S1 Slope water management in 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 slope water management in CHASM

We developed a new functionality in 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 onto slopes.

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):

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

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. 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 the city wide drainage systems (Ortuste, 2012). Considering that the slope cross section is represented with a mesh of columns and cells, 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, any surface runoff is deleted because this surface water process 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 slope cells underneath the house and the amount of rain intercepted by the roof is discharged onto the slope cells to the sides of the house. If the roof is dual pitch, half of the intercepted rain is discharged upslope and half downslope, and it is equal to the rainfall rate multiplied by half of the roof area. This means that the surface water infiltrating into 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

1

30   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.

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
35   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
40   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
45   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

50

The landslide occurred in a 6 m high cut slope, at the bottom of the slope. 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
55   demonstrates how the model 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 GeoSlope to evaluate how the two models 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 CHAMS 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 GeoSlope

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) m³ m^{-3}; Ln (Residual water content – θres) m³m^{-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 respectively in Fig. S1.a and S1.b (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 in CHASM appears to be lower than the FS in 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 distinct hydrological and stability response. We therefore use a stochastic framework to perform a back analysis that

100   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.
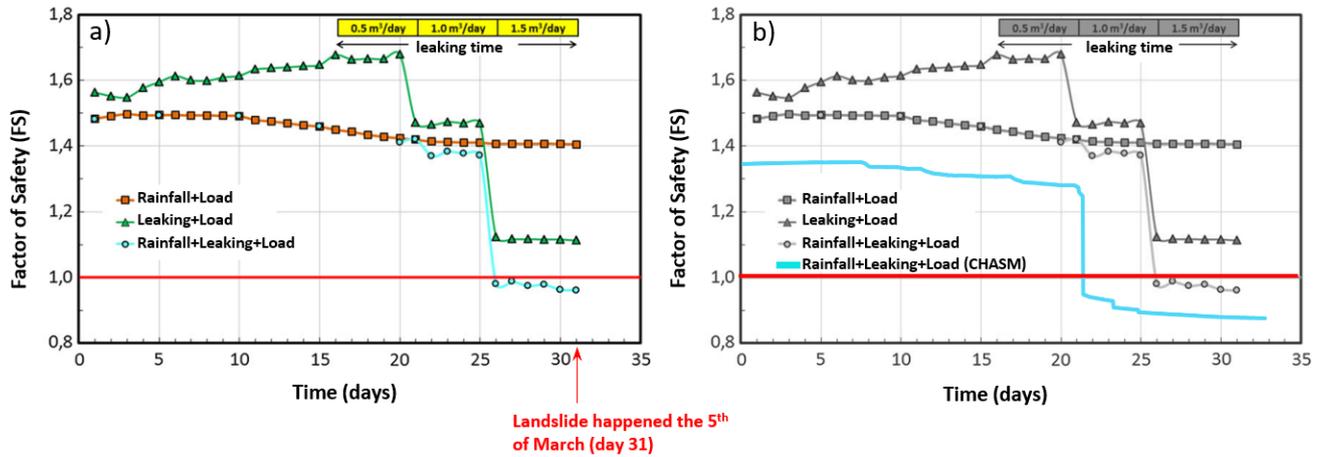
Figure S1: (a) results obtained by Mendes et al. using GeoSlope. (b) Results obtained with CHASM (blue line) in comparison with
105   **Mendes et al. for the case where rainfall, leakage and load are considered.**

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.
110   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
115   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 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
120   obtained with GeoSlope (black line).
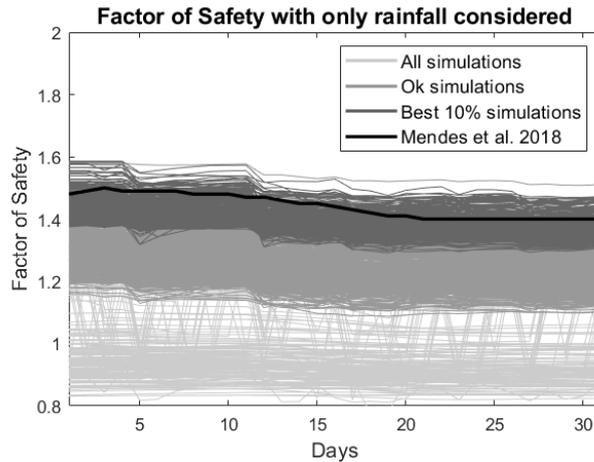
Figure S2: Factor of Safety calculated considering only rainfall and load of the tank for both CHASM and GeoSlope (this latter is referred as Mendes et al. 2018, and it corresponds to the orange line in figure S1a). The different FS trends of CHASM ("All" "Ok" and "Best 10%" simulations) are obtained using the ranges of S3 in the corresponding colour.

