Response to Reviewer #1

We thank Reviewer 1 for taking the time to read our paper. We think the Reviewers' comments can be addressed in a revised manuscript as follow. The comments made by the Reviewer are in black, and our responses are in blue. The italics highlights the additions made to the revised manuscript. Line numbers refer to the revised manuscript.

This paper studies the effect of informal housing on the slope stability using an improved mechanistic model CHASM (Combined Hydrology and Stability Model). This is an interesting topic, although this is rarely accounted for landslide hazard assessment. Hence, the reviewer does not suggest the current manuscript for publication. The manuscript needs a major revision.

Some comments for the revision.

(1) The study site map should be added to the C1 NHESSD Interactive comment Printer-friendly version Discussion paper reviewed manuscript.

Authors' reply: The methodology applied allows evaluation of the probability of failure of slopes for which there is scarce and/or uncertain data. Rather than referring to a specific site with measured geometry, urbanisation, soil and rainfall data, we stochastically generate tens-of-thousands of possible slope cross-sections that represent the population of slopes that might be observed in the case study region (using data from literature and previous fieldwork). Adding a map could therefore be misleading. However, to render the concept of stochastic generation of slopes clearer to the reader we modified the text of the introduction as follows (L.83):

"A sample of tens-of-thousands of rainfall events and slopes was stochastically generated from these distributions and simulated in CHASM."

Additionally, we replaced the term "site" with "case study" to avoid any confusion. Furthermore, we added, as suggested, more information on the type of climate (humid tropical) and on the type of soil and weathering grade usually found in the region in section 2.1, where other information about the geological setting are given.

We replaced the title of section 2.1 (L.137) with:

"Case study: Saint Lucia, Eastern Caribbean"

We added more explicit information on the climate and geological settings, in section 2.1 (L.138):

"Saint Lucia is an Eastern Caribbean island with a humid tropical climate. The main landslide trigger is rainfall, and shallow rotational landslides dominate on both steep and shallow slopes (Van Westen, 2016; Anderson and Holcombe, 2013). The geology is almost entirely comprised of volcanic bedrock and deep volcanic deposits. Due to the tropical climate, these volcanic parent materials are subjected to deep weathering, which decreases their strength and increases landslide susceptibility. The strata of a typical slope cross section comprise weathered residual soils overlying decomposed rock and volcanic bedrock. These three types of strata typically correspond respectively to the weathering Grade V-VI, Grade II-IV and Grade I-II, of the Hong Kong Geotechnical Engineering Office weathering grade classification (GEO, 1988). There is a high variability in terms of engineering soils, but they can broadly classify as fine grained soils such as silty clays, clayey silts and sandy clays (DeGraff, 1985). The combination of tropical climate, steep topography and volcanic geology render the region particularly susceptible to rainfall-triggered landslides. Furthermore, landslide risk is increased by informal housing which occupy steep slopes and employ unregulated engineering practices (WB/GFDRR 2012, p. 226-235)".

Reference: DeGraff, J. F.: Landslide hazard on St. Lucia, West Indies- Final Report: Washington D. C., Organization of American States, 1985.

(2) The thickness of the soil layer is crucial to the model calculation. How to consider the question in the improved modeling

Authors' reply: We agree with the reviewer on the importance of the thickness of the soil layers. Indeed, our results confirm this point via both the sensitivity analysis and CART. Soil thickness is considered in our modelling as an uncertain input factor and it is stochastically varied within a reasonable range, deduced from previous fieldwork. We are thus not completely sure about what the Reviewer means. No modifications were made in the main manuscript to address this comment.

(3) According to the reviewer's knowledge, the point water sources from informal housing may be closely related to preferential flow practically. Does your new model take into account the preferential flow?

Authors' reply: We did not include preferential flows in CHASM+. We agree with the Reviewer that leaking pipes and buried tanks can induce soil pipe erosion in response to increasing water inputs. We suggest that the wide range of hydraulic conductivity values sampled in the stochastic modelling approach could be assumed to account for the potential effects of increased hydraulic conductivity at the grid-cell resolution (i.e as a lumped or effective permeability that averages the K_{sat} for the preferential flow paths and the K_{sat} of the surrounding soil in each cell). Based on our benchmarking study of the new CHASM+ point-water source functionality (see Supplementary material) we believe that the current CHASM+ representation is sufficient to depict landslide initiation due to flow accumulation from the point water source. We added this discussion in the supplement (S1.1, L.39) as follow:

"Leaking pipes and buried tanks can induce soil pipe erosion in response to increasing water inputs. This could be simulated for example with a dual permeability model, but then it would be difficult to implement the pore pressure calculated into the slope stability model (Bogaard and Greco, 2016). Furthermore, the inclusion of preferential flows requires the definition of additional input factors which may be difficult in data-scarce contexts. So, given the spatial scale, the purpose of the analysis and the data available, the current CHASM+ representation can be considered sufficient to depict landslide initiation due to flow accumulation around the point water source".

Reference: "Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdiscip. Rev. Water, 3(3), 439–459, doi:10.1002/wat2.1126, 2016."

We also added in the same section (L.15) a comment about the fact that a dynamic change in the hydraulic properties due to the water leaked is anyway already taken into account by the model, given the way that CHASM represents hydrological processes.

"When water is added into the cell, the moisture content increases. The unsaturated hydraulic conductivity, which depends on the moisture content, also increases and is iteratively calculated with the Millington-Quirk formulation (Millington and Quirk, 1959). The maximum value is reached when soil is saturated (saturated hydraulic conductivity, fixed by the user).

Response to Reviewer #2

We thank Reviewer 2 for taking the time to read our paper and helpful suggestions to clarify the manuscript. The comments made by the Reviewer are in black, and our responses are in blue. The italics highlights the additions made to the reviewed manuscript. Line numbers refer to the revised manuscript.

General comment:

The manuscript (MS) deals with modeling possible impacts of informal urbanization on the hydrologic and geo-mechanical response of hillslopes, also with the aim at understanding which of the factors of such an urbanization process are the most detrimental for slope stability. The modeling is built as an extension of a previously released model (CHASM). I really enjoyed reading the MS, which is well written and structured. The supplementary material explains in detail the CHASM+ model and other aspects of the MS, and it is really an added value to the main text.

This paper studies the effect of informal housing on the slope stability using an improved mechanistic model CHASM (Combined Hydrology and Stability Model). This is an interesting topic, although this is rarely accounted for landslide hazard assessment. Hence, the reviewer does not suggest the current manuscript for publication. The manuscript needs a major revision.

From a general standpoint, the conclusion that slope cutting is the most detrimental among the other factors included in the modeling could be somewhat expected/or reached without the use of the massive modeling in the paper. However, I think that the main contribution given by this MS is that the model enables to QUANTIFY the response of the hillslope to the most important factors of informal urbanization and that it presents the application of some interesting statistical techniques to resume and communicate the main results of the modeling. Processes are represented in a somewhat simplified manner, but still the resulting model is quite complex and has several input parameters. Perhaps one could argue about some of the choices made in the model and the definition of the parameters' probability distributions (see also referee 1), but my opinion is that the authors have made all those choices in the most reasonable manner possible. For all the reasons above, I finally think this is a very good work, and my opinion is that that the MS can be accepted after minor revisions. In the following I provide just some suggestions to improve it.

Specific comments:

(1) L 83 The MS "promises" that somehow the modeling exercise will take into account climate change. I think this is quite weak in the analysis presented. The authors should discuss a little if climate change projections could be used to define future values of rainfall based on Representative concentration scenarios and simulations by Regional/Global climate models, and mention literature on the subject: e.g. https://doi.org/10.1016/j.jhydrol.2016.02.007, https://doi.org/10.1016/J.JHYDROL.2018.10.036

Authors' reply: We understand where the Reviewer is coming from with this comment as we have adopted a perhaps less common approach to account for climate change in our modelling. In the approach we adopt, sometimes referred to as 'bottom-up' (Groves and Lempert, 2007; Wilby and Dessai, 2010), we do not choose a single climate projection scenarios to define future values of rainfall and propagate them through the modelling chain ('top-down' approach) but rather we uniformly increase the severity of observed rainfall events and use CART to find those combinations of rainfall (and other uncertain input factors) that would produce unwanted outcomes (slope failure in our case). We therefore explore the feasible rainfall space widely, rather than focusing on the potentially more likely space covered (in terms of rainfall intensity and duration) of one or more scenarios (even though we will include this scenario space). In this way we can 1) quantify the effects of other uncertainties (such as uncertain soil properties) compared to climate uncertainty; 2) identify for which values of rainfall intensity and duration landslide hazard starts to significantly increase. These threshold values may then be compared to GCMs projections for a specific place, in order to assess the chances that they could be exceeded in the future. We recognise the point

is not clear in the manuscript and it requires a better explanation. We will include in the introduction the above discussion and references as follow (L.84):

"By this approach the possible effects of climate change were explored widely, instead of focusing on one (or a few) climate projection scenarios (such as those provided by downscaled generalised circulation models) propagated through the modelling chain (Groves and Lempert, 2007; Wilby and Dessai, 2010). This strategy can be extended to include the exploration of both feasible climate as well as feasible land use futures (Singh et al., 2014)".

Reference: Groves, D. G. and Lempert, R. J.: A new analytic method for find- ing policy-relevant scenarios, Global Environ. Chang., 17, 73– 85, doi:10.1016/j.gloenvcha.2006.11.006, 2007.

Wilby, R. L. and Dessai, S.: Robust adaptation to climate change, Weather, 65, 180–185, doi:10.1002/Wea.543, 2010.

Singh, R., Wagener, T. Crane, R. Mann, M. E. and Ning L.: A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA, Water Resour. Res., 50, 3409–3427, doi:10.1002/2013WR014988, 2014

We also added a figure (Fig.2 and related text) representing the feasibility space used for the stochastic generation of the rainfall events. The figure should clarify how future climate events (i.e. extreme combinations of rainfall intensity and duration) are considered in the analysis (L. 207)

"The model is forced with rainfall events which are specified in terms of their duration (in hours) and hourly intensity. The aim is to create both rainfall events that have been observed in the past, and rainfall events that might occur in the future (e.g. with higher intensity and duration than observed historically). To constrain the rainfall variability space, we use the intensity-duration-frequency (IDF) relationships derived from a Gumbel analysis of 40-years of daily rainfall data from weather stations across the island by Klohn-Crippen (1995) (Fig.2). From these IDFs we derive a range of rainfall intensity between 0 and 200 mm h-1, and a range of rainfall duration between 0 and 72 h. We then sample independently from the two uniform distributions, thus obtaining combinations of intensity and duration that might have been observed in the past (light grey area in Fig. 2) or not (dark grey area in Fig. 2)."

(2) LL 198-200 The water table height is varied between 0 and 90 % of the slope height. This seems a quite wide range. Perhaps the reasons for this choice could be better explained.

Authors' reply: The wide range aims to represent the variability across the ensemble of slopes that can be found in our study region. We then use CART to define thresholds of water table height above which slope failure is more likely to occur. These threshold values can be then compared with levels of water table height of a particular slope and in a particular moment to assess its landslide probability. We better specify the stochastic generation of the water table in section 2.4, L. 204:

"This water table height is varied between 0% and 90% of the slope height (H in Fig. 3), to account for its variability across the region and for the variability of the initial soil moisture conditions due to antecedent rainfall events."

(3) L 234 Perhaps a reference explaining the Latin Hypercube sampling C2 NHESSD Interactive comment Printer-friendly version Discussion paper technique can be useful for readers.

Authors' reply: L.242 *"(McKay et al., 1979)"*

Reference: "McKay, M. D., Beckman, R. J. and Conover, W. J.: Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21(2), 239–245, doi:10.1080/00401706.1979.10489755, 1979".

(4) Section 4.2 and LL 263-275 of the supplement: The objectives of the multi-optimization are quite unusual. Perhaps in this case, an optimization based on ROC (receiver operating characteristics) analysis (i.e.: True and false positives/negatives) could have been employed and would have been more meaningful. At least, literature in the subject should be mentioned: e.g. https://doi.org/10.1007/s10346-020-01420-8, https://doi.org/10.1029/2012JF002367, https://doi.org/10.5194/hess-18-4913-2014

Authors' reply: Optimization based on ROC analysis could have been an option, though we think our approach is also suitable given that our aim was to essentially identify the two parameters of the minimum rainfall threshold line. We already cite a study that employs ROC analysis and a review where it is mentioned (Staley et al. 2013 and Segoni et al. 2018). We make a more explicit reference in L. 239 of the supplement:

"An alternative to this approach could be to use a (single-objective) optimization based on ROC (receiver operating characteristics), where false positives and negatives (represented in this case by the simulated landslides below the threshold and simulated stable slopes above the threshold) are minimised (Gariano et al., 2015; Staley et al., 2013)".

Reference: Gariano, S. L., Brunetti, M. T., Iovine, G., Melillo, M., Peruccacci, S., Terranova, O., Vennari, C. and Guzzetti, F.: Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy, Geomorphology, 228, 653–665, doi:10.1016/j.geomorph.2014.10.019, 2015.

(5): Fig S1 (supplement): Panel (a) is repeated in panel (b), so perhaps it could be removed. Possibly add to the plot the rainfall time series (cumulated sum).

Authors' reply: We have modified Figure S1 as suggested

(6): Section S1. Perhaps the case of houses WITH gutters should be explained.

Authors' reply: We explained also the case WITH gutters in L37 of section S1.1:

"If gutters are present the rainwater intercepted by the roof is deleted, consequently decreasing the rainfall rate infiltrating into the slope".

Technical corrections L60 (supplement) and L137

Authors' reply: We have addressed these typographic and grammatical errors and check the whole manuscript and supplement.

Track changes version of the main manuscript: please note that there might be some minor differences between the .pdf uploaded and the text below (some of the changes have been accepted by mistake but they do not regard the reviewers' comments).

Including informal housing in slope stability analysis – an application to a data-scarce location in the humid tropics

Elisa Bozzolan[1,2], Elizabeth Holcombe[1,2], Francesca Pianosi[1,2], Thorsten Wagener[1,2]

[1]Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK
[2]Cabot Institute, University of Bristol, Bristol, BS8 1TR, UK

Correspondence to: Elisa Bozzolan (elisa.bozzolan@bristol.ac.uk)

Abstract

Empirical evidence from the humid tropics shows that informal housing can increase the occurrence of rainfall-triggered landslides. However, informal housing is rarely accounted for in landslide hazard assessments at community or larger scales. We include informal housing influences (vegetation removal, slope cutting, house loading and point water sources) in a slope stability analysis. We extend the mechanistic model CHASM (Combined Hydrology and Stability Model) to include leaking pipes, septic tanks, and roof gutters. We ~~test~~ apply CHASM+ in a region of the humid tropics, using a stochastic framework to account for uncertainties related to model parameters and drivers (includ~~ing.~~ climate change). We find slope cutting to be the most detrimental construction activity for slope stability and we quantify its influence and those of other destabilising factors. When informal housing is present, more failures (+85%) are observed in slopes ~~with~~ that would otherwise have had low landslide susceptibility, and for high intensity, short duration precipitations. As a result, the rainfall threshold for triggering landslides is lower when compared to non-urbanised slopes, and comparable to those found empirically for similar urbanised regions. Finally, low cost-effective 'low regrets' mitigation actions are suggested to tackle the main landslide drivers identified in the study area. The proposed methodology and rainfall threshold calculation are suitable for data scarce contexts, i.e. when not ~~much~~ limited field measurements or landslide inventories are available.

1 Introduction

Global and regional landslide records reveal an increase in rainfall- and human-triggered landslides during the last century, mainly in economically developing countries with rapid population growth and urbanisation (Kirschbaum, et al., 2015; Froude and Petley, 2018). This increase might be partly due to continuing improvements in landslide recording, but it also indicates the growing impact of climate and urban pressure on landslide occurrence (Larsen, 2008). Understanding the mutual interactions between the natural and urban environment becomes particularly relevant in the humid tropics where high intensity and duration rainfall events are the main landslide triggers and urban expansion is poorly regulated (Lumb, 1975; UN-Habitat, 2015). The natural landslide susceptibility of these regions coupled with the lack of urban planning and regulations can increase risk, not only in terms of vulnerability and exposure but also in terms of hazard.

Evidence from low income urban settlements in the humid tropics reveals a link between poorly regulated urban construction activities and increasing landslide hazard. Potential anthropogenic landslide drivers include slope cutting and filling for house and road construction (Sidle and Ziegler, 2012; Smyth and Royle, 2000), slope degradation with clearance of forested areas (Gerrard and Gardner, 2006; Vanacker et al., 2003), and inadequate drainage networks, unplanned redirection of storm runoff and poorly maintained septic systems (Diaz, 1992; Anderson et al. 2008). In this paper, we use the term 'informal housing' to

refer to the combination of these urban modifications which influence slope stability by altering its geometry, hydrology and material strength (Figure 1).

However, informal housing is usually neglected or not quantified in landslide hazard assessment at community and larger scales. There are two main reasons for this: lack of reporting and the highly localised scale and heterogeneous nature of human landslide drivers. A landslide is defined as triggered by human activities when there is a direct (and easily recognisable) connection with the failure process (e.g. during mining activities). Landslides of this type are small and often not recorded (Petley, 2012).When considering rainfall triggered-landslides, human landslide drivers are often either not considered or not distinguished from the natural drivers (SafeLand, 2011). Urban construction activities are localised and even if they contribute to land instability, they remain difficult to observe either in situ (e.g. leaking pipes) or via satellite images. For these reasons, there are numerous site specific analysis that investigate the influence of urban construction activities for individual slopes with known soil and rainfall trigger characteristics (e.g. Preuth et al., 2010; Zhang et al., 2012), but there are few studies that explore the influence of informal housing more widely, for different combinations of human landslide drivers, soils, slope geometry and rainfall triggers. This limits the transferability of the findings from slope to larger scales where less detailed data is available.

Empirical-statistical and heuristic methods have been used in regional studies to link informal housing to the spatial and temporal occurrence of landslides. For example, precipitation and landslide records have been analysed in relation to lithology and land use change (Alewell and Meusburger, 2008; Gerrard and Gardner, 2006), or in relation to soil type, and type of settlement (Smyth and Royle, 2000). Here, most of the recorded landslides were found to be associated with poorly regulated construction techniques, water management and land degradation. Rainfall thresholds for triggering landslides were observed to depend on the proportion of impervious surfaces (Diaz, 1992). However, these analyses did not enable the differentiation of the relative role of natural and human landslide drivers precluding the translation of the results into actions at slope/engineering scale (Anderson et al., 2013; Maes et al., 2017).

Mechanistic slope hydrology and stability models can be used to represent the landslide drivers for historical, current and potential future climate conditions (e.g. Ciabatta et al., 2016; Almeida et al., 2017). If these models included the effect of informal housing, the analysis of different combinations of slope, urban and climate properties could lead to assess the relative role of natural and urban properties on triggering landslides and to identify the conditions at which urban construction activities become most detrimental. This could be a useful information for engineers to prioritise slopes that are currently at risk, to identify those at higher risk to be impacted in the future, and to deduce appropriate hazard mitigation or preparedness actions. The inclusion of informal housing in slope stability analysis could also lead to considerations about the reliability of rainfall thresholds for triggering landslides within highly urbanised communities, since they might be underestimating the level of the hazard (Mendes et al., 2018).

