Global Sensitivity Analysis of a Dam Breaching Model: To Which Extent Is Parameter Sensitivity Case-Dependent?

V. Schmitz[1] ⓘ, **M. Arnst[2]**, **K. El Kadi Abderrezzak[3,4]** ⓘ, **M. Pirotton[1]**, **S. Erpicum[1]** ⓘ, **P. Archambeau[1]** ⓘ, **and B. Dewals[1]** ⓘ

[1]Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium, [2]Computational and Stochastic Modeling, University of Liège, Liège, Belgium, [3]National Laboratory for Hydraulics and Environment (LNHE), EDF R&D, Chatou, France, [4]Saint Venant Laboratory for Hydraulics, Chatou, France

Abstract Failure of dams and dikes often leads to devastating consequences in protected areas. Numerical models are instrumental tools to assess flood risk and guide emergency management, but numerous uncertainties affect model outcomes. Identifying uncertain model input parameters that induce high uncertainties in model outputs is essential. This paper focuses on two model outputs: the maximum breach discharge and the time to reach this peak. Using our implementation of a simplified physically based dam breaching model developed by Wu (2016), a global sensitivity analysis was conducted based on Sobol indices of total order. Unlike in most previous studies, many different configurations (twenty-seven), both at laboratory and field scales, were considered. For each of them, input variables were ranked according to the significance of the contribution of their uncertainty to the output variability, and the dependency between reference configurations and sensitivity analysis results was highlighted. Depending on the considered case study, input parameter uncertainties with the largest impact on output variability were different, and so was the magnitude of output uncertainties. We demonstrate that sensitivity analysis results obtained for a specific configuration cannot be transferred as it to other configurations. Finally, sensitivity and uncertainty quantification results were combined in a decision tree to determine which input parameter uncertainty is the most critical in each configuration and what standard deviation in the output variables is expected. The global sensitivity analysis procedure presented here may apply to a wide variety of models in environmental sciences.

1. Introduction

With the surge of extreme meteorological events and intensification of urbanization downstream of hydraulic structures, the need for predicting failure of dams and dikes has become of paramount importance for establishing emergency response procedures (Zhong et al., 2021). To this end, numerical models are instrumental tools for simulating the embankment breaching process. Existing models can be classified into three categories. First, statistical (or parametric) models are based solely on regression analysis of data from past events or laboratory campaigns. They describe some breaching parameters (e.g., final breach width, failure duration or maximum breach discharge) as a function of dam or reservoir properties. These simple, computational-efficient models may lack generality because they entirely rely on data from specific cases without considering underlying physics (Chen et al., 2019; De Lorenzo & Macchione, 2014; Lee, 2019). Conversely, distributed physically based models can describe the phenomenon in greater detail, as they solve the flow and sediment governing equations using a computational mesh of the domain. Their results may be accurate but only if reliable data is available and if physical processes are numerically well represented, for example, erosion of non-homogeneous dam material, slope failure or 3D-flow patterns (Cantero-Chinchilla et al., 2019; Onda et al., 2019; Pheulpin et al., 2020; Shimizu et al., 2020). Additionally, the time required to run distributed physically based models can be substantial (ASCE, 2011). Simplified physics-based models offer a good trade-off (Wu, 2013). Without spatially distributing the flow description nor the embankment morphology, they enable simulating hydraulic and dam breach variables (e.g., time-evolution of breach discharge and dimensions) by describing selected physical processes (Li et al., 2020; Peter et al., 2018; Tsai et al., 2019; Wu, 2013; Zhong et al., 2017). Simplified physics-based models are computationally efficient and enable uncertainty analyses that require many runs, for example, the Monte Carlo method. In this paper, breaching of homogeneous dams made of non-cohesive material is computed using our implementation of the physics-based lumped fully coupled hydro-morphodynamic model developed by Wu (2016).

In all numerical models, input variables can be subject to uncertainties, which may emerge from multiple sources (Kiureghian & Ditlevsen, 2009; Peter et al., 2018). Depending on their origin, they are generally classified as parametric or modeling uncertainties. The first group gathers uncertainties on physical parameters. They may originate either from inaccuracies in measurements or from the natural variability of parameters (e.g., spatial distribution of geometrical parameters or soil composition of embankments). Modeling uncertainties can be encountered when uncertainties arise from the model structure itself (Gupta et al., 2012), or when using parameters tied to parametrization of complex physical phenomena (e.g., regression parameters). Assigning purely deterministic values to uncertain parameters can be suboptimal and alternative characterizations may be preferred, such as variation intervals or probabilistic distributions.

