Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale

Vicente Medina (1), Marcel Hürlimann (1), Zizheng Guo (1, 2*), Antonio Lloret (1), Jean Vaunat (1)

1 Division of Geotechnical Engineering and Geosciences; Department of Civil and Environmental Engineering, UPC BarcelonaTECH; 08034 Barcelona, Spain
2 Department of Engineering Geology and Geotechnical Engineering, Faculty of Engineering, China University of Geosciences; 430074 Wuhan, China

* Corresponding author: Zizheng Guo (cuggzz@cug.edu.cn)

Abstract: Rainfall-induced landslides represent an important threat in mountainous areas. Therefore, a physically-based model called “Fast Shallow Landslide Assessment Model” (FSLAM) was developed to calculate large areas (> 100 km²) with a high-resolution topography in a very short computational time. FSLAM applies a simplified hydrological model and the infinite slope theory, while the two most sensitive soil properties regarding slope stability (cohesion and friction angle) can be stochastically included. The model has five necessary input raster files including information of soil properties, vegetation, elevation and rainfall. The principal output is the probability of failure (PoF) map. The Principality of Andorra was selected as case study, where FSLAM was successfully applied and validated using the existing landslide inventory. The PoF raster file of Andorra (including 19 million cells) was calculated in only 2 minutes. Therefore, an accurate calibration of the input parameters was easy, which strongly improved the final outcomes.

Keywords: Shallow slides; Rainfall; Susceptibility assessment; Physically-based model; ROC-analysis; Pyrenees
1. Introduction

Landslides represent one of the most significant geomorphological hazards in mountainous regions and characterise an important risk for people and infrastructures (Petley, 2012; Kirschbaum et al., 2015; Froude and Petley, 2018). Especially shallow landslides that are triggered by rainfall exceeding a certain threshold, can cause considerable losses (Caine, 1980; Guzzetti et al., 2007; Hungr et al., 2014). It is therefore a fundamental task for stakeholders and research institutions to perform an adequate landslide susceptibility, hazard and risk assessment (Guzzetti et al., 1999; Crozier, 2005; Fell et al., 2008; Segoni et al., 2020).

The susceptibility analysis identifies the prone areas where landslides can initiate and propagate (Guzzetti et al., 2005, 2006), and is the starting point of each hazard and risk assessment (Corominas et al., 2014). One of the results obtained from this task is the landslide susceptibility map, which is generally used for the management of territories or the land use planning (Fell et al., 2008; Chen et al., 2019). It is evident that approaches applied to landslide susceptibility modelling are highly essential for the reliability of the resulting map, and even a small increment in prediction accuracy may have a large impact.

Numerous models have been developed for the purpose of landslide susceptibility assessment. They can be divided into four main categories: expert-based models, physically-based models, statistical models, and machine learning models (Guzzetti et al., 1999; Huang et al., 2017; Sezer et al., 2017; Broeckx et al., 2018; Reichenbach et al., 2018). Among these models, the expert-based models mainly use the expert opinion to build the explanatory variables. However, their opinions are highly subjective and therefore the analysed process is difficult to be replicated by other users to different areas (Kirschbaum et al., 2016; Hearn and Hart, 2019). The statistical and machine learning models are normally considered as data-driven approaches, both of which focus on the analysis of the landslide influencing factors using past and present landslide datasets (Goetz et al., 2015; Huang and Zhao, 2018; Zêzere et al., 2017). Although in some cases the data-driven approaches have been considered more accurate than other approaches, they ignore the complex physical processes involved in landslide initiation. On the contrary, physically-based models can take geotechnical characteristics of landslides into account, and normally quantify the slope stability by combining the infinite slope stability approach and hydrological assumptions. Hence, one important advantage of physically-based models
is to calculate slope stabilities by using physical properties that control geomorphological processes, which better reflects the mechanism of landslide occurrences. The physically-based technique has been frequently adopted for the susceptibility assessment of shallow landslides and a list of available codes is given in Table 1. Most models use Mohr Coulomb theory to achieve geotechnical modelling, such as \textit{SHALSTAB} and \textit{TRIGRS}, while a few models use other theories, including \textit{STEP-TRAMM}, \textit{SCOOPS 3D} and \textit{R.ROTSTAB}. Regarding the hydrological modelling, some models consider both lateral and vertical flows of groundwater (e.g., \textit{GEOtop-FS} and \textit{HIRESSS}) but most models only involve one of them (e.g., \textit{SINMAP} model).

In this study, we developed a new physically-based model, because the quantitative analysis of future climate and vegetation changes was one of the requirements of the code and only physically-based models are able to include different rainfall scenarios and the effect of root strength due to vegetation. However, physically-based models have also limitations like the ones referring to the determination of the soil properties, which generally includes a large uncertainty (e.g. Tofani et al., 2017). In particular, the site-specific data generally is seldom available at regional scale (Carrara et al., 2008). Consequently, the regional-scale application of physical models for landslide susceptibility zonation is increasingly being an operational challenge. Another drawback of physically-based models is the high computational cost, especially when a comprehensive approach for the rainfall infiltration in unsaturated soil is used (e.g. the incorporation of the so-called Richards equation; (Iverson, 2000)).

In our new model, the geotechnical model is based on the infinite slope theory, whereas the hydrological model combines the lateral and vertical flows to calculate the water table. The input parameters consider various aspects that have impacts on slope stabilities, including digital elevation model (\textit{DEM}), soil properties, vegetation and rainfall. Finally, a stochastic approach is applied for soil properties given their uncertainty of spatial distribution.

The main objectives of this work include: (i) the presentation of a novel physically-based model, called Fast Shallow Landslide Assessment Model (\textit{FSLAM}), (ii) the evaluation of the different input parameters by a sensitivity analysis, and (iii) the application and evaluation of the proposed model at regional scale. Herein, the definition of regional scale stands for areas with an extension of more than 100 km$^2$, like our study area of Andorra, which covers about 470 km$^2$. 
Table 1 Summary of existing physically-based models and comparison with the FSLAM model presented in this study.

<table>
<thead>
<tr>
<th>Code</th>
<th>Geotechnical Model*</th>
<th>Lateral flow</th>
<th>Vertical flow</th>
<th>Input-output</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Deterministic approach</td>
<td>Stochastic approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHALSTAB</td>
<td>MC-inf</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>SINMAP</td>
<td>MC-inf</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>GEotop-FS</td>
<td>MC-inf</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRIGRS</td>
<td>MC-inf</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>SLIP</td>
<td>MC-inf</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>HIRESSS</td>
<td>MC-inf</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHIA_Landslide</td>
<td>MC-inf</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>STEP-TRAMM</td>
<td>FB-hex</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>SCOOPS 3D</td>
<td>Col-gr</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R.ROTSTAB</td>
<td>Col-gr</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSLAM</td>
<td>MC-inf</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>


2. Study area

2.1 General settings

The Principality of Andorra is situated in the middle of the Pyrenees (Figure 1) and covers a total area of approximately 470 km². The country has a typical high-mountain morphology with the capital of Andorra la Vella situated at about 1000 m asl. The morphology of Andorra has been strongly shaped by glacial and fluvial processes. The resulting main valleys include the two valleys Valira d’Orient and Valira del Nord that unify at the capital Andorra la Vella to form the Gran Valira valley (Figure 1). The peaks that surround these valleys have altitudes between 2500 and 3000 m asl.

