Addressing challenges in uncertainty quantification. The case of geohazard assessments

Ibsen Chivata Cardenas[1], Terje Aven[1], Roger Flage[1]

[1]Department of Safety, Economics and Planning, University of Stavanger, Stavanger, 4021, Norway

Correspondence to: Ibsen Chivata Cardenas (ibsen.chivatacardenas@uis.no)

Abstract. We analyse some of the challenges in quantifying uncertainty when using geohazard models. Despite the availability of recently developed, sophisticated ways to parameterise models, a major remaining challenge is constraining the many model parameters involved. Additionally, there are challenges related to the credibility of predictions required in the assessments, the uncertainty of input quantities, and the conditional nature of the quantification, making it dependent on the choices and assumptions analysts make. Addressing these challenges calls for more insightful approaches yet to be developed. However, as discussed in this paper, clarifications and reinterpretations of some fundamental concepts and practical simplifications may be required first. The research thus aims to strengthen the foundation and practice of geohazard risk assessments.

1 Introduction

Uncertainty quantification (UQ) helps determine the uncertainty of a system's responses when some quantities and events in such a system are unknown. Using models, the system's responses can be calculated analytically, numerically, or by random sampling (including the Monte Carlo method, rejection sampling, Monte Carlo sampling using Markov chains, importance sampling, and subset simulation) (Metropolis and Ulam, 1949; Brown, 1953; Ulam, 1961; Hastings, 1970). Sampling methods are frequently used because of the high-dimensional nature of hazard events and associated quantities. Sampling methods result in less expensive and more tractable uncertainty quantification than analytical and numerical methods. In the sampling procedure, specified distributions of the input quantities and parameters are sampled, and respective outputs of the model are recorded. This process is repeated as many times as required to achieve the desired accuracy (Vanmarcke, 1984). Eventually, the distribution of the outputs can be used to calculate probability-based metrics, such as expectations or probabilities of critical events. Model-based uncertainty quantification using sampling is now more often used in geohazard assessments, e.g., Uzielli and Lacasse (2007), Wellmann and Regenauer-Lieb (2012), Rodríguez-Ochoa et al. (2015), Pakyuz-Charrier et al. (2018), Huang et al. (2021), Luo et al. (2021), Sun et al. (2021a).

This paper considers recent advances in UQ and analyses some remaining challenges. For instance, we note that a major problem persists, namely constraining the many parameters involved. Only some parameters can be constrained in practice based solely on historical data (e.g., Albert, Callies, and von Toussaint, 2022). Another challenge is that model outputs are conditional on the choice of model parameters and the specified input quantities, including initial and boundary conditions. For example, a geological system model could be specified to include some geological boundary conditions (Juang et al., 2019). Such systems are usually time-dependent and spatial in nature and may involve, e.g., changing conditions (e.g., Chow, Li, and Koh, 2019). Incorporating uncertainties related to such conditions complicates the modelling and demands further data acquisition. Next, models could accurately reproduce data from past events but may be inadequate for unobserved outputs or predictions. This might be the case when predicting, e.g., extreme velocities in marine turbidity currents, which are driven by emerging and little-understood soil and fluid interactions (Vanneste et al., 2019). Overlooking these challenges implies that the quantification will only reflect some aspects of the uncertainty involved. These challenges are, unfortunately, neither exhaustively nor clearly discussed in the geohazard literature. Options and clarifications addressing these challenges are underreported in the field. Analysing these challenges can be useful in treating uncertainties consistently and providing meaningful results in an assessment. This paper's objective is to bridge the gap in the literature by providing an analysis and clarifications enabling a useful quantification of uncertainty.

It should be emphasised that, in this paper, we consider uncertainty quantification in terms of probabilities. Other approaches to measure or represent uncertainty have been studied by, for example, Zadeh (1968), Shafer (1976), Ferson and Ginzburg (1996), Helton and Oberkampf (2004), Dubois (2006), Aven (2010), Flage et al. (2013), Shortridge, Aven, and Guikema (2017), Flage, Aven, and Berner (2018), and Gray et al. (2022a,b). These approaches will not be discussed here. The discussion about the complications in UQ related to computational issues generated by sampling procedures is also beyond the scope of the current work.

The remainder of the paper is as follows. In Section 2, based on recent advances, we describe how uncertainty quantification using geohazard models can be conducted. Next, some remaining challenges in UQ are identified and illustrated. Options to address the challenges in UQ are discussed in Section 3. A simplified example, further illustrating the discussion, is found in Section 4, while the final section provides some conclusions.

2 Quantifying uncertainty using geohazard models

In this section, we make explicit critical steps in uncertainty quantification (UQ). We describe a general approach to UQ that considers uncertainty as the analysts' incomplete knowledge about quantities or events. The UQ approach described is restricted to probabilistic analysis. Emphasis is made on the choices and assumptions usually made by analysts.

A geohazard model can be described as follows. We consider a system (e.g., debris flow) with a set of specified input quantities $\boldsymbol{X}$ (e.g., sediment concentration, entrainment rate) whose relationships to the model output $\boldsymbol{Y}$ (e.g., runout volume, velocity, or height of flow) can be expressed by a set of models $\boldsymbol{\mathcal{M}}$. Analysts *identify* or *specify* $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{\mathcal{M}}$. A vector $\boldsymbol{\Theta}_{m}$ (including, e.g., friction, viscosity, turbulence coefficients) parameterises a model m in $\boldsymbol{\mathcal{M}}$. The parameters $\boldsymbol{\Theta}_{m}$ determine specific functions among a family of potential functions modelling the system. Accordingly, a model m can be described as a multi-output function with, e.g., $\boldsymbol{Y}$ = {runout volume, velocity, height of flow}. Based on Lu and Lermusiaux (2021), we can write:

$$m: \boldsymbol{X}_{s,t} \times \boldsymbol{\Theta}_{m} \rightarrow \boldsymbol{Y}_{s,t} \quad (1)$$

$$m \equiv (\boldsymbol{E}_{m}, \boldsymbol{SG}_{m}, \boldsymbol{BC}_{m}, \boldsymbol{IC}_{m}) \quad (2)$$

Realisations of $\boldsymbol{Y}$ are the model responses y when elements in $\boldsymbol{X}$ take the values x at a spatial location $s \in \mathbf{S}$ and a specific time $t \in \boldsymbol{T}$, and parameters $\theta_{m} \in \boldsymbol{\Theta}_{m}$ are used. In expression (1), $\boldsymbol{X} \subset \mathbb{R}^{d_{\boldsymbol{X}}}$ is the set of specified input quantities, $\boldsymbol{T} \subset \mathbb{R}^{d_{\boldsymbol{T}}}$ is the time domain, $\boldsymbol{S} \subset \mathbb{R}^{d_{\boldsymbol{S}}}$ is the spatial domain, $\boldsymbol{\Theta}_{m} \subset \mathbb{R}^{d_{\boldsymbol{\Theta}_{m}}}$ corresponds to a parameter vector, and $\boldsymbol{Y} \subset \mathbb{R}^{d_{\boldsymbol{Y}}}$ is the set of model outputs. To consider different dimensions, d = {1, 2, or 3}. The system is fully described if m is *specified* in terms of a set of equations $\boldsymbol{E}_{m}$ (e.g., conservation equations), the spatial domain geometry $\boldsymbol{SG}_{m}$ (e.g., extension, soil structure), the boundary conditions $\boldsymbol{BC}_{m}$ (e.g., downstream flow), and the initial conditions $\boldsymbol{IC}_{m}$ (e.g., flow at $t = t_0$), see expression (2).

Probabilities reflecting analysts' uncertainty about input quantities are specified in uncertainty quantification. Such distributions are then sampled many times, and the distribution of the produced outputs can be calculated. The output probability distribution for a model m can be denoted as $f(y|x,\theta_{m},m)$, for realisations y, x, θ_{m}, m of $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{\Theta}_{m}$, and $\boldsymbol{\mathcal{M}}$, respectively.

Betz (2017) has suggested that the parameter set is fully described by a parameter vector $\boldsymbol{\Theta}$, Eq. (3):

$$\boldsymbol{\Theta} = \{\boldsymbol{\Theta}_{m}, \boldsymbol{\Theta}_{X}, \boldsymbol{\Theta}_{\varepsilon}, \boldsymbol{\Theta}_{o}\} \quad (3)$$

in which, $\boldsymbol{\Theta}_{m}$ refers to parameters of the model m, $\boldsymbol{\Theta}_{\boldsymbol{X}}$ are parameters linked to the input $\boldsymbol{X}$, $\boldsymbol{\Theta}_{\varepsilon}$ is the vector of the output-prediction error ε, and $\boldsymbol{\Theta}_{o}$ is the vector associated with observation/measurement errors. More explicitly, to compute an overall joint probability distribution, we may have the following distributions:

- $f(y|x,\theta_{m},m)$ is the distribution of $\boldsymbol{Y}$ when $\boldsymbol{X}$ takes the values x, and parameters $\theta_{m} \in \boldsymbol{\Theta}_{m}$ and a model $m \in \boldsymbol{\mathcal{M}}$ are used to compute y;
- $f(x|\theta_{X},m)$ is the conditional distribution of $\boldsymbol{X}$ given the parameters $\theta_{\boldsymbol{X}} \in \boldsymbol{\Theta}_{\boldsymbol{X}}$ and the model m. Note that each m defines which elements in $\boldsymbol{X}$ are to be considered in the analysis;
- $f(x|\hat{x},\theta_{o})$ is a distribution of $\boldsymbol{X}$ given the observed values $\hat{\boldsymbol{X}} = \hat{x}$ and the observation/measurement error parameters $\theta_{o} \in \boldsymbol{\Theta}_{o}$;

- additionally, one can consider $f(y^*|y,\theta_\varepsilon,\mathfrak{m})$, which is a distribution of $\boldsymbol{Y}^*$, the future system's response, conditioned on the model output y and the output-prediction error vector $\theta_\varepsilon \in \boldsymbol{\Theta}_\varepsilon$. The output-prediction error ε is the mismatch between the model predictions and non-observed system responses y^*. ε is used to correct the imperfect model output y (Betz, 2017; Juang et al., 2019).