125

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.

130 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

135 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,

140 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. S1b and the light grey lines in Fig. S2). The uniform suction gradient assumed in

145 CHASM is physically unrealistic, but it is used for the initial distribution of water moisture content across cells. The

6

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.
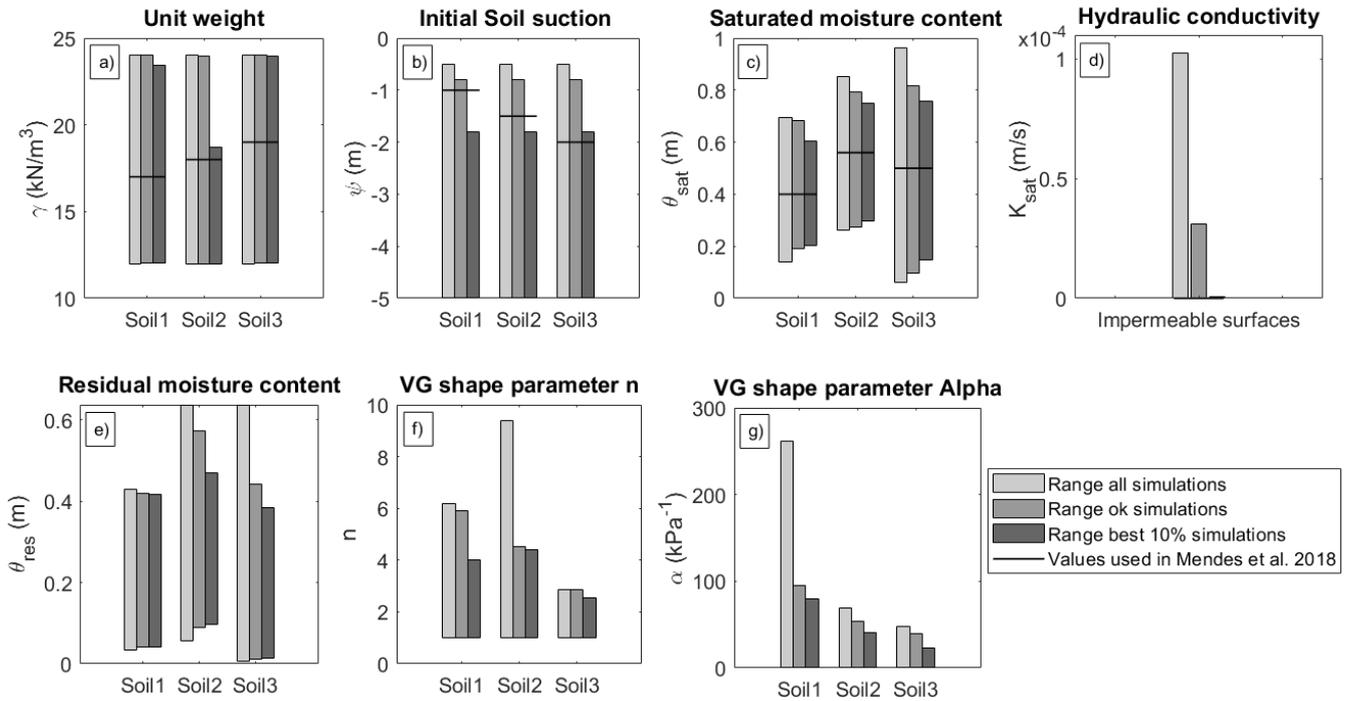
150

Figure S3: The input factors used in CHASM are varied within the ranges defined in Table 1. 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). 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 1). Note as in plot (d) the black line is at zero level.

