Response to Referee #1

Palacio Cordoba, J., Mergili, M., and Aristizábal, E.: Probabilistic landslide susceptibility analysis in tropical mountainous terrain using the physically based r.slope.stability model, Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-223, in review, 2019.

We are grateful for the constructive comments on our paper. we are pleased to know that our article is well structured and of interest to the scientific community. We have included the suggestions in the revised manuscript. Please find below our responses in detail. Changes in the manuscript are highlighted in blue letter.

- I find a figure with the SHALSTAB and SHIA_Landslide models to be lacking.
In Figure 6, we show the results obtained with the models SHALSTAB and SHIA_Landslide. The SHALSTAB map was modified to 3 classes. considering a single class for "unstable" and "unconditionally unstable". for the construction of the maps is considered the range found according to the distance to the perfect classification. (L303 - 306).

- The statistical indexes of the table 4 are not defined in the "Data and procedure" section.
The statistical indexes of Table 4 are now defined. (L269-273).

- ROC curves for SHALSTAB and SHIA_Landslide models are missing. This validation technique is the most robust and, in my opinion, the most appropriate to elucidate which method has the best performance.
The ROC curve was added for the SHIA Landslide model as suggested. For the SHALSTAB model the ROC curve could not be realized as this model does not yield a factor of safety or any continuous value for each cell, like the other models evaluated. In SHALSTAB, each cell is classified as unconditionally stable, unconditionally unstable, or potentially unstable – a concept which is, unfortunately, incompatible with the concept of the ROC curve.

-The elaboration of ROC curves of each model needs a more detailed explanation. I am not sure how the A, B, C and D curves presented in figure 5 have been calculated.
The explanation was added and modified in the "Data and procedure" section. (L263 - 275)

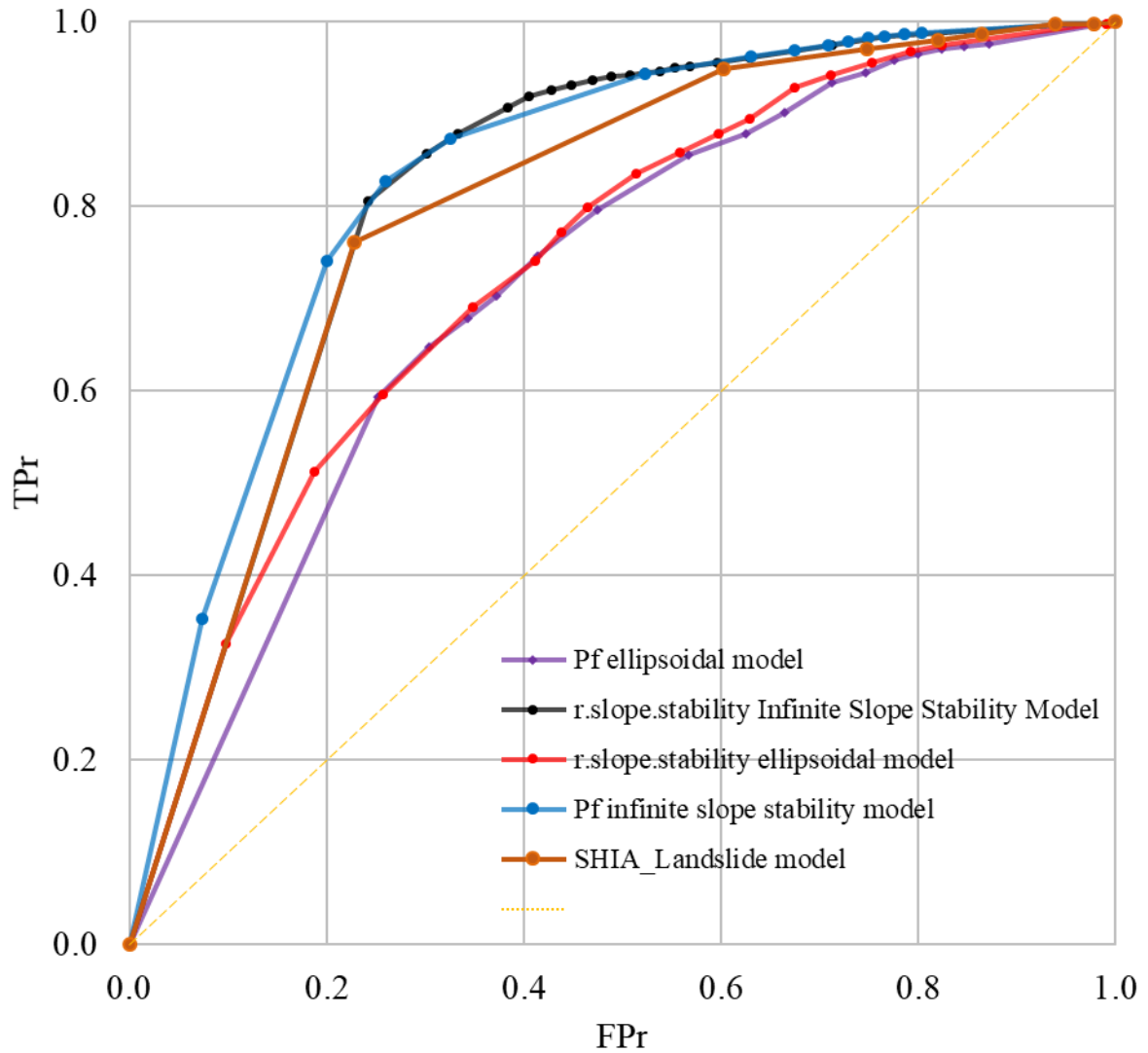
- I am surprised that authors do not discuss about some of the most relevant publications about the comparison of stability analysis methods implemented in GIS environments. I encourage authors to read and discuss the results obtained by Cervi et al. (2010) and Zizioli et al. (2013) among others.
We find the discussion by Cervi et al. (2010) on the different models very interesting, especially the discussion on SHALSTAB and considerations of the model (L376-377):
"Related to this aspect, Cervi et al., (2010) expose the limitations of SHALTAB due to the hydrological assumptions (flow in steady-state conditions)."

Also the paper of *Zizioli et al. (2013)* was very helpful (L105-106):
"In terms of scale or spatial resolution, physically-based models are suggested to be applied to finer-scale study areas and their results are strongly influenced by the level of detail in the input data (Zizioli et al., 2013)"

OTHER REMARKS:
-Line 46 delete "because of its equatorial location", it is a repetition.
Line 198. Change "Source: Adapted from (Aristizábal et al., 2016; Qiu et al., 2007)" by "Adapted from Aristizábal et al., (2016) and Qiu et al., (2007)".
-Line 316. Change "071" by "0.71".

Thank You, these errors were corrected!

-Figure 5. I think that the figure is confused. Please change the style of the figure to understand it at a glance.

-Change A, B, C and D codes by the name of the models. Use similar colors according to the type of the models.
We have updated Fig. 5 accordingly, also including the ROC curve for the result obtained with the SHIA_Landslide model.

Probabilistic landslide susceptibility analysis in tropical mountainous terrain using the physically based r.slope.stability model

Johnnatan Palacio Cordoba[1], Martin Mergili[2,3], Edier Aristizábal[4]

[1] Departamento de Ingeniería Civil, Facultad de Minas, Universidad Nacional de Colombia, sede Medellín, Colombia.
[2] Geomorphological Systems and Risk Research, Department of Geography and Regional Research, University of Vienna, Universitätsstraße 7, 1010 Vienna, Austria
[3] Institute of Applied Geology, University of Natural Resources and Life Sciences (BOKU), Peter-Jordan-Straße 82, 1190 Vienna, Austria
[4] Departamento de Geociencias y Medio Ambiente, Facultad de Minas, Universidad Nacional de Colombia, sede Medellín, Colombia

Correspondence: Johnnatan Palacio Cordoba (japalacioc@unal.edu.co)

Abstract. Landslides triggered by rainfall are very common phenomena in complex tropical environments such as the Colombian Andes, one of the regions of South America most affected by landslides every year. Currently in Colombia, physically based methods for landslide hazard mapping are mandatory for land use planning in urban areas. In this work, we perform probabilistic analyses with r.slope.stability, a spatially distributed, physically based model for landslide susceptibility analysis, available as an open-source tool coupled to GRASS GIS. This model considers alternatively the infinite slope stability model or the 2.5D geometry of shallow planar and deep-seated landslides with ellipsoidal or truncated failure surfaces. We test the model in the La Arenosa catchment, northern Colombian Andes. The results are compared to those yielded with the corresponding deterministic analyses and with other physically based models applied in the same catchment. Finally, the model results are evaluated against a landslide inventory using a confusion matrix and Receiver Operating Characteristic (ROC) analysis. The model performs reasonably well, the infinite slope stability model showing a better performance. The outcomes are, however, rather conservative, pointing to possible challenges with regard to the geotechnical and geo-hydraulic parameterization. The results also highlight the importance to perform probabilistic instead of – or in addition to – deterministic slope stability analyses.

1. Introduction

Landslides cause substantial human and economic losses every year (Kjekstad and Highland, 2009; Petley, 2012; Schuster and Highland, 2001). According to Dilley et al. (2005), the worldwide area exposed to landslides is around 3.7 million km^2, where 66 million people live in the 820,000 km^2 identified as the high-risk zone. Although economic losses tend to concentrate in industrialized and developed countries, the numbers of human fatalities and affected persons are highest in densely populated, less developed countries (Petley, 2012; Sepúlveda and Petley, 2015).

Landslides triggered by earthquakes and rainfall are a frequent phenomenon in mountainous terrain (Keefer et al., 1987; Van Westen et al., 2008; Varnes, 1978). Petley (2008) mentioned that in 2007, 89.6% of the fatalities due to landslides worldwide were related to rainfall-triggered landslides. In tropical environments and complex terrain such as the Colombian Andes, a high percentage of landslides are triggered by heavy or prolonged rainfall (Van 40 Westen and Terlien, 1996; Terlien, 1998).

Colombia, located in the north-western corner of South America, exhibits complex geographical and hydro-climatological features arising from its tectonic setting and equatorial location. The mountainous configuration of Colombia is the result of the Caribbean Plate moving south-westward relative to the South American Plate and the eastward subduction of the Nazca Plate beneath the northern Andes along the western margin of Colombia 45 (Kellogg et al., 1995; Taboada et al., 2000; Trenkamp et al., 2002). Related to the hydro-climatological conditions and because of its equatorial location, rainfall in Colombia is highly intermittent in space and time, influenced by the atmospheric circulation patterns over the neighboring tropical Pacific Ocean and the Caribbean Sea and the combined hydro-climatic and ecological dynamics of the Amazon and Orinoco basins (Poveda et al., 2007).