If, for example, the parameters $\boldsymbol{\Theta}_{\mathfrak{m}}$ are poorly known, a prior distribution $\pi(\theta_{\mathfrak{m}}|\mathfrak{m})$ weighing each parameter value $\theta_{\mathfrak{m}}$ for a model $\mathfrak{m}$ is usually *specified*. A prior is a subjective probability distribution quantified by expert judgement representing uncertainty about the quantities prior to considering data (Raices-Cruz, Troffaes, and Sahlin, 2022). When some measurements $\mathcal{d} = \{\hat{\boldsymbol{Y}} = \hat{y}, \hat{\boldsymbol{X}} = \hat{x}\}$ are available, such parameter values $\theta_{\mathfrak{m}}$, or their distributions $\pi(\theta_{\mathfrak{m}}|\mathfrak{m})$, can be constrained by back-analysis methods. Note that measurements $\mathcal{d}$ form part of different sources of data $\boldsymbol{\mathcal{D}}$, i.e., $\mathcal{d} \in \boldsymbol{\mathcal{D}}$. Back analysis methods include matching experimental measurements $\hat{y}$ and calculated model outputs y using different assumed values $\theta'_{\mathfrak{m}}$. Values for $\theta_{\mathfrak{m}}$ can be calculated as follows (based on Liu et al., 2022):

$$\theta_{\mathfrak{m}} = argmin[\hat{y}\text{-}y(\hat{x},\theta'_{\mathfrak{m}})] \quad (4)$$

The revision or updating of the prior $\pi(\theta_{\mathfrak{m}}|\mathfrak{m})$ with measurements $\mathcal{d}$ to obtain a posterior distribution denoted $\pi(\theta_{\mathfrak{m}}|\mathcal{d},\mathfrak{m})$ is also an option in back analysis. The updating can be calculated as follows (based on Juang et al., 2019; Liu et al., 2022):

$$\pi(\theta_{\boldsymbol{m}}|\mathcal{d},\mathfrak{m}) = \frac{\boldsymbol{\mathcal{L}}(\theta_{\boldsymbol{m}}|\mathcal{d})\pi(\theta_{\boldsymbol{m}}|\mathfrak{m})}{\int \boldsymbol{\mathcal{L}}(\theta_{\boldsymbol{m}}|\mathcal{d})\pi(\theta_{\boldsymbol{m}}|\mathfrak{m})\mathrm{d}\theta_{\boldsymbol{m}}} \quad (5)$$

where $\mathcal{L}(\theta_m|\mathcal{d}) = f(\mathcal{d}|\theta_m)$ is a likelihood function, i.e., a distribution that weighs $\mathcal{d}$ given $\theta_{\mathfrak{m}}$.

Similarly, we can constrain any of the distributions above, e.g., $f(y|x,\theta_{\mathfrak{m}},\mathfrak{m})$, or $f(x|\theta_X,\mathfrak{m})$ to obtain $f(y|x,\theta_{\mathfrak{m}},\mathcal{d},\mathfrak{m})$ and $f(x|\theta_X,\mathcal{d},\mathfrak{m})$, respectively.

For a geohazard problem, it is often possible to *specify* several competing models, e.g., distinct geological models with diverse boundary conditions, see Eq. 2. If the available knowledge is insufficient to determine the best model, different models $\mathfrak{m}$ can be considered. The respective overall output probability distribution is computed as (Betz, 2017; Juang et al., 2019):

$$f(y|x,\boldsymbol{\Theta},\boldsymbol{\mathcal{D}},\boldsymbol{\mathcal{M}}) = \sum f(y|x,\theta,\mathcal{d},\mathfrak{m})\omega(\mathfrak{m}|\boldsymbol{\mathcal{D}},\boldsymbol{\mathcal{M}}) \quad (6)$$

$$f(y|x,\theta,\mathcal{d},\mathfrak{m}) = \int f(y|x,\theta,\mathfrak{m})\pi(\theta|\mathcal{d},\mathfrak{m})d\theta \quad (7)$$

In Eq. (6), $\omega(\mathfrak{m}|\boldsymbol{\mathcal{D}},\boldsymbol{\mathcal{M}})$ is a distribution weighing each model $\mathfrak{m}$ in $\boldsymbol{\mathcal{M}}$.

The various models $\boldsymbol{\mathcal{M}}$, their inputs $\boldsymbol{X}$, parameters $\boldsymbol{\Theta}$, outputs $\boldsymbol{Y}$, and experimental data $\mathcal{d}$ can be coupled all together through a Bayesian network, as has been suggested by Sankararaman and Mahadevan (2015) or Betz (2017). One possible configuration of a network coupling some elements in $\boldsymbol{\mathcal{M}}$, $\boldsymbol{X}$, $\boldsymbol{\Theta}$, $\boldsymbol{Y}$, $\boldsymbol{Y^*}$ is illustrated in Figure 1.

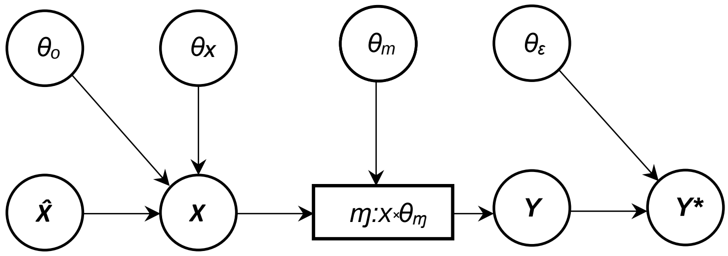


Figure 1: A configuration of a network coupling some elements in $\boldsymbol{\mathcal{M}}$, $\boldsymbol{X}$, $\boldsymbol{\Theta}$, $\boldsymbol{Y}$, $\boldsymbol{Y^*}$

The previous description of a general approach to UQ considers uncertainty as that reflected in the analysts' incomplete knowledge about quantities or events. In UQ, to measure or describe uncertainty, subjective probabilities can be used and constrained using observations $\mathcal{d}$. It is also explicitly shown that model outputs are conditional on observations $\mathcal{d}$ made available and models $\boldsymbol{\mathcal{M}}$ chosen by analysts. Analysts might also select several parameters $\boldsymbol{\Theta}$ and initial and boundary conditions, $\boldsymbol{BC}_{m}$ and $\boldsymbol{IC}_{m}$. Based on the above description, in the following, we analyse some of the challenges that arise when conducting UQ.

As mentioned, back-analysis methods help constrain some elements in $\boldsymbol{\Theta}$. However, given the considerable number of parameters (see expressions 1-3) and data scarcity, constraining $\boldsymbol{\Theta}$ is often only achieved in a limited fashion. Back-analysis is further challenged by the potential dependency among $\boldsymbol{\Theta}$ or $\boldsymbol{\mathcal{M}}$ and between $\boldsymbol{\Theta}$ and $\boldsymbol{SG}_{m}$, $\boldsymbol{BC}_{m}$, $\boldsymbol{IC}_{m}$. We also note that back analysis, or, more specifically, inverse analysis, faces problems regarding non-identifiability, non-uniqueness, and instability. Non-identifiability occurs when some parameters do not drive changes in the inferred quantities. Non-uniqueness arises because more than one set of fitted or updated parameters may adequately reproduce observations. Instability in the solution arises from errors in observations and the non-linearity of models (Carrera and Neuman, 1986). Alternatively, in specifying a joint distribution $f(x,\theta)$ to be sampled, analysts may consider the use of, e.g., Bayesian networks (Albert, Callies, and von Toussaint 2022). However, under the usual circumstance of a lack of information, establishing such a joint distribution is challenging and requires that analysts encode many additional assumptions (e.g., prior distributions, likelihood functions, independence, linear relationships, normality, stationarity of the quantities and parameters considered), see e.g., Tang, Wang, and Li (2020); Sun et al. (2021b); Albert, Callies, and von Toussaint (2022); Pheulpin, Bertrand, and Bacchi (2022). A more conventional choice is that x or θ are *specified* using the maximum entropy principle (MEP), to specify the least *biased* distributions possible on the given information (Jaynes, 1957). Such distributions are subject to the system's physical constraints based on some available data. The information entropy of a probability distribution measures the amount of information contained in the distribution. The larger the entropy, the less information is provided by the distribution. Thus, by maximising the entropy over a suitable set of probability distributions, one finds the least informative distribution in the sense that it contains the least amount of information consistent with the system's constraints. Note that a distribution is sought over all the candidate distributions subject to a set of constraints. The MEP has been questioned since its validity and usefulness lie in the proper *choice* of physical constraints (Jaynes, 1957; Yano 2019). Doubts are also raised regarding the potential information loss when using the principle. Analysts usually strive to use all available knowledge and avoid unjustified information loss (Christakos, 1990; Flage, Aven, and Berner, 2018).

Options to address the parametrisation challenge also include surrogate models, parameter reduction, and model learning (e.g., Lu and Lermusiaux, 2021; Sun et al., 2021b; Albert, Callies, and von Toussaint, 2022; Degen et al. 2022; Liu et al., 2022). Surrogate models are learnt to replace a complicated model with an inexpensive and fast approximation. Parameter reduction is achieved based on either principal component analysis or global sensitivity analysis to determine which parameters significantly impact model outputs and are essential to the analysis (Degen et al., 2022; Wagener, Reinecke, and Pianosi, 2022). Remarkably, versions of the model learning option do not need any prior information about model equations $\boldsymbol{E}_{m}$ but require local verification of conservation laws in the data d (Lu and Lermusiaux, 2021). These approaches still require large data sets sourced systematically, which is a frequent limitation in geohazard assessments. More importantly, however, is that, like many models, the *credibility* of unobserved surrogate model outputs can always be questioned, since, for instance, records may miss crucial events (Woo, 2019). Models may also fail to reproduce outputs caused by recorded abrupt changes (e.g., extreme velocities of turbidity currents) (Alley, 2004). An additional point is the issue of incomplete model response, which refers to a model not having a solution for some combinations of the specified input quantities (Cardenas, 2019; van den Eijnden, Schweckendiek, and Hicks, 2021).