Uncertainty analyses were developed to gain insights into model response to variations in input parameters (Borgonovo & Plischke, 2016). They are generally divided into three main steps (Arnst & Ponthot, 2014; Helton et al., 2006): (a) uncertainty characterization, (b) uncertainty propagation, and (c) exploitation of the results in accordance with the study goals. In Step 1, probability distributions are attributed to input variables. In Step 2, a numerical model is used to propagate those uncertainties to the model outputs. In Step 3, a common approach consists in leading a sensitivity analysis to better understand the relation that exists between output and input uncertainties. Sensitivity analyses are defined as local or global (Reed et al., 2022). In the first case, uncertainties in the model outputs are evaluated by slight modifications in the input values around a reference configuration. Although being computationally efficient, this approach is not suitable for nonlinear models as it explores a limited fraction of the input space and often assumes that input variables do not interact in the model (Rakovec et al., 2014; Saltelli and Annoni, 2010; Saltelli et al., 2019). To tackle those limitations, global sensitivity analyses are led to reveal the global impact of input uncertainties on model outputs by exploring the entire input space of interest.

Four main sensitivity analysis settings were listed by Saltelli et al. (2004), namely factor (or input parameter) prioritization, factor fixing, variance cutting and factor mapping. When the pursued objective is to identify the uncertain input parameters that induce the most critical variability in model outputs, one refers to factor prioritization. Conversely, factor fixing aims at pointing out input parameters whose variability has a negligible impact on model outputs, allowing to discard their influence in subsequent analyses. Variance cutting is used to assess the minimum number of input parameters that should be fixed to limit the uncertainty on model outputs below a prescribed threshold. Finally, factor mapping identifies which input values lead to output values contained in a specific range of the output space. This approach allows highlighting combinations of input values that lead to non-physical results (Pianosi et al., 2016; Spear & Hornberger, 1980).

This paper highlights the dependency between global sensitivity analysis results and test configuration through factor prioritization and factor fixing approaches. Within this context, variance-based methods are convenient as they rank inputs or subsets of inputs as a function of the influence of their uncertainty on the variability of resulting outputs (Saltelli et al., 2008). Among them, the use of Sobol indices of total order is widespread (Saltelli et al., 2010; Sobol, 1993). Those indicators illustrate the portion of the global output variance caused by the uncertainty on a specific input subset, including the interactions of this subset with any other input subset.

While previous studies were often limited to the analysis of a single case study (Tables S1 and S2 in Supporting Information S1), we apply here a global sensitivity analysis based on Sobol indices of total order to twenty-seven embankment configurations, in both laboratory and field scales. The results of this study highlight that the input parameter uncertainties contributing most to the output variability differ considerably from one configuration to the other. The magnitude of output uncertainties also varies with the considered configuration. Here, we demonstrate that sensitivity analysis results obtained for a specific dam breaching case may not be transferred to any other case.

The remaining of the paper is organized as follows: the computational model and the simulated dam configurations are introduced in Section 2. The input sampling methods and the global sensitivity analysis technique are described in Section 3. In Section 4, results of the sensitivity analysis are presented and discussed for twenty-seven tests in laboratory and field-scale configurations. Conclusions are drawn in Section 5.

2. Physical Data and Methods

This section starts with the description of our implementation of a model initially developed by Wu (2016), that is, the numerical model on which the sensitivity analysis was applied (Section 2.1). It is composed of three interacting modules: (a) a hydrodynamic module, (b) a sediment transport module and (c) a morphodynamic module,
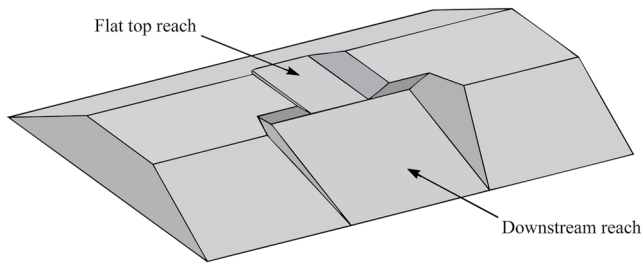
Figure 1. Scheme of the embankment geometry used in Wu's (2016) model.

in which the breach geometry is described as the combination of a flat top reach corresponding to the top of the breach and a downstream reach located on the dike floodplain face (Figure 1). A brief description of the considered experimental configurations is provided in Section 2.2.

2.1. Computational Model

Based on initial and boundary conditions, the model hydrodynamic module evaluates the water level in the main channel and on the different parts of the embankment, the breach discharge, and the resulting shear stress using two main assumptions: critical flow on the dam crest and uniform flow on the downstream face. Those values are then introduced in the sediment transport module to compute both suspended-load and bed load concentrations in the flow through the breach assuming non-equilibrium sediment transport. Using the eroded material volume, the new dam geometry is generated by the morphodynamic module. The procedure is repeated by feeding the hydrodynamic module with the updated dam geometry. At the start of the computation, the geometry of an initial notch in the dam crest needs to be defined by the user, as the model is not able to predict the location of breach initiation.

The computational model was coded using *Matlab* software and explicit resolution schemes were used. The time step was 0.5 s for all tests. The present section summarizes the three modules of the numerical model, and detailed flow charts are provided in Supporting Information S1 (Figures S1, S2 and S3).