From a geological point of view, Andorra is located in the Axial Pyrenees, where bedrock is composed of igneous and metamorphic Paleozoic rocks folded, intruded and metamorphosed during the Hercinian orogeny (Zwart, 1965). Surficial formations mostly cover the bedrock and include
principally glacial and peri-glacial deposits, but also torrential and colluvial materials are common. The climate of Andorra is affected by a combination of different factors including the vicinity of the Mediterranean Sea, the west winds from the Atlantic Ocean and the orographic effect. The average temperature in Andorra la Vella is about 10°C, while January is the coldest month with a mean minimum temperature of -1°C and July-August are the hottest months with a mean maximum temperature of 23°C. The average annual precipitation ranges from about 800 mm in the valley floor at Andorra la Vella up to 1300 mm at the highest peaks (Atlas Climatic Digital of Andorra, 2020). The largest rainfall amounts are registered during summer months. In the Pyrenees, shallow slope failures and debris flows are generally triggered by two types of precipitations: on one side, prolonged rainfall episodes during autumn, winter and spring (maybe associated with snowmelt); and, on the other side, high-intensity and short rainstorms in summer (Corominas et al., 2002; Abancó et al., 2016; Shu et al., 2019).

Figure 1: The Andorra study area and the landslide inventory. a) Location of Andorra in the Pyrenees. b) Slope angle map calculated by the 5m DEM. The red dots indicate the initiation points of shallow slides and debris flows included in the inventory.

2.2 The landslide inventory
As mentioned in the previous section, Andorra has been regularly affected by rainfall-induced
landslides, but unfortunately, detailed information is only available for the most recent events. However, landslide events are better registered than in most other areas of the Pyrenees, because elements at risk are more frequent in Andorra. That’s why Andorra represents an adequate study area for the evaluation of new FSLAM model.

In this study, an inventory of shallow landslides and debris flows was created including data from different sources. Principally, the inventory can be divided into two main types of data. First, there are the historic events, which mainly include shallow slides and some debris flows associated to the catastrophic MORLEs in autumn 1937 and 1982. These two MORLEs affected large areas of the Pyrenees, but no detailed inventory task was performed. In particular the 1937 MORLE lacks of data and only some specific works performed an interpretation of the 1948 aerial photographs to obtain the location of the slope failures in Andorra (e.g. Hürlimann et al., 2006). In contrast, the 1982 MORLE was investigated with more effort and two different landslide inventories were created in some parts of Andorra for scientific purpose using interpretation of aerial photographs and some field work (Landé, 1984; Baeza, 1994).

Second, there are the more recent events, mainly debris flows, which occurred during the last two or three decades. Detailed information on these events is available due to technical reports and scientific papers (e.g. CENMA, 2008; Portilla et al., 2010). In addition, an academic work was performed by the interpretation of recent Google Earth images, which provided the locations of additional slope failures in remote areas of Andorra (Gelonch Roca, 2014).

Finally, the inventory included 164 initiation points of shallow slides and debris flows (Figure 1). Some events were mapped by polygons, but all these polygons were transformed into points that indicate the position, where the slope failure initiated. The resulting overall density for the Andorra inventory is 0.35 landslide per square kilometre.

In a first step, the initiation points were analysed to determine the governing factors of landslide occurrence regarding geomorphology and land use and land cover (LULC) (Figure 2). The available digital elevation model (DEM) with a cell size of 5 m was used to calculate the elevation, slope angle and the contributing drainage area of each initiation point (Figure 2a, b and d). The analysis of the morphologic factors showed that the landslides started at elevations between 1000 and 2600 m asl with a large peak of density (initiation points per square-kilometre) between 2400 and 2600 m asl. The slope
angle at the initiation points has a maximum between 30 and 40°, which fits with other studies in the Pyrenees (Shu et al., 2019) or other mountain ranges (Ruff and Czurda, 2008; Bordoni et al., 2019). The flow accumulation or the contributing drainage area of the initiation point range from 100 m² to 1 km². Finally, the effect of LULC on landslide initiation (Figure 2c) shows that highest density of initiation points is located in scree and bare rock (maybe weathered bedrock or rocky channels with small amount of loose bedrock). The effect of root strength, which is proposed to be larger in forest, than in shrubs and grassland (Rickli and Graf, 2009; Reichenbach et al., 2014; Shu et al., 2019), was not observed in our inventory. This may be related to the explanation that 2012 map of land use and land cover was utilized to determine their influence on slope stability. However, the LULC certainly have strongly changed during the last hundred years and many inventory points are corresponding to the 1937 and 1982 MORLEs. A cross-check with the existing 1948 aerial photographs showed that especially in 1937, large areas of the present forest were covered by shrubs or grassland.

Figure 2: Analysis of the Andorra landslide inventory and effect of some governing factors of landslide initiation. a) elevation, b) slope angle, c) land use and land coverage (LULC), d) flow accumulation.
3. Methods and data

3.1 General aspects

The main assumption underlying the FSLAM model is to consider the rainfall as the landslide triggering mechanism. The rainfall effect appears in the soil mechanics as the pore pressure reducing the effective stresses that contribute to destabilize the soil.

Slope stability models mostly involve two sub-models: on one side a mechanical or geotechnical model, and on the other side a hydrological model focussing to the water flow. Regarding the geotechnical model, the infinite slope theory has been widely used when dealing with shallow slope failures computed at regional scale (see Table 1). Regarding the hydrological model, there is no agreement and no clear consensus. In general, two main approaches have been proposed that we herein call “lateral flow” and “vertical flow”. Iverson (2000) published a rigorous mathematical justification of the vertical flow preponderance, but there are evidences of slope failures in convergent areas related to lateral flows (e.g. Reneau and Dietrich 1987). The FSLAM-model incorporates both approaches and assume that lateral flow occurs in a mid to long term timescale and the vertical flow in a short term (Papa et al., 2013). In other words, the lateral flow includes the antecedent rainfall in order to define the steady-state condition of the water table, while the vertical flow represents the water infiltration during a specific rainfall event (see below for detailed explanations).

3.2 The geotechnical model

The selected mechanical model is the infinite slope theory (Lambe and Whitman, 1979; Pack et al., 1998), which computes the factor of safety ($FS$) by

$$FS = \frac{C}{\gamma_{m} z \cos \theta \sin \theta} + \left( 1 - \left( \frac{h}{z} \right) \left( \frac{\rho_{w}}{\rho_{s}} \right) \frac{(\tan \phi)}{(\tan \theta)} \right)$$

(1)

where $C$ is the cohesion, $\rho_{s}$ is the density of the saturated soil, $\rho_{w}$ is the density of water, $\theta$ is the terrain slope, $\phi$ the internal friction angle, $h$ is the water table depth and $z$ is the soil depth (Figure 3). Both $h$ and $z$ are measured in the vertical direction. In addition, the cohesion is calculated as follows:

$$C = C_s + C_r$$

(2)

where $C_s$ is the effective cohesion related to the soil matrix and $C_r$ is the apparent cohesion produced by the root strength.
It should be mentioned that the term $h/z$ is the average saturation degree and must be always lower or equal to one, hence, strictly speaking it should be $\min(h/z, 1)$. To simplify, this will not appear in the future equations.

![Geotechnical model applied in FSLAM](image)

Figure 3: Geotechnical model applied in FSLAM. The infinite slope model includes a soil layer of a given thickness with an impermeable bedrock at its lower limit.