However, the use of data intensive mechanistic models can be challenging in data scarce locations, such as in low income urban settlements. The more complex the model, the more data required to set its parameters and model forcing, and the more uncertainties might be introduced into the analysis. Sources of uncertainties can relate to slope and soil properties, urban features as well as to a limited understanding of physical processes or future scenarios (epistemic uncertainties) (see Beven et al., 2018a, for a review of this issue). Many researchers have assessed the impact of uncertainties related to slope properties (e.g. Cho, 2007) and future climate (e.g. Ciabatta et al., 2016) on slope stability at different scales. However, to the best of our knowledge, there are no analyses that consider both sources of uncertainties when modelling informal housing in landslide hazard assessment. Urban construction activities are either considered separately (e.g. slope cutting or pipes leaking) (e.g. El-Ramly et al., 2006) or the slope properties are varied using discrete conservative values under fixed rainstorm conditions (Anderson et al., 2008; Holcombe et al., 2016). This separation might overlook significant changes of the slope's behaviour for combinations of urban constructions activities and/or slope/soil/rainfall properties that have not been considered but are still likely to occur.

Almeida et al.~~,~~ (2017) ~~has~~ demonstrated how mechanistic landslide models can consider ~~both~~ uncertainties due both to poorly defined slope properties and to potential future climate changes. The mechanistic model CHASM (Combined Hydrology and Stability Model) was used in a Monte Carlo framework and applied in Saint Lucia, in the Eastern Caribbean, where data support is limited but landslide hazard is particularly high. The ~~probability distributions of~~ uncertainties in slope and soil properties were characterised through probability distributions ~~are~~ extrapolated from available data and literature, while the rainfall properties were varied ~~widely and~~ uniformly across wide ranges, also considering rainfall intensity-duration combinations that~~which~~ were not observed in the past but that might ~~be~~ occur~~observed~~ in the future. A sample of tens-of-thousands of rainfall events and slopes ~~were~~ was stochastically generated from these distributions and simulated in CHASM. By this approach the ~~Potential~~possible effects of climate change were ~~therefore considered~~explored widely, instead of focusing on one (or a few ~~more~~) ~~not by choosing a single~~ climate projection scenarios (such as those provided by downscaled generalised circulation models ~~GCMs~~) ~~to~~ propagated through the modelling chain ~~('top-down' approach), but rather by uniformly increasing the severity of the rainfall event, i.e. increasing the rainfall intensity and duration~~ ~~('bottom-up' approach~~ (Groves and Lempert, 2007; Wilby and Dessai, 2010). This strategy can be extended to include the exploration of both feasible climate as well as feasible land use futures (Singh et al., 2014). ~~). The rainfall events and a population of tens-of-thousands slopes that might be observed in the region in the present or in the future were stochastically generated and simulated in CHASM.~~ Statistical and data mining algorithms were then used by Almeida et al. (2017) to quantify the relative role of the input factors (and thus their uncertainties) on the stability of the simulated slopes as well as to identify critical thresholds in slope properties and rainfall drivers likely to lead to slope failure. In this study we extend the work of Almeida et al. (2017) by including informal housing into such a slope stability analysis. We consider the same location of the humid tropics and the same core model, CHASM, but with new functions to represent the mechanistic influences of informal housing. CHASM is a two-dimensional model which has a relative low data requirement for a mechanistic model even with the inclusion of the new informal housing functions. In addition to the original ability to represent the mechanical and hydrological effects of vegetation and the effects of slope cutting and loading, we have added the effects of point water sources resulting from leaking septic tanks, water supply pipes, and houses without roof gutters. By varying both the natural and urban factors, we aim to identify under which slope and climate conditions landslide hazard is significantly increased by the presence of informal housing and how this information can be used for deducing landslide mitigation measures Thus, for our humid tropical case study scenario we aim to address the following questions:

1. How can we identify which informal urban housing characteristics are most detrimental to slope stability?
2. How is the rainfall threshold for triggering landslides modified when informal housing is considered?
3. Which landslide mitigation strategies and practices can be deduced from the analysis for current and potential future scenarios of urbanisation and rainfall?

The proposed methodology is suitable for data scarce contexts, i.e. when not much field measurements or landslide inventories are available. If applied in countries with similar natural/climate/urban characteristics (so with similar input space variability) we might expect similar slope stability responses and thresholds. Conversely, a change in (part of) the input data (or their probability distributions) to reflect a different urban landslide context could potentially produce quite different outputs (Wagener and Pianosi 2019).

Figure 1: Examples of informal housing affecting land stability. The image (a) and (b) show examples of unsupported cut slopes, respectively in Saint Lucia (Caribbean) and Dumsi Pakha (Kalimpong, India). The image (c) shows the effect of lack of water management in an informal community of Saint Lucia (Caribbean). From the blog AGU landslides: https://blogs.agu.org/landslideblog/2016/03/14/managing-urban-landslides-1/ ~~From the blog Save the Hills (http://savethehills.blogspot.com/)~~ and the community based project Mossaic (Anderson and Holcombe, 2013)

2. Method

We want to analyse the relative role of informal housing on slope stability under different natural and climate conditions. The methodology we introduce here entails the following steps:

- Choose a model that represents the main instability mechanisms of the case study area. We are interested in representing the rainfall-triggered landslides and the informal housing of Saint Lucia (Caribbean). We therefore use the mechanistic model CHASM which represents both the hydrology-stability routing, but also vegetation, slope cutting, and various forms of water management.
- Define the inputs factors necessary to run the model and their variability space. In our case study, the input factors are the parameters defining the slope soil, geometry, urban characteristics, as well as rainfall forcing data. Each input factor is assumed to be a random variable and its range of variability is determined by a probability distribution. The probability distributions can be defined based on the physical meaning of the input factors, available data and/or existing literature. We use information gathered both from fieldwork in Saint Lucia and also from literature.
- Create synthetic combinations of input factors by stochastically sampling from their probability distributions and run CHASM to generate an equivalent number of model outputs. We select the minimum Factor of Safety (FoS) and the slip surface where the minimum FoS is calculated as summary output variables to analyse. We repeat the stochastic sampling with and without including the urban properties among the input factors, in order to facilitate considerations about the role of informal housing on land stability.
- Identify the input factors that most influence slope stability using global sensitivity analysis (Wagener and Pianosi, 2019). In particular, we use a regional sensitivity analysis approach (RSA, Hornberger and Spear, 1981) to identify which input factors are most influential in leading to slope failure.
- Identify parameters' thresholds beyond which the slopes become unstable. The threshold of an input factor over/below which failure is predicted might depend on the value of the other input factors (e.g. slopes with higher slope angles require higher soil strength to maintain stability). Machine learning is a set of methods that computers use to understand trends from data, also considering their mutual interactions. We use CART (Classification and Regression Trees) to develop a set of decision rules that predict for which combination of soil, geometry, urbanisation and rainfall input values a particular slope is more likely to fail.

In the following paragraphs we are going to describe in detail the tools and the data used to implement our analysis on the island of Saint Lucia.

2.1 The ~~study site~~case study: Saint Lucia, Eastern Caribbean

Saint Lucia is an Eastern Caribbean island with a humid tropical climate. The main landslide trigger is rainfall, and shallow rotational landslides dominate on both steep and shallow slopes (Van Westen, 2016; Anderson and Holcombe, 2013). The ~~region is made up~~geology is almost entirely comprised of volcanic bedrock and ~~rocks, with~~ deep volcanic deposits. Due to the tropical climate, these volcanic ~~bedrock~~parent materials ~~is~~are subjected to deep weathering, which decreases their ~~soil~~ strength and increases landslide susceptibility. The strata of a typical slope cross section ~~is composed of~~comprise ~~soil strata of~~ weathered residual soils overlying decomposed rock and volcanic bedrock. These three ~~layers~~types of strata typically correspond respectively to the weathering Grade V-VI, Grade II-IV and Grade I-II, of the Hong Kong Geotechnical Engineering Office ~~(1988)~~ weathering grade classification (GEO, 1988). There is a high variability in terms of engineering soils, but they can broadly classify as fine grained soils such as silty clays, clayey silts and sandy clays (DeGraff, 1985). The combination of tropical climate, steep topography and volcanic geology render the region particularly susceptible to rainfall-triggered landslides. Furthermore, landslide risk is increased by informal housing which occupy steep slopes and employ unregulated engineering practices (WB/GFDRR 2012, p. 226-235). Various sources of information on the slope, soil, rainfall and urban properties of this region are available from previous studies by government engineers and planners, the local water company and consultants (e.g. CHARIM, 2015; Mott MacDonald, 2013; Klohn-Crippen, 1995), and from community based projects for the improvement of slope stability with surface water drainage works (Anderson and Holcombe, 2013). In this project, estimates of soil strength properties are based on direct shear tests of local soils (Anderson and Kemp, 1985; DIWI, 2002; Holcombe, 2006), and secondary data sources on similar volcanic tropical residual soils such as those in Hong Kong (Anderson, 1982; Anderson and Howes, 1985). Information about soil type, soil depth, type of house construction, cut slope angles and the management of surface runoff and waste water on slopes was based on community-based mapping and elicitation of local expert knowledge undertaken by Anderson and Holcombe (2013) who co-developed these datasets with residents, government, and local experts.

2.2 CHASM: a mechanistic model for rainfall-triggered landslides

CHASM (Combined Hydrology and Stability Model) is a 2-D mechanistic model which analyses dynamic slope hydrology and its effect on slope stability over time. A full description of the model can be found in Anderson and Lloyd (1991) and Wilkinson et al. (2002a,b). Here we briefly describe its hydrology and stability components, whereas the representation of the urban properties is detailed in Paragraph 2.3. In CHASM the slope cross section is represented with a regular mesh of columns and cells. Hydrological and geotechnical parameters are specified per cell, while the initial hydrological conditions define the position of the water table, and the matric suction of the top cell of each column. The dynamic forcing for CHASM is rainfall specified in terms of intensity and duration. For each computational time step (usually 10–60s), a forward explicit finite difference method is used to solve the Richard's (1-D, vertical flow) and Darcy's (2-D flow) equations which regulate respectively the unsaturated and saturated groundwater flow. At the end of each simulation hour the resulting soil pore water pressures (positive and negative) are used as input for the slope stability analysis which implements a Bishop's simplified circular limit equilibrium method (Bishop, 1955) and uses the coordinates of the slope surface. An automated search algorithm identifies the location of the slip surface with the minimum factor of safety, FoS, which is given as output at the end of each simulated hour. In a validation exercise in Hong Kong, CHASM shown an accuracy of 72.5% (Anderson, 1990) which is comparable to the performances of other models used for the stability analysis (e.g. Formetta et al., 2014). CHASM has been employed in Malaysia, Indonesia, Eastern Caribbean, and New Zealand, to propose landslide mitigation measures, as well as to identify land instability drivers along roads and in urban and rural areas (Brooks et al., 2004; Lloyd et al., 2001). Almeida et al. (2017) used CHASM stochastically in a Monte Carlo framework to account for uncertainties in both slope properties and future climate scenarios.

2.3 A new functionality in CHASM: urban point water sources

The new CHASM+ can now not only represent slope cutting, additional (house and tank) load, and vegetation removal, but also the presence or absence of roof gutters on houses and localised water leakages from buried septic tanks and superficial water supply pipe networks. Slope cuttings are represented by a corresponding change in slope geometry; additional loads are simulated by appropriately increasing the unit weight of the soil underneath the loading object (i.e. house and tanks); vegetation, which is removed during the urbanisation process, is represented through rainfall interception, evapotranspiration, root water uptake, vegetation surcharge, and increased permeability and soil cohesion due to the root network (see Wilkinson et al., 2002b). Pipes above ground and buried tanks can be added to the slope, with specified dimensions and leakage rates. Pipe leakage is accounted for as additional surface water which infiltrates into the slope according to the infiltration capacity of the soil. If water exceeds the infiltration capacity of the soil, it is stored as ponding water. If the ponding water exceeds the maximum water detention capacity (set at 10 mm), the water excess is removed (no runoff considered). Leakage from tanks is added to the water moisture content in the soil cells underneath the tank. Where houses are present, rainfall is intercepted by the roof. If roof gutters are not included, the intercepted rainwater is discharged onto the slope cells adjacent to the house, in accordance to the roof type (double or single pitch). More details on the new functionality and its benchmarking against another model are given in the supplementary document that accompanies this paper (S1.1~~.~~ and S1.2).

2.4 Definition of the input factors and their probability distributions

We use thirty input factors to characterise our case study area in CHASM+. These factors fall into the following categories: slope profile geometry, soil geotechnical and hydrological properties, urban characteristics, initial hydrological conditions and rainfall properties. Table 1 reports the full list of these input factors and the probability distributions that define their range of variability, while Fig. 4~~3~~ shows an example of slope derived from a combination of input factors.

The slope geometric properties consist~~s~~ of the natural slope angle and height, and the material thickness. Slope angles vary between 20 and 45 degrees to represent typical scenarios of informal housing on moderate and steep slopes. 45° is considered the highest slope angle on which a settlement can be located without some form of engineered slope stabilisation measures. The cross-sectional profile is discretised into three parallel layers of materials to represent the typical weathering profile of volcanic parent material, with a layer of residual soil at the surface (layer 1), underlain by a layer of weathered material (layer 2), and then unweathered bedrock (layer 3). Ranges of material thickness and geotechnical properties are derived from previous field work and lab tests, as described in section 2.1.

The height of the water table is defined as an initial hydrological condition. This water table height is varied between 0% and 90% of the slope height (H in Fig. 3), to account for its variability across the region and~~/or~~ ~~-~~for the variability of the initial soil moisture conditions due ~~by~~to antecedent rainfall events. ~~the uncertainties relating to the initial soil moisture conditions, including the potential effect of antecedent rainfall events.~~

The model is forced with rainfall events which are specified in terms of their duration (in hours) and hourly intensity. The aim is to create both rainfall events that have been observed in the past, and rainfall events that might occur in the future (e.g. with higher ~~rainfall~~ intensity~~ies~~ and duration than ~~historically~~ observed~~s~~ historically). To constrain the rainfall ~~feasibility~~variability space,~~are based on~~ we use the intensity-duration-frequency (IDF) relationships derived from a Gumbel analysis of 40-years of daily rainfall data from weather stations across the island by Klohn-Crippen (1995) (Fig.2a). From these IDFs we deriv~~fin~~e a range~~s~~ of rainfall intensity ~~from~~ between 0 ~~to~~ and 200 mm h^{-1}, and a range of rainfall duration ~~from~~ between 0 ~~to~~ and 72 h. ~~which are~~We then sample~~d~~ independently from the two uniform distributions~~. In this way~~, thus obtaining combinations of ~~rainfall~~ intensity and duration both observed (light grey area in Fig. 2b) and not observed ~~in the past~~ (dark grey area in Fig. 2b) in the past~~ are stochastically generated. we take into account our poor knowledge about future climate by including any possible future design storm in the sampling .~~ Prior to the initiation of the rainfall event we include 168 ~~simulation~~ hours (7 days) of simulation with rainfall intensity equal to zero. This ensures a redistribution of water moisture in the unsaturated zone

of the slope and allows hydrological equilibrium with steady state seepage to be established. A further 168 hours of zero rainfall simulation are added after the storm in order to consider the ground water response after the rainfall event.

Informal housing is represented by four urban properties: slope cutting, absence of roof gutters, vegetation removal, and leaking pipes and tanks. While the angle of the cut slope is varied according to its probability distribution, the vegetation, roof gutters and water leakage are defined as present (option 1 = yes) or absent (option 0 = no) (Fig.~~2~~3). The cut slope angle is varied between 39 and 89 degrees with a maximum cut slope height equal to 4m. We represent the maximum number of cut slopes that can accommodate a house that is 4m wide (+1m of surrounding space) on a slope that is 70m long. We therefore obtain either five or six cut slopes and a corresponding number of houses on each slope depending on the angle of the cut slope. The house width and house load (8 kN m^{-2}) are not varied, and correspond to the size and load of informal houses constructed with shallow concrete strip or block foundations, wooden walls and sheet-metal roofing that are typically observed in Saint Lucia (Holcombe et al., 2016). When vegetation is present on the original non-urbanised slope, it is removed on the surface of the cuts for the urban scenario. The vegetation properties used represent a tropical forest cover, a sensitive choice for this study site (see Holcombe et al. 2016 and online supplementary material, Table S5). These properties are kept fixed throughout the sampling, therefore the effect of different types of vegetation on slope stability is not analysed. Both the tank and the pipe leakage rate are assumed to be half of 4.2e^{-6} m^3 s^{-1}, which corresponds to the estimated leakage of 15% of the total water supply for low income households in Saint Lucia (Anderson and Holcombe 2013). When present, the leak is maintained constant during the simulation time.

The input factors that define the discretisation of the model, such as the cell size of 1m x 1m and the computational time step of 60 s (both used by CHASM's dynamic hydrology functions), and the slip search grid location and dimensions, are not varied. These values are chosen because they typically ensure numerical stability relating to the mass balance of the moisture in the domain and thus a minimum number of failed model's runs. A smaller cell-size would enable a more detailed representation of the slope hydrology, but it would require smaller time step to preserve the moisture content mass balance and numerical stability. Smaller time steps would result in significantly longer simulation time. The resolution chosen is therefore a trade-off between acceptable accuracy and calculation time. The influence of the variation of these two discretisation parameters on slope stability is not explored.

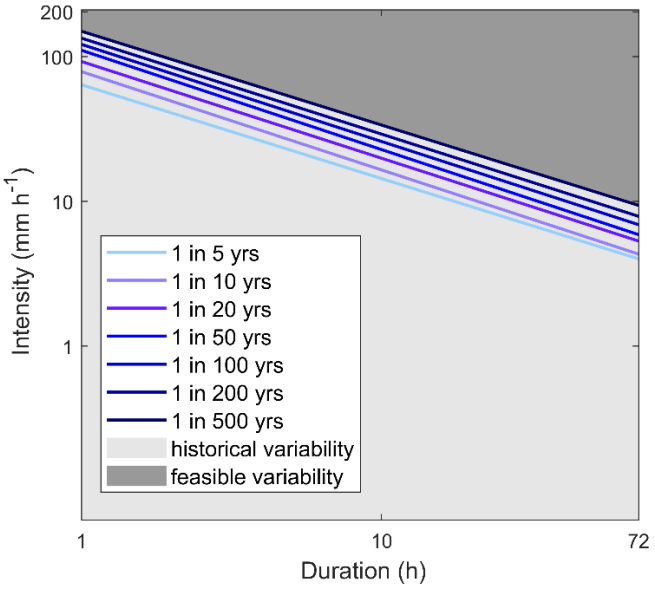


Figure 2: Rainfall intensity–duration–frequency (IDF) curves for Saint Lucia developed by Klohn-Crippen (1995) using Gumbel analysis of 40 years of daily rainfall data from 15 rainfall gauges. The light grey section includes rainfall events from observed data (below IDF curves); the dark grey section represents combinations of rainfall intensity-duration not recorded in the past but that might occur in the future (above IDF curves).

