multi-Risk sciEnce for resilienT commUnities undeR a changiNg climate

Codice progetto MUR: **PE00000005** – CUP LEADER PARTNER C93C22005160002

Deliverable title: From deterministic to stochastic prediction in a weighted uncertainty range

Deliverable ID: 2.4.9

Due date: November 30th 2024

Submission date: November 30th 2024

AUTHORS

Simone Bizzi (UniPD); Andrea Brenna (UniPD); Xue Chen (UniPD); Ebrahim Ghaderpour (Sapienza); Valeria Lo Presti (UniPA); Viviana Mangraviti (UniPD); Nicolò Parrino (UniPA); Alberto Armigliato (UniBO); Silvia Ceramicola (OGS); Melania De Falco (UniNA); Carlo Esposito (Sapienza); Giovanni Forte (UniNA); Lucia Mele (UniNA); Fabio Rollo (Sapienza); Rita Tufano (UniNA); Filippo Zaniboni (UniBO)

1 Technical references

Project Acronym	RETURN
Project Title	multi-Risk sciEnce for resilienT commUnities undeR a changiNg climate
Project Coordinator	Domenico Calcaterra UNIVERSITA DEGLI STUDI DI NAPOLI FEDERICO II domcalca@unina.it
Project Duration	December 2022 – November 2025 (36 months)

Deliverable No.	DV 2.4.9
Dissemination level*	PU
Work Package	WP4 - Trigger-based multiple geohazard scenarios
Task	T 2.4.4 - Reliability and uncertainness of statistical solutions. Uncertainty assessment methods, based on back analysis of event distribution, for ensemble and single process as well as for coupled/cascade multiple triggers.
Lead beneficiary	UniPD
Contributing beneficiary/ies	UniPD, UniRoma1, UniPA, UniNA, OGS, UniBO

* PU = Public

PP = Restricted to other programme participants (including the Commission Services)

RE = Restricted to a group specified by the consortium (including the Commission Services)

CO = Confidential, only for members of the consortium (including the Commission Services)

1.1 Document history

Version	Date	Lead contributor	Description
0.1		S. Bizzi (UniPD); A. Brenna (UniPD); X. Chen (UniPD); E. Ghaderpour (Sapienza); V. Lo Presti (UniPA); V. Mangraviti (UniPD); N. Parrino (UniPA); A. Armigliato (UniBO); S. Ceramicola (OGS); M. De Falco (UniNA); C. Esposito (Sapienza); G. Forte (UniNA); L. Mele (UniNA); F. Rollo (Sapienza); R. Tufano (UniNA); F. Zaniboni (UniBO)	Final draft for Spoke review
0.2	25/11/2024	S. Martino, F. Bozzano (Sapienza); D. Calcaterra, D. Di Martire (UniNA)	Critical review and edits for approval
1.0	30/11/2024	S. Bizzi (UniPD); A. Brenna (UniPD); X. Chen (UniPD); E. Ghaderpour (Sapienza); V. Lo Presti (UniPA); V. Mangraviti (UniPD); N. Parrino (UniPA); A. Armigliato (UniBO); S. Ceramicola (OGS); M. De Falco (UniNA); C. Esposito (Sapienza); G. Forte (UniNA); L. Mele (UniNA); F. Rollo (Sapienza); R. Tufano (UniNA); F. Zaniboni (UniBO)	Final version

2 ABSTRACT

This deliverable presents the third output of Task 2.4.4, part of the WP 2.4 "Trigger-based multiple geohazard scenarios" focused on analysing the reliability and uncertainty of statistical methods in assessing ground instabilities under various environmental conditions.

In geoscience applications, data scarcity often arises due to limitations in instrumentation, high costs, and challenging weather conditions. Even when measurements are available, significant gaps—both temporally and spatially—can still exist. In these situations, researchers typically rely on physical and qualitative methods to model ground instability. As a result, quantifying uncertainties associated with both the data and the models used is often challenging. In such cases, conducting reliable sensitivity analyses to identify the most influential input data and parameters becomes essential.

In DV 2.4.8, three toolchain examples were presented from different environments—mountain, plain, and marine. In each case, the authors made an initial attempt to apply the Uncertainty Workflow introduced in DV 2.4.7. As the literature suggests, these studies highlight that quantifying the uncertainties involved in numerical or analytical simulations of ground instabilities events is inherently complex due to the numerous factors that must be considered.

This document will focus on the critical role of uncertainty analysis in modelling ground instabilities. We will draw on examples from both our own case studies and the broader literature. Our goal is to provide specific guidance and recommendations on performing Global Sensitivity Analysis in the context of ground instability, offering a roadmap for future research and development.

3 Table of contents

List of Figures

4 First Chapter: Introduction

This deliverable represents the third output of the Task 2.4.4 research having as its topic "Reliability and uncertainness of statistical solutions. Uncertainty assessment methods, based on back analysis of event distribution, for ensemble and single process as well as for coupled/cascade multiple triggers". The task is part of the WP 2.4 "Trigger-based multiple geohazard scenarios". As mentioned in the Executive Working Plan (Milestone 2.1), the Task 2.4.4 "is focused on the definition of contexts and indicators mostly affected by uncertainty (DV 2.4.7, completed in May 2024), the coding of procedures for the assessment of such uncertainties (DV 2.4.8), and the determination of uncertainty ranges with stochastic and/or deterministic methods (DV 2.4.9)". Deliverables DV 2.4.8 and DV 2.4.9 collectively conclude Project Milestone 2.4, titled "Proof of Concept for Seamless Integration of Projections and Uncertainty Assessment".

4.1 Uncertainty Workflow from DV 2.4.7

With the aim of summarizing the main findings of the previous deliverable on uncertainty topic (DV 2.4.7), the conceived general (i.e., valid for both for marine and terrestrial LEs and the relative working tools), workflow proposed is here reported. The workflow was conceived to provide best practices and concrete steps to deal with the main and different sources of uncertainty to which the ground instabilities are subjected. This workflow was used in this deliverable as a starting point to identify what are the steps already implemented in the three environments and the aspects that needs to be improved instead.

According to the workflow (Figure 1), the first source of uncertainty concerns the available input data and is usually related to measurement techniques. The second source of uncertainty is related to the modelling framework used. Usually, the uncertainty related to all the different kind of available input data should be first identified and evaluated. Then, probabilistic methods can be used to study the propagation of input uncertainties and model uncertainties to arrive at defining a confidence interval for instability.

However, aware that estimating uncertainty ranges for all input data is often unfeasible for lack of data, resource, and time, a sensitivity analysis is proposed as an intermediate step with the aim of identifying the input data and model parameters that are more sensitive to the model framework, i.e., that affect more significantly the model outputs. In doing so, it is possible to identify a subset of input parameters for which uncertainty estimation is more meaningful and worth the use of available resources. For this purpose, a method based on Global Sensitivity Analysis will be presented in more detail in the following paragraph. Additionally, uncertainties (in input, models, and output) should be defined depending on the application and the study scale. High spatial or temporal resolution of measurements might be not necessary when the study focuses on regional scale, whereas in local scale precision and accuracy of measurements matters. In local, regional, and global scales there is a trade-off between "user accuracy" (i.e., showing the reality based on field observations) and "producer accuracy" (based on classification point of view). These thoughts should be properly considered when operationally developing the different tool chains that will constitute the project rationale. Furthermore, clearly stating uncertainties and assumptions is fundamental to avoid potential misunderstandings in the transfer of information to stakeholders and among domains of expertise and communities of practice. In all these cases, Global Sensitivity Analysis is a general and suitable approach to investigate how variations of outputs of a model are linked to variations of input data.

