Dealing with deep uncertainties in landslide modelling for disaster risk reduction under climate change

Susana Almeida[1], Elizabeth Holcombe[1], Francesca Pianosi[1], Thorsten Wagener[1,2]

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

Correspondence to: Susana Almeida (susana.almeida@bristol.ac.uk)

Abstract. Landslides have large negative economic and societal impacts, including loss of life and damage to infrastructure. Slope stability assessment is a vital tool for landslide risk management, but high levels of uncertainty often challenge its usefulness. Uncertainties are associated with the numerical model used to assess slope stability and its parameters, with the data characterising the geometric, geotechnic and hydrologic properties of the slope, and with hazard triggers (e.g., rainfall). Uncertainties associated with many of these factors are also likely to be exacerbated further by future climatic and socio-economic changes, such as increased urbanisation and resultant land use change. In this study, we illustrate how numerical models can be used to explore the uncertain factors that influence potential future landslide hazard using a bottom-up strategy. Specifically, we link the Combined Hydrology And Stability Model (CHASM) with sensitivity analysis and Classification And Regression Trees (CART) to identify critical thresholds in slope properties and climatic (rainfall) drivers that lead to slope failure. We apply our approach to a slope in the Caribbean, an area that is naturally susceptible to landslides due to a combination of high rainfall rates, steep slopes, and highly weathered residual soils. For this particular slope, we find that uncertainties regarding some slope properties (namely thickness and effective cohesion of top soil) are as important as the uncertainties related to future rainfall conditions. Furthermore, we show that 89% of the expected behaviour of the studied slope can be characterised based on only two variables – the ratio of top soil thickness to cohesion and the ratio of rainfall intensity to duration.

1 Introduction

Landslide hazards and their impacts on human lives, economies and infrastructure present a growing challenge globally Dai et al. (2002). Approximately 300 million people worldwide are estimated to be exposed to landslides (Dilley et al., 2005) with the greatest absolute economic losses experienced in Europe and the overwhelming majority of injuries and fatalities in Asia, South and Central America and the Caribbean (UNU, 2006; Petley, 2012). The susceptibility of slopes to landslides is known to increase with urban development activities such as vegetation removal, construction of earthworks (cuts and fills), buildings and roads, and altered drainage (Smyth and Royle, 2000). Climate change is likely to increase the occurrence of rainfall

triggered landslides, because a warmer atmosphere tends to produce more high intensity rainfall events (Ibsen and Brunsden, 1996; Borgatti and Soldati, 2010; Crozier, 2010; Kendon et al., 2014).

Landslide hazard assessment forms the basis for disaster risk reduction decisions such as the design of physical landslide hazard mitigation measures, planning controls and early warning systems for hazard avoidance, vulnerability reduction (resilience) approaches, or insurance. The spatial scale and the purpose of the hazard assessment, as well as the data available, determine which methods or slope stability models can be applied. Available assessment methods include inventory-based susceptibility mapping or regional forecasting, and statistical, heuristic and physically-based modelling (either spatially distributed or site specific) (Soeters and van Westen, 1996; Dai and Lee, 2001). These methods require data in some or all of the following categories: i) *inventories* of past landslide locations, types and triggers; ii) *preparatory factors* determining the inherent susceptibility of a slope to landslides, such as slope geometry (slope angles and heights, convergence/divergence, soil and parent material depths), the geotechnical and hydrological properties of slope material, and land-use; and iii) *triggering factors* such as rainfall intensity, duration and frequency.

The challenge of acquiring this data in sufficient quantity, quality and resolution is currently hindering the production of "actionable" landslide hazard information for decision-makers (Aitsi-Selmi et al., 2015). With the exception of particular regions such as Hong Kong, there is generally a lack of systematic landslide data collection even in the most landslide-prone countries (Corominas et al., 2014). Especially high frequency, low intensity events may be missed out due to their small spatial scale and impact. Existing databases are therefore often spatially and temporally biased or incomplete. Lack of data on past landslides limits our ability to build inventory-based and statistical models for predicting likely locations, timing or impacts of future landslides. In comparison, the advantage of physically-based slope stability models is their smaller reliance on observations of past events and an ability to mechanistically represent the preparatory and triggering processes driving slope failure. This latter characteristic also allows them to assess the impact of higher intensity rainfall events than a region might have experienced in the past, or of possible urbanisation scenarios. Landslide hazard researchers and civil engineers currently employ physically-based models to diagnose existing stability conditions, design new slopes and determine landslide probabilities.

The mechanistic ability of physically-based models comes at a cost. The detailed representation of slope processes in the model requires detailed information on slope properties such as soils and topography. The more complex, high resolution and comprehensive the representation of slope processes in the model, the more data regarding the physical site characteristics is required. Parsimonious models are therefore often selected that are consistent with data availability and with the required level of process representation. Even then, data is rarely available in sufficient detail, and this introduces uncertainty into the model parameterisation. Sources of uncertainty include those associated with slopes geometries and material strata depths (Lumb, 1975; Corominas et al., 2014), soil properties (Cho, 2007; Beven and Germann, 2013), and a limited understanding of how measured variables relate to model parameters (the commensurability issue - Beven, 1989; Wagener and Gupta, 2005; Beven, 2006). The lack in accuracy of forcing boundary conditions, such as the temporal and areal variability of historical rainfall, introduces further uncertainty that needs to be considered (Minder et al., 2009; von Ruette et al., 2014). Studies have assessed

the impact of uncertainties associated with such models (Hall et al., 2004; Arnone et al., 2016) for both site-specific and spatially distributed landslide hazard predictions – though such information is rarely used to support risk reduction decision-making.

Additional uncertainty is introduced by the poorly known potential implications of future climate or land use change. Such uncertainties are different from the ones listed above, because they cannot be easily characterised by probability distributions that different experts could agree on, i.e. they are often called *deep uncertainties* (Bankes, 2002; Lempert et al., 2003; Stein and Stein, 2013). The standard approach to dealing with deep uncertainties is through scenario-led strategies, also called top-down approaches. In these approaches, climate change projections of a general circulation model (GCM) are downscaled to derive local rainfall intensities and frequencies, as well as other climatic variables. These scenarios are then used to drive a slope stability model in a top-down manner (Collison et al., 2000; Melchiorre and Frattini, 2012). Studies that have attempted to quantify the uncertainty in the estimation of climate variables derived in this manner have found them to be prohibitively large in many cases (e.g., Collins et al., 2012; Ning et al., 2012). As a result, one usually finds that the uncertainty in the final predicted impacts is also quite large and that a wide range of possible outcomes is feasible, which is of little practical use for decision-making or for identifying an "optimal" management solution (Bankes, 2002; Wilby and Dessai, 2010; Hallegatte et al., 2012; Herman et al., 2014). To reduce the range of possible outcomes, a smaller subset of the many possible potential future scenarios can be selected. However, this approach is problematic as arbitrary selection of scenarios or of the downscaled simulations will undermine the credibility of the results (e.g., Kim et al., 2015).

Given that such deep uncertainties are unavoidable, a shift from "top-down" to "bottom-up" approaches has been suggested to derive actionable information for decision-makers (Groves and Lempert, 2007; Wilby and Dessai, 2010; Singh et al., 2014; Ray and Brown, 2015). While "top-down" approaches simulate system behaviour under potential future conditions in a predictive manner (e.g., to estimate probability of slope failure given one or more climate change scenarios), "bottom-up" approaches focus on exploring the vulnerabilities of the system, i.e. on finding those combinations of factors values that would produce unwanted outcomes (e.g., slope failure). Bottom-up approaches are therefore stakeholder driven since they start with the stakeholder who has to define what threshold separates acceptable from unwanted outcomes. A wide range of possible values of the uncertain factors can then be considered (i.e. propagated through the model) and mapped onto the regions of vulnerability of the output space. We can use statistical data-mining algorithms to quantify the link between inputs and outputs, i.e. the mapping stage. Commonly used algorithms to implement the mapping required for bottom-up approaches include Friedman and Fisher's (1999) Patient Rule Induction Method (PRIM) and Classification and Regression Trees (CART) developed by Breiman et al. (1984). The bottom-up strategy is very similar to the problem of mapping in Global Sensitivity Analysis, where one tries to understand which parts of the input factor space produce a particular model output, e.g. output values above a certain threshold (e.g., Saltelli et al., 2008; Pianosi et al., 2016).

In this study, we apply a bottom-up approach to a landslide hazard assessment model in order to address the following three questions:

1. Can we identify the dominant preparatory and triggering factors driving slope instability (i.e. slope geometry, geotechnical and hydrological parameters or rainfall drivers) in data scarce locations?

2. Can deep uncertainty in future landslide triggers, such as climate change, exceed other uncertainties?

3. What are the implications of uncertainty for data acquisition and assessment of future hazard?

2 Model and study site

We select a physically-based model and study site to represent the landslide hazard assessment and risk reduction challenges typical of data-scarce and resource limited locations (Fig. 1). The model is representative of those used by civil engineers and slope stability modellers. It can be parameterised using study site data and has a track-record of successful application in data-scarce locations. The methodology we develop around this case study is transferable to other locations and other types of physically-based slope stability assessment models.

2.1 An urban study site in the humid tropics

The selected study site is situated on the lower slope of a ravine on a hillside in the city of Castries, Saint Lucia (Eastern Caribbean), in which informal construction of houses has led to increased landslide risk. The slope section is approximately 50 m in height and has an overall angle of about 30°. The material strata comprise up to 6 m of residual soils overlying a similar depth of decomposed rock (weathering grades III-IV; after GEO, 1988) overlying volcanic bedrock. As is typical for the humid tropics, the dominant landslide trigger in the Caribbean is rainfall (Lumb, 1975; De Graff et al., 1989), and shallow rotational slides are the most common type of landslide in the deep weathered tropical residual soils on the mid to lower slopes of steep hillsides (Migon, 2010). Such locations tend to be the only land available to the most socio-economically vulnerable families, so even small landslide events can have a high societal impact (UNISDR, 2015). These "everyday disasters" are also increasingly seen as indicators of risk accumulation (low disaster resilience) and represent a potential hindrance to national economic development (Bull-Kamanga et al., 2003).