~~daily in this study~~

2.5 Creation of synthetic combinations of input factors and model simulation

We use Latin Hypercube sampling (McKay et al. 1979) to generate 10 000 different combinations of the 30 independently varying input factors shown in Table 1. ~~Figure 3~~ Figure 4 illustrates ~~how each slope is~~one example of a slope defined ~~based~~ by a single combination of ~~on~~ these input factors ~~combination~~. Due to the randomness of the process, checks are undertaken to ensure that realistic combinations of factors are generated; if not, they are discarded (around 70% of the times) and replaced by another randomly generated, feasible combination These "feasibility" checks are reported at the footnote of Table 1 (letters a, b, c, d and f). The stochastically generated simulations~~combinations of input factors~~ are then run in CHASM+ using the high performance computer BlueCrystal Phase 3 at the University of Bristol. The outputs considered for each simulation are the minimum Factor of Safety (FoS) and the slip surface where the minimum FoS is calculated. We divide the completed simulations ~~in~~ according to whether ~~failed or stable according to the value of~~ the minimum FoS ~~(failed if FoS~~is less than 1 (slope predicted to have failed, i.e. a landslide) or greater than 1 (slope is predicted stable). We exclude the simulations ~~that~~ whether the slope is predicted failed before the start of the rainfall event, which represent inherently unstable slopes (for example steep slopes with deep soil thickness). We repeat the same procedure with and without including the urban properties. We therefore obtain two sets of model's outputs: 10 000 representing urbanised slopes, and 10 000 representing non-urbanised slopes.

Table 1: Input factors of CHASM and their probability distributions

Parameter	Symbol/Unit		Range values		
Slope geometric properties:			**Layer 1 ***	**Layer 2 ***	**Layer 3 ***
Slope angle	δ [degrees]	U (20,45)			
Thickness of layer	H [m]		U (1,6)	U (1,6)	
Soil properties:					
Effective cohesion [a]	c [kPa]		Ln (2.3688, 0.5698)	Ln (3.4121, 0.5774)	80
Effective friction angle [b]	φ [degrees]		Ln (3.2937, 0.2092)	Ln (3.1559, 0.3251)	60
Dry unit weight [c]	γd [kN m^{-3}]		U (16,18)	U (18, 20)	23
Saturated moisture content [d]	VG θsat [m^3 m^{-3}]		N (0.413, 0.074)	N (0.413, 0.074)	N (0.413, 0.074)
Residual moisture content [d]	VG θres [m^3 m^{-3}]		Ln (-1.974, 0.376)	Ln (-1.974, 0.376)	Ln (-1.974, 0.376)
VG alpha parameter [d]	VG α [m^{-1}]		Ln (1.264, 1.076)	Ln (1.264, 1.076)	Ln (1.264, 1.076)
VG n parameters [d]	VG n		Ln (0.364, 0.358)	Ln (0.364, 0.358)	Ln (0.364, 0.358)
Saturated Hydraulic Conductivity	Ksat [m s^{-1}]		Ln (-11.055, 0.373)	Ln (-13.357, 0.373)	1xe-8
Initial hydrological condition					
Water table height [e]	DWT [%]	U (0,90)			
Rainfall properties					
Rain intensity	I [m h^{-1}]	U (0 0.2)			
Rain duration	D [h]	Ud (1 72)			
Urban properties:					
Cut slope angle [f]	β [degrees]	N (65.18, 12.61)			
Roof gutters [g]	-	Ud (0 1)			
Vegetation [h]	-	Ud (0 1)			
Septic tank and Pipe leak [i]	Qt/p [m^3 s^{-1}]	Ud (0 1)			

U = Uniform distribution; Ud = Discrete uniform; N = Normal distribution; Ln = Log-normal distribution.
*Layer 1: Residual Soil, Weathering Grade V-VI; Layer2: Weathered material Grade III–IV; Layer3: bedrock Grade I–II; Weathering grades defined according to GEO (1988).

[a] Effective cohesion > 0. Effective cohesion c (layer 3) > c (layer 2)> c (layer 3).

[b] Effective friction angle > 0. Effective friction angle φ (layer 3) > φ (layer 2)> φ (layer 3). φ < 90 degrees

[c] γs =γd +2, where γs is the saturated unit weight. γd (layer 3) > γd (layer 2) > γd (layer 1)

[d] Values from Hodnett and Tomasella (2002) for Sandy Clay Loam material. We impose n > 1; θsat > θres; θres > 0.
VG: Van Genuchten parameters for defining suction moisture characteristics curve.

[e] Water table height is defined as a percentage of slope height measured to the toe of the slope.

[f] Slope of the cut forced to be between 39 and 89 degrees, and it is always greater than natural slope angle

[g] 0 stands for house without rain gutters; 1 stands for house with rain gutters. Roof type = double pitch

[h] Vegetation presence: 0 no vegetation; 1 insert vegetation in the spare spaces.

[i] The leak of the septic tank is equal to the leak of the pipe. When 0 is selected there is no leak, whilst with 1 there are both. The leak rate is constant and equal to 4.2e-6 m^3 s^{-1}

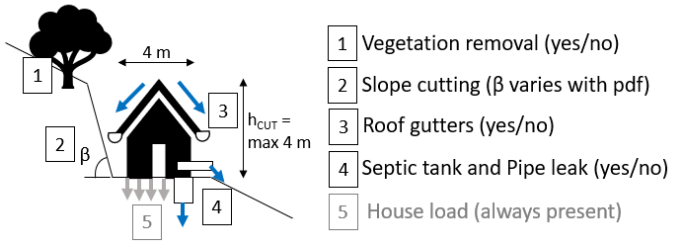


Figure 3: Urban properties of informal housing included in the slope stability analysis. Each house corresponds to a cut on the slope. Cut slope angle varies according to its probability distribution, defined in Table 1. Vegetation, roof gutters, leaking tanks/pipes are stochastically inserted or not. The house on the cut slope is always present and its load is not varied. The height of the cut slope varies relatively to the cut slope angle, but it is forced to be maximum 4 metres

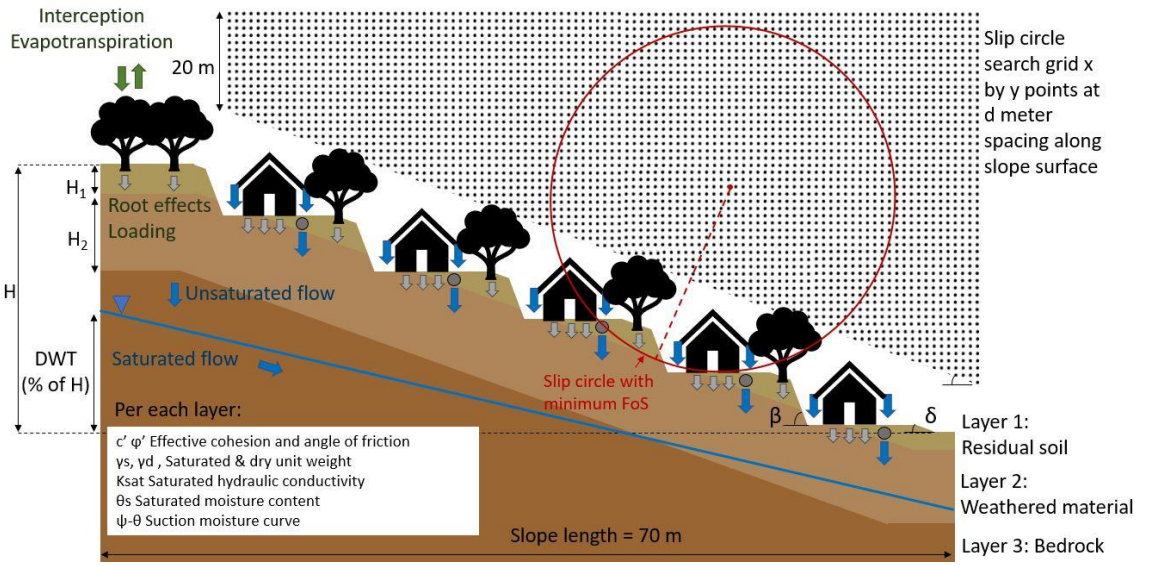


Figure 4: Example of slope generated by stochastically sampling from the ranges of input factors specified in Table 1. H is the slope height resulting from the fixed slope length and varying slope angles. The dimensions of the slip circle search grid are fixed, with initial height of 20 m, and width equal to the slope length. The grid extends downslope parallel to the slope as shown.

2.6 Regional sensitivity analysis (RSA) and Classification And Regression Trees (CART)

Global sensitivity analysis is a set of statistical techniques that evaluate how the variations of a model's outputs can be attributed to the variations of the model's input factors. In this case we want to identify which input factors - among geometry, soil, hydrology, rainfall and urban properties - have the strongest impact on slope stability. Since in our case the model output is binary, as simulated slopes are categorised as failed (if FoS<1) or stable (FoS>1), we use the Regional Sensitivity Analysis (RSA) approach (Hornberger and Spear, 1981) which is particularly suitable when dealing with categorical outputs. In the RSA approach, the cumulative marginal distribution of each input factor is computed for each output category, i.e. the stable slopes and the failed ones. If the distributions significantly separate out, it is taken as evidence that the model output (slope stability) is significantly affected by variations of the considered input factor. The level of separation between the cumulative distributions can be formally measured with the Kolmorov-Smirnov (KS) statistic and used as sensitivity index. The confidence intervals of the sensitivity indexes can be estimated via bootstrap technique. The bootstrap randomly draws N samples (with replacement) from the available data, to compute N KS statistics for each input factor. The magnitude of fluctuations of the KS statistic from one sample to another represents the level of confidence in the estimation of the sensitivity indexes. For this application, we use the SAFE toolbox (Pianosi et al. 2015) to perform RSA and to calculate the sensitivity indices and their confidence intervals by bootstrapping technique.

Classification And Regression Tree (CART) analysis is a supervised machine learning method which we use to predict critical thresholds in input factors over/below which a particular slope is more likely to fail (Breiman et al., 1984). In this analysis, the predictor model takes the form of a binary tree. Starting from the whole set of simulations, CART finds the best possible input

factor (e.g. slope angle rather than rainfall intensity) and the best possible value of that input factor (e.g. slope angle greater or less than 30°) that divide the simulations into stable and failed simulations. This process is recursively repeated, creating at every split two branches and two ("child") nodes of the tree. In choosing the best splitter the model seeks to maximise the "purity", i.e. to maximise the number of stable or failed simulations at the two generated nodes. Amongst the different measures of purity available, we use the Gini purity index defined as:

$$1 - \sum_{i=1}^{m} p^2(i) \quad (1)$$

where m is the number of categories for the output (in this case two: stable or failed) and $p(i)$ is the fraction of simulations in the node belonging to category i. The Gini purity index is 0 when all the simulations in the considered node belong to the same category (a "pure" node, i.e. all stable or failed). The splitting process typically continues until all final leaf nodes show Gini purity indices below a chosen threshold. The final nodes express the prediction for the corresponding branch. While a high number of nodes increases predicting accuracy, it also makes the model more difficult to interpret and generalise to other datasets (i.e. the problem of overfitting). A pruning technique can be applied to avoid this overfitting and to identify an acceptable trade-off between predictive power and number of nodes. In this analysis, we build a CART ~~using~~ within the Matlab Statistics and Machine Learning Toolbox (Matlab R2018a), using the K-fold cross-validation to better estimate its predictive power. In particular, we use 10-fold cross-validation, which randomly divide the original dataset (10 000 simulations) into 10 sub-groups. 9 sub-groups are used to construct 10 CARTs, while the remaining sub-group is used to test the CARTs performance. The average value of the ten misclassification errors so obtained, represents the cross-validation error which can be used to select suitable pruning levels. To reduce the number of nodes without increasing the misclassification errors, auxiliary variables can be used to combine correlated input factors. Auxiliary variables can simplify the tree's structure (by using fewer combined input factors) and potentially modify the input space in a way that the division of failed and stable simulations is more effective (see rotation of the coordinate systems in Dalal et al., 2013). Three auxiliary variables will be used in this analysis: the ratio of soil thickness and effective soil cohesion of layer 1; the ratio between rainfall intensity and duration introduced (both introduced by Almeida et al. 2017) and a weighted combination of natural and cut slopes angles. These variables will be described in the results (CART analysis) section and supplementary material (S2).

3 Results

In this section we analyse the 10 000x2 outputs generated by CHASM+ for the urbanised and non-urbanised slope scenarios. As previously mentioned, we split the simulations into stable and failed according to the value of the minimum FoS (respectively greater or less than one). As a first analysis we compare the percentage of failed slopes against stable slopes for each of the urban properties. Figure 5~~4~~ shows that the presence of cut slopes significantly influences the percentage of predicted slope failures: the steeper the cut slope angle, the higher the percentage of failed slopes. Vegetation removal and roof gutters instead have a negligible role in dividing the two sets. Last, septic tanks and leaking pipes have some effect, generating about 10% more failed slopes when present.

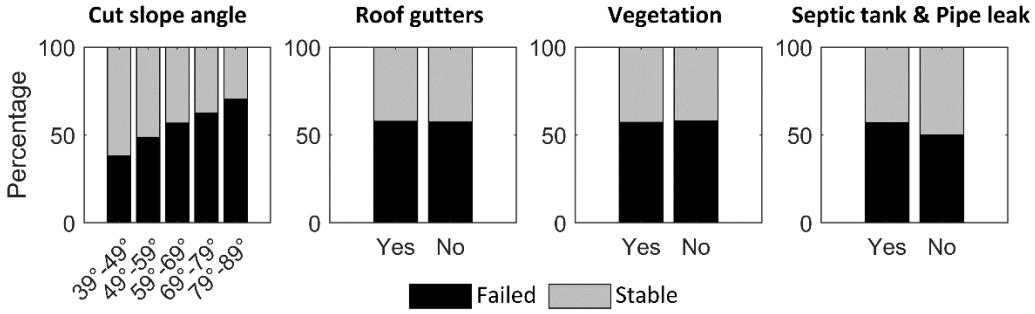


Figure 5: Percentage of predicted stable and failed slopes per each urban property. An urban property will be influencing slope stability if the percentage of the predicted failures changes with the variation of that urban property.

3.1 Regional Sensitivity Analysis

We then perform RSA on both sets of urbanised and non-urbanised slope simulations, calculating the cumulative marginal distributions of the failed and stable simulations for each input factor. The maximum distance between the two distributions (KS statistic) is computed and used as a sensitivity index. A high value of the sensitivity index suggests that the variation of that input factor significantly influences slope stability. The results are shown in Fig. 65, for both urbanised and non-urbanised slopes. Figure 65 shows that slope stability is insensitive to many input factors, and highly sensitive to few, namely effective cohesion and thickness of the layer 1 (residual soil), slope angle, and rain intensity and duration. These sensitive input factors represent the main landslide drivers. The sensitivity indices of the urban properties (in orange) are consistent with the findings of Fig. 54, where only the variation of cut slope angle influences slope stability. When looking at the comparison between urbanised and non-urbanised slopes, it appears that the urban presence decreases the sensitivity indices of all the input factors, except for the effective cohesion of layer 1 and the rainfall intensity.

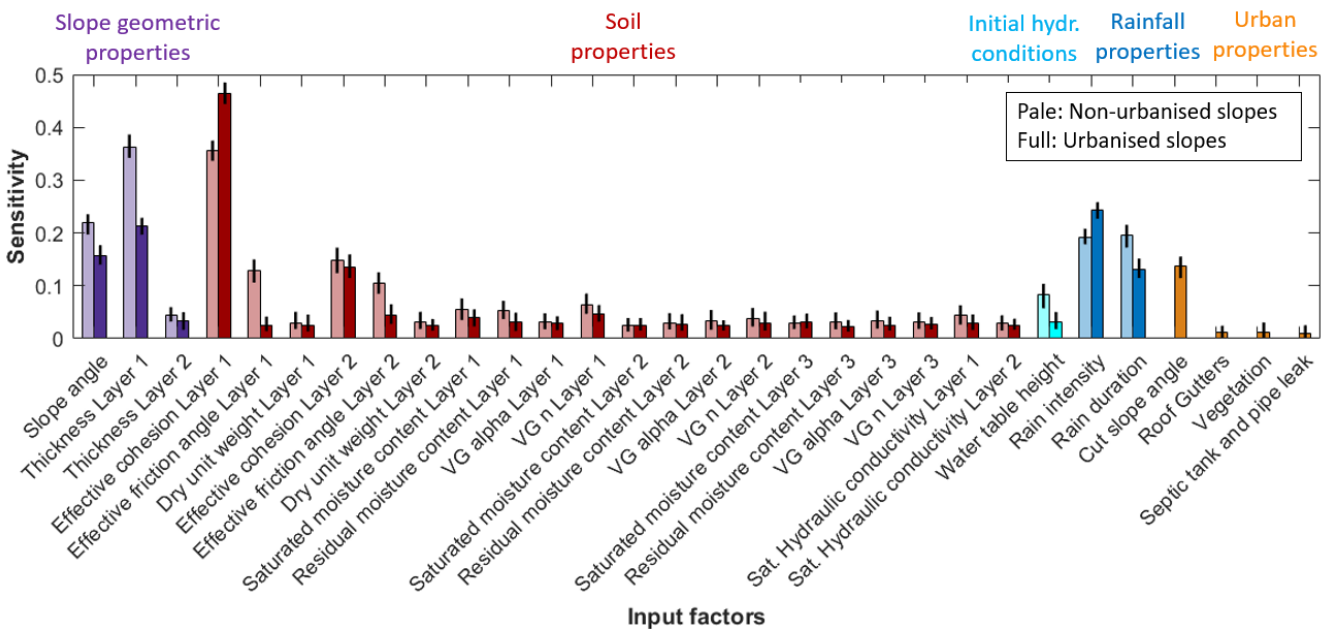


Figure 6: Sensitivity index for each input factor in the urbanised (full colour) and not urbanised (pale colour) cases. The bars correspond to the mean value of sensitivity for each input factor calculated with bootstrapping, while the black vertical lines at the top of the bars represent the confidence interval (Number of bootstrap resampling N = 100; significance level for the confidence intervals 0.05).