Shallow landslides triggered by rainfall are very common phenomena in tropical environments such as the 50 Colombian Andes, where hillslopes are characterized by deep weathering profiles and are subjected to periods of intense tropical rainfall (Aristizábal, 2013). In Colombia, landslide-prone regions are densely populated. As a consequence, hundreds of fatalities are associated with landslides triggered by rainfall every year (Sanchez and Aristizábal, 2018). Therefore, there is a strong social and economic need to include landslide susceptibility and hazard zoning in land use planning, to reduce landslide fatalities and economic losses.

55 According to the Emergency Events Database (2011), Colombia is one of the South American countries with most landslides. In the period between 1901 and 2017, 45 landslide disasters were registered with 3,619 fatalities, 78,395 people affected, and economic losses of 2.4 million USD. In the Global Landslide Catalog, Colombia has 87 entries with a total of 464 deaths (Kirschbaum et al., 2015) from 2007 to 2013. The Latin America and Caribbean landslide database has compiled a record of 110 fatal landslides in Colombia between the years 2004 and 2013 with a total of 60 880 deaths, where Colombia is the country with the second highest number of fatal landslides in Latin America and Caribbean, only Brazil shows a slightly higher number of fatal landslides (119) (Sepúlveda and Petley, 2015). Mergili et al. (2015) expose that Colombia, according to EM-DAT, presents a victim/event ratio of 77.34 (3171/41) and is only exceeded by Peru and Ecuador in this respect. However, the EM-DAT database only considers events with high numbers of fatalities: actually, the real number of fatal landslides is much higher (Aristizábal and Gómez, 65 2007; Mergili et al., 2015).

Landslide susceptibility assessment can be determined by qualitative or by quantitative methods (Corominas et al., 2014). Qualitative methods correspond to knowledge-driven approaches based entirely on the judgment of experts using geomorphological criteria in the field (van Westen et al., 2000). Quantitative approaches are subdivided into data-driven methods and physically-based models. Statistical or data-driven methods evaluate the relationship 70 between landslides and causative factors to predict the landslide spatial probability (Carrara, 1983; Gorsevski et al., 2000; Lee, 2005; Lee and Pradhan, 2007; Süzen and Doyuran, 2004). Physically based models for landslide susceptibility and hazard assessment of detailed areas include the interaction between hydrology, topography, soil

properties, and in some cases, vegetation in order to understand and predict the location and timing of landslide occurrence. Such models generally compute slope stability, using the Factor of Safety (*FoS*). *FoS* is given by the dimensionless ratio between the resisting forces and the driving forces (Lam and Fredlund, 1993). Most of the physically based models available in the literature build on the limit equilibrium concept and the assumption of a planar slope of infinite length with a potential failure surface parallel to the topographic surface (Chen and Chameau, 1983; Lam and Fredlund, 1993). However, the infinite slope stability approach is proposed only for shallow, planar sliding surfaces in friction-dominated soils and fails to capture the complexity of deep-seated landslide phenomena (Bishop, 1954; Carson & Kirkby, 1972; Crozier, 1986; Duncan & Wright, 2005).

Limit equilibrium models have been extended to three-dimensional (3D) failure surfaces: geometric shapes such as spheres or ellipsoids represent non-planar slip surfaces in a much better way and are important to consider in areas of complex lithological conditions or for soils with high cohesion values. The first 3D slope stability model was presented by Baligh and Azzouz (1975). Later, Chen and Chameau (1983) developed a method to analyse cohesive and frictional slopes with different pore water conditions. Dennhardt and Forster (1985) proposed a method using an ellipsoidal slip surface. Kalatehjari and Ali (2013) carried out a review of different 3D analysis models in which they exposed the fact that many of the methods considered the slope and slip surface as symmetrical shapes in order to determinate the static condition of equilibrium. Hovland (1997) presented a method for cohesive and frictional soils based on the Fellenius method (Fellenius, 1927): in this approach, the forces that act between columns are disregarded and *FoS* is determined by normal and shear forces that act at the bases of the columns (Lam and Fredlund, 1993).

Although *FoS* represents a quantitative – and seemingly objective – approach to evaluate slope stability, it has to be used carefully. Authors such a Kirsten (1983) mention that different values of FoS could be obtained from slopes with equal probability of failure. However, several software packages use this concept for 3D slope stability analysis, e.g. STAB3D (Baligh and Azzouz, 1975), 3D-PCSTABL (Thomaz, 1986), CLARA (O Hungr, 1988), and TSLOPE3 (Pyke, 1991). Most of these models include some limitations reducing the accuracy of *FoS* obtained (Stark, 2003). One of the most important limitations is that they were designed to analyses individual landslides or slopes; they are not appropriate for regional or catchment-scale slope stability analyses (Mergili et al., 2014b). A few 3D (or, strictly speaking, 2.5D) slope stability models in Geographic Information Systems (GIS) have been used for landslide susceptibility mapping (Carrara and Pike, 2008; Qiu et al., 2007; Xie et al., 2003). Recently, the r.slope.stability model, a C- and Python-based raster module in the open-source software GRASS GIS (GRASS Team, 2019) has been proposed. The r.slope.stability model considers the 2.5D geometry of the sliding surface for analyzing a number of randomly selected potential sliding surfaces that are ellipsoidal or truncated in shape (Mergili et al., 2014a, b), and also offers an implementation of the infinite slope stability model.

In terms of scale or spatial resolution, physically-based models are suggested to be applied to finer-scale study areas and their results are strongly influenced by the level of detail in the input data (Zizioli et al., 2013), whereas data-driven approaches are recommended for broader-scale landslide susceptibility analyses (van Westen et al., 2006; Corominas et al., 2014). In fact, the implementation of data-driven methods or physically-based models for the incorporation of landslide hazard mapping into land use planning has been regulated in several countries. In

110 Colombia, according to the Decree 1807/2014, the implementation of deterministic or probabilistic physically-based methods is obligatory in urban and urban expansion areas, whereas statistical and knowledge-driven models are only permitted for rural areas.

In order to meet this requirement, we have to obtain a more detailed knowledge on the suitability of the available physically-based landslide susceptibility models, with the final goal to use model ensembles in order to obtain a

115 broader, more robust picture of the landslide susceptibility conditions. We think that r.slope.stability could be a candidate to enrich such an ensemble, considering shallow planar and deep-seated ellipsoidal failure surfaces. Consequently, the scientific aims of the present study are (i) to evaluate the suitability of r.slope.stability for physically-based landslide susceptibility mapping in tropical mountainous terrain, and (ii) to identify its fit with other potentially suitable models, helping to learn about strengths, limitations, and uncertainties.

120 Following these aims, we present a probabilistic analysis of slope stability in GIS for modelling landslide susceptibility in a tropical and mountainous environment using the r.slope.stability model. This model is evaluated using a landslide inventory prepared after a major and destructive rainfall-triggered multi-landslide event in the La Arenosa catchment on 21 September 1990. A quantitative performance evaluation of the model by ROC analysis is carried out. The results are compared with those obtained through the corresponding deterministic analyses with

125 r.slope.stability, with SHALSTAB (Dietrich and Montgomery, 1998) and with SHIA_Landslide (Aristizábal et al., 2016), which represents a new model developed for tropical mountainous terrain.

2. Study area

The La Arenosa catchment, with an area of 9.9 km^2, is located on the north-western side of the Colombian Andes, at 1000–1900 m above sea level (Velásquez and Mejía, 1991; Aristizábal et al., 2016). The climate is tropical

130 humid with a mean annual temperature of 23°C and a mean annual precipitation of 4,300 mm. However, precipitation is highly variable between the different seasons, and between different years. The annual cycle of precipitation shows a bimodal period of rainfall (inter-annual scale) with rainfall peaks in the months of April (450 mm) and October (600 mm) (IDEAM, 2010). Rainfall often occurs in the afternoon and at night in the form of heavy rainstorms or cloudbursts of short duration (Aristizábal et al., 2015; Garcia, 1995). Most of the landslides in

135 the catchment are triggered by such events, even though it is often not the rainfall itself, but the response of the groundwater level – also depending on its initial state due to antecedent rainfall conditions – which is decisive for the occurrence of landslides (Fressard et al., 2016).

Although the natural vegetation of the La Arenosa catchment would correspond to very humid premontane forest, all the primary forest has been removed, and the lands are exclusively dedicated to agricultural use. In the highest

140 and steepest parts of the basin, the predominance of coffee crops, sugar cane, pastures and very small areas of secondary forest is maintained. This situation is considered a conditional cause of landslide occurrence.

Residual soils which have evolved from granodiorite rocks covered by slope deposits and fluvio-torrential deposits are characteristic for the area. Slope deposits cover approx. 15% of the catchment. Strong in situ weathering occurs due to chemical decomposition in the humid tropical climate (Velásquez and Mejía, 1991). Indicators of rapid,

145 extensive and progressive spheroidal decomposition of the granite are observed down to an average depth of 30 m.

The saprolite is fairly well graded. Its texture is described as sandy silt or silty sand with some gravel and a small fraction of clay. Relict joints in the parent rock are preserved in the saprolite. They facilitate preferential flow and therefore strongly influence the observed hydraulic conductivity of the surrounding soil matrix (INTEGRAL, 1990; Aristizábal et al., 2015). The matrix-supported slope deposits are formed by boulders of granite, residual soils, and vegetation debris (Aristizábal et al., 2015). Slope deposits generally accumulate at foot slopes or in gullies. Those usually poorly consolidated deposits are the consequence of past landslides. Their content in cobbles and boulders is high, and natural soil pipes are common (Velásquez and Mejía, 1991).

On 21 September 1990, the La Arenosa catchment was strongly affected by a rainfall event of high intensity and short duration. In less than 3 hours, 208 mm of precipitation, with a maximum intensity of 90 mm.h^{-1}, was recorded within the study area, triggering approximately 800 landslides. Based on the intensity-duration-frequency (IDF) curve, a return period of 200 years was estimated for this event (Velásquez and Mejía, 1991).

The strong rainfall in the catchment, imposed upon a general saturation of the soils in a rugged topography, triggered a series of almost simultaneous landslides of rotational or translational type in the catchment (Hermelin et al., 1992; Garcia, 1995). However, there is no distribution map which permits the differentiation between translational and rotational landslides. The population was strongly affected: 20 fatalities were counted and 260 people had to be evacuated, 27 houses were destroyed and 30 damaged, and so were several bridges (Hermelin et al., 1992; Aristizábal et al., 2016). Estimates arrived at more than 6 million US$ of total loss.

 INTEGRAL (1990) with Velásquez and Mejía (1991) analysed a set of aerial images and conducted a detailed field survey to produce a detailed landslide inventory for the event. However, it was not feasible to generate a complete inventory of landslides for the entire catchment as aerial photographs and topographic maps were not available for an area of approx. 2 km^2. Only the area covered by the landslide inventory was considered for this study, corresponding to an area of 7.6 km2.