2.1.1. Hydrodynamic Module

The hydrodynamic module (Figure S1 in Supporting Information S1) computes first the water level in the main channel using mass balance:

$$\frac{dz_s}{dt} = [Q_{\text{in}} - Q_b - Q_d - Q_{\text{out}}(z_s)]\frac{1}{A_{\text{res}}},$$ (1)

with z_s the water level in the main channel, A_{res} the water area in the main channel, Q_{in} the inflow discharge, Q_b the discharge through the breach, Q_d the drain discharge and Q_{out} the outflow discharge at the downstream end of the main channel. If the water level is lower than the breach bottom elevation, z_b, the breach discharge remains zero. Otherwise, the breach discharge is computed as

$$Q_b = c_1 b_{\text{top}}(z_s - z_b)^{3/2} + c_2 m(z_s - z_b)^{5/2},$$ (2)

with c_1 and c_2 two weir coefficients, b_{top} the breach bottom width on the flat top reach (Figure 1) and m the breach side slope. Since non-cohesive dams are considered here, m can also be expressed as $m = (\tan \varphi_r)^{-1}$, where φ_r is the dam material repose angle considered similar for both wet and dry material.

The water levels on the flat top and downstream reaches (Figure 1) were computed assuming critical flow on the dam crest and uniform flow over the downstream face.

$$\text{Flat top reach} : h_{\text{top}} = \max\left(\frac{2}{3}(z_s - z_b); 0\right),$$
$$\text{Downstream reach} : Q_b = \frac{1}{n}AR^{2/3}S_d^{-1/2},$$ (3)

with A the flow area over the considered reach, R the hydraulic radius, S_d the slope of the breach downstream side (H:V) and n the Manning roughness coefficient computed based on Strickler's formula:

$$n = \max\left(d_{50}^{1/6}/A_n; n_{\text{min}}\right),$$ (4)

where d_{50} is the median size of the embankment material, A_n an empirical coefficient and n_{min} the minimum value allowed for n, as defined by Wu (2016).

On both reaches, the bed shear stress can then be calculated as

$$\tau_b = \frac{\rho g n^2 Q_b^2}{A^2 R^{1/3}},$$ (5)

where ρ is the water density and g the gravitational acceleration.

2.1.2. Sediment Transport Module

The bed shear stress is introduced in the sediment transport module (Figure S2 in Supporting Information S1) to evaluate the eroded material volume on each reach. The effective shear stress is computed as follows:

$$\tau_e = \tau_b' + \lambda_0 \tau_c \frac{\sin \phi}{\sin \varphi_r}, \tag{6}$$

with ϕ the angle between a horizontal plane and the considered reach, that is, $\phi = 0$ on the flat top reach and $\phi = \tan^{-1}(1/S_d)$ on the downstream slope. $\tau_c = \theta_{cr}(\rho_s - \rho)gd_{50}$ is the critical shear stress involving the critical Shields parameter, θ_{cr}, and the sediment density, ρ_s. τ_b' is the grain shear stress based on an empirical coefficient $A_{n'}$ and writes

$$\tau_b' = \left(\frac{n'}{n}\right)^{3/2}\tau_b \quad \text{with } n' = \max\left(\frac{d_{50}^{1/6}}{A_{n'}}; n_{\min}\right). \tag{7}$$

λ_0 is a correction factor introduced by Wu (2016):

$$\lambda_0 = 1 + \lambda_{0,a}\left(\frac{\tau_b'}{\tau_c}\right)^{\lambda_{0,b}} e^{2\sin\phi/\sin\varphi_r}, \tag{8}$$

with $\lambda_{0,a}$ and $\lambda_{0,b}$ two empirical coefficients.

The equilibrium sediment concentration on both reaches is made of two contributions:

$$\text{Suspended load}: C_* = \frac{1}{C_a^* \rho_s}\left(\frac{U^3}{gR\omega_s}\right)^{C_b^*} * \left[1 + \left(\frac{1}{C_c^*}\frac{U^3}{gR\omega_s}\right)^{C_d^*}\right]^{-1},$$

$$\text{Bed load}: q_b^* = q_{b,a}^*\left(\max\left[\frac{\tau_e}{\tau_c} - 1; 0\right]\right)^{q_{b,b}^*}\sqrt{\left(\frac{\rho_s}{\rho} - 1\right)gd_{50}^3}, \tag{9}$$

with $\left\{C_a^*, C_b^*, C_c^*, C_d^*\right\}$ and $\left\{q_{b,a}^*, q_{b,b}^*\right\}$ empirical coefficients obtained through experimental data fitting (Wu et al., 2000; Zhang, 1961), U the flow velocity on the considered reach and ω_s the sediment settling velocity computed using the formula of Wu and Wang (2006), which is based on sediment density, ρ_s, sediment median size, d_{50}, and Corey shape factor, S_p.

The actual sediment concentration at the outlet of each reach, $C_{t,\text{out}}$, is linked to the actual concentration at the inlet, $C_{t,\text{in}}$, through the following formula:

$$C_{t,\text{out}} = C_{t*} + (C_{t,\text{in}} - C_{t*})\exp\left(-\frac{\Delta}{L_s}\right), \tag{10}$$

with $C_{t*} = C_* + B_w q_b^*/Q_b$ the total equilibrium sediment concentration and B_w the width of the water surface. Δ is the length of the considered reach and $L_s = \lambda B_w$ is the mixing length characterizing the adjustment of sediment from a non-equilibrium state to the equilibrium state. λ is a numerical parameter to be adjusted in each simulation.