We further explore the change in sensitivity caused by urbanisation by plotting the percentage of failed slopes for the main landslide drivers (Fig. 76). The figure shows how this percentage varies for the urbanised (black bars and lines) and non-urbanised cases (green bars and lines). In general, urbanised slopes produce more failures than non-urbanised slopes though

they both display similar trends: an increased percentage of predicted landslides when we would expect the slope to become more susceptible (e.g. when slope angles are higher) or the trigger more severe (when rainfall intensity and duration are larger). For example, in Fig. 7~~6~~b the percentage of failed slopes in the non-urbanised case, linearly increases from ~5% (for soil thickness 1–2m) to ~50% (thickness of 5–6m). In the same figure, urbanised slopes show higher failure rates for all values, though the greatest increase occurs for soil thicknesses less than 4 metres (up to +46% for category 2–3m). This means that the most significant increase in number of landslides occurs for thin soil thicknesses, i.e. on slopes less susceptible to failure when non-urbanised. The same can be said for slope angles less than 25 degrees and rainfall duration less than 10 hours, where percentages of slope failures passes from less than 15% to more than 40% when urbanisation is introduced (Fig 7~~6~~ a,c). In the lower plots instead, more urban landslides are observed on slopes that show high percentage of failures also when urbanisation is not present (+43% for low values of soil cohesion, Fig. 7~~6~~d; +35%, for high rainfall intensities Fig. 7~~6~~e).The difference in failure rates with variations of input factors also explains the change in sensitivity found in Fig. 6~~5~~: when urbanised, slope's response varies less (less sensitive) to variations of the input factors in the upper plots (whose sensitivity indices gets smaller), and more (more sensitive) to variations of the input factors of the lower plots (whose sensitivity indices gets larger).

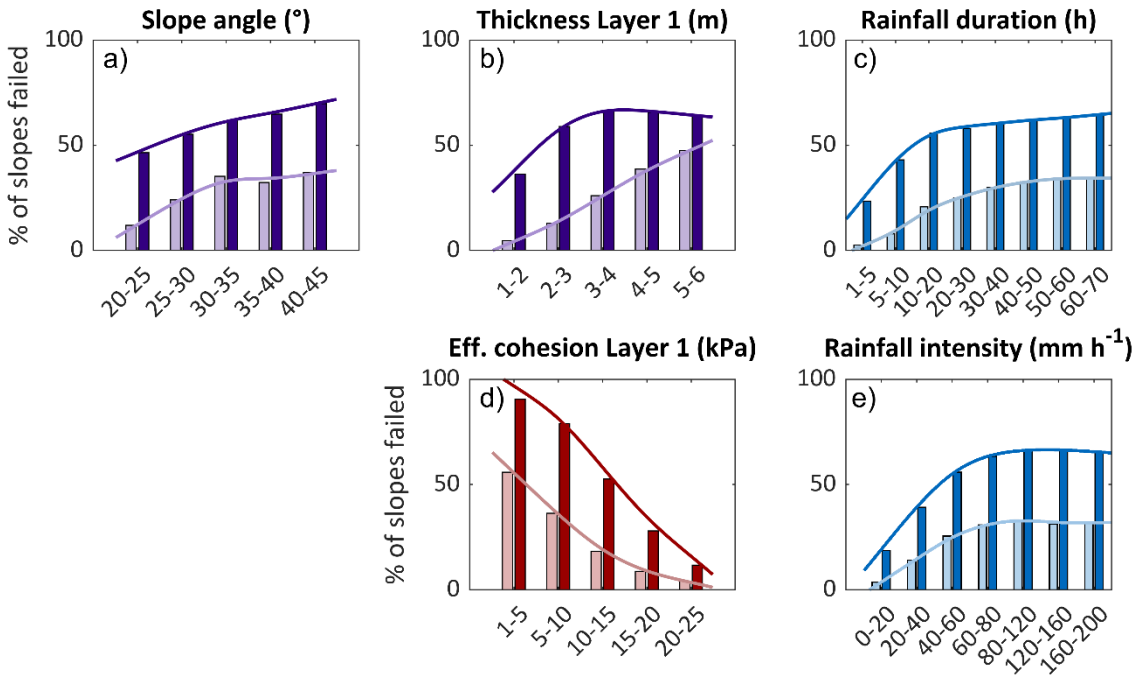
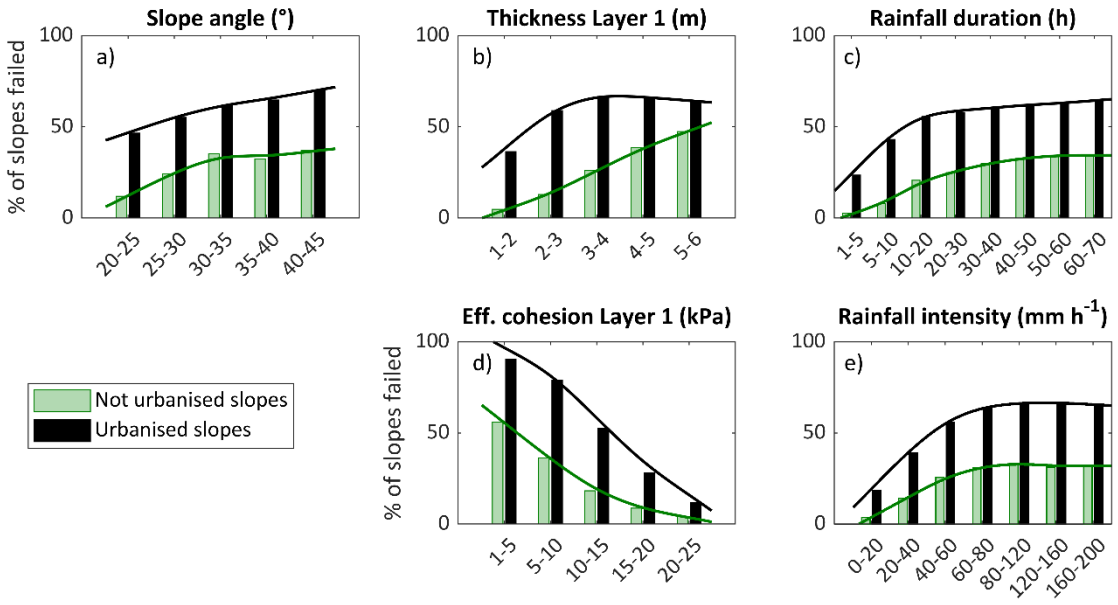


Figure 7: Percentage of slope failures for urbanised and non-urbanised slopes for different categories of input factors. Throughout, urbanised slopes show higher failure rates than non-urbanised slopes. In the upper plots (a), (b), (c), the distribution of failure rates for urbanised slopes are more uniform for variations of input factors than the non-urbanised case, while in the lower plots (d), (c) it is more pronounced. The upper plots represent the input factors whose sensitivity indexes are smaller when urbanisation is introduced in Fig. 6, while the lower plots show the input factors whose sensitivity gets smaller when urbanisation is introduced in Fig. 6



595

~~Figure 6: Percentage of slope failures for urbanised and non-urbanised slopes for different categories of input factors. Throughout, urbanised slopes show higher failure rates than non-urbanised slopes. In the upper plots (a), (b), (c), the distribution of failure rates for urbanised slopes are more uniform for variations of input factors than the non-urbanised case, while in the lower plots (d), (c) it is more pronounced. The upper plots represent the input factors whose sensitivity indexes are smaller when urbanisation is~~

600 **~~introduced in Fig.5, while the lower plots show the input factors whose sensitivity gets smaller when urbanisation is introduced in Fig.5~~**

3.2 CART analysis

We use the CART analysis to formalise the critical thresholds of input factors above/below which slopes are most likely to be

605 predicted as stable or failed. ~~Figure 7~~Figure 8 represents the two trees for the non-urbanised (a) and urbanised case (b). As expected, the best predictor selected in CART are the same input factors previously identified as most influential (Fig. 6~~5~~). The boxes with double colour represent the auxiliary variables that combine correlated input factors: the ratio between effective cohesion and thickness of layer 1 to account their counterbalancing effect on slope stability (i.e. slope with more cohesive soil can be thicker without experiencing failure); the negative ratio between the logarithm of rainfall intensity and rainfall duration,

610 which represent the slope of the rainfall threshold for triggering landslides; and the weighted average of the natural and the cut slope angles, to account that slope susceptibility can significantly increase for low natural slope angles but high cut slopes angles (see Section 2 of the supplementary document for details about the auxiliary variables and the change in model's performance when they are not considered). Using these few predictors, both trees correctly classify more than 85% of the simulations as stable or failed (details about the pruning in Section 3 of the supplementary document). Each branch of the tree

615 shows the paths and thresholds of input factors that lead to slopes most likely to fail (black branch), or most likely to not fail (grey branch). At the end of each branch the black/grey bar shows the fraction of failed and stable simulations, while the thickness of the branch is proportional to the number of simulations following that path. For example, in the tree resulting from non-urbanised slopes (left hand side), the thickest grey line shows that more than 50% of simulated slopes resulted stable 91.2% of the times for ratios of cohesion/thickness of layer 1 greater than 2.5 kPa m^{-1}. The thick black branch instead shows

620 that the greatest proportion of simulations predicted as failed occurred for ratios of cohesion/thickness of layer 1 less than 2.5 kPa m^{-1}, rainfall intensity duration ratios (-log(I)/log(D)) greater than 0.9 m h^{-2} and slope angles greater than 25 degrees.

In the trees resulting from non-urbanised slopes (right hand side), the black branch leading to the majority of failures is similar to the non-urbanised tree, but it presents higher splitting thresholds: from the top, the split happens for ratios of cohesion/thickness of layer 1 less than 4.9 (instead of 2.5) and for rainfall intensity/duration ratio 1.06 (instead of 0.9). The

625 branch then leads to the majority of failures for values of effective cohesion of layer 1 less than 12.6 kPa, regardless of the

natural slope angle. Higher threshold in cohesion/thickness ratios indicates that when urbanisation is present, more failures occur on slopes with higher soil cohesion and/or thinner soil layers than non-urbanised slopes (compatible with Fig. 76b and 76d), while higher rainfall intensity duration ratios suggest that more failures occur for higher rainfall intensity and/or lower rainfall durations when compared to non-urbanised slopes (as shown in Fig. 76c and 76e). Finally, going back to the top and looking at the grey thick branch of the urbanised tree, it can be noted that a ratio between the effective cohesion and the thickness of layer 1 greater than 4.9 ensured 95% of slope stability only when the weighted slope angle is less than 48 degrees.

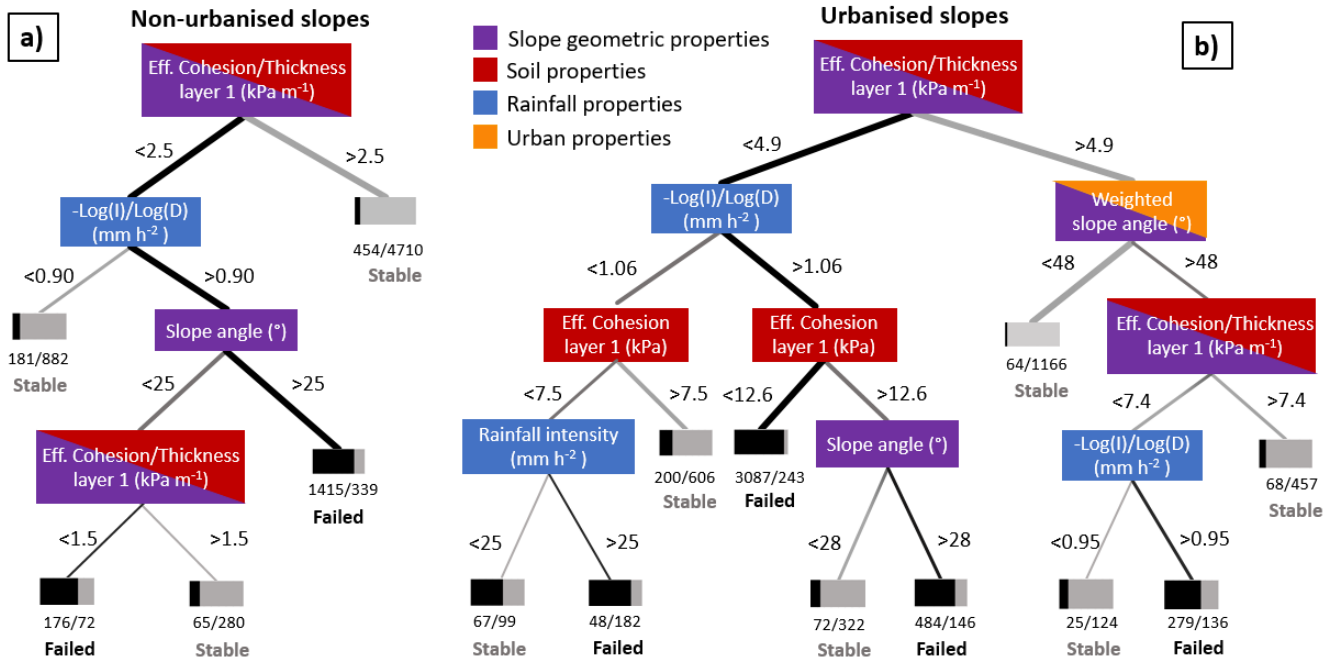


Figure 8: Classification tree of slope response for non-urbanised slopes (a) and urbanised slopes (b). Black branches represent the paths that lead to simulations predicted as failed, while grey branches lead to simulations predicted as stable. The bar under each leaf shows the proportion of simulations that resulted as failed (black) or stable (grey) for that leaf. The thickness of the branch is proportional to the number of simulations following that path. Note as 14% and 22% of the simulated slopes have been excluded respectively for the non-urbanised and urbanised case, because failed before the start of the precipitation.

4 Discussion

4.1 Slope cutting is the urban construction activity most detrimental for slope stability

In this analysis, slope cutting is the urban construction activity with the strongest effect on slope stability's response (Fig.54 and Fig.65). Figure 76 indicates that when urbanisation is present, more slopes failures are observed, mainly on slopes with relatively low slope angles, and with low values of both soil (layer 1)topsoil thickness and of layer 1 and low cohesion values of layer 1 (Fig. 76b,d, also reflected by higher effective cohesion/ thickness ratios in CART in Fig.87b). This is interpreted as caused by cut slopes: when cut slope angles are steep, a higher effective cohesion, and thus a higher soil strength, is required to maintain stability, regardless the natural slope angles; when soil layerss intersect the cuts, small low soil strength resulting from thin soil layers is not sufficient to einsure slope stability even on thin, therefore less susceptiblelandslide-prone, soil layers. The intersection interaction between the depth of soil layer 1 and the cut slope geometry s is deduced from Fig.76b: almost 50% more failures are observed- for a thickness of layer 1 smaller than the slope's height (4 m), i.e. when the layer intersects interface of soil (layer 1) and weathered material (layer 2) outcrops in the cut slope face (as illustrated in Figure 3Fig. 4). For these slopes, visual inspection reveals that the slip surface is generally located between layer 1 (residual soil, weathering grade V – VI) and layer 2 (weathered material, grade III – IV). This is explained by the different soil strength of the two layers, which constrains the slip surface within the weaker layer 1, and the different hydraulic conductivities. As rainfall infiltrates, the lower hydraulic conductivity of the underneath underlying weathered material leads to a progressive

accumulation of water, promoting a perched ~~ground~~water table. The raised pore water pressure decreases the effective soil strength and consequently the stability of the soil layer. Part of the increase in pore water pressure might be caused by the presence of water leakages at the top of the cut slope. However, the low sensitivity of the slope response to leakage (Fig. 6) does not allow for more considerations.

Slope cutting is therefore considered, in this analysis, the most detrimental practice for slope stability. This result is consistent with studies carried out in the humid tropics at regional scales, for which slope cutting was identified as one of the major cause of landslides (e.g. Brand et al., 1984; Froude and Petley, 2018; Holcombe et al., 2016). Cuts with slope angles greater than 60° are also considered at particular high risk (e.g. Cheng, 2009), while excess of pore water pressure was shown to be a dominant process in triggering shallow failures on cut slopes (Anderson, 1983). CHASM therefore successfully captures these physical mechanisms, confirming, despite the uncertainties, the governing role of soil properties and soil thickness in determining slope equilibrium. The other urban construction activities considered, seem to have a less significant role on landslide hazard. Previous studies found that vegetation can be both beneficial and detrimental for slope stability (Wu et al. 1979; Collison and Anderson, 1996). Here we find that its effect is negligible, probably due to its limited presence in urbanised slopes (trees are left at the crest of each cut slope where they add loading and may actually be detrimental to the local cut slope stability). Also, adding roof gutters does not seem to decrease the number of slopes failed. However, in the scenarios generated here we have only reached a maximum of 30% slope coverage by houses, i.e. about 30% of impervious surface (5–6 households on 70 metres slope), due to our inclusion of cut slopes for every house. Evidence shows that roof guttering effectiveness become evident only when the house coverage is above 50%, and thus a considerable portion of rain does not infiltrate into the slope (Anderson and Holcombe, 2013). On the other hand, leak from septic tanks and pipes lead to 10% more failures despite the low house coverage. When higher house densities are considered, the lack of water management might become even more significant (Di Martire et al., 2012).

4.2 The rainfall threshold for triggering landslides is lower when informal housing is included

We found that when slopes are urbanised, the most significant increase in the percentage of slopes failed occurs for rainstorm events with high intensity (>20mm h^{-1}) and low duration (<20 h) (Fig. 7~~6~~c and 7~~6~~e). Accordingly, our CART analysis identifies a higher threshold of rainfall intensity-duration ratio to divide the stable and failed slopes in the urbanised case (Fig. 8~~7~~b). In landslide analysis, so-called minimum rainfall thresholds are defined as the combinations of rainfall intensity (I) and duration (D) above which we would expect landslides starting to occur. These thresholds are generally expressed by a power law relationship $I = \gamma D^{\alpha}$ (Guzzetti et al. 2007), and they are constructed based on inventories of observed landslides and the rainfall that triggered them (e.g. Caine, 1980; Larsen and Simon, 1993; Guzzetti et al., 2007). Many countries in the humid tropics have limited empirical data on landslides, and therefore it would be useful to be able to generate such thresholds from stochastic analyses of the type we performed here. To demonstrate how this could be done, we applied a multi-objective optimisation method to our sample of stochastically generated slopes (details about our approach in Section 4 of the supplementary document). We do not use the more commonly employed frequentist methods (Brunetti et al., 2010; Melillo et al., 2018), because the high frequency of slopes failed for high intensity and high duration events would strongly bias the position of the threshold. Figure 9~~8a and~~and 9~~8~~b show the calculated thresholds on a log-log scale, respectively for the urbanised and not urbanised case (red lines). In both cases, 99.9% of the failed simulations fall above them. The thresholds present the typical descending trend found in empirical analysis, for which lower rainfall intensities are needed to trigger a landslide when rainfall durations increase. The fact that this trend can be replicated by our synthetic simulations indicates that CHASM+ and our stochastic modelling framework are giving realistic hydrological and stability responses to the rainfall forcing.

