiman, L.: Random Forests, Mach Learn, 45, 5–32,
699 https://doi.org/https://doi.org/10.1023/A:1010933404324, 2001.

700 Capretti, P. and Battisti, A.: Water stress and insect defoliation promote the
701 colonization of Quercus cerris by the fungus Biscogniauxia mediterranea, For
702 Pathol, 37, 129–135, https://doi.org/10.1111/J.1439-0329.2007.00489.X, 2007.

703 Cascini, L., Cuomo, S., and Guida, D.: Typical source areas of May 1998 flow-
704 like mass movements in the Campania region, Southern Italy, Eng Geol, 96, 107–
705 125, https://doi.org/10.1016/j.enggeo.2007.10.003, 2008.

706 Celico, F., Naclerio, G., Bucci, A., Nerone, V., Capuano, P., Carcione, M.,
707 Allocca, V., and Celico, P.: Influence of pyroclastic soil on epikarst formation:
708 A test study in southern Italy, Terra Nova, 22, 110–115,
709 https://doi.org/10.1111/J.1365-3121.2009.00923.X, 2010.

710 Comegna, L., Damiano, E., Greco, R., Guida, A., Olivares, L., and Picarelli, L.:
711 Field hydrological monitoring of a sloping shallow pyroclastic deposit, Canadian
712 Geotechnical Journal, 53, 1125–1137, https://doi.org/10.1139/cgj-2015-0344,
713 2016.

714 Cowpertwait, P. S. P., O'Connell, P. E., Metcalfe, A. V., and Mawdsley, J. A.:
715 Stochastic point process modelling of rainfall. I. Single-site fitting and validation,
716 J Hydrol (Amst), https://doi.org/10.1016/S0022-1694(96)80004-7, 1996.

717 Damiano, E. and Olivares, L.: The role of infiltration processes in steep slope
718 stability of pyroclastic granular soils: laboratory and numerical investigation,

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

719 Natural Hazards, 52, 329–350, https://doi.org/10.1007/s11069-009-9374-3,
720 2010.

721 Damiano, E., Olivares, L., and Picarelli, L.: Steep-slope monitoring in
722 unsaturated pyroclastic soils, Eng Geol, 137–138, 1–12,
723 https://doi.org/10.1016/j.enggeo.2012.03.002, 2012.

724 Damiano, E., Greco, R., Guida, A., Olivares, L., and Picarelli, L.: Investigation
725 on rainwater infiltration into layered shallow covers in pyroclastic soils and its
726 effect on slope stability, Eng Geol, 220, 208–218,
727 https://doi.org/10.1016/j.enggeo.2017.02.006, 2017.

728 de Amorim, R. C. and Hennig, C.: Recovering the number of clusters in data sets
729 with noise features using feature rescaling factors, Inf Sci (N Y), 324, 126–145,
730 https://doi.org/10.1016/J.INS.2015.06.039, 2015.

731 De Vita, P., Agrello, D., and Ambrosino, F.: Landslide susceptibility assessment
732 in ash-fall pyroclastic deposits surrounding Mount Somma-Vesuvius:
733 Application of geophysical surveys for soil thickness mapping, J Appl Geophy,
734 59, 126–139, https://doi.org/10.1016/j.jappgeo.2005.09.001, 2006.

735 Feddes, R. A., Kowalik, P., Kolinska-Malinka, K., and Zaradny, H.: Simulation
736 of field water uptake by plants using a soil water dependent root extraction
737 function, J Hydrol (Amst), 31, 13–26, https://doi.org/10.1016/0022-
738 1694(76)90017-2, 1976.

739 Fiorillo, F., Guadagno, F., Aquino, S., and De Blasio, A.: The December 1999
740 Cervinara landslides: Further debris flows in the pyroclastic deposits of
741 Campania (Southern Italy), Bulletin of Engineering Geology and the
742 Environment, https://doi.org/10.1007/s100640000093, 2001.

743 Gao, S. and Shain, L.: Effects of water stress on chestnut blight, Canadian journal
744 of forest research, 25, 1030–1035, 1995.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

Greco, R. and Gargano, R.: A novel equation for determining the suction stress of unsaturated soils from the water retention curve based on wetted surface area in pores, Water Resour Res, 51, 6143–6155, https://doi.org/10.1002/2014WR016541, 2015.