In 2011 a landslide hazard reduction project was implemented at the study site location using a community-based approach – Management of Slope Stability in Communities (Mossaic) – developed by Anderson and Holcombe (2006). At the time, estimates of geotechnical, hydrological and geometrical parameters for the study site were derived from a combination of: topographic maps, site surveys, interviews with residents to estimate soil strata depths and weathering grades (based on their experiences of excavating the soils to construct house foundations), elicitation of local engineering knowledge of soils, and information from shear box and ring infiltrometer testing of similar soils in Saint Lucia. Soil properties were also benchmarked against extensive triaxial and permeameter test data for similar undisturbed tropical residual soils in Hong Kong (GCO, 1982).

2.2 A physically based model for rainfall-triggered landslides

Deterministic physically-based modelling of slope stability has previously been carried out at the study site using the Combined Hydrology And slope Stability Model (CHASM) to diagnose landslide drivers and estimate the benefit:cost ratio of landslide mitigation (Holcombe et al., 2012). In a validation exercise in Hong Kong CHASM was shown to be numerically robust and

5 capable of correctly classifying 77% of failed slopes and 68% of stable slopes for a specified rainfall event (Anderson, 1990). CHASM has been extensively used by slope stability researchers and practitioners in Malaysia, Indonesia, the Eastern Caribbean, United Kingdom and New Zealand (Anderson et al., 1997; Lloyd et al., 2001; Wilkinson et al., 2002a; Wilkinson et al., 2002b)

A brief overview of CHASM is given here – full descriptions of the numerical scheme and principle equations can be found

10 in Anderson and Lloyd (1991) and Wilkinson et al. (2002b). CHASM represents the slope cross-section as a regular two-dimensional mesh of columns and cells with geotechnical and hydrological parameters specified for each soil type (Fig. 2). Initial hydrological conditions are the matric suctions in the top cells of each column and the water table position. Subsequent dynamic forcing conditions are rainfall events of specified intensities and durations imposed on the top cells. For each hydrological time-step (usually 10-60 seconds) a forward-explicit finite difference scheme is used to solve Richards' equation

15 (Richards, 1931) and Darcy's Law (Darcy, 1856) for rainfall infiltration, unsaturated and saturated groundwater flows. Cell moisture conditions, pressure heads and unsaturated hydraulic conductivities are updated at each time-step using soil moisture characteristic curves and the Millington-Quirk procedure (Millington and Quirk, 1959). At the end of each simulation hour the pressure head fields are used to calculate pore water pressures (positive and negative) for input to a two-dimensional Limit Equilibrium Method (LEM) calculation of slope stability. In LEM analysis the slope Factor of Safety (F) is calculated as the

20 ratio of destabilising forces to resisting forces for a potential landslide slip surface location, such that F<1 indicates failure. In CHASM either Bishop's simplified circular method of slices (Bishop, 1955) or Janbu's non-circular method (Janbu, 1954) is implemented using an automated search algorithm to identify the location of the slip surface with the minimum value of F for that hour.

3 Methods

25 Our study aims to advance understanding of the critical uncertainties driving rainfall-triggered landslides. To this end, the simulation model, CHASM, is run with 10000 different combinations of values for the uncertain input factors. Simulations are performed using the BlueCrystal Phase 3 high-performance cluster at the University of Bristol, which contains 16 x 2.6 GHz SandyBridge cores. Each simulation is classified as corresponding to stable or unstable slope based on the resulting slope Factor of Safety, F, (model output) being above or below 1 at any stage during the simulated time period. We first perform a

30 preliminary visual analysis of the simulations to identify influential factors that lead to slope failure. We then apply Classification And Regression Trees (CART) to formally determine combinations of factors and their thresholds values for

which the model predicts slope failure in a bottom-up strategy. The Matlab SAFE toolbox (Pianosi et al., 2015) and the CART functions in the Matlab Statistics and Machine Learning Toolbox (Mathworks, 2015) are used to perform our analysis.

3.1 Characterisation of uncertain input factors

The study site slope cross-section, as represented in CHASM, is illustrated in Fig. 2. Slope input factors (slope geometry, geotechnical and hydrological properties) are assumed to be random variables characterised by different statistical distributions. These distributions and their respective statistical parameters have been obtained from different sources as described in Sect. 2.1, and summarised in Table 1. For each set of CHASM input factors generated, checks are undertaken to ensure particular combinations of the input factor values selected are physically realistic (for example, that saturated unit weight is greater than unsaturated unit weight). Model discretisation parameters – such as the cell size (1m x 1m), hydrological time-step (60 seconds), and slip search grid location and dimensions – and physical and mathematical constants are not varied. The dynamic hydrological scheme of CHASM requires the specification of hourly rainfall intensities to drive the dynamic hydrological component of CHASM for the selected hydrological time-step (60 seconds). While the uncertainties in slope properties are characterised by probability distributions based on past experience with the model and the study area, uncertainties related to potential future rainfall are deeply uncertain. We therefore represent our lack of knowledge by varying rainfall intensity-duration combinations widely to ensure that any feasible future design storm in a changing climate is captured in our sampling. The ranges of rainfall intensity and duration used in this analysis are based on Intensity-Duration-Frequency (IDF) relationships derived for the design of the Roseau Dam in Saint Lucia (Fig. 3). Engineering consultants, Klohn-Crippen (1995), applied Gumbel analysis of 40-years of daily rainfall data from weather stations around the island to estimate the intensities and durations of 1:5 to 1:500-year return periods events. From these IDF relationships we define ranges of possible rainfall intensities of 0 to 200 mm.h^{-1} and durations of 0 to 72 h, which we sample independently and uniformly.

For this experiment the first 168 hours of rainfall forcing are set to an intensity of 0 mm.h^{-1} so that water table, cell moisture conditions and slope stability (indicated by F) can reach states of steady seepage and equilibrium. Rainfall of the selected intensity and duration is then imposed on the slope to determine the stability of the slope for that storm event. In the absence of information on typical hourly rainfall rates for events longer than 1 hour, rainfall intensity is assumed to be uniform across the rainfall duration sampled. The storm is followed by a further 168 hours of zero-rainfall simulation time to consider the continued effects of the groundwater response on slope stability.

3.2 Classification And Regression Trees (CART)

While different algorithms have been used to implement the mapping step in bottom-up approaches in past studies, we have chosen to use Classification And Regression Trees (CART). Comparison between the most popular algorithms, CART and PRIM, did not show either algorithm to be superior (Lempert et al., 2008), while CART hast he advantages of simplicity and an ability to work with minimal input from the user. CART is a machine-learning method for constructing prediction models from data (Breiman et al., 1984). In our application, such model takes the form of a binary tree, where a categorical dependent

outcome (i.e. slope failure versus stability) is predicted from a set of continuous independent variables (i.e. slope and design storm properties). The tree is composed of nodes and branches. At each node, an if-then condition is applied to one of the independent variables (e.g. "slope angle above or below 30 degrees") to generate two different branches. Several strata of nodes and branches compose the tree up to a stratum of terminal nodes (leaf nodes) where a prediction of the categorical

5   dependent outcome is made (e.g., "slope fails").

A CART is constructed through a recursive algorithm applied to a sample of independent variables (inputs) and associated categorical outcomes (outputs). In this study, the input/output sample is generated by the Monte Carlo simulation of CHASM. The algorithm automatically selects which input variable and threshold value to use at every node. The selection is based on maximising the homogeneity of the output samples in the subsequent nodes. The level of homogeneity can be measured by

10   different criteria. Here, we use the Gini impurity measure, which is defined in Eq. (1) as:

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

where, m is the number of categories for the output (2 in our case), and $p(i)$ is the fraction of output samples in the node belonging to category i. A Gini impurity index equal to 0 corresponds to a pure node, i.e. a node where all output samples belong to the same category. In general, pure nodes cannot be obtained, but the algorithm seeks to minimise the Gini impurity

15   index.

A key question in constructing a classification tree is that of the optimal sizing (i.e. number of nodes and branches). Increasing the size of the tree generally increases its predictive accuracy over the sample used for its construction, but it might reduce its ability to generalise to new samples (overfitting). Moreover, a tree with simpler structure might be easier to interpret and communicate. Once a CART has been constructed, a machine learning technique called "pruning" can be applied to reduce

20   the size of the tree by removing sections that provide little classification power. More details about the pruning technique can be found, for example, in Hastie et al. (2009).

4 Results

In this section, we analyse the outputs from the 10000 CHASM simulations to determine which factors control slope stability and whether critical thresholds exist beyond which slope failure will occur.

25   We perform an initial evaluation of the factors controlling slope stability by comparing the marginal distributions of the input factors that cause slope failure with those that do not. The approach is generally referred to as Regional Sensitivity Analysis (first proposed by Spear and Hornberger, 1980; for a general introduction see Pianosi et al., 2016). We split the 10000 model simulations into two sub-sets: those that produce slope failure (F<1) and those that simulate a stable slope (F>1). For each input factor, the (marginal) empirical cumulative distribution functions of the two sub-sets are computed. Figure 4 shows these

30   distributions for a selection of factors. The grey dashed line is the distribution corresponding to a stable slope and the black solid line is the one corresponding to slope failure. The grey and black lines clearly deviate (i.e. the two distributions are

different) for thickness of top soil, effective cohesion of top soil, rainfall intensity and rainfall duration, implying that these factors are important determinants of slope failure. For all other factors (including those not shown in Fig. 4), the two lines essentially overlap, implying that the factors have no direct influence on determining failure (although they still may have influence through interactions, see Saltelli et al., 2008). Besides identifying the influential factors, Fig. 4 also provides

5   information on the ranges of influential factors associated with slope failure. For example, the top left panel shows that the model is unlikely to predict failure when thickness of top soil is smaller than 3 m, as the black solid line is nearly flat towards the smallest values of the original range of this factor. This initial analysis suggests that relatively few input factors have a significant impact on slope stability. Furthermore, it also provides initial guidance about which values of those factors are more likely to produce failure.

