Accounting for Soils Parameter Uncertainty in a Physically-Based and Distributed Approach for Rainfall-Triggered Landslides

E. Arnone[1], Y. G. Dialynas[2], L. V. Noto[1], R. L. Bras[2]

[1] Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali, Università degli Studi di Palermo, Palermo, ITALY

[2] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Correspondence to: E. Arnone (elisa.arnone@unipa.it)

Abstract

In this study we propose a probabilistic approach for coupled distributed hydrological-hillslope stability models that accounts for soil parameters uncertainty at basin scale. The geotechnical and soil retention curve parameters are treated as random variables across the basin and theoretical probability distributions of the Factor of Safety (FS) are estimated. The derived distributions are used to obtain the spatio-temporal dynamics of probability of failure, in terms of parameters uncertainty, conditioned to soil moisture dynamics. The framework has been implemented in the tRIBS-VEGGIE (Triangulated Irregular Network (TIN)-based Real-time Integrated Basin Simulator - VEGetation Generator for Interactive Evolution) - Landslide model and applied to a basin in the Luquillo Experimental Forest (Puerto Rico) where shallow landslides are common. In particular, the methodology was used to evaluate how the spatial and temporal patterns of precipitation, whose variability is significant over the basin, affect the distribution of probability of failure, through event scale analyses. Results indicate that hyetographs where heavy precipitation is near the end of the event lead to the most critical conditions in terms of probability of failure.

Keywords: Landslides; Rainfall; Distributed Hydrological Modeling; Parameters Uncertainty; Probability of Failure; Puerto Rico.

1 Introduction

Rainfall-triggered landslides are among the most common types of landslides, which every year cause fatalities, damage to properties and economic losses all over the world. Methods available in the scientific literature for rainfall-triggered hazards analyses are various and can be empirical, statistical or deterministically based. The first two types of methods can be included within a broader category in which the relation between rainfall and landslide hazard is analyzed based on past observations and lead to the derivation of landslide susceptible areas or rainfall thresholds (e.g. Glade, 1998; Guzzetti et al., 1999; Guzzetti et al., 2007; Ponziani et al., 2013; Stefanini, 2004). Whereas, a class of coupled distributed hydrological-stability models, to which our approach belongs, can be used to evaluate, dynamically, the risk of shallow rainfall-triggered landslide hazards at catchment scale (deterministic approach). In fact, failures result from interdependent spatio-temporal dynamics which include many of the hydrological processes (rainfall, evapotranspiration, infiltration, etc..) as well as vegetation surcharge, root strength, soil moisture conditions. The practice consists in estimating a spatially distributed Factor of Safety (FS) as a function of the basin hydrological response, which is evaluated in terms of soil moisture and groundwater fields (Arnone et al., 2011;Burton and Bathurst, 1998;Capparelli et al., 2002;Montgomery and Dietrich, 1994;Rosso et al., 2006;Simoni et al., 2008;Wu and Sidle, 1995). Mechanical and hydrological soil properties play a crucial role in such an evaluation, and the importance of appropriately modeling soil water dynamics has been clearly demonstrated in some studies (Lanni et al., 2009;Lepore et al., 2013).

A limitation of using physically-based and spatial distributed models is the high numbers of model parameters whose reliable estimation is not always possible in a natural catchment (Beven, 1993;Beven and Binley, 1992). The inability to fully characterize hydrological and geotechnical behavior of soil may have a significant impact on model results; indeed, the uncertainty in model parameter evaluation has been recognized as an important cause of the mismatch between simulated and observed distributions of landslides across the catchment (Burton et al., 1998).

Burton et al., 1998 provides an in depth analysis on the effect of the spatial variation of soil parameters (i.e. soil strength, soil depth and slope) on landslide occurrence, starting from an accurate dataset of parameters with high frequency of spatial measurements. The analysis, conducted on a selected area of a catchment, has provided interesting insights into the

statistical relationships among the parameters and has demonstrated that the effects of spatial variability of model parameters can be significant.

To account for this uncertainty, FS can be computed within a probabilistic framework, by considering soil parameters as random variables with given probability distributions. This practice has received considerable attention in the geotechnical engineering literature where we find different methodologies for modeling and analyzing the uncertainty related to the shear strength parameters (i.e. soil cohesion and friction angle) at hillslope scale (Abbaszadeh et al., 2011;Ray and Baidya, 2011;Malkawi et al., 2000). Based on similar approaches, some studies have been conducted for basin scale applications using coupled hydrological-stability models (Frattini et al., 2009;Melchiorre and Frattini, 2012;Pack et al., 1998;Simoni et al., 2008). In those applications, the probability of FS, conditioned on the soil moisture, is dynamically estimated across the basin, whereas the probability distributions of the shear strength parameters are time independent.

However, the uncertainty of soil hydrological properties, which may be predominant in the case of unsaturated conditions and at basin scale, is still neglected in most published literature. Of particular interest and importance are the soil retention curve parameters that are most significant in determining the soil matric suction contribution to the stability of slopes.

If one presumes that soil parameters are random variables then the probability distribution of FS could potentially be derived numerically, analytically, or through analytical approximations. The Monte Carlo simulation method uses independent sets of soil properties, generated through a priori assigned probability distributions (Abbaszadeh et al., 2011;Malkawi et al., 2000) at fixed topographic (i.e. slope) and hydrological (i.e. soil moisture) conditions to obtain a solution. However, such an approach may have significant computational cost for basin scale applications, since the conditions change in time and space. The FS probability distribution can be analytically derived in the case where only geotechnical parameters (i.e., cohesion and friction angle) are considered as time independent random variables and the infinite slope model is used for the slope stability analysis (Frattini et al., 2009;Simoni et al., 2008). When all soil retention curve parameters are also assumed to be random, analytical derivation of FS distribution is not tractable. Then, the First Order Second Moment (FOSM) method (Benjamin and Cornell, 1970) is commonly used to estimate analytical approximations of the spatio-temporal FS statistics (i.e. mean and variance), that can be used to fit a theoretical probability distribution for FS and estimate the

spatio-temporal dynamics of probability of failure. To our knowledge, a first example of using the FOSM method in a hydrological-stability model is given by Arnone et al. (2014). In that effort, the lognormal distribution is used to describe the FS occurrences. However, when the function is highly nonlinear as the case of the FS equation, a better approximation to the mean can be obtained by using higher order terms in the Taylor series expansion of the FS expression.

In this paper we propose an advanced methodology to take into account the uncertainty associated to the soil parameters in the computation of FS, by using higher order expansions of the non linear FS expression, thus improving on the results of Arnone et al., (2014). The parameters of the FS equation are thus treated as random variables and consequently the associated distributions of FS are estimated. Parameters of the soil water retention curve are treated as random variables as well as the mechanical parameters. However, the uncertainty derived from the spatial variability across the same soil type is not taken into account. The framework is implemented in the tRIBS-VEGGIE (Triangulated Irregular Network (TIN)-based Real-time Integrated Basin Simulator - VEGetation Generator for Interactive Evolution) - Landslide module (Lepore et al., 2013). The proposed methodology is applied to the Rio Mameyes Basin, located in the Luquillo Experimental Forest in Puerto Rico that is particularly susceptible to shallow landslides due to both the morphology and the numerous intense rainfall events that hit the area. Precipitation at the site is strongly correlated to the topography (Daly et al., 2003;Garcia-Martino et al., 1996) and the extreme events are also characterized by different types of temporal distributions (Bonnin et al., 2006). Such spatial and temporal patterns may have different impacts on the hydrologic response (Ogden and Julien, 1993;Singh, 1997;Wilson et al., 1979) and hence may initiate landslides events differently (D'Odorico et al., 2005). The basin has been the subject of previous landslides analysis with the tRIBS-VEGGIE-Landslide model in Lepore et al., 2013 work, which we will refer to for the description of most of the data. In addition to the description and implementation of the new probabilistic framework of tRIBS-VEGGIE-Landslide model, this paper also aims to assess the effects of different temporal distribution of precipitation on the hillslope stability and on the consequent landslide hazards across the basin.

2 Methods

2.1 Coupled Hydrological-Stability Model

The methodology is built around the tRIBS-VEGGIE-Landslide model (Lepore et al., 2013) which was developed by embedding a limit equilibrium analysis into the eco-hydrological model tRIBS-VEGGIE (Ivanov et al., 2008).The model inherits from tRIBS the capability to simulate most of the hydrological processes (e.g. infiltration, evapotranspiration, interception, lateral redistribution, soil moisture dynamics), by explicitly considering the spatial variability of land-surface properties as well as in precipitation field. Topography is described by means of an irregular triangulated spatial mesh which allows for the use of variable computational elements and thus for the increase in accuracy only in the most critical areas of the basin (Vivoni et al., 2004). The vegetation module (VEGGIE) simulates the plant physiology, and in particular the biophysical energy processes (e.g. transpiration), biophysical hydrologic processes (e.g. unsaturated zone flow) and biochemical processes (e.g. photosynthesis, plant respiration) (Ivanov et al., 2008). A detailed description of model can be found in (Ivanov et al., 2008). For understanding this paper it is nevertheless important to review the following elements of tRIBS-VEGGIE-Landslide:

- The infiltration module is based on a numerical approximation of the one dimensional Richards' equation (Hillel, 1980), which provides the moisture transfer in the subsurface within an element. Subsurface and surface moisture is then laterally redistributed among the elements along the direction of steepest descent in a rate depending on the unsaturated hydraulic conductivity of the receiving cell.
- As a consequence of the Richard's equation resolution scheme, soil moisture is estimated in a multi-layer scheme parallel to the slope surface, with a number of layers equal to 25.
- The Brooks and Corey (1964) (*BC*) parameterization scheme is used to model the soil retention curve and the unsaturated hydraulic conductivity, as a function of saturated hydraulic conductivity in the normal to the soil surface direction, air entry bubbling pressure, and pore-size distribution index.
- In this work only the effects of vegetation on soil moisture are considered, not the mechanical effects.

The coupled model is capable to dynamically compute the factor of safety, $FS(t)$, across a basin as a response of the soil moisture dynamics, by applying the infinite slope analysis. The implemented equation is the following (Lepore et al., 2013):

$$FS(t) = \frac{c'}{\gamma_s z_n \sin\alpha} + \frac{\tan\phi}{\tan\alpha} + \frac{\gamma_w \psi_b}{\gamma_s z_n} \cdot \left(\frac{\theta(t) - \theta_r}{\theta_s - \theta_r} \right)^{1-\frac{1}{\lambda}} \cdot \frac{\tan\phi}{\sin\alpha} \quad (1)$$

where $FS(t)$ is the time dependent factor of safety (hereinafter simply FS), c' is the effective soil cohesion, γ_s is the total unit weight of soil (varying with soil moisture), γ_w is the water unit weight, z_n is the soil depth measured along the normal direction to the slope; α is the slope angle, ϕ is the soil friction angle, ψ_b is the air entry bubbling pressure, λ is the pore-size distribution index, $\theta(t)$ is the time depended volumetric water content (hereinafter simply θ), and θ_r and θ_s are the residual and saturated soil moisture contents, respectively. ψ_b, λ, θ_r and θ_s are the Brooks-Corey equation parameters used to represent the soil retention curve. As a result of the multi-layer representation of soil moisture, the final products of the module are dynamic maps of instability areas as well as dynamic FS profiles at selected areas.

2.2 Probabilistic Framework

A probabilistic framework was developed to take into account the soil parameters uncertainty that characterizes slope stability analyses. The methodology consists of i) treating the soil hydrological and geotechnical parameters of the FS equation (Eq. 1) as random variables, ii) estimating the FS moments through the FOSM method and, finally, iii) computing the FS distribution by fitting a theoretical distribution to the estimated moments. The framework upgrades the approach presented in Arnone et al. (2014) by improving the accuracy of the approximation and by using a more appropriate theoretical probability distribution for the Factor of Safety.