155 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

160 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

170 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

175 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.
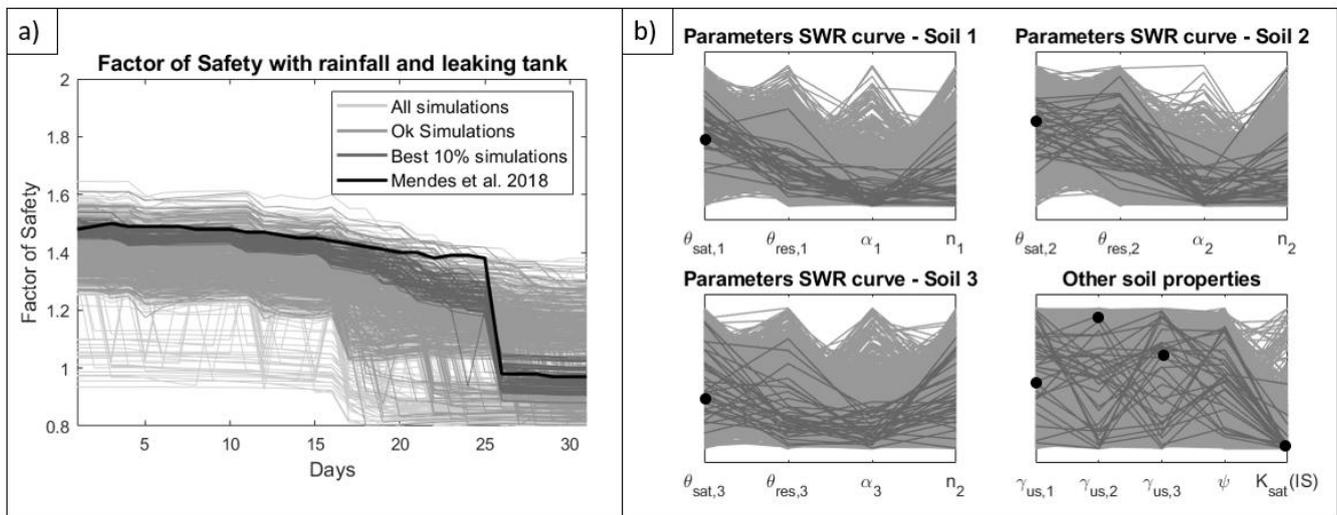
Figure S4: a) shows Factor of Safety calculated considering both rainfall and the leaking tank with CHASM and GeoSlope (referred as Mendes et al. 2018); b) shows the parallel plots of the ok (lighter grey) and best performing (darker grey) simulations.
180 **The lines identify how the input factors are 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

185 validation error is obtained for pruning level 123 and 143 respectively ("absolute minimum"), 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