10   CART analysis provides a systematic way to quantify the thresholds separating slope failure and stability. Figure 5 shows the classification tree obtained from the same set of simulations (see Appendix A for details on pruning). CART results are consistent with the initial analysis of input distributions, as evidenced by the fact that the main factors emerging in the classification tree of Fig. 5 are the same that were shown as influential in Fig. 4. From the 28 analysed factors (26 model parameters plus the two design storm properties), five factors alone (thickness of top soil, effective cohesion of top soil, rainfall

15   intensity, rainfall duration and initial depth of water table) are sufficient to correctly classify 89% of the simulations. Figure 5 also illuminates the critical thresholds in slope properties and rainfall drivers that separate slope failure and stability. For example, the leftmost branch shows that if the thickness of the top soil is less than 2.9 m and effective cohesion of this stratum is below 2.1 kPa, the landslide model tends to predict slope failure. This happens regardless of what values rainfall intensity and duration take. The black/grey-shaded bar at the end of each node visualises the fraction of input factor combinations that

20   produce failure/stability respectively. It therefore provides a visual indicator of the predictive performance of the tree at each node and shows the high predictive performance achieved.

Effective cohesion and thickness of top soil appear multiple times in the same branch in Fig. 5 (for example, at the leftmost and the rightmost branches), which may indicate that these two factors interact with each other. This suspicion is confirmed if we look at the scatter plot of the effective cohesion of top soil samples versus the thickness of top soil samples for simulations

25   that lead to slope failure (black) and stability (grey) (Fig. 6a). The triangular pattern in Fig. 6a clearly indicates that these two factors interact to produce slope failure, i.e. a slope with more cohesive soil can be thicker without experiencing failure. Visualising the thresholds identified from the CART analysis (Fig. 5) in this scatter plot (red dashed line in Fig. 6b) shows the inability of CART to characterise this interaction, i.e. CART attempts to reproduce the interaction through a sequence of vertical and horizontal separations. One can approach this problem either by rotating the axes of this graph or by creating a

30   new auxiliary variable combining the interacting factors (Dalal et al., 2013). We create a new auxiliary variable, the ratio between effective cohesion and thickness of the top soil, because we believe that it is a physically meaningful variable. We then generate a new tree based on the original factors plus the new auxiliary variable. The resulting tree, pruned for a similar error as the original tree, is shown in Fig. 7. It still classifies 89% of the simulations correctly but using a much simpler structure than the tree in Fig. 5. Figure 7 shows that when the ratio of cohesion to thickness of the top soil is above 2.0 kPa.m^{-1}, CHASM

predicts a stable slope most of the time (4611 simulations correspond to a stable slope and only 385 simulations produce slope failure). When the top soil cohesion to thickness ratio is below 2.0 kPa.m^{-1} then rainfall characteristics need to be taken into account to predict slope stability. Fig. 6c also shows that the new separation line (red dashed) between failure and stability is now more consistent with the underlying scatter plot.

5   We further expect rainfall intensity and duration to interact in the context of slope stability. Indeed, numerous field observations in the literature, and empirical relationships used in regional landslide early warning systems worldwide, show that both high-intensity/short-duration rainfall combinations and low-intensity/long-duration combinations can result in slope failure (Lumb, 1975; Crosta, 1998; Martelloni et al., 2012). Empirical rainfall intensity-duration (I-D) thresholds are typically obtained by plotting observed landslide events according to their triggering rainfall intensity and duration, and are generally linear, with a

10   negative gradient, when logarithmic scales are used. This log-log relationship can also be seen in Fig. 8, which shows a separation between the rainfall I-D combinations that can trigger landslides (black dots) and those that do not (grey crosses). We therefore create a second auxiliary variable – the negative ratio between the logarithm of rainfall intensity and the logarithm of rainfall duration ($-\log(I)/\log(D)$) – and repeat the CART analysis to see whether it is possible to further simplify the tree. The resulting tree (Fig. 9) is pruned based on a similar error to previous trees (Figs. 5 and 7), but with a significantly simplified

15   structure. We consider this final classification tree as the most effective output of the CART analysis, since its simplicity facilitates communication with stakeholders, while preserving the same information content.

5 Discussion

We have shown that the application of a mechanistic model such as CHASM in a combined Monte Carlo and CART analysis framework can help identify the dominant preparatory and triggering factors driving slope instability in a data scarce location.

20   Our method goes beyond previous studies by accounting for both site-specific preparatory factors (geometrical, geotechnical and hydrological conditions) and future uncertain rainfall triggers. Results for our study site indicate that targeted geotechnical data acquisition could help to constrain uncertainties in cohesion and soil depths; and the effects of different rainfall intensities and durations should be represented to capture both current and potential future hazard scenarios.

5.1 Preparatory and triggering factors driving slope instability

25   For our study site the ratio of the top soil strata's effective cohesion to strata thickness is shown to be a dominant factor in the stability of the slope. Inspection of selected model outputs indicates that the critical slip surface generally tends to be located within this top strata (weathering grade V-VI) with the deepest part of the slip circle at the interface with the weathered material strata (grade III-IV), as illustrated in Fig. 2. The slip surface location within the top soil strata is explained by the site-specific parameters (Table 1) in which: i) the higher strength weathered material strata constrains the failure surface to within the

30   weaker soil strata above; and ii) as rainfall infiltrates, the lower hydraulic conductivity of the weathered material is likely to cause loss of soil matric suction and thus a reduced apparent strength at the soil/weathered material interface.

Of the two top soil strata strength parameters, it is effective cohesion rather than angle of friction, that shows up in the CART analysis because the study site slope is likely to be cohesion controlled rather than friction controlled – i.e. the slope angle is typically greater than the friction angle. Figure 6 shows that when the top soil strata is thin it can remain stable even for very low values of effective cohesion, whereas, thicker soils tend to require a much higher effective cohesion for stability. This result is in keeping with other physically-based modelling studies of the relationships between the geometry of slopes, strata and shallow landslides in cohesion controlled slopes (Frattini and Crosta, 2013; Milledge et al., 2014). Ignoring the effects of water table location and pore water pressures, the greater self-weight of soil at the base of a thicker soil stratum generates higher shear stresses and requires greater shear resistance for stability than in a shallower soil.

The importance of cohesion for the stability of tropical residual soil slopes, such as our study site, and its inclusion in stability analysis is the subject of an ongoing debate amongst geotechnical engineering, researchers and practitioners. Laboratory analysis of the shear strength of remoulded clays show that the value of the effective cohesion parameter is affected by measurement uncertainties (Parry, 2004) and that peak cohesion is lost with seasonal cycles of dilatancy (Take and Bolton, 2011). The known sensitivity of slope stability to cohesion and the uncertainties associated with its measurement thus lead some engineers to adopt a highly conservative approach to the value of effective cohesion used in slope design – often assigning it a value of zero (Schofield, 2006). Yet, landslide hazard assessment scientists and engineering practitioners in the humid tropics argue for its inclusion as an observable strength parameter in the analysis of existing slopes comprised of undisturbed tropical residual soils which exhibit relict structures from the weathered parent rock (Burland et al., 2008). For the study site, our modelling approach supports the inclusion of non-zero values of effective cohesion in the analysis to account for its observable stability in the field. Our results also indicate that any data acquisition strategy for this site should target both soil thickness and effective cohesion values (both undisturbed and remoulded) to improve landslide hazard predictions.

The second most important factor in the stability of our study site is the nature of rainfall events in terms of their intensity and duration. As noted in Sect. 4, when the predicted failed and stable slopes are plotted on log-log axes of associated rainfall intensities and durations (Fig. 8) a negative linear threshold is found above which landslides are more likely to occur. This relationship is observed in landslide inventories which are widely used to generate empirical regional rainfall intensity-duration (I-D) thresholds of the form: $I = a_1 D^{-a_2}$, where a_1 and a_2 are parameters specific to a site or region (e.g., Larsen and Simon, 1993; Guzzetti et al., 2007). In the absence of empirical data, physically-based models may be used to generate synthetic thresholds using Monte Carlo methods (e.g., Peres and Cancelliere, 2014). Our method thus demonstrates both the importance of representing the dynamic hydrological processes involved in triggering landslides, while also providing a starting point for generating site-specific rainfall I-D thresholds in data scarce locations.

We have to reiterate that our results are valid within the context of the assumptions made in our study. For example, changing a particular input factor distribution may influence the importance of other factors (Sect. 3.1 and Table 1). This is of course not a limitation that is specific to our study, but simply the constraint of any model-based analysis, which can in any case be reduced by additional analysis. The approach presented here provides a means for identifying dominant landslide preparatory and triggering factors, guiding data acquisition, updating the analysis and refining the hazard assessment.

5.2 Can deep uncertainty in future rainfall exceed other uncertainties?

In this study, we have evaluated how uncertainty about slope characteristics and future rainfall change may influence risk of slope failure. Our findings have demonstrated that, for our study site, physical slope properties, namely effective cohesion and thickness of the top soil, are significant drivers of slope stability. For this slope, these physical properties have a more significant impact on landslide hazard risk than variability in future rainfall intensity and duration.

Our findings have a number of important implications for landslide hazard research. With few exceptions (e.g., Melchiorre and Frattini, 2012), previous studies typically have analysed the impacts of uncertainty related to slope characteristics and future climate independently. For example, Dehn and Buma (1999), Collison et al. (2000) and Ciabatta et al. (2016) consider the impacts of climate change on slope stability, but ignore uncertainty around soil properties; while Rubio et al. (2004) account for uncertainty introduced by slope hydrology and geotechnical properties, but ignore uncertainty relating to design storms. However, our results suggest that the failure to consider both sources of uncertainty simultaneously may lead to a significant underestimation of slope vulnerability under potential climate change. Furthermore, we have demonstrated how physically-based models, like CHASM, can be utilised to rapidly assess the impacts of multiple interacting and uncertain drivers of landslide occurrence in ways that would not be possible using simpler statistical models (e.g., Dixon and Brook, 2007).