In bypassing the described challenges when quantifying uncertainty, simplifications are usually enforced, sometimes unjustifiably, in the form of assumptions, denoted here by $\mathcal{A}$. The set $\mathcal{A}$ can include one or more of the assumptions listed in Table 1. Note that the set of assumptions can be increased with those assumptions imposed by using specific models $\mathcal{M}$ (e.g., conservation of energy, momentum, or mass, Mohr-Coulomb's failure criterion).

Table 1. Some enforced assumptions in UQ for geohazard assessments

Predictions (non-observed outputs) of $\boldsymbol{Y^*}$ are credible, despite models only reproducing responses based on historical data $d=\{\hat{\boldsymbol{Y}}=\hat{y},\boldsymbol{X}=\hat{x}\}$, $d \in \boldsymbol{\mathcal{D}}$.
A model has a solution for any combination of the specified input quantities $\boldsymbol{X}$.
Elements in $\boldsymbol{X}$ are fully specified.
Elements in $\boldsymbol{X}$ are mutually independent.
The joint distribution $f(x,\theta)$ distributes according to the maximum entropy principle.
If measurements are available, some specified input quantities $\boldsymbol{X}$ are set to specific values $x=\hat{x}$.
Specified input quantities $\boldsymbol{X}$ are set to constant values x_0, that is $\boldsymbol{X}=x_0$.
Some θ are set to specific point values and are mutually independent.
Some θ are independent of $\boldsymbol{SG}_{m}$, $\boldsymbol{BC}_{m}$, $\boldsymbol{IC}_{m}$.
$\boldsymbol{SG}_{m}$, $\boldsymbol{BC}_{m}$, $\boldsymbol{IC}_{m}$ are set to be constant.
When some data d is available in the form of measurements $\{\hat{y}, \hat{x}\}$, likelihood functions $\mathscr{L}[m(\theta\|d)]$ are mutually independent.

3 Addressing the challenges in uncertainty quantification

From the previous section, we saw that it is very difficult in geohazard assessments to meet data requirements for the ideal parameterisation of models. Further, we have noted that, although fully parameterised models could potentially be accurate at reproducing data from past events, these may turn out to be inadequate for unobserved outputs. We also made explicit that predictions are not only conditional on $\boldsymbol{\Theta}$ but possibly also on $\boldsymbol{SG}_{m}$, $\boldsymbol{BC}_{m}$, $\boldsymbol{IC}_{m}$, see Eq. (1-7). Ultimately, assumptions made also condition model outputs. More importantly, note that when only some model input quantities or parameters can be updated using data d, the modelling will only reflect some aspects of the uncertainty involved. If the above challenges remain unaddressed, UQ lacks credibility. To address such challenges and provide increased credibility, clarifications and reinterpretation of some fundamental concepts and practical simplifications may be required, which are discussed in the following. Table 2 shows the major challenges found and how they are addressed in related literature, while in Table 3, some clarifications or considerations put forward by us are displayed. The discussion in this section builds on previous analysis by Aven and Pörn (1998), Apeland, Aven, and Nilsen (2002), Aven and Kvaløy (2002), Nilsen and Aven (2003), Aven and Zio (2013), Khorsandi and Aven (2017), and Aven (2019).

Table 2. Major challenges and options to address them in geohazard assessments

Challenges, CH	Options to address the challenges, O
Challenges related to the model outputs and system responses	
CH1. *Model outputs* ***Y*** *lack credibility* since these are outputs not recorded in the data $\mathscr{d}$, ***CH2***. A model does *not have a solution* for a feasible *combination of the specified input quantities* ***X***.	***O1***. *Credibility* of predictions is *judged* in terms of physical consistency checks (Wagener, Reinecke, and Pianosi, 2022) and by examining the ability of models to reproduce disruptive changes recorded in the data (Alley, 2004). ***O2***. Predictions by *Bayesian forecasting methods*. Based on a prior distribution for *y*, a posterior distribution of *y* is obtained by including the information provided by the model prediction in the form of model likelihood (Montanari and Koutsoyiannis, 2012).
Challenges related to input quantities	
CH3. Data available $\mathscr{d}$ may *not include all the crucial historical events or disruptive changes,* ***CH4***. Some input quantities remain *unknown (unidentified)* to analysts during an assessment, ***CH5***. *The distribution f(x) or the bounds* of *x* are *unknown,* ***CH6***. Some input quantities ***X*** may be mutually *dependent.*	***O3***. Using the *maximum entropy principle (MEP)* to specify the distributions based on the *choice* of physical constraints of the phenomena involved. To reduce unjustified information loss, constraining the distributions by data including data other than measurements (Jaynes, 1957; Christakos, 1990; Betz, 2017; Yano 2019). ***O4***. *Counterfactual analysis* in which alternative events to observed facts, including disruptive changes, are *assumed* to obtain alternative system responses using models (Pearl, 1993; Woo, 2019). ***O5***. To specify input distributions, an exhaustive investigation of input uncertainty using the *assumptions deviation approach* (Aven, 2013).
Challenges related to the parameters and models	
CH7. The *distribution f(θ)* or the *bounds* of *θ* are *unknown,* ***CH8***. Some *θ* may be *dependent* on $\boldsymbol{SG_{\mathscr{m}}}$, $\boldsymbol{BC_{\mathscr{m}}}$, *or* $\boldsymbol{IC_{\mathscr{m}}}$, ***CH9***. *Likelihood functions* $\mathscr{L}[\mathscr{m}(\theta\vert\mathscr{d})]$ may be mutually *dependent,* ***CH10***. *Models* $\mathscr{m}$ *in* $\boldsymbol{\mathscr{M}}$ may be mutually *correlated.*	***O3***. Using the *MEP* as described above. ***O6***. A joint distribution of $\boldsymbol{\Theta}$, $\boldsymbol{SG_{\mathscr{m}}}$, $\boldsymbol{BC_{\mathscr{m}}}$, $\boldsymbol{IC_{\mathscr{m}}}$, ***X*** for each $\mathscr{m}$, can be specified by encoding other assumptions (e.g., prior distributions, likelihood functions, independence, linear relationships, normality, stationarity) in *Bayesian networks* (Albert, Callies, and von Toussaint, 2022). ***O7***. Using *surrogate models*, *parameters reduction*, and *model learning* (Lu and Lermusiaux, 2021; Albert, Callies, and von Toussaint, 2022).

Table 3. Some clarifications and considerations to address the challenges in UQ

C1. *Uncertainty* refers to lack of knowledge about *quantities* or *events*.

C2. Models are *simplifications*, *mainly used for understanding* the performance of the system and *approximating* its *responses*. Models are part of the knowledge of the system, and they do not introduce uncertainty.

C3. The focus is on quantifying the *uncertainty of the system responses* rather than on the accuracy of a model reproducing recorded data.

C4. Predictions are *conditional* on the *model(s)* chosen and the *assumptions* made by analysts.

C5. The *specification* of the joint distribution *f(x,θ) cannot solely rely on* the use *of the maximum entropy principle* but on the full scrutiny of *background knowledge* $\boldsymbol{\mathscr{K}}$.

C6. Some elements in the parameter set $\boldsymbol{\Theta}$ *are not properties of the system* as such, and there could not be uncertainty about them.

C7. Analysts may *choose* a model or a set of models which are *believed* or *judged* to be the *best credible models*.

Among the clarifications, we consider a major conceptualisation suggested by the literature, which is the definition of uncertainty. Uncertainty refers to incomplete information or knowledge about a quantity or the occurrence of an event (Society for Risk Analysis, 2018). In Table 3, we denote this clarification as *C1*. Embracing this definition has some implications for uncertainty quantification using geohazard models. We use these implications to address the major complications and challenges. For instance, if uncertainty is measured in terms of probability, one such implication is that analysts are discouraged from using so-called frequentist probabilities. We note that frequentist probabilities do not measure uncertainty or lack of knowledge. Rather such probabilities reflect frequency ratios representing fluctuation or variation in the outcomes of quantities. Frequentist probabilities are of limited use because these assume that quantities vary in large populations of identical settings, a condition which can be justified only for rather few geohazard quantities. The often one-off nature of many geohazard features and the impossibility of verifying or validating data by, e.g., a large number of repeated tests, make it difficult to develop such probabilities. Thus, a more meaningful and practical approach suggests to measure uncertainty by the use of knowledge-based (also referred to as judgemental or subjective) probabilities (Aven 2019). A knowledge-based probability is an expression of the degree of belief in the occurrence of an event or quantity by a person assigning the probability conditional on the available knowledge $\boldsymbol{K}$. Such knowledge $\boldsymbol{K}$ includes not only data in the form of measurements d made available, but also other data sources in $\boldsymbol{D}$. The models $\boldsymbol{M}$ chosen for the prediction and the modelling assumptions $\boldsymbol{A}$ made by analysts are also part of $\boldsymbol{K}$. Accordingly, to describe uncertainty about quantities, probabilities are assigned based on $\boldsymbol{K}$ and, therefore, those probabilities are conditional on $\boldsymbol{K}$. In the previous section, we have made evident the conditional nature of the uncertainty quantification (i.e., the probabilities) on measured data d and models $\boldsymbol{M}$, and wrote the expression $f(y|x,\boldsymbol{\Theta},\boldsymbol{D},\boldsymbol{M})$ for the overall output probability distribution (see Eq. 6). If assumptions $\boldsymbol{A}$ are also acknowledged as a conditional argument of the uncertainty quantification, we write more explicitly $f(y|x,\boldsymbol{\Theta},\boldsymbol{D},\boldsymbol{M},\boldsymbol{A})$, or equivalently $f(y|x,\boldsymbol{\Theta},\boldsymbol{K})$. We can therefore write:

$$f(y|x,\boldsymbol{\Theta},\boldsymbol{K}) = f(y|x,\boldsymbol{\Theta},\boldsymbol{D},\boldsymbol{M},\boldsymbol{A}) \quad (8)$$

The meaning of this expression is explained next. If, in a specific case, we would write $f(y|x,\boldsymbol{\Theta},\boldsymbol{K}) = f(y|x,\theta,\boldsymbol{D})$, it means that $\boldsymbol{D}$ summarises all the knowledge that analysts have to calculate y given (realised or known) x and θ. Accordingly, the full expression in Eq. 8 implies that to calculate y, and given knowledge of x and θ, the background knowledge includes $\boldsymbol{D}$, $\boldsymbol{M}$, $\boldsymbol{A}$**.** Note that $\boldsymbol{K}$ can also be formed by observations, justifications, rationales, and arguments, thus, Eq. 8 can be further detailed to include these aspects of $\boldsymbol{K}$. Structured methods exist to assign knowledge-based probabilities (see, e.g., Apeland, Aven, and Nilsen, 2002; Aven 2019). Here we should note, however, that since models form part of the available background knowledge $\boldsymbol{K}$, models can also inform these knowledge-based probability assignments. It follows that, based on knowledge-based input probabilities, an overall output probability distribution calculated using models is also subjective or knowledge-based (Jaynes, 1957). Some of the implications of using knowledge-based probabilities are described throughout this section.