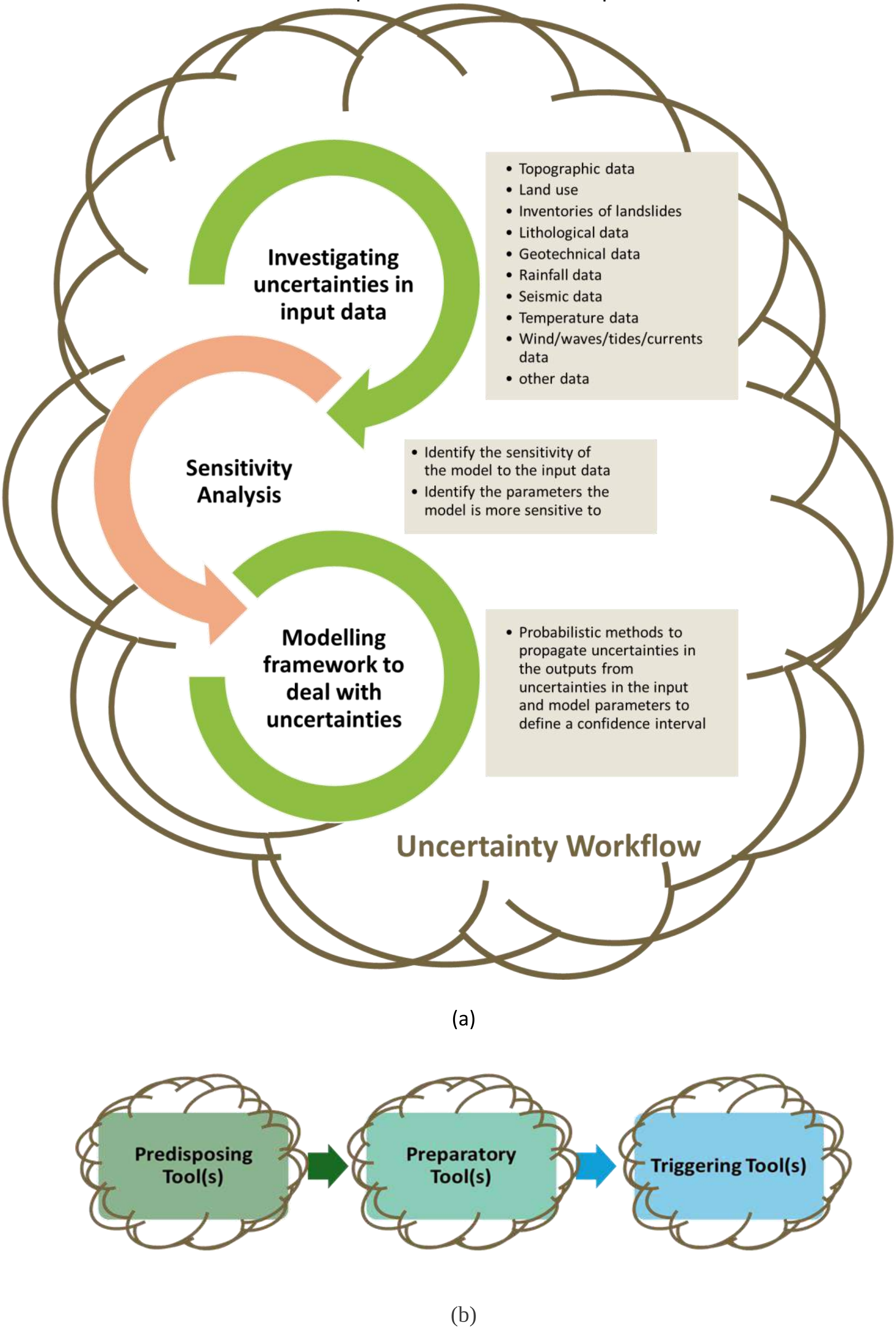

Figure 1 – Uncertainty Workflow introduced in DV 2.4.7: (a) workflow to identify the contexts subjected to uncertainty and (b) the Uncertainty Workflow applied to the tool chains included in the Proof of Concepts

The framework for tool chains dedicated to assessing specific types of ground instabilities, as exemplified in tool chains outlined in DVs 2.3.1, 2.3.3, 2.4.1, 2.4.3, and 2.4.5, entails a sequential arrangement of tools across successive stages. These stages initially consider the predisposition of a given terrain for potential ground instability occurrences, followed by evaluating the preparatory conditions, and ultimately, the triggering mechanisms of such phenomena. The general schema, illustrated in Figure 1a, must be systematically applied to each working tool within a tool chain. Typically, this begins with tool(s) designed to assess predisposition, progressing to those focused on evaluating ground instability preparation and triggering (as depicted in Figure 1b). Only the pertinent input data required by the specific tool within the tool chain should be considered. It is noteworthy that the same input data may be utilized by different tools operating at distinct stages of a tool chain. For instance, slope measurements may serve as input data for tools assessing slope susceptibility to landslide occurrences, as well as for tools evaluating the triggering and propagation of such ground instabilities. Nonetheless, the degree of uncertainty associated with this data may vary between stages due to differences in data sources (e.g., Digital Terrain Models with varying spatial resolutions and accuracy) and the differential sensitivity of each modelling approach to specific parameters.

4.2 Introduction of GSA

4.2.1 What is GSA

Global Sensitivity Analysis (GSA) is a term describing a set of mathematical techniques to investigate how the variation in the output of a numerical model can be attributed to variations of its inputs. GSA can be applied for multiple purposes, including:

- to apportion output uncertainty to the different sources of uncertainty of the model, e.g. unknown parameters, measurement errors in input forcing data, etc. and thus prioritise the efforts for uncertainty reduction;
- to investigate the relative influence of model parameters over the predictive accuracy and thus support model calibration, verification and simplification;
- to understand the dominant controls of a system (model) and to support model-based decision-making.

Mathematically, given the above definitions, we can assume that one can always resort to the general formulation

$$y = g(x) = g(x_1, x_2, \dots, x_n)$$

where y is the output, $x = [x_1, x_2, \dots, x_n]$ is the vector of input factors, which belongs to the input variability space x, and g is the function that maps the input factors into the output, see Figure 2. This input-output relation is sometimes referred to as response surface or model's response, rather than 'model'. Since model's response function g is often unavailable, a numerical procedure is available to evaluate it for any given combination of input factor values (Pianosi et al. 2016)

4.2.2 Types of Sensitivity Analysis

- Local and global sensitivity analysis

Local sensitivity analysis considers the output variability against variations of an input factor around a specific value, while global sensitivity analysis (or GSA) considers variations within the entire space of variability of the input factors (Pianosi et al., 2016).

- Quantitative and Qualitative sensitivity analysis

Quantitative sensitivity analysis refers to methods where each input factor is associated with a quantitative and reproducible evaluation of its relative influence, normally through a set of sensitivity indices (or 'importance measures'). Qualitative sensitivity analysis, instead, sensitivity is assessed qualitatively by visual inspection of model predictions or by specific visualization tools like, for instance, tornado plots, scatter (or dotty) plots or representations of the posterior distributions of the input factors (Pianosi et al., 2016).

- One-At-a-Time (OAT) and All-At-a-Time (AAT)

In OAT methods, output variations are induced by varying one input factor at a time, while keeping all others fixed. In AAT methods, output variations are induced by varying all the input factors simultaneously, and therefore the sensitivity to each factor considers the direct influence of that factor as well as the joint influence due to interactions (Pianosi et al., 2016).

4.2.3 Purposes (settings) of Sensitivity Analysis

- Ranking (or Factor Priorization) aims at generating the ranking of the input factors according to their relative contribution to the output variability.
- Screening (or Factor Fixing) aims at identifying the input factors, if any, which have a negligible influence on the output variability.
- Mapping aims at determining the region of the input variability space that produces significant, e.g. extreme, output values (Pianosi et al., 2016).