To date, only Melchiorre and Frattini (2012) have attempted to analyse slope failure considering both uncertainty arising from slope properties and future climate. Their research used a top-down approach to quantify the impacts of multiple uncertainties. Specifically, they used Monte Carlo simulations and sensitivity analysis to assess the impacts of uncertainty in soil properties on predictions of slope stability for a pre-defined set of precipitation scenarios. We conclude that deep uncertainty due to potential future climate characteristics is not yet fully considered in landslide hazard assessment, and, as a result, policy recommendations may lead to undesirable outcomes given the difficulty in predicting impacts of climate change on future rainfall. In contrast, bottom-up approaches, such as that proposed in this study, consider a much wider range of possible system drivers (rainfall) and other uncertainties, without introducing assumptions regarding the probability of future precipitation. This insight is particularly relevant for small islands like the Caribbean, where the variety of different processes that contribute to rainfall change, some of which are poorly resolved by GCMs, make it very difficult to provide projections of future rainfall (Seneviratne et al., 2012).

5.3 Implications for data acquisition and assessment of future hazard

Bottom-up approaches to natural hazard risk assessment provide valuable knowledge to inform management decisions and to target data acquisition, especially in situations where resources may be limited. This study has shown how dominant physical slope properties driving landslide occurrence for a particular class of slope can be identified. As a result, decision-makers may seek to provide funding for targeted data acquisition to reduce uncertainty about the values of these parameters (e.g., thickness and effective cohesion of top soil) in order to improve understanding of how likely a slope is to fail. Decision-makers could also use the knowledge gained from CART to target management practices to improve slope stability (e.g., improving slope

drainage or land cover) or to develop options to mitigate consequences of slope failure (e.g., restricting development in high-risk areas). In this regard, the methodology developed in this study could be used to quantify the impact of different management options on landslide occurrence thresholds, providing an objective measure of the potential value of each strategy that can support debate amongst disaster risk reduction practitioners.

5   CART also has the distinct advantage that trees can provide useful knowledge for decision-makers even when uncertainties about future changes in climate are large, as is the case with rainfall projections in the Caribbean islands that are predominant areas where landslides occur globally. Any available climate information can be used *a posteriori* to assess the plausibility of threshold tipping points being crossed, and support the discussion whether improved management is required to improve slope stability in the future. Ultimately, the decision to implement any given action depends on how risk averse stakeholders are and

10   how much they are willing to spend on an adaptation strategy. However, CART provides the tools and information to enable managers to make informed choices that are robust under a wide range of plausible future conditions, reducing the risk of wasted investment and/or unanticipated negative outcomes.

6 Conclusions

In this study, we used a combination of physical-based modelling and empirical CART analysis to quantify the importance of

15   different sources of uncertainty when predicting landslide hazards for an example case study in the Caribbean. Contrary to common assumptions, our findings have showed that prediction of landslide occurrence may be more strongly influenced by uncertainty related to physical slope properties (e.g., cohesion and thickness of the top soil) than by (deep) uncertainty associated with future changes in rainfall patterns due to climate change. We suggest that failure to account for uncertainty related to both slope properties and climate change therefore will lead to a significant underestimation of landslide risks and

20   associated impacts on human populations.

The methodology developed in this paper has demonstrated that bottom-up approaches, implemented here using CART, can provide valuable information for assessment of landslide hazards even in data sparse environments. Our bottom-up approach illuminates dominant drivers of slope instability, enabling stakeholders and decision-makers to target data acquisition to reduce model prediction uncertainty. Moreover, CART analysis provides estimates of critical rainfall thresholds at which slope failure

25   is predicted to occur. Using this knowledge, decision-makers can assess whether it is likely these threshold tipping points being crossed in the future given available climate change information, and they can determine if improved management may be required to ensure long-term slope stability in the face of climate change.

The approach presented is valid within the context of the model used, including its assumptions, and any other assumptions made. The results are transferrable to other regions where the assumptions made here hold true. Future work will focus on

30   expanding these assumptions, e.g. by varying other slope characteristics such as slope angle and height to reflect a wider region. The general approach to bottom-up analysis presented here is transferrable to other hazard models.

Appendix A: Tree pruning

To avoid overfitting of data we pruned the tree using cross validation. Figure A.1 shows the cross validation error (i.e. the proportion of misclassified data on the validation dataset) for different pruning levels. The smallest cross validation error is obtained for pruning level 23, which corresponds to a tree with 51 leaf nodes. However, Fig. A.1 shows that choosing a much simpler tree, for example one with 25 leaf nodes (pruning level 31), also results in a very small error. Simpler trees are easier to interpret, and, therefore, in this study we select the tree with 25 leaf nodes, which has a cross validation error equal to 0.119.

Acknowledgements

This work was supported by the Natural Environment Research Council [Consortium on Risk in the Environment: Diagnostics, Integration, Benchmarking, Learning and Elicitation (CREDIBLE); grant number NE/J017450/1]. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc. The SAFE Toolbox for Global Sensitivity Analysis is freely available for research purposes from www.bris.ac.uk/cabot/resources/safe-toolbox/.

References

Aitsi-Selmi, A., Blanchard, K., Al-Khudhairy, D., Ammann, W., Basabe, P., Johnston, D., Ogallo, L., Onishi, T., Renn, O., Revi, A., Roth, C., Peijun, S., Schneider, J., Wenger, D., and Murray, V.: UNISDR STAG 2015 Report: Science is used for disaster risk reduction 2015.

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, L.: Sustainable landslide risk reduction in poorer countries, Proceedings of the Institution of Civil Engineers - Engineering Sustainability, 159, 23-30, doi: 10.1680/ensu.2006.159.1.23, 2006.

Anderson, M. G., Kemp, M. J., and Lloyd, D. M.: Hydrological design manual for slope stability in the tropics, Transport Research Laboratory, Crowthorne, UK, 58 pp., 1997.

Anderson, M. G. and Lloyd, D. M.: Using a Combined Slope Hydrology Stability Model to Develop Cut Slope Design Charts, P I Civil Eng Pt 2, 91, 705-718, doi: 10.1680/iicep.1991.17486, 1991.

Arnone, E., Dialynas, Y. G., Noto, L. V., and Bras, R. L.: Accounting for soils parameter uncertainty in a physically based and distributed approach for rainfall-triggered landslides, Hydrol Process, 30, 927-944, doi: 10.1002/hyp.10609, 2016.

Bankes, S. C.: Tools and techniques for developing policies for complex and uncertain systems, Proceedings of the National Academy of Sciences, 99, 7263-7266, doi: 10.1073/pnas.092081399, 2002.

Beven, K.: Changing ideas in hydrology - the case of physically-based models, J. Hydrol., 105, 157-172, doi: 10.1016/0022-1694(89)90101-7, 1989.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18-36, doi: 10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Germann, P.: Macropores and water flow in soils revisited, Water Resour Res, 49, 3071-3092, doi: 10.1002/wrcr.20156, 2013.

Bishop, A. W.: The use of the Slip Circle in the Stability Analysis of Slopes, Géotechnique, 5, 7-17, 1955.

Borgatti, L. and Soldati, M.: Landslides and climatic change. In: Geomorphological Hazards and Disaster Prevention, Alcántara-Ayala, I. and Goudie, A. S. (Eds.), Cambridge University Press, Cambridge, 2010.

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

Bull-Kamanga, L., Diagne, K., Lavell, A., Leon, E., Lerise, F., MacGregor, H., Maskrey, A., Meshack, M., Pelling, M., Reid, H., Satterthwaite, D., Songsore, J., Westgate, K., and Yitambe, A.: From everyday hazards to disasters: the accumulation of risk in urban areas, Environ Urban, 15, 193-203, doi: 10.1177/095624780301500109, 2003.

Burland, J. B., Dean, E. T. R., Gudehus, G., Muhunthan, B., and Collins, I. F.: Discussion: Interlocking, and peak and design strengths., Géotechnique, 58, 527-532, doi: geot.2008.58.6.527, 2008.

Cho, S. E.: Effects of spatial variability of soil properties on slope stability, Eng. Geol., 92, 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., in press, doi: 10.1016/j.jhydrol.2016.02.007, 2016.

Collins, M., Chandler, R. E., Cox, P. M., Huthnance, J. M., Rougier, J., and Stephenson, D. B.: Quantifying future climate change, Nat Clim Change, 2, 403-409, doi: 10.1038/Nclimate1414, 2012.

Collison, A., Wade, S., Griffiths, J., and Dehn, M.: Modelling the impact of predicted climate change on landslide frequency and magnitude in SE England, Eng. Geol., 55, 205-218, doi: 10.1016/S0013-7952(99)00121-0, 2000.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervas, J., and Smith, J. T.: Recommendations for the quantitative analysis of landslide risk, B Eng Geol Environ, 73, 209-263, doi: 10.1007/s10064-013-0538-8, 2014.

Crosta, G.: Regionalization of rainfall thresholds: an aid to landslide hazard evaluation, Environ. Geol., 35, 131-145, doi: 10.1007/s002540050300, 1998.

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

Dai, F. C. and Lee, C. F.: Frequency–volume relation and prediction of rainfall-induced landslides, Eng. Geol., 59, 253-266, doi: 10.1016/S0013-7952(00)00077-6, 2001.

Dai, F. C., Lee, C. F., and Ngai, Y. Y.: Landslide risk assessment and management: an overview, Eng. Geol., 64, 65-87, doi: 10.1016/s0013-7952(01)00093-x, 2002.

Dalal, S., Han, B., Lempert, R., Jaycocks, A., and Hackbarth, A.: Improving scenario discovery using orthogonal rotations, Environ Modell Softw, 48, 49-64, doi: 10.1016/j.envsoft.2013.05.013, 2013.

Darcy, H.: Les Fontaines Publiques de la Ville de Dijon., Dalmont, Paris, 1856.

De Graff, J. V., Bryce, R., Jibson, R. W., Mora, S., and Rogers, C. T.: Landslides: their extent and economic significance in the Caribbean. In: Landslides: Extent and Economic Significance, Brabb, E. E. and Harrod, B. L. (Eds.), A. A. Balkema, Rotterdam, 1989.