According to the left column in Table 2, the focus of the challenges relates to the model outputs, more specifically predictions (*CH1* and *CH2*), input quantities (*CH3-CH6*), parameters (*CH7-CH9*), and models (*CH10*). We recall that uncertainty quantification helps determine the system's response uncertainty based on specified input quantities. Accordingly, an assessment focuses on the potential system's responses. The focus is often on uncertainty about future non-observed responses $\boldsymbol{Y^*}$, which are approximated by the model output $\boldsymbol{Y}$, considering some specified input quantities ***X***. We recall that ***Y**** and ***X**** are quantities that are unknown at the time of the analysis but will take some value in the future, and possibly become known. Thus, during an assessment, ***Y**** and ***X**** are the uncertain quantities of the system since we have incomplete knowledge about ***Y**** and ***X*.*** Accordingly, the output-prediction error ε, the mismatch between the model prediction values y, and the non-observed system's response values y^*, can only be specified based on the scrutiny of ***K*.**

There is another consequence of considering the definition of uncertainty put forward in *C1*, which links uncertainty solely to quantities or events. The consequence is that models, as such, are not to be linked to uncertainty. Models are merely mathematical artefacts. Models, per se, do not introduce uncertainty, but they are likely inaccurate. Accordingly, another major distinction is to be set in place. We recall that models, by definition, are simplifications, approximations of the system being analysed. They express or are part of the knowledge of the system. Models should therefore be solely used for understanding the performance of the system rather than for illusory perfect predictions. In Table 3, we denote the latter clarification as *C2*.

Regarding the challenges *CH1* and *CH2*, we should note that geohazard analysts are often more interested in predictions rather than known system outputs. For instance, predictions are usually required to be calculated for input values not contained in the validation data. We consider that predictions are those model outputs not observed or recorded in the data, i.e., extrapolations out of the range of values covered by observations. Thus, the focus is on quantifying the uncertainty of the system's responses rather than on the accuracy of a model reproducing recorded data. This is the clarification *C3* in Table 3. Considering this, models are yet to provide accuracy in reproducing observed outputs but, more importantly, afford credibility in predictions. Such credibility is to be assessed mainly in terms of judgements, since conventional validation cannot be conducted using non-observed outputs. Recall that model accuracy usually relates to comparing model outputs with experimental measurements (Roy and Oberkampf, 2011; Aven and Zio, 2013) and is the basis for validating models. Regarding the credibility of predictions, Wagener, Reinecke, and Pianosi (2022) have reported that such credibility can be mainly judged in terms of the physical consistency of the predictions. Such consistency is judged by checks rejecting physically impossible representations of the system. The credibility of predictions may also include the verification of the ability of models to accurately reproduce disruptive changes recorded in the data (Alley, 2004). However, as we have made explicit in the previous section, model predictions are conditional on a considerable number of critical assumptions and choices made by analysts (see Table 1 and clarification C4 in Table 3). Therefore, predictions can only be as good as the quality of the assumptions made. The assumptions could be wrong, and the examination of the impact of these deviations on the predictions must be assessed. To provide credibility of predictions, such assumptions and choices should be justified and scrutinised, ref. option *O5* in Table 2. Option *O5* addresses the challenge *CH1*; however, when conducting UQ, *O5* has a major role when investigating input uncertainty, which is discussed next.

A critical task in UQ is the quantification of input uncertainty. Input uncertainty may originate when crucial historical events or disruptive changes are missing in the records, *CH3*. Some critical input quantities may also remain unidentified to analysts during an assessment, *CH4*. Analysts can unintendedly fail to identify relevant elements in ***X**** due to insufficiencies in data or limitations of existing models. For example, during many assessments, trigger factors that could bring a soil mass to failure could remain unknown to analysts (e.g., Hunt et al. 2013; Clare et al. 2016; Leynaud et al. 2017; Casalbore et al. 2020). UQ requires simulating sampled values from ***X***, and elements in ***X*** can be mutually dependent. However, the joint distribution of ***X***, namely $f(x)$, is often also unknown. This is the challenge *CH6*. Considering the potential challenges *CH3* to *CH6*, to specify *f(x)*, we cannot solely rely on using the maximum entropy principle (MEP). The MEP may fail to advance an exhaustive uncertainty quantification in the input, e.g., by missing relevant values not recorded in the measured data. This would undermine the quality of predictions and, therefore, uncertainty quantification. Recall that the MEP suggests using the least informative distribution among candidate distributions constrained solely on measurements. Using counterfactual analysis, as described in Table 2, is an option. However, the counterfactual analysis will also fail to provide quality predictions since this analysis focuses on counterfactuals (alternative events to observed facts d, $d \in \boldsymbol{D}$) and little on the overall knowledge available $\boldsymbol{K}$. Note that the knowledge $\boldsymbol{K}$ about the system includes, e.g., the assumptions made in the UQ, such as those shown in Table 1. Further note that such assumptions relate not only to data but also to input quantities, modelling, and predictions. Thus, it appears that the examination of these assumptions should be at the core of UQ in geohazard assessments, as suggested in Table 2, option *O5*. The risk assessment of deviations from assumptions was originally suggested by Aven (2013) and exemplified by Khorsandi and Aven (2017). An assumption deviation risk assessment evaluates different deviations, their associated probabilities of occurrence, and the effect of the deviations. A major distinctive feature of the assumption deviation risk assessment approach is the evaluation of the credibility of the knowledge $\boldsymbol{K}$ supporting the assumptions made. Another feature of this approach is questioning the justifications supporting the potential for deviations. The examination of $\boldsymbol{K}$ can be achieved by assessing the justifications for the assumptions made, the amount and relevance of data or information, the degree of agreement among experts, and the extent to which the phenomena involved are understood and can be modelled accurately. Justifications might be in the form of direct evidence becoming available, indirect evidence from other observable quantities, supported by modelling results, or possibly inferred by assessments of deviations of assumptions. This approach is succinctly demonstrated in the following section. Accordingly, we suggest specifying *f(x)* in terms of knowledge-based probabilities in conjunction with investigating input uncertainty using the assumptions deviation approach. This is identified as consideration *C5* in Table 3.

Another point to consider is that when uncertainty is measured in terms of knowledge-based probabilities, analysts should be aware of what *conditionality* means. If, for example, a quantity X_2 is conditional on a quantity X_1, this implies that increased knowledge about X_1 will change the uncertainty about X_2. The expression that denotes this is conventionally written as $X_2|X_1$. Analysts may exploit this interpretation when specifying, e.g., the joint distribution $f(x,\theta)$. For example, when increased knowledge about a quantity X_1 will not result in increased knowledge about another quantity X_2, analysts may simplify the analysis according to the scrutiny of $\boldsymbol{K}$, meaning that a distribution $f(y|x_{1,}x_2)$ to be specified may reduce to $f(y|x_1)f(y|x_2)$ according to probability theory. Apeland, Aven, and Nilsen (2002) have illustrated how conditionality in the setting of knowledge-based probabilities can inform the specification of a joint distribution.

The parameterisation problem, which involves the challenges *CH7* to *CH9* in Table 2, warrants exhaustive consideration. Addressing these challenges also requires some reinterpretation. To start, note that parameters are coefficients determining specific functions among a family of potential functions modelling the system. Those parameters constrain a model's output. Recall that y, as realisations of $\boldsymbol{Y}$, are the model output when $\boldsymbol{X}$ takes the values x, and some parameters $\theta \in \boldsymbol{\Theta}$, and models $\eta \in \boldsymbol{M}$ are used. Thus, as shown in the previous section, any output y is conditional on θ, and so is the uncertainty attached to y^*. We may also distinguish two types of parameters. We may have parameters associated with a property of the system. There exist other parameters that are merely artefacts in the models and are not properties of the system. As suggested, if uncertainty can solely be attached to events or quantities, we may say that parameters that are not properties of the system are not to be linked to any uncertainty. This is identified as clarification *C6* in Table 3. For example, analysts may consider that, the parameters not being part of the system as such are those linked to: the output-prediction error ε, the vector associated with observation/measurement errors $\boldsymbol{\Theta}_o$, and the overall attached hyperparameters linked to probability distributions (including priors, likelihood functions). Analysts may consider the latter parameters as modelling artefacts, so it is questionable to attach uncertainty to them. Thus, focused on the uncertainty of the system responses rather than model inaccuracies, uncertainty is to be assigned to those parameters that represent physical quantities. Fixed single values can be assigned to those parameters that are not properties of the system. To help identify those parameters to which some uncertainty can be linked, we can scrutinise, e.g., the physical nature of these. In fixing parameters to a single value, we can still make use of back analysis procedures, as mentioned previously. Analysts may have some additional basis to specify parameter values when the background knowledge available $\boldsymbol{K}$ is scrutinised. $\boldsymbol{K}$ can be examined to verify that not only data measurements but other sources of data, models, and assumptions made strongly support a specific parameter value. Based on this interpretation, setting the values of the parameters that are not properties of the system to a single value reduces the complications in quantifying uncertainty considerably. It also follows that analysts are encouraged to make explicit that model outputs are conditional on these fixed parameters, as well as on the model or models chosen, as we have shown in the previous section. The latter also leads us to argue that the focus of UQ is on the uncertainty of the system response rather than the inaccuracies of the models. This implies in a practical sense that in geohazard assessments, when parameters are clearly differentiated from specified input quantities, and models providing the most credible predictions are chosen, uncertainty quantification can then proceed. This parsimonious modelling approach is identified as consideration *C7* in Table 3. This latter consideration addresses, to an extent, the challenge ***CH10.***

In the following section, we further illustrate the above discussion by analysing a documented case in which UQ in a geohazard assessment was informed by modelling using sampling procedures.