The FOSM method is the most widely used approximate method in engineering design for the analytical estimation of the mean and variance of a random function and it is based on a Taylor series expansion. Consider a function of variables, $X_1,\ldots,X_n$: $Y=y(X_1,\ldots,X_n)$. The approximation used here expands up to second order terms in the Taylor series, as opposed to Arnone et al., (2014), resulting in:

$$E[Y] = y(\mu_{X1},\ldots,\mu_{Xn}) + \frac{1}{2}\sum_{i=1}^{n} Var[X_i]\frac{d^2y}{dX_i^2} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov[X_i, X_j]\frac{d^2y}{dX_i dX_j} \quad (2)$$

where E[·] denotes the expectation, $Var[\cdot]$ denotes the variance, $Cov[X_i, X_j]$ denotes the covariance between X_i and X_j and n is the number of random variables. Eq. (2) has been used in engineering studies for the estimation of the mean *FS* (Abbaszadeh et al., 2011).

The variance of the random function Y is approximated by the FOSM method as following:

$$Var[Y] = E\left[\left(\sum_{i=1}^{n}\left(X_i - \mu_{X_i}\right)\frac{\partial y}{\partial X_i}\right)^2\right] = \sum_{i=1}^{n}\left(\frac{\partial y}{\partial X_i}\right)^2 Var[X_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{\partial y}{\partial X_i}\right)\left(\frac{\partial y}{\partial X_j}\right)Cov[X_i, X_j] \quad (3)$$

Derivatives of y in Eqs. (2) and (3) are estimated around the mean values μ_{Xi}.

Given the marginal (means and variances) and joint statistics (covariances) of the six assumed random variables (c, ϕ, θ_r, θ_s, ψ_b, and λ), Eqs. (2) and (3) analytically approximate the mean, and variance of *FS,* $E[FS]$ and $Var[FS]$ respectively. Eq. (1) is continuously differentiable with respect to each assumed random variable, a condition required for the analytical implementation of the FOSM method. First, second and mixed derivatives of Eq. (1) with respect to each of the six parameters were analytically derived for a total of 15 derivatives. In particular, we obtained 6 first derivatives for the 6 variables, 3 second derivatives for θ_r, θ_s, and λ (the others are null) and 6 mixed derivatives of the correlated random variables (i.e. θ_r-θ_s, θ_r-ψ_b, θ_r-λ, θ_s-ψ_b, θ_s-λ, ψ_b-λ). Each term is thus a function of the topographic characteristics and the soil moisture spatial and temporal dynamics (θ), leading the statistics $E[FS]$ and $Var[FS]$ that are dependent on both space and time, via changes of the moisture and topography.

Given the two FS moments, it is then possible to fit a two-parameter distribution to the factor of safety. Given the FS distribution it is possible to compute the soil moisture dependent probability of failure as the cumulative probability that FS is less than a given critical value, FS_{crit}, $\Pr[FS<FS_{crit}]$. This critical value is the one that delimits stable and unstable conditions, and rigorously it is equal to 1, based on the definition of FS (eq. 1). However, under conservative conditions, it could be set to values greater than 1 (e.g. 1.2). In this work we assumed FS_{crit} =1. Probability of failure can be thus evaluated across the basin and used to assess shallow landslide hazards. The best-fit theoretical *FS* distribution can be identified by

Monte-Carlo experiments with fixed hydrologic and topographic conditions (see Section 4). The normal distribution has been commonly used to describe the *FS* as a random variable (e.g., (Abbaszadeh et al., 2011;Malkawi et al., 2000;Simoni et al., 2008). However, in those studies only the geotechnical parameters were considered as random variables. Nevertheless, other studies (e.g. (Arnone et al., 2014;Frattini et al., 2009)) have found that the *FS* may exhibit significant positive skewness, with the tail of the distribution located on the right, with considerable probability mass concentrated at the left side of the distribution. In order to capture the asymmetry of the *FS* distribution, works by Arnone et al., 2014 and Frattini et al., 2009 used the lognormal distribution, that was verified for different soil formations.

Under saturated conditions, matric suction vanishes and *FS* is independent of the BC parameters, becoming a linear function of the normally distributed soil cohesion and friction angle. In this case, the *FS* distribution is analytically derived and the use of an approximate method is not required. More precisely, by applying the convolution integral (Feller, 1971) it can be shown that the *FS* becomes also a normally distributed random variable, and the *FS* statistics (i.e., $E[FS]$ and $Var[FS]$) can be analytically derived:

$$E[FS] = \frac{\mu_c}{\gamma_s h_s \sin\alpha} + \left(1 - \frac{\gamma_w h_w}{\gamma_s h_s}\right) \frac{\mu_{\tan\phi}}{\tan\alpha} \tag{4}$$

$$Var[FS] = \frac{\sigma_c^2}{(\gamma_s h_s \sin\alpha)^2} + \left(1 - \frac{\gamma_w h_w}{\gamma_s h_s}\right)^2 \frac{\sigma_{\tan\phi}^2}{(\tan\alpha)^2} \tag{5}$$

where μ_c, σ^2_c, and $\mu_{\tan\phi}$, $\sigma^2_{\tan\phi}$ are the statistics of cohesion and friction angle respectively, h_s the soil thickness and h_w the water level. Given the *FS* distribution, the probability of failure for saturated conditions is then computed as the cumulative probability that *FS* is less than 1:

$$\Pr(FS < 1) = \frac{1}{2}\left[1 + erf\left(\frac{1 - \mu_{FS}}{\sqrt{2}\sigma_{FS}}\right)\right] \qquad \textit{saturated conditions} \tag{6}$$

where *erf* is the error function.

The framework described above was implemented in the tRIBS-VEGGIE-Landslide model to dynamically evaluate the landslide hazard at basin scale. In particular, probabilities of failure are computed for each time step, voronoi element and soil vertical layer (*i*). The probability of *FS* being equal or lower than a critical value at the i^{th} layer (here referred to as event E_i,

corresponding to the Probability of Failure at the i^{th} layer, PrF_i) is computed, assuming both independent and not mutually exclusive events among different soil layers. This hypothesis implies that the occurrence of an event in a layer does not affect nor exclude the occurrence of another event in a different layer.

At each timestep, the model estimates the spatial map of the probability of landslide occurrence at each soil column, here referred as *Probability of Failure of the soil Column* (*PrFC*). The *PrFC* is equal to the probability of the union of events E_i for the entire soil column. Furthermore, at each soil column the probability of the landslide depth is dynamically estimated. The probability that the plane of failure is located at the i^{th} layer (here referred as *Probability of Plane of Failure at i^{th} layer*, $PrPF_i$) is given by the joint probability of *FS* being equal or lower than FS_{crit} at layer *i*, while there is no failure above that layer (mutually exclusive).

The landslide depth corresponding to the maximum value of $PrPF_i$ represents the most probable depth of failure. Thus, the model is able to evaluate when and where failure is most probable to occur within each soil column and across the basin.

3 Case Study

The island of Puerto Rico is located in the northeastern Caribbean and is the smallest island of the Greater Antilles. With its roughly rectangular shape (Figure 1), the island is characterized topographically by flat coastal areas and two mountain ranges, the Cordillera Central, which spans east to west with a peak of 1,338 m, and the Sierra de Luquillo which will be the focus of this study, with Pico del Este, or East Peak, at 1075 m. The Sierra de Luquillo was chosen for this study as it contains the *Luquillo Experimental Forest* (LEF) which is part of both the *Long-Term Ecological Research* (LTER) and of the *Critical Zone Observatory* (CZO) networks. This study will focus on the Rio Mameyes basin (hereto referred to as "Mameyes Basin"), in the northeast of the Sierra de Luquillo and within the LEF boundaries (Figure 1). The LEF has been a focal point for studies in landslide impacts on ecology, geomorphology, biology, disturbance and recovery of vegetation (Myster et al., 1997;Scatena and Lugo, 1995;Shiels et al., 2008;Walker and Shiels, 2008;Walker et al., 1996).

The Mameyes basin has an area of 16.7 km^2, and is characterized by a rapid change in elevation from 104.2 m to 1046 m across a horizontal distance of 3 km. An analysis of the slope distribution derived from a 30 m resolution US Geological Survey (USGS) National

Elevation Dataset (NED) DEM (*Digital Elevation Model*) (Figure 1a, c) reveals 10% of the basin area with slopes greater than 30° and 30% of the basin area with slopes greater than 25°. The climate of the island is controlled by the easterly trade winds from the Atlantic ocean and the pronounced topography (Garcia-Martino et al., 1996). The basin falls in the windward climatic region of the Island, making it one of the wettest basins on the island. Because of the strong gradient in elevation, rainfall, cloudiness and temperature vary consistently throughout the basin. The mean annual precipitation (MAP) varies from approximately 3000 mm, measured at an elevation of 352 m (Bisley Tower), to 5000 mm at higher elevations (Lepore et al., 2012).

The Luquillo forest is characterized by a mix of lower montane wet tropical, wet subtropical and rain forest (Ewel and Whitmore, 1973). There are three major forest types: the Tabonuco forest (*Dacryodes excelsa*) in the wet subtropical and subtropical rain forest life zones, typically within the 150-600 m elevation range; the Colorado forest (*Cyrilla racemiflora*) in the lower montane wet and rain forest life zones, within 600 and 900 m, and the Dwarf (cloud) forest, above 900 m (Waide et al., 1998).

The Mameyes basin has been previously modeled in terms of spatio-temporal dynamics of hillslope stability by Lepore et al., (2013) using the tRIBS-VEGGIE-Landside model. Therefore, most of the needed data, data description and model setup are described in detail in Lepore et al., 2013. For the sake of clarity, we briefly describe the main model data and parameters in the following section.

3.1 Data and model parameters

The tRIBS-VEGGIE-Landside model operates at hourly scale and the required model inputs are mainly meteorological data, soil data, and soil and ecological parameters.

The meteorological data used were obtained from the Bisley Tower located within the basin (lat. 18.31, long. 65.74, 352 m) (Figure 1a), which measures many of the needed input data at hourly resolution (wind speed and direction, air temperature, cloud cover, relative humidity, rainfall, incoming shortwave radiation).

The 30m DEM was also used to derive the Triangulated Irregular Network (TIN) that describe the topography within the model. The soil data were obtained from the USDA Forest Service's International Institute of Tropical Forestry of San Juan; in particular, the

USDA/NRCS (Huffaker, 2002) report provides a detailed description of the 12 soil complexes that characterize the basin. As described in Lepore et al., 2013, the 12 soil types were further simplified into four soil types (Figure 1b) according to the USDA soil classification. The bedrock is located at a depth of 8m or deeper (Simon et al., 1990) and is assumed not to affect a shallow slope failure mechanism.

The main hydrological soil properties (i.e. saturated hydraulic conductivity, anisotropy ratio) were obtained through a validation/confirmation procedure based on soil moisture data (Lepore et al., 2013). In particular, nine soil moisture hourly series measured at a 30 cm depth were used for the model confirmation. The measurements were taken at three locations, within an area close to the Bisley Tower and each with three time-domain reflectometry (TDRs) Campbell Scientific Model CS616 instruments (Lepore et al., 2013). tRIBS-VEGGIE defines the anisotropy ratio as the ratio of the saturated hydraulic conductivities in the directions parallel to the slope and normal to the slope and thus it is partially responsible of the lateral subsurface flux transfer. In this work, the anisotropy ratio is 100 (Lepore et al., 2013).

The BC soil retention parameters, i.e. θ_r, θ_s, ψ_b, and λ, and their statistics were estimated from a generalized soil properties database available in the literature (Table 1). In particular, Brakensiek et al. (1981) suggested transformations of the BC parameters to normality and reported their statistical properties for different soil types. McCuen et al. (1981) reported BC parameters statistics and demonstrated the variation that the BC parameters exhibit across different soil textural classes. Rawls et al. (1982) provided BC marginal statistical properties. Meyer et al. (1997) and Flores et al. (2010) reported marginal distributions of BC parameters by applying parameter equivalency relationships between BC and van Genuchten (1980) soil retention parameters. As in Arnone et al. (2014), we used statistical properties (i.e., means, variances and cross-correlation coefficients) reported by Brakensiek et al. (1981) and McCuen et al. (1981) and the corresponding transformations of BC parameters to normal random variables (Table 1), which allowed the use of the joint Normal distribution for the BC soil retention parameters.