A total of 699 landslides were identified and mapped in the La Arenosa catchment. According to Velásquez and Mejía (1991), most of the landslides started as shallow translational slides and transformed into debris flows, from very to extremely rapid with high water content (Fig. 1). The landslide bodies were small with respect to the flow length, and the slip surfaces were parallel to the slope surface. The majority of the landslides initiated within residual soils in hollows and open slopes with a slope steepness ranging from 35° to 42° (Velásquez and Mejía, 1991). The same authors described that the depth of failure surface was less than 3 m and corresponded to the contact of the residual soil with the saprolite.

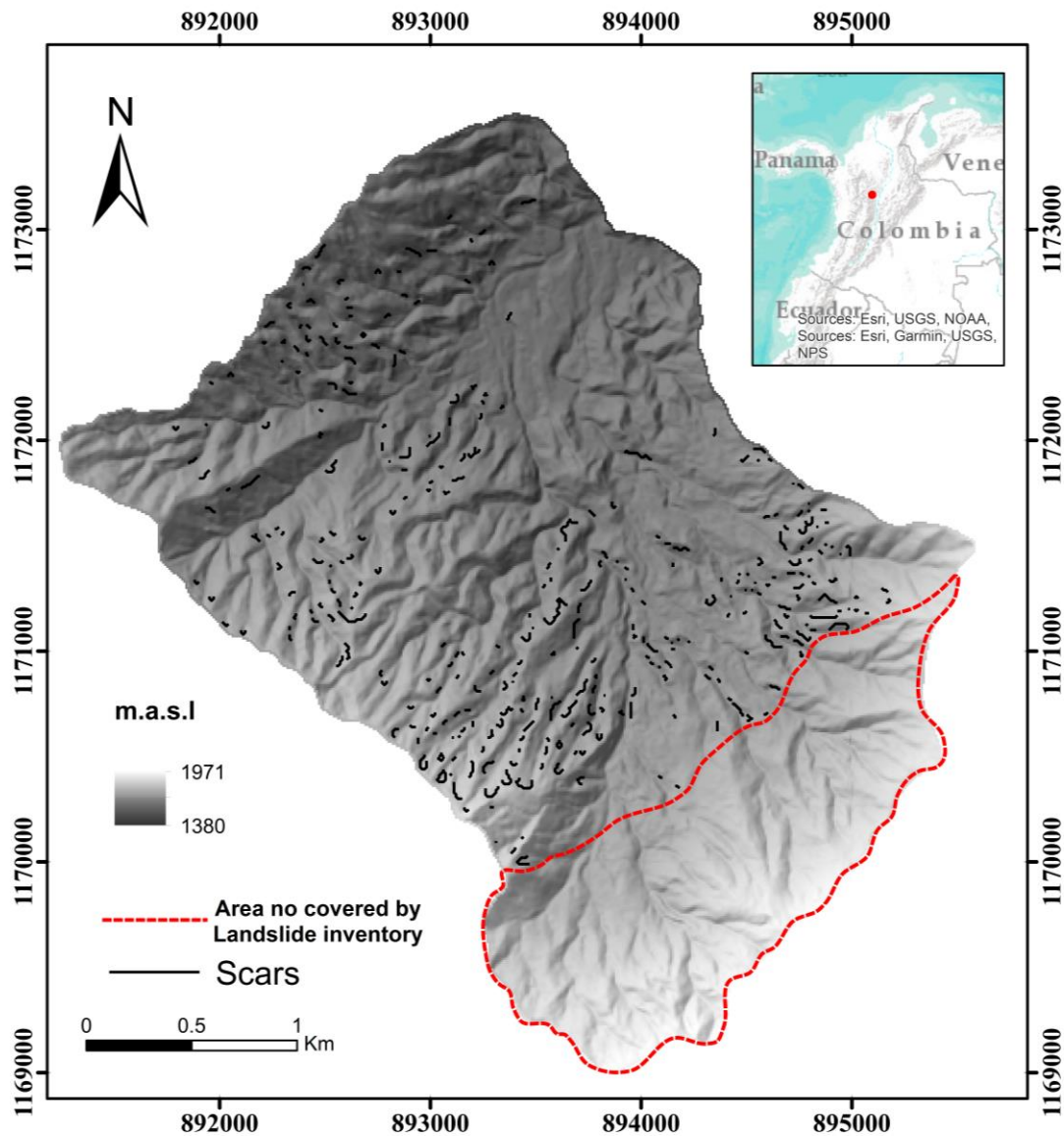



Figure 1. Landslides scars inventory database according to landslide inventory. The area with a red line does not have landslide inventory. The plot of a nonlinear surface. Source: Adapted from (INTEGRAL, 1990; Velásquez and Mejía, 1991)

3. The r.slope.stability model

The r.slope.stability model is a GIS-based, free, and open-source slope stability modelling software (www.slopestability.org) developed by Mergili et al. (2014a, b) as a C- and Python- based raster module of the GRASS GIS software package (GRASS Team, 2019). The tool is able to take into account planar failure with an infinite slope stability module and a slip surface module.

For shallow and planar landslides, r.slope.stability includes a classic infinite slope stability approach. For the infinite slope stability analysis, S acts parallel to the shear plane and seepage is considered parallel to the slope. For ellipsoid-shaped slip surfaces, in contrast, S is generally not parallel to the shear plane of the columns, even if it is parallel to the slope (Mergili et al., 2014b). The infinite slope stability model is run independently from the ellipsoidal failure surface analysis: FoS_{inf} for each raster cell is calculated according to Eq. (2).

$$FoS_{inf} = \frac{c' \cdot A + G' \cos ß \cdot \tan \varphi'}{G' \cos ß + S}, \qquad (2)$$

where β is the slope angle of the slip surface (corresponding to the inclination of the terrain).

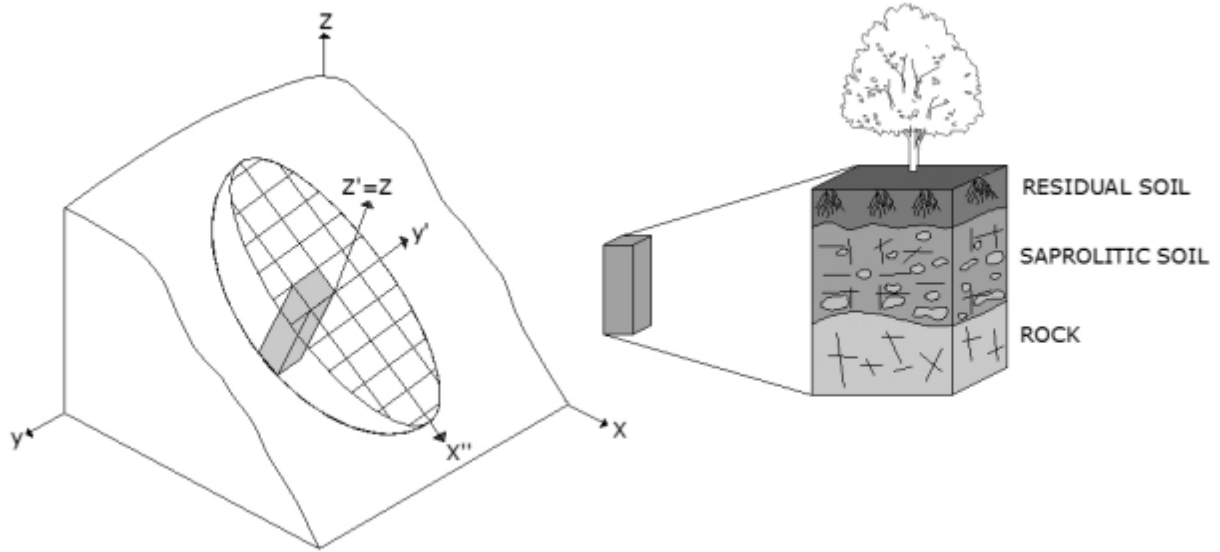
The slip surface model considers the 2.5D geometry of the sliding surface and evaluates FoS or the probability of slope failure (P_f) for many randomly selected potential ellipsoidal or truncated slip surfaces (Fig. 2). Each raster cell can be affected by various slip surfaces and is characterized by a unique value of FoS or P_f for each raster cell of the study area. Thereby, the lowest value of FoS or the highest value of P_f out of the values for all the sliding surfaces touching the cell is considered relevant. The model permits the users to impose restrictions with respect to the width, length, and depth of the ellipsoids (Mergili et al., 2014a, b).

The slip surface model used in r.slope.stability represents a revision and extension of the 2.5D sliding surface model of Hovland (1997) (Xie et al., 2003). The calculation of FoS is based on the basic principle of equilibrium (Eq. 1).

$$FoS = \frac{\sum (c' \cdot A + (G' \cos \beta_c + N_s) \tan\varphi') \cos\beta_m}{\sum (G' \sin \beta_m + T_s) \cos\beta_m}, \qquad (1)$$

where c' is the effective cohesion (N m^{-2}), G' is the weight of the moist soil (N), β_c is the inclination of the slip surface, φ' is the effective internal friction angle, β_m is the apparent dip of the sliding surface in direction of the aspect α, N_s and T_s (N) are the contributions of the seepage force to the normal force and the shear force, respectively, and A (m^2) is the slip surface area assigned to each column. Inter-column forces and external forces (e.g seismic loading) are neglected (Mergili et al., 2014a, b). The slip surface model is further based on the model of King (1989), in which the direction of the seepage force (S) corresponds with the direction of the hydraulic gradient, approximated by slope and aspect of the groundwater table.

Both the infinite slope stability model and the slip surface model can be used for a probabilistic analysis, applying a range of geotechnical parameters (c', φ') and the depth of the failure surface. Thereby, rectangular, normal, log-normal, or exponential probability density functions are employed. The result is a probability of failure (P_f), representing the fraction of tested parameter combinations yielding $FoS<1$ (Mergili et al., 2014a).

Note that, in the present work, with probabilistic model, we always refer to the random variation of the geotechnical parameters. The slip surface model includes a probabilistic component where the dimensions of the ellipsoids are randomly varied for the computation of FoS.

220 **Figure 2.** The typical ellipsoid used, slip surface with a model column in r.slope.stability, typical weathering profile of tropical environments and complex terrains. Source: Adapted from Aristizábal et al. (2016) and Qiu et al. (2007).

4. Data and procedure

The input of the r.slope.stability model consists of a Digital Terrain Model (DTM), spatial datasets of the mechanic and hydraulic characteristics of the study area, and finally the restraints imposed to the model by the user, based on

225 the knowledge of the study area. A DTM with a spatial resolution of 10 x 10 m was provided by the Instituto Geografico Agustin Codazzi. A soil thickness map was built using interpolation, employing an empirical relationship between soil thickness and slope angle in the study area (Aristizábal, 2013; Catani et al., 2010; Thiery et al., 2019). The computed residual soil depth ranges from 1 m to 2.8 m.