4 Case analysis

To further describe the proposed considerations, we analyse a case reported in the specialised literature. The case deals with the quantification of uncertainty of geological structures, namely uncertainty about the subsurface stratigraphic configuration. Conditions in the subsurface are highly variable, whereas site investigations only provide sparse measurements. Consequently, subsurface models are usually inaccurate. At a given location, subsurface conditions are unknown until accurately measured. Soil investigation at all locations is usually impractical and uneconomical, and point-to-point condition variation cannot be known (Vanmarcke, 1984). Such uncertainty means significant engineering and environmental risk to, e.g., infrastructure built on the surface. One way to quantify this uncertainty is by calculating the probability of every possible configuration of the geological structures (Tacher et al. 2006; Thiele et al. 2016; Pakyuz-Charrier et al. 2019). Sampling procedures for UQ are helpful in this undertaking. We use an analysis and information from Zhao et al. (2021), which refer to a site located in the Central Business District, Perth, Western Australia, where 6 boreholes were executed. The case has been selected taking into account its simplicity to illustrate the points of this paper, but at the same time, it provides details to allow some discussion. Figure 2 displays the system being analysed.

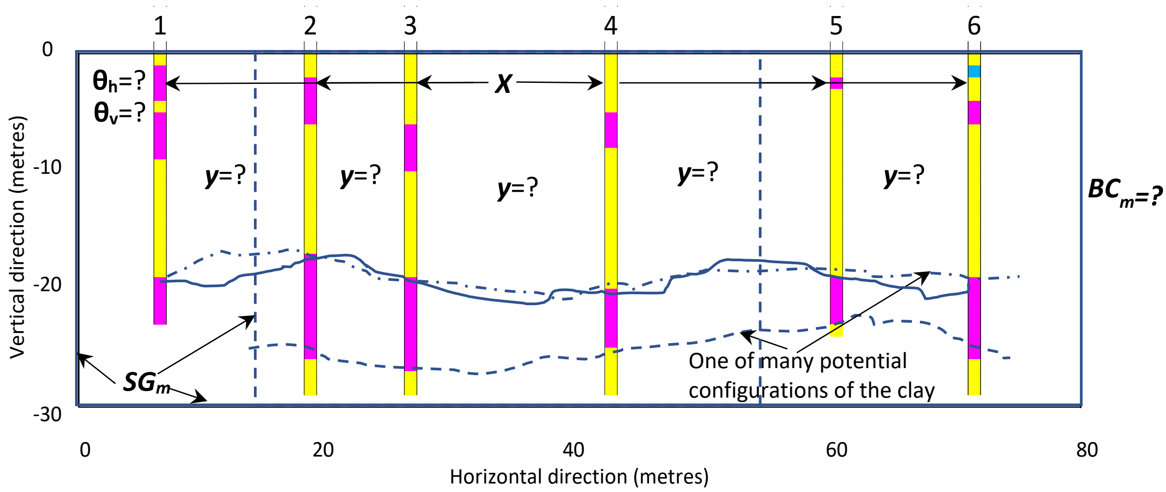


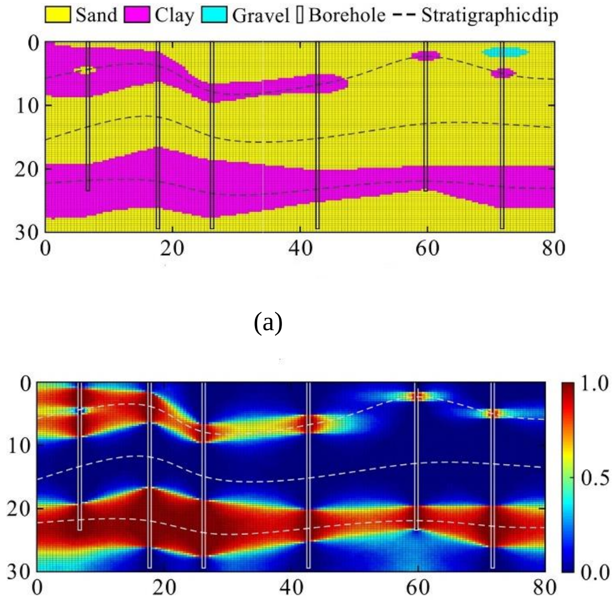
Figure 2: Borehole logs in colours and longitudinal section reported by Zhao et al. (2021) located in the Central Business District, Perth, Western Australia. The records correspond to information on six boreholes. Three types of materials are revealed by the boreholes, including sand (yellow), clay (magenta), and gravel (blue).

In the system under consideration, a particular material type to be found in a non-bored point, a portion of terrain not penetrated during soil investigation, is unknown and thus uncertain. The goal is to compute the probability of encountering a given type of soil at these points. Zhao et al. (2021) focus on calculating the probabilities of encountering clay in the subsurface. The approach advocated was a sampling procedure to generate many plausible configurations of the geological structures and evaluate their probabilities. In a non-penetrated point in the ground, to calculate the probability of encountering a given type of soil c, $p(\boldsymbol{y}=c)$, Zhao et al. (2021) used a function that depends on two correlation parameters, namely the horizontal and vertical scale of fluctuation θ_h and θ_v. Note that spatial processes and their properties are conventionally assumed as spatially correlated. Such spatial variation may presumably be characterised by correlation functions, which depend on a scale of fluctuation parameter. The scale of fluctuation measures the distance within which points are significantly correlated (Vanmarcke, 1984). Expression 9 describes the basic components of the model chosen by Zhao et al. (2021) (specific details are given in the Appendix to this paper):

$$mj: \boldsymbol{X}_{s^{\times}} \boldsymbol{\Theta}_{mj} \rightarrow \boldsymbol{Y}_s \rightarrow p(y=c) \qquad (9)$$

where $\boldsymbol{X}$ is the collection of all specified quantities at borehole points s_x which can take values x from the set {sand, clay, gravel}, according to the setting in Figure 2. $\boldsymbol{Y}$ is the collection of all model outputs with values y at non-borehole points s_y. Probabilities $p(\boldsymbol{y}=c)$ are computed based on the sampling of the values y and x, and a chosen model using the parameters θ_h=11,1 and θ_v=4,1 metres, $\theta_h, \theta_v \in \boldsymbol{\Theta}_{mj}$. Using the maximum likelihood method, the parameters were determined based on the borehole data revealed at the site. In determining parameters, the sampling from uniform and mutually independent distributions of θ_h and θ_v was the procedure advocated. The system is further described by a set of equations $\boldsymbol{E}_{mj}$ (a correlation function and a probability function), the spatial domain geometry sg_{mj} (a terrain block of 30 x 80 metres), and the boundary conditions bc_{mj} (the conditions at the borders). More details are given in the Appendix to this paper. Since this system is not considered time-dependent, the initial conditions $\boldsymbol{IC}_{mj}$ were not specified.

The summary results reported by Zhao et al. (2021) are shown in Figure 3. In Figure 3, the most probable stratigraphic configuration, along with the spatial distribution of the probability of the existence of clay, is displayed. The authors focused on this sensitive material, which likely represents a risk to the infrastructure built on the surface.



(a)

(b)

Figure 3: Zhao et al. (2021) findings shown in their Figure 9. (a) Most probable stratigraphic configuration. (b) Spatial distribution of the probability of the existence of clay. Reprinted from Engineering Geology, 238, Zhao et al. 2021, Probabilistic characterisation of subsurface stratigraphic configuration with modified random field approach, Pages 106138, Copyright (2021), with permission from the COPYRIGHT OWNER: Elsevier. Distances in metres.

Zhao et al. (2021) stated that "*characterisation results of the stratigraphic configuration and its uncertainty are consistent with the intuition and the state of knowledge on site characterisation*". Next, throughout Zhao et al.'s (2021) analysis, the following assumptions were enforced (Table 4), although these were not explicitly disclosed by the authors.

Table 4. Assumptions enforced by Zhao et al., (2021)

Predictions (non-observed outputs) are credible.
Likelihood functions $\mathscr{L}[ɱ(\theta|d)]$ were set to be mutually independent.
For the determination of parameters and model, $f(\theta_h,\theta_v)$ distributes according to the maximum entropy principle.
θ_h and θ_v are mutually independent.
Specified elements $\boldsymbol{X}$ are complete.
$\boldsymbol{X}$ only takes values from the set {sand, clay, gravel}.
$\boldsymbol{X}$ were set to the measured values, i.e., $x = \hat{x}$ (no inaccuracies in data).
θ are independent of $sg_{ɱ}$ and $bc_{ɱ}$.
$sg_{ɱ}$, $bc_{ɱ}$ were set to be constant.

Unfortunately, the authors did not report enough details on how the majority of these assumptions are justified. We should note, however, that providing these justifications was not the objective of their research. Yet, here we analyse how assumptions can be justified by scrutinising $\mathcal{K}$ and using some elements of the assumption deviation approach described in the previous section. Table 5 summarises the analysis conducted and only reflects the most relevant observations and reservations we identified. Accordingly, the information in Table 5 may not be exhaustive, but is still useful for the desired illustration. Table 5 displays some of our observations related to the credibility of the knowledge $\mathcal{K}$. The examination of $\mathcal{K}$ is achieved by assessing the amount and relevance of data or information, the extent to which the phenomena involved are understood and can be modelled accurately, the degree of agreement among experts, and the justifications for the assumptions made. Observations regarding the justifications for potential deviations from assumptions also form part of the analysis.