### 3.3 Hydrological modelling

Hydrological modelling is the task in charge of setting the value of $h$ in Equation (1). As mentioned above, the FSLAM-model incorporates two different approaches: the first one focuses on the mid to long-term timescale and applies the lateral flow method to determine the increase of the water table ($h_a$) associated with the antecedent precipitation ($P_a$). The second approach is considering the effect of the short-term rainfall event and applies the vertical flow methods to calculate the increase the water table ($h_e$) related to the specific event rainfall ($P_e$). The final position of the water table ($h$) is calculated by

$$h = h_a + h_e$$ (3)

During the mid to long-term timescale, the antecedent rainfall ($P_a$) contributes to the groundwater through the recharge ($q_a$) and finally determines the value of $h_a$ at a kind of steady condition (Figure 4a). The recharge is a reduced percentage of the precipitation due to the runoff and evapotranspiration,
and can also be described as the effective water infiltration into the soil layer. In any case, the groundwater flow requires careful evaluation of the antecedent rainfall to quantify properly this recharge (Papa et al., 2011). The quantitative relation between the rainfall and the groundwater recharge is a difficult task in hydrology and is addressed in water balance models. However, the FSLAM-model does not include water balance algorithms; hence it requires available data to set the most adequate water recharge value. In summary, the antecedent rainfall ($P_a$) introduced in FSLAM as an input parameter represents a long term effective infiltration with the units of mm/d.

The calculation of $h_a$ is performed by the methodology introduced by Montgomery and Dietrich (1994), which was applied in many models (e.g. Dietrich and Montgomery, 1998; Pack et al., 1998):

$$h_a = \left(\frac{a}{b}\right) \frac{q_a}{K \sin \theta \cos \theta} \left(\frac{\rho_w}{\rho_s}\right)$$

where $a$ is the drainage area, $b$ is the cell size, $K$ is the horizontal hydraulic conductivity and $q_a$ is the effective infiltration rate due to antecedent rainfall.

On the opposite and regarding the event rainfall ($P_e$), the FSLAM-model calculates the increase of the water table ($h_e$) by the vertical flow approach neglecting the recharge contribution from upward areas.
A dimensionless analysis justifying this approach can be found in Iverson (2000), who compared the timescales for the vertical and horizontal water flows. Vertical flow triggering mechanism has been applied in several models (e.g. Šimunek et al., 1998; Baum et al., 2008). The vertical flow deals with the vadose zone, where the unsaturated soil flow is represented by the Richards equation having complex physics. Models that include the Richards equation require many input parameters and an extensive calibration. There is also a computational challenge in resolving these physics at regional scale. Simplified approximations have been developed to overcome these difficulties (Iverson, 2000; Baum et al., 2002; Savage et al., 2003, 2004). It should be mentioned that Iverson (2000) assumes no mass flow in the vertical direction, but just a pressure wave propagation. This assumption is based on the hypothesis that the soil is close to saturation before the triggering event. The complexity and the high time consumption of the calculations strongly reduce the applicability of such models at regional scale, especially when many scenarios must be included. However, the goal of the present study was to provide a model that fulfils the requirement of being fast at regional scale, even with a high resolution topography. Therefore, a simplified version of the vertical water flow is proposed. While traditionally a dynamic equation is used to capture the landslide triggering moment (Iverson, 2000; Baum et al., 2010; Papa et al., 2013), herein a static approach is selected instead. Therefore, the FSLAM-model does not track the destabilization process, but it assesses the final value of the factor of safety. The total infiltration related to the rainfall event is translated instantaneously to the increase of water table considering the soil porosity:

\[ h_e = \frac{q_e}{n} \]  

(5)

where \( q_e \) is the storm event infiltration, and \( n \) is the soil porosity. To get the value of \( q_e \), the event rainfall (\( P_e \)) must be converted into groundwater recharge. Both parameters, \( P_e \) and \( q_e \) have the units of millimetres. The method selected herein is the event oriented SCS-CN model (USDA, 1986a). This model has been developed to compute the surface runoff associated with a storm event, but to do it, it computes implicitly the infiltration. The success of the model is due to its simplicity, just requiring one parameter, named the curve number (CN). Finally, the event infiltration \( q_e \) is computed by the SCS-CN model as:

\[ q_e = P_e - \frac{(P_e - (5080/CN - 51))^2}{P_e + 4(5080/CN - 51)} \]  

(6)

Then, all the terms related to the two different rainfall approaches are merged and the final equation...
of the factor of safety is expressed as:

\[
FS = \frac{c}{\gamma_p z \cos \theta \sin \theta} + \left(1 - \left(\frac{\alpha}{h_{Kz \sin \theta \cos \theta} + \frac{g \gamma}{\tan \theta}}\right)\frac{\tan \varphi}{\tan \theta}\right)
\]  

Last but not least, it should be mentioned that there are fully 3D models with no previous assumption on the flow direction, able to capture both flows, vertical and lateral ones (e.g. Simoni et al., 2008). However, its application to regional scale is difficult due to its computational cost.

### 3.4 Model flowchart

The principal input data of FSLAM contain five raster files and two text files. The raster files include information on: i) the soil characteristics and the corresponding raster is called SOIL, ii) the Land Use and Land Cover (LULC), iii) the Digital Elevation Model (DEM), iv) the antecedent rainfall (RAIN_ANT) and, v) the event rainfall (RAIN_EVENT) (Figure 5). All the raster files were treated by the open source project on geographical information systems called QGIS (2019).

The parameters related to the soil as well as to the land use and land cover properties are defined in two text files called “soil.dat” and “hmtu.dat”. The file soil.dat provides the soil properties (soil cohesion, friction angle, density, hydraulic conductivity and soil thickness) and the hydrologic soil group (HSG) (USDA, 1986a, 2007a) for each soil class used in the SOIL raster file. The hmtu.dat is linked to the LULC raster and includes information on the root cohesion (Cr) for each land use class.

In addition, the hmtu.dat file merges the two raster files SOIL and LULC in order to create the HMTU raster, which divides the study area into different Hydrological-Mechanical Terrain Units (HMTU).

This unification of different information on soil and LULC is necessary, because the hydrological model requires the curve number (CN). In order to provide this unification, a combination matrix is provided in hmtu.dat. This file contains a table, similar to the one used in the SCS-CN model, where for every HSG and land use a CN value is provided (USDA, 1986a).

The DEM raster is used to compute the slope angle (\(\theta\)) and the flow accumulation (\(a\)) in each cell. Finally, two raster files are required for the precipitation input: antecedent rainfall (RAIN_ANT) and event rainfall (RAIN_EVENT). Both of them contain spatially distributed information. As described above, the RAIN_ANT raster provides the \(P_a\) - value in the study area, which is finally coinciding with the effective infiltration (\(q_a\)). The RAIN_EVENT contains the precipitation of the triggering rainfall (\(P_e\))
and is combined with the CN in order to obtain the event infiltration ($q_e$) (Figure 5).

Figure 5: Flow chart of the FS-calculation used in FSLAM.