4.2.4 How does Global Sensitivity Analysis

Global sensitivity analysis (GSA) unravels the parameter space in order to provide robust sensitivity measures in the presence of nonlinearity and interactions among the parameters compared to the local sensitivity analysis (Wainwright et al., 2014)

GSAs, though robust, can be computationally expensive, because they need sampling parameter sets. Although many approximation models are proposed to reduce the computational cost, they tend to introduce additional model assumptions and response surface fittings, which are not universally applicable.

GSA can be categorised into the Morris design, meta-modelling, regression-based, and variance-based approaches. Many models are used to conduct GSA of process-based crop models, such as Sobol' method, Fourier amplitude sensitivity test (FAST), extended FAST (EFAST), and random-based-design FAST (Rathnappriya et al., 2022)

4.2.5 How does Global Sensitivity Analysis (GSA) work?

Suppose the goal is to test how the uncertainty of four model inputs (or assumptions) influences the variability of the model output.

The input factor is any element that can be changed before running the model. In general, input factors could be equations implemented in the model, set-up choices needed for the model execution on a computer, parameters and input data (such as slope inclination, geotechnical properties…). The input factors could be continuous and discrete variables, or the distribution of an input. The output can be any variable that is obtained after the model's execution (such as the liquefaction potential index or the tsunami amplitude at the coast analyzed in DV 2.4.8).

Before evaluating the model, the inputs will be simulated within their range of variability, followed by running the model so that all four inputs vary simultaneously in each simulation (Input Sampling step). For every output of interest a probability distribution is obtained, after which sensitivity analysis with the method of choice is performed, which allows to obtain a set of sensitivity indices for each output (i.e. one per input, which shows the relative influence input factors have on the output) (Figure 2) (Noacco et al., 2020).

In most of ground instabilities modelling, such as slope instability modelling, the model reliability to represent a specific process can be validate on historical events, but their applications are mostly to forecast future hazards (scenarios based) and then cannot be validated. In such cases, studying how uncertainties of the scenario-based conditions are propagated in the outputs of the models is fundamental to provide meaningful hazard information.

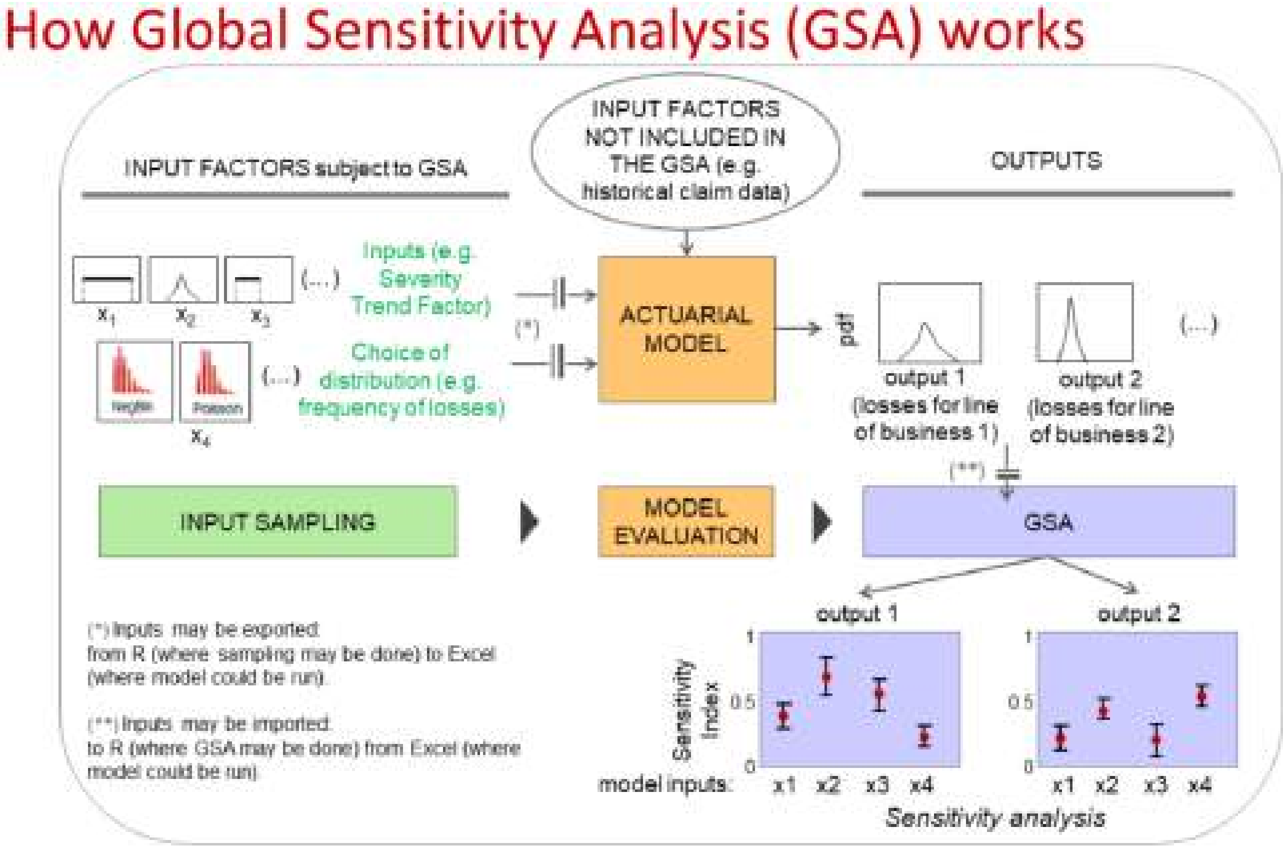

Figure 2 – Steps of how GSA works (Noacco et al., 2020).

4.2.6 Sensitivity Analysis and Uncertainty Analysis

Sensitivity Analysis (SA), especially Global Sensitivity Analysis (GSA), is closely linked to Uncertainty Analysis (UA) in numerical model assessments. UA quantifies output uncertainty (e.g., the variation of liquefaction potential index), while GSA attributes this uncertainty to specific input factor (e.g., the peak ground acceleration) (Saltelli, 2008). Both analyses often use Monte Carlo simulations to propagate uncertainty. Some methods, like the Generalized Likelihood Uncertainty Estimation (GLUE) strategy, were developed from Regional Sensitivity Analysis concepts, highlighting the close relationship between UA and GSA (Beven and Freer, 2001). In practice, they complement each other: GSA benefits from UA to ensure output variability aligns with acceptable behaviour, while UA can incorporate sensitivity indices for added insights with minimal extra effort.

5 Second Chapter: The SAFE toolbox as an operative toolbox to implement Global Sensitivity Analysis

5.1 Introduction of the SAFE toolbox

Pianosi et al. (2015) present a Matlab/Octave toolbox for the application of GSA, called SAFE (Sensitivity Analysis For Everybody). It is specifically designed to conform with several principles that reflect the authors' view on "good practice" in GSA, namely: (i) the application of multiple GSA methods as a means to complement and validate individual results; (ii) the assessment and revision of the user choices made when applying each GSA method, especially in relation to the robustness of the estimated sensitivity indices; and (iii) the use of effective visualisation tools

The first release of the SAFE Toolbox includes the Elementary Effects Test (EET, or Morris method), Regional Sensitivity Analysis (RSA), Variance-Based Sensitivity Analysis (VBSA, or Sobol' method), the Fourier Amplitude Sensitivity Test (FAST), Dynamic identifiability analysis (DYNIA) and a novel density-based sensitivity method (PAWN). The Toolbox also offers several visual tools including scatter (dotty) plots, parallel coordinate plot and the visual test for validation of screening. The basic steps of GSA are shown in Figure 3.