The eroded volume is computed on each reach using the following equation, in which p denotes the dike material porosity:

$$\frac{dV_b}{dt} = \frac{Q_b(C_{t,\text{in}} - C_{t,\text{out}})}{1 - p}. \tag{11}$$

2.1.3. Morphodynamic Module

In line with Wu (2016), uniform erosion on all breach faces is assumed. The morphodynamic module (Figure S3 in Supporting Information S1) computes the breach depth variation on each reach as

$$\frac{dz_b}{dt} = \frac{1}{A_{e,\text{tot}}}\frac{dV_b}{dt} \quad \text{with } A_{e,\text{tot}} = A_{\text{bottom}} + 2A_{\text{side}}, \tag{12}$$

where A_{bottom} and A_{side} are the breach bottom and breach side areas of the considered reach, respectively.

The breach bottom width variation on the flat top reach is computed as

$$\frac{db_{\mathrm{top}}}{dt} = 2\frac{dz_b}{dt}\left(\frac{1}{\sin\varphi_r} - \frac{1}{\tan\varphi_r}\right), \tag{13}$$

whilst the breach top width variation on this reach writes

$$\frac{dB_{\mathrm{top}}}{dt} = \frac{2}{\sin\varphi_r}\frac{dz_b}{dt}. \tag{14}$$

On the downstream reach, Wu (2016) assumed that the breach width variation is influenced by the ratio between the breach width on the flat top reach and on the downstream reach. He introduced a correction factor c_b expressed as

$$c_b = \min\left[1, \max\left(0, c_{b,\mathrm{coef}}\frac{b_{\mathrm{top}}}{b_{\mathrm{D/S}}} - (\mathrm{c_{b,coef}} - 1)\right)\right], \tag{15}$$

where $c_{b,\mathrm{coef}}$ is an empirical parameter and b_{top} and $b_{\mathrm{D/S}}$ are the breach bottom widths on the flat top and the downstream reach, respectively.

Breach width variations on the downstream reach write

$$\frac{db}{dt} = 2\frac{dz_b}{dt}\left(\frac{\max(c_b, \cos\varphi_r)}{\sin\varphi_r} - \frac{1}{\tan\varphi_r}\right) \text{ and } \frac{dB}{dt} = \frac{2c_b}{\sin\varphi_r}\frac{dz_b}{dt}. \tag{16}$$

To close the loop, updated geometrical variables are finally introduced in the hydrodynamic module and a new iteration is performed.

2.2. Test Cases

The sensitivity analysis was applied to various experimental tests to assess how the configuration influences the relation between uncertainties on input and output variables. Frank (2016) investigated experimentally the spatial dam breach process due to overtopping for dams made of homogeneous non-cohesive material. A total of forty-five three-dimensional embankment breach tests in frontal configuration and laboratory conditions were conducted. The influence of several parameters was analyzed, including dam scaling, inflow discharge, sediment grain size, main channel width, dam cross-section, and reservoir volume (Table 1). In the latter case, the pumping system was regulated to simulate a reservoir of a larger volume than the real physical reservoir. For each test, hydrographs (breach, inflow, and drainage discharges) and the upstream water level were recorded. Twenty-seven laboratory experiments performed by Frank (2016) were first considered. Table 2 summarizes the non-dimensional characteristics of these tests on which an uncertainty analysis was applied while dimensional data are gathered in Table S3 in Supporting Information S1. Since Frank (2016) considered half embankments, parameter values have been adapted in this study to represent entire dams. To be consistent with other configurations, both dam side slopes were set equal to 1:2 (V:H) in Test 28, whilst equal to 1:1.5 in Frank's (2016) original experiments.

To assess the impact of geometric scale on the results of sensitivity analysis, an upscaling was applied to the laboratory tests to obtain field-scale configurations: dike heights were multiplied by a factor 10, while non-dimensional parameters were kept constant. The objective here is to show how dependencies between uncertainties in model inputs and model outputs evolve with the geometric scale. It is not intended to investigate how physical processes differ between the laboratory and field scales, but merely to assess potential differences in model behavior when the geometric scale is varied.