### 3.5 Global patterns of the model

The equation of the FSLAM-model can finally be split into three different components:

$$FS = C_1 - q_a C_2 - q_e C_3$$  \hspace{1cm} (8)$$

where the definition of the three components is:

$$C_1 = \frac{c}{\rho \rho_s z \cos \theta \sin \theta} - \tan \frac{\phi}{\tan \theta}$$

$$C_2 = \left( \frac{\alpha}{\beta} \right) \frac{1}{K_h z \sin \theta \cos \theta} \left( \frac{\rho_w}{\rho_s} \right) \left( \frac{\tan \phi}{\tan \theta} \right)$$

$$C_3 = \frac{1}{n z} \left( \rho_w \right) \left( \frac{\tan \phi}{\tan \theta} \right)$$

(9)

The term $C_1$ corresponds to the stabilising term, equivalent to dry conditions. It includes the contribution from cohesion and internal friction angle. The term $C_2$ corresponds to the destabilization provoked by the antecedent rainfall, being its sign negative. The third term $C_3$ corresponds to the
destabilization triggered by the event rainfall, having negative sign as well. The final factor of stability of each cell is obtained from the superposition of the three contributions. This means that the result of the model is a superposition of three maps (Equation (8)): on one side the dry stability conditions and on the other side the contributions of the antecedent and the event rainfall, which are weighted using the corresponding recharge values $q_a$ and $q_e$. If the term $C_1$ is lower than 1, the cell is unconditionally unstable, and no rainfall is necessary to destabilize the cell. On the other side, unconditionally stable conditions are given, when term $C_1$ is large enough, because the terms $C_2$ and $C_3$ are constrained due to the maximum saturation by:

$$q_aC_2 + q_eC_3 \leq \left( \frac{\rho_w}{\rho_s} \right) \left( \frac{\tan \varphi}{\tan \theta} \right)$$  (10)

### 3.6 Stochastic approach for soil properties

The proposed model includes eight parameters related to the physical properties that depend on the soil type and vegetation cover: $C_s$, $\varphi$, $z$, $K$, $n$, $\rho_s$, $C_r$ and $CN$. In addition, there are two parameters related to the rainfall: $P_a$ and $P_e$. The water density and the gravity cannot be selected, because they are assumed to be constant. It is known that especially the determination of the soil properties at regional scale is a challenge. To overcome this difficulty stochastic models have been proposed. In these stochastic approaches, the input parameters have a statistical distribution, hence the factor of safety inherits the statistical properties and has its own probability distribution.

The result of applying stochastic parameters to the calculation of the factor of safety has different consequences. If the stochastic parameters are statistically independent, the expected value for the factor of safety is the same as the one obtained by a deterministic approach. Hence, when analysing the factor of safety, the stochastic approach is only justified, if some statistical dependence exists between the stochastic parameters. If there are no statistical dependence, the stochastic model still allows us to obtain the FS variance. When the stochastic approach is used, the term probability of failure ($PoF$) must be introduced, which correspond to the probability of having a $FS$-value lower than 1. The stochastic approach enables to include the uncertainty that is associated with the definition of the input parameters by different distributions of their values. Figure 6 shows this phenomenon referring to the cohesion of two types of soils. The first soil called “A” has a larger uncertainty and a
higher mean value than the second soil named “B”. If a deterministic approach would be applied using the mean values of the soils as input value, soil A would have a higher $FS$-value than soil B. But when the stochastic approach is applied and the Probability of Failure ($PoF$) is compared, soil A has a higher $PoF$-value than soil B.

Figure 6. Probability distribution of the factor of safety for two soil properties. The probability of Failure ($PoF$) using a stochastic approach and the factor of safety ($FS$) for a deterministic approach are indicated for the two soils. See text for additional explanations.

The incorporation of the stochastic approach into the model is not straightforward, since the equations for the $FS$-calculation combines the parameters by sums, quotients and products. Summing several terms of statistical distributions in one equation gives rise to a known statistical distribution. In contrast, the product or quotient of statistical distributions terms in one equation does not have a general solution. Hence, it is possible to consider stochastic distributions for the terms in the equation that are summed up. Following the work of Simoni et al. (2008), the cohesion and the internal friction angle belong to two separated terms that are added (see below in Section 4 for a detailed justification, why friction angle and cohesion was also selected in the present study). If the statistical distribution ($\nu$) of these parameters is Gaussian, the mean ($\mu_{FS}$) and the standard deviation ($\sigma_{FS}$) of the $FS$-distribution are Gaussian, too. Then, the resulting $FS$-distribution can be easily calculated by:

$$
\nu\left(\mu_{FS}, \sigma_{FS}^2\right) = \frac{\nu\left(\mu_C, \sigma_C^2\right)}{\gamma_s z \cos \theta \sin \theta} + \left(1 - \left(\frac{h}{z}\right)\left(\frac{\gamma_w}{\gamma_s}\right)\left(\frac{\nu\left(\mu_{\tan\theta}, \sigma_{\tan\theta}^2\right)}{\tan \theta}\right)\right)
$$

(11)
where the parameters $\mu_{FS}$, and $\sigma_{FS}^2$ are computed by:

\[ \mu_{FS} = \frac{\mu_{\text{tan}q}}{D} + \frac{\mu_C}{A} \]  
(12)

\[ \sigma_{FS}^2 = \frac{\sigma_{\text{tan}q}^2}{D^2} + \frac{\sigma_C^2}{A^2} \]  
(13)

\[ A = \frac{zy_y \sin 2\theta}{2} \]  
(14)

\[ D = \frac{\tan \theta}{1 - \left(\frac{1}{2}\right) \left(\frac{t_w}{y_y}\right)} \]  
(15)

It is important to stress that the option of the stochastic approach for cohesion and friction angle has an analytical solution. In contrast, the incorporation of three or more stochastic parameters implies moving to the Monte Carlo integration approach, which would strongly increase the computational time. A preliminary version of such a fully stochastic model showed that the number of runs per integral reach more than $10^{10}$.

### 4. Results

#### 4.1 Model verification and sensitivity analysis

**4.1.1 Model verification**

To perform the verification of the FSLAM code, a homogenous slope with the size of $100 \times 100$ m and a constant slope angle of $25^\circ$ was created. The cell size of the input raster was 10 m. The factor of safety ($FS$) of each cell in the slope was calculated changing the input parameters of our model. The ranges of the input values and their default values were selected using standard literature and expert criteria (USDA, 1986b, 2007b) (Table 2). In parallel, the $FS$-values of a selected cell were calculated manually, and compared with the ones obtained by FSLAM. The results of this model verification show that all $FS$-values calculated manually and by the FSLAM code are the same indicating that the code correctly simulates the stability conditions.
Table 2 Values of input-parameters selected during the model verification and the sensitivity analysis.

<table>
<thead>
<tr>
<th>Parameter type</th>
<th>Parameter</th>
<th>Unit</th>
<th>Values</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainfall</td>
<td>$P_a$</td>
<td>mm/d</td>
<td>0/0.01/0.1/0.2/1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>$P_e$</td>
<td>mm</td>
<td>0/50/100/150</td>
<td>100</td>
</tr>
<tr>
<td>Geometry</td>
<td>$\theta$</td>
<td>°</td>
<td>20/25/30/35/40</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>$a$</td>
<td>m²</td>
<td>100/800/10000</td>
<td>800</td>
</tr>
<tr>
<td>Soil</td>
<td>$\varphi$</td>
<td>°</td>
<td>20/25/30/35/40</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>$C_s$</td>
<td>kPa</td>
<td>0/3/25/50</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>log$K$</td>
<td>m/s</td>
<td>-10/-8/-6/-4</td>
<td>-6</td>
</tr>
<tr>
<td></td>
<td>$z$</td>
<td>m</td>
<td>0.5/1/2/5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$n$</td>
<td>-</td>
<td>0.2/0.25/0.3/0.35/0.4</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>$\rho_s$</td>
<td>kg/m³</td>
<td>1800/2000/2200</td>
<td>2000</td>
</tr>
<tr>
<td>Soil and LULC</td>
<td>$CN$</td>
<td>-</td>
<td>30/50/70/100</td>
<td>70</td>
</tr>
<tr>
<td>LULC</td>
<td>$C_r$</td>
<td>kPa</td>
<td>0/3/10/20/30</td>
<td>0</td>
</tr>
</tbody>
</table>

4.1.2 Sensitivity analysis of input parameters

The specific goal of this task includes: i) determining the importance of each input parameter on the resulting $FS$-value, and ii) selecting the parameter(s) that should be included into the model by a stochastic approach. We used the same value range of the input parameters as selected during the model verification, but added the slope angle and drainage area. The default values of each parameter are listed in Table 2.