With regard to the geotechnical parameters, Simon et al. (1990) and Lohnes and Demirel (1973) reported values for cohesive strength and friction angle for some of the geological units of the study area, and illustrated the expected high variability of these two quantities. In this study, as in Lepore et al., (2013), we assumed spatially homogenous properties over the entire basin following the predominant geological unit (Simon et al., 1990). The uncertainty

characterizing estimates of soil cohesion and friction angle has been thoroughly studied in the literature. Lumb (1966) suggests that c' and ϕ can be described by the Normal distribution. This assumption has been widely used in the literature (Abbaszadeh et al., 2011;Frattini et al., 2009;Melchiorre and Frattini, 2012;Simoni et al., 2008;Tobutt, 1982;Wu and Kraft, 1967;Langejan, 1965;Malkawi et al., 2000;Rackwitz, 2000). Fredlund and Dahlman (1972), Lumb (1974) and Schultze (1975) provided statistical properties of geotechnical parameters. Matsuo and Kuroda (1974) and Lumb (1974) suggested that correlation between cohesion and friction angle is negligible, and independence between the two random variables has been assumed in past studies (Abbaszadeh et al., 2011;Christian et al., 1994;Dettinger and Wilson, 1961;Malkawi et al., 2000;Yucemen MS et al., 1975). Therefore, here we assume that c' and ϕ are independent Normal random variables. The associated statistics of the geotechnical parameters are reported in Table 1. Incorporating marginal distributions consistent with the literature is an advantage of the implemented framework over simpler approaches (e.g. over Pack et al. (1998) who assume input parameters which follow a uniform distribution).

As for the vegetation characteristics, only the Tabonuco forest is considered, because it is the predominant vegetation type of the basin and it is present where both the meteorological and the soil moisture measurements were taken (Lepore et al., 2013). The forest is modeled as broadleaf evergreen tropical (BET) class, with vegetation height of 20m and a LAI of $6 m^2 m^{-2}$ (Wang et al., 2003;Weaver and Murphy, 1990). The root component has been modeled with a rectangular density function through a depth of 40 cm (typical of the Tabanuco forest). Other parameters used in the vegetation modeling have been taken from the literature (Ivanov et al., 2008a, b; Wang et al., 2003;Weaver and Murphy, 1990).

4 Monte Carlo Simulations

As previously mentioned, the selection of the probability distribution was based on Monte Carlo experiments for various hydrological and topographical conditions (i.e. soil moisture and slope values and soil depth). First, values of *BC* parameters, soil cohesion, and friction angle are sampled from the corresponding assumed theoretical distributions (as explained section 2.2), given the moments reported in Table 1. Then, the *FS* is estimated (Eq.1) for each set of sampled values, and for the fixed hydrological and topographical conditions. The results are then used to obtain the empirical distribution of *FS*. The procedure was repeated for different sets of slope, soil depth and soil moisture, and for all different soil types

characterizing the area, i.e., clay, clay-loam, silty-clay, and sandy-loam. The empirical FS distributions were compared to different theoretical fits, based both on Kolmogorov-Smirnov goodness of fit statistics and visually. The graphical method of Quantile-Quantile (QQ) plots (Wilk and Gnanadesikan, 1968) was used for visual comparison. We are specifically interested in reproducing the values of FS probability around 1, which is the FS_{crit} for failure. Thus, it is important that the theoretical quantiles be as close as possible to the empirical ones around values of FS=1.

As a further improvement with respect to the work by Arnone et al., (2014), several two-parameter theoretical distributions were compared against the empirical FS distribution as well as different combinations of fixed values of slope, depth of failure, and volumetric water content. For the sake of brevity, here we show the comparison of only three probability distributions against the empirical FS distribution (Figure 2), i.e., the normal, lognormal, and inverse gamma distributions, for the case of slope, depth of failure, and volumetric water content equal to 40°, 1000 mm, and 0.3 mm^3/mm^3, respectively, and for hydrological and soil properties given in Table 1. The inverse gamma distribution, which constitutes a special case of Pearson V distribution (Pearson, 1895), is a non-negative skewed distribution, characterized by a heavy tail on the right.

The perfect agreement between theoretical and empirical quantiles is depicted by the diagonal gray line in the QQ plots of Figure 2. This figure demonstrates that the normal distribution (purple dots) is the worst in terms of reconstruction of the empirical quantiles whereas the best results correspond to the inverse (red dots) and the lognormal (blue dots) distributions. Given that the goal is to identify the distribution that provides the most accurate estimation of FS quantiles close to the critical value (1 in this case), the inverse gamma distribution can be considered as the most appropriate distribution for the estimation of the probability of failure. Similar results were obtained for other combinations of slope, depth of failure, and volumetric water content.

The inverse gamma distribution was then used in tRIBS-VEGGIE-Landslide for the dynamic computation of the FS distribution at each cell. The probability of failure for unsaturated conditions is then computed as the cumulative probability that FS is less than 1:

$$\Pr(FS<1)=\frac{\Gamma\left(\alpha,\frac{\beta}{FS}\right)}{\Gamma(\alpha)} \qquad \textit{unsaturated conditions} \qquad (7)$$

where α and β are the parameters of the distribution, defined as a function of the mean and variance of FS:

$$\alpha = \frac{E[FS]^2}{Var[FS]} + 2 \qquad for\ \alpha > 2 \tag{8}$$

$$\beta = E[FS] \cdot (\alpha - 1) \tag{9}$$

Note that the results are strictly dependent on the assumed moments of the random variables (i.e., values given in Table 1).

5 Rainfall Analysis

In order to evaluate the effect of the temporal rainfall distribution and provide generalized landslide probability maps for the study area, different synthetic hyetographs representative of extreme rainfall observations were used to force the model.

The hyetographs were derived from the precipitation frequency estimates provided by the NOAA Atlas 14 Volume 3 (Bonnin et al., 2006) for Puerto Rico and the U.S. Virgin Islands, and available at the data sever *http://hdsc.nws.noaa.gov/hdsc/pfds.* The estimates are provided for several durations and return periods and are based on statistical analysis of the annual maximum series. In particular, the procedure makes use of regional frequency analysis based on L-moments method (Hosking and Wallis, 1997) for selecting and parameterizing probability distributions. The quantiles are provided at a spatial resolution of 3 arc-seconds (~ 80 m x 80 m) obtained by spatially interpolating the mean annual maximum of the series at each station and duration. In particular, data were spatially interpolated by using the PRISM (*Parameter-elevation Regressions on Independent Slopes Model*) model (Daly et al., 2002;Daly et al., 2003;Daly and Neilson, 1992), a hybrid statistical-geographic tool for mapping climate data that generates spatial distribution of estimated climatic parameters based on the correlation between point data and other geographic and climatic information. This procedure takes into account the topographic effect on climate and in particular precipitation patterns, which are known to be particularly significant in Puerto Rico (Daly et al., 2003;Garcia-Martino et al., 1996). For details reader can refer to NOAA report (Bonnin et al., 2006) . The result of this analysis is a spatial distribution of the total precipitation at given duration and return period, as specified above.

NOAA also provides the normalized temporal distribution of precipitation of durations of 1, 6, 12, 24, 96 hours. The temporal distributions are expressed in probabilistic terms as cumulative percentages of precipitation and duration at various percentiles. The data were also subdivided into four categories (identified as 'quartiles' in the NOAA report (Bonnin et al., 2006) based on where the most precipitation occurred in the distribution. For example, first-quartile group consist of cases where the greatest percentage of the total rainfall precipitates during the first quarter of the time period. This procedure led to four different hyetographs which characterize the area (Bonnin et al., 2006).

For our application, we selected a return period of 100 years and duration of 24 hours, which represents a typical event duration that may cause initiation of landslide events according to Larsen and Simon's (1993) work. Figure 3 shows the corresponding time distributions for each category, which are representative of four different types of precipitation events (Rosso, 2002): in Q1 most of the rainfall precipitates at the beginning of the event (typical of heavy storms), whereas in Q2 and Q3, the peak of precipitation is expected around the middle part of the event (frontal precipitations); finally, in Q4 most of the rainfall volume falls at the end of the event (tropical cyclones).

For the chosen duration, the most recurrent type of precipitation follow the Q1 pattern (32%);, 27% follow Q2, 22% is of Q3 type and 19% is Q4 (for more details see NOAA report, A. 1-3).

Given the temporal distribution and the spatial distribution of total precipitation at selected duration and return period it is then possible to derive, cell by cell, the corresponding hyetographs by simply distributing the total volume through the duration according to the curves showed in Figure 3. As an example, Figure 4 shows the obtained hyetographs at a selected location (in particular, where the Bisley tower is located) for the four types of hyetographs.

The spatially-averaged precipitation volume for this 24 hr, 100yr event is around 500 mm; in particular, figure 5 shows the map of the total rainfall estimate of hyetograph Q1. Precipitation is heavier at higher elevations (west and south corners of the basin), with values up to 615 mm, and lower in the extreme north-east area (close to the outlet zone), with values up to 485 mm (Figure 5). The resulting spatial coefficient variation is 0.05. It is worth

highlighting that the four types of rain storms have a stationary spatial pattern responding only to elevation, i.e. no dynamic propagation of precipitation has been considered.

6 Model application

As mentioned above, four 24 hours, 100 year return period events with different temporal distributions were used to force the model. The initial conditions of the basin were identified using a spin-up procedure, which consists of running the model continuously for a long period (in this case one year) so that the model reaches equilibrium (Lepore et al., 2013). Equal initial conditions were thus imposed to the 4 storms; the effect of different initial conditions is not analyzed in this work. Results will be analyzed in a time window of 48 hours, which includes 24 hours of no rainfall after the end of the event to account for the soil moisture redistribution effects.'

The model output includes: the time series and depth profiles of volumetric soil moisture (θ), probability that FS≤1 (PrF_i) and probability that the plane of failure is located at a given depth ($PrPF_i$), for a given element. Spatially the model provides the distribution of the probability that failure occurs at any depth within the element-soil column ($PrFC$) and the distribution of the most probable depth of failure.

In order to provide a comprehensive description of the model variables, the following section will discuss the results of a single model application, i.e. the results of the model forced with a single rainfall series whereas the analysis of the effect of different rainfall temporal distributions is discussed in section 6.2.

6.1 Time series model output

A time series output is given in Figure 6, which shows the response to the rainfall type Q1 (panel a) at element scale (the time window includes a period before the event, useful to assess the initial conditions in terms of soil moisture profiles, which are different across the basin). The selected element falls within the clay-loam soil type and has a slope value of ~52°, and thus can be considered as relatively steep. The element is located upstream and has a small contributing area of about 1.3 ha. The soil moisture distribution (panel b) shows

moderately fast dynamics within the shallow layers of the column, down to 600 mm of depth where the soil reaches full saturation at the time the rainfall peaks. Although this part of the basin is characterized by high value of hydraulic conductivity (see Table 1), saturation is not reached throughout the entire column because a significant portion of water is laterally redistributed, at a rate which depends on the steepness and which is governed by the anisotropy ratio coefficient in the model (see Lepore et al., 2013 for additional analyses). Panel *c* shows the probability of failure (PrF_i) for a selected time at each soil depth. This type of probability is representative of the likelihood of failure occurrence at each depth conditioned on the local soil moisture content (time variant) and soil weight (time invariant). At *t1* the probability of failure is zero along the entire column and θ corresponds to the initial conditions, with the soil water content uniform with depth. At *t2* the soil moisture rapidly increases to a depth of 600 mm where the PrF_i reaches its peak (~0.58). Apart from the soil moisture conditions, the cohesion (which is invariant) and the soil weight are the other factors that affect the final result; therefore the probability of failure is generally very low at shallow depths, due to the low weight. At *t3* the probability reaches its maximum value (~0.65) at a depth of 800 mm where both the contribution of moisture and soil weight are significant, and also at *t4*. Panels *d* and *e* show the probability that the failure occurs at a given depth At each layer, the $PrPF_i$ depends on the probability that the failure does not occur in the above layers (see definition in section 2.2). Consequently, the most probable planes of failure (i.e., where $PrPF_i$ is high) are located at the medium depth layers where the effects of soil moisture and soil weight are significant, which, in this case, is around 300-500 mm at the time of rainfall peak (dark red area in panel d). Deeper layers have lower probability of becoming a plane of failure, as shown in the panel *d* for timestep *t3* and *t4*. Finally, panel *d* also shows how the high risk of failure occurrence is prolonged in time and the probability gradually decreases at shallower layers and increases at deeper layers, as long as the soil moisture conditions are close to saturation.