3. Uncertainty Quantification Data and Methods

This section first provides a general description of uncertainty analysis procedures (Section 3.1). Uncertain input variables are then characterized in the numerical model, and uncertainty quantification methods are presented for

Table 1
Summary of Parameters Involved in the Numerical Model

	Symbol	Description
Input parameters	S_u, S_d	Upstream and downstream slopes of the dam, respectively
	L_k	Dam crest length
	h_d	Dam height
	φ_r	Sediment repose angle
	ρ_s	Sediment density
	d_{50}	Sediment median size
	p	Dam material porosity
	Q_{in}	Inflow discharge
Model parameters	c_1, c_2	Weir efficiency coefficients involved in the breach discharge evaluation (Equation 2)
	$A_n, A_{n'}, n_{\min}$	Parameters in Strickler's formula: $n = \max\left(\frac{d_{50}^{1/6}}{A_n}; n_{\min}\right)$; $n' = \max\left(\frac{d_{50}^{1/6}}{A_{n'}}; n_{\min}\right)$, with n and n' total and effective Manning's coefficients, respectively
	θ_{cr}	Critical Shields parameter
	$\lambda_{0,a}, \lambda_{0,b}$	Coefficient and exponent involved in the effective shear stress computation (Damgaard et al., 1997)
	S_p	Corey shape factor
	$C_{*a}, C_{*b}, C_{*c}, C_{*d}$	Regression coefficients involved in the suspended load concentration (Equation 9)
	$q_{b,a}^*, q_{b,b}^*$	Regression coefficients involved in the bed load transport capacity formula (Equation 9)
	λ	Empirical coefficient involved in mixing length computation
	$c_{b,\text{coef}}$	Artificial breach widening limitation coefficient (Equation 15)

independent and dependent input variables (Section 3.2). Using the standard Monte Carlo method, uncertainty is propagated through the computational model and a global sensitivity analysis based on Sobol indices is presented (Section 3.3).

3.1. Uncertainty Analysis Procedures

3.1.1. Step 1: Uncertainty Characterization of Input Variables

Model predictions are directly affected by input uncertainties, highlighting the need for systematically quantifying them to reliably predict physical processes, such as embankment failures (ASCE, 2011; Froehlich, 2008; Pappenberger & Beven, 2006). Different approaches exist to quantify uncertainties on physical parameters, depending on data availability and modeling objectives, that is, description of a specific case or general representation of typical embankment failures. In the latter case, when the data set is large, maximum likelihood estimation is often used to define parameter distributions (Rossi, 2018). In contrast, when scarce data is available, Bayesian inference is preferred (Saltelli et al., 2004). When information on parameters of interest is too limited or not available at all (e.g., model parameters), plausible parameter distributions may be defined based on modeler expertise. Dependency between input variables is seldom considered, while this may influence input probability distributions significantly (Baroni & Tarantola, 2014; Da Veiga et al., 2009, 2021; Jacques et al., 2006; Li et al., 2010; Pheulpin et al., 2022). In the present case, nineteen independent input variables and two sets of dependent input variables are involved. Specific embankment configurations are considered, for which almost no information on parameter distributions is available. Therefore, probability density functions (PDF) associated to each input were mostly generated based on expert judgment, with a special care for dependent input variables (see Sections 3.2.1 and 3.2.2).

3.1.2. Step 2: Uncertainty Propagation

The uncertainty in input variables must be propagated through the numerical model to assess its impact on the outputs. This procedure is crucial to properly interpret model outcomes, since it determines the variation range of model predictions (Pappenberger & Beven, 2006). It is most often implemented using a Monte Carlo sampling

Table 2
Frank's (2016) Experimental Parameters

	Test ID	L_r/h_d	l_r/h_d	A_r/h_d^2	h_d (mm)	L_k/h_d	$z_{n,\mathrm{ini}}/h_d$	B_{ini}/h_d	d_{50}/h_d	d_{50} (mm)	Q_{in} (l/s)	h_{cr}/h_d
Scaling	11				150				5.7×10^{-3}	0.86	2.62	
	10	4.07	10/3	13.57	300	1/3	1/3	4/3	5.8×10^{-3}	1.75	14.8	9.4×10^{-2}
	8				600				6.3×10^{-3}	3.78	83.6	
	12				150				5.7×10^{-3}	0.86	1.31	
	13	4.07	10/3	13.57	300	1/3	1/3	4/3	5.8×10^{-3}	1.75	7.4	5.9×10^{-2}
	14				600				6.3×10^{-3}	3.78	41.8	
	15				150					0.43	2.62	
	16	4.07	10/3	13.57	300	1/3	1/3	4/3	2.9×10^{-3}	0.86	14.8	9.4×10^{-2}
	17				600					1.75	83.6	
Discharge	18										4.4	2.6×10^{-2}
	19	11.47	20/3	76.47	300	1/3	1/3	4/3	5.8×10^{-3}	1.75	8.8	4.2×10^{-2}
	20										18.4	6.8×10^{-2}
	21										35.2	10.5×10^{-2}
Sediment	22								2.9×10^{-3}	0.86		
	20	11.47	20/3	76.47	300	1/3	1/3	4/3	5.8×10^{-3}	1.75	18.4	6.8×10^{-2}
	23								1.3×10^{-2}	3.78		
Initial breach width	20							4/3				
	24							5/3				
	25	11.47	20/3	76.47	300	1/3	1/3	2	5.8×10^{-3}	1.75	18.4	6.8×10^{-2}
	26							8/3				
	27							4				
Crest length	28					0						
	20	11.47	20/3	76.47	300	1/3	1/3	4/3	5.8×10^{-3}	1.75	18.4	6.8×10^{-2}
	29					4/3						
Reservoir water surface area	40			76.4								
	36			299								
	41	11.47	20/3	367	300	1/3	1/15	4/15	5.8×10^{-3}	1.75	0	0
	42			521								
	43			1410								
	44			2298								