The results of the sensitivity analysis are plotted by graphs, where the ranges of the obtained $FS$-values are indicated for each input parameter and the following three rainfall scenarios (Figure 7): i) variable value of $P_a$ and $P_e = 0$ mm, ii) variable $P_e$ and $P_a = 0$, and iii) variable values of both $P_a$ and $P_e$. The Factor of Safety that was calculated by using the default values of each parameter, is clearly visible in the different plots and has the following value for each rainfall scenario: i) $FS = 1.14$, ii) $FS = 1.32$, and iii) $FS = 0.96$.

As already detected in other studies (e.g. Alonso, 1976; El-Ramly et al., 2002), the results show that the most important input parameter is cohesion, herein represented by $C_s$ and $C_r$. The $FS$-ranges...
obtained by varying the cohesion are significantly larger than the ones calculated by changing the values of other parameters. When only the antecedent rainfall is included in the model (Figure 7a), the ranges of the resulting FS-values confirm that cohesion is the most important parameter. The other parameters have more or less the same influence, except the density, which has a very limited effect. It is important to understand that a FS-value of 0.75 for the parameters \( P_a, a, z \) and \( \log K \) stands for a totally saturated soil layer, since the default values of the slope angle is 25º, of friction angle 35º and of cohesion 0 kPa. It can also be observed that the hydraulic conductivity plays an important role for permeable soils and the soil layer turns completely saturated for K-values lower than \( 10^{-8} \) m/s applying the default values.

When only the event rainfall is included in the model (Figure 7b), the results show again the important effect of cohesion. Regarding the other parameters, the friction angle strongly effects the stability conditions, while the slope angle and the soil depth have also a rather large influence. In contrast, the other parameters like porosity, curve number and especially density have a minor importance. A notable outcome is the fact that the influence of the event rainfall is much less than the one of the antecedent rainfall considering the parameter ranges and default values selected in this study. Since the reduction of the safety factor is directly related to the rise of water table and the resulting ratio between water table and soil depth (\( h/z \); see Equation (5) and (6)), an event rainfall of 100 mm increases the water table only by 19 cm using the default values of \( CN = 70 \) and \( n = 0.35 \). In fact, this augment only produces a rise of the water table up to a third of the soil layer considering a soil thickness of 1 m. Thus, the reduction of the stability conditions is small.

When the last rainfall scenario is analysed (Figure 7c), which includes variable inputs of both rainfalls \( (P_a \) and \( P_e) \), the outcomes of the previous scenarios are confirmed. The cohesions is the most important parameter and the antecedent rainfall has higher influence than the event rainfall. In addition, the friction angle, the soil thickness and the hydraulic conductivity are other important soil parameters, while porosity, density or the curve number have a minor influence. Finally, the plot also shows that the two morphologic parameters (drainage area and slope angle) play an important in the stability conditions.
Figure 7: Results of the sensitivity analysis of all the parameters included in the FSLAM-code. The three plots present the ranges of factor of safety for three rainfall scenarios: a) changing the antecedent rainfall and neglecting the event rainfall ($P_e = 0$ mm), b) changing the event rainfall and neglecting the antecedent rainfall ($P_a = 0$ mm), and c) changing both the event rainfall and the antecedent rainfall. The parameter value is added next to each dot, while the default value is underlined.

A complementary analysis was performed to identify the most important parameters of the model. We assumed no previous knowledge of the equation governing the model. The same data that was used in the sensitivity analysis is included by means of a factorial regression, reaching an R-value of 0.996 (Montgomery, 2013). For every parameter in the model, a t-Student distribution is used to test the null hypothesis of the parameter having no effect on the results. Finally, a Pareto chart is constructed and a significance level of 0.99 is plotted (Figure 8). As already observed in the sensitivity analysis, the most important parameters are the cohesions, followed by the internal frictional and the soil depth.

In conclusion, the results of Figure 7 and Figure 8 show that efforts should be focussed on the correct definition of the cohesion. To deal with the uncertainty in these parameters we can take advantage of the stochastic approach. The internal friction angle was rated in the third place and finally included in
the stochastic approach, because of its simple incorporation in the simulation (see Section 2).

Figure 8: Pareto plot showing the standardised effect of the parameters included in the sensitivity analysis. The dashed line indicates the significance level of 0.99, which has a standardised effect of 2.7.

4.2 Susceptibility assessment

4.2.1 Modelling strategy

In this section, the general modelling strategy and the selection of the input data are described. The modelling procedure for the susceptibility assessment include two major parts. First, we calculated the susceptibility maps for the entire study area of Andorra representing the landslide susceptibility by the probability of failure in each cell. The resulting maps were compared with the existing inventory. During this task, the values of both the most important soil properties and the root cohesion were calibrated using the outcomes of the sensitivity analysis. A receiver operating characteristic (ROC) analysis was carried out to compare the performance of the different simulation results. Second, we analysed the influence of the two input rainfalls (antecedent and event rainfall) on the resulting susceptibility maps.

As stated in the description of the FSLAM-code and illustrated in Figure 5, five raster files and two text files containing soil and LULC properties are the most important input data. The pre- and post-processing of all the files were carried out in the open source project of QGIS (2019).
A Digital Elevation Model (DEM) with a 5 m cell size was selected for this study (Figure 9a) and downloaded from the SIGMA project website of the Andorran Government (SIGMA, 2020). The information of Land Use and Land Cover (LULC) was downloaded from the same website as a vector shapefile and then transformed into a raster. The most recent LULC map of 2012 was selected and the original LULC classes were reclassified into nine (Figure 9b). The information on soil properties was the most difficult to obtain, because no geotechnical map of Andorra is available. Finally, the geomorphologic map at 1:50000 scale was used in this study and downloaded from the SIGMA website as vector shapefile. Again, the initial geomorphologic units were reclassified and transformed into a raster file (Figure 9c).

Regarding the rainfall input data, two raster files are necessary: the antecedent and the event rainfall (Figure 5). Due to the lack of detailed meteorological differences over the study area, both raster files had constant values in the entire simulation domain. On one side, the antecedent rainfall was estimated by the water balance data corresponding to the period between 1974 and 2018 (Andorra Environmental Ministry, 2015). The average monthly recharge values for July, August and September are 1, 12 and 5 mm, respectively. The summer season was selected, because many recent shallow slides and debris flows have been triggered during this time of the year (e.g. Portilla et al., 2010). Finally, the value of 1 mm/d was used for the entire input raster file. This estimate of the antecedent effective infiltration would correspond to very humid month with a total recharge of 30 mm. On the other side, the event rainfall was assumed to be 100 mm. This value corresponds to a rainstorm with an approximate return period of 10 to 20 years considering a rainfall duration of 24 hours (Lopez Gumucio, 2015). For a comparison, the debris flows triggered in summer 2008 had a cumulative rainfall of about 65 mm in 3 hours (Portilla et al., 2010). In summary, hydro-meteorological time-series in Andorra are not very detailed and the two rainfall inputs are approximated, but they represent rather well the antecedent and triggering conditions of recent shallow landslides and debris flows in Andorra.
In the next step, the values of the two main input text files (soil.dat and hmtu.dat) were determined. This is a really complex and difficult task, since no detailed in-situ measurements are available that cover the entire study area and the heterogeneous conditions. Therefore, our modelling strategy applied an iterative approach and focussed on the parameters with larger influence on the factor of safety (see sensitivity analysis). Special attention was taken to the two cohesion values (\(C_s\) and \(C_r\)). The final values of the input files soil.dat and hmtu.dat are listed in Table 3 and Table 4. The uncertainty of the soil properties and root strength is a common drawback of physically-based models and was also considered in this study. Nevertheless, the selected values or value ranges of the input parameters
reflect rather well the real conditions (see the susceptibility maps in the following section) and also reasonably fit with the data published in literature.