Greco, R., Comegna, L., Damiano, E., Guida, A., Olivares, L., and Picarelli, L.: Hydrological modelling of a slope covered with shallow pyroclastic deposits from field monitoring data, Hydrol Earth Syst Sci, 17, 4001–4013, https://doi.org/10.5194/hess-17-4001-2013, 2013.

Greco, R., Comegna, L., Damiano, E., Guida, A., Olivares, L., and Picarelli, L.: Conceptual Hydrological Modeling of the Soil-bedrock Interface at the Bottom of the Pyroclastic Cover of Cervinara (Italy), Procedia Earth and Planetary Science, https://doi.org/10.1016/j.proeps.2014.06.007, 2014.

Greco, R., Marino, P., Santonastaso, G. F., and Damiano, E.: Interaction between Perched Epikarst Aquifer and Unsaturated Soil Cover in the Initiation of Shallow Landslides in Pyroclastic Soils, Water (Basel), 10, 948, https://doi.org/10.3390/w10070948, 2018.

Greco, R., Comegna, L., Damiano, E., Marino, P., Olivares, L., and Santonastaso, G. F.: Recurrent rainfall-induced landslides on the slopes with pyroclastic cover of Partenio Mountains (Campania, Italy): Comparison of 1999 and 2019 events, Eng Geol, 288, 106160, https://doi.org/10.1016/j.enggeo.2021.106160, 2021.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Reviews of Geophysics, 52, 218–242, https://doi.org/10.1002/2013RG000443, 2014.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

769 Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning
770 Data Mining, Inference, and Prediction, 2nd ed., Springer Series on Statistics,
771 Stanford, 1–763 pp., 2008.

772 Lloyd, S. P.: Least Squares Quantization in PCM, IEEE Trans Inf Theory, 28,
773 1982.

774 Lu, N. and Likos, W. J.: Suction Stress Characteristic Curve for Unsaturated Soil,
775 Journal of Geotechnical and Geoenvironmental Engineering,
776 https://doi.org/10.1061/(asce)1090-0241(2006)132:2(131), 2006.

777 Marino, P., Comegna, L., Damiano, E., Olivares, L., and Greco, R.: Monitoring
778 the Hydrological Balance of a Landslide-Prone Slope Covered by Pyroclastic
779 Deposits over Limestone Fractured Bedrock, Water (Basel), 12, 3309,
780 https://doi.org/10.3390/w12123309, 2020a.

781 Marino, P., Santonastaso, G. F., Fan, X., and Greco, R.: Prediction of shallow
782 landslides in pyroclastic-covered slopes by coupled modeling of unsaturated and
783 saturated groundwater flow, Landslides, https://doi.org/10.1007/s10346-020-
784 01484-6, 2020b.

785 McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb,
786 T., Plaut, J., Sperry, J., West, A., Williams, D. G., and Yepez, E. A.: Mechanisms
787 of plant survival and mortality during drought: Why do some plants survive while
788 others succumb to drought?, New Phytologist, 178, 719–739,
789 https://doi.org/10.1111/J.1469-8137.2008.02436.X, 2008.

790 Neyman, J. and Scott, E. L.: Statistical Approach to Problems of Cosmology,
791 Journal of the Royal Statistical Society: Series B (Methodological),
792 https://doi.org/10.1111/j.2517-6161.1958.tb00272.x, 1958.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

Nieber, J. L. and Sidle, R. C.: How do disconnected macropores in sloping soils facilitate preferential flow?, Hydrol Process, 24, 1582–1594, https://doi.org/10.1002/hyp.7633, 2010.

Olivares, L. and Picarelli, L.: Shallow flowslides triggered by intense rainfalls on natural slopes covered by loose unsaturated pyroclastic soils, Geotechnique, https://doi.org/10.1680/geot.2003.53.2.283, 2003.

Pagano, L., Picarelli, L., Rianna, G., and Urciuoli, G.: A simple numerical procedure for timely prediction of precipitation-induced landslides in unsaturated pyroclastic soils, Landslides, 7, 273–289, https://doi.org/10.1007/s10346-010-0216-x, 2010.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Ottlé, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, Hydrol. Earth Syst. Sci, 24, 1485–1509, https://doi.org/10.5194/hess-24-1485-2020, 2020.