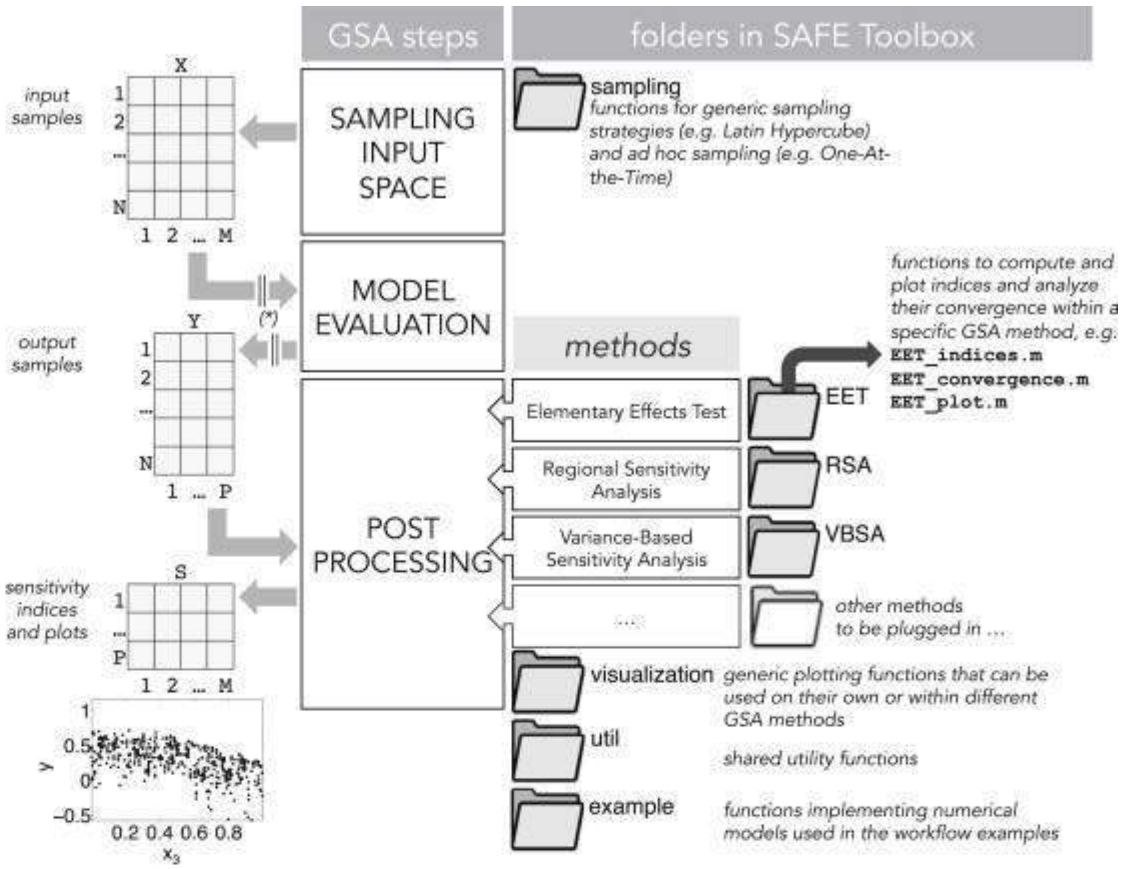

Figure 3 – The three basic steps of GSA. On left hand side of this Figure, the variables that each step takes as input and/or delivers as output: a matrix X of N randomly sampled input combinations (each made up of M components, M being the number of model inputs subject to GSA); a matrix Y of output samples (that can have P > 1 columns when evaluating the sensitivity of multiple model outputs); a matrix S of sensitivity indices. The asterisk indicates where variables may be exported/imported from/into Matlab to another computing environment (Pianosi et al., 2015).

5.2 How does SAFE toolbox work: A Case Study of Uncertainties Analysis in Landslides

As discussed in the DV. 2.4.7, uncertainties in ground instabilities models usually arise from the model parameters, geometric, geotechnic, and hydrologic data, and hazard triggers, like rainfall. These uncertainties may worsen when these models are used to forecast future processes and related hazards due to changing climate and socio-economic conditions, such as urbanization and land use shifts. In such context the use of GSA for ground instabilities modelling result unavoidable nowadays.

Almeida et al. (2017) has already demonstrated how numerical models, using a bottom-up approach, can assess these uncertainties in future landslide hazards. By combining the Combined Hydrology and Stability Model (CHASM) model with sensitivity analysis and Classification and Regression Trees (CART) (Breiman, 2017), they identify key thresholds in slope properties and rainfall that lead to failure. Their study, applied to a Caribbean slope, highlights those uncertainties in topsoil thickness and cohesion are as critical as uncertainties in future rainfall.

The CHASM simulation model is run with 10,000 combinations of 28 uncertain input factors (including slope height, slope angle, thickness of strata, depth of initial water table, saturated hydraulic conductivity, saturated soil moisture content, van Genuchten suction-moisture curve α, van Genuchten suction-moisture curve n^c, residual soil moisture content, initial surface suction, dry unit weight, effective cohesion, effective friction angle), generated through random sampling from probability distributions representing slope properties and wide-ranging uniform distributions for rainfall intensity and duration. A preliminary visual analysis identifies key factors leading to slope failure. Then, Classification and Regression Trees (CART) are used to formally identify factor combinations and threshold values that predict slope failure. The analysis is performed using the Matlab SAFE toolbox and CART functions from the Matlab Statistics and Machine Learning Toolbox.

The 10,000 model simulations were divided into two subsets: those resulting in slope failure (F<1) and those with stable slopes (F>1). Empirical cumulative distribution functions were calculated for each input factor. Figure 4 shows that topsoil thickness, effective cohesion, rainfall intensity, and duration significantly influence slope failure, as indicated by the clear deviation between stable and failure distributions. For other factors, the distributions overlap, suggesting no direct influence on failure, though interactions may still play a role.