Figure 7a shows the response to the same rainfall type Q1 of a silty-clay element with a slope value of ~21°, gentle relative to the previous case. Note that the hyetograph over this element is slightly different than that over the previously discussed element, as a result of the spatially distributed precipitation. The element is located in the downstream flatter part of the basin, with a contributing area of about 2.5 ha. In this case, the initial soil moisture profile (panel *b*) is wetter than the previous case (due to slightly higher convergence of fluxes and lower redistribution to adjacent to the cells) and is not homogeneous with depth, with dryer

conditions at shallower layers, due to the evapotranspiration processes. At the rainfall peak the column almost reaches full saturation. In fact, the combination of gentler topography and medium-high hydraulic conductivity results in faster soil moisture dynamics within the column but slower among the adjacent cells, since the lateral redistribution is slower for gentler slopes.

In terms of probability of failure, the $PrPF_i$ distribution (panel *c*) depicts a restricted risk of failure limited to depths between 600 and 800 mm and to the peak of the storm. Panel *d* shows the probability of failure associated with the entire column (*PrFC*). This variable takes into account the model assumption that failure may occur anywhere in the column, instead of occurring exclusively at a preselected depth. Here the maximum value is reached at the peak of the storm and is strictly dependent on the fully saturated conditions.

6.2 Effect of rainfall temporal distribution

The basin response to the four different hyetographs described in Section 5 is analyzed here in terms of landslide probability. Variables are dynamically evaluated by the model, which produces a spatial distribution of the most probable areas of failure at each computational time step. The most severe scenario associated with each event is evaluated as the maximum value of the probability of failure anywhere in the column, *PrFC,* recorded at each voronoi cell over the entire run time (i.e. peak of the curve shown in figure 7d).

Maps showing *PrFC* for each rainfall forcing are shown in Figure 8.

Across the four maps, the likelihood of failure occurrence is particularly high in the steepest part of the basin (yellow regions). The dark green regions depict the areas where the probability of failure is zero or close to zero, which correspond mostly to the flat areas. In the remaining part of the basin, the probability of failure is the result of the interaction of rainfall type with basin soils and morphology.

In order to compare the results better, we define five classes of landslide hazards at equal intervals: very low (*VL*), 0-0.2; low (*L*), 0.2-0.4; medium (M), 0.4-0.6; high (*H*), 0.6-0.8; very high (*VH*), 0.8-1.0. Differences are then quantified in terms of relative frequency distribution over the basin. Results across the four rainfall events (Figure 9a) indicate that events Q2 and Q3 resulted in the safest scenarios, showing the highest frequency in the *VL* class and the lowest in the remaining classes. Q4 is the rainfall event that provides the most severe landslide hazard, with the highest relative frequency within the *VH* landslide probability class

and the lowest within the *VL* class. Event Q1 shows an intermediate behavior between Q4 and Q2-Q3, with a frequency distribution similar to Q2 and Q3 within the most hazardous class, and similar to Q4 for *H*, *M* and *L* landslide probability classes; class *VL* is slightly less frequent during events Q2 and Q3, but more frequent for Q4.

The above analysis highlights a clear and strong effect of the rainfall type on the landslide probability; this effect depends on the different interactions between rainfall forcing, topography and soil characteristics, which induce different soil moisture redistribution and, in particular, whether full saturation is achieved or not. In fact, saturation leads to the loss of suction with a significant increase of probability of failure. The Q1 and Q4 events induce saturation in a larger portion of the basin, leading to higher failure probability, although the timing of the maximum landslide probability is different over the run time. This is depicted in Figure 9b that shows the frequency distribution of the time of occurrence of the maximum value of *PrFC*. The peak of the distributions clearly follows the peak of the rainfall intensity for each hyetograph, with 1 hr delay (at time 7, 11, 17 and 23 hour respectively). Q1 and Q4 exhibit the highest peaks of frequency. Moreover, for Q1, the effects of the rainfall are more prolonged. *PrFC* is still high after 13 hours from the rainfall peak. Q3 is the event that provides the highest probability of failure almost throughout the duration of the precipitation event. These results can be particularly interesting in terms of basin risk management, since the high probability of failure is significantly different for the four cases, despite the same total volume in 24 hr. The higher peack of frequency of Q4 with respect to Q1 highlights the importance of the moisture conditions prior the rainfall peak. In fact, although Q1 and Q4 have comparable storm peaks, the response with respect to landslide occurrence does differ because soil moisture conditions at time of the rainfall peak are not the same for the two scenarios.

In order to understand how the rainfall forcing interacts differently with the soils of the basin, the responses in terms of soil moisture are analyzed below in conjunction with the main basin characteristics (i.e. topography and soil types).

Figure 9c shows the relative frequency distribution of the average (in depth) soil moisture in the root zone (equal to 1 m) reached at time of the maximum value of *PrFC* across the basin. Clearly, Q4 shows the highest occurrence of elements at saturation (here equal to 0.55 mm^3/mm^3).

In order to identify which portion of the basin experiences the most differences in probability of failure, we analyze the spatial distribution of the difference in maximum *PrFC* between the events Q4 and Q3 (which are the two configurations that differ the most), denoted as ΔQ_4Q_3 and evaluated for each single element. The map is depicted in Figure 10 and it reports only the differences greater than a threshold, fixed at 0.05; in many cases, this difference is greater than 0.5 (blue). Since soil moisture redistribution processes are driven by topography (i.e. slope) and hydrological soil properties (i.e. hydraulic conductivity and anisotropy), one might expect this difference to follow somehow similar patterns to either the soils (Figure 10a) or distributions of slope(Figure 10b).

The map in Figure 10a highlights that ΔQ_4Q_3 values greater than the threshold are distributed all over the basin, falling within all the soil types and highlighting the absence of a clear spatial pattern related to the soil characteristics. Instead, Figure 10b clearly shows that the spatial distribution of ΔQ_4Q_3 values greater than the threshold fall within the part of basin characterized by a particular range of slope, i.e. between 0.20 and 0.45 rad (~10° and 25°), which is the only interval shown in the map.

The reason for this result is the 'partition' of the contribution to the probability of failure of soil moisture and soil weight. Flat areas, either saturated or not, are certainly stable (or 'unconditionally stable', as defined by Montgomery and Dietrich (1994)); steep areas, although they may not reach the full saturation, can easily have high probability of failure, given the significant component of destabilizing forces due to the steepness ('unconditionally unstable' as defined by Montgomery and Dietrich (1994)). Therefore, as previously discussed, rainfall type distributions influence mostly areas where both factors, soil moisture and soil weight, are predominant. In this case, those are areas of medium slope (from 10° and 25°).

Figure 11 provides a clear confirmation of these results. Panel *a* shows the relationship, for each element, between slope and the max value of *PrFC* for Q1, Q3 and Q4 (Q2 is omitted because of its similarity to Q3). Clearly, the *PrFC* increases with slope following a well-defined trend (at given soil moisture and for fixed hydrological parameters, probability of failure only changes with slope according to Eq. 1). At a given slope (i.e. within a given voronoi cell) Q4 results are shifted up with respect to Q1 (blue) and Q3 (red), i.e. provides higher max *PrFC*. Panel *b* shows the distribution of ΔQ_4Q_3 values with the slope, separately for each soil type. The variable jumps to high values within the slope range 0.2 – 0-45, except

for the clayey soil where differences are not that significant, due to the much slower soil moisture dynamics (i.e. low hydraulic conductivity).

6.3 Discussion

The described results show the capability of the proposed methodology in accounting for uncertainty of soil parameters and in evaluating how the temporal variability of storms may influence the initiation of landslide occurrences in terms of probability of failure.

The model produces different types of 'probabilities of failure', thus overcoming the uncertainty related to the assumption of the depth of failure. The most probable depth of failure and the probability of failure of the soil column (generated anywhere within the column) are the two main representative outputs. The results demonstrate that, in agreement with the definition of shallow landslides, the most probable failure surfaces occurred at depths between 300 and 1000 mm.

The precipitation fields in the case study area are strongly variable both in space and in time. The analysis of the influence of different hyetograph shapes indicates that heavy rainfall concentrated during the latter part of event (as the case of Q4) leads to higher probability of failure across the basin, making Q4 the most critical type of event. Such a result agrees with the outcomes of D'Odorico et al., (2005) work which demonstrated that hyetographs with the peak near the end of the storm produce peak pressure heads higher than uniform hyetographs, decreasing the return period of rainfall events causing landsliding. Although events Q1 and Q4 have the same maximum rainfall intensity (~55mm/hr at the selected location reported in Figure 3), Q1 results in a less serious scenario because the Q1 hyetograph has the maximum rainfall intensity closer to the beginning of the storm. These outcomes depend on how the rainfall distribution affects the hydrological basin response, which is mainly driven by the hydraulic-hydrological soil properties and topography. The spatial analysis of the differences in the maximum *PrFC* between Q3 and A4 ($\Delta Q_4 Q_3$), clearly demonstrates that the shape of the hyetograph, and in particular the timing of the peak, controls the resulting stability mainly at medium slopes (from 10° and 25°). As previously discussed, at these slopes, both the soil moisture and soil weight contribute to instability. Events Q2 and Q3, which are characterized by a lower peak rainfall intensity (~45mm/hr for the location shown in Figure 3) and a less skewed temporal distribution, result in a much larger portion of the basin having low probability of failure.

In order to evaluate the consistency of the model results, the classified map of probability of failure obtained for Q4 is compared with the map of historical landslide scars (Figure 12). The study area is historically characterized by a high frequency of landslide occurrence and plenty of events have been mapped over the years. Landslide scars reported in Figure 12 are obtained from the work of Larsen (2012), which mapped historical landslide scars estimated to represent landslide activity for a 50 year period (1945-1995). Landslides were mapped from aerial photographs (1951, 1962, 1972, 1974, 1979, 1990 and 1995), but the storms that triggered landslide scars were not precisely identified because no historical records of the events exist (Larsen, 2012). According to Larsen and Simon (1993), the 61% of the rainfall triggered landslides were likely triggered by the tropical disturbances that struck the central mountain of Puerto Rico between 1960 and 1990. Moreover, the reported map includes different type of movements, i.e. shallow soil slip, debris avalanche, debris flow, slump). Soil slips are the second most common type of scar after debris flow (Larsen, 2012) and both these landslide types are characteristic of high-intensity, short-duration storms. Figure 12 reports also some landslide scars identified from more recent satellite images (2006-2014) obtained from Google Earth.

Visually the areas of high and very high probability of landslide occurrence are consistent with areas of past landslide zones. For instance, the northwest, south, and southeast parts of the basin correspond to areas dense in landslide scars which are adequately reproduced by the model. Other areas classified as probable to fail are not in the historical record of scars(central north-eastern part of the basin). Part of the southwest strip of the watershed perimeter is mostly classified as stable by the model, while in the past it did experience various failure events.

It is clear that such a comparison cannot be interpreted as model validation because a rigorous validation of dynamic approaches requires a functional database which records the locations of failure associated with the exact timing of the failure, the corresponding meteorological data, and the topographical data prior to the failure events, information that is not available. Moreover, the map reported by Larsen (2012) includes various type of landslides (according to Larsen, (2012) mainly debris flow) that are not of the nature represented in this work. About 43% of historical landslide mapped in the Mameyes basin were associated with road construction and maintenance (Larsen, 2012), anthropogenic disturbance that is not accounted for in this study. Nevertheless, the rough comparison between historical landslides and model

results demonstrate the model capability in identifying the most critical landslide areas of the basin.

7 Conclusions

This study proposed a probabilistic framework, implemented in an existing rainfall-triggered shallow landslide model, able to take into account the uncertainty of soil parameters. We presented a methodology based on a modified version of the First Order Second Moment (FOSM) method, which has been implemented into the physically-based tRIBS-VEGGIE-Landslide model (Lepore et al., 2013). The framework allows to dynamically evaluate the Factor of Safety (FS) probability distribution conditioned on the soil moisture content. In contrast to similar approaches, this methodology treats all parameters of the FS equation (i.e. soil cohesion, friction angle, and soil retention curve parameters) as random variables. The probabilistic approach provides a sound and suitable framework to evaluate the probability of landslide occurrence by accounting for the uncertainty associated with the difficulty in estimating the soil hydrological and geotechnical parameters.

The methodology was used in the Rio Mameyes basin, located in Puerto Rico, to also evaluate the effect of spatial and different temporal distributions of precipitation on shallow landslide occurrence over the basin. Four different synthetic and spatially distributed hyetographs at a given rainfall volume were used to carry out the analysis.