8

(A.V.), shown in Fig. 7a and 7b 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. S8).
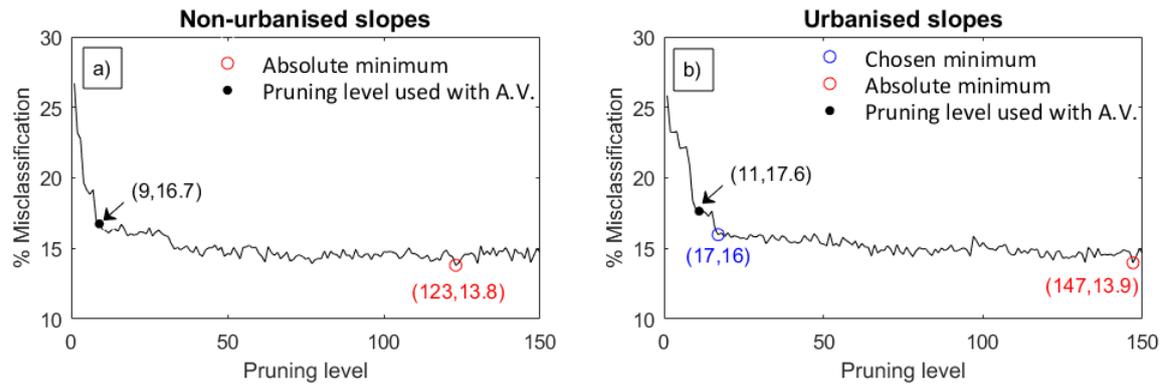
Figure S5: Cross-validation error of the CART for increasing pruning level for non-urbanised (a) and urbanised (b) slopes. The cross-validation error is computed by randomly dividing the dataset in 10 subgroups. Ten trees are then constructed by using 9 subgroups as training set. The excluded subgroup 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 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 (Fig7 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.
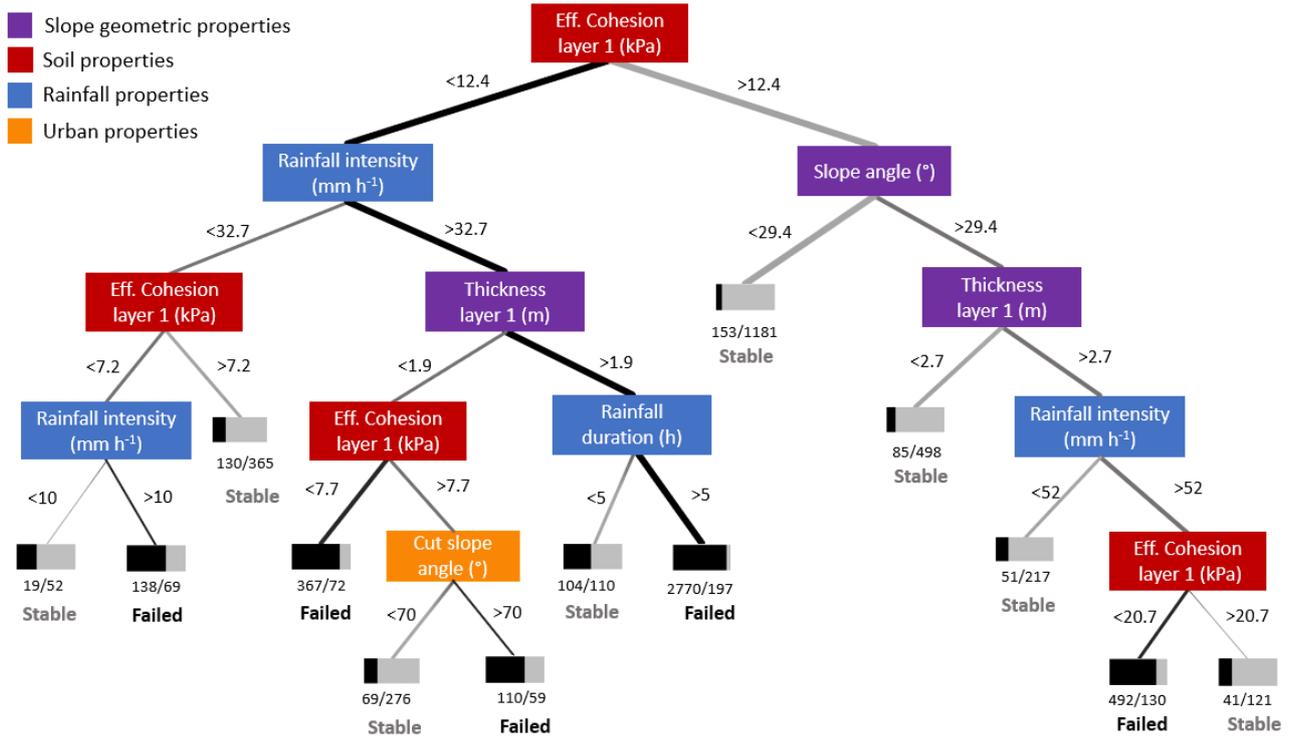
Figure S6: CART tree obtained for urbanised slopes without considering auxiliary variables. Black branches represent the paths that lead to simulations predicted as failed, while grey branches lead to simulations predicted as stable. The bar under each leaf shows the proportion of simulations that resulted 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. S7). The weights are represented by the sensitivity indices reported in Fig. 5 of the paper (w1 = 0.15 for slope angle; w2 = 0.13 for cut slope angle)

$$\text{Weighted Slope Angle} = \frac{w1*(\text{Slope angle}) + w2*(\text{Cut slope angle})}{w1+w2} \tag{S7}$$