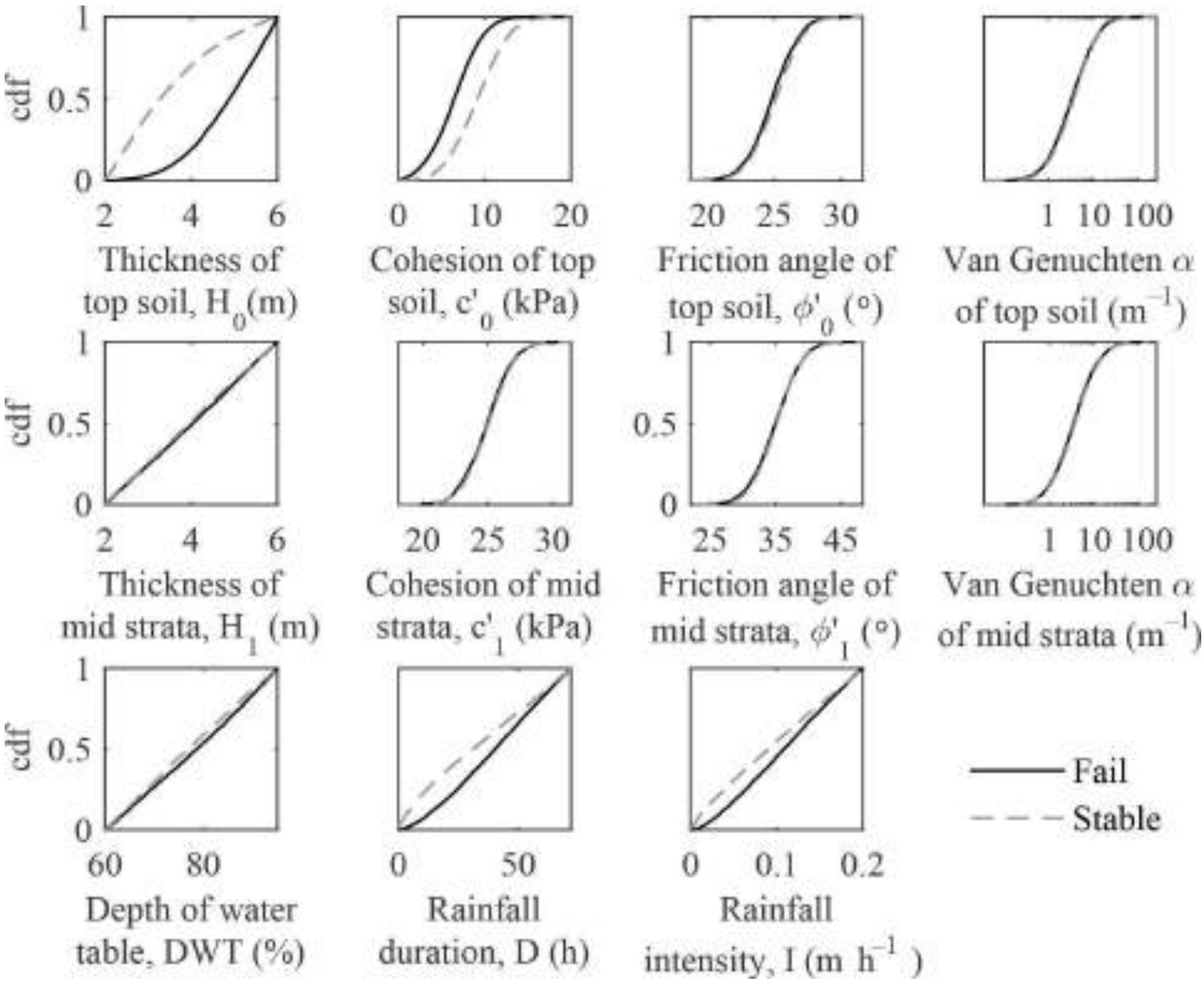

Figure 4 – Cumulative probability distributions (cdf) of slope failure and stability predicted by CHASM for several different input factors. Note that the Van Genuchten suction–moisture curve is shown in logarithmic scale (Almeida et al., 2017).

This study evaluated how uncertainty in slope characteristics and future rainfall changes influence the risk of slope failure. The findings show that for the study site, physical properties such as effective cohesion and topsoil thickness are key factors in slope stability, having a greater impact on landslide risk than variations in future rainfall intensity and duration. Extending this research to other slopes could further explore the interactions between soil depth, permeability, rainfall patterns, and slide depth. The results highlight the importance of considering both slope properties and climate uncertainty together, as neglecting either may significantly underestimate landslide susceptibility. Additionally, the study demonstrates the value of using physically based models, like CHASM, to assess the complex interactions between multiple drivers of landslide occurrence, surpassing the capabilities of simpler statistical models.

6 Third Chapter: How uncertainties matter in the three environments analyzed in DV 2.4.8

In DV 2.4.8, three examples of toolchains developed in the three environments (mountain, plain and marine) were presented. In all the three cases, the authors made a first attempt in applying the Uncertainty Workflow (UW) proposed in DV 2.4.7. As already found in the literature, from these studies emerged that the quantification of uncertainties affecting the numerical/analytical simulation of natural events is usually very complex, due to the many factors entering the assessment.

All the three environments have highlighted the large uncertainty that characterizes the parametrization of input data and model coefficients. In the Marine environment (Chapter 3 of DV 2.4.8), for instance, a key role is played by bathymetric data. Their uncertainty usually relates to the techniques, and thus acoustic parameters, used to acquire the geophysical dataset. When using information derived by swath bathymetry and acoustic sub-bottom profiling data, the uncertainty can be linked to several factors including data resolution. Spatial resolution affects the level of detail that can be captured. Higher resolution provides finer details, but may introduce more noise, while lower resolution reduces detail, potentially missing critical features. Vertical and horizontal resolution affect the ability to distinguish small depth changes and fine features in sub-bottom layers. Coarser resolution can lead to an underrepresentation of complex topography or stratigraphy. In addition to that, simulating marine landslides requires estimating landslide volume reconstruction. This parameter is assessed using morphological considerations which are partly subjective and can introduce a large degree of uncertainty, whose quantification is difficult. Concerning model parameters accounting for the landslide dynamics, the friction coefficient is usually adjusted to fit the observed landslide deposit and the landslide run-out (when available) amplifying bathymetry and landslide volume uncertainties just discussed.

Topographic related uncertainty plays a significant role also in mountain and plain environments to characterize slopes and morphology in general. However, in these environments other factors are also playing a major role in affecting numerical simulations. In particular, the assessment of hydraulic conditions of slopes due to rainfall is a key parameter in modelling slope failure. There is intrinsic uncertainty in rainfall measures. Often rainfall is measured locally and assumed to be homogeneously distributed over an area. It is quite difficult to quantify the potential error of such an assumption, especially in this era of climate change, where rainfall anomalies tend to occur more often than in the past and in small portions of a landscape (Borga et al., 2019, 2022).

Moreover Chapter 1 of DV 2.4.8 has shown how physically based analyses of landslide dynamics are related to the variability of geotechnical (unit volume weight, soil cohesion, root cohesion, friction angle) and geometric (mobilizable soil thickness) parameters. These parameters to be estimated require ad hoc monitoring, which is resource demanding. For modelling purposes and needs often local information is scaled to a larger area introducing a degree of uncertainty which is of difficult estimation. The amount of seepage and then, the water table location is strongly affected too by hydraulic parameters assigned to soil layers and those also play a significant role in landslide modelling.

Chapter 2 of DV 2.4.8 has discussed how the water table plays a key role also in modelling liquefaction in plain environments. Its estimation requires ad hoc monitoring networks and associated modelling, which often are not available. For this reason, analysts use expert based designed scenarios, as in Chapter 2, to assess how uncertainty related in water table values affect liquefaction potential.

Finally making robust assumptions on peak ground acceleration (PGA) is challenging and yet this parameter plays a key role in all the ground instabilities simulations provided in DV 2.4.8.

Adding to these large uncertainties in input data and model parameters estimations, even the choice of the model structure to be used to simulate the process of interest is not always defined. In the plain environment we have seen how the assessment of mechanical parameters to deduce soil resistance are obtained from indirect approaches through empirical formulations, and many alternative methods exist. In the mountain environment “native” outputs of the susceptibility assessment are featured by ROC curves (and related metrics) or Detection Rate Curves (depending on the technique adopted) that allow users to somehow quantify the uncertainty (specificity and sensitivity of the predictive model) related to susceptibility classes once the latter are estimated from the continuous results. The marine case study has described how the computational grid realization itself introduces some approximations, since the codes need a regularly spaced grid that is realized through interpolation of the raw data.

The case studies developed in DV2.4.8 have proven that multiple sources of uncertainties exist and play a pivotal role in our ability to simulate ground instabilities in different environments. In Chapter 1 of DV 2.4.8 landslide modelling displacement considerably changes varying PGA and the semi-empirical relationships linking PGS with Ky (see Figure 5.10 in the DV 2.4.8). Chapter 2 shows that simply creating a few scenarios of ground table levels and PGA plausible values generated significant differences in mapping the potential of liquefaction in the area studied (see Figure 6.11 in the DV 2.4.8). The analysts also comment on the difficulty to state what PGA or water table values would be more realistic, or plausible to assess potential critical situations in the future. In Chapter 3, maximum tsunami amplitude on an Italian coast largely changes simply multiplying the initial mass thickness, whose assessment is characterized by a considerable uncertainty, by a factor ranging the interval 0.25 - 2 (see Figure 7.7 in the DV 2.4.8).

In such contexts, we believe GSA is essential to provide robust simulations of ground instabilities in decision making context. We also care to highlight its relevance to improve our understanding of the processes functioning. The sensitivity analyses provided in the marine case study (Chapter 3 DV 2.4.8) show how changes in the parameters affects model outputs and propose an empirical functions that relate the tsunami amplitude with the volume of the landslides, the slope gradient (underlining the importance of the initial inclination of the landslide, affecting the initial acceleration of the mass and consequently the wave generation), the slide sea-depth and its distance from the coast (equation 1 in Chapter 3 DV 2.4.8). Empirical functions of this type disclose processes functioning which are not obvious and depend on high nonlinear and complex modelling routines.