The use of the proposed physically based approach allowed the simulation of the complex nonlinear interactions between rainfall forcing and soil basin characteristics (mainly topography and soil types) in evaluating the hydrological and slope stability basin responses. The main results indicated that i) hyetographs exhibiting a peak towards the end of the event, typical of tropical cyclones, can be most catastrophic in terms of rainfall-triggered landslide occurrence. ii) Given equal rainfall intensity peaks, events with peaks occurring early in time result in lower probability of failure across the basin. iii) Precipitation with maximum intensities around the middle part of the event (frontal precipitations) and with a lower peak result in lower probabilities of failure; although in those cases, relatively high probabilities of failure persisted throughout the whole event. Finally, iv) differences in probability of failure due to the different shape of the hyetographs were observed at intermediate slopes where the stability is controlled by the combined impact of soil moisture and soil weight, in contrast to steeper or flatter slopes. The overlap between historical landslide events, mainly relative to

the 20th century, and the region of the maximum probability of failure has confirmed the model capability to identify the areas at dense frequency of landslides, even if this comparison cannot be considered a rigorous validation procedure.

The proposed approach is computationally feasible, as opposed to numerical probabilistic analyses at basin scale that require prohibitive numbers of model runs. Some modeling assumptions could be still limiting: (i) the assumption of normal distribution of cohesion and friction angle, which is widely used; (ii) the assumption of independence between cohesion and friction angle; (iii) the slope geometry of the slope can be very restrictive in some cases, as for all the models based on the infinite slope models; (iv) the spatial variability of parameters which is not taken into account in this discussion; (v) the fact that the dynamic propagation of storms was not considered in this case study.

A more accurate characterization of natural variability of soil parameters could be developed by taking into account the spatial correlation that soil geotechnical and hydrological properties may exhibit. This aspect could be achieved in the future by introducing geostatistical techniques, able to model the spatial correlation of such parameters, within the proposed framework. Moreover, the mechanical role of vegetation dynamics could be also included in a future development of the probabilistic methodology to quantify the effects on landslide occurrence.

Acknowledgements

This work has been supported by the Regional Sicilian Government (Italy) “Linea di intervento 4.1.1.1 - POR FESR Sicilia 2007-2013” under Project SESAMO - SistEma informativo integrato per l’acquisizione, geStione e condivisione di dati AMbientali per il supportO alle decisioni, by a subcontract from the University of Pennsylvania’s Luquillo Critical Zone Observatory effort funded by the National Science Foundation (project EAR0722476), and by a subcontract from the University of New Hampshire’s Luquillo Critical Zone Observatory effort funded by the National Science Foundation (project EAR1331841).

References

Abbaszadeh, M., Shahriar, K., Sharifzadeh, M., and Heydari, M.: Uncertainty and Reliability Analysis Applied to Slope Stability: A Case Study From Sungun Copper Mine, Geotechnical and Geological Engineering, 29, 581-596, 10.1007/s10706-011-9405-1, 2011.

Arnone, E., Noto, L., Lepore, C., and Bras, R.: Coupled stability and eco-hydrological model to predict shallow landslides, The 2th Italian Workshop on Landslides, Naples, 2011, 2011.

Arnone, E., Dialynas, Y. G., Noto, L. V., and Bras, R. L.: Parameter Uncertainty in Shallow Rainfall-triggered Landslide Modeling at Basin Scale: A Probabilistic Approach, Procedia Earth and Planetary Science, 9, 101-111, http://dx.doi.org/10.1016/j.proeps.2014.06.003, 2014.

Benjamin, J. R. A., and Cornell, C. A. A.: Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill Ryerson, Limited, 1970.

Beven, K., and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279-298, 10.1002/hyp.3360060305, 1992.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Advances in Water Resources, 16, 41-51, http://dx.doi.org/10.1016/0309-1708(93)90028-E, 1993.

Bonnin, G., Martin, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D.: Precipitation Frequency for Puerto Rico and the US Virgin Islands - NOAA Atlas 14 Volume 3, NOAA/NWS/Office of Hydrologic Development, Hydrometeorological Design Studies Center, Silver Spring, MD, USA, 2006.

Brakensiek, K. L., Engleman, R. L., and Rawls, W. J.: Variation within texture classes of soil water parameters., ASAE, 24, 335-339, 1981.

Brooks, R. H., and Corey, A. T.: Hydraulic properties pf porous media, Hydrology Paper, Civil Engineering Dep., 1964.

Burton, A., Arkell, T. J., and Bathurst, J. C.: Field variability of landslide model parameters, Environmental Geology, 35, 100-114, 10.1007/s002540050297, 1998.

Burton, A., and Bathurst, J. C.: Physically based modelling of shallow landslide sediment yield at a catchment scale, Environmental Geology, 35, 89-99, 1998.

Capparelli, G., Biondi, D., Luca, D. L. D., and Versace, P.: Hydrological and complete models for forecasting landslides triggered by rainfalls, Integration The Vlsi Journal, 2002.

Christian, J., Ladd, C., and Baecher, G.: Reliability Applied to Slope Stability Analysis, Journal of Geotechnical Engineering, 120, 2180-2207, doi:10.1061/(ASCE)0733-9410(1994)120:12(2180), 1994.

D'Odorico, P., Fagherazzi, S., and Rigon, R.: Potential for landsliding: Dependence on hyetograph characteristics, Journal of Geophysical Research: Earth Surface, 110, F01007, 10.1029/2004jf000127, 2005.

Daly, C., and Neilson, R. P.: A digital topographic approach to modeling the distribution of precipitation in mountainous terrain, in: Interdisciplinary Approaches in Hydrology and Hydrogeology, edited by: Jones, M. E., and Laenen, A., American Institute of Hydrology, Minneapolis, MN, 437-454, 1992.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P.: A knowledge-based approach to the statistical mapping of climate, Climate Research, 22, 99-113, 2002.

Daly, C., Helmer, E. H., and Quinones, M.: Mapping the climate of Puerto Rico, Vieques and Culebra, International Journal of Climatology, 23, 1359-1381, 2003.

Dettinger, M., and Wilson, J.: First order analysis of uncertainty in numerical models of groundwater flow part: 1. Mathematical development, Water Resources Research, 17, 149-161, 1961.

Ewel, J. J., and Whitmore, J. L.: The ecological life zones of Puerto Rico and the U.S. Virgin Islands., Institute of Tropical forestry, Rio Piedras, Puerto Rico, U.S. Forest Service Research Paper, 1973.

Feller, W.: An introduction to probability theory and its applications, Wiley, 1971.

Flores, A. N., Entekhabi, D., and Bras, R. L.: Reproducibility of soil moisture ensembles when representing soil parameter uncertainty using a Latin Hypercube–based approach with correlation control, Water Resources Research, 46, W04506, 10.1029/2009wr008155, 2010.

Frattini, P., Crosta, G., and Sosio, R.: Approaches for defining thresholds and return periods for rainfall-triggered shallow landslides, Hydrological Processes, 23, 1444-1460, 10.1002/hyp.7269, 2009.

Fredlund, D. F., and Dahlman, A. E.: Statistical Geotechnical Properties of Glacial Lake Edmonton Sediments, in: Statistics and Probability in Civil Engineering, edited by: Press, H. K. U., Oxford University Press, London, 1972.

Garcia-Martino, A. R., Warner, G. S., Scatena, F. N., and Civco, D. L.: Rainfall, Runoff and Elevation Relationships in the Luquillo Mountains of Puerto Rico, Caribbean Journal of Science, 32, 413-424, 1996.

Hillel, D.: Fundamentals of soil physics, Academic Press, New York, NY, USA, 1980.

Hosking, J. R. M., and Wallis, J. R.: Regional Frequency Analysis, an approach based on L-moments, Cambridge University Press, Cambridge, UK, 1997.

Huffaker, L.: Soil Survey of Caribbean National Forest and Luquillo Experimental Forest, Commonwealth of Puerto Rico, USDA, Natural Resources Conservation Service, Washington, DC, 2002.

Ivanov, V., Bras, R. L., and Vivoni, E. R.: Vegetation-Hydrology Dynamics in Complex Terrain of Semiarid Areas: II. Energy-Water Controls of Vegetation Spatio-Temporal Dynamics and Topographic Niches of Favorability, Water Resources Research, 44, W03430, 10.1029/2006WR005595, 2008.

Langejan, A.: Some aspects of the safety factor in soil mechanics considered as a problem of probability, Sixt Int. Conf. Soil Mechanics and Foundation Eng, Montreal, Quebec, 1965, 500-502,

Larsen, M. C., and Simon, A.: A rainfall intensity-duration threshold for landslides in a humid- tropical environment, Puerto Rico, Geografiska Annaler, Series A, 75 A, 13-23, 1993.

Larsen, M. C.: Landslides and sediment budgets in four watersheds in eastern Puerto Rico, in: Water quality and landscape processes of four watersheds in eastern Puerto Rico. US Geol Surv Prof Pap 1789, edited by: Stallard, S. F. M. a. R. F., 153-178, 2012.

Lepore, C., Kamal, S. A., Shanahan, P., and Bras, R. L.: Rainfall-Induced Landslide Susceptibility Zonation of Puerto Rico, Environmental Earth Sciences, doi:10.1007/s12665-011-0976-1, 2012.

Lepore, C., Arnone, E., Noto, L. V., Sivandran, G., and Bras, R. L.: Physically based modeling of rainfall-triggered landslides: a case study in the Luquillo forest, Puerto Rico, Hydrol. Earth Syst. Sci., 17, 3371-3387, 10.5194/hess-17-3371-2013, 2013.

Lohnes, R. A., and Demirel, T.: Strength and structure of laterites and lateritic soils, Engineering Geology, 7, 13-33, 1973.

Lumb, P.: The Variability of Natural Soils, Canadian Geotechnical Journal, 3, 74-97, 10.1139/t66-009, 1966.

Lumb, P.: Applications of statistics in soil mechanics, in: Soil mechanics. New Horizons, edited by: JK, L., London, 1974.

Malkawi, A. I., Hassan, W. F., and Abdulla, F. A.: Uncertainty and reliability analysis applied to slope stability, Structural Safety, 22, 161-187, 2000.

Matsuo, M., and Kuroda, K.: Probabilistic Approach to Design of Embankments, Soils and Foundations, 14, 1-17, 1974.

McCuen, R. H., Rawls, W. J., and Brakensiek, D. L.: Statistical analysis of the Brooks-Corey and the Green-Ampt parameters across soil textures, Water Resources Research, 17, 1005-1013, 10.1029/WR017i004p01005, 1981.

Melchiorre, C., and Frattini, P.: Modelling probability of rainfall-induced shallow landslides in a changing climate, Otta, Central Norway, Climatic Change, 113, 413-436, 10.1007/s10584-011-0325-0, 2012.

Meyer, P. D., Rockhold, M. L., and Gee, G. W.: Uncertainty analyses of infiltration and subsurface flow and transport for SDMP sites, Pacific Northwest National, Laboratory, Division of Regulatory Applications, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission Washington, DC, 1997.

Montgomery, D. R., and Dietrich, W. E.: A physically based model for the topographic control on shallow landsliding, Water Resources Research, 30, 1153-1171, 1994.

Myster, R. W., Thomlinson, J. R., and Larsen, M. C.: Predicting landslide vegetation in patches on landscape gradients in Puerto Rico., Landscape Ecology, 12, 299-307, 1997.

Ogden, F. L., and Julien, P. Y.: Runoff sensitivity to temporal and spatial rainfall variability at runoff plane and small basin scales, Water Resources Research, 29, 2589-2597, 10.1029/93wr00924, 1993.

Pack, R. T., Tarboton, D. G., and Goodwin, C. N.: SINMAP, a stability index approach to terrain stability hazard mapping. SINMAP user's manual, Terratech Consulting Ltd, 1998.

Pearson, K.: Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material, Philosophical Transactions of the Royal Society of London. A, 186, 343-414, 10.2307/90649, 1895.

Rackwitz, R.: Reviewing probabilistic soils modelling, Computers and Geotechnics, 26, 199-223, http://dx.doi.org/10.1016/S0266-352X(99)00039-7, 2000.