Table 3 Parameter values used in the soil.dat input file for the Andorra study area.

<table>
<thead>
<tr>
<th>Soil class</th>
<th>$C_s$-min/max (kPa)</th>
<th>$\phi$-min/max (°)</th>
<th>$h$ (m)</th>
<th>$K$ (m/s)</th>
<th>$n$ (-)</th>
<th>$\rho$ (kg/m³)</th>
<th>HSG (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>colluvium</td>
<td>0/5</td>
<td>25/35</td>
<td>1.5</td>
<td>$1.6 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>B</td>
</tr>
<tr>
<td>lacustric</td>
<td>0/10</td>
<td>20/30</td>
<td>3</td>
<td>$3.85 \times 10^{-6}$</td>
<td>0.3</td>
<td>2000</td>
<td>C</td>
</tr>
<tr>
<td>fluvial floodplain</td>
<td>0/1</td>
<td>35/45</td>
<td>3</td>
<td>$6.25 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>A</td>
</tr>
<tr>
<td>anthropic</td>
<td>0/5</td>
<td>25/35</td>
<td>1.5</td>
<td>$1.6 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>B</td>
</tr>
<tr>
<td>scree</td>
<td>0/1</td>
<td>35/45</td>
<td>1.5</td>
<td>$6.25 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>A</td>
</tr>
<tr>
<td>glacial</td>
<td>0/5</td>
<td>25/35</td>
<td>1.5</td>
<td>$1.6 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>B</td>
</tr>
<tr>
<td>torrential fans</td>
<td>0/1</td>
<td>35/45</td>
<td>2</td>
<td>$6.25 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>A</td>
</tr>
<tr>
<td>landslides</td>
<td>0/5</td>
<td>25/30</td>
<td>3</td>
<td>$1.6 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>B</td>
</tr>
<tr>
<td>other areas</td>
<td>0/5</td>
<td>25/35</td>
<td>1.5</td>
<td>$1.6 \times 10^{-5}$</td>
<td>0.3</td>
<td>2000</td>
<td>B</td>
</tr>
</tbody>
</table>

Table 4 Parameter values used in the hmtu.dat input file for the Andorra study area.

<table>
<thead>
<tr>
<th>LULC</th>
<th>$C_r$-min/max (kPa)</th>
<th>CN-A (-)</th>
<th>CN-B (-)</th>
<th>CN-C (-)</th>
<th>CN-D (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>forest</td>
<td>0/5</td>
<td>36</td>
<td>60</td>
<td>73</td>
<td>79</td>
</tr>
<tr>
<td>shrubs</td>
<td>0/3</td>
<td>43</td>
<td>65</td>
<td>76</td>
<td>82</td>
</tr>
<tr>
<td>grassland</td>
<td>0/2</td>
<td>49</td>
<td>69</td>
<td>79</td>
<td>84</td>
</tr>
<tr>
<td>bare rock</td>
<td>0/0</td>
<td>77</td>
<td>86</td>
<td>91</td>
<td>94</td>
</tr>
<tr>
<td>scree</td>
<td>0/0</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>bare soil</td>
<td>0/0</td>
<td>77</td>
<td>86</td>
<td>91</td>
<td>94</td>
</tr>
<tr>
<td>farmland</td>
<td>0/1</td>
<td>56</td>
<td>68</td>
<td>80</td>
<td>84</td>
</tr>
<tr>
<td>water</td>
<td>999/999</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>urban area</td>
<td>0/1</td>
<td>90</td>
<td>92</td>
<td>96</td>
<td>98</td>
</tr>
</tbody>
</table>

### 4.2.2 Comparison between susceptibility maps and inventory

First, the probability of failure ($PoF$) was calculated for two scenarios: i) no rainfall input, and ii) the incorporation of both antecedent and event rainfall (Figure 10a and b). The effect of the rainfall is clearly visible by an important increase of areas with higher $PoF$-values. In addition, the cells with unconditionally stable and unconditionally unstable conditions are presented in Figure 10c. Most unconditionally stable areas are situated in the valley floor or in forested areas with low slope angle, while the unconditionally unstable areas are located in steep terrains with grassland or no vegetation.
Figure 10: Comparison between susceptibility maps and the Andorra landslide inventory showing the probability of failure (PoF) calculated by FSLAM under different scenarios. a) The PoF map for $P_\sigma=0$ mm/d and $P_e=0$ mm. b) The PoF map when $P_\sigma=1$ mm/d and $P_e=100$ mm. c) The map indicating the cells that are unconditionally stable and unconditionally unstable. The yellow rectangles in a) indicate the specific areas analysed in Figure 14.

Second, the 164 source points included in the landslide inventory were analysed in detail. The PoF-values of the 164 cells with initiation points were compared with 5000 randomly selected cells in the entire map using again the two rainfall scenarios described above. In order to correctly compare the
inventory with the randomly selected points, the number of points in the two datasets were normalized and finally presented by the percentage of the points versus the PoF (Figure 11). While the scenario with no rainfall input shows maximum percentages for low failure probability ($PoF \leq 0.2$) in both datasets, the maximum for the inventory points changes to high $PoF$-values ($PoF > 0.9$), when the rainfall scenario is applied.

![Figure 11: Evaluation of the FSLAM-model for the Andorra landslide inventory and analysis of the effect of rainfall. The inventory points are compared with randomly selected points under the “no rainfall” scenario (a) and the “rainfall” scenario (b).](image)

Third, a detailed ROC-analysis was performed using the rainfall scenario of $Pa = 1$ mm/d, $Pe = 100$ mm and the 164 inventory points. Special attention was paid to the effect of the three stochastic input parameters ($C_s$, $C_r$, and $\phi$). On one side, the standard plot comparing false-positive ($FP$) rate with true positive ($TP$) was drawn and the area under the curve ($AUC$) was calculated (Figure 12a). On the other side, the threshold of $PoF$ was compared with the accuracy calculated (Figure 12b).
The ROC-curves and the corresponding AUC-values show that the performance of the model improves, when friction angle and especially cohesion values are included by the stochastic approach (Figure 12a). Although the AUC-values are rather similar, these results justify the necessity of including these parameters by the stochastic option.

When comparing the PoF-threshold with the accuracy, the results show that the accuracy is rather similar for PoF-thresholds that range between 0.1 and 0.8. This means that the PoF-threshold used to determine the raster cells as stable or unstable could be selected in this range. There are accuracy maximums for each curve in this range, but fixing the adequate threshold is a decision that must be carefully analysed. Because the accuracy is an index that assigns the same weight to positive and negative conditions, high values mean success in classifying both conditions. In this study, the 0.5 probability threshold was finally selected. This value is equivalent to the median of the factor of safety and is also often applied in risk management to distinguish the cells as stable or unstable (e.g. Camilo et al., 2017; where maximum AUC was 0.71 and the best PoF - threshold was at 0.5).