The La Arenosa catchment is basically composed of two soil types, alluvial and residual, with properties strongly

230 related to the parental material (IGAC, 2007). Alluvial soils cover 6.7% of the total area; they correspond to quaternary deposits composed of alluvial sediments of moderate depth, limited by the presence of fragments of rock and gravel. Residual soils cover 93.3% of the total area; they are derived from igneous rocks such as granites and quartz-diorites. The residual soils are medium to fine textured, well-drained, in some cases characterized by gravel or stones in the profile. These soils have deep weathering profiles depending on parent rock lithology and local

235 conditions, partly reaching down to a depth of 100 m (Aristizábal et al., 2005; Suarez, 1998). The geotechnical parameters were obtained based on studies and laboratory tests carried out in La Arenosa by Velásquez and Mejía (1991). For the residual soils, cohesion values range between 5 kPa and 12.5 kPa, whereas the internal friction angle of the soil ranges from 16° to 24°. The dry unit weight ranges from 14.3 to 14.9 kN.m^{-3}. No geotechnical laboratory tests are available for the alluvial deposits; however, they show very gentle slopes generally not prone to

240 landsliding; the cohesion and friction angle values were therefore assumed based on literature values (Ameratunga et al., 2016; Aristizábal, 2013; Aristizábal et al., 2015; Geotechdata, 2013).

Table 1. Geotechnical parameters of La Arenosa catchment from laboratory tests. γ_d = specific weight dry of soil, c'=effective cohesion, φ'= effective angle of internal friction and θ_s =saturated water content minimum and maximum values for c' and φ' presented

	Alluvial soil	Residual soil
θ_s (%)	25	40
γ_d (kN/m^3)	17	14.9
c' (kN/m^2)	1 (0.6-1.4)	5 (3-7)
φ' (°)	34 (30-38)	24 (21-27)
Depth (m)	2.5-2.8	1.2-2.8

Adapted from (Aristizábal, 2013; Aristizábal et al., 2016)

The r.slope.stability model is applied with the probabilistic approach and deterministic approach for comparison with other models tested in the catchment. Both are used in combination with the infinite slope stability model and the slip surface model, resulting in a total of four model runs. Rectangular probability density functions for c', φ', and the soil depth d (m) are considered for the probabilistic analysis. The rectangular distribution is suitable for representing random variables which have known upper and lower bounds and an equal likelihood of occurring anywhere between these bounds (Fenton and Griffiths, 2008). For each parameter (c', φ', and d) we use a sample size of ten values, which are randomly selected from the ranges in Table 1.

With the infinite slope stability model, *FoS* for each raster cell is calculated with respect to the bottom of the soil. The analysis with truncated ellipsoidal failure surfaces is performed in a procedure with 1000 simulated surfaces touching each cell. Preliminary tests have indicated that this value represents a good compromise between computational efficiency and accuracy of the results. Each ellipsoidal slip surface is defined by the coordinates of the centre and the variable lengths of the three half-axes, the aspect, the inclination and the offset centre above the terrain, the failure surfaces include widths between 10 and 100 m and lengths between 10 and 200 m, with a maximum truncated depth of 2.8 m.

Quantitative evaluation of the empirical adequacy of the r.slope.stability model results are accomplished through a confusion matrix and an ROC analysis for the continuous output. Each raster cell of both the observed data (inventory) and the predicted data (model result) is assigned to one of two classes: stable (no landslide mapped or $FoS \geq 1$, respectively) or unstable (landslide mapped or $FoS < 1$, respectively) (Fawcett, 2005). The overlay of the classes is known as confusion matrix, where each raster cell is assigned to one of four classes: true positive (observed unstable cell is predicted as unstable), true negative (observed stable cell is predicted as stable), false positive (observed stable cell is predicted as unstable) and false negative (observed unstable cell is predicted as stable).

The statistical indexes measuring the performance (Table 4) are the true positive rate or hit rate (*TPr*) is defined as the ratio between the true positives and the observed positives. The true negative rate or specificity (*TNr*) is the ratio between the true negatives and the observed negatives. The false positive rate or false alarm rate (*FPr*) is

defined as the ratio between the false positives and the observed negatives, and the positive predictive value, also called the precision, is the ratio between the true positives and the total predicted positives (Aristizábal et al., 2015, 2016). Evaluation only considers the area covered by the landslide inventory. The ROC analysis plots *TPr* against

275 *FPr* for various threshold levels of P_f.

r.slope.stability is compared with the SHALSTAB and SHIA_Landslide models to evaluate the consistency of the results. The SHALSTAB model, developed by Montgomery and Dietrich (1994), applies a topographic index to estimate the saturation of the soil as a function of rainfall infiltration. This procedure builds on the assumption that surface topography can be used as a main indicator of landslide susceptibility (Aristizábal et al., 2015). The model

280 employs the hydrological model TOPOG which uses steady-state rainfall and an infinite slope approach for the geotechnical component (O'Loughlin, 1986). SHIA_Landslide is a physically based and conceptual model, developed by Aristizábal et al. (2016), for computing positive pore pressure changes as well as the resulting changes in *FoS* due to rainfall infiltration, coupling a distributed hydrological model with a classical infinite slope stability model.

285 **5.   Results**

The results using the deterministic analysis with the infinite slope stability model in r.slope.stability are shown in Figure 3A, whereas the results obtained with the slip surface model are shown in Figure 3B, both of them in terms of *FoS*. Table 2 shows the confusion matrix calculated by comparing the deterministic analysis results with the scars in the landslide inventory map. For the infinite slope stability model, unstable conditions with *FoS* < 1 are

290 shown for 79.2% of the catchment area, whereas only 10.5% show acceptably stable conditions with *FoS* values > 1.5; these areas correspond to the lower parts of the catchment formed by alluvial sediments with very gentle slopes. With regard to the slip surface model, 84% of the catchment area show *FoS* < 1, and only 5.8% show acceptable stability conditions with *FoS* > 1.5.

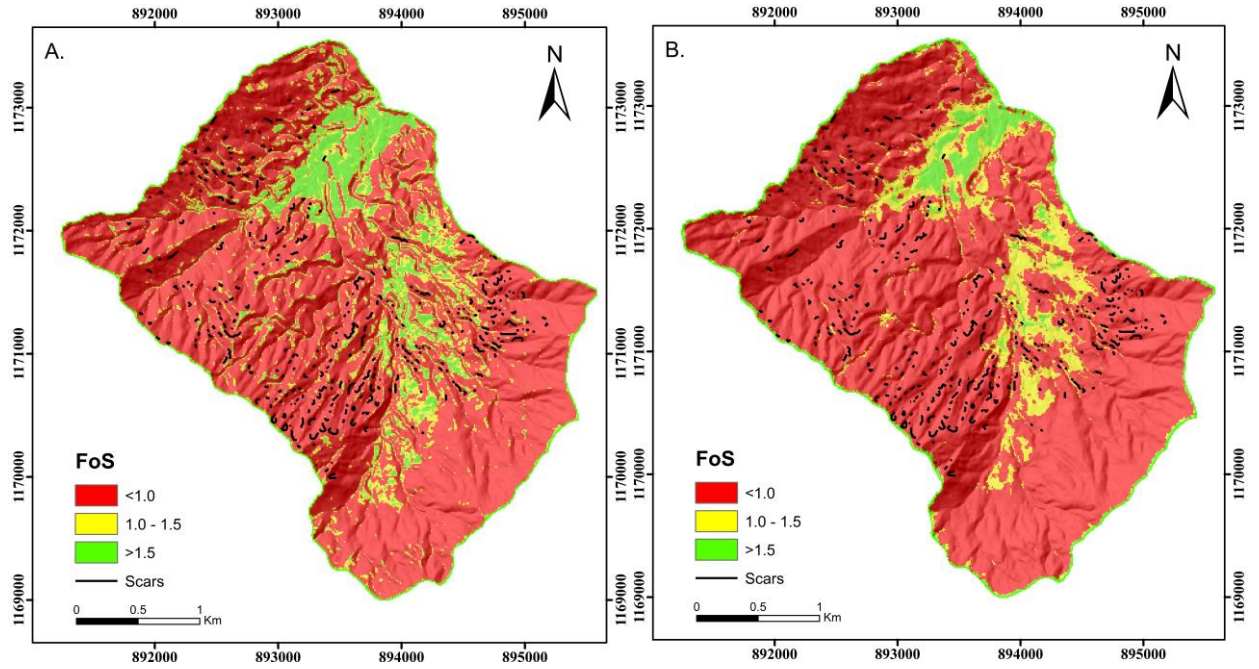



Figure 3. Landslide susceptibility classified according to *FoS*. Deterministic analyses of r.slope.stability (infinite slope stability model) (**A**) and r.slope.stability (ellipsoid-based model) (**B**).

Figure 4 illustrates the results of the probabilistic component of r.slope.stability used with the infinite slope stability model (Fig. 4A) and with the slip surface model (Fig. 4B) in terms of P_f. This probability is computed as the proportion of parameter combinations predicting $FoS < 1.0$ at a specific raster cell. Table 3 shows the confusion matrix calculated by comparing the probabilistic analysis results with the scars in the landslide inventory map. To define the critical P_f thresholds, the distance to perfect classification parameter (r) proposed by Medina-cetina and Cepeda (2010) is used:

$$r = \sqrt{(FPr)^2 + (1 - TPr))^2} \qquad (3)$$

The threshold values yielding the lowest value of r, indicating the best model performance, are used to discriminate between predicted positive and predicted negative cells. We are fully aware that this is a purely statistical optimization approach, not necessarily meaningful from a geotechnical point of view – an issue that will be further elaborated in the Discussion.

295

300

305

310

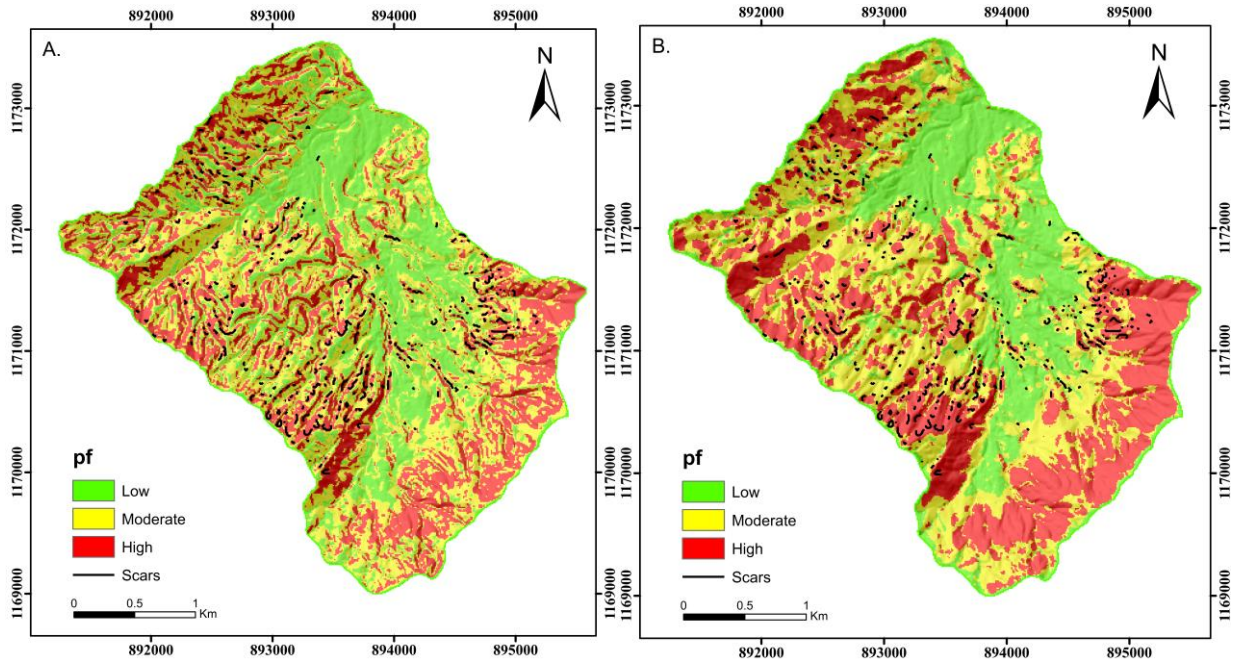



Figure 4. Probability of failure for r.slope.stability (infinite slope stability model) (**A**) and r.slope.stability (ellipsoid-based model) (**B**).