Not surprisingly, the observations in our analysis concentrate on the predictions' credibility. Recall that UQ focuses on the system's response, approximated by model predictions (considerations *C2* and *C3* in Table 3). For example, although using correlations is an accepted practice and a practical simplification, correlation functions appear counterintuitive to model geological structures or domains. Further, correlation functions do not help much in understanding the system (consideration *C2* in Table 3). Recall that such structures are mainly disjoint domains linked to a finite set of possible categorical quantities (masses of soil or rock) rather than continuous quantities. Next, the variation of such structures can occur by abrupt changes in materials, thus the use of smoothed correlation functions to represent them requires additional consideration. Moreover, the physical basis of the correlation functions is not clear and physical models based on deposition processes may be suggested (e.g., Catuneanu et al., 2019). We should note a potential justification for the deviation from the assumption regarding the credibility of predictions. This is because knowledge from additional sources such as surface geology, sedimentology, local geomorphic setting, and structural geology was not explicitly taken into account in quantifying uncertainty. The revision of this knowledge can contribute to reducing the probability of deviation in predictions. Based on the observations in Table 5, we can conclude that there is potential to improve the credibility of predictions.

The choices made by Zhao et al. (2021) regarding the use of parameters with fixed values together with the choice for a single best model can be highlighted. These choices illustrate the points raised in considerations *C6* and *C7* (Table 3). The maximum likelihood method supported these choices; a back analysis method focused on matching measurements and calculated model outputs using different assumed values for θ_h and θ_v. We highlight that a model judged to be the best model was chosen. This includes the specification of a particular spatial domain geometry in $\boldsymbol{SG}_{m_j}$. Investigating the impact of the variation of $\boldsymbol{SG}_{m_j}$ was considered unnecessary. There was no need to specify several competing models, which is in line with our consideration labelled as *C7* in this paper.

Zhao et al. (2021) investigated the joint distribution $f(x)$, which was sampled to calculate probabilities. However, someone can suggest that the joint distribution $f(x,\theta,sg_{m_j},bc_{m_j})$ could have been produced. Nevertheless, we can argue that establishing such a joint distribution is challenging and requires, in many instances, that analysts encode many additional assumptions (e.g., prior distributions, likelihood functions, independence, linear relationships, normality, stationarity of the quantities and parameters considered).

Table 5. Examination of supporting knowledge $\mathcal{K}$ and justifications for the potential deviation of assumptions

The amount and relevance of data or information	The extent to which the phenomena involved are understood, and accurate models exist	The degree of agreement among experts	Justifications for the assumptions made	Justifications supporting the potential deviations
Assumption: Predictions of ***Y**** are credible.				
The analysis is only based on borehole information; however, such investigation is exceptionally exhaustive, 6 boreholes.	The physical basis for using correlations is dubious and models based on the deposition process can be considered. Variation of geological structures can occur by abrupt changes, thus the use of smoothed functions to represent them requires additional consideration. Global rather than local correlation between spatial quantities has been used, possibly misrepresenting geological structures variation.	The use of correlations is an accepted practice in the field.	Exhaustive borehole information.	The knowledge of surface geology, sedimentology, local geomorphic setting, and structural geology was not explicitly incorporated into UQ.
Assumption: Likelihood functions $\mathscr{L}[\eta(\theta\|d)]$ were set to be mutually independent.				
Assumption: $f(\theta_h,\theta_v)$ distributes according to the maximum entropy principle and θ_h and θ_v are mutually independent.				
Data of the six boreholes has been used to calibrate the model chosen. However, knowledge of surface geology, sedimentology, local geologic/geomorphic setting, and structural geology was not explicitly incorporated into the analysis.		Based on the maximum likelihood method, a model judged to be the best model was chosen.	Dependency between θ_h and θ_v, cannot be supported by general knowledge and such dependency hardly can be enforced.	An increased revision of $\mathcal{K}$, could have been useful to specify *f(x)* and *f(θ)* providing a richer information than that suggested by the maximum entropy principle, regarding how ***X*** or ***Θ*** take values.
Assumption: Specified elements ***X*** are complete.				
		Input quantities were considered fully specified.		
Assumption: Specified input quantities ***X*** were set to measured values $x=\hat{x}$.				
Errors during surveys may have resulted in horizontal positioning inaccuracies.		Data was judged to be accurate.	There is usually not data basis to calculate these errors.	

A more crucial observation derived from the analysis of potential deviations of assumptions might considerably impact the credibility of predictions. This observation comes from revisiting the knowledge sources of Zhao et al.'s (2021) analysis, available from https://australiangeomechanics.org/downloads/. Another type of sensitive material was revealed by other soundings in the area, more specifically, silt. Depending on the revision of $\mathcal{K}$, this fourth suspected material could be analysed in an extended uncertainty quantification of the system. Note that the specified input quantities ***X*** were originally assumed to take values *x* from the set {*sand*, *clay*, *gravel*}. Such an assumption was based on the records of six boreholes which were believed accurate. The latter illustrates the relevance of consideration *C5* in Table 3.

Another choice by Zhao et al. (2021) is that they disregarded the possibility of incorporating measurement errors of the borehole data into the UQ, probably because this data was judged to be accurate. We recall in this respect that these errors reflect the inaccuracy of the measurements rather than the uncertainty about the system. As stated for consideration *C6* (Table 3), we can hardly justify attaching uncertainty to measurement error parameters since measurement errors are not a property of the system. The same can be said for the parameters θ_h and θ_v, which are not properties of the system. Note that their physical basis is questioned. We should note, however, that assuming global coefficients for the parameters θ_h and θ_v is an established practice (Vanmarcke, 1984, Lloret-Cabot et al., 2014, Juang et al., 2019). It can be pointed out that uncertainty quantification in this kind of system is, to an extent, sensitive to the choice of scale of fluctuation values (Vanmarcke, 1984). It can also be argued that using a global rather than local correlation between spatial quantities can misrepresent geological structure variation. Accordingly, further examination of the existing knowledge $\mathcal{K}$ justifies some assessment of the impact of assuming a local rather than global scale of fluctuation.

Overall, the Zhao et al. (2021) analysis is, to an extent, based on the previously suggested definition of uncertainty, ref. the consideration *C1* in Table 3.

We should stress that Zhao et al.'s (2021) uncertainty quantification refers specifically to the ground model described at the beginning of this section. In other words, the probabilities displayed in Figure 3b are conditional on the parameters chosen (θ_h=11,1 and θ_v=4,1 metres), the model selected (described by Eq. 9, A-1 and A-2 in the Appendix to this paper), the specified spatial domain geometry sg_{η} (a terrain block of 30 x 80 metres), and ultimately the assumptions made (listed in Table 4). This information is to be reported explicitly to the users of the results. This reflects the clarification *C4* in Table 3.

Regarding the consideration of subjective probabilities, there has been some agreement on their use in this kind of UQ since Vanmarcke (1984). However, the use of knowledge-based probabilities in the extension described here is recommended, given the illustrated implications to advance UQ (as discussed in the previous section and stated in consideration *C5*). For example, increased examination of $\mathcal{K}$ might have resulted in using a more informative distribution $f(\theta_h,\theta_v)$ than the uniform distribution. The increased examination of $\mathcal{K}$ might have led to different values for θ_h and θ_v, as well as a different model. Recall that the selection of the model and determination of parameters were based on the maximum likelihood method, which only uses measured data $\mathcal{d}$.

In our analysis of Zhao's et al. (2021) assessment, the examination of supporting knowledge $\mathcal{K}$ resulted essentially in:

(i) judging the credibility of predictions;

(ii) providing justifications for assessing assumption deviations by considering the modelling of a fourth material;

(iii) considering additional data other than the borehole records, such as surface geology, sedimentology, local geomorphic setting, and structural geology;

(iv) analysing the possibility of distinct geological models with diverse spatial domain geometry and local correlations; and

(v) ultimately, further examining the existing $\boldsymbol{K}$.

5 Conclusions

In this paper, we have discussed challenges in uncertainty quantification (UQ) for geohazard assessments. Beyond the parameterisation problem, the challenges include assessing the quality of predictions required in the assessments, quantifying uncertainty in the input quantities, and considering the impact of choices and assumptions made by analysts. Such challenges arise from the common-place situation of limited data and the one-off nature of geohazard features. If these challenges are kept unaddressed, UQ lacks credibility. Here, we have formulated seven considerations that may contribute to providing increased credibility in the quantifications. For example, we proposed understanding uncertainty as lack of knowledge, a condition that can only be attributed to quantities or events. Another consideration is that the focus of the quantification should be more on the uncertainty of the system response rather than the accuracy of the models used in the quantification. We drew attention to the clarification that models, in geohazard assessments, are simplifications used for predictions approximating the system's responses. We have also considered that since uncertainty is only to be linked to the properties of the system, models do not introduce uncertainty. Inaccurate models can, however, produce poor predictions and such models should be rejected. Then, an increased examination of background knowledge will be required to quantify uncertainty credibly. We also put forward that there could not be uncertainty about those elements in the parameter set that are not properties of the system. The latter also has pragmatic implications, including how the many parameters in a geohazard system could be constrained in a geohazard assessment.

We went into detail to show that predictions, and in turn UQ, are conditional on the model(s) chosen together with the assumptions made by analysts. We identified limitations of measured data to support the assessment of the quality of predictions. Accordingly, we have proposed that the quality of UQ needs to be judged based also on some additional crucial tasks. Such tasks include the exhaustive scrutiny of the knowledge coupled with the assessment of deviations of those assumptions made in the analysis.

Key to enacting the proposed clarifications and simplifications is the full consideration of knowledge-based probability. Considering this type of probability will help overcome the identified limitations of the maximum entropy principle or counterfactual analysis to quantify uncertainty in input quantities. We have exposed that the latter approaches are prone to produce unexhausted uncertainty quantification due to their reliance on measured data, which can miss crucial events or overlook relevant input quantities.

Code and data availability. The research reported in this paper did not generate any data. No code was developed.