Dehn, M. and Buma, J.: Modelling future landslide activity based on general circulation models, Geomorphology, 30, 175-187, doi: 10.1016/S0169-555x(99)00053-7, 1999.

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A. L., Arnold, M., Agwe, J., Buys, P., Kjekstad, O., Lyon, B., and Yetman, G.: Natural Disaster Hotspots: A Global Risk Analysis, The World Bank, Washington, D.C., 2005.

Dixon, N. and Brook, E.: Impact of predicted climate change on landslide reactivation: case study of Mam Tor, UK, Landslides, 4, 137-147, doi: 10.1007/s10346-006-0071-y, 2007.

Frattini, P. and Crosta, G. B.: The role of material properties and landscape morphology on landslide size distributions, Earth and Planetary Science Letters, 361, 310-319, doi: 10.1016/j.epsl.2012.10.029, 2013.

Friedman, J. and Fisher, N.: Bump hunting in high-dimensional data, Statistics and Computing, 9, 123-143, doi: 10.1023/A:1008894516817, 1999.

GCO: Mid-Levels Study: Report on Geology, Hydrology and Soil Properties, Geotechnical Control Office, Public Works Department Hong Kong, Hong Kong, 1982.

5 GEO: Guide to Rock and Soil Descriptions. Geoguide 3, Civil Engineering and Development Department, The Government of the Hong Kong, Special Administrative Region, Hong Kong, PR China, 1988.

Groves, D. G. and Lempert, R. J.: A new analytic method for finding 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, P. C.: Rainfall thresholds for the initiation of landslides in central and 10 southern Europe, Meteorology and Atmospheric Physics, 98, 239-267, doi: 10.1007/s00703-007-0262-7, 2007.

Hall, J. W., Rubio, E., and Anderson, M. G.: Random sets of probability measures in slope hydrology and stability analysis, Zeitschrift Fur Angewandte Mathematik Und Mechanik, 84, 710-720, doi: 10.1002/zamm.200410146, 2004.

Hallegatte, S., Shah, A., Lempert, R., Brown, C., and Gill, S.: Investment decision making under deep uncertainty application to climate change, The World Bank, Washington, DC, 2012.

15 Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd, Springer, New York, 2009.

Herman, J. D., Zeff, H. B., Reed, P. M., and Characklis, G. W.: Beyond optimality: Multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty, Water Resour Res, 50, 7692-7713, doi: 10.1002/2014WR015338, 2014.

20 Holcombe, E., Smith, S., Wright, E., and Anderson, M. G.: An integrated approach for evaluating the effectiveness of landslide risk reduction in unplanned communities in the Caribbean, Nat Hazards, 61, 351-385, doi: 10.1007/s11069-011-9920-7, 2012.

Ibsen, M. L. and Brunsden, D.: The nature, use and problems of historical archives for the temporal occurrence of landslides, with specific reference to the south coast of Britain, Ventnor, Isle of Wight, Geomorphology, 15, 241-258, doi: 10.1016/0169-555x(95)00073-E, 1996.

25 Janbu, N.: Application of composite slip surface for stability analysis, European Conference on the Stability of Earth Slopes, Stockholm, Sweden, 43-49, 1954.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nature Clim. Change, 4, 570-576, doi: 10.1038/nclimate2258, 2014.

30 Kim, H. G., Lee, D. K., Park, C., Kil, S., Son, Y., and Park, J. H.: Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios, Environmental Earth Sciences, 73, 1385-1400, doi: 10.1007/s12665-014-3775-7, 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. and Simon, A.: A rainfall intensity-duration threshold for landslides in a humid-tropical environment, Puerto Rico, Geografiska Annaler. Series A, Physical Geography, 75, 13-23, doi: 10.2307/521049, 1993.

Lempert, R., Bryant, B. P., and Bankes, S. C.: Comparing algorithms for scenario discovery, 2008.

Lempert, R. J., Popper, S. W., and Bankes, S. C.: Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis, RAND, Santa Monica, CAMR-1626-RPC, 2003.

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, Proceedings of the Institution of Civil Engineers - Civil Engineering, 144, 129-134, doi: 10.1680/cien.2001.144.3.129, 2001.

Lumb, P.: Slope failures in Hong Kong, Quarterly Journal of Engineering Geology and Hydrogeology, 8, 31-65, doi: 10.1144/gsl.qjeg.1975.008.01.02, 1975.

Martelloni, G., Segoni, S., Fanti, R., and Catani, F.: Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides, 9, 485-495, doi: 10.1007/s10346-011-0308-2, 2012.

Mathworks: Matlab Statistics and Machine Learning Toolbox Release 2015a. Mathworks Inc., Natwick, MA, USA, 2015.

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

Migon, P.: Mass movement and landscape evolution in weathered granite and gneiss terrains, 23, 33-45, doi: 10.1144/EGSP23.4, 2010.

Milledge, D. G., Bellugi, D., McKean, J. A., Densmore, A. L., and Dietrich, W. E.: A multidimensional stability model for predicting shallow landslide size and shape across landscapes, Journal of Geophysical Research: Earth Surface, 119, 2481-2504, doi: 10.1002/2014jf003135, 2014.

Millington, R. J. and Quirk, J. P.: Permeability of Porous Media, Nature, 183, 387-388, doi: 10.1038/183387a0, 1959.

Minder, J. R., Roe, G. H., and Montgomery, D. R.: Spatial patterns of rainfall and shallow landslide susceptibility, Water Resour Res, 45, doi: 10.1029/2008wr007027, 2009.

Ning, L., Mann, M. E., Crane, R., Wagener, T., Najjar, R. G., and Singh, R.: Probabilistic Projections of Anthropogenic Climate Change Impacts on Precipitation for the Mid-Atlantic Region of the United States, J Climate, 25, 5273-5291, doi: 10.1175/Jcli-D-11-00565.1, 2012.

Parry, R. G. H.: Mohr Circles, Stress Paths and Geotechnics, 2nd edition, Spon Press, 2004.

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

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

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T.: Sensitivity analysis of environmental models: a systematic review with practical workflow, Environ Modell Softw, 79, 214-232, doi: 10.1016/j.envsoft.2016.02.008, 2016.

Pianosi, F., Sarrazin, F., and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environ Modell Softw, 70, 80-85, doi: 10.1016/j.envsoft.2015.04.009, 2015.

Ray, P. A. and Brown, C. M.: Confronting climate uncertainty in water resources planning and project design : the decision tree framework., World Bank, Washington, D.C., 2015.

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics-J Gen. Appl. Phys., 1, 318-333, doi: 10.1063/1.1745010, 1931.

Rubio, E., Hall, J. W., and Anderson, M. G.: Uncertainty analysis in a slope hydrology and stability model using probabilistic and imprecise information, Comput Geotech, 31, 529-536, doi: 10.1016/j.compgeo.2004.09.002, 2004.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis. The Primer, John Wiley & Sons, Ltd, Chichester, UK, 2008.

Schofield, A. N.: Interlocking, and peak and design strengths, Géotechnique, 56, 357-358, doi: 10.1680/geot.2006.56.5.357, 2006.

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M. (Eds.), Cambridge University Press, Cambridge, 2012.

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.

Smyth, C. G. and Royle, S. A.: Urban landslide hazards: incidence and causative factors in Niteroi, Rio de Janeiro State, Brazil, Appl Geogr, 20, 95-117, doi: 10.1016/S0143-6228(00)00004-7, 2000.

Soeters, R. and van Westen, C. J.: Slope instability recognition, analysis and zonation. In: Landslides, investigation and mitigation, Special Report 247, Transportation Research Board, National Research Council, 1996.

Spear, R. C. and Hornberger, G. M.: Eutrophication in Peel Inlet .2. Identification of Critical Uncertainties Via Generalized Sensitivity Analysis, Water Res, 14, 43-49, doi: 10.1016/0043-1354(80)90040-8, 1980.

Stein, S. and Stein, J. L.: Shallow versus deep uncertainties in natural hazard assessments, Eos, Transactions American Geophysical Union, 94, 133-134, doi: 10.1002/2013EO140001, 2013.

Take, W. A. and Bolton, M. D.: Seasonal ratcheting and softening in clay slopes, leading to first-time failure, Géotechnique, 61, 757-769, doi: 10.1680/geot.9.P.125, 2011.

UNISDR: Global Assessment Report on Disaster Risk Reduction, United Nations, 2015.

UNU (United Nations University): Landslides. Asia has the most; Americas, the deadliest; Europe the costliest; Experts seek ways to mitigate landslide losses; Danger said growing due to climate change, other causes, news release January 18 2006, MR/E01/06/rev1, 2006.

von Ruette, J., Lehmann, P., and Or, D.: Effects of rainfall spatial variability and intermittency on shallow landslide triggering patterns at a catchment scale, Water Resour Res, 50, 7780-7799, doi: 10.1002/2013WR015122, 2014.

Wagener, T. and Gupta, H. V.: Model identification for hydrological forecasting under uncertainty, Stoch Env Res Risk A, 19, 378-387, doi: 10.1007/s00477-005-0006-5, 2005.

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

Wilkinson, P. L., Anderson, M. G., and Lloyd, D. M.: An integrated hydrological model for rain-induced landslide prediction, Earth Surf Proc Land, 27, 1285-1297, doi: 10.1002/esp.409, 2002a.

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 Modell Softw, 17, 333-344, doi: 10.1016/S1364-8152(01)00078-0, 2002b.