For the infinite slope stability model a minimum value of $r = 0.31$ is obtained for $P_f = 0.96$, whereas for the slip

315 surface model a minimum value of $r = 0.46$ is obtained for $P_f = 0.99$. For the infinite slope stability model, unstable conditions with $P_f > 0.96$ are shown just for 30.5% of the catchment area, whereas 69.5% show stable conditions ($FoS \geq 1.0$). 36.4% of the catchment display values of $P_f > 0.99$ according to the slip surface model, whereas 63.6% display stable conditions. Unstable hillslopes, according to this criterion, are located mostly in the southern portion of the catchment, where no landslide inventory is available.

320

Table 2. The confusion matrix for the deterministic analysis

Classifier	r.slope.stability (ellipsoid-based model)			
	Pixel	Area (m²)	Total percentage	Partial percentage
Observed landslide areas				
TP	2089	208900	2.7%	95%
FN	100	10000	0.1%	5%
Observed non-landslide areas				
TN	17314	1731400	22.7%	23%
FP	56838	5683800	74.5%	77%

Classifier	r.slope.stability (Infinite slope stability model)			
	Pixel	Area (m^2)	Total percentage	Partial percentage
Observed landslide areas				
TP	2154	215400	2.8%	98%
FN	35	3500	0.05%	2%
Observed non-landslide areas				
TN	17326	1732600	22.7%	23%
FP	56826	5682600	74.4%	77%

325   According to the confusion matrix, the deterministic analysis of r.slope.stability correctly predicts 98% and 95% of the observed landslide areas with the infinite slope stability model and the slip surface model, respectively. The other 2% and 5% are predicted as stable but did experience landslides according to the inventory. However, for the observed non-landslide areas, only 23% are correctly predicted as stable by the infinite slope model and slip surface model, whereas the other 77%, which are predicted as unstable, did not fail according to the inventory. The

330   deterministic model is more efficient in correctly classifying slopes where landslides occurred and less efficient at classifying slopes on which landslides did not occur. With the thresholds of r applied in this study, the confusion matrix for the probabilistic analysis shows a correct prediction of 83% and 65% of the observed landslide areas with the infinite slope stability model and the slip surface model, respectively. The other 17% and 35% are erroneously predicted as stable according to the inventory. For the observed non-landslide areas 74% and 70% are

335   correctly predicted as stable by the infinite slope model and slip surface model, respectively, whereas only 26% and 30% are predicted as unstable, but did not fail.

Table 3. The confusion matrix for the probabilistic analysis

Classifier	r.slope.stability (ellipsoid-based model)			
	Pixel	Area (m^2)	Total percentage	Partial percentage
Observed landslide areas				
TP	1416	141600	1.9%	65%
FN	773	77300	1.0%	35%
Observed non-landslide areas				
TN	51623	5162300	67.6%	70%
FP	22529	2252900	29.5%	30%

Classifier	r.slope.stability (Infinite slope stability model)			
	Pixel	Area (m^2)	Total percentage	Partial percentage
Observed landslide areas				
TP	1812	181200	2.4%	83%

FN	377	37700	0.5%	17%
Observed non-landslide areas				
TN	54890	5489000	71.9%	74%
FP	19262	1926200	25.2%	26%

340 The area under the ROC curve (AUC) values indicate a good ability of the probabilistic results to distinguish between susceptible and less susceptible areas (Fig. 5). The infinite slope stability model yields areas under the ROC curve of 0.82 and 0.83 for the deterministic and probabilistic analyses, respectively, whereas the slip surface model performs worse, but still fair (0.73 and 0.71).

345 **Table 4.** Statistical indexes measuring the performance of r.slope.stability and the other models

	Hit rate (0-100)	False alarm rate (0-100)	Specificity (0-100)	Precision
r.slope.stability (P_f ellipsoidal based model)	65	30	70	0.06
r.slope.stability (P_f infinite slope stability model)	83	26	74	0.08
r.slope.stability (FoS infinite slope stability model)	98	77	23	0.04
r.slope.stability (Fos ellipsoid-based model)	95	77	23	0.04
SHIA_Landslide	77	24	76	0.07
SHALSTAB	91	42	58	0.05

Adapted from (Aristizábal et al., 2015; Aristizábal et al., 2017)

Table 4 summarizes the statistical indices measuring the performance and prediction of the r.slope.stability model compared to SHALSTAB and SHIA_Landslide for the La Arenosa catchment. Figure 5 focuses on the performance

350 of the area under curve ROC for the r.slope.stability model compared to SHIA_Landslide in the La Arenosa catchment, the better yield corresponds to the infinite slope stability model.

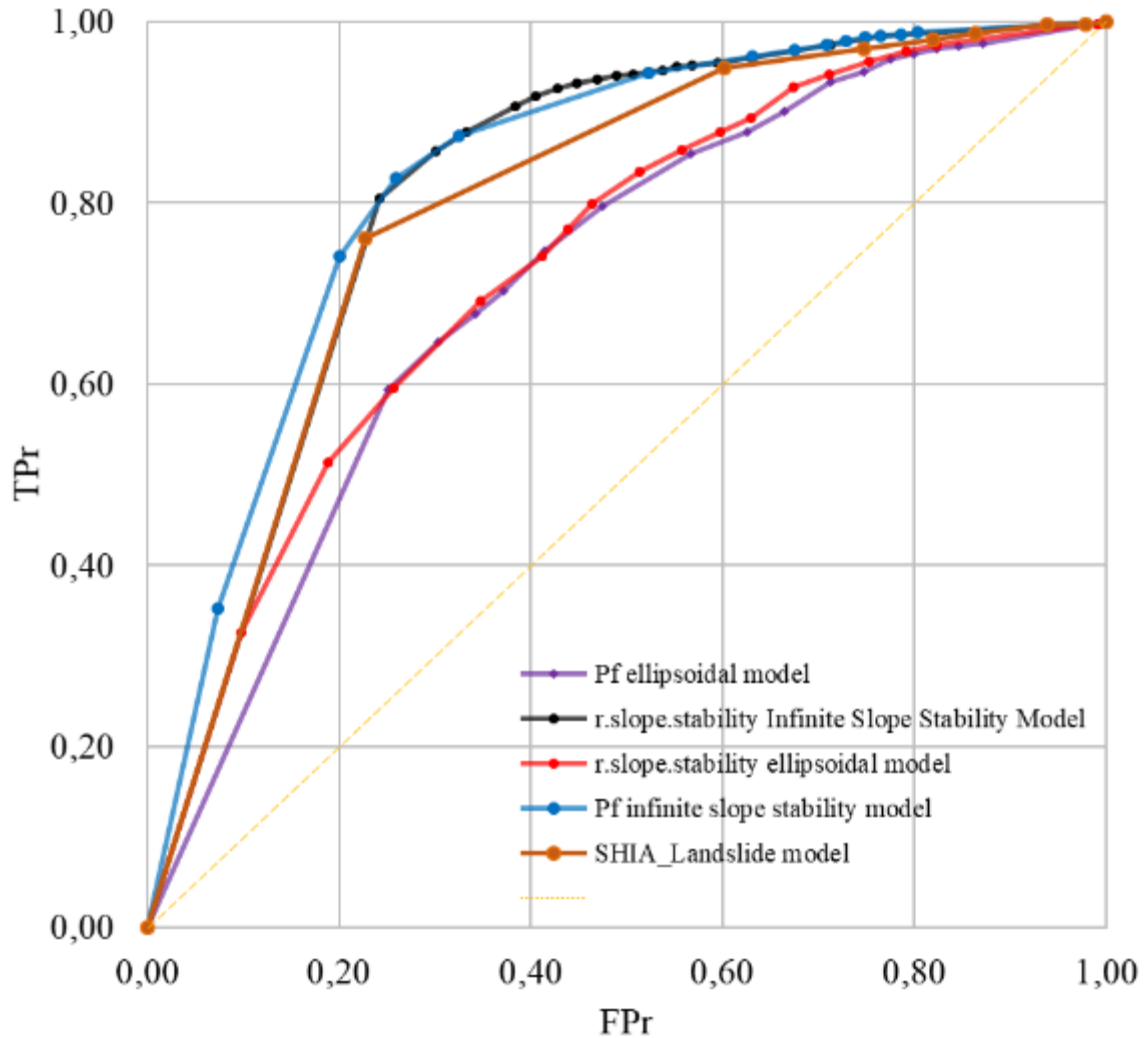



Figure 5. ROC curve models, r.slope.stability (infinite slope stability model) deterministic curve (0.83) and P_f (0.82) for r.slope.stability (ellipsoid-based model) deterministic curve (0.73) and P_f (0.71), SHIA_Landslide (0.77).

Figure 6 compares the results of r.slope.stability with SHALSTAB and SHIA_Landslide for a specific area of the catchment. SHALSTAB shows more areas classified as unconditionally unstable and unstable displaying similarities to the deterministic analysis of r.slope.stability, whereas SHIA_Landslide and the probabilistic analysis of r.slope.stability tend to show fewer areas with $FoS < 1$ or with high values of P_f. In summary, the deterministic analysis of the r.slope.stability model shows considerably higher hit rates, compared to the probabilistic analysis of r.slope.stability, SHALSTAB and SHIA_Landslide. However, both deterministic analyses of r.slope.stability show the highest false alarm rates. The major reasons for this result – and for the similar patterns yielded with SHIA_Landslide – are most likely an underestimate of the geotechnical stability of the material, the neglect of the effects of vegetation, and/or the overestimate of water saturation. Geotechnical testing is often performed on material which is disturbed in the one or the other way, so that the resulting parameters do not necessarily represent

the natural conditions over larger areas. Roots could lead to some degree of stabilization, and it could also be the vegetation retaining sufficient water to avoid full saturation of the soil throughout the catchment.