Addressing uncertainty by GSA forces the analyst to: (i) state input data uncertainty distributions; (ii) identify the most important parameters of the model applications; (iii) provide envelopes of simulations instead of deterministic ones. Such ingredients provide the basis to handle uncertainty in ground instability modelling. In this discipline, as debated so far, uncertainty sources are multiple and notable, and it is not always possible to quantify all of them. Such context should not discourage uncertainty analysis, on the contrary should call for more stringent principles analysts should apply to explicitly declare all uncertainty sources in their model applications, stating clearly if those can be addressed or not. Uncertainty and sensitivity analysis on model outputs should also be required as a standard procedure in science and management applications. Workflow provided in DV 2.4.7 and reported in Figure 1 provide a roadmap to such a goal. Maybe it should be more common to declare that given the input data quality available and or the model adequacy, we are not able to provide robust simulations of specific processes. We admit that this is rarely happening. These principles should be standardized in any applications to advance either in science or management.

7 Fourth chapter: conclusions, limiting factors and future improvements in the RETURN project

In the context of geoscientific applications addressing issues of ground instability, the emphasis placed on uncertainty quantification and sensitivity analysis to identify the most influential parameters affecting the outcomes of various approaches (i.e., tools and toolchains in the frame of RETURN project) has historically been lower compared to other scientific disciplines, such as hydrology (e.g., Song et al., 2015). This disparity highlights a significant gap within the geoscience community; a gap that, given the critical importance of these topics, merits concerted efforts to bridge.

The work conducted as part of deliverables DV 2.4.8 and DV 2.4.9, encompassing both practical applications and theoretical explorations of uncertainty and sensitivity analysis, has underscored several key insights, which are briefly outlined below.

Despite the relatively simplified nature, a few deterministic scenarios generated by the model developers, of the applications conducted in DV 2.4.8, the practical use of the considered tools and toolchains demonstrates that uncertainty associated with input data and model parameters/assumptions can lead to highly variable predictive outputs. This finding arises from analysing quantitative outputs generated under different scenarios across various case studies and environmental settings. However, it is worth noticing that the amount of work to be done to bring the case studies presented here to fully address a proper Global Sensitivity Analysis (GSA), as proposed in the first sections of this DV, is considerable. GSA requires: i) to explicitly parametrize distribution function to assess input uncertainty; ii) to run Monte Carlo simulations of model runs to generate uncertainty envelopes of the model outputs which are statistically robust and, when necessary, including conditional probability between different input data and model parameter uncertainties; iii) to define statistical metrics to quantify how much a certain parameter affects the model results compared to the others. Why doing GSA and what are the benefits in environmental modelling has been largely demonstrated (Wagener et al, 2019). Maybe these are more standard protocols in the realm of hydrology, but ground instability models have been investigated as well (Almeida et al., 2017), as discussed in this deliverable (see section 5.2). Tools and guidelines to applied GSA are available (Pianosi et al., 2016). One example is the SAFE toolbox, discussed in detail in the Second chapter of this deliverable. However, reaching the capacity to apply GSA in Return will require several conditions to be met, including an increased awareness of these themes within the geoscientific community and the enhancement of related technical and scientific competencies.

Reasoning on the Return experience a few limiting factors appear to be relevant: i) uncertainty in input data and model parameters can be significant and the scientists do not always have enough knowledge to properly quantify it; ii) ground instability models are often use to forecast future hazards and quantify the range of plausible future conditions (for predisposing, preparatory, and triggering events) is particularly challenging and subjective; iii) lack of proper expertise in the different geoscientific communities to properly handle uncertainty issues in their modelling exercises.

In geosciences applications, there is a lack of access to datasets due to lack of instruments, costs, weather conditions, etc. In such cases, researchers usually rely on physical and qualitative approaches. On many occasions, there are available measurements, but there are also significant gaps/missing values both temporally and spatially. Therefore, it is not always feasible to perform a reliable sensitivity analysis to determine the most influential input data and parameters. For instance, for the plain environment (Chapter 2 DV 2.4.8) we have seen as a large source of uncertainty concerns the input parameters adopted to characterize the soil resistance. In fact, it can be assessed using three different in-situ investigation techniques, i.e. CPT, SPT, Vs. These methods carry on their own uncertainty in the assessment of mechanical parameters, as the latter are obtained from indirect approaches through empirical formulation.

Another degree of complexity concerns the availability of multiple datasets for the same application. For example, there are several digital elevation models (DEMs) that can be derived for a region of interest, from satellite, field survey, LiDAR, aircraft, etc. If the application is about investigating slope stability, it is important to select a reliable DEM that can also be used for sensitivity analysis. Uncertainty in the DEM measurements, including its spatial resolution, can play a significant role in the analysis. There are also occasions when the data are incomplete, and researchers attempt to apply gap-filling models or data fusion models to reconstruct more reliable data with higher resolution.

Data generation in general in geosciences is experiencing an epochal transformation due to technological advances in the last decade. Our ability to monitor the environment is unprecedented from a range of platforms (e.g., satellites, and drones), and sensors (radar, multi spectral, hyper spectral, Lidar etc..) and that is transforming forever geosciences and our ability to monitor and then model earth processes including ground instabilities (Piégay et al., 2020, and Bizzi et al., 2024). These technological transformations should be grasped as an opportunity to more routinely include uncertainty estimations, nowadays more commonly available, in our data acquisition protocols.

GSA and uncertainty analysis are not new at all in ground instabilities modelling. For landslide susceptibility mapping, we have already discussed in a previous section Almeida et al. (2017) who have applied the SAFE Toolbox for landslide modelling. Gaidzik and Ramírez-Herrera (2021) also demonstrated that the most appropriate input data (e.g., landslide inventory type, raster resolution of topographic data, number of landslide-causing factors) and techniques (i.e., data sampling method) need to be selected after a detailed assessment of the input data, their quality, and resolution, as well as the purpose of the susceptibility mapping. They showed that landslide susceptibility models based on 1 m resolution LIDAR derived digital terrain models were more precise and showed higher prediction accuracy than those developed using 15 m resolution digital elevation models. Furthermore, PREVIEW is a European Commission FP6 Integrated Project with the aim of developing, at a European level, innovative geo-information services for atmospheric, geophysical and man-made risks. Within this framework, the Landslides Platform Service 2 (forecasting of shallow rapid slope movements) has developed an integrated procedure for the forecasting and warning of distributed shallow landslides to be used for civil protection purposes. In this frame, Segoni et al. (2009) carried out a sensitivity analysis on how much each input parameter of the slope stability module weighs in the Factor of Safety value. Their method is based on the partial derivative error propagation method, see Figure 5.