Appendix

In this Appendix, the necessary details of the original analysis made by Zhao et al. (2021) are given. The following are the basic equations $\boldsymbol{E}_{m}$ used by these authors.

$$p(\boldsymbol{y} = c) \sim \frac{\sum_{x_s \times \boldsymbol{Y}} \rho_{x=c,y=c}}{\sum_{c=1}^{C} \sum_{x_s \times \boldsymbol{Y}} \rho_{x=c,y=c}} \qquad \text{(A-1)}$$

$$\rho_{xy} = exp\left(-\pi \frac{\overline{s_x s_y}}{\theta_h} - \pi \frac{|s_x s_y}{\theta_v}\right) \tag{A-2}$$

where ***X*** is the collection of all specified quantities at borehole points, which take values *x*. ***Y*** is the collection of all outputs at non-borehole points with values *y*. ρ_{xy} is the value of correlation between a quantity value *x* at a penetrated point $s_x \in \boldsymbol{S}_x$ and the value *y* at a non-penetrated point $s_y \in \boldsymbol{S}_y$. $\overline{s_x s_y}$ is the horizontal distance between points s_x and s_y, while $|s_x s_y$ is the vertical one. θ_h and θ_v are the horizontal and vertical scales of fluctuation, respectively. Each material class considered is associated exclusively with an element in the set of integers $\{1,2,\ldots,C\}$. $p(\boldsymbol{y}=c)$ is the probability of encountering a type of material *c* in a point s_y. Such probability is initially approximated using Eq. A-1. More accurate probabilities are computed based on the repeated sampling of the joint distribution *f*(*x*,*y*), which was approximated using Eq. A-1. Eq. A-1, described in short, approximates probabilities as the ratio of the sum of correlation values, calculated for a penetrated point in the set S_x and the set of non-penetrated points S_y for a given material *c*, to the sum of correlation values for all points and all materials.

Based on data collected at borehole locations, the selection of the type of correlation function and the scales of fluctuation took place using the maximum likelihood method. The authors considered three types of correlation functions, namely squared exponential, single exponential, and second-order Markov. In this case, the likelihood function $\mathscr{L}(\theta_m|d) = f(d|\theta_m)$ represents the likelihood of observing d at borehole locations, given the spatial correlation structure θ_m. The squared exponential function yielded the maximum likelihood when the horizontal and vertical scales of fluctuation were set to 11.1 and 4.1 metres, respectively. Hence, the squared exponential function correlation, whose expression is Eq. A-2 in this Appendix was selected. Eq. A-3 and A-4 correspond to the single exponential and the second-order Markov functions, respectively.

$$\rho_{xy} = exp\left(-2 \frac{\overline{s_x s_y}}{\theta_h} - 2 \frac{|s_x s_y}{\theta_v}\right) \tag{A-3}$$

$$\rho_{xy} = \left(1 + 4 \frac{\overline{s_x s_y}}{\theta_h}\right)\left(1 + 4 \frac{|s_x s_y}{\theta_v}\right) exp\left(-4 \frac{\overline{s_x s_y}}{\theta_h} - 4 \frac{|s_x s_y}{\theta_v}\right) \tag{A-4}$$

Sumplement. The paper has no suplement.

Authors contributions. ICC Conceptualisation, Methodology, Writing- Original draft preparation, Investigation, Validation; TA Investigation, Supervision, Writing- Reviewing and Editing; RG Investigation, Supervision, Writing- Reviewing and Editing.

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

Acknowledgements. This research is funded by ARCEx partners and the Research Council of Norway (grant number 228107).

References

Albert, C. G., Callies, U., and von Toussaint, U.: A Bayesian approach to the estimation of parameters and their interdependencies in environmental modeling, Entropy, 24(2), 231. doi:10.3390/e24020231, 2022.

Alley, R. B. Abrupt climate change, Sci. Am., 291(5), 62-69. doi:10.1126/science.1081056, 2004.

Apeland, S., Aven, T., and Nilsen, T.: Quantifying uncertainty under a predictive, epistemic approach to risk analysis, Reliab. Eng. Syst. Saf., 75(1), 93-102. doi:10.1016/S0951-8320(01)00122-3, 2002.

Aven, T.: On the need for restricting the probabilistic analysis in risk assessments to variability, Risk Anal., 30(3), 354-360. doi:10.1111/j.1539-6924.2009.01314.x, 2010.

Aven, T.: Practical implications of the new risk perspectives, Reliab. Eng. Syst. Saf., 115, 136-145. doi:10.1016/j.ress.2013.02.020, 2013.

Aven, T.: The science of risk analysis: Foundation and practice. Routledge, London. doi:10.4324/9780429029189, 2019.

Aven, T., and Kvaløy, J. T.: Implementing the Bayesian paradigm in risk analysis, Reliab. Eng. Syst. Saf., 78(2), 195-201. doi:10.1016/S0951-8320(02)00161-8, 2002.

Aven, T., and Pörn, K.: Expressing and interpreting the results of quantitative risk analyses, Review and discussion. Reliab. Eng. Syst. Saf., 61(1-2), 3-10. doi:10.1016/S0951-8320(97)00060-4, 1998.

Aven, T., and Zio, E.: Model output uncertainty in risk assessment, Int. J. Perform. Eng., 29(5), 475-486. doi:10.23940/ijpe.13.5.p475.mag, 2013.

Betz, W.: Bayesian inference of engineering models (Doctoral dissertation, Technische Universität München), 2017.

Brown, G.W.: Monte Carlo methods, Modern Mathematics for the Engineers, 279-303. McGraw-Hill, New York, 1956.

Cardenas, I.: On the use of Bayesian networks as a meta-modelling approach to analyse uncertainties in slope stability analysis, Georisk, 13(1), 53–65. doi:10.1080/17499518.2018.1498524. 2019.

Carrera, J., and Neuman, S.: Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms, Water Resour. Res., 22(2), 211–227. doi:10.1029/WR022i002p00211, 1986.

Casalbore, D., Passeri, F., Tommasi, P., Verrucci, L., Bosman, A., Romagnoli, C., and Chiocci, F.L.: Small-scale slope instability on the submarine flanks of insular volcanoes: the case-study of the Sciara del Fuoco slope (Stromboli), Int. J. Earth Sci., 109, (8), 2643-2658. doi:10.1007/s00531-020-01853-5, 2020.

Catuneanu, O., Abreu, V., Bhattacharya, J. P., Blum, M. D., Dalrymple, R. W., Eriksson, P. G., ... and Winker, C.: Towards the standardisation of sequence stratigraphy, Earth-Sci. Rev., 92(1-2), 1-33. doi:10.1016/j.earscirev.2008.10.003, 2009.

Chow, Y.K., S. Li, and Koh. C.G.: A particle method for simulation of submarine landslides and mudflows, Paper presented at the 29th International Ocean and Polar Engineering Conference, June 16–21, in Honolulu, Hawaii, USA, ISOPE-I-19-594, 2019.

Christakos, G.: A Bayesian/maximum-entropy view to the spatial estimation problem, Math. Geol., 22(7), 763-777. doi:10.1007/BF00890661, 1990.

Clare, M.A., Clarke, J.H., Talling, P.J., Cartigny, M.J., and Pratomo, D.G.: Preconditioning and triggering of offshore slope failures and turbidity currents revealed by most detailed monitoring yet at a fjord-head delta, Earth Planet. Sci. Lett., 450, 208-220. doi:10.1016/j.epsl.2016.06.021, 2016.

Clare, M. A., Vardy, M. E., Cartigny, M. J., Talling, P. J., Himsworth, M. D., Dix, J. K., ... and Belal, M.: Direct monitoring of active geohazards: Emerging geophysical tools for deep-water assessments, Near Surf. Geophys., 15(4), 427-444. doi:10.3997/1873-0604.2017033, 2017.

Degen, D., Veroy, K., Scheck-Wenderoth, M., and Wellmann, F.: Crustal-scale thermal models: Revisiting the influence of deep boundary conditions, Environ. Earth Sci., 81(3), 1-16. doi:10.1007/s12665-022-10202-5, 2022.

Dubois, D.: Possibility theory and statistical reasoning, Comput. Stat. Data Anal., 51(1), 47-69. doi:10.1016/j.csda.2006.04.015, 2006.

Ferson, S., and Ginzburg, L. R.: Different methods are needed to propagate ignorance and variability, Reliab. Eng. Syst. Saf., 54(2-3), 133-144. doi:10.1016/S0951-8320(96)00071-3, 1996.

Flage, R., Baraldi, P., Zio, E., and Aven, T.: Probability and possibility-based representations of uncertainty in fault tree analysis, Risk Anal., 33(1), 121-133. doi:10.1111/j.1539-6924.2012.01873.x, 2013.

Flage, R., Aven, T., and Berner, C. L.: A comparison between a probability bounds analysis and a subjective probability approach to express epistemic uncertainties in a risk assessment context–A simple illustrative example, Reliab. Eng. Syst. Saf., 169, 1-10. doi:10.1016/j.ress.2017.07.016, 2018.

Gray, A., Ferson, S., Kreinovich, V., and Patelli, E.: Distribution-free risk analysis, Int. J. Approx. Reason., 146, 133-156. doi:10.1016/j.ijar.2022.04.001, 2022a.

Gray, A., Wimbush, A., de Angelis, M., Hristov, P. O., Calleja, D., Miralles-Dolz, E., and Rocchetta, R.: From inference to design: A comprehensive framework for uncertainty quantification in engineering with limited information, Mech. Syst. Signal Process., 165, 108210. doi:10.1016/j.ymssp.2021.108210, 2022b.

Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, Biometrika, 87, 97-109. doi:10.2307/2334940, 1970.

Helton, J. C., and Oberkampf, W. L.: Alternative representations of epistemic uncertainty, Reliab. Eng. Syst. Saf., 1(85), 1-10. doi:10.1016/j.ress.2011.02.013, 2004.

Huang, L., Cheng, Y. M., Li, L., and Yu, S. Reliability and failure mechanism of a slope with non-stationarity and rotated transverse anisotropy in undrained soil strength, Comput. Geotech., 132, 103970. doi:10.1016/j.compgeo.2020.103970, 2021.