370   **6.   Discussion**

In this work, the r.slope.stability tool was applied to the La Arenosa catchment, where the models SHALSTAB and SHIA_Landslide had been tested before (Aristizábal et al., 2015, 2016). An ideal model performance simultaneously maximizes *TPr* and minimizes *FPr*. In the La Arenosa catchment, the failure area associated with the rainfall event under investigation corresponds just to 2.2% of the whole catchment. Considering this situation,
375   SHALSTAB tends to predict more unstable areas for this specific rainstorm, increasing the prediction capacity of the model but at the same time increasing the false positive error. Related to this aspect, Cervi et al., (2010) expose the limitations of SHALTAB due to the hydrological assumptions (flow in steady-state conditions). By contrast, the SHIA_Landslide model shows a strong capacity for prediction and a very low *FPr*. In the case of r.slope.stability, the probabilistic analysis with the infinite slope stability model replicates the September 21, 1990 event with a very
380   good reliability, in a much more successful way compared to SHALSTAB, and to the deterministic r.slope.stability approach; and a similar performance of SHIA_Landslide with a higher hit rate and a slightly higher value of *FPr*.
An important advantage of r.slope.stability compared to SHALSTAB and SHIA_Landslide is the possibility of carrying out a probabilistic analysis in terms of considering ranges of the key model parameters. Besides r.slope.stability, there are several physically-based models which provide a probabilistic module, such as TRIGRS
385   (Baum et al., 2008) and ALICE (Thiery et al., 2017). Much more work has to be done in this direction. Measuring geotechnical and hydraulic parameters for large areas is difficult, time-consuming, and expensive, and there is an inherent variability in parameters associated with lithology and soil formation processes (Canli et al., 2017). Similarly, soil thickness shows a very high variability and uncertainty. This means there is a lot of uncertainty related to the horizontal and vertical natural variations of soil hydraulic and geotechnical parameters (Christian and
390   Baecher, 2001). In general, soil properties show a pseudo-random pattern rather than a constant value. Additionally, landslides are more complex than their representation in the physically-based models adopted, and the geometrical and mechanical parameters that control slope stability are not known with sufficient accuracy (Griffiths et al., 2012; Guzzetti, 2016).
The conventional deterministic approach neglects uncertainties in the slope stability analysis. Although the *FoS*
395   computation is more likely to identify areas prone to slope failure during a given hydro-meteorological event rather than to predict the exact locations of specific landslides (Baum et al., 2010), *FoS* is often not a reliable indicator of the slope stability conditions because it is – in terms of interpretation – a binary value derived from several uncertain parameters (Chowdhury, 2009). Thus, considering that physically based models are very sensitive to soil properties and soil depth, the probability distribution of failure is a much better indicator of the slope stability
400   conditions.
Probabilistic analyses permit the inclusion of natural soil variations in the analysis, but also the mechanism of failure is fundamental for obtaining adequate results. In the case of the September 21, 1990 rainstorm, the infinite

slope stability analysis of the r.slope.stability model using planar shallow failure surfaces shows a much better performance. The results obtained show that the failure geometry is most appropriately approximated by shallow

405   planar surfaces, which is confirmed by the reference information: Even though Hermelin et al. (1992) and Garcia (1995) mentioned that both shallow and deep-seated landslides were observed after the studied rainfall event, Velásquez and Mejía (1991) underlined the dominance of shallow landslides. This also agrees with the type of landslides often triggered by short, heavy rainfall events causing a rapid increase in pore pressure (Crosta, 1998). Such landslides are characterized by small and shallow, slope-parallel failure planes (depth of 0.3–2 m) (Anderson

410   and Sitar, 1995). The displaced material, by means of processes of static liquefaction and rapid reduction of shear strength in undrained conditions, develops into flows that spread downward (Anderson and Sitar, 1995; Wang and Sassa, 2003). This type of landslide is described by Hungr et al. (2014) as debris avalanche. In cases where failures occur on slip surfaces with curved shapes, the assumption of ellipsoid-shaped slip surfaces would be expected to provide better results. This indicates the necessity of evaluating a priori the mechanism of failure and then

415   employing the most appropriate model, or – if both mechanisms are relevant – evaluate the model results with different sub-datasets of the inventory. In the present study, due to the lacking distinction of landslide mechanisms in the inventory at the level of individual landslides, we applied both the infinite slope stability model and the slip surface model in order to cover a broad range of mechanisms. The result of the evaluation against the entire landslide inventory was in line with the generally reported dominance of shallow translational landslides.

420   The current version of r.slope.stability does not permit variation of the saturation level, meaning that the analysis has to be carried out in either dry or saturated conditions. In the r.slope.stability model, including the role of the infiltration process under saturated conditions in particular, could strongly improve the model performance in terms of being more effective in considering local hydrological conditions which govern slope instability processes (Mergili et al., 2014a, b), given that the required data are available.

425

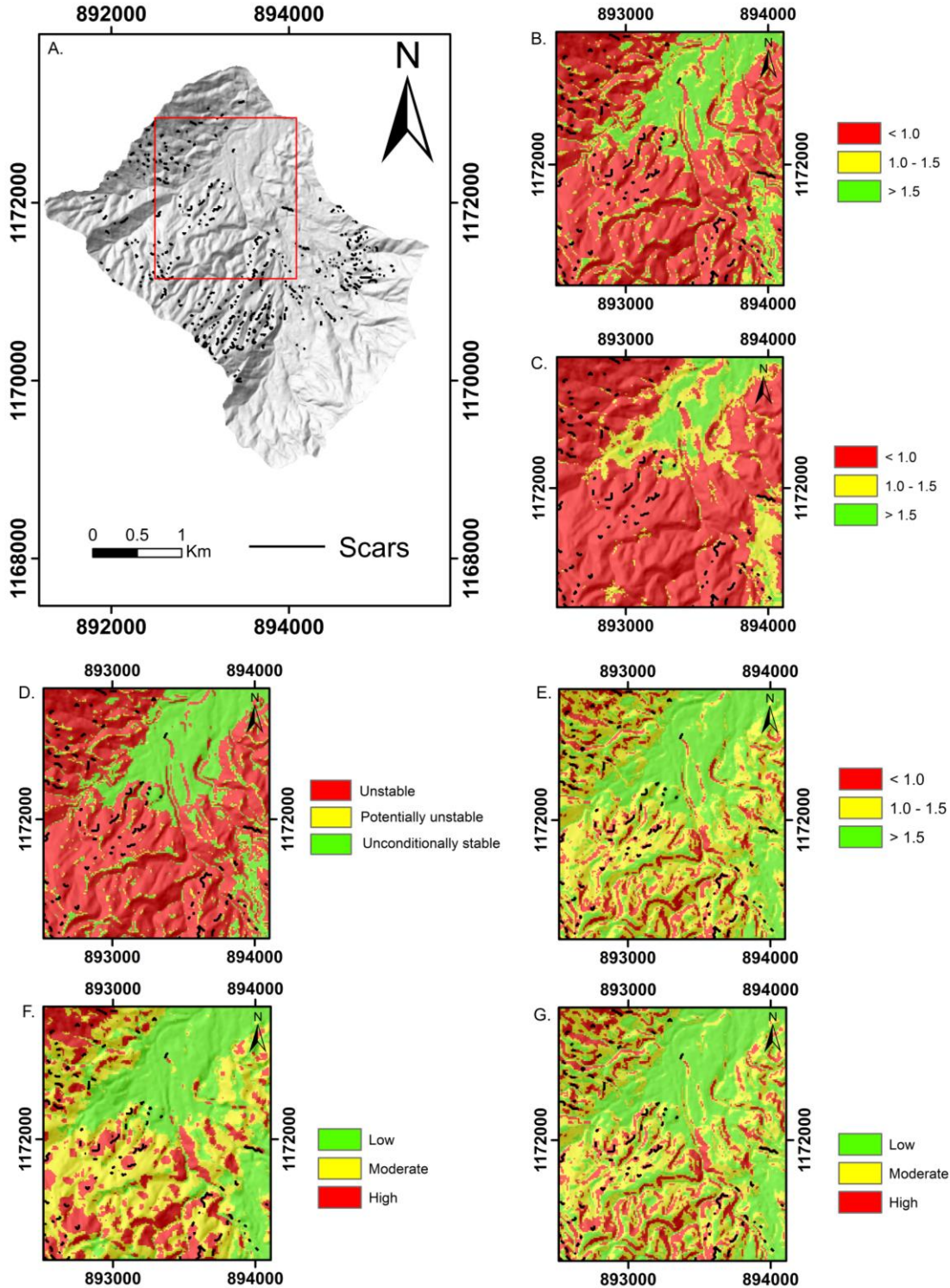



Figure 6. Comparison of results. (**A**) Analysis area (**B**) *FoS* obtained with r.slope.stability (infinite slope stability model) (**C**) *FoS* obtained with r.slope.stability (slip surface model) (**D**) Analysis with SHALSTAB (Aristizábal et al., 2015) (**E**) *FoS* obtained with SHIA_Landslide (Aristizábal et al., 2016) (**F**) P_f obtained with r.slope.stability (ellipsoid-based model) (**G**) P_f obtained with r.slope.stability (infinite slope stability model).

7.   Conclusions

In this work, we have presented the results of r.slope.stability for the La Arenosa catchment in the Colombian Andes. r.slope.stability is a 2.5D slope stability model capable of dealing with shallow and deep-seated landslides triggered by rainfall. The model was evaluated with a set of observed landslides triggered by the 21 September 1990 rainstorm, for which different slope stability models had been previously applied. Considering the probabilistic analyses performed, r.slope.stability shows a high hit rate, suggesting an acceptable prediction capacity for failure areas (83%–65%) for the infinite slope stability model and the slip surface model, respectively. The false alarm rate is relatively low for the infinite slope stability model (26%) and for the slip surface model (30%). The areas under the ROC curve yielded by the probabilistic approach are 0.83 for the infinite slope stability model and 0.71 for the slip surface model. These results clearly suggest a higher performance for the assumption of shallow, planar failure surface (infinite slope stability) model than for the deep-seated slip surface model, a finding which is in line with the physical characteristics of the observed landslides. Despite the generally good model performance, the results were far too conservative, compared to the observations, meaning that either (1) the assumption of the saturation patterns was inappropriate; or (2) the geotechnical parameters fed into the model are not representative for the study area. The same challenges were identified for the SHALSTAB model. Future studies shall further elaborate on this issue.

Compared to many other models, r.slope.stability has the advantage that it supports the derivation of a failure probability in terms of considering ranges of the key model parameters, instead of fixed values. Since in any landslide susceptibility analysis it is necessary to consider that soil parameters and their spatial variability are highly uncertain, the computation of failure probabilities in addition to *FoS* is highly recommended. In Colombia, hazard mapping is mandatory for use land planning in urban areas. Results like the slope failure probability (Fig. 6F and Fig. 6G) are considered suitable for this purpose. However, they only represent one step on a long way:

(i)   Much of the landslide risk in the Arenosa catchment is not so much related to the failure of unstable slopes, but rather to the downslope propagation of the mobilized material as debris flows (Velásquez & Mejía, 1991). Therefore, coupling of the slope stability model results with mass flow simulation tools is absolutely necessary (Mergili et al., 2017; Bout et al., 2018).