1   Table 1: CHASM parameters and their distributions.

Parameter	Symbol/Unit	Values		
		Grade V-VI[a] residual soil (0)	Grade III-IV[a] weathered material (1)	Grade I-II[a] bedrock (2)
Slope height	H (m)		U(48,49)	
Slope angle	δ (degrees)		U(27,30)	
Thickness of strata	H_i (m)	U(2,6)	U(2,6)	
Depth of initial water table[b]	DWT (%)		U(60,95)	
Saturated hydraulic conductivity	K_s (m.s^{-1})	ln N(-11.055,0.373)	ln N(-13.357,0.373)	1×10^{-8}
Saturated soil moisture content	θ_s (m^3.m^{-3})		N(0.413,0.074)	
Van Genuchten suction-moisture curve α	α (m^{-1})		ln N(1.264,1.076)	
Van Genuchten suction-moisture curve n[c]	n (-)		ln N(0.364,0.358)	
Residual soil moisture content[d]	θ_r (m^3.m^{-3})		ln N(-1.974,0.376)	
Initial surface suction	Ψ_{init}		U(-1.5,-0.5)	
Dry unit weight[e]	γ_d (kN.m^{-3})	U(16,18)	U(18,20)	23
Effective cohesion[f]	c' (kPa)	N(8,3.3)	N(25,1.7)	80
Effective friction angle	ϕ' (degrees)	N(25,1.7)	N(35,3.3)	60

[a] Weathering grades defined according to GEO (1988)

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

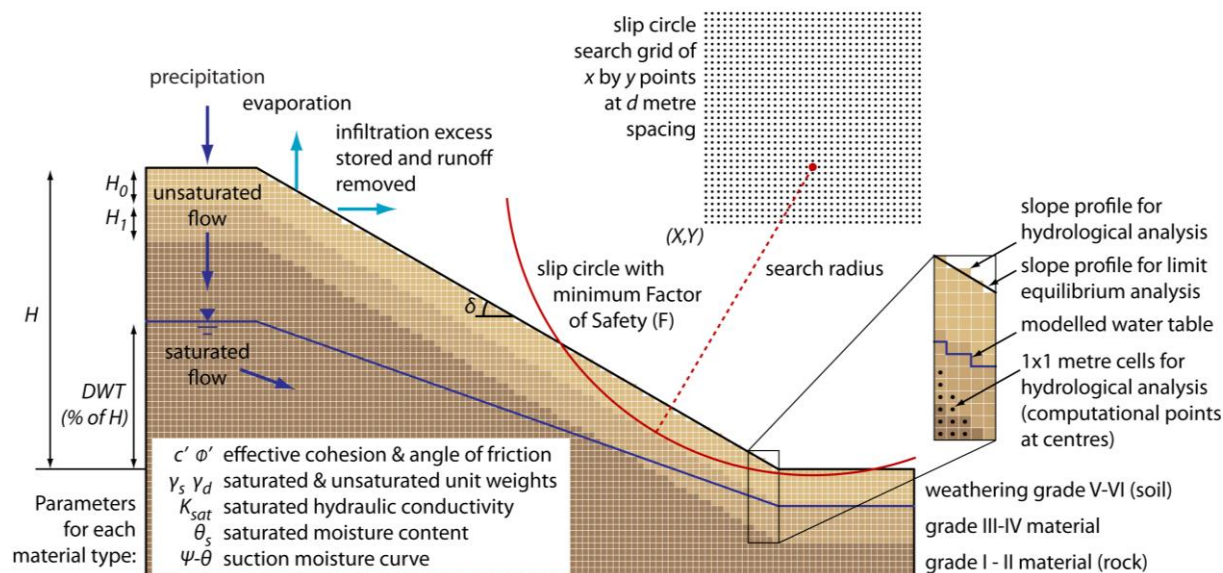
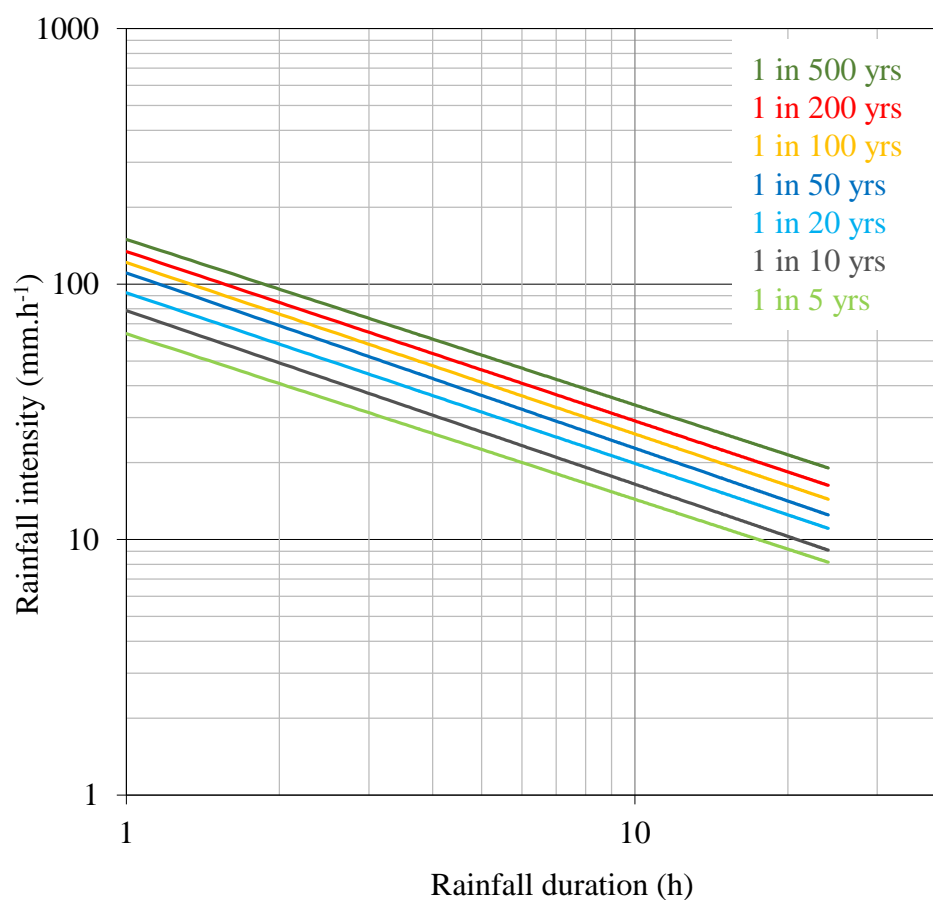
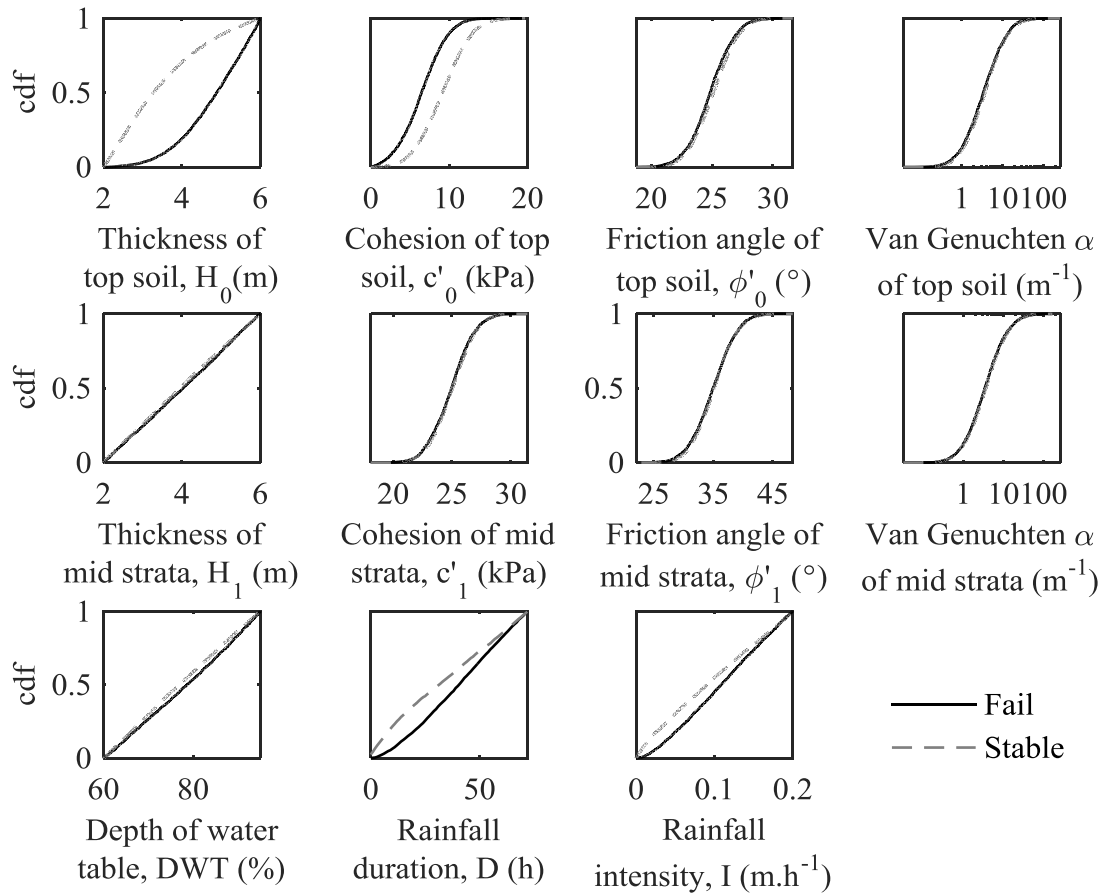
[c] n is always greater than 1

[d] θ_r is always greater than 0

[e] $\gamma_s = \gamma_d + 2$, where γ_s is the saturated unit weight

[f] Effective cohesion is always greater than 0

U – Uniform distribution; N – Normal distribution; ln N – lognormal distribution

2 **Figure 1: Typical informal housing on a landslide-prone slope in the Eastern Caribbean (photograph by Holcombe, 2007).**

2 **Figure 2: Conceptual model of the experimental slope. CHASM uses the Bishop method of slices to perform slope stability analysis,**
3 **which involves a numerical search for the slip surface.**

2 **Figure 3: Rainfall Intensity Duration Frequency curves for Saint Lucia developed by Klohn-Crippen (1995) using Gumbel analysis**
3 **of 40 years of hourly rainfall data from 15 rainfall gauges.**