The higher the intensity and/or the duration of the rainfall event, the more slope failures occur in both cases. However, when informal housing is present, more failures are observed for rainfall duration less than 10 hours (short events - Larsen and Simon, 1993). This pushes down the intercept of the rainfall threshold, as reflected in the change in the coefficients of the

power law equations (reported in each figure). The slope of the threshold line (i.e. the exponent of the power law) is also steeper in the urbanised case, implying the presence of more failures for lower rainfall intensities throughout the duration axis. These results are compatible with the increase of small scale landslides previously commented (failure depths less than cut slope's height): to reach saturation at shallow depths, relatively low rainfall intensities and durations can be sufficient to initiate slope failure. Figure 98c confirms this assumption: when slopes are urbanised (black dots), failures tend to occur with smaller radius of slip surface and for higher values of intensity/duration ratio. The findings reflect the empirical evidence in low income communities which report a high frequency of small scale landslides, particularly associated with cut slopes, for high intensity and short duration events ('the everyday disasters'- Bull-Kamanga et al., 2003). Finally, we compare our results with the empirical rainfall threshold proposed by Larsen and Simon (1993) for Puerto Rico, which is based on landslide inventories that also include failures observed on slopes modified by construction activities (mainly slope cuts for road network, see Larsen and Parks, 1997). When informal urbanisation is considered, the two thresholds are almost overlapping (Fig. 98b). This reinforces both the potential of using mechanistic models within a stochastic framework to generate synthetic thresholds in data scarce locations, and the possibility to use the resulting thresholds for regions of the humid tropics with similar geophysical, climatic and urban properties.

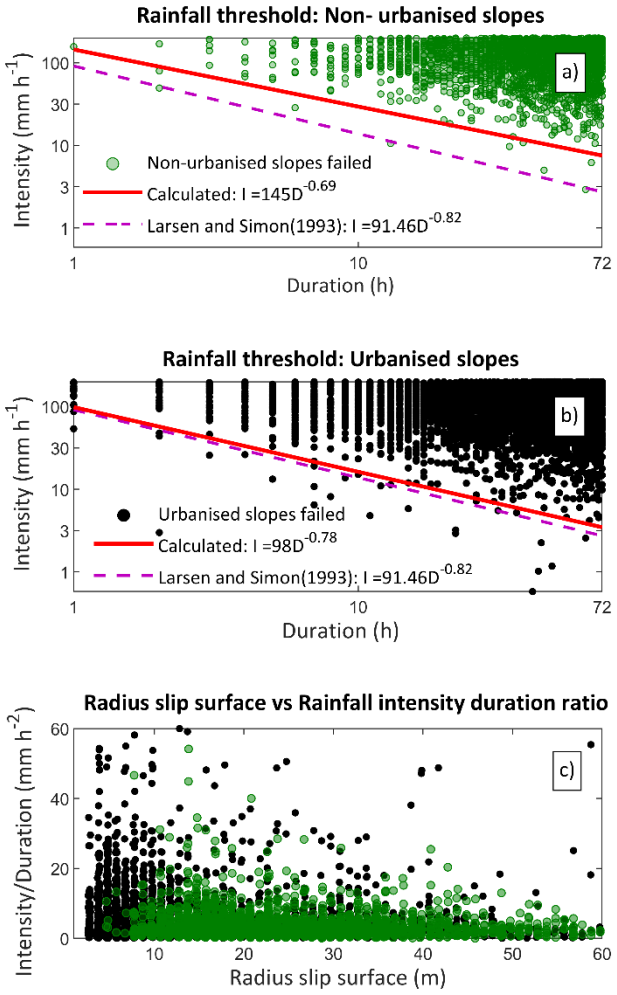


Figure 9: In figure (a) and (b) the red line represent the minimum rainfall thresholds calculated from our stochastic sample (99.9% of the failed slopes in the sample are above the thresholds). Figure (c) represents the radius of the slip surfaces of the recorded landslides plotted against the corresponding triggering rainfall intensity/duration ratio. Note as in (a) and (b) the x and y axis are in logarithmic base 10 scale, but the notation is linear for an easier readibility.

4.3 Guidelines for landslide mitigation actions to tackle the main instability drivers

The identification of the main instability drivers and their thresholds can contribute to create objective rules to classify slopes as hazardous in a region with scarce data availability. For example, in Saint Lucia our analysis suggests that slopes with

effective cohesion of layer 1 less than 12kPa and thickness less than 2.5m are particularly at risk for rainfall events with intensity/duration ratio greater than 1.06 m h^{-2} (Fig. 87b). These rules can shape look up tables or priority ranking to classify man made slopes as dangerous (Anderson and Lloyd, 1991; Cheng, 2009). Figure 65 shows that only few input factors particularly influence slope stability with or without urbanisation. These are effective cohesion and thickness of the layer 1 (residual soil), natural and cut slope angles, and rain intensity and duration. The other input factors might have a smaller direct or indirect effect, but they are not dominant. This is an expected finding in global sensitivity analysis (Wagener and Pianosi 2019), despite different outputs (e.g. the timing of the failure) might be sensitive to different input factors (e.g. variations in the moisture suction curves, as demonstrated in the supplementary material S1.2). The identification of these main landslide drivers helps addressing data acquisition efforts, while the comparison between urbanised and not urbanised simulations quantifies the different relative role (e.g. weight) of preparatory factors in landslide susceptibility assessment when informal urbanisation is present. For example, a weighted average of natural and cut slope angle can be used to identify areas (not) at risk.

All the results presented are subjected to the assumptions made in our study. The large variation of some of the input factors can lead to overestimating the hazard. Almeida et al. (2017), for example, varied the slope angles between 27 and 30 degrees (instead of between 20 and 45 degrees), and hence found a lower value of the cohesion/thickness ratio to separate stable and failed slopes than we found (in the non-urbanised case). Data acquisition can help to reduce these uncertainties. However, when data are not available, our approach allows for the identification of so called 'low regrets' mitigation measures, i.e. actions that have a positive impact on slope stability regardless of the uncertain factors. According to our analysis, the most effective action would be avoiding slope cutting, since it resulted as the most detrimental urban construction activity for slope stability. However, this is of scarce utility since informal housing often outstrips urban regulations (Fekade, 2000). Better hazard awareness and construction practices should be therefore suggested. These include for example reducing surface water infiltration on slopes, especially when the topsoil layers intersect the cut slope and the resulting perched water tables reduce shear strength in a critical location. Slope surface and subsurface drainage can be designed to reduce the infiltration of rainwater to a level that, in effect, reduces the total rainfall intensity below the rainfall threshold calculated. Another cost-effective landslide mitigation strategy can be the planting of deep rooting grasses, shrubs or small trees which increases slope strength (e.g. slope cohesion) in the top couple of metres of soil and also reduces soil moisture content though root-water uptake and evapotranspiration (Holcombe et al., 2016; Ng et al., 2011; Wilkinson et al., 2002).

Finally, Fig. 98b shows that when slopes are urbanised, high intensity, short duration rainfall events lead to an increased number of small-scale landslides (failure depths less than 4m, Fig. 76b and radius of slip failure less than 10m, Fig. 98c). Future climate change could potentially increase the frequency of intense precipitation events (e.g. O'Gorman and Schneider, 2009), and therefore the occurrence of these type of landslides in informal communities. However, if small scale failures produced by anthropogenic factors are neglected in the calculation of rainfall thresholds, also current rainstorms events could be excluded as triggering factors (Crozier, 2010; Mendes et al., 2018). Small scale, high frequency landslide events might not lead to major disasters, but they are increasingly seen as indicators of risk accumulation, detrimental for disaster resilience and economic development (Bull-Kamanga et al., 2003). For this reason, these types of landslides deserve a greater attention from the scientific community.

5. Conclusions

We include informal housing into slope stability analysis using a newly extended version of the mechanistic model CHASM in a Monte Carlo framework. In this way, we consider uncertainties due to both poorly known slope properties and potential future changes in urban and climate conditions. We demonstrate that informal housing increases landslide hazard and that slope cutting is the most detrimental construction activity, when compared to vegetation removal, lack of roof gutters and

presence of water leaks. The presence of informal housing also modifies the relative role that natural slope angle, soil cohesion and soil thickness have in maintaining slopes stable, with increased hazard occurrence for low values of these three main landslide drivers. CART analysis identifies the thresholds of input factors separating stable and unstable slopes. These thresholds can be used as objective criteria for guiding local engineers in identifying slopes at risk, deducing landslide mitigation actions, as well as targeting data acquisition to reduce model prediction uncertainty. Moreover, this analysis allows for the estimation of critical rainfall thresholds at which slope failure is predicted to occur. This rainfall threshold is lower when informal housing is present, with an increased number of small scale landslides (+85%, with failure depth less than 4m and radius of slip surface less than 10m) for high intensity and short duration events. The rainfall threshold resulting from the urbanised slopes is comparable to the one proposed by Larsen and Simon (1993) for the region of Puerto Rico, suggesting its potential validity also for other similar (data scarce) regions of the humid tropics.

Future work will seek to vary the properties that were kept constant in this study, such as the degree of urbanisation and house dimensions, to evaluate their significance for slope stability. This might confirm the importance of household water management such as roof-guttering, leaking water supply pipes and septic tanks when the number of households is increased. Analysis of slopes where slope cutting is replaced by other possible construction techniques (such as houses suspended on pile foundations) can identify whether the construction of future hillside settlements could be done in a manner less detrimental to slope stability. Different bioengineering techniques to mitigate hazard likelihood could also be modelled and their effectiveness evaluated. Finally, we seek to transfer the thresholds found in our CART analysis into spatial scale susceptibility maps in order to identify slopes at higher risk within low income urban settlements. This would confirm whether the areas suggested to be most hazardous correspond to areas where more landslides have been observed.

Supplement. The supplement related to this article is available online at:

Author contribution. EB performed background research, performed the computations and analysis, and wrote the paper. EH, FP and TW supervised the entire study in all stages, discussed the results and contributed to the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Existing MATLAB codes from Susana Almeida and Rose Hen Jones have been used for this analysis. We thank Prof. Dave Petley (University of Sheffield) for giving permission to use one of the pictures for Kalimpong, India (http://savethehills.blogspot.com/). Funding: The first author was supported by an EPSRC PhD studentship. Partial support to ~~TW~~ Thorsten Wagener was provided by a Royal Society Wolfson Research Merit Award [WM170042]. ~~FP~~ Francesca Pianosi is partially funded by the Engineering and Physical Sciences Research Council (EPSRC) "Living with Environmental Uncertainty" Fellowship [EP/R007330/1].

References

Alewell, C. and Meusburger, K.: Impacts of anthropogenic and environmental factors on the occurrence of shallow landslides in an alpine catchment (Urseren Valley, Switzerland), Nat. Hazards Earth Syst. Sci., 8(3), 509–520 [online] Available from: http://hal.archives-ouvertes.fr/hal-00299525, 2008.

Almeida, S., Ann Holcombe, E., Pianosi, F. and Wagener, T.: Dealing with deep uncertainties in landslide modelling for disaster risk reduction under climate change, Nat. Hazards Earth Syst. Sci., 17(2), 225–241, doi:10.5194/nhess-17-225-2017, 2017.

Anderson, M., Holcombe, L., Flory, R. and Renaud, J. P.: Implementing low-cost landslide risk reduction: A pilot study in

unplanned housing areas of the Caribbean, Nat. Hazards, 47(3), 297–315, doi:10.1007/s11069-008-9220-z, 2008.

Anderson, M. G: Mid-Levels Study. Report on Geology, Hydrology and Soil Properties, Geotechnical Control Office, Civil Engineering Services Department, Hong Kong, 1982.

Anderson, M. G.: Road-cut slope topography and stability relationships in St Lucia, West Indies, Appl. Geogr., 3(2), 105–114, doi:10.1016/0143-6228(83)90033-4, 1983.

Anderson, M. G.: A feasibility study in mathematical modelling of slope hydrology and stability, Report to Geotechnical Control Office Civil Engineering Services Department, Hong Kong, 1990.

Anderson, M. G. and Holcombe, E.: Community-based landslide risk reduction: managing disasters in small steps, The World Bank., 2013.

Anderson, M. G. and Howes, S.: Development and application of a combined soil water-slope stability model, Q. J. Eng. Geol. Hydrogeol., 18(3), 225–236, 1985.

Anderson M.G. and Kemp M.J.: The prediction of pore water pressure conditions in road cut slopes, St Lucia, West Indies. Overseas Development Agency, London, UK, Final Technical Report R3426, 1985.

Anderson, M. G. and Lloyd, D. M.: USING A COMBINED SLOPE HYDROLOGY-STABILITY MODEL TO DEVELOP CUT SLOPE DESIGN CHARTS., Proc. Inst. Civ. Eng., 91(4), 705–718, 1991.

Anderson, M. G., Holcombe, E., Holm-Nielsen, N. and Della Monica, R.: What Are the Emerging Challenges for Community-Based Landslide Risk Reduction in Developing Countries?, Nat. Hazards Rev., 15(2), 128–139, doi:10.1061/(asce)nh.1527-6996.0000125, 2013.

Beven, K., Almeida, S., Aspinall, W. P., Bates, P. D., Blazkova, S., Borgomeo, E., Freer, J., Goda, K., Hall, J., Phillips, J. C. and others: Epistemic uncertainties and natural hazard risk assessment-Part 1: A review of different natural hazard areas, Nat. hazards earth Syst. Sci., 18(10), 2741–2768, 2018a.

Beven, K. J., Aspinall, W. P., Bates, P. D., Borgomeo, E., Goda, K., Hall, J. W., Page, T., Phillips, J. C., Simpson, M., Smith, P. J., Wagener, T. and Watson, M.: Epistemic uncertainties and natural hazard risk assessment - Part 2: What should constitute good practice?, Nat. Hazards Earth Syst. Sci., 18(10), 2769–2783, doi:10.5194/nhess-18-2769-2018, 2018b.

Bishop, A. W.: The use of the Slip Circle in the Stability Analysis of Slopes, Géotechnique, 5(1), 7–17, doi:10.1680/geot.1955.5.1.7, 1955.

Brand, E. W., Premchitt, J. and Phillipson, H. B.: Relationship between rainfall and landslides in Hong Kong, in Proceedings of the 4th International Symposium on Landslides, vol. 1, pp. 377–384., 1984.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and regression trees, CRC Press, 1984.

Brooks, S. M., Crozier, M. J., Glade, T. W. and Anderson, M. G.: Towards Establishing Climatic Thresholds for Slope Instability: Use of a Physically-based Combined Soil Hydrology-slope Stability Model, Pure Appl. Geophys., 161(4), 881–905, doi:10.1007/s00024-003-2477-y, 2004.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D. and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10(3), 447–458, 2010.

Bull-Kamanga, L., Diagne, K., Lavell, A., Leon, E., Lerise, F., MacGregor, H., Maskrey, A., Meshack, M., Pelling, M., Reid, H. and others: From everyday hazards to disasters: the accumulation of risk in urban areas, Environ. Urban., 15(1), 193–204, 2003.

CHARIM, 2015. Caribbean Handbook on Risk Information Management. Available at: http://www.charim.net/

Caine, N.: The rainfall intensity-duration control of shallow landslides and debris flows, Geogr. Ann. Ser. A, Phys. Geogr., 62(1–2), 23–27, 1980.

Cheng, P. F. K.: The New Priority Ranking Systems for Man-made Slopes and Retaining Walls, Special Project Report (SPR) 4/2009, Geotech. Eng. Off. (GEO), Hong Kong, 6–14, 2009.

Cho, S. E.: Effects of spatial variability of soil properties on slope stability, Eng. Geol., 92(3–4), 97–109,

doi:10.1016/j.enggeo.2007.03.006, 2007.

Ciabatta, L., Camici, S., Brocca, L., Ponziani, F., Stelluti, M., Berni, N. and Moramarco, T.: Assessing the impact of climate-change scenarios on landslide occurrence in Umbria Region, Italy, J. Hydrol., 541, 285–295, doi:10.1016/j.jhydrol.2016.02.007, 2016.

Collison, A. J. C., Anderson, M. G. and Lloyd, D. M.: Impact of vegetation on slope stability in a humid tropical environment: A modelling approach, Proc. Inst. Civ. Eng. Water, Marit. Energy, 112(2), 168–175, doi:10.1680/iwtme.1995.27662, 1995.

Crozier, M. J.: Geomorphology Deciphering the effect of climate change on landslide activity : A review, Geomorphology, 124(3–4), 260–267, doi:10.1016/j.geomorph.2010.04.009, 2010.

Dalal, S., Han, B., Lempert, R., Jaycocks, A. and Hackbarth, A.: Improving scenario discovery using orthogonal rotations, Environ. Model. Softw., 48, 49–64, 2013.

DeGraff, J. F.: Landslide hazard on St. Lucia, West Indies- Final Report: Washington D. C., Organization of American States, 1985.

Diaz, V. J.: Landslides in the squatter settlements of Caracas; towards a better understanding of causative factors, Environ. Urban., 4(2), 80–89, 1992.

Di Martire, D., De Rosa, M., Pesce, V., Santangelo, M. A. and Calcaterra, D.: Landslide hazard and land management in high-density urban areas of Campania region, Italy, Nat. Hazards Earth Syst. Sci., 12(4), 905–926, doi:10.5194/nhess-12-905-2012, 2012

DIWI Consult: Materials Report. RDP-001 Soufriere-Choiseul. DIWI Consult International GmbH, Consulting Services for the Roads Development Programme, Ministry of Communications, Works, Transport & Public Utilities, Government of Saint Lucia, 2002

El-Ramly, H., Morgenstern, N. R. and Cruden, D. M.: Probabilistic assessment of stability of a cut slope in residual soil, Risk Var. Geotech. Eng. Inst. Civ. Eng., (1), 197–204, doi:10.1680/ravige.34860.0020, 2006.

Fekade, W.: Deficits of formal urban land management and informal responses under rapid urban growth, an international perspective, Habitat Int., 24(2), 127–150, doi:10.1016/S0197-3975(99)00034-X, 2000.

Formetta, G., Rago, V., Capparelli, G., Rigon, R. and Versace, P.: Integrated Physically based system for modeling landslide susceptibility, Procedia Earth Planet. Sci., 9, 74–82, doi:10.1016/j.proeps.2014.06.006, 2014.

Froude, M. J. and Petley, D. N.: Global fatal landslide occurrence from 2004 to 2016, , (2012), 2161–2181, 2018.

GEO: Geoguide 3- Guide to Rock and Soil Descriptions, Geotech. Eng. Off., 177, 1988.

Gerrard, J. and Gardner, R.: Relationships Between Landsliding and Land Use in the Likhu Khola Drainage Basin, Middle Hills, Nepal, Mt. Res. Dev., 22(1), 48–55, doi:10.1659/0276-4741(2002)022[0048:rblalu]2.0.co;2, 2006.

Groves, D. G. and Lempert, R. J.: A new analytic method for find–ing policy-relevant scenarios, Global Environ. Chang., 17, 73– 85, doi:10.1016/j.gloenvcha.2006.11.006, 2007.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, , 267, 239–267, doi:10.1007/s00703-007-0262-7, 2007.

Hessel, R.: Effects of grid cell size and time step length on simulation results of the Limburg soil erosion model (LISEM), , 3049(February 2004), 3037–3049, doi:10.1002/hyp.5815, 2005.

Hodnett, M. G. and Tomasella, J.: Marked differences between van Genuchten soil water-retention parameters for temperate and tropical soils : a new water-retention pedo-transfer functions developed for tropical soils, , 108, 155–180, 2002.

Holcombe E.A.: Modelling Landslide Risk on Highway Cut Slopes in Developing Countries, PhD thesis, University of Bristol, UK, 2006.