Rawls, W. J., Brakensiek, D. L., and Saxton, K. E.: Estimation of Soil Water Properties, Transactions American Society of Agricultural Engineers, St. Joseph, MI, 25, 1316-2320, 1982.

Ray, A., and Baidya, D.: Probabilistic Analysis of a Slope Stability Problem, Indian Geotechnical Conference, Kochi, 2011, K-250,

Rosso, R.: Manuale di protezione idraulica del territorio, Cusl, 2002.

Rosso, R., Rulli, M. C., and Vannucchi, G.: A physically based model for the hydrologic control on shallow landsliding, Water Resources Research, 42, 16, 10.1029/2005WR004369, 2006.

Scatena, F. N., and Lugo, A. E.: Geomorphology, disturbance, and soil and vegetation of two subtropical wet steepland watersheds of Puerto Rico, Geomorphology, 13, 199-213, 1995.

Schultze, E.: Some Aspects Concerning the Application of Statistics and Probability to Foundation Structures, 2nd International Conference on Applications of Statistics and Probability in Soil and Structural Engineering,, Aachen, 1975, 457-494,

Shiels, A. B., West, C. A., Weiss, L., Klawinski, P. D., and Walker, L. R.: Soil factors predict initial plant colonization on Puerto Rican landslides, Plant Ecology, 195, 165-178, 2008.

Simon, A., Larsen, M. C., and Hupp, C. R.: The role of soil processes in determining mechanisms of slope failure and hillslope development in a humid-tropical forest eastern Puerto Rico, Geomorphology, 3, 263 - 286, 1990.

Simoni, S., Zanotti, F., Bertoldi, G., and Rigon, R.: Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS, Hydrological Processes, 22, 532-545, 10.1002/hyp.6886, 2008.

Singh, V. P.: Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph, Hydrological Processes, 11, 1649-1669, 10.1002/(sici)1099-1085(19971015)11:12<1649::aid-hyp495>3.0.co;2-1, 1997.

Tobutt, D. C.: Monte Carlo Simulation methods for slope stability, Computers & Geosciences, 8, 199-208, http://dx.doi.org/10.1016/0098-3004(82)90021-8, 1982.

van Genuchten, M. T.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1, Soil Sci. Soc. Am. J., 44, 892-898, 10.2136/sssaj1980.03615995004400050002x, 1980.

Vivoni, E. R., Ivanov, V. Y., Bras, R. L., and Entekhabi, D.: Generation of triangulated irregular networks based on hydrological similarity, Journal of Hydrologic Engineering, 9, 288-302, 2004.

Waide, R. B., Zimmerman, J. K., and Scatena, F. N.: Controls of Primary Productivity: Lessons from the Luquillo Mountains in Puerto Rico, Ecology, 79, 31-37, 1998.

Walker, L. R., Zarin, D. J., Fetcher, N., Myster, R. W., and Johnson, A. H.: Ecosystem development and plant succession on landslides in the Caribbean, Biotropica, 28, 566-576, 1996.

Walker, L. R., and Shiels, A. B.: Post-disturbance erosion impacts carbon fluxes and plant succession on recent tropical landslides, Plant Soil, 313, 205-216, 2008.

Wang, H., Hall, C. A. S., Scatena, F. N., Fetcher, N., and Wu, W.: Modeling the spatial and temporal variability in climate and primary productivity across the Luquillo Mountains, Puerto Rico, Forest Ecology and Management, 179, 69-94, 2003.

Weaver, P. L., and Murphy, P. G.: Forest Structure and Productivity in Puerto Rico's Luquillo Mountains, Biotropica, 22, 69-82, 1990.

Wilk, M. B., and Gnanadesikan, R.: Probability Plotting Methods for the Analysis of Data, Biometrika, 55, 1-17, 10.2307/2334448, 1968.

Wilson, C. B., Valdes, J. B., and Rodriguez-Iturbe, I.: On the influence of the spatial distribution of rainfall on storm runoff, Water Resources Research, 15, 321-328, 10.1029/WR015i002p00321, 1979.

Wu, T., and Kraft, L.: The Probability of Foundation Safety, Journal of the Soil Mechanics and Foundations Division, 93, 213-231, 1967.

Wu, W., and Sidle, R.: A distributed slope stability model for steep forested basins, Water Resources Research, 31, 2097-2110, 1995.

Yucemen MS, Tang WH, and AHS., A.: Long term stability of soil slopes - a reliability approach, Second Int. Conf. Applications of Statistics and Probability to Soil and Structural Eng., Aachen, Germany., 1975,

Table 1.Hydrological and mechanical soil properties and their statistics for the four soil types present in the Mameyes basin. See text for description and origin of data. Note that shear strength parameters are assumed homogenous across the basin.

Parameter	Description	Units	Clay – Loam	Sandy – Loam	Silty - Clay	Clay
K_s	Saturated hydraulic conductivity	[mm/hr]	50.0	50.0	50.0	10.0
$\mu_{\theta S}$	Mean of saturated soil moisture, θ_S	[mm^3/mm^3]	0.56	0.55	0.55	0.53
$\mu_{\theta R}$	Mean of residual soil moisture, θ_R	[mm^3/mm^3]	0.075	0.041	0.051	0.09
μ_{λ}	Mean of pore-size distribution index, λ	[-]	0.200	0.322	0.127	0.130
$\mu_{\psi b}$	Mean of air entry bubbling pressure, ψ_b	[mm]	-250	-150	-340	-370
$\mu_{c'}$	Mean of soil effective cohesion, c'	[N/m^2]	*3000*	*3000*	*3000*	*3000*
μ_{ϕ}	Mean of soil friction angle, ϕ	[°]	*25*	*25*	*25*	*25*
$\sigma_{c'}$	Standard deviation of c'	[N/m^2]	*1200*	*1200*	*1200*	*1200*
σ_{ϕ}	Standard deviation of ϕ	[°]	*2.5*	*2.5*	*2.5*	*2.5*
$\sigma_{\psi b}$	Standard deviation of ψ_b	[mm]	290	210	390	600
$\sigma_{\theta S}$	Standard deviation of θ_S	[mm^3/mm^3]	0.054	0.076	0.064	0.040
$\sigma_{\theta R}$	Standard deviation of θ_R	[mm^3/mm^3]	0.007	0.004	0.022	0.011
σ_{λ}	Standard deviation of λ	[-]	0.113	0.145	0.094	0.098
$\rho_{\psi b-\theta S}$	Coefficient of correlation ψ_b-θ_S	[-]	0	0	0	-0.216
$\rho_{\psi b-\theta R}$	Coefficient of correlation ψ_b-θ_R	[-]	0.203	0	0	0.154
$\rho_{\psi b-\lambda}$	Coefficient of correlation ψ_b-λ	[-]	0.151	0.274	0	0.128
$\rho_{\theta S-\theta R}$	Coefficient of correlation θ_S-θ_R	[-]	0.307	0	0	0
$\rho_{\theta S-\lambda}$	Coefficient of correlation θ_S-λ	[-]	0.168	0	0	0
$\rho_{\theta R-\lambda}$	Coefficient of correlation θ_R-λ	[-]	0.429	0.518	0.476	0.442

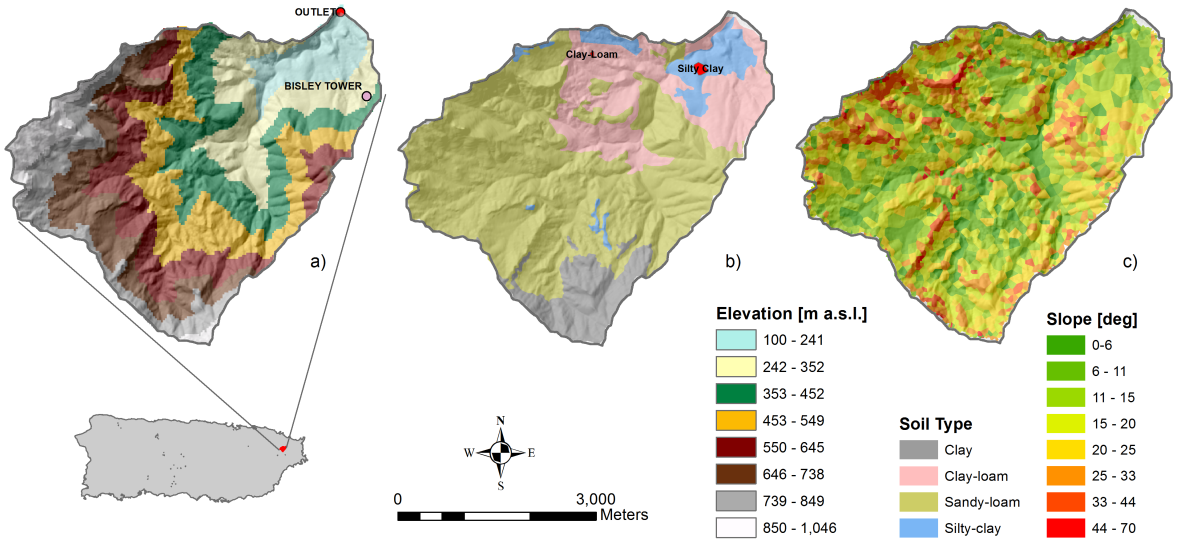

Figure 1. Mameyes basin characteristics: Digital Elevation Model (a), soil type distribution (b) and slope distribution (c).

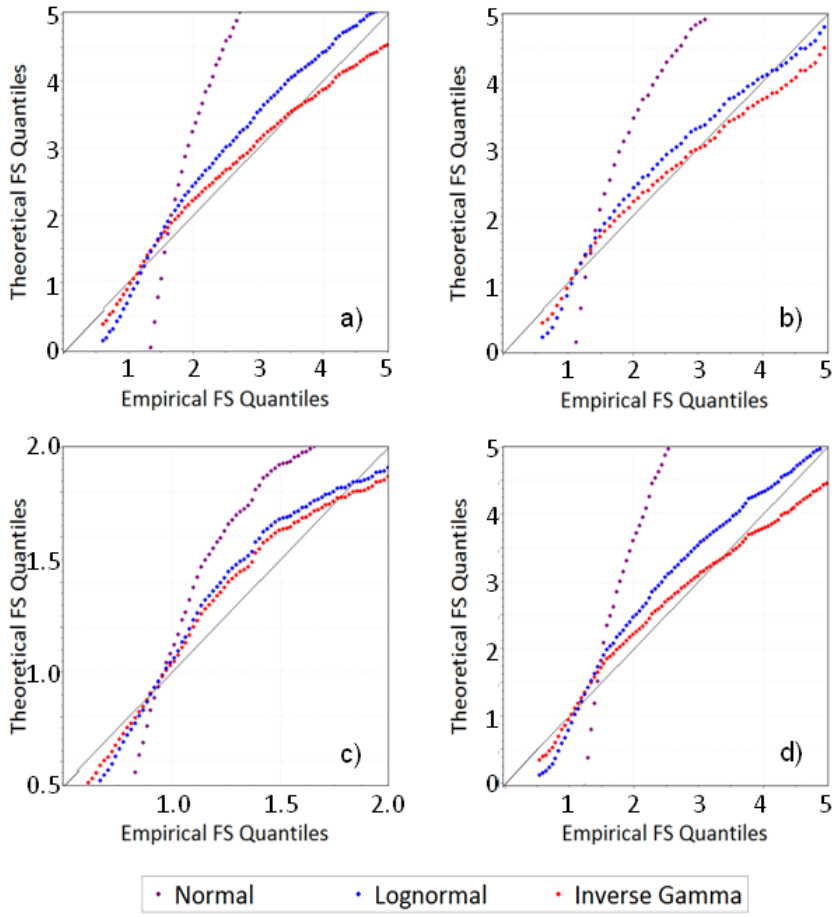

Figure 2. Comparison of the empirical FS quantiles vs the theoretical FS quantiles of Normal (purple), Lognormal (blue), and Inverse Gamma (red) distributions, for (a) clay, (b) clay-loam, (c) sandy-loam, and (d) silty-clay. The values of slope, depth of failure surface and volumetric water content correspond to 40°, 1000 mm, and 0.3 mm^3/mm^3, respectively, while the associated statistical properties are reported on Table 1.