2 **Figure 4: Cumulative probability distributions (cdf) of slope failure and stability predicted by CHASM for several different input**
3 **factors. Note that the Van Genuchten suction-moisture curve α is shown in logarithmic scale.**

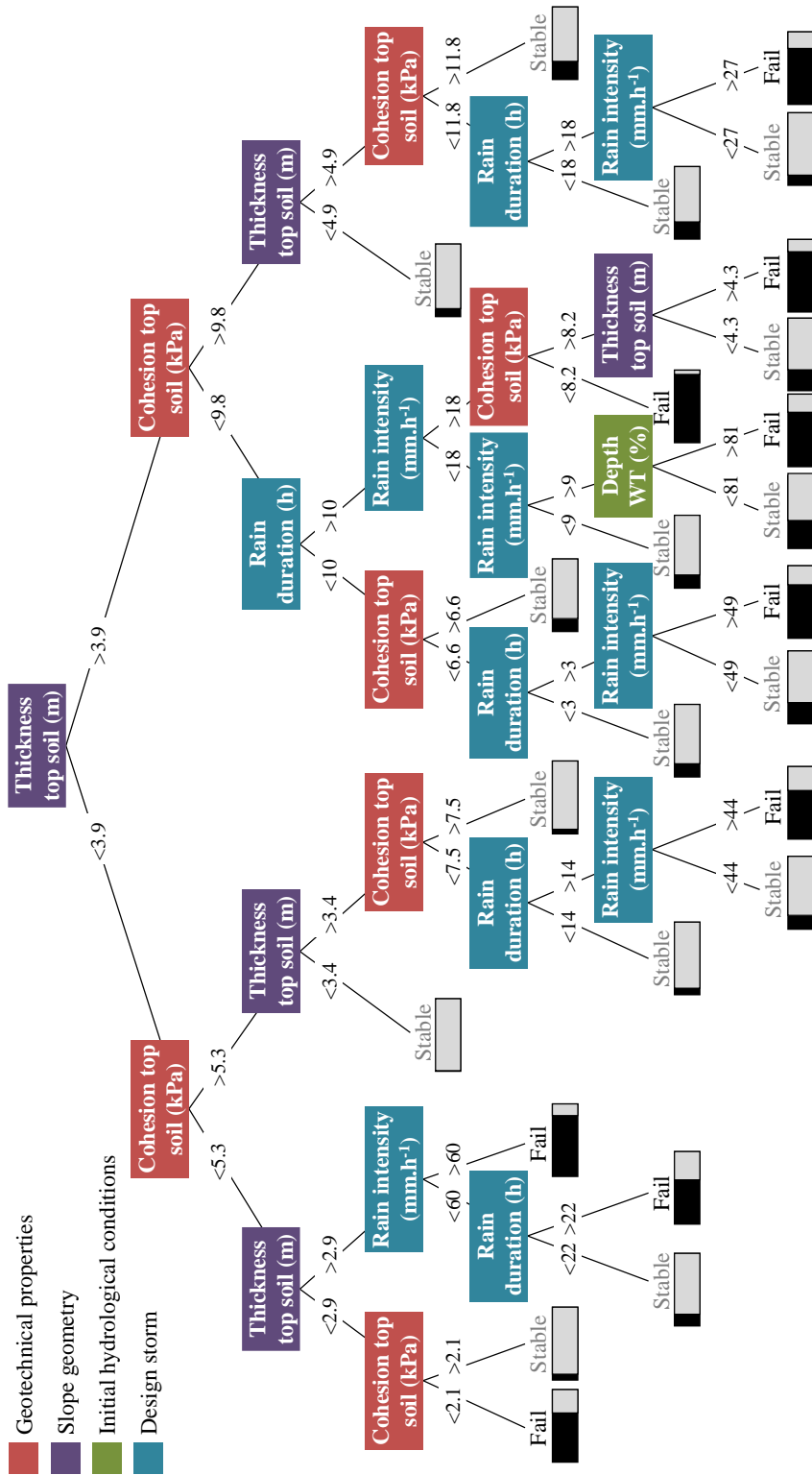



Figure 5: Classification tree of slope response. Each interior node corresponds to one of the analysed uncertain input factors (model parameters, boundary conditions and design storm properties). The bar under each leaf shows the proportion of simulations that resulted in slope failure (black) or stability (grey) for that leaf.

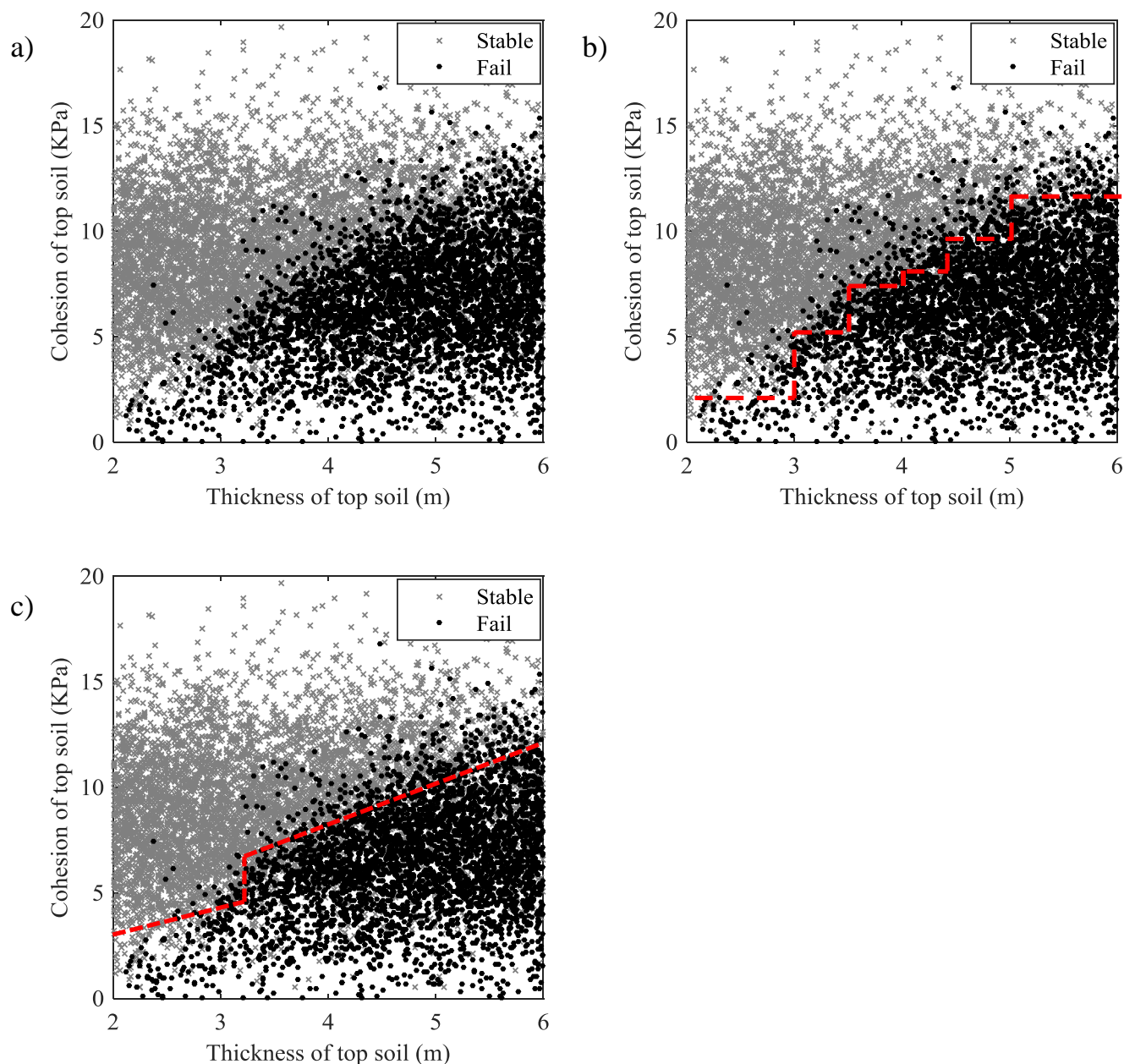
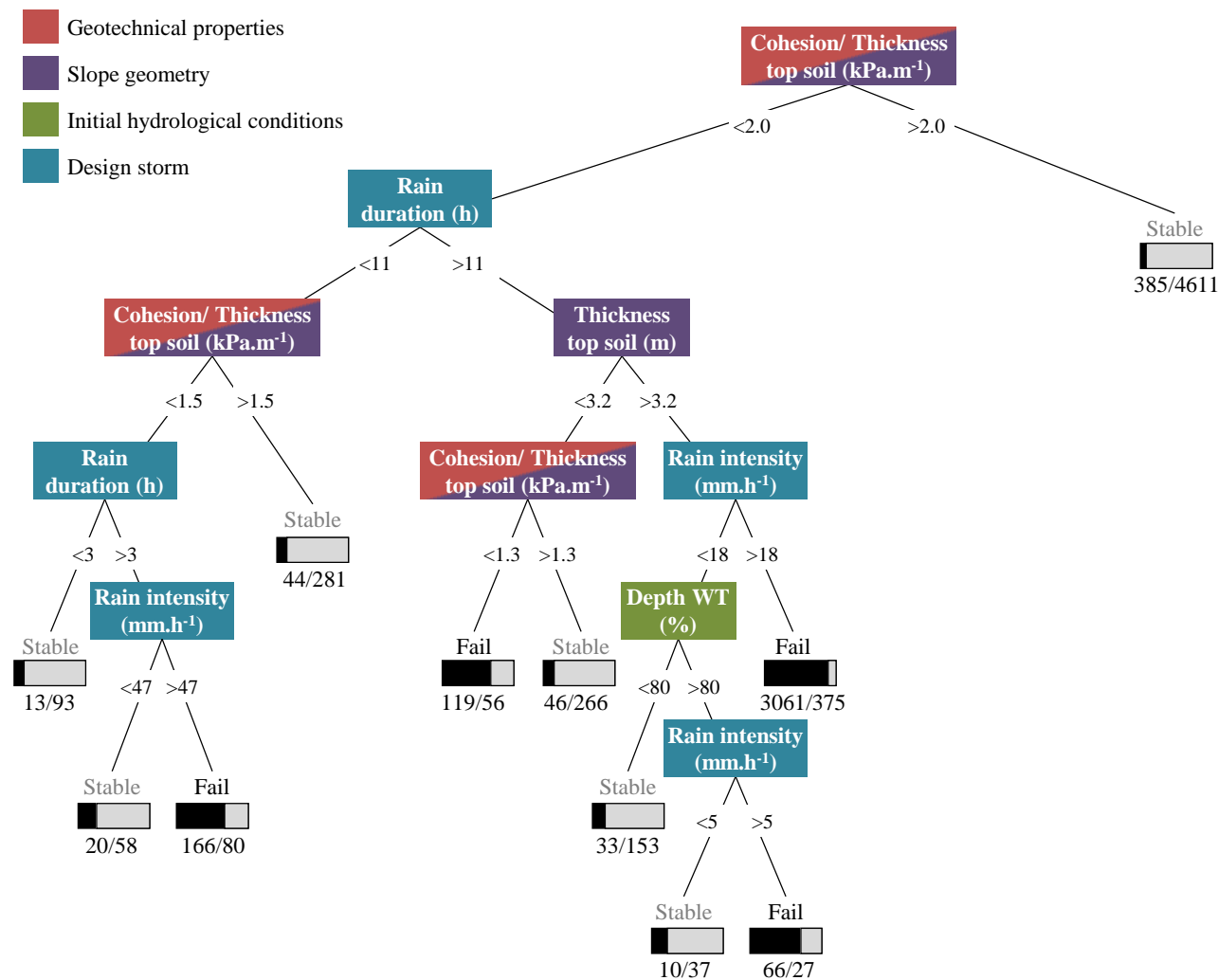