Paulik, C., Dorigo, W., Wagner, W., and Kidd, R.: Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network, International Journal of Applied Earth Observation and Geoinformation, 30, 1–8, https://doi.org/10.1016/J.JAG.2014.01.007, 2014.

Peres, D. J. and Cancelliere, A.: Derivation and evaluation of landslide-triggering thresholds by a Monte Carlo approach, Hydrol Earth Syst Sci, 18, 4913–4931, https://doi.org/10.5194/hess-18-4913-2014, 2014.

Pirone, M., Papa, R., Nicotera, M. V., and Urciuoli, G.: Soil water balance in an unsaturated pyroclastic slope for evaluation of soil hydraulic behaviour and

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

818 boundary conditions, J Hydrol (Amst),
819 https://doi.org/10.1016/j.jhydrol.2015.06.005, 2015.

820 Ponce, V. M. and Hawkins, R. H.: Runoff Curve Number: Has It Reached
821 Maturity?, J Hydrol Eng, 1, 11–19, https://doi.org/10.1061/(ASCE)1084-
822 0699(1996)1:1(11), 1996.

823 Revellino, P., Guerriero, L., Gerardo, G., Hungr, O., Fiorillo, F., Esposito, L.,
824 and Guadagno, F. M.: Initiation and propagation of the 2005 debris avalanche at
825 Nocera Inferiore (Southern Italy), Italian Journal of Geosciences,
826 https://doi.org/10.3301/IJG.2013.02, 2013.

827 Richards, L. A.: Capillary conduction of liquids through porous mediums, J Appl
828 Phys, https://doi.org/10.1063/1.1745010, 1931.

829 Rodriguez-Iturbe, I., Febres De Power, B., and Valdes, J. B.: Rectangular pulses
830 point process models for rainfall: analysis of empirical data, J Geophys Res,
831 https://doi.org/10.1029/JD092iD08p09645, 1987.

832 Rolandi, G., Bellucci, F., Heizler, M. T., Belkin, H. E., and De Vivo, B.: Tectonic
833 controls on the genesis of ignimbrites from the Campanian Volcanic Zone,
834 southern Italy, Mineral Petrol, 79, 3–31, https://doi.org/10.1007/s00710-003-
835 0014-4, 2003.

836 Rousseeuw, P. J.: Silhouettes: A graphical aid to the interpretation and validation
837 of cluster analysis, J Comput Appl Math, 20, 53–65,
838 https://doi.org/10.1016/0377-0427(87)90125-7, 1987.

839 Shuttleworth, W. J.: Evaporation, in: Handbook of Hydrology, edited by:
840 Maidment, D. R., McGraw-Hill, New York, NY, USA, 1993.

841 Tromp-Van Meerveld, H. J. and McDonnell, J. J.: Threshold relations in
842 subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope, Water
843 Resour Res, 42, 2410, https://doi.org/10.1029/2004WR003778, 2006.

https://doi.org/10.5194/egusphere-2022-1078
Preprint. Discussion started: 16 January 2023

844 Tufano, R., Formetta, G., Calcaterra, D., and De Vita, P.: Hydrological control
845 of soil thickness spatial variability on the initiation of rainfall-induced shallow
846 landslides using a three-dimensional model, Landslides,
847 https://doi.org/10.1007/s10346-021-01681-x, 2021.

848 Twarakavi, N. K. C., Sakai, M., and Šimůnek, J.: An objective analysis of the
849 dynamic nature of field capacity, Water Resour Res, 45,
850 https://doi.org/10.1029/2009WR007944, 2009.

851 van Genuchten, M. Th.: A Closed-form Equation for Predicting the Hydraulic
852 Conductivity of Unsaturated Soils1, Soil Science Society of America Journal, 44,
853 892, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

854 Wicki, A., Lehmann, P., Hauck, C., Seneviratne, S. I., Waldner, P., and Stähli,
855 M.: Assessing the potential of soil moisture measurements for regional landslide
856 early warning, Landslides, 17, 1881–1896, https://doi.org/10.1007/S10346-020-
857 01400-Y/FIGURES/11, 2020.

858 Williams, P. W.: The role of the epikarst in karst and cave hydrogeology: a
859 review, Int J Speleol, 37, 1–10, https://doi.org/10.5038/1827-806X.37.1.1, 2008.

860