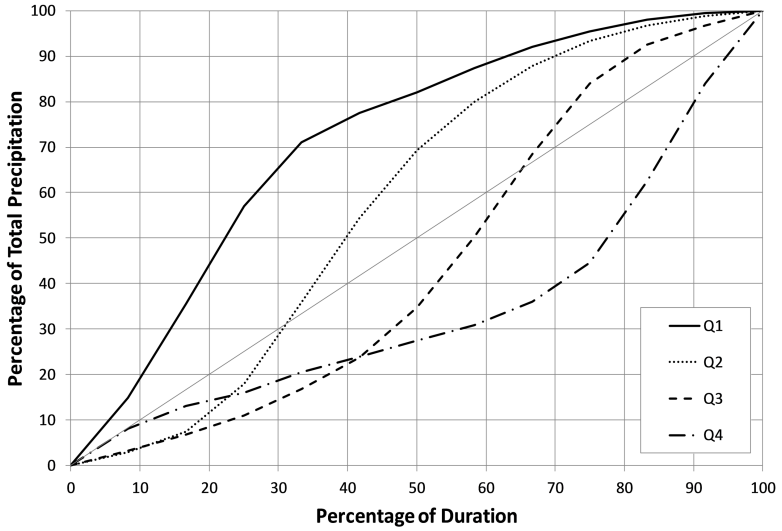

Figure 3. Different types of rainfall temporal distributions of 24 hours duration.

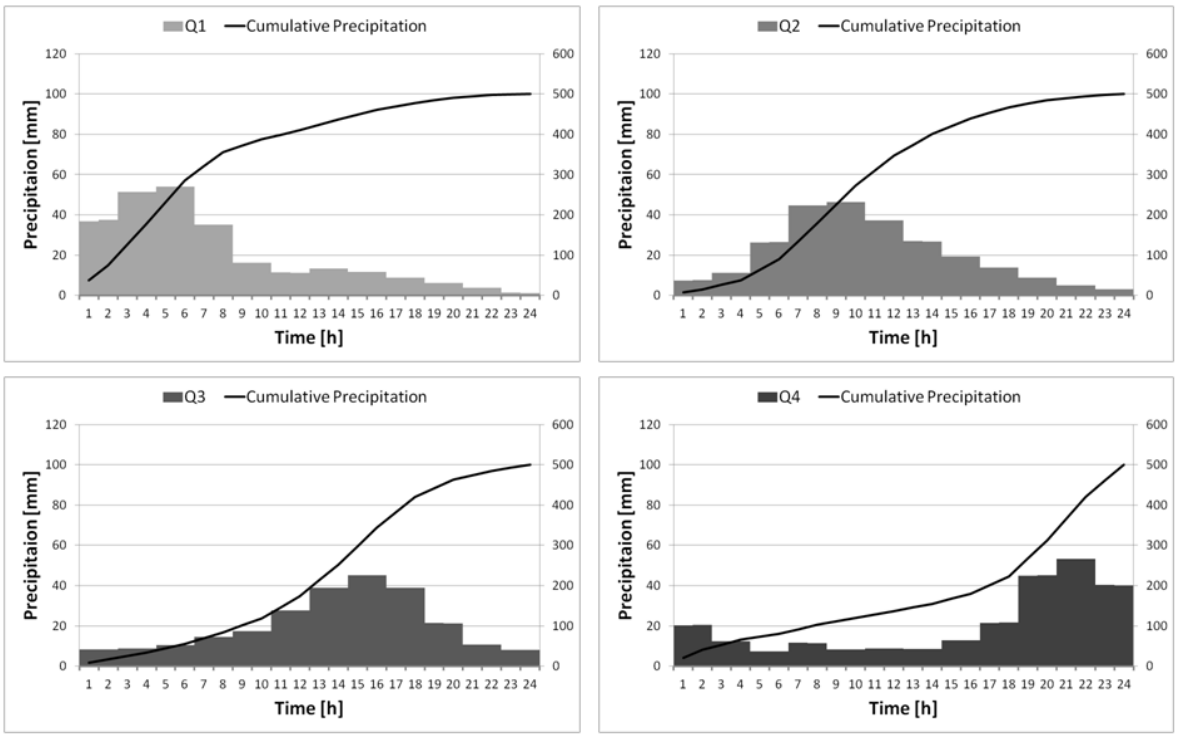

Figure 4. Hyetographs at Bisley tower location (indicated in figure 1a) for each type of rainfall events. Total precipitation volume is about 500 mm.

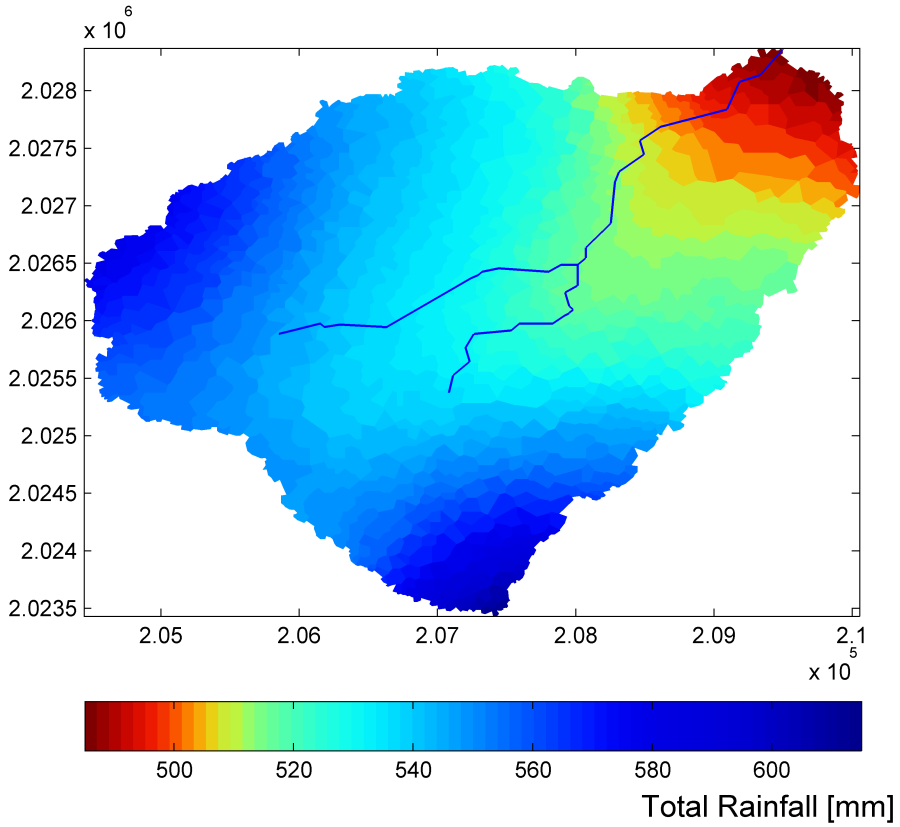

Figure 5. Spatial distribution of total rainfall estimate of hyetograph Q1. Total rainfall ranges from ~485 mm of the zone approaching to the basin outlet, to ~615 mm of the highest zones. The spatial coefficient of variation is 0.05.

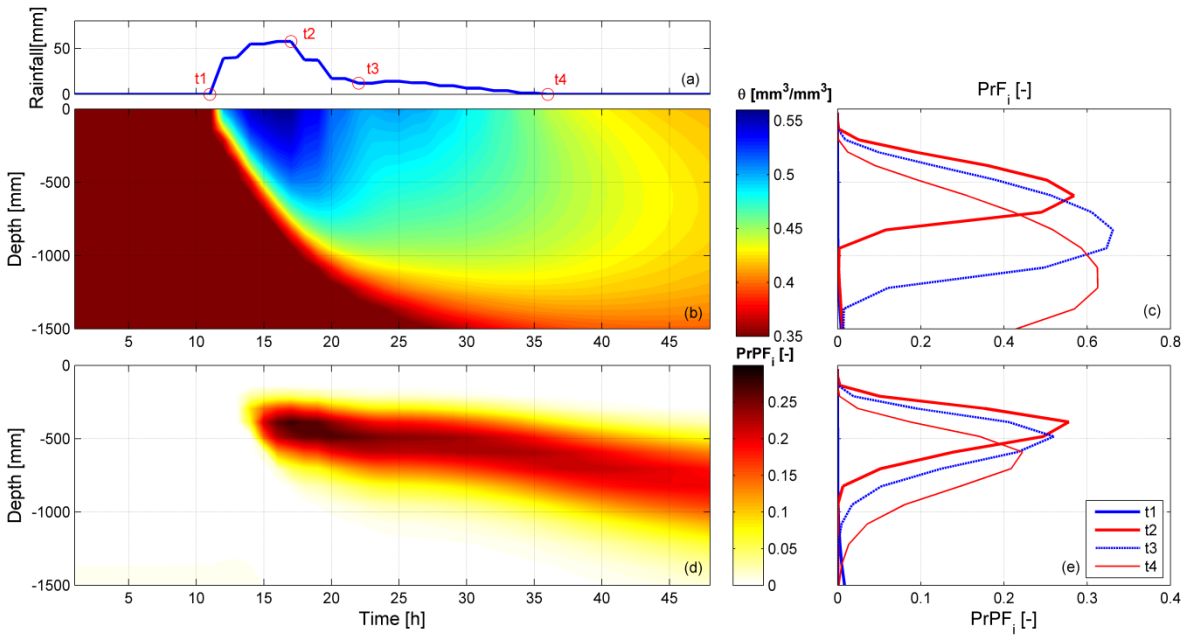

Figure 6. Time series model output response of a clay loam element (52°): rainfall type Q1 (a); volumetric soil content (θ) profile and time series (b); probability of failure at each layer at selected times, PrF_i (c); profile and time series of probability of plane of failure at given depth, $PrPF_i$ (d); $PrPF_i$ at selected times (e).

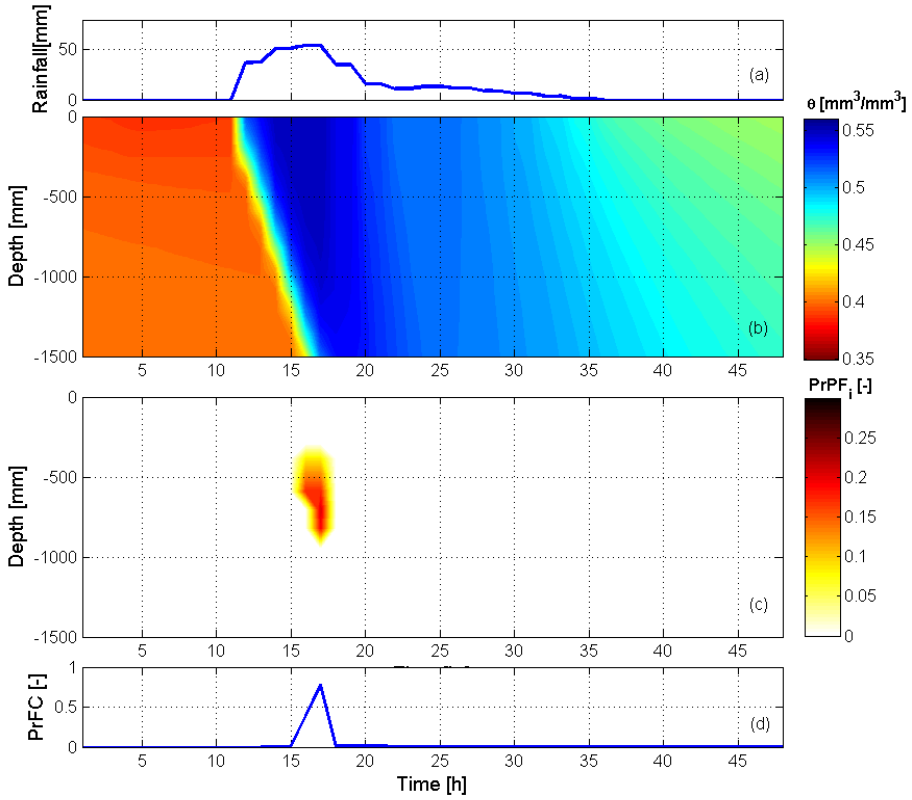

Figure 7. Time series model output response of a silty-clay element with gentler slope (~21°): rainfall type Q1 (a); volumetric soil moisture content (θ) profile and time series (b); profile and time series of probability of plane of failure at given depth, $PrPF_i$ (c); time series of the probability of failure of the column, *PrFC* (d) (see section 2.2).