Holcombe, E. A., Beesley, M. E. W., Vardanega, P. J. and Sorbie, R.: Urbanisation and landslides: hazard drivers and better

practices, Proc. Inst. Civ. Eng. - Civ. Eng., 169(3), 137–144, doi:10.1680/jcien.15.00044, 2016.

Hornberger, G. M. and Spear, R. C.: An approach to the preliminary analysis of environmental systems, J. Environ. Manage., 12(1), 7–18, 1981.

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

Klohn-Crippen: Roseau Dam and ancillary works. Tropical storm Debbie, final report on hydrology, Unpublished report held by WASCO, Saint Lucia, 1995.

Larsen, M. C.: Rainfall-triggered landslides, anthropogenic hazards, and mitigation strategies, Adv. Geosci., 14, 147–153, doi:10.5194/adgeo-14-147-2008, 2008.

Larsen, M. C. and Parks, J. E.: How wide is a road? The association of roads and mass-wasting in a forested montane environment, Earth Surf. Process. Landforms J. Br. Geomorphol. Gr., 22(9), 835–848, 1997.

Larsen, M. C. and Simon, A.: A rainfall intensity-duration threshold for landslides in a humid-tropical environment, Puerto Rico, Geogr. Ann. Ser. A, Phys. Geogr., 75(1–2), 13–23, 1993.

Lloyd, D. M., Anderson, M. G., Hussein, A. N., Jamaludin, A. and Wilkinson, P. L.: Preventing landslides on roads and railways: A new risk-based approach, Proc. Inst. Civ. Eng. Civ. Eng., 144(3), 129–134, doi:10.1680/cien.144.3.129.39903, 2001.

Lumb, P.: Slope failures in Hong Kong, Q. J. Eng. Geol. Hydrogeol., 8(1), 31–65, doi:10.1144/GSL.QJEG.1975.008.01.02, 1975.

Maes, J., Kervyn, M., de Hontheim, A., Dewitte, O., Jacobs, L., Mertens, K., Vanmaercke, M., Vranken, L. and Poesen, J.: Landslide risk reduction measures, Prog. Phys. Geogr., (iv), 030913331668934, doi:10.1177/0309133316689344, 2017.

Di Martire, D., De Rosa, M., Pesce, V., Santangelo, M. A. and Calcaterra, D.: Landslide hazard and land management in high-density urban areas of Campania region, Italy, Nat. Hazards Earth Syst. Sci., 12(4), 905–926, doi:10.5194/nhess-12-905-2012, 2012.

McKay, M. D., Beckman, R. J. and Conover, W. J.: Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21(2), 239–245, doi:10.1080/00401706.1979.10489755, 1979.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A. and Guzzetti, F.: Environmental Modelling & Software A tool for the automatic calculation of rainfall thresholds for landslide occurrence, Environ. Model. Softw., 105, 230–243, doi:10.1016/j.envsoft.2018.03.024, 2018.

Mendes, R. M., de Andrade, M. R. M., Graminha, C. A., Prieto, C. C., de Ávila, F. F. and Camarinha, P. I. M.: Stability analysis on urban slopes: case study of an anthropogenic-induced landslide in São José dos Campos, Brazil, Geotech. Geol. Eng., 36(1), 599–610, 2018.

Mott MacDonald: Landslide Risk Assessment for Saint Lucia's Primary Road Network, Hurricane Tomas Rehabilitation and Reconstruction Final Feasibility Report, , (September), 258, 2013.

Ng, S. L., Chu, L. M., Li, L. and Qin, J.: Performance assessment of slope greening techniques in Hong Kong, Asian Geogr., 28(2), 135–145, 2011.

Nicolet, P., Foresti, L., Caspar, O. and Jaboyedoff, M.: Shallow landslide's stochastic risk modelling based on the precipitation event of August 2005 in Switzerland: Results and implications, Nat. Hazards Earth Syst. Sci., 13(12), 3169–3184, doi:10.5194/nhess-13-3169-2013, 2013.

O'Gorman, P. A. and Schneider, T.: The physical basis for increases in precipitation extremes in simulations of 21st-century climate change, Proc. Natl. Acad. Sci., 106(35), 14773–14777, doi:10.1073/pnas.0907610106, 2009.

Peduzzi, P., Chatenoux, B., Dao, H., De Bono, A., Deichmann, U., Giuliani, G., Herold, C., Kalsnes, B., Kluser, S., Løvholt, F. and others: The Global Risk Analysis for the 2009 Global Assessment Report on Disaster Risk Reduction, Www-Fourier.Ujf-Grenoble.Fr, 1–6 [online] Available from: http://www-fourier.ujf-

grenoble.fr/~mouton/Publis_HDR_applis/Peduzzi-The_Global_Risk_Analysis_for_the_2009_GAR-149.pdf, 2009.

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

Pianosi, F., Sarrazin, F. and Wagener, T.: Environmental Modelling & Software Short communication A Matlab toolbox for Global Sensitivity Analysis, Environ. Model. Softw., 70, 80–85, doi:10.1016/j.envsoft.2015.04.009, 2015.

Premchitt, J.: Rain-induced landslides in Hong Kong 1972-1992, , (June), 2016.

Preuth, T., Glade, T. and Demoulin, A.: Geomorphology Stability analysis of a human-in fl uenced landslide in eastern Belgium, Geomorphology, 120(1–2), 38–47, doi:10.1016/j.geomorph.2009.09.013, 2010a.

Preuth, T., Glade, T. and Demoulin, A.: Stability analysis of a human-influenced landslide in eastern Belgium, Geomorphology, 120(1–2), 38–47, doi:10.1016/j.geomorph.2009.09.013, 2010b.

SafeLand: Deliverable 1.6: Analysis of Landslides triggered by anthropogenic factors in Europe, SafeLand European Project Living with Landslide Risk in Europe: Assessment, Effects of Global Changes, and Risk Management Strategies. Edited by Nadim, F., Høydal, Ø, Haugland, H. and McLean, A., 2011.

Segoni, S., Piciullo, L. and Gariano, S. L.: A review of the recent literature on rainfall thresholds for landslide occurrence, Landslides, 15(8), 1483–1501, doi:10.1007/s10346-018-0966-4, 2018.

Sidle, R. C. and Ziegler, A. D.: The dilemma of mountain roads, Nat. Geosci., 5(7), 437–438, doi:10.1038/ngeo1512, 2012.

Smyth, C. G. and Royle, S. A.: Urban landslide hazards: Incidence and causative factors in Niteroi, Rio de Janeiro state, Brazil, Appl. Geogr., 20(2), 95–118, doi:10.1016/S0143-6228(00)00004-7, 2000.

UN-Habitat: The challenge of slums: global report on human settlements 2003, Manag. Environ. Qual. An Int. J., 15(3), 337–338, 2003.

UN-Habitat: Habitat Iii Issue Papers 22 – Informal Settlements, United Nations Conf. Hous. Sustain. Urban Dev., 2015(May), 0–8, doi:http://dx.doi.org/10.3402/gha.v5i0.19065, 2015.

Van Westen, C.J.: National Scale Landslide Susceptibility Assessment for Saint Lucia. CHARIM Caribbean Handbook on Risk Information Management, World Bank GFDRR, ACP-EU Natural Disaster Risk Reduction Program. http://www.charim.net/sites/default/files/handbook/maps/SAINT_LUCIA/SLULandslideReport.pdf, 2016.

Vanacker, V., Vanderschaeghe, M., Govers, G., Willems, E., Poesen, J., Deckers, J. and De Bievre, B.: Linking hydrological, infinite slope stability and land-use change models through GIS for assessing the impact of deforestation on slope stability in high Andean watersheds, Geomorphology, 52(3–4), 299–315, doi:10.1016/S0169-555X(02)00263-5, 2003.

Wagener, T. and Pianosi, F.: What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling, Earth-Science Rev., 194(April), 1–18, doi:10.1016/j.earscirev.2019.04.006, 2019.

Wilkinson, P. L., Anderson, M. G., Lloyd, D. M. and Renaud, J. P.: Landslide hazard and bioengineering: Towards providing improved decision support through integrated numerical model development, Environ. Model. Softw., 17(4), 333–344, doi:10.1016/S1364-8152(01)00078-0, 2002.

Wilkinson, P. L., Anderson, M. G. and Lloyd, D. M.: AN INTEGRATED HYDROLOGICAL MODEL FOR RAIN-INDUCED LANDSLIDE PREDICTION, , 1297(September), 1285–1297, doi:10.1002/esp.409, n.d.

World Bank: Disaster risk management in Latin America and the Caribbean Region: GFDRR country notes (English). Washington, DC: World Bank. http://documents.worldbank.org/curated/en/826811468010903390/Disaster-risk-management-in-Latin-America-and-the-Caribbean-Region-GFDRR-country-notes, 226 pp., 2012.

Wilby, R. L. and Dessai, S.: Robust adaptation to climate change, Weather, 65, 180–185, doi:10.1002/Wea.543, 2010.

Wu, T. H., McKinnell, W. P. and Swanston, D. N.: Strength of Tree Roots and Landslides on Price of Wales Island, Alaska., Can Geotech J, 16(1), 19–33, doi:10.1139/t79-003, 1979.

Zhang, F., Liu, G., Chen, W., Liang, S., Chen, R. and Han, W.: Human-induced landslide on a high cut slope: a case of repeated

failures due to multi-excavation, J. Rock Mech. Geotech. Eng., 4(4), 367–374, doi:10.3724/SP.J.1235.2012.00367, 2012.

975

980

985

990

995

1000

1005

1010

Supplement reviewed. Highlighted the changes related to the reviewers comments. Please refer to the .pdf uploaded as supplement for the fully reviewed version.

S1 Slope water management in the Combined Hydrology And Stability Model, CHASM

Section 1 is divided in two parts: the first describes the new functionality developed in CHASM representing slope water management (S1.1); the second part illustrates its benchmark against another slope stability software (S1.2).

S1.1 Description of urban slope water management in CHASM+

We have developed new functionality in the CHASM code (which we are calling 'CHASM+') which is now able to simulate three additional processes: water leaks from buried septic tanks, leaks from superficial pipes, and the effect of houses without roof gutters discharging rainwater from their roofs onto the slope.

Leaking Septic tanks: the user can determine the position of the tanks, their dimensions (width and depth), the leakage rate ($m^3\ s^{-1}$) and the type of leakage (local or evenly distributed). Considering that the slope cross section is represented with a mesh of columns and cells, a tank will occupy some of these cells according to its dimensions and position. These cells are modelled as being impermeable and heavier than the surrounding soil. The water leakage is added to the moisture content of the cells underneath the tank, through the following Water Balance equation (S1):

$$\frac{\partial \theta}{\partial t} = \frac{\partial (Q + Q_{leak})}{\partial z} \quad \text{(S1)}$$

where, θ is the moisture content, changing over time according to the water flow Q, and Q_{leak} represents the water leaked by the tank, which is constant throughout the simulation time. When water is added into the cell, the moisture content increases. The unsaturated hydraulic conductivity, which depends on the moisture content, also increases and is iteratively calculated with the Millington-Quirk formulation (Millington and Quirk, 1959). The maximum value is reached when soil is saturated (saturated hydraulic conductivity, fixed by the user).

Note, buried leaking pipes can also be simulated by using this option by not considering the load of the tank.

Leaking Pipe on the slope surface: we want to simulate pipes discharging water onto the slope surface. This can be due by low pipe maintenance or when water collectors are poorly designed and not properly connected to formal drainage or sewerage system (Ortuste, 2012). In CHASM the slope cross section is represented by a two-dimensional mesh of columns and cells for the purposes of the hydrology calculations. Therefore, the water leaked by the pipe is added to the surface water of the column of the slope where the pipe is positioned. This water infiltrates into the slope according to the infiltration capacity of the top cell of that column, which is a function of its hydraulic conductivity. The water that does not infiltrate because exceeds the infiltration capacity, is stored on the surface as ponding water. The maximum storage of ponding water is determined by the user as detention capacity of that cell. If the ponding water exceeds this value, it is removed from the calculation because surface water runoff is not included in the CHASM hydrology scheme. The leakage when present is constant throughout the simulation time.

Houses without gutters: Where houses are present, rainfall does not reach the top cell of the slope underneath the house, and the amount of rain intercepted by the roof is calculated and discharged onto (added to) the top cells of the slope to the sides of the house. If the roof is dual pitch, half of the intercepted rain is discharged upslope and half downslope of the house, and it is equal to the rainfall rate multiplied by half of the roof area. This means that the surface water being added to the cells immediately adjacent to the house is the sum of the rainfall that would fall in that cell plus the intercepted rainfall discharged from the roof. The same calculation is used for the mono pitch roof, but in this case the rainwater that falls on the roof is entirely discharged downslope or upslope of the house and it is equal to the rainfall rate multiplied by the whole roof area. The

surface water will then infiltrate into the slope as described for the case of the leaking pipe. If gutters are present, the rainwater intercepted by the roof is deleted, consequently decreasing the rainfall rate infiltrating into the slope.

Leaking pipes and buried tanks can induce soil pipe erosion in response to increasing water inputs. This could be simulated for example with a dual permeability model, but then it would be difficult to implement the pore pressure calculated into the slope stability model (Bogaard and Greco, 2016). Furthermore, the inclusion of preferential flows requires the definition of additional input factors which may be difficult in data-scarce contexts. So, given the spatial scale, the purpose of the analysis and the data available, the current CHASM+ representation can be considered sufficient to depict landslide initiation due to flow accumulation around the point water source.

S1.2 Benchmarking the new slope water point source functions in CHASM+

To benchmark the new functionality introduced, we compare the results obtained with CHASM+ with an example found in the literature. Mendes et al. (2018) analysed the natural and anthropogenic drivers of a rainfall-triggered landslide event happened in the city of Sao Jose' dos Campos (Brazil) on March 5th, 2016. Their analysis demonstrated that rainfall could only have initiated the observed landslide if combined with water tank leakage at the top of the slope. Mendes et al. (2018) used the Seep/w and Slope/w modules of the Geo-Slope software to analyse the hydrology and the stability component of the landslide. Rainfall records for the time of the landslide were available from nearby weather stations and were used to reproduce a daily accumulated rainfall graph for the 31 days prior to the landslide occurrence. These 31 days were used in the simulation to predict the landslide. The soil properties of the three soil layers used for the analysis are reported Table S1. Their characterisation was based on in situ inspection and on previous studies carried out in the same location (Mendes, 2014 and Mendes and Filho, 2015). The soil water retention curves (SWRC) and the conductivity functions are estimated from this data. The slope was 55 m high with an average slope angle of 40 degrees. The boundary conditions were set according to field observations and to considerations made by Rahardjo et al., (2007) in regarding the position of the water table and retaining walls at the bottom of the slope.

Table S1: geotechnical and hydrological characterisation of the soil layers of the failed slope from Mendes et al. (2018), re-adapted according to the unit of measure used in this analysis

Layer in soil profile	Depth sample (m)	USCS*	Bulk specific weight (kN m^{-3})	Effective cohesion (kPa)	Effective internal friction angle (°)	Hydraulic conductivity (m s^{-1})	Initial pore water pressure (m)
Soil 1	0.5	SM	17	10	33	2.8 e-6	-1
Soil 2	3.0	CL-ML	18	15	35	1.15 e-6	-1.5
Soil 3	6.5	SM	19	21	37	1.3 e-5	-2

*USCS: Unified Soil Classification System

The observed landslide occurred in a 6 m high cut slope, at the bottom of the hillslope. A water tank of 1000 litres capacity (10 kN m^{-2}) was found to be leaking just above the cut slope. Since the leakage rate at the moment of the failure was unknown, a linearly increasing leak of 0.5, 1.0 and 1.5 m^3 per day was assumed from day 16 until day 31. The Factor of Safety (FS) was then calculated for three cases: i) including just the rainfall, ii) including just the leaking tank, and iii) including both. The analysis by Mendes at al. (2018) demonstrated how their modelling approach predicts failure (FS<1) for the condition where the rainfall is combined with the leaking tank, but not for rainfall alone.

We want to emulate the above analysis with the new extended version of CHASM for the cases: i) including just rainfall and case iii) including both rainfall and the water leakage. With this aim, the analysis entails the following steps:

1) CHASM+ is compared to the GeoSlope models to evaluate how the two modelling approaches differ in the process representation and input factor specification. Some of the input factors not specified in Mendes et al. (2018) but necessary to run CHASM+ will be assumed.
2) CHASM+ is run using both the input factors specified in Mendes et al. (2018) and the input factors assumed in step 1. The results obtained in this (deterministic) simulation are compared to the results presented in Mendes et al. (2018)
3) CHASM+ is run stochastically, where the input factors specified in Mendes et al. (2018) are fixed and the input factors assumed are stochastically varied within reasonable ranges. This allows to take into account the uncertainties introduced by the different input factors specifications.

1) Comparing CHASM+ with the GeoSlope model(s)

The two models CHASM+ and GeoSlope present similarities and differences with respect to their process representation and in their specification and implementation of the input factors. Both models are based on limit equilibrium method of slices; they can represent unsaturated and saturated soil conditions using the Darcy equations; and they allow to define a grid of slip surface centres to analyse trial slips with different minimum factor of safety. GeoSlope operates on finite elements meshes for computing soil stresses with two-dimensional seepage. CHASM+ employs a forward explicit finite difference method to analyse the effective stresses at each computational node, with two-dimensional seepage on saturated soil conditions and one dimensional seepage on unsaturated soil conditions. Table S2 reports the differences in the governing equations and input factors specifications. The input factors not specified in Mendes et al. (2018) but necessary to run CHASM+ are assumed. These assumed values are fixed for the deterministic analysis (step 2) and varied within ranges for the stochastic analysis (step 3).