Figure 6: a) Sampled values of effective cohesion of top soil versus thickness of top soil. Black points correspond to simulations resulting in slope failure, while grey crosses correspond to simulations resulting in a stable slope. b) Same as panel a, with the additional red dashed line representing the thresholds derived from the classification tree in Fig. 5. c) Same as panel a, with the additional red dashed line representing the thresholds derived from the classification tree in Fig. 9.

2 **Figure 7: Classification tree of slope response, when adding an auxiliary variable (ratio between effective cohesion and thickness of**
3 **top soil) among uncertain input factors. The bar under each leaf shows the proportion and total number of simulations leading to**
4 **slope failure (black bar, left hand number) or a stable slope (grey bar, right hand number).**

2 **Figure 8: Sampled values of rainfall intensity versus rainfall duration when effective cohesion over thickness of top soil is smaller**
3 **than 2.0 kPa.m^{-1}.**

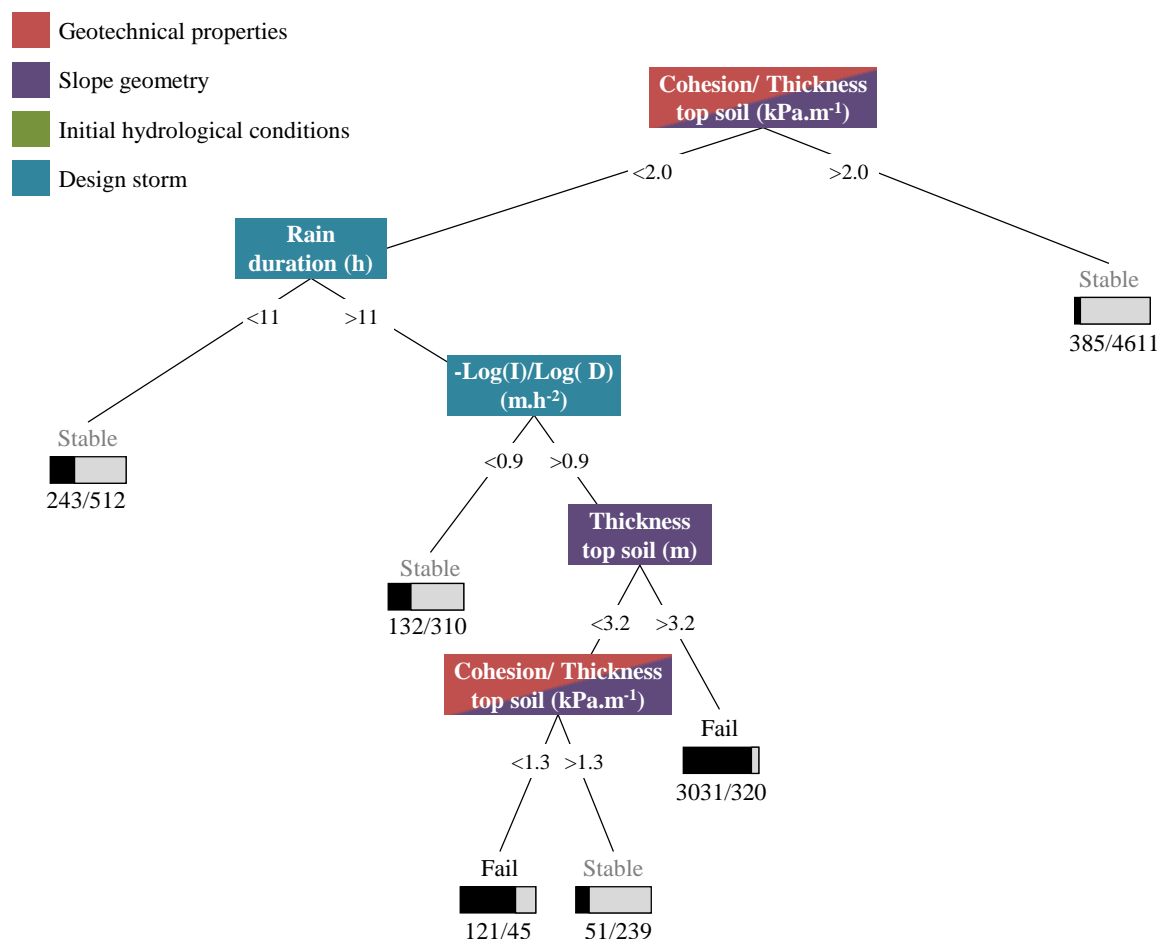



Figure 9: Classification tree of slope response, when adding two auxiliary variables: the ratio between effective cohesion and
thickness of top soil, and the negative ratio between the logarithm of rainfall intensity and the logarithm of rainfall duration. The
bar under each leaf shows the proportion and total number of simulations leading to slope failure (black bar, left hand number) or
a stable slope (grey bar, right hand number).

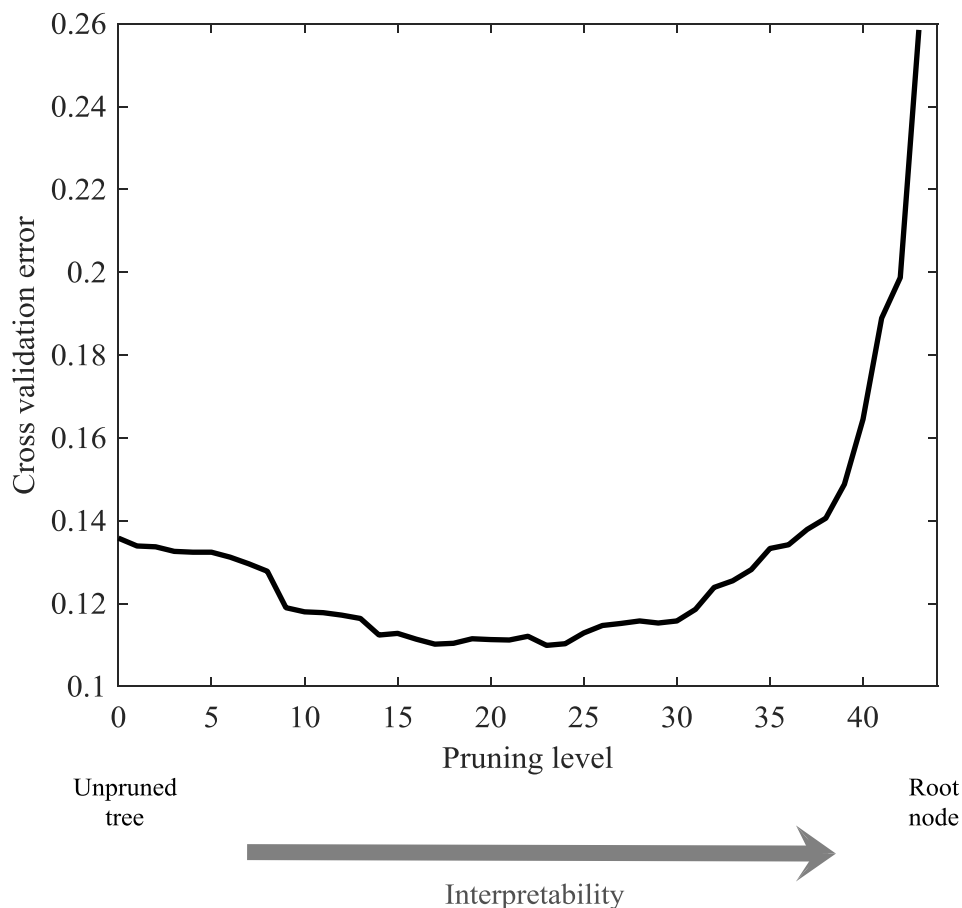



Figure A.1. Cross validation error of the classification tree for increasing pruning level. The cross validation error is computed by splitting the available training data (10000 CHASM model simulations in our case) at random into 10 subsets. Ten different classification trees are then constructed, where each tree is trained using 9 of the 10 subsets and its misclassification rate on the remaining subset is evaluated. The average of these 10 misclassification rates is the cross validation error (black line). The minimum cross validation error is equal to 0.110 and is obtained at pruning level 23, which corresponds to 51 leaf nodes. For the pruning level 31 (25 leaf nodes) the cross validation error is 0.119.