Figure 12: Results of the ROC-analysis for the Andorra landslide inventory comparing different options of stochastic input parameters. a) False Positive (FP) - rate versus True Positive (TP) – rate of ROC curves. b) Threshold of Probability of Failure (PoF) versus accuracy curves.

Finally, the stochastic approach was compared with the deterministic one using the confusion matrix (Table 5). Different indexes were calculated including True Positive Rate (TPR), False Positive Rate
(FPR) and Accuracy (ACC). ACC and TPR are clearly better for the stochastic models, but deterministic model obtains similar results for FPR.

In conclusion, the performance analysis and the comparison of the different approaches show that the best approach is the one including cohesion and friction angle as stochastic parameters. However, it should be taken into account that none of the presented indexes to analyse the confusion matrix could be assumed as globally accepted. Depending on the final application of the models, it is better to focus on the false positives (e.g. for early warning systems) or on the false negatives (risk mapping).

Table 5 Confusion Matrix of the ROC-analysis for the Andorra landslide inventory comparing different approaches.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Stochastic parameters</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>TPR</th>
<th>FPR</th>
<th>ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastic</td>
<td>$C_s$, $C_r$, $\varphi$</td>
<td>100</td>
<td>3740</td>
<td>1260</td>
<td>64</td>
<td>0.610</td>
<td>0.252</td>
<td>0.679</td>
</tr>
<tr>
<td>Stochastic</td>
<td>$C_r$, $\varphi$</td>
<td>132</td>
<td>3021</td>
<td>1979</td>
<td>32</td>
<td>0.805</td>
<td>0.396</td>
<td>0.705</td>
</tr>
<tr>
<td>Stochastic</td>
<td>$C_r$, $\varphi$</td>
<td>139</td>
<td>2644</td>
<td>2356</td>
<td>25</td>
<td>0.848</td>
<td>0.471</td>
<td>0.688</td>
</tr>
<tr>
<td>Stochastic</td>
<td>$\varphi$</td>
<td>154</td>
<td>2045</td>
<td>2833</td>
<td>10</td>
<td>0.939</td>
<td>0.581</td>
<td>0.674</td>
</tr>
<tr>
<td>Stochastic</td>
<td>$C_s$, $C_r$</td>
<td>100</td>
<td>3740</td>
<td>1260</td>
<td>64</td>
<td>0.610</td>
<td>0.252</td>
<td>0.679</td>
</tr>
<tr>
<td>Deterministic</td>
<td>No</td>
<td>10</td>
<td>2955</td>
<td>2045</td>
<td>15</td>
<td>0.061</td>
<td>0.409</td>
<td>0.326</td>
</tr>
</tbody>
</table>

4.2.3 Effect of antecedent and event rainfall

The effect of the two rainfall inputs (antecedent and event rainfall) was analysed in detail using the same soil and vegetation parameter values as in the previous simulations (see Table 3 and Table 4). First, the percentage of the entire study area with critical stability conditions ($PoF > 0.5$) was calculated for different combinations of rainfall inputs (Figure 13). The results show that an increase of the antecedent rainfall stronger affects the area with critical slope stability than an increase of the event rainfall. This outcome can be demonstrated by the following examples: Assuming an event rainfall of 50 mm, the percentage of cells with a $PoF > 0.5$ raises from 10.8 % for $P_a = 0 \text{ mm/d}$ to 22.3 % for $P_a = 1 \text{ mm/d}$. In contrast and assuming an antecedent rainfall of 0 mm/d, the percentage of cells with critical stability conditions only increases from 8.2 % for $P_e = 0 \text{ mm}$ to 13.0 % for $P_e = 100 \text{ mm}$. 
Figure 13: Effect of the two rainfall inputs (antecedent ($P_a$) and event rainfall ($P_e$)) on the percentage of the study area in Andorra with critical stability conditions ($P_{OF} > 0.5$).

Second, the effect of the two rainfall inputs was analysed by comparing the resulting $P_{OF}$-maps of two selected zones of Andorra (Figure 14). The interpretation was easier in these smaller zones and provided new insights regarding the consequences of the two rainfall inputs. Finally, the following four rainfall scenarios were compared: i) no rainfall, ii) only $P_a$, iii) only $P_e$, and iv) $P_a$ and $P_e$. The comparison between scenario ii (second raw in Figure 14) with scenario iii (third raw in Figure 14) clearly visualize the difference of the two rainfall inputs. Since the effect of the antecedent rainfall is directly related to the drainage area of each cell, cells of larger $P_{OF}$-values are associated with morphologic features of concave plan curvatures. In such morphologic features, like channels, the slope stability is lower because of the higher water table. In contrast, the event rainfall is only related to the vertical infiltration and does not depend on the drainage area. Therefore, the position and extension of areas with critical stability conditions keeps constant between the initial map with no rainfall input (first raw in Figure 14). The only change due to the event rainfall is an increase of the $P_{OF}$-value in each individual cell of the map.
The maps of Figure 14 also show that there are inventory points with low PoF-values even applying both rainfall inputs. The experience of Andorra revealed two principal explanations: i) On one side, there is a problem in the location of the inventory points and the inventory points related to historic events were not located with enough precision. On the other side, the soil properties (especially the root and soil cohesion) could not have been correctly defined due to their large uncertainty.
Figure 14: Probability of failure (PoF) maps for two selected zones and four different rainfall scenarios: a) and e) $P_a = 0$ mm/d, $P_e = 0$ mm, b) and f) $P_a = 0$ mm/d, $P_e = 100$ mm, c) and g) $P_a = 1$ mm/d, $P_e = 0$ mm, d) and h) $P_a = 1$ mm/d, $P_e = 100$ mm. See Figure 10 for the exact location of the two selected zones.

5. Discussion

Complex physically-based models are often used to assess rainfall-induced landslide susceptibility at regional scale, but computational cost is normally high (e.g., Alvioli et al., 2016; Pourghasemi and Rahmati, 2018). Although modern computational hardware has been strongly improved, the computational time can still need several days for a relatively small area at high temporal and spatial resolution (Rossi et al., 2013). This drawback is mainly related to the calculation of the rainfall infiltration, the consequent definition of pore water pressure and the definition of the final water table in the soil layer. This problem was resolved in FSLAM by the application of two flow approaches, herein called “lateral flow” and “vertical flow”. The lateral flow approach includes the antecedent rainfall in order to determine a steady water table condition, which represent the initial state of soil saturation previous to the triggering rainfall. Then, the effect of the event rainfall is calculated by the vertical flow approach, which determines the increase of water table by a simple methodology without applying time-dependent transient calculations. This simplification may be a drawback, because no transitory conditions can be computed. However, the differences in the final water table level inside the soil layer are mostly rather small between the two modelling procedures (e.g. Healy and Cook,
and computational cost of our simplified model is much lower. A comparison between the computational cost of *FSLAM* and the well-known *TRIGRS*-code (Baum et al., 2008) in the Andorra study area shows that our new model needs more than 50 times less using the same input data and a standard personal computer. Indeed, the computational time for *FSLAM* was 2 minutes for one run of the entire study area including about 19 million cells. In contrast, *TRIGRS* needed 112 minutes selecting the Iverson (2000) approach for wet initial conditions and 210 minutes applying the Gardner (1958) approach for unsaturated initial conditions. All the calculations were performed by a computer with one 8-cores 1.8 GHz CPU and 8 GB of RAM.