Table S2: differences between GeoSlope and CHASM+. The table specifies both the input factors used for the first deterministic comparison (step 2) and the space of variability for the input factors in the stochastic analysis (step 3)

	GeoSlope	CHASM+	Assumed values for the deterministic analysis	Assumed ranges for the stochastic analysis
Initial suction	Assigned per each cell with different values per soil type	Assigned only at the top cell and interpolated linearly until reaching the water table (where suction = 0 m)	-2 m at the top cell	U (-5; -0.5) m
Soil water retention curves	Specified by the parameters derived from lab tests	Use Van Genuchten model to calculate the soil water retention curve.	Hodnett and Tomasella (2002) Soil 1 = sandy clay loam Soil 2 = silty clay Soil 3 = loam	Varied according to the standard deviation suggested by Hodnett and Tomasella (2002) *
Unsaturated Hydraulic conductivity	Calculated with Van Genuchten interpolation	Calculated with the Millington-Quirk equation (Millington and Quirk, 1959)		
Unit weight	Only bulk specific weight specified	Need to specify both dry and saturated unit weight	Dry unit weight specified. Saturated unit weight = (dry unit weight + 2) kN m^{-3}	Dry unit weight: Soil 1 = Soil 2 = Soil 3 U (12; 24) kN m^{-3}
Impermeable surfaces	Applied impermeable barriers (software available option)	Obtained by decreasing the soil permeability of the cells occupied by the tank and walls	10e-13 m s^{-1}	Ln (-11.654 0.898) m s^{-1}
Tank leaking	Linearly increasing	Constant throughout the simulation time	Function modified to reproduce the same linear increment in the water leakage	As in the deterministic analysis

U = Uniform distribution; Ud = Discrete uniform; N = Normal distribution; Ln = Log-normal distribution.
VG: Van Genuchten parameters for defining suction moisture characteristics curve.
*Probability distributions assumed: N (Saturated water content – θsat) $m^3 m^{-3}$; Ln (Residual water content – θres) $m^3 m^{-3}$; Ln (VG α parameter) m^{-1}; Ln (VG n parameter)

2) Deterministic analysis: CHASM+ predicts lower slope stability than GeoSlope

The slope presented in Mendes et al. (2018) is reproduced in CHASM+, maintaining the same geometry, initial hydrological conditions, leak rate from the water tank and daily accumulated rainfall specified in the paper (the 31 days prior to the landslide occurrence). We use the soil properties reported in Table S1, and Table S2 (column: 'assumed values for the deterministic analysis'). The Factor of Safety (FS) predicted by GeoSlope and CHASM+ under these conditions, are presented in Fig. S1 (for CHASM+ only case (iii) is shown). Both the models predict an early failure: in GeoSlope, the FS falls below 1 the 26th day (5 days before the landslide occurrence) while CHASM+ predicts failure the 21st day (10 days before). Furthermore, the FS calculated with CHASM+ appears to be lower than the FS calculated with GeoSlope throughout the whole simulation time. This might indicate that the assumed input factors used in CHASM+ or/and the different numerical implementation could have led to a different hydrological and stability response. We therefore use a stochastic framework to perform a back analysis that explores which combination of input factors allow CHASM+ to give similar results to GeoSlope, and if this combination is physically consistent with the observed landslide event and data.

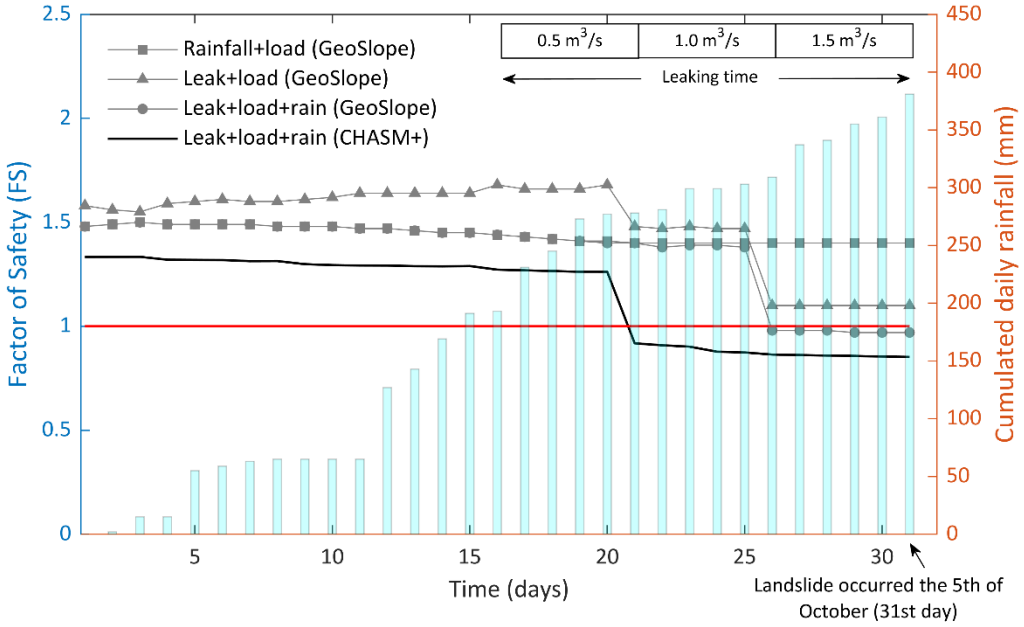


Figure S10: In grey the results obtained by Mendes et al. using GeoSlope for the three cases analysed (with and without rain and leaking tank); in black the results obtained with CHASM+ for the case where rainfall, leakage and load are considered. The light blue bars represent the cumulated rainfall of the 31 days preceding the landslide.

3) Stochastic analysis: CHASM+ presents consistent results with GeoSlope

To explore what it could have led CHASM+ to have a different stability response in step 1, the input factors specified in Mendes et al. are kept fixed while the input factors assumed are varied within reasonable ranges. The fixed factors are: the slope geometry; the tank leakage and load; rain frequency and intensity; and the soil properties of Table S1 and not part of Table S2. SWR curves, initial soil suction, soil unit weight and the hydraulic conductivity representing impermeable surfaces are varied according to the ranges specified in Table S2 (last column). 10 000 different combinations of these input factors are

created by stochastically sampling from those ranges (for a description of the stochastic method used, refer to the Methodology Section of the main manuscript).

Figure S2 shows the comparison between the simulations obtained with CHASM+ and with GeoSlope when only rainfall is considered. The values of input factors that create an hydrological response not compatible with the initial conditions used in Mendes et al (2018) are identified through sensitivity analysis (for example, suction values corresponding to levels of initial water table higher than Mendes et al., leading to an early failure – light grey lines in the figure). Only the simulations that do not use these values are considered, and they are called "ok simulations". The best performing 10% CHASM+ simulations are identified by comparing the Root Mean Square Error (RMSE) between the FSs obtained with CHASM+ (dark grey lines) and the FS obtained with GeoSlope (black line).

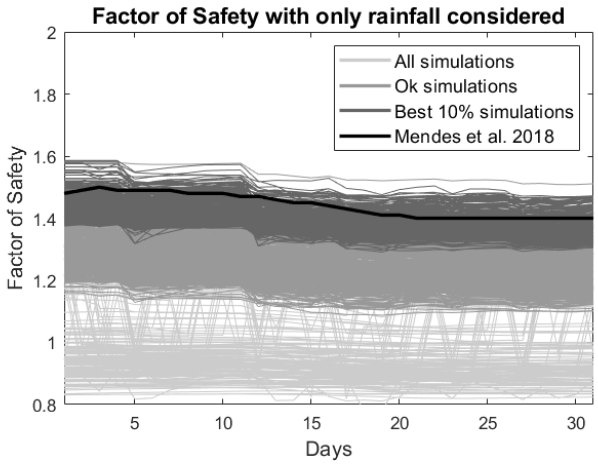


Figure S2: Factor of Safety calculated considering only rainfall and load of the tank for both CHASM+ and GeoSlope (this latter is referred as Mendes et al. 2018, and it corresponds to the grey line with square markers in Fig. S1).

Figure S3 shows how the ranges of the varied input factors are differently constrained when obtaining the "ok" and the best performing simulations. The bars represent the ranges of the input factors. If the bars reduce in size, part of the values of the given range has not been used to create the corresponding response. For example, the best performing simulations never use values of initial soil suction equal to -1 m (Fig. S3b). The black horizontal lines represent the values used by Mendes et al. 2018 (present only for the upper plots). These values are amongst those used to produce the best performing simulations in CHASM+ (i.e. they are within the dark grey bars), except for the initial soil suction. CHASM+ performs best with low saturated hydraulic conductivity values when representing impermeable surfaces which is physically consistent (the value used in Mendes et al. 2018 is assumed to be equal to 0 m s^{-1} for impermeable surfaces, Fig. S3d), and with low values of the Van Genuchten (VG) parameters defining the SWR curves (saturated moisture content θ_{sat}, residual moisture content θ_{res}, and parameters n, α, Fig. S3c,e,f,g). Low values of the VG parameters correspond to steeper SWR curves, a preferred condition for the hydrological numerical stability in CHASM+. The initial soil suction values used to obtain the best performing simulations ranges between -5 m to -2 m (Fig. S3b). These values are lower than those used in GeoSlope. The difference is due by the assumptions governing the initial water content distribution in CHASM+, which is determined by the cell resolution 1x1 m of the slope, and by the suction gradient. In the first time step, the initial suction, defined at the top cells of the slope, linearly decreases until reaching the water table. The matric suction for each cell is therefore calculated by dividing the surface suction into the number of cells above the water table. When the initial suction is low (i.e. closer to 0) and the SWR are smooth (i.e. with little changes of saturated water content for different suction values), more cells at the proximity of the water table result close to saturation, and the water level can increase up to 5 - 6 meters. High water table heights can intersect the cut slope and lead to an early failure (Fig. S1 and the light grey lines in Fig. S2). The uniform suction gradient assumed in CHASM+ is physically unrealistic, but it is used for the initial distribution of water moisture content across cells. The hydrological equilibrium is then regulated by the Richard's equation for unsaturated soil. However, this assumption leads to a

different initial hydrological condition when compared to GeoSlope. High values of initial suctions are therefore necessary to maintain the water table levels in CHASM+ in the same position simulated in GeoSlope.

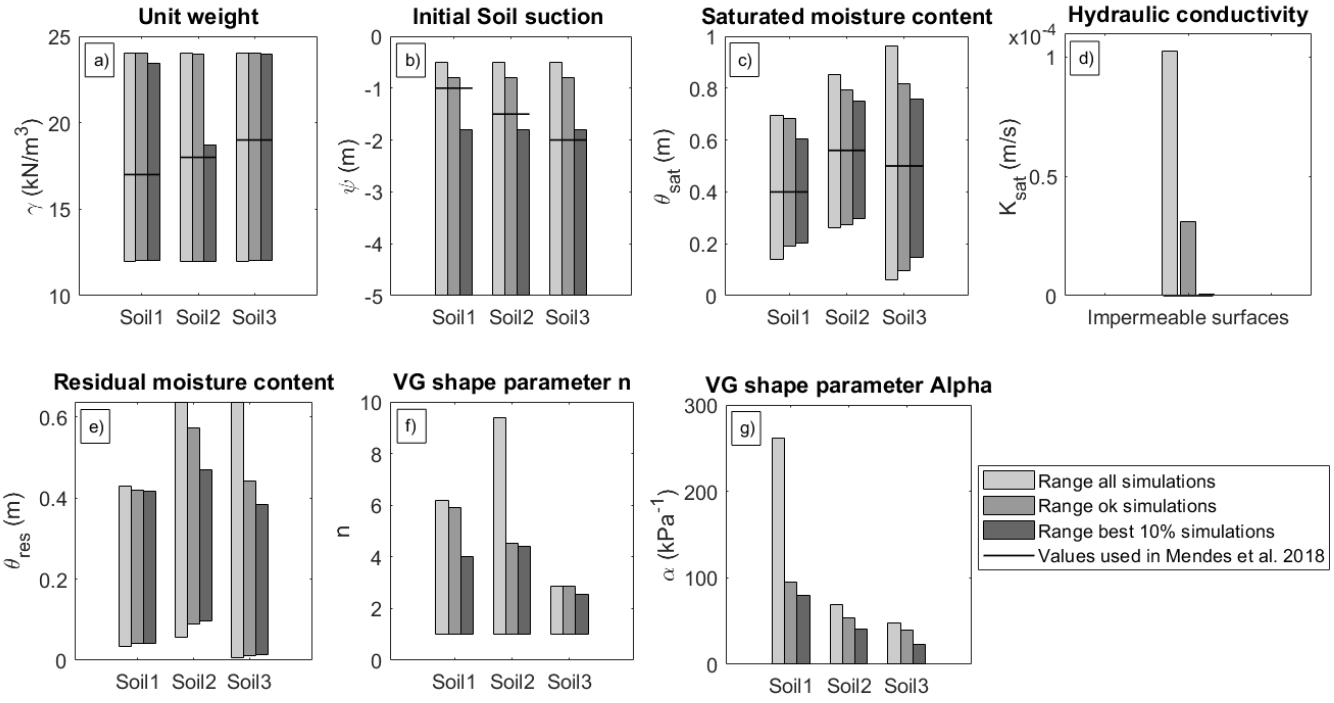


Figure S3: The input factors used in CHASM+ are varied within the ranges defined in Table S2. The whole ranges are represented in this figure as "range all simulations" (light grey bars). These are compared to the ranges of values that produce ok and best performing simulations (darker greys) which correspond to the FS trends shown in Fig. S1 in the same colour. The black horizontal lines reported on the upper plots (a,b,c,d) identify the discrete values used by Mendes et al. 2018 in the GeoSlope analysis (see Table S1). Note as in plot (d) the black line is at zero level.

For the second case the leaking tank is also considered. Other 10 000 simulations are created by sampling from the ranges previously identified as those producing the best performing simulations. Figure S4a shows the calculated FSs. This time, the number of ok simulations differ from the total number of simulations of just 4%. This is because the values of the input factors that were not compatible with the assumptions of the model (i.e. initial soil suction set too low) were excluded in the initial ranges. Amongst the ok simulations, CHASM+ predicts slope failures (FS<1) for a variety of different times (from Day 17 to Day 31). We want to explore which are the combinations of input factors that produce a most similar response to GeoSlope (dark grey lines in Fig. S4a, i.e. best 10% performing simulations). The parallel plots in Fig. S4b show the distribution of the input factors within their variability range. Ranges are standardised to allow for comparison across the factors. Each line corresponds to a simulation. The darker lines identify the combinations of input factors corresponding to the 10% best performing simulations and thus to the "correct" timing of the failure. If the dark lines concentrate in a subrange, that factor is influencing the distinction between ok and best performing simulations. This is evident for the VG alpha parameter for the three soil types, and the hydraulic conductivity of the cells representing the impermeable surfaces (IS). Values of hydraulic conductivity close to 0 are consistent with the representation of impermeable surfaces. Low values of alpha correspond to steep SWR curves. The other VG parameters counterbalance their effect to obtain the same result (van Genuchten, 1980). For example, when the saturated water content of soil 2 is high, the corresponding residual water content is low. Steep SWR curves means that the water content of the soil increases slower with the decrease of soil suction. This explains their influence on the timing of the failure. The values used by Mendes et al. (black dots) are all part of the lines that corresponds to the best performing simulations of CHASM+ and therefore they are values used to create similar responses to GeoSlope. Furthermore, with these combinations of input factors, CHASM+ predicts the same failure position as GeoSlope (not shown). We have therefore demonstrated that using the sets of input factors identified as best performing, we can create similar responses to

GeoSlope, a widely used dynamic slope hydrology and stability software. We use this analysis as an evidence that CHASM+ can correctly represent leakages from buried tanks.

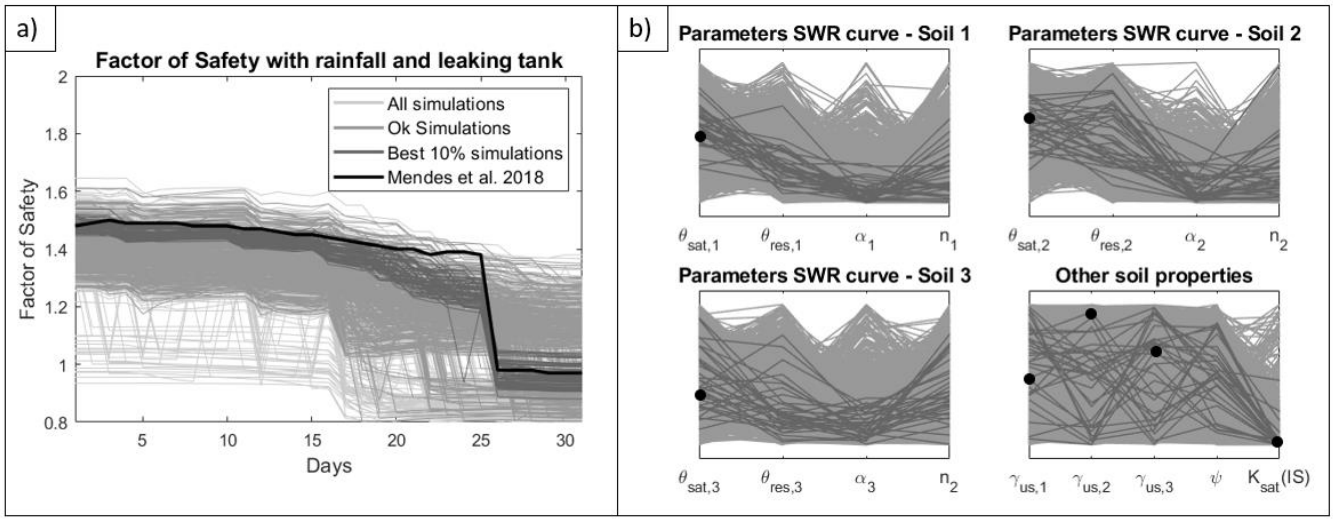


Figure S4: a) shows Factor of Safety calculated considering both rainfall and the leaking tank with CHASM+ and GeoSlope (referred as Mendes et al. 2018); b) shows the parallel plots of the ok (lighter grey) and best performing (darker grey) simulations. The lines identify how the input factors are distributed within their variability ranges. The black dots are the values used by Mendes et al. (2018)

S2 CART performance without auxiliary variables

Figure S5 shows the percentage of misclassified simulations (i.e. the cross-validation error) for different pruning levels for the non-urbanised (a) and the urbanised case (b) when auxiliary variables are not considered. In these cases, the minimum validation error is obtained for pruning level 123 and 143 respectively ("absolute minimum in red"), which correspond to trees with 219 and 269 nodes. The arrows in the figures point to the pruning levels used to construct the CARTs with auxiliary variables (A.V.), shown in Fig. 8a and 8b of the main manuscript. If the auxiliary variables were not considered, the misclassification errors at these pruning levels would be respectively 16.7% and 17.6%, instead of 13.4% and 14.4% (as shown in Fig. S7).

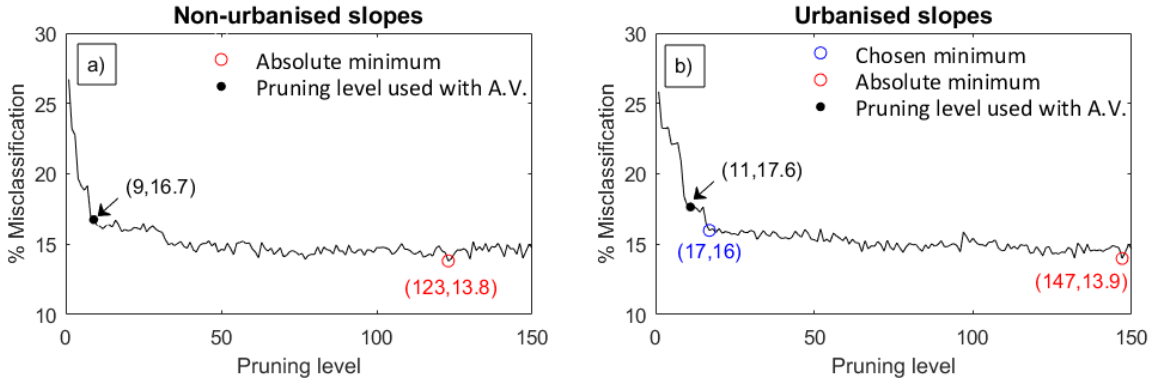


Figure S5: Cross-validation error of the CART for increasing pruning level for non-urbanised (a) and urbanised (b) slopes. The cross-validation error is computed by randomly dividing the dataset in 10 subgroups. Ten trees are then constructed by using 9 subgroups as training set. The excluded subgroup is used to calculate the misclassification error (in percentage). The average value of the ten misclassification errors so obtained gives the cross-validation error (at given pruning level). The "chosen minimum" (in blue) represents the pruning level and corresponding misclassification error to build the CART in Fig. S6; the pruning level used to build the trees reported in the paper (Fig8 a,b) and the corresponding misclassification error resulted without considering auxiliary variables (A.V.) are reported in black.