Note. L_r = reservoir length; l_r = reservoir width; A_r = reservoir area; h_d = dam height; L_k = dam crest length; $z_{n,\mathrm{ini}}$ = initial notch depth; B_{ini} = initial notch width; d_{50} = median grain size; Q_{in} = inflow discharge; h_{cr} = critical flow depth.

method (Metropolis & Ulam, 1949). Many simulations are run using different sets of input variable values generated according to their probability distributions. As the number of runs increases, approximations of the model output statistical moments converge, for example, mean and variance (Janssen, 2013). The classical Monte Carlo sampling method is computationally expensive, making it suitable for fast models only. To reduce the number of runs required to reach a given accuracy on output statistical moments, advanced Monte Carlo sampling methods were developed (Hou et al., 2019). One of the most frequently used is the Latin Hypercube sampling method (McKay et al., 1979). Instead of selecting input values purely based on their probability distributions, this method favors sets of input values located in underrepresented areas of the input space (Helton & Davis, 2003). When considering smooth but computationally expensive models containing a limited number of input variables, a surrogate model (or metamodel) is often fitted to the original numerical model to mimic the relationship between input and output variables while improving computational efficiency (Jin et al., 2001; Queipo et al., 2005; Sudret, 2008). An alternate method is the point estimate method initially introduced by Rosenblueth (1975).

Based on a limited number of model evaluations, it aims at computing output statistical moments using statistical moments of the input variables. In this paper, uncertainty propagation was performed using our implementation of the physically based fully coupled hydro-morphodynamic model developed by Wu (2016). Thanks to the computational efficiency of this model, a regular Monte Carlo sampling method was adopted here.

3.1.3. Step 3: Results Exploitation

Uncertainty propagation results are generally exploited in two different ways. On the one hand, uncertainty quantification focuses on the output variability characterization, including output main statistical moments, for example, mean and variance, or complete probability distribution (Walker et al., 2003). On the other hand, sensitivity analyses highlight the relationships between uncertainties on specific input and output variables (Borgonovo & Plischke, 2016). Among them, global sensitivity analyses are generally preferred as derivatives and other local methods provide an incomplete picture of model response over the range of variability in the model inputs (Saltelli et al., 2019). In this context, screening techniques are widespread, with Morris method being the most common (Morris, 1991). This method evaluates two indicators. The first one reflects the mean output sensitivity to the variation of a specific input variable, while the second one describes the nonlinear behavior of this dependency within the input space (Iooss & Lemaître, 2015). Regression-based analyses provide an algebraic expression of the relationships between outputs and input variables (Helton et al., 2006). One of the simplest and most popular approaches is the least squares linearization technique, which consists in a multi-linear regression between model output and input variables. A coefficient, characterizing the relative contribution of each input uncertainty to the output variability, is assigned to each input. The larger this coefficient, the larger the impact of the considered input on the output variability. An example of more advanced regression-based analysis is the permutation feature importance, which relies on machine learning regression and evaluates how permutations of input parameters alter model predictions (Breiman, 2001). Moment-independent methods are usually preferred when the focus is set on the entire output distribution, especially when the output distribution is highly skewed or when it is multi-modal. The delta method computes sensitivity indicators that represent the normalized expected shift in the output distribution induced by uncertainties on a specific input variable (Borgonovo, 2006, 2007; Pianosi & Wagener, 2015). The use of a variance-based sensitivity analysis is shown to be particularly general in its applicability and in its capacity to reflect nonlinear processes and the effects of interactions among variables (Hall et al., 2009). The Sobol index of first order represents the output variability caused by a specific uncertain input subset. In contrast, the Sobol index of total order also includes the interactions of this subset with any other input subsets. Both indices aim at assessing the relative significance of input variables uncertainties (Iooss & Lemaître, 2015). It allows spotting the most critical input uncertainties as well as the non-influential ones, that is, factor prioritization and factor fixing.

Tables S1 and S2 in Supporting Information S1 provide an overview of existing studies applying uncertainty analysis to dam or dike breach models. In each of these studies, only a single case study was considered. Part of them solely focused on uncertainty quantification, for example, Abdedou et al. (2020), Froehlich (2008), Froehlich and Goodell (2012), Peter et al. (2018), Vorogushyn et al. (2011) and Westoby et al. (2015). Others performed global sensitivity analysis. Among them, Alhasan et al. (2016) and Bellos et al. (2020) used a Morris screening technique, whilst Ahmadisharaf et al. (2016), Goeury et al. (2022) and Sattar (2014) performed regression-based analyses. Goeury et al. (2022) also used the moment-independent delta method because of the presence of multi-modal output distributions. Variance-based methods, such as Sobol indices, were used by Kalinina et al. (2020), Pheulpin et al. (2020) and Tsai et al. (2019), but also considering each only a single case study. In contrast, in the present work, Sobol indices of total order were computed for each input subset in twenty-seven embankment configurations both at laboratory and field scales.