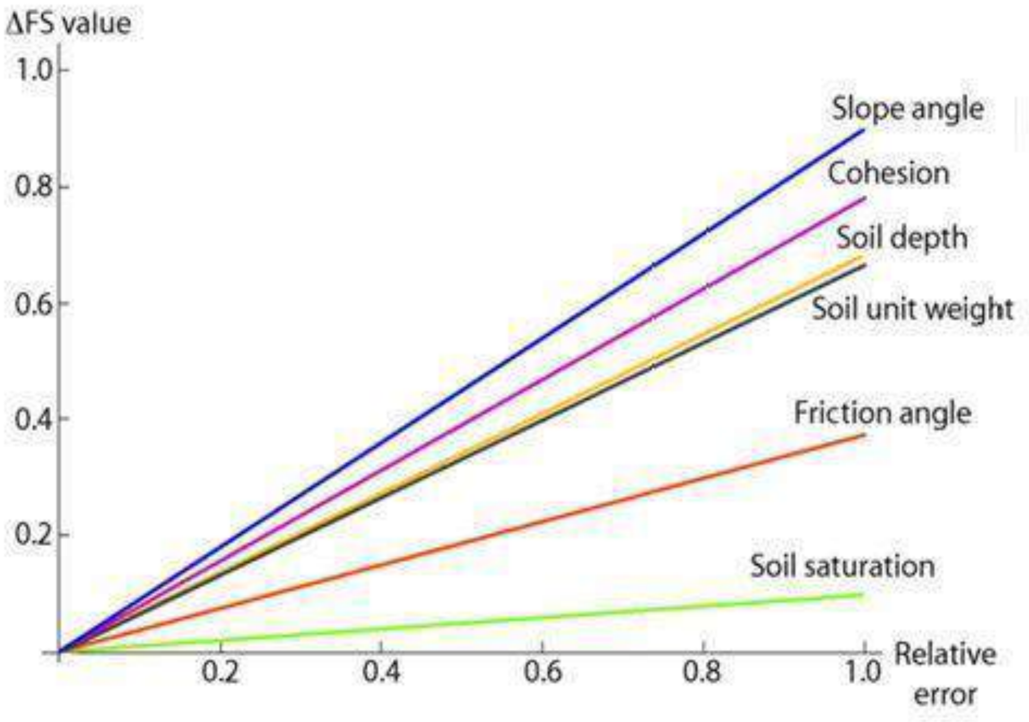

Figure 5 – The graph shows how much the input parameter of the slope stability module weighs in the Factor of Safety value (y-axis).

Kaur et al. (2024) studied the performance of several machine learning algorithms, such as Random Forest (RF), XGBoost, Naive Bayes (NB), and K-Nearest Neighbour (KNN) for landslide susceptibility mapping. They showed that XGBoost with an accuracy of 91% has out-performed other machine learning models, RF (88%), NB (87%), and KNN (82%). Then they performed a SA on the input parameters and showed how they impacted the outputs for their study region. They found in their study region that the inclination of slope, elevation, distance to thrust, road, topographic wetness index (TWI) and slope aspect are the most sensitive factors as small variation in these factors results in great change in the model output, see Figure 6.

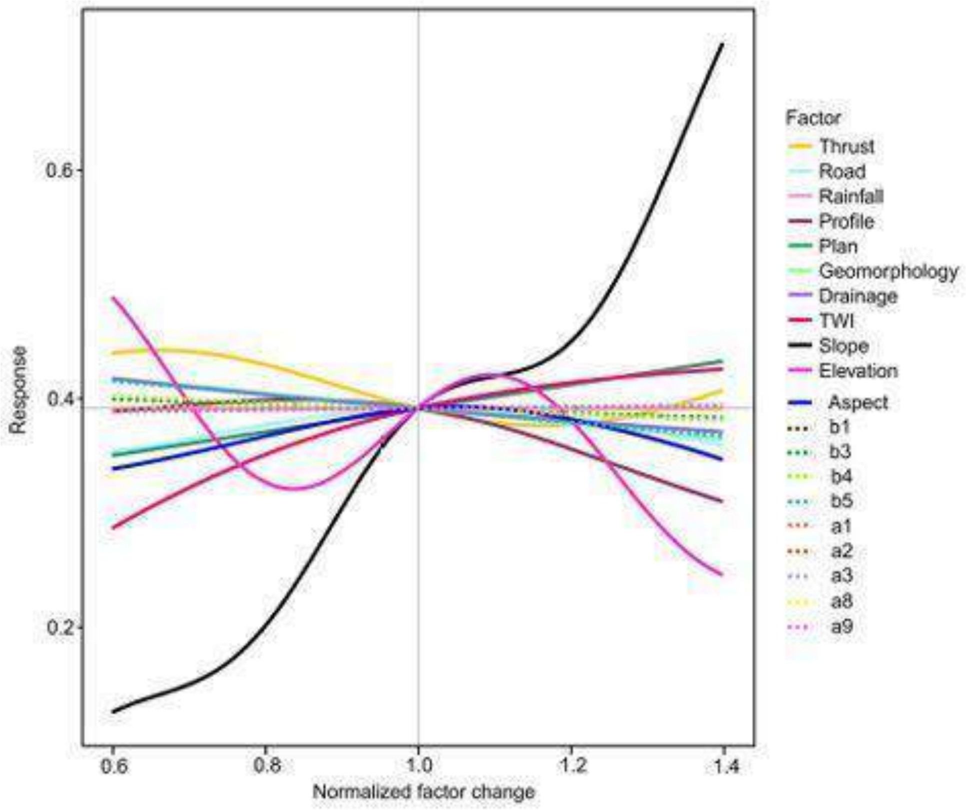

Figure 6 – Sensitivity analysis plot of XGBoost model response against normalized change in each conditioning factor.

The lack of knowledge about input and model parameter uncertainties is particularly evident when considering phenomena such as submarine landslides and seafloor liquefaction, as in offshore environments it is complex to observe gravitational processes before, during and after their occurrences for directly studying their preconditioning, preparatory and triggering factors. Moreover, as demonstrated by recent catastrophic events like the 2018 Palu (Indonesia) tsunami (Omira et al., 2019), where multiple cascading hazards were involved, parameter uncertainty plays a crucial role in hazard assessment (Goda et al., 2019). Indeed, modelling uncertainty is becoming the standard for tsunamis generated by earthquakes, adopting an approach that is similar to the seismic hazard assessment (see for a review Behrens et al., 2021). For tsunamis generated by landslides the procedure is less developed, due also to the different characteristics of the phenomenon itself. A preliminary approach, addressing both landslide-tsunamis as well as earthquake-tsunamis, is presented in the work by Grezio et al. (2012). In this study, the authors have developed a procedure to estimate the probability of exceeding a tsunami run-up of 0.5 meters in one year for the area of the Straits of Messina.

Ground instabilities models providing operative frameworks to address uncertainty exist, however these research experiences are not yet standard procedures. These types of applications require specific expertise in statistics and modelling that is not available in all research groups. Algorithms are often not suitable to implement GSA and uncertainty analysis. Coding efficient algorithms for model simulation is resource demanding, requires specific expertise and is often a neglected task in many research applications. At present, this level of harmonization appears to be incomplete in the tools and toolchains developed as part of VS2 of the RETURN project. During the digitization and further development of these tools, dedicated efforts will be required to achieve, even partially, this objective. Addressing these practical and technical challenges will prepare the way for the consistent application of more advanced GSA methodologies, thereby enhancing the reliability of model predictions and reinforcing the capacity to manage the uncertainties inherent in geoscientific modelling.

Our recommendation reiterates the importance of applying a workflow to address uncertainty as described in Figure 1 and presented in DV 2.4.7. In this DV we have documented the importance of uncertainty analysis in ground instabilities modelling with examples from our own case studies and from literature. Specific guidance and recommendations related to more operatively include GSA in ground instabilities modelling are provided within this deliverable, serving as a roadmap for future developments. By prioritizing these efforts, the geoscientific community can better integrate state-of-the-art uncertainty and sensitivity analysis tools into routine workflows.

In conclusion, we would like to emphasize once again that, in the study of ground instabilities, regardless of the environment considered or the specific type of phenomenon, there is often significant uncertainty associated with the input data and the parameters used in the various models (or tools) applied. The models considered here are designed to provide forecasts, predictions regarding the occurrence and future behaviours of ground instabilities. Consequently, the evaluation of the results obtained cannot generally be validated, e.g. through back analysis procedures (with historical data). A viable solution, therefore, is to rely on expert judgment (as done in the applications reported in DV 2.4.8) to select plausible scenarios and to repeat simulations while varying the input data and model parameters. These scenarios should be, as much as possible, quantified through probabilistic functions, so the outputs will provide envelopes of possible future conditions, providing a wider and more meaningful description of the future hazards. An effective approach to address these challenges and increase the robustness of our results and assessment of the relative uncertainty is represented by Global Sensitivity Analysis, which enables the identification of the most critical data and parameters to focus further efforts on and provides an estimate of the uncertainty associated with the obtained outputs.