Figure S6 shows the CART obtained without considering the auxiliary variables (pruning level 17 and 16% misclassification error - "Chosen minimum" in Fig. S5b). The thickest branches of the tree show for which critical thresholds of the input factors the majority of simulated slopes failed (black branch) or did not fail (grey branch). The majority of failed simulations in this

case, occur for values of effective cohesion of layer 1 less than 12.4 kPa, rainfall intensities greater than 32.7 mm h^{-1}, thicknesses of layer 1 (residual soil) more than 1.9 m, and rainfall durations greater than 5 h.

Almeida et al. (2017) showed how cohesion and thickness of layer 1 as well as rainfall intensity and duration interact to produce slope failures. Two auxiliary variables were introduced: the ratio between effective cohesion and thickness of layer 1 and the negative ratio between the logarithm of rainfall intensity and rainfall duration. The misclassification error was similar (11%) with and without auxiliary variables, but the resulting trees had a much simpler structure. In this analysis, the misclassification error decreases of 1.91% (from 17.64% to 15.73% at pruning level 11) when these two auxiliary variables are considered.

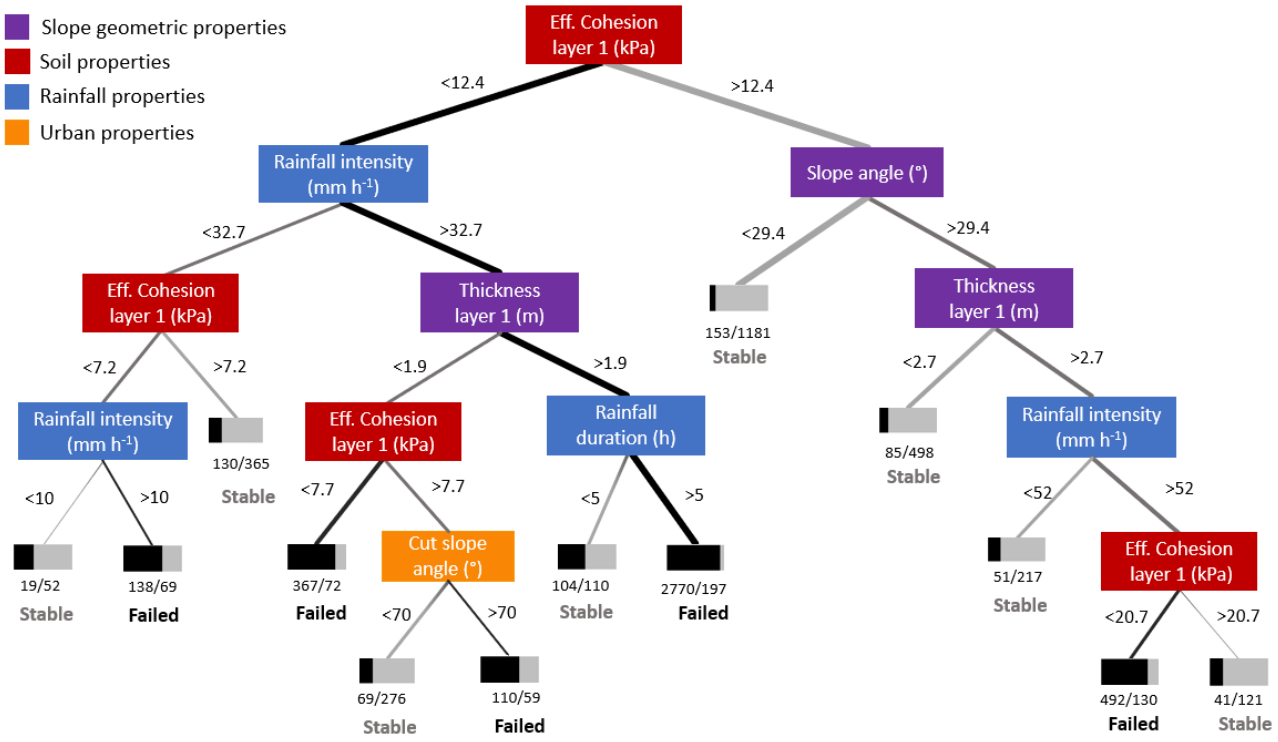


Figure S6: CART tree obtained for urbanised slopes without considering auxiliary variables. Black branches represent the paths that lead to simulations predicted as failed, while grey branches lead to simulations predicted as stable. The bar under each leaf shows the proportion of simulations that resulted as failed (black) or stable (grey) for that leaf. The thickness of the branch is proportional to the number of simulations following that path. The pruning level used is 17, with 17.6% simulations misclassified (Fig. S5b).

We introduce a third auxiliary variable: a weighted average of the natural and the cut slope angles (Eq. S2). The weights are represented by the sensitivity indices reported in Fig. 6 of the paper (w1 = 0.15 for slope angle; w2 = 0.13 for cut slope angle)

$$\text{Weighted Slope Angle} = \frac{w1*(\text{Slope angle})+w2*(\text{Cut slope angle})}{w1+w2} \quad \text{(S2)}$$

Weighted slope angles consider that slope susceptibility can significantly increase for low natural slope angles but high cut slopes angles. We use the sensitivity indexes as weights to reflect that the natural slope angles resulted more influential than cut slope angles. An averaged sum of the two input factors would result from equal weights. In this last case, the reduction in misclassification error would be 0.3%. When the sensitivity indices are considered as weights, the reduction increases to 1.3% (from 15.73% found introducing the first two auxiliary variables to 14.4%). The weighted slope angle presented in Eq. (S7) is therefore better performing and it is used for the CART analysis.

S3 CART pruning

We use cross-validation to avoid overfitting. Figure S7 shows the percentage of misclassified simulations (i.e. the cross-validation error) for different pruning levels for the not urbanised and the urbanised case. The minimum validation error is

obtained for pruning level 83 and 69 respectively ("absolute minimum" in red), which correspond to trees with 145 and 116 nodes. We choose much simpler trees with pruning level 9 and 11 ("chosen minimum"). These correspond to cross-validation error of 13.4% and 14.4% respectively for the two cases.

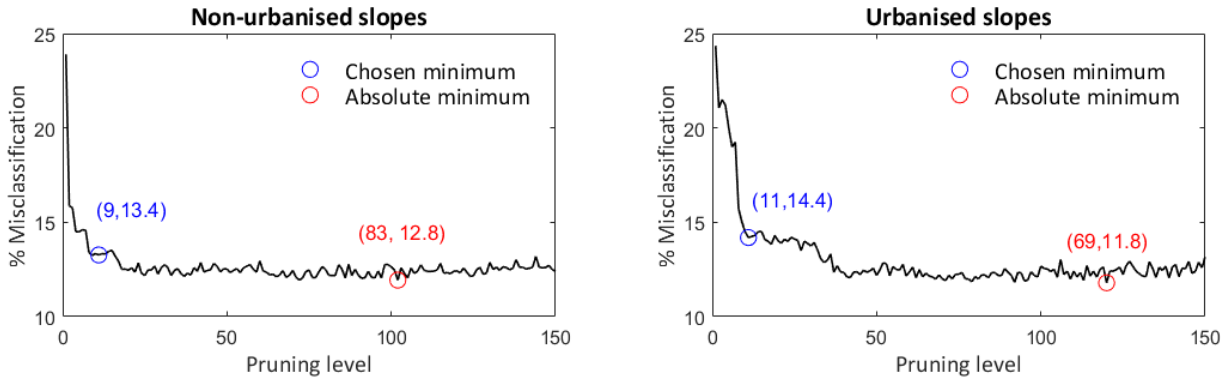


Figure S7: Cross-validation error of the CART for increasing pruning level. The cross-validation error is computed by randomly dividing the dataset in 10 subgroups. Ten trees are then constructed by using 9 subgroups as training set. The excluded subgroup is used to calculate the misclassification error (in percentage). The average value of the ten misclassification errors so obtained gives the cross-validation error (at given pruning level).

S4 Calculation of the rainfall threshold by multi-objective optimisation

Figure S8 shows the slopes simulated as failed (black) and stable (grey), plotted on log-log axes of associated rainfall intensities (I) and durations (D). The plot shows a descending trend according to which landslides are more likely to occur for high-intensity short-durations rainfall events, and for long-duration low-intensity rainfall events. This relationship is observed in landslide inventories and it is widely used to generate rainfall empirical thresholds for landslides prediction and landslide warning systems (see Segoni et al., 2018, for a review on the topic). Intensity duration thresholds are the most common type of thresholds that can be found in literature (Guzzetti et al., 2007), and they identify the intensity-duration combinations below which landslides are not expected to occur. Intensity duration thresholds are generally expressed by a power low $I = \gamma D^{\alpha}$ (Guzzetti et al., 2007) which in logarithmic axis becomes:

$$\log_{10}(I) = \gamma - \alpha \log_{10}(D) \quad \text{(S3)}$$

i.e. a linear equation where γ (the intercept) and α (the slope) are parameters specific to the site considered.

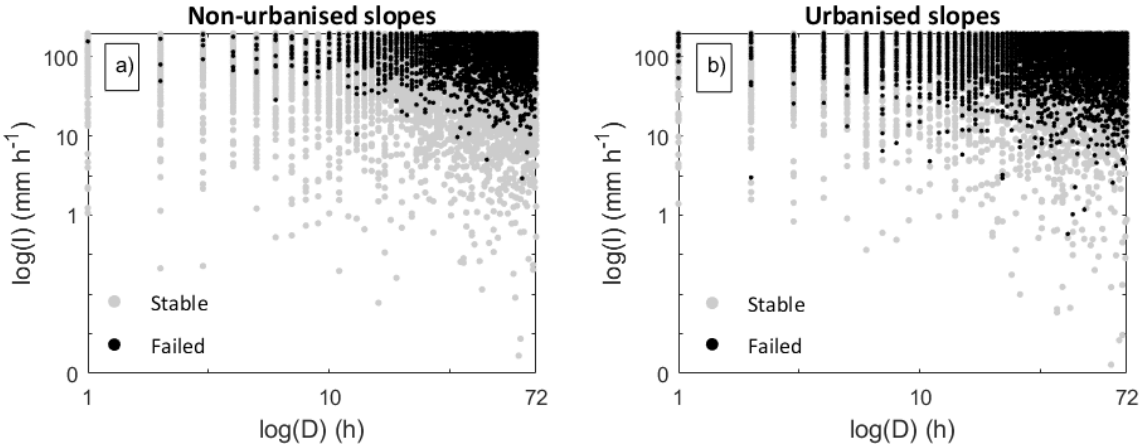


Figure S8: Combinations of rainfall intensities and durations resulted into stable (grey dots) or failed (black dots) slopes, for the non-urbanised (a) and urbanised (b) case. The plots show how the recorded landslides follow the typical descending trend found in empirical rainfall thresholds. The x and y axis are in logarithmic base 10, but the notation is linear for an easier readability.

To formalise the threshold that best divide failed from stable slopes, the parameters γ and α of Eq. S3 need to be evaluated. Different methods have been suggested to calculate these two parameters (see Table 3 in Segoni et al. 2018). Amongst these, statistical methods are widely employed because they provide objective and reproducible results (Brunetti et al., 2010; Perruccacci et al. 2012; Staley, 2013; Melillo et al. 2015; Piciullo et al. 2017; Perruccacci et al. 2017; Melillo et al. 2018). Frequentist methods showed to give satisfactory results for large datasets and allowed the definition of multiple minimum thresholds based on different exceedance levels (Brunetti et al. 2010; Perruccacci et al. 2012; Melillo et al. 2018). This can be useful in setting different landslide warning levels, each based on different probability of landslides occurrence. However, frequentist methods result unsuitable for analysing our synthetic dataset because of the high frequency of slopes failed for high intensity and high duration events (which are usually not recorded in reality) would strongly bias the position of the threshold. We therefore suggest a new approach that employs:

- the combinations of rainfall intensity and durations resulted in landslides (black dots in Fig. S8)
- a multi-objective optimisation algorithm for the estimation of the two parameters γ and α of Eq. S3.

The multi-objective optimisation involves minimising or maximising multiple objective functions subject to a set of constrains. In this case, we want to draw a threshold line in the form of Eq. S3 which identifies the space where landslides are recorded. This translates into choosing parameters γ and α of Eq. S3 that satisfy the following two contrasting objectives:

1) maximise the number of (simulated) failed slopes falling above the threshold line (Fig. S9a)
2) minimise the area above the threshold line (Fig. S9b)

To constrain the search to realistic values of rainfall intensity and duration, the optimisation only explores values of γ and α within upper and lower boundaries specified as:

$$\gamma\ [-0.5; -2] \quad \text{(S4a)}$$

$$\alpha\ [0.05; 2] \quad \text{(S4b)}$$

The range of α so defined includes typical slope values of empirical rainfall thresholds (Guzzetti et al., 2007), while the range of γ is designed to include all the rainfall intensities simulated. To perform the multi-objective optimisation, we used the generic algorithm implemented in the "gamultiobj" function of the Matlab Optimisation Toolbox (R2018a). As any multi-objective optimiser, it produces a set of Pareto-optimal solutions that realise different optimal trade-offs of the two objectives. In this case, 13 possible optimal combinations of (γ, α) are obtained, and among them we (subjectively) chose the one that gives a threshold line with 99.9% of failed simulations above it or, in other words, with 0.1% landslide probability below it. This is the threshold line reported in Fig. 98a,b of the main manuscript. A different choice could be made to determine the threshold line for any exceedance probability level. An alternative to this approach could be to use a (single-objective) optimization based on ROC (receiver operating characteristics), where false positives and negatives (represented in this case by the simulated landslides below the threshold and simulated stable slopes above the threshold) are minimised (Gariano et al., 2015; Staley et al., 2013).

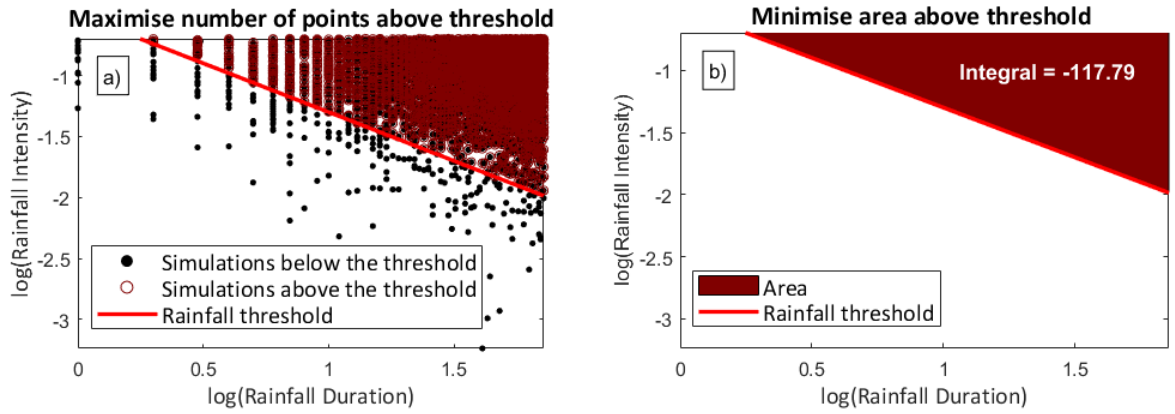


Figure S9: Illustration of the two objectives functions used in the optimisation, for a given threshold line: (a) maximise the number of failed slopes above the threshold and (b) minimise the area above the threshold.

References

Almeida, S., Ann Holcombe, E., Pianosi, F. and Wagener, T.: Dealing with deep uncertainties in landslide modelling for disaster risk reduction under climate change, Nat. Hazards Earth Syst. Sci., 17(2), 225–241, doi:10.5194/nhess-17-225-2017, 2017.

Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdiscip. Rev. Water, 3(3), 439–459, doi:10.1002/wat2.1126, 2016.Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D. and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10(3), 447–458, 2010.

Gariano, S. L., Brunetti, M. T., Iovine, G., Melillo, M., Peruccacci, S., Terranova, O., Vennari, C. and Guzzetti, F.: Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy, Geomorphology, 228, 653–665, doi:10.1016/j.geomorph.2014.10.019, 2015.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, , 267, 239–267, doi:10.1007/s00703-007-0262-7, 2007.

Hodnett, M. G. and Tomasella, J.: Marked differences between van Genuchten soil water-retention parameters for temperate and tropical soils : a new water-retention pedo-transfer functions developed for tropical soils, , 108, 155–180, 2002.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L. and Guzzetti, F.: An algorithm for the objective reconstruction of rainfall events responsible for landslides, Landslides, 12(2), 311–320, doi:10.1007/s10346-014-0471-3, 2015.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A. and Guzzetti, F.: Environmental Modelling & Software A tool for the automatic calculation of rainfall thresholds for landslide occurrence, Environ. Model. Softw., 105, 230–243, doi:10.1016/j.envsoft.2018.03.024, 2018.

Mendes, R.M.: Study of experimental fields monitoring geotechnical and climatic variables for the implementation of the warning systems for landslides in the State of Sao Paulo. Research Report (Grant number 11/22577-2). IOP Publishing Physics, available from: http://www.bv.fapesp.br/36971, 2014.

Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdiscip. Rev. Water, 3(3), 439–459, doi:10.1002/wat2.1126, 2016.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D. and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10(3), 447–458, 2010.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, , 267, 239–267, doi:10.1007/s00703-007-0262-7, 2007.

Mendes, R. M., de Andrade, M. R. M., Graminha, C. A., Prieto, C. C., de Ávila, F. F. and Camarinha, P. I. M.: Stability analysis on urban slopes: case study of an anthropogenic-induced landslide in São José dos Campos, Brazil, Geotech. Geol. Eng., 36(1), 599–610, 2018.

Millington, R. J. and Quirk, J. F.: Permeability of Porous Media, Nature, 183, 387–388, 1959.

Ortuste, F. R.: Living without Sanitary Sewers in Latin America: The Business of Collecting Fecal Sludge in Four Latin American Cities, World Bank, (March), 2–4 [online] Available from: https://openknowledge.worldbank.org/handle/10986/17332, 2012.

Rahardjo, H., Ong, T. H., Rezaur, R. B. and Leong, E. C.: Factors controlling instability of homogeneous soil slopes under rainfall, J. Geotech. Geoenvironmental Eng., 133(12), 1532–1543, 2007.

Segoni, S., Piciullo, L. and Gariano, S. L.: A review of the recent literature on rainfall thresholds for landslide occurrence, Landslides, 15(8), 1483–1501, doi:10.1007/s10346-018-0966-4, 2018.