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 S8 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"), 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.
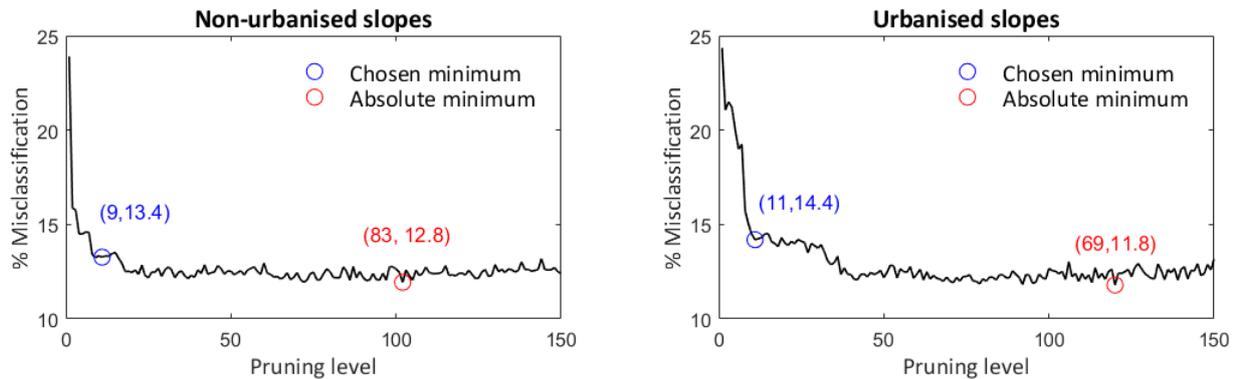
Figure S8: Cross-validation error of the CART for increasing pruning level. The cross-validation error is computed by randomly dividing the dataset in 10 subgroups. Ten trees are then constructed by using 9 subgroups as training set. The excluded subgroup is used to calculate the misclassification error (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 S9 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:

$$\log_{10}(I) = \gamma - \alpha \log_{10}(D) \tag{S10}$$

i.e. a linear equation where γ (the intercept) and α (the slope) are parameters specific to the site considered.
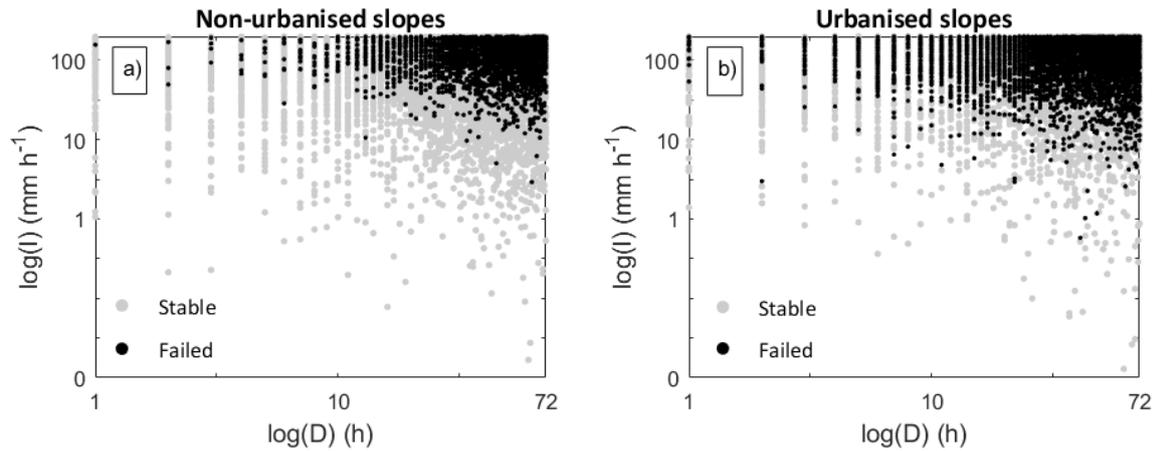
Figure S9: Combinations of rainfall intensities and durations resulted into stable (grey dots) or failed (black dots) slopes, for the non-urbanised (a) and urbanised (b) case. The plots show how the recorded landslides follow the typical descending trend found in empirical rainfall thresholds. The x and y axis are in logarithmic base 10, but the notation is linear for an easier readability.

250 To formalise the threshold that best divide failed from stable slopes, the parameters γ and α of Eq. (S10) 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

255 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:

260   -   the combinations of rainfall intensity and durations resulted in landslides (black dots in Fig. S9)

  -   a multi-objective optimisation algorithm for the evaluation of the two parameters γ and α of Eq. (S10).