Hunt, J.E., Wynn, R.B., Talling, P.J., Masson, D.G.: Frequency and timing of landslide-triggered turbidity currents within the Agadir Basin, offshore NW Africa: Are there associations with climate change, sea level change and slope sedimentation rates? Mar. Geol., 346, 274-291. doi:10.1016/j.margeo.2013.09.004, 2013.

Jaynes, E. T.: Information theory and statistical mechanics, Phys. Rev., 106(4), 620. doi:10.1103/PhysRev.106.620, 1957.

Juang, C. H., Zhang, J., Shen, M., and Hu, J.: Probabilistic methods for unified treatment of geotechnical and geological uncertainties in a geotechnical analysis, Eng. Geol, 249, 148-161. doi:10.1016/j.enggeo.2018.12.010, 2019.

Khorsandi, J., Aven, T.: Incorporating assumption deviation risk in quantitative risk assessments: A semi-quantitative approach, Reliab. Eng. Syst. Saf., 163, 22-32. doi:10.1016/j.ress.2017.01.018, 2017.

Leynaud, D., Mulder, T., Hanquiez, V., Gonthier, E., Régert, A.: Sediment failure types, preconditions and triggering factors in the Gulf of Cadiz, Landslides, 14(1), 233-248. doi:10.1007/s10346-015-0674-2, 2017.

Liu, Y., Ren, W., Liu, C., Cai, S., and Xu, W.: Displacement-based back-analysis frameworks for soil parameters of a slope: Using frequentist inference and Bayesian inference, Int. J. Geomech., 22(4), 04022026. doi:10.1061/(ASCE)GM.1943-5622.0002318, 2022.

Lloret-Cabot. M., Fenton, G.A., Hicks, M.A.: On the estimation of scale of fluctuation in geostatistics, Georisk 8(2):129-140. doi:10.1080/17499518.2013.871189, 2014.

Lu, P., and Lermusiaux, P. F.: Bayesian learning of stochastic dynamical models, Phys. D: Nonlinear Phenom., 427, 133003. doi:10.1016/j.physd.2021.133003, 2021.

Luo, L., Liang, X., Ma, B., and Zhou, H.: A karst networks generation model based on the Anisotropic Fast Marching Algorithm, J. Hydrol., 126507. doi:10.1016/j.jhydrol.2021.126507, 2021.

Metropolis, N., and Ulam, S.: The Monte Carlo method, J. Am. Stat. Assoc., 44(247), 335-341. doi:10.1080/01621459.1949.10483310, 1949.

Montanari, A., and Koutsoyiannis, D.: A blueprint for process-based modeling of uncertain hydrological systems, Water Resour. Res., 48(9). doi:10.1029/2011WR011412, 2012.

Nilsen, T., and Aven, T.: Models and model uncertainty in the context of risk analysis, Reliab. Eng. Syst. Saf., 79(3), 309-317. doi:10.1016/S0951-8320(02)00239-9, 200.

Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J., and Jessell, M.: Monte Carlo simulation for uncertainty estimation on structural data in implicit 3-D geological modeling, a guide for disturbance distribution selection and parameterisation, Solid Earth, 9(2), 385-402. doi:10.5194/se-9-385-2018, 2018.

Pearl, J.: Comment: graphical models, causality and intervention, Statist. Sci., 8(3), 266-269. https://www.jstor.org/stable/2245965, 1993.

Pheulpin, L., Bertrand, N., and Bacchi, V.: Uncertainty quantification and global sensitivity analysis with dependent inputs parameters: Application to a basic 2D-hydraulic model, LHB, 108(1), 2015265. doi:10.1080/27678490.2021.2015265, 2022.

Raíces-Cruz, I., Troffaes, M. C., and Sahlin, U.: A suggestion for the quantification of precise and bounded probability to quantify epistemic uncertainty in scientific assessments, Risk Anal., 42(2), 239-253. doi:10.1111/risa.13871, 2022.

Roy, C. J., and Oberkampf, W. L.: A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing, Comput. Methods Appl. Mech. Eng., 200(25-28), 2131-2144. doi:10.1016/j.cma.2011.03.016, 2011.

Rodríguez-Ochoa, R., Nadim, F., Cepeda, J. M., Hicks, M. A., and Liu, Z.: Hazard analysis of seismic submarine slope instability, Georisk, 9(3), 128-147. doi:10.1080/17499518.2015.1051546, 2015.

Sankararaman, S., and Mahadevan, S.: Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems, Reliab. Eng. Syst. Saf., 138, 194-209. doi:10.1016/j.ress.2015.01.023, 2015.

Shafer, G.: A mathematical theory of evidence. In A mathematical theory of evidence. Princeton university press. 197.

Shortridge, J., Aven, T., and Guikema, S.: Risk assessment under deep uncertainty: A methodological comparison, Reliab. Eng. Syst. Saf., 159, 12-23. doi:10.1016/j.ress.2016.10.017, 2017.

Society for Risk Analysis Society for Risk Analysis Glossary Available from: SRA-Glossary-FINALpdf Accessed on 25 March 2021, 2018.
Society for Risk Analysis, Society for Risk Analysis glossary. https://www.sra.org/wp-content/uploads/2020/04/SRA-Glossary-FINAL.pdf., 2018 (Accessed 25 June 2021).
Sun, X., Zeng, X., Wu, J., and Wang, D.: A Two-stage Bayesian data-driven method to improve model prediction, Water Resour. Res., e2021WR030436. doi:10.1029/2021WR030436, 2021b.
Sun, X., Zeng, P., Li, T., Wang, S., Jimenez, R., Feng, X., and Xu, Q.: From probabilistic back analyses to probabilistic run-out predictions of landslides: A case study of Heifangtai terrace, Gansu Province, China, Eng. Geol, 280, 105950. doi:10.1016/j.enggeo.2020.105950, 2021a.
Tacher, L., Pomian-Srzednicki, I., Parriaux, A.: Geological uncertainties associated with 3-D subsurface models, Comput. Geosci., 32(2), 212-221. doi:10.1016/j.cageo.2005.06.010, 2006.
Tang, X. S., Wang, M. X., and Li, D. Q.: Modeling multivariate cross-correlated geotechnical random fields using vine copulas for slope reliability analysis, Comput. Geotech., 127, 103784. doi:10.1016/j.compgeo.2020.103784, 2020.
Thiele, S.T., Jessell, M.W., Lindsay, M., Wellmann, J.F., Pakyuz-Charrier, E.: The topology of geology 2: Topological uncertainty, J. Struct. Geol., 91,74-87. doi:10.1016/j.jsg.2016.08.010, 2016.
Ulam, S. M.: Monte Carlo calculations in problems of mathematical physics, Modern Mathematics for the Engineers, 261-281. McGraw-Hill, New York. 1961.
Uzielli, M., and Lacasse, S.: Scenario-based probabilistic estimation of direct loss for geohazards, Georisk, 1(3), 142-154. doi:10.1080/17499510701636581. 2007.
van den Eijnden, A. P., Schweckendiek, T., and Hicks, M. A.: Metamodelling for geotechnical reliability analysis with noisy and incomplete models, Georisk, 1-18. doi:10.1080/17499518.2021.1952611, 2021.
Vanneste, M., F. Løvholt, D. Issler, Z. Liu, N. Boylan, and Kim. J.: A novel quasi-3D landslide dynamics model: from theory to applications and risk assessment. Paper presented at the Offshore Technology Conference, May 6–9, in Houston, Texas, OTC-29363-MS. doi:10.4043/29363-MS, 2019.
Vanmarcke, E.H.: Random fields: Analysis and synthesis. Cambridge, MA: The MIT Press, 1984.
Wagener, T., Reinecke, R., and Pianosi, F. On the evaluation of climate change impact models. Wiley Interdiscip. Rev. Clim. Change, e772. doi:10.1002/wcc.772, 2022.
Wellmann, J. F., and Regenauer-Lieb, K.: Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models, Tectonophysics, 526, 207-216. doi:10.1016/j.tecto.2011.05.001, 2012.
Woo, G.: Downward counterfactual search for extreme events, Front. Earth Sci., 340. doi:10.3389/feart.2019.00340, 2019.
Yano, J. I.: What is the Maximum Entropy Principle? Comments on “Statistical theory on the functional form of cloud particle size distributions”, J Atmos Sci, 76(12), 3955-3960. doi:10.1175/JAS-D-18-0223.1, 2019.
Zadeh, L. A.: Probability measures of fuzzy events, J. Math. Anal. Appl., 23(2), 421-427. doi:10.1016/0022-247X(68)90078-4, 1968.
Zhao, C., Gong, W., Li, T., Juang, C. H., Tang, H., and Wang, H.: Probabilistic characterisation of subsurface stratigraphic configuration with modified random field approach, Eng. Geol, 288, 106138. doi:10.1016/j.enggeo.2021.106138, 2021.

Table 1. Some enforced assumptions in UQ for geohazard assessments

Table 2. Major challenges and options to address them in geohazard assessments

Table 3. Some clarifications and considerations to address the challenges in UQ

Table 4. Assumptions enforced by Zhao et al., (2021)

Table 5. Examination of supporting knowledge Қ and justifications for the potential deviation of assumptions

Figure 1: A configuration of a network coupling some elements in *Ɱ*, *X*, *Θ*, *Y*, *Y**

Figure 2: Borehole logs in colours and longitudinal section reported by Zhao et al. (2021) located in the Central Business District, Perth, Western Australia. The records correspond to information on six boreholes. Three types of materials are revealed by the boreholes, including sand (yellow), clay (magenta), and gravel (blue).

Figure 3: Zhao et al. (2021) findings shown in their Figure 9. (a) Most probable stratigraphic configuration. (b) Spatial distribution of the probability of the existence of clay. Reprinted from Engineering Geology, 238, Zhao et al. 2021, Probabilistic characterisation of subsurface stratigraphic configuration with modified random field approach, Pages 106138, Copyright (2021), with permission from the COPYRIGHT OWNER: Elsevier. Distances in metres.