8 References

Almeida, S., Holcombe, E.A., Pianosi, F., & Wagener, T. 2017. Dealing with deep uncertainties in landslide modelling for disaster risk reduction under climate change. Natural Hazards and Earth System Sciences, 17, 225-241. https://doi.org/10.5194/nhess-17-225-2017

Bizzi, S., Bozzolan, E., Marra, F., e Cavalli, M. Quantifying earth surface processes via remote sensing technologies. In Climate and Anthropogenic Impacts on Earth Surface Processes in the Anthropocene, edited by Achim A. Beylich, Daniel Vázquez Tarrío, Dongfeng Li, Marc Oliva, e Mario Morellón Marteles, 105–19. Elsevier, 2025. https://doi.org/10.1016/B978-0-443-13215-5.00010-3.

Beven, K., & Freer, J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of hydrology, 249, 11-29.

Borga, M., Comiti, F., Ruin, I., Marra, F., 2019. Forensic analysis of flash flood response. WIRES Water 6 (2), e1338. https://doi.org/10.1002/ wat2.1338.

Borga, M., Marra, F., Gabella, M., 2022. Rainfall estimation by weather radar. In: Morbidelli, R. (Ed.), Rainfall: Modeling, Measurement and Applications. Elsevier

Breiman, L. 2017. Classification and regression trees. Routledge.

Erharter, G. H., Lacasse, S., Tschuchnigg, F., Tentschert, E., Becker, D., & Phoon, K. K. 2024. A consistent terminology to communicate ground-related uncertainty. Engineering Geology, 342, 107744.

Feizizadeh, B., Jankowski, P., & Blaschke, T. 2014. A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. Computers & Geosciences, 64, 81-95.

Gaidzik, K., & Ramírez-Herrera, M.T. 2021. The importance of input data on landslide susceptibility mapping. Scientific Reports, 19334. https://doi.org/10.1038/s41598-021-98830-y

Ghaderpour, E., Masciulli, C., Zocchi, M., Bozzano, F., Scarascia Mugnozza, G., & Mazzanti, P. 2024. Estimating Reactivation Times and Velocities of Slow-Moving Landslides via PS-InSAR and Their Relationship with Precipitation in Central Italy. Remote Sensing. 16, 3055. https://doi.org/10.3390/rs16163055

Ghaderpour, E., Mazzanti, P., Bozzano, F., & Scarascia Mugnozza, G. 2024. Trend Analysis of MODIS Land Surface Temperature and Land Cover in Central Italy. Land, 13, 796. https://doi.org/10.3390/land13060796

Goda, K., Mori, N., Yasuda, T., Prasetyo, A., Muhammad, A., Tsujio, D., 2019. Cascading Geological Hazards and Risks of the 2018 Sulawesi Indonesia Earthquake and Sensitivity Analysis of Tsunami Inundation Simulations. Front. Earth Sci. 7, 261. https://doi.org/10.3389/feart.2019.00261

González-Vida, J.M., Macías, J., Castro, M.J., Sánchez-Linares, C., De La Asunción, M., Ortega-Acosta, S., Arcas, D., 2019. The Lituya Bay landslide-generated mega-tsunami – numerical simulation and sensitivity analysis. Nat. Hazards Earth Syst. Sci. 19, 369–388. https://doi.org/10.5194/nhess-19-369-2019

Grezio, A., Babeyko, A., Baptista, M.A., Behrens, J., Costa, A., Davies, G., Geist, E.L., Glimsdal, S., González, F.I., Griffin, J., Harbitz, C.B., LeVeque, R.J., Lorito, S., Løvholt, F., Omira, R., Mueller, C., Paris, R., Parsons, T., Polet, J., Power, W., Selva, J., Sørensen, M.B., Thio, H.K., 2017. Probabilistic Tsunami Hazard Analysis: Multiple Sources and Global Applications. Reviews of Geophysics 55, 1158–1198. https://doi.org/10.1002/2017RG000579

Grilli, S.T., Taylor, O.-D.S., Baxter, C.D.P., Maretzki, S., 2009. A probabilistic approach for determining submarine landslide tsunami hazard along the upper east coast of the United States. Marine Geology 264, 74–97. https://doi.org/10.1016/j.margeo.2009.02.010

Kaur, R., Gupta, V. & Chaudhary, B.S. 2024. Landslide susceptibility mapping and sensitivity analysis using various machine learning models: a case study of Beas valley, Indian Himalaya. Bull Eng Geol Environ 83, 228. https://doi.org/10.1007/s10064-024-03712-y

Li, T., Zhang, L., Gong, W., & Tang, H. 2023. Cyclic freezing-thawing induced rock strength degradation, crack evolution, heave and settlement accounted for by a DEM model, International Journal of Rock Mechanics and Mining Sciences, 170, 105498. https://doi.org/10.1016/j.ijrmms.2023.105498

Micallef, A., Ribó, M., Canals, M., Puig, P., Lastras, G., Tubau, X., 2014. Space-for-time substitution and the evolution of a submarine canyon-channel system in a passive progradational margin. Geomorphology 221, 34–50. https://doi.org/10.1016/j.geomorph.2014.06.008

Noacco, V., Pianosi, F., & Wagener, T. 2020. Workflow to apply Global Sensitivity Analysis to an actuarial pricing model. In: https://safe-insurance.uk/workflow_GSA_actuarial_example.html.

Omira, R., Dogan, G.G., Hidayat, R., Husrin, S., Prasetya, G., Annunziato, A., Proietti, C., Probst, P., Paparo, M.A., Wronna, M., Zaytsev, A., Pronin, P., Giniyatullin, A., Putra, P.S., Hartanto, D., Ginanjar, G., Kongko, W., Pelinovsky, E., Yalciner, A.C., 2019. The September 28th, 2018, Tsunami In Palu-Sulawesi, Indonesia: A Post-Event Field Survey. Pure Appl. Geophys. 176, 1379–1395. https://doi.org/10.1007/s00024-019-02145-z

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., & Wagener, T. 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software, 79, 214-232.

Pianosi, F., Sarrazin, F., & Wagener, T. 2015. A Matlab toolbox for global sensitivity analysis. Environmental Modelling & Software, 70, 80-85. https://doi.org/10.1016/j.envsoft.2015.04.009

Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., Dufour, S., Liébault, F., Ruiz-Villanueva, V., e Slater, L.. Remotely Sensed Rivers in the Anthropocene: State of the Art and Prospects. Earth Surface Processes and Landforms 45, fasc. 1 (2020): 157–88. https://doi.org/10.1002/esp.4787

Sabeti, R., Heidarzadeh, M., 2022. Numerical Simulations of Tsunami Wave Generation by Submarine Landslides: Validation and Sensitivity Analysis to Landslide Parameters. J. Waterway, Port, Coastal, Ocean Eng. 148, 05021016. https://doi.org/10.1061/(ASCE)WW.1943-5460.0000694

Saltelli, A. 2008. Global Sensitivity Analysis: the Primer. John Wiley & Sons.

Segoni, S., Leoni, L., Benedetti, A. I., Catani, F., Righini, G., Falorni, G., Gabellani, S., Rudari, R., Silvestro, F., & Rebora, N. 2009. Towards a definition of a real-time forecasting network for rainfall induced shallow landslides, Nat. Hazards Earth Syst. Sci., 9, 2119–2133, https://doi.org/10.5194/nhess-9-2119-2009.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., & Xu, C., 2015. Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. Journal of hydrology, 523, 739-757, https://doi.org/10.1016/j.jhydrol.2015.02.013.

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