The multi-objective optimisation involves minimising or maximising multiple objective functions subjected to a set of constrains. In this case, we want to draw a threshold line in the form of Eq. S10 which identifies the space where landslides are recorded. This translates into choosing parameters γ and α of Eq. S10 that satisfy the following two contrasting objectives:

265   1)   maximise the number of (simulated) failed slopes falling above the threshold line (Fig. S11a)

  2)   minimise the area above the threshold line (Fig. S11b)

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:

$$\gamma \, [-0.5; -2] \tag{S12a}$$

$$\alpha \, [0.05; 2] \tag{S12b}$$

The range of α so defined includes typical slope values of empirical rainfall thresholds (Guzzetti et al., 2007), while the range

270 of γ is designed to include all the rainfall intensities simulated. 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. We choose the (γ, α) set defining a threshold with 99.9% of failed simulations above it (reported in Fig. 8a,b of the main manuscript) i.e. a probability of recording landslides below this threshold equal to 0.1%. The method can be used to determine thresholds for any exceedance probability level. In this analysis, the multi-objective optimisation was

275 performed the "gamultiobj" function of the Matlab Optimisation Toolbox (R2018a).
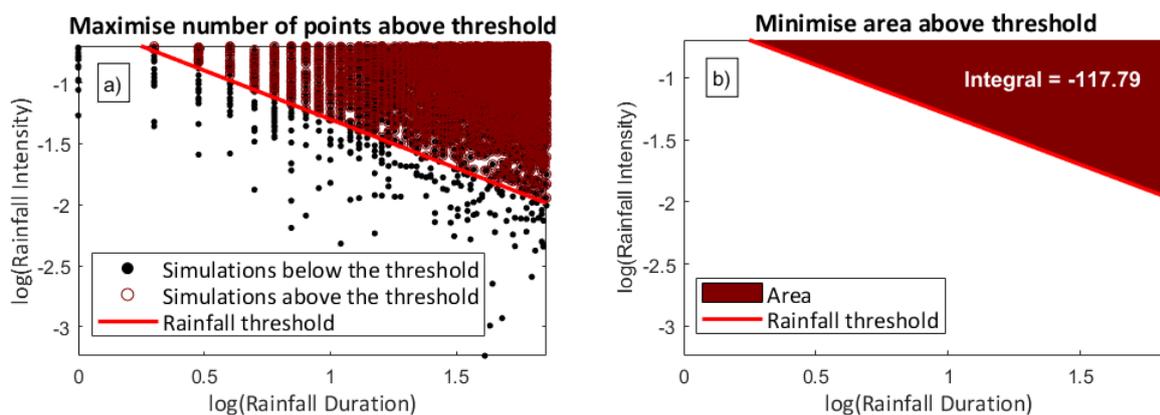
Figure S11: Illustration of the two objectives functions used in the optimisation, for a given threshold line: (a) maximise the number of failed slopes above the threshold and (b) minimise the area above the threshold.

References

280 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.

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.

285 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

290 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.

295 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.

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.

300 Mendes R.M., Filho V.M.: Real-time monitoring of climactic and geotechnical variables during landslides on the slopes of Serra do Mar and Serra da Mantiqueira (Sao Paulo State, Brazil). Engineering. doi:10.4236/eng.2015. 73012, 2015.

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.

305 Peruccacci, S., Brunetti, M. T., Luciani, S., Vennari, C. and Guzzetti, F.: Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy, Geomorphology, 139–140, 79–90, doi:10.1016/j.geomorph.2011.10.005, 2012.

Peruccacci, S., Brunetti, M. T., Gariano, S. L., Melillo, M., Rossi, M. and Guzzetti, F.: Rainfall thresholds for possible landslide occurrence in Italy, Geomorphology, 290(June), 39–57, doi:10.1016/j.geomorph.2017.03.031, 2017.

310 Piciullo, L., Gariano, S. L., Melillo, M., Brunetti, M. T., Peruccacci, S., Guzzetti, F. and Calvello, M.: Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides, Landslides, 14(3), 995–1008, doi:10.1007/s10346-016-0750-2, 2017.

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.

315 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.

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M. and Laber, J. L.: Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California, Landslides, 10(5), 547–562, doi:10.1007/s10346-012-0341-9, 2013.

320 Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils 1, Soil Sci. Soc. Am. J., 44(5), 892–898, 1980.