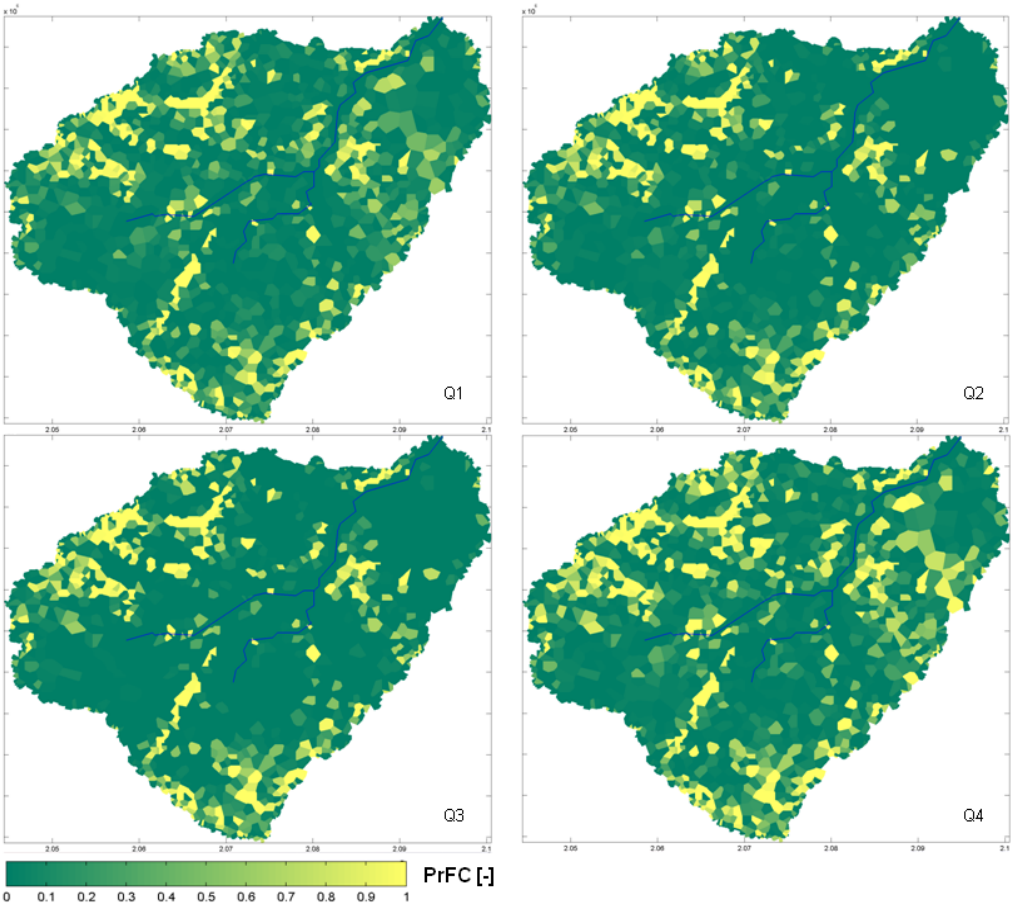

Figure 8. Spatial distribution of the maximum values of *PrFC* recorded at each pixel across the run time for each rainfall type.

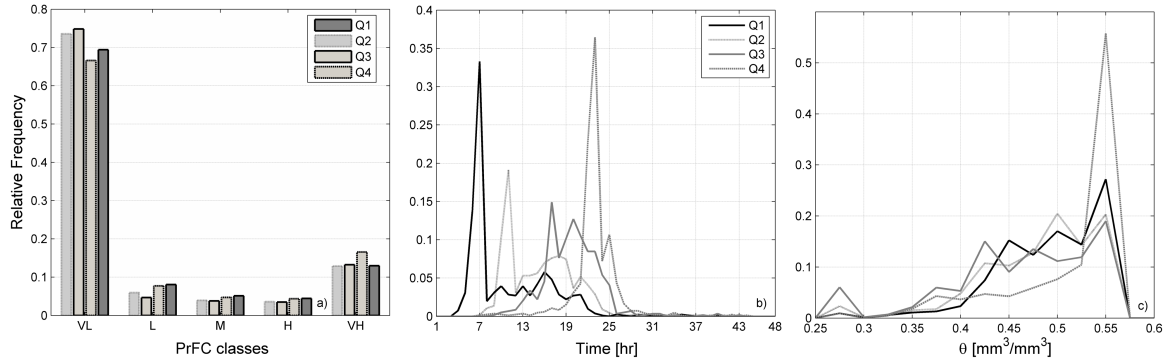

Figure 9. Relative frequency distribution across the basin of (a) the *PrFC* occurrence for the 4 rainfall events, (b) the time of occurrence of the maximum value of *PrFC* and (c) 1m depth average soil moisture at time of maximum *PrFC* values.

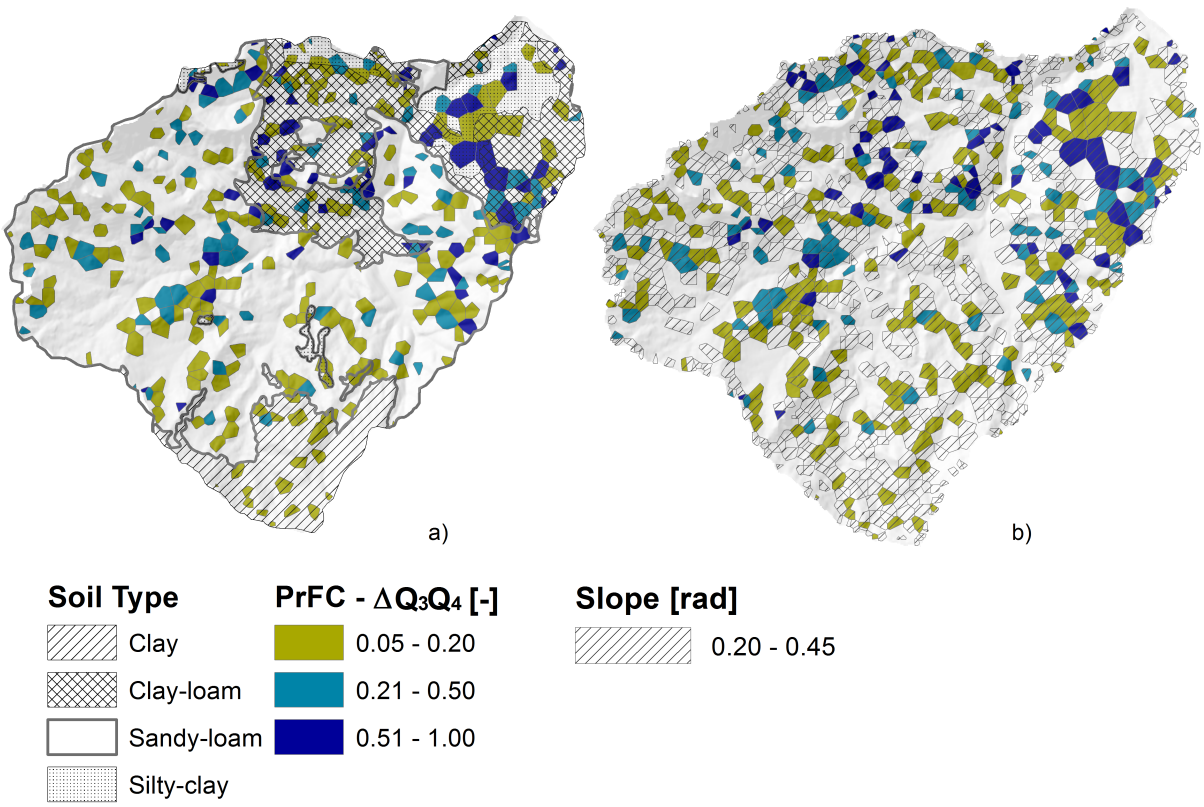

Figure 10. Spatial distribution of the difference of maximum *PrFC* between Q4 and Q3 ($\Delta Q_4 Q_3$) overlapped with the maps of soil type (a) and slope range 0.20 and 0.45 rad (~10° and 25°) (b).

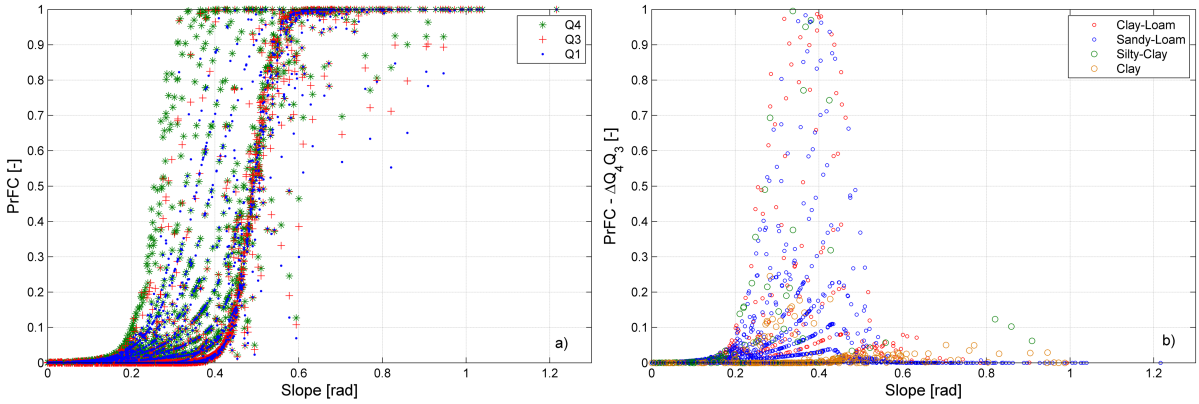

Figure 11. Distribution of maximum *PrFC* with slope for Q1, Q3 and Q4 (Q2 is omitted because it is similar to Q3). (a); distribution of difference of maximum *PrFC* between Q4 and Q3 ($\Delta Q_4 Q_3$) with slope (b).

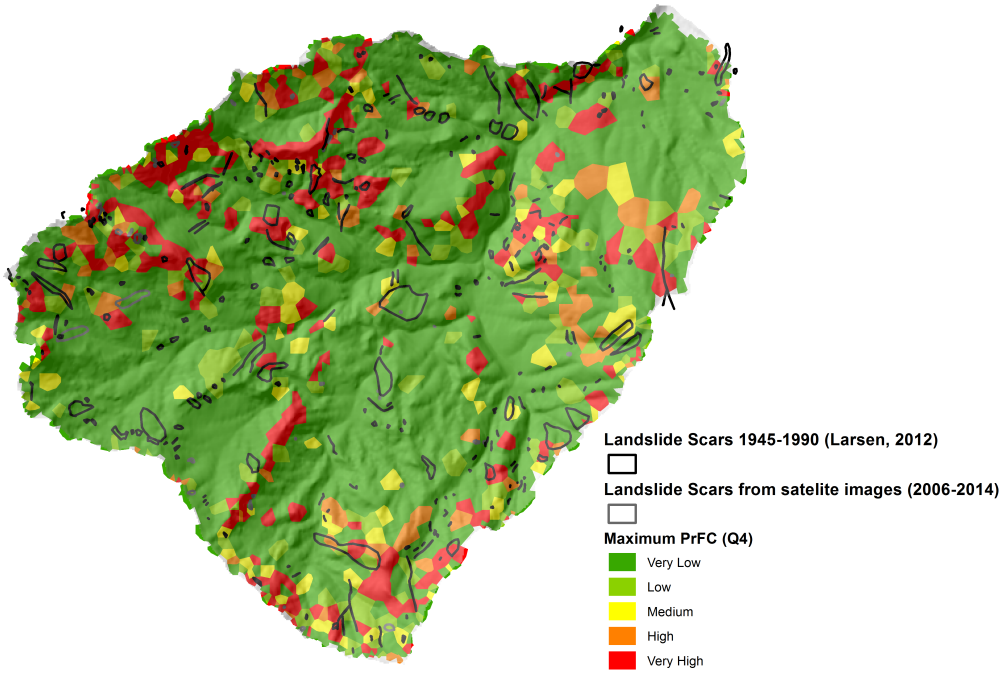

Figure 12. Comparison between historical landslide scars caused by intense or prolonged storms during the 20th century (Larsen, 2012) and observed through recent satellite images (from Google Earth) with the classified maximum *PrFC* of Q4.