(ii)   Slope failure probabilities – or impact indicator indices (Mergili et al., 2017) – neither take into account the dimension of time nor the fine-scale patterns of the geotechnical characteristics of the catchment. Consequently, those results have to be interpreted in a relative, qualitative rather than an absolute, quantitative sense. The combination of these maps with expert judgement is therefore essential to define suitable thresholds separating the study area into different zones with their individual hazard levels and recommendations for action.

Acknowledgements. The authors are grateful to the anonymous reviewers for their valuable suggestions.

References

Ahmed, A., Ugai, K., and Yang, Q. Q.: Assessment of 3D Slope Stability Analysis Methods Based on 3D Simplified Janbu and Hovland Methods. *International Journal of Geomechanics*, *12*(2), 81–89. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000117, (2012).

470   Aleotti, P., and Chowdhury, R.: Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the Environment*, *58*(1), 21–44, (1999).

Ameratunga, J., Sivakugan, N., and Das, B. M.: *Correlations of Soil and Rock Properties in Geotechnical Engineering*. https://doi.org/10.1007/978-81-322-2629-1, (2016).

Anderson, S. A., and Sitar, N.: Analysis of Rainfall-Induced Debris Flows. *Journal of Geotechnical Engineering*, 475   *121*(12), 544–552. https://doi.org/10.1061/(ASCE)0733-9410(1995)121, (1995).

Aristizábal, E.: *SHIA _ Landslide: Developing a physically based model to predict shallow landslides triggered by rainfall in tropical environments.* Universidad Nacional de Colombia, (2013).

Aristizábal, E., García, E., and Martínez, H.: Susceptibility assessment of shallow landslides triggered by rainfall in tropical basins and mountainous terrains. *Natural Hazards*, *78*(1), 621–634. https://doi.org/10.1007/s11069-480   015-1736-4, (2015).

Aristizábal, E., García, E., and Martínez, H.: Modelling Shallow Landslides Triggered by Rainfall in Tropical and Mountainous Basins. *Springer International Publishing AG*, (2017).

Aristizábal, E., and Gómez, J.: Inventario De Emergencias Y Desastres En El Valle De Aburrá. Originados Por Fenómenos Naturales Y Antropicos En El Periodo 1880-2007. *Gestión y Ambiente*, *10*, 17–30, (2007).

485   Aristizábal, E., Roser, B., and Yokota, S.: Tropical chemical weathering of hillslope deposits and bedrock source in the Aburrá Valley, northern Colombian Andes. *Engineering Geology*, *81*(4), 389–406. https://doi.org/10.1016/j.enggeo.2005.08.001, (2005).

Aristizábal, E., Vélez, J. I., Martínez, H. E., and Jaboyedoff, M.: SHIA_Landslide: a distributed conceptual and physically based model to forecast the temporal and spatial occurrence of shallow landslides triggered by 490   rainfall in tropical and mountainous basins. *Landslides*, *13*(3), 497–517. https://doi.org/10.1007/s10346-015-0580-7, (2016).

Baligh, M. M., and Azzouz, A. S.: End Effects of Stability of Cohesive Slopes. *Geotechnical Engineering Division, ASCE*, *101*(11), 1105–1117, (1975).

Baum, R. L., Godt, J. W., and Savage, W. Z.: Estimating the timing and location of shallow rainfall - induced 495   landslides using a model for transient , unsaturated infiltration. *Journal of Geophysical Research*, *115*(F3), F03013. https://doi.org/10.1029/2009JF001321, (2010).

Bishop, A. W.: The use of the Slip Circle in the Stability Analysis of Slopes. *Géotechnique*, *5*(1), 7–17. https://doi.org/10.1680/geot.1955.5.1.7, (1954).

Bout, B., Lombardo, L., van Westen, C. J., & Jetten, V. G.: Integration of two-phase solid fluid equations in a 500   catchment model for flashfloods, debris flows and shallow slope failures. *Environmental Modelling and Software*, *105*, 1–16. https://doi.org/10.1016/j.envsoft.2018.03.017, (2018).

Canli, E., Loigge, B., and Glade, T.: Spatially distributed rainfall information and its potential for regional landslide

early warning systems. *Natural Hazards*, *91*, 103–127. https://doi.org/10.1007/s11069-017-2953-9, (2017).

Carrara, A.: Multivariate models for landslide hazard evaluation. *Journal of the International Association for Mathematical Geology*, *15*(3), 403–426. https://doi.org/10.1007/BF01031290, (1983).

Carrara, A., and Pike, R. J.: GIS technology and models for assessing landslide hazard and risk. *Geomorphology*, *94*(3–4), 257–260. https://doi.org/10.1016/j.geomorph.2006.07.042, (2008).

Carson, M., & Kirkby, M.: Hillslope Form and Process. In *Cambridge University*. https://doi.org/10.1017/S0016756800000546, (1972).

Catani, F., Segoni, S., & Falorni, G.: An empirical geomorphology-based approach to the spatial prediction of soil thickness at catchment scale. *Water Resources Research*, *46*(5), 1–15. https://doi.org/10.1029/2008WR007450, (2010).

Cervi, F., Berti, M., Borgatti, L., Ronchetti, F., Manenti, F., & Corsini, A.: Comparing predictive capability of statistical and deterministic methods for landslide susceptibility mapping: A case study in the northern Apennines (Reggio Emilia Province, Italy). *Landslides*, *7*(4), 433–444. https://doi.org/10.1007/s10346-010-0207-y, (2010).

Chen, R. H., and Chameau, J.: Three-dimensional limit equilibrium analysis of slopes. *Géotechnique*, *33*(1), 31–40. https://doi.org/10.1680/geot.1983.33.1.31, (1983).

Chowdhury, R., Flentje, P. and Bhattacharya, G.(1): Geotechnical Slope Analysis, CRC Press, London, https://doi.org/10.1201/9780203864203, (2009).

Christian, J. T., and Baecher, G. B.: Discussion of "Factors of Safety and Reliability in Geotechnical Engineering." *Journal of Geotechnical and Geoenvironmental Engineering*, *127*. https://doi.org/10.1061/(asce)1090-0241(2001)127, (2001).

Corominas, J., Van Westen, C. J., Frattini, P., Cascini, L., Malet, J.-P., Fotopoulou, S., Catani, F., … Smith, J. T.: Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment*, *73*(2), 209–263. https://doi.org/10.1007/s10064-013-0538-8, (2014).

Crosta, G.: Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology*, *35*(2–3), 131–145. https://doi.org/10.1007/s002540050300, (1998).

Crozier, M.: *LANDSLIDES: CAUSES, CONSEQUENCES AND ENVIRONMENT*. (R. K. & Paul, Ed.). https://doi.org/10.1111/j.1745-7939.1989.tb01143.x, (1986).

de Lima Neves Seefelder, C., Koide, S., and Mergili, M.: Does parameterization influence the performance of slope stability model results? A case study in Rio de Janeiro, Brazil. *Landslides*, (July), 1–13. https://doi.org/10.1007/s10346-016-0783-6, (2016).

Dennhardt, M., and Forster, W.: Problems of Three Dimensional Slope Stability. *The 11th International Conference in Soil Mechanics and Foundation Engineering, Part 2*, 427–431. San Francisco, (1985).

Dietrich, W. E., and Montgomery, D. R..: SHALSTAB: a digital terrain model for mapping shallow landslide potential. *NCASI (National Council of the Paper Industry for Air and Stream Improvement) Technical Report*, *1998*, (1998).

Dilley, M., Chen, R. S., Deichmann, U., Lener-Lam, A. L., Arnold, M., Agwe, J., … Yetman, G.: Natural Disaster

540   Hotspots A Global Risk Analysis. In *World Bank*. https://doi.org/10.1080/01944360902967228, (2005).

Duncan, J. M., & Wright, S. G.: *Soil Strength and Slope Stability*. New York: John Wiley & Sons, (2005)

EM-DAT, (The International Disaster Database).: *The OFDA/CRED International Disaster Database*, (2011).

Fawcett, T.: *An introduction to ROC analysis*. *35*(6), 299–309. https://doi.org/10.1016/j.patrec.2005.10.010, (2005).

Fellenius.: Erdstatische Berechnungen mit Reibung und Kohäsion (Adhäsion) und unter Annahme
545   kreiszylindrischer Gleitflächen. *W. Ernst and Sohn*, (1927).

Fenton, G. A., and Griffiths, D. V.: Risk Assessment in Geotechnical Engineering. In *Wiley*. https://doi.org/10.1002/9780470284704, (2008).

Fressard, M., Olivier, M., Thiery, Y., Davidson, R., & Lissak, C.: Multi-method characterisation of an active landslide: Case study in the Pays d'Auge plateau (Normandy, France). *Geomorphology*, *270*.
550   https://doi.org/10.1016/j.geomorph.2016.07.001, (2016).

Garcia, G.: Daños Ocurridos en la central hidroelectrica de Calderas por la avalancha en la quebrada La Arenosa. *I Taller Latinoamericano Reducción de Los Efectos de Los Desastres Naturales En Infraestructura Energetica.*, 13, (1995).

Geotechdata.: GeoParameter. Retrieved from http://www.geotechdata.info/parameter/parameter.html, (2013).

555   Gorsevski, P.V.; Gessler, P.;Foltz, R. B.: Spatial Prediction of Landslide Hazard Using Discriminant Analysis and GIS. *4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4); Problems, Prospects and Research Needs*, (2000).

GRASS Team.: GRASS Development Team 2011. Geographic Resources Analysis Support System (GRASS) Software. Open Source Geospatial Foundation Project. Retrieved from http://grass.osgeo.org, (2019).

560   Griffiths, D. V., Huang, J., and Fenton, G. A.: Risk assessment in geotechnical engineering: Stability analysis of highly variable soils. *Geotechnical Special Publication*, (226 GSP), 78–101. https://doi.org/10.1061/9780784412138.0004, (2012).

Guzzetti, F.: Forecasting natural hazards, performance of scientists, ethics, and the need for transparency. *Toxicological and Environmental Chemistry*, *98*(9), 1043–1059.
565   https://doi.org/10.1080/02772248.2015.1030664, (2016).

Hermelin, M., Mejía, O., and Velásquez, E.: Erosional and depositional features produced by a convulsive event, San Carlos, Colombia, September 21, 1990. *Bulletin of the International Association of Engineering Geology*, *45*(1), 89–95. https://doi.org/10.1007/BF02594908, (1992).

Hovland, H. J.: Three-dimensional slope stability analysis method. *Geotechnical Engineering Division, ASCE*,
570   *103*(9), 971–986, (1997).