Similar, rather long computational times were observed in other studies, when physically-based models were applied to large areas by ordinary personal computers even supercomputers (Rossi et al., 2013; Mergili et al., 2014). For instance, Rossi et al (2013) used SP6 supercomputer with 512 CPUs and HIRESSS model to simulate slope stability in southern Italy (approximately 3 million cells), and one hour was taken. A short computational time is crucial for civil protection purpose and the implementation of the model in an operational early warning system at regional scale (Piciullo et al., 2018; Guzzetti et al., 2020). The second major challenge in the regional susceptibility analysis deals with the uncertainty in the definition of the appropriate soil parameters (e.g., Cho, 2007; Tofani et al., 2017). Through the sensitivity analysis and the Pareto plot of the input parameters, we found that the slope stability is most sensitive to the changes of both soil and root cohesion. The friction angle has a minor influence on the stability conditions, but is more important than the other parameters. Other soil properties or even the rainfall have a smaller effect on the factor of safety considering the ranges of parameter values selected in this study. These outcomes coincide with the ones obtained in previous investigations (e.g. Alonso, 1976; El-Ramly et al., 2002).

The problem regarding the uncertainty of the input parameters was resolved in *FSLAM* by the option to incorporate the cohesion and friction angle as stochastic parameters. Our experience from the sensitivity analysis and simulations in Andorra revealed that special attention should be paid to the correct definition of the two cohesion values ($C_s$ and $C_r$). Because of their importance on the stability and the uncertainty of their correct value regarding regional scales, these two parameters must be introduced by a stochastic distribution. The increase of computational cost due to the incorporation of the cohesion and friction angle by a stochastic approach is acceptable, since it has an analytical solution
A test revealed that the computational time without a stochastic approach does not reduce the computational cost significantly. In contrast, a fully stochastic model, which means that all the eight input parameters related to the soil and land use properties would be included by a stochastic distribution, does not have a general analytical solution and needs more than $10^{10}$ runs per integral. Hence, it would strongly increase the computational time and would not be appropriate at regional scale. Moreover, the stochastic component from rainfall also has an impact on the probability of failure (Peres and Cancelliere, 2014, 2016). The reason why the uncertainty of the rainfall is not included in the present model is twofold. First, the sensitivity analysis revealed that the influence of rainfall seems to be less important than the one of other input parameters (Figure 8). Second, the modeling scenario and the available extreme events analysis did not include the confidence intervals computations or any other uncertainty analysis. However, it may be interesting to include a more extended analysis of the precipitation uncertainty in a future version of FSLAM incorporating the stochastic component to the hydrological part of the model.

The application to Andorra confirmed the outcomes of previous studies that the uncertainty of the terrain characteristics is an important aspect to be considered. Two main information are necessary to overcome this problem of the correct determination of the terrain input parameters: i) maps containing an accurate characterisation of both the soil and vegetation type, and ii) correct values of the mechanical soil parameters (especially soil and root cohesion). While land cover maps are generally available with a good resolution, detailed geotechnical maps of the soil properties are normally not available at regional scale and information from geological or other maps must be adapted. This uncertainty in the input data is one of the big open issue regarding the advances in regional landslide susceptibility assessment, in particular when dealing with physically-based models. Therefore, a fast numerical code is very helpful in order to rapidly perform an iterative calibration of the input parameters. Regarding the model accuracy, we evaluated its performance by comparing the susceptibility maps and the available landslide inventory without setting training and validation sets separately. Probably better results might be obtained using a training set from the inventory, although there is the drawback associated with the small number of landslides in the dataset. On the other hand, the advantage in not calibrating with a training set is the fact that the model can be applied in areas with no inventory, but obviously the uncertainty will be higher.
6. Conclusions

The susceptibility assessment of rainfall-induced shallow slope failures at regional scale (> 100 km²) is a complex task. In the present study, a novel model called FSLAM was proposed aiming to achieve fast assessment of landslide susceptibility on large areas.

The sensitivity analysis and Pareto plot revealed that the cohesion ($C_r$ and $C_s$) and the friction angle ($\phi$) are the most important input parameters. Thus, they are included into the model as stochastic input parameters. In contrast, the other five properties ($z$, $K$, $n$, $CN$ and $\rho_s$) have a minor importance on the stability conditions and were selected as deterministic input parameters. In addition, the sensitivity analysis showed that the antecedent rainfall has a higher influence than the event rainfall considering the parameter ranges and default values selected in this study.

The FSLAM model was applied and validated in Andorra, which covers an area of about 470 km² with total 164 recorded landslide initiation points. The calculations were carried out for about 19 million cells with a pixel size of 5x5 m². The simulations showed that the writing of the final raster file lasted far longer than the stability calculations. A comparison of computational time showed that the stochastic stability calculations lasted only 3 seconds, while the entire run including the writing of the raster file lasted 2 minutes.

The comparison between susceptibility maps and the Andorran landslide inventory provided satisfactory results considering the ROC-analysis. The model performance was examined quantitatively and an AUC-value of 0.77 was calculated. When different thresholds of PoF were compared, the changes of the accuracy value were limited. This result reveal that the model is robust and a PoF – threshold of 0.5 seems to be adequate, when stable and unstable conditions have to be separated during a risk assessment. This study shows that the physically-based model FSLAM is an effective tool to forecast the susceptibility of rainfall-induced landslides at regional scale in a very fast way. The presented model still has drawbacks and improvements may be necessary. However, the goal was to develop a fast and simple model, and therefore, simplifications like the one adopted in hydrological calculations related to the vertical water flow into the soil during the event rainfall, are justified.

Last but not least, FSLAM can be applied in the future to other mountainous regions where shallow landslides exist in order to assess their landslide susceptibility. It represents also a perfect model for
studies with many different scenarios like the ones dealing with climate change or land use and land
cover changes. The resulting maps can be easily included into landslide risk management and
mitigation strategies, and subsequently help decision-makers to choose preventive measures for land
use planning.

Supplementary material

The FSLAM-code is available at github platform (https://github.com/smucphy/fslam). It is open to the
community and binaries are also provided, including test data. The code is written in FORTRAN and
solutions for Visual Studio 2013 are provided as well.

Acknowledgements

The study was funded by the national research project called “Slope mass-wasting under climate
change” (smucphy.upc.edu) granted by the Spain Government (project reference number BIA 2015-
67500-R) and co-funded by AEI/FEDER, UE. Zizheng Guo acknowledges the financial support of
China Scholarship Council for his research at UPC BarcelonaTECH.

References

Abancó, C., Hürlimann, M., Moya, J., Berenguer, M., 2016. Critical rainfall conditions for the initiation of
https://doi.org/10.1016/j.jhydrol.2016.01.019

Geotechnique 26, 453–472. https://doi.org/10.1680/geot.1977.27.2.254

Alvioli, M., Spiga, D., Baum, R., 2016. Evaluation of the parallel performance of the TRIGRS v2.1 model for
rainfall-induced landslides. PeerJ. https://doi.org/10.7287/peerj.preprints.2206

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

Baeza, C., 1994. Evaluación de las condiciones de rotura y de la movilidad de los deslizamientos superficiales
mediante el uso de técnicas de análisis multivariante, Ph.D.-Thesis. Technical University of Catalonia,
Barcelona.


Service.