3.2. Input Distributions and Sampling Methods

Uncertain input variables are gathered in Table 1. They are divided into two types, namely model parameters that need to be set based on experimental observations and input parameters, for example, material characteristics and dam geometry parameters.

Values of the different input variables are selected randomly according to their respective probability distributions to feed the numerical model. It is therefore necessary to define a probability distribution function (PDF), or multivariate probability distribution assigned to each input variable or group of input variables, respectively.

The approach differs for variables whose values are assumed independent from each other and variables with dependent values. Both procedures are introduced hereafter.

3.2.1. Independent Input Variables

A Beta distribution has been assigned as a PDF to all independent variables due to its high versatility, wide-spread use in sensitivity analyses in environmental sciences and beyond (e.g., Benke et al., 2008, Hall et al., 2005, Kalinina et al., 2020; Mokhtari & Frey, 2005), and its finite support. This last feature is of great interest as it allows narrowing the analysis to behavioral input samples, for example, avoiding negative values for uncertain physical inputs (Wagener & Pianosi, 2019). It is expressed as

$$\text{PDF}(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \quad \text{with} \quad x = \frac{\tilde{x} - \tilde{x}_{\min}}{\tilde{x}_{\max} - \tilde{x}_{\min}} \in [0; 1], \tag{17}$$

where x is the normalized form of the considered variable $\tilde{x}$ that varies within the interval $[\tilde{x}_{\min}; \tilde{x}_{max}]$, α and β two real and positive parameters to be fixed and B the beta function.

In this work, we determined the values of parameters α and β from two features of the input variable x. First, we enforced the PDF to exhibit a maximum at a reference value, x_{ref}, so that

$$\frac{d\text{PDF}(x; \alpha, \beta)}{dx}\Big|_{x=x_{\text{ref}}} = 0 \quad \text{with } x_{\text{ref}} \in]0; 1[. \tag{18}$$

The second feature we considered is the variance of the variable x. However, due to the approximate knowledge of this characteristic, a qualitative approach based on expert knowledge is used here. One of the two beta distribution parameters is fixed manually while the other is determined using Equation 18. PDF characteristics of each uncertain input variable of the present study are presented in Table S4 in Supporting Information S1.

Note that the choice of a particular type of PDF and its parameters may have an influence on the global sensitivity analysis results. This aspect is discussed in Section 4.2.

3.2.2. Dependent Input Variables

A multivariate probability distribution is needed to represent input dependency properly and generate relevant input samples. The typical case of regression coefficients $\{C_a^*, C_b^*, C_c^*, C_d^*\}$ and $\{q_{b,a}^*, q_{b,b}^*\}$ obtained through experimental data fitting (Wu, 2016) is handled here. The presented methods (Methods 1 and 2) aim at generating sets of regression coefficients that lead to meaningful regression curves with regards to experimental observations (Figure S4 in Supporting Information S1). No explicit expression of the multivariate probability distribution is provided through these methods.

3.2.2.1. Method 1: Sample Generation Based on Bootstrapping

Let an empirical data vector containing n experimental points be written as

$$\boldsymbol{d} = \{d_1, d_2, \ldots, d_n\} \text{ with } d_i = (x_{\exp,i}, y_{\exp,i}). \tag{19}$$

Based on these n experimental points, the regression function, f, can be expressed as

$$y_{\text{reg}} = f(\boldsymbol{c}, x), \tag{20}$$

with $\boldsymbol{c}$ the vector containing the regression coefficients.