Hungr, O.: CLARA: Slope Stability Analysis in Two or Three Dimensions. *O. Hungr Geotechnical Research, Inc., Vancouver, B.C,* (1988).

Hungr, Oldrich, Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update. *Landslides*, *11*(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y, (2014).

575   IDEAM.: *Climagramas Antioquia*. Retrieved from http://atlas.ideam.gov.co/basefiles/antioquia_graf.pdf, (2010).

IGAC.: *Estudio general de suelos y zonificación de tierras del departamento de Antioquia.* (p. 207). p. 207.

Bogotá: Instituto Geográfico Agustín Codazzi, (2007).

INTEGRAL.: *Informe sobre daños en la central de calderas por la avalancha ocurrida en l quebrada La Arenosa el 21 de septiembre de 1990 y su reparación. Internal Report Interconexión eléctrica SA*, (1990).

580 Kalatehjari, R., and Ali, N.: A review of three-dimensional slope stability analyses based on limit equilibrium method. *Electronic Journal of Geotechnical Engineering*, *18 A*(March), 119–134, (2013).

Keefer, D. K., Wilson, R. C., Mark, R. K., Brabb, E. E., William, M., Ellen, S. D., … Zatkint, R. S.: Real-Time During Warning Landslide Rainfall Heavy. *Science*, *238*, 921–925, (1987).

Kellogg, J. N., Vega, V., Stailings, T. C., and Aiken, C. L. V.: Tectonic development of Panama, Costa Rica, and
585 the Colombian Andes: Constraints from Global Positioning System geodetic studies and gravity. *Geological Society Of*, 75–90. https://doi.org/10.1130/SPE295-p75, (1995).

King, G. J. W.: Revision of effective-stress method of slices. *Géotechnique*, *39*(3), 497–502, (1989).

Kirschbaum, D., Stanley, T., and Zhou, Y.: Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, *249*, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016, (2015).

590 Kirsten, H. A. D.: Significance of the Probability of Failure in Slope Engineering. *Civil Engineer in South Africa*, *25*(1), (1983).

Kjekstad, O., and Highland, L.: Economic and Social Impacts of Landslides. In K. Sassa & P. Canut (Eds.), *Landslides – Disaster Risk Reduction* (pp. 573–587). https://doi.org/10.1007/978-3-540-69970-5, (2009).

Lam, L., and Fredlund, D. G.: A general limit equilibrium model for three-dimensional slope stability analysis.
595 *Canadian Geotechnical Journal*, *30*(6), 905–919. https://doi.org/10.1139/t93-089, (1993).

Lee, S.: Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis. *Geosciences Journal*, *9*(1), 63–71, (2005).

Lee, S., and Pradhan, B.: Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides*, *4*(1), 33–41. https://doi.org/10.1007/s10346-006-0047-y, (2007).

600 Medina-cetina, Z., and Cepeda, J.: Probabilistic classification of local rainfall-thresholds for landslide triggering. *EGU General Assembly*, *12*, (2010).

Mergili, M., and Fellin, W.: Three-Dimensional Modelling of Rotational Slope Failures with GRASS GIS. *Landslide Science and Practice*, *3*, 359–366. https://doi.org/10.1007/978-3-642-31445-2, (2013).

Mergili, M., Marchesini, I., Alvioli, M., Metz, M., Schneider-muntau, B., Rossi, M., and Guzzetti, F.: A strategy
605 for GIS-based 3-D slope stability modelling over large areas. *Geoscientific Model Development*, *7*(6), 2969–2982. https://doi.org/10.5194/gmd-7-2969-2014, (2014)a.

Mergili, M., Marchesini, I., Rossi, M., Guzzetti, F., and Fellin, W.: Spatially distributed three-dimensional slope stability modelling in a raster GIS. *Geomorphology*, *206*, 178–195. https://doi.org/10.1016/j.geomorph.2013.10.008, (2014)b.

610 Mergili, M., Santiago, C., and Moreiras, S.: Causas, características e impacto de losprocesos de remoción en masa, en áreascontrastantes de la región Andina. *REVISTA COLOMBIANA DE GEOGRAFÍA*, *24*, 113–131, (2015).

Mergili, M., Fischer, J. T., Krenn, J., & Pudasaini, S. P.: r.avaflow v1, an advanced open-source computational framework for the propagation and interaction of two-phase mass flows. *Geoscientific Model Development*,

10(2), 553–569. https://doi.org/10.5194/gmd-10-553-2017, (2017).

615

Montgomery, D. R., and Dietrich, W. E.: A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, *30*(4), 1153–1171. https://doi.org/10.1029/93WR02979, (1994).

O'Loughlin, E. M.: Prediction of Surface Saturation Zones in Natural Catchments by Topographic Analysis. *Water Resources Research*, *22*(5), 794–804. https://doi.org/10.1029/WR022i005p00794, (1986).

620 Petley, D.: the global occurrence of fatal landslides in 2007. *Geophysical Research Abstracts, EGU General Assembly 2008*, *10*, (2008).

Petley, D.: Global patterns of loss of life from landslides. *Geology*, *40*(10), 927–930. https://doi.org/10.1130/G33217.1, (2012).

Poveda, G., Vélez, J. I., Mesa, O. J., Cuartas, A., Barco, J., Mantilla, R. I., … Quevedo, D. (2007).: Linking long-

625 term water balances and statistical scaling to estimate river flows along the drainage network of Colombia. *Hydrologic Engineering*, 44, (2007).

Pyke, R.: TSLOPE3: Users Guide. *Taga Engineering Systems and Software,Lafeyette, CA.,* (1991).

Qiu, C., Xie, M., and Esaki, T.: *Application of GIS Technique in Three-Dimensional Slope Stability Analysis*. *Computational Mechanics, 3*(1), 703-712. https://doi.org/10.1007/978-3-540-75999-7_116, (2007).

630 Sanchez, O., and Aristizábal, E.: Spatial and temporal patterns and socioeconomic impact of landslides in Colombia. In EGU General Assembly 2018 (Ed.), *European Geosciences Union* (Vol. 20). Viena, Austria, (2018).

Schuster, R. L., and Highland, L. M.: Socioeconomic Impacts of Landslides in the Western Hemisphere. In *U.S. Geological Survey*, (2001).

635 Sepúlveda, S. A., and Petley, D. N.: Regional trends and controlling factors of fatal landslides in Latin America and the Caribbean. *Natural Hazards and Earth System Science*, *15*(8), 1821–1833. https://doi.org/10.5194/nhess-15-1821-2015, (2015).

Stark, T. D.: Three-Dimensional Slope Stability Methods in Geotechnical Practice. *51st Annual Geotechnical Engineering Conference*, 1–33. https://doi.org/10.1061/(ASCE)1090-0241(1998)124:11(1049), (2003).

640 Suarez, J.: Deslizamientos Y Estabilidad De Taludes En Zonas Tropicales. In *Instituto de Investigaciones sobre Erosión y Deslizamientos.* Bucaramanga, Colombia.: Universidad Industrial de Santander, (1998).

Süzen, M. L., and Doyuran, V.: Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Engineering Geology*, *71*(3–4), 303–321. https://doi.org/10.1016/S0013-7952(03)00143-1, (2004).

645 Taboada, A., Rivera, L. A., Fuenzalida, A., and Cisternas, A.: *Geodynamics of the northern Andes: Subductions and intracontinental deformation (Colombia).Tectonics, 19(5),* 787-813. *https://doi.org/10.1029/2000TC900004*, (2000).

Terlien, M. T. J.: The determination of statistical and deterministic hydrological landslide-triggering thresholds. *Environmental Geology*, *35*(2–3), 124–130. https://doi.org/10.1007/s002540050299, (1998).

650 Thiery, Y., Lacquement, F., & Marcot, N.: Landslides triggered in weathered crystalline rocks of moderate

latitudes: A case study in Mediterranean environment (The Maures Massif, France). *Engineering Geology*, *248*, 164–184, (2019).

Thiery, Y., Reninger, P.-A., Lacquement, F., Raingeard, A., Lombard, M., & Nachbaur, A.: Analysis of Slope Sensitivity to Landslides by a Transdisciplinary Approach in the Framework of Future Development: The Case of La Trinité in Martinique (French West Indies). *Geosciences*, *7*, 135. https://doi.org/10.3390/geosciences7040135, (2017).

Thomaz, J. E.: A General Method for Three Dimensional Slope Stability Analysis : Informational Report. In *Purdue University*. https://doi.org/10.5703/1288284314113, (1986).

Trenkamp, R., Kellogg, J. N., and Freymueller, J. T.: Wide plate margin deformation, southern Central America and northwestern South America, CASA GPS observations. *Journal of South American*, (2002).

Van Westen, C. J., Castellanos, E., and Kuriakose, S. L.: Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, *102*(3–4), 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010, (2008).

Van Westen, C. J., Soeters, R., and Sijmons, K.: Digital geomorphological landslide hazard mapping of the Alpago area, Italy. *ITC Journal*, *2*(1), 51–60. https://doi.org/10.1016/S0303-2434(00)85026-6, (2000).

Van Westen, C. J., and Terlien, M. T. J.: An Approach Towards Deterministic Landslide Hazard Analysis in GIS . a Case Study from Manizales (colombia). *Earth Surface Processes and Landforms*, *21*(June), 853–868. https://doi.org/10.1002/(SICI)1096-9837(199609)21, (1996).

Van Westen, C. J., van Asch, T. W. J., and Soeters, R.: Landslide hazard and risk zonation - Why is it still so difficult? *Bulletin of Engineering Geology and the Environment*, *65*(2), 167–184. https://doi.org/10.1007/s10064-005-0023-0, (2006).

Varnes, D. J.: Slope Movement Types and Processes. *Transportation Research Board Special Report*, (176), 11–33. https://doi.org/In Special report 176: Landslides: Analysis and Control, Transportation Research Board, Washington, D.C., (1978).

Velásquez, M. E., and Mejía, R.: *Procesos y depósitos asociados al aguacero de septiembre 21 de 1990 en el Área de San Carlos*. Medellín, (1991).

Wang, G., and Sassa, K.: Pore-pressure generation and movement of rainfall-induced landslides: Effects of grain size and fine-particle content. *Engineering Geology*, *69*(1–2), 109–125. https://doi.org/10.1016/S0013-7952(02)00268-5, (2003).

Xie, M., Esaki, T., Zhou, G., and Mitani, Y.: Geographic Information Systems-Based Three-Dimensional Critical Slope Stability Analysis and Landslide Hazard Assessment. *Journal of Geotechnical and Geoenvironmental Engineering*, *129*(12), 1109–1118. https://doi.org/10.1061/(ASCE)1090-0241(2003)129:12(1109), (2003).

Zizioli, D., Meisina, C., Valentino, R., & Montrasio, L.: Comparison between different approaches to modeling shallow landslide susceptibility: A case history in Oltrepo Pavese, Northern Italy. *Natural Hazards and Earth System Sciences*, *13*(3), 559–573. https://doi.org/10.5194/nhess-13-559-2013, (2013).