k samples of length n are generated with replacement from the experimental points of vector $\boldsymbol{d}$. This procedure, called *bootstrapping* (Davison & Hinkley, 1997), leads to k vectors of length n: $\{\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^k\}$. For each vector, the value of the parameters contained in $\boldsymbol{c}$ is determined to best fit the experimental data using least square minimization method in a trust-region algorithm, leading to k sets of regression coefficients: $\{\boldsymbol{c}^1, \boldsymbol{c}^2, \ldots, \boldsymbol{c}^k\}$. As illustrated in Figure S4 in Supporting Information S1, each set $\boldsymbol{c}^i$ can be considered as a meaningful random sample of the considered dependent inputs with regards to experimental observations. This physically meaningful method was used to generate results presented in Section 4. As a comparison, a different sampling method (Method 2) is presented hereafter.
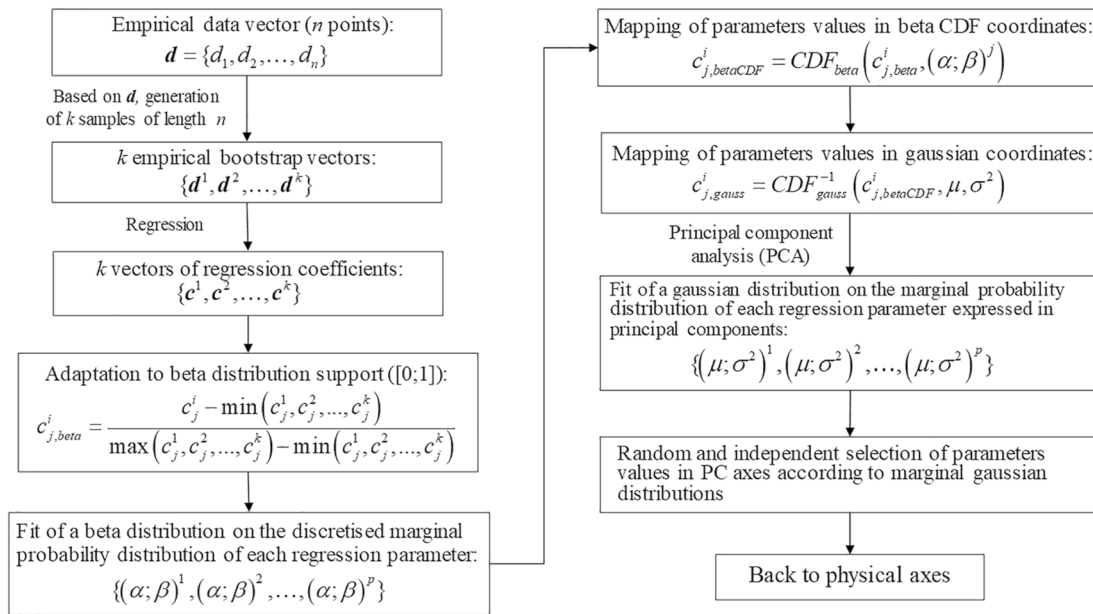
Figure 2. Flow chart of Method 2.

3.2.2.2. Method 2: Sample Generation Based on Parameters Decoupling

Based on a limited number of regressions k, this method aims at decoupling dependent regression parameters and generating independent marginal PDF. Conversely to Method 1, sets of regression coefficients can then be generated without performing any additional regression, reducing the computation time. In this context, a principal component analysis (PCA) is used to minimize the correlation between variables. However, zero correlation implies decoupling in the sole case where parameters follow a Gaussian distribution. This condition is met through a change of variables. Special care should be taken to avoid alteration of parameters dependency through the procedure. A flow chart of the method is provided in Figure 2.

First, k sets of p regression coefficients $\{c^1, c^2, \ldots, c^k\}$ are generated according to the procedure related to Method 1. Then, parameter values are adapted to fit a beta distribution support ([0;1]), so that

$$c^i_{j,\text{beta}} = \frac{c^i_j - \min(c^1_j, c^2_j, \ldots, c^k_j)}{\max(c^1_j, c^2_j, \ldots, c^k_j) - \min(c^1_j, c^2_j, \ldots, c^k_j)}, \tag{21}$$

where c^i_j is a value of the jth regression parameter observed in $\{c^1, c^2, \ldots, c^k\}$.

A beta distribution is then fitted on the discretized marginal probability distribution of each regression parameter using maximum likelihood estimates of the beta distribution parameters α and β (Fisher, 1912). The fitted beta distributions are then used to map the regression parameters in beta CDF coordinates, so that

$$c^i_{j,\text{betaCDF}} = \text{CDF}_{\text{beta}}\left(c^i_{j,\text{beta}}, (\alpha; \beta)^j\right), \tag{22}$$

with $(\alpha;\beta)^j$ the fitted parameters of the beta distribution related to the jth regression parameter.

The resulting regression parameters are then transposed in Gaussian coordinates:

$$c^i_{j,\text{gauss}} = \text{CDF}^{-1}_{\text{gauss}}\left(c^i_{j,\text{betaCDF}}, \mu, \sigma^2\right), \tag{23}$$

with μ and σ^2 the mean and the variance of the Gaussian distribution. For the sake of simplification, it was arbitrarily decided to use $\mu = 0$ and $\sigma^2 = 1$ in this study. Note that this choice has no implication on the results.

A principal component analysis (PCA) is then applied on the resulting data to minimize covariance. At this point, data approximately follow a Gaussian distribution. Then, Gaussian distributions are fitted on the data expressed

in the principal axes using maximum likelihood estimates of the Gaussian distribution parameters μ and σ^2. Based on those PDF, parameter values are selected randomly and independently before being transposed in the initial physical axes to constitute the desired samples.

Although this method does not provide an explicit form of the multivariate probability distribution, it generates independent PDF for each regression coefficients. The computation time required for the sampling procedure is thus reduced. Nonetheless, those PDF are approximations of the multivariate probability distribution, and the accuracy of the results should thus be assessed systematically.

In this work, weir efficiency coefficients (c_1, c_2) are assumed dependent and varying proportionally. Their probability distributions are indirectly characterized through a linear mapping, so that

$$\begin{cases} c_1 = c_{\text{eff}} c_1^{\text{ref}} \\ c_2 = c_{\text{eff}} c_2^{\text{ref}} \end{cases} \text{with } c_{\text{eff}} \text{ a random variable with } c_{\text{eff}}^{\text{ref}} = 1. \tag{24}$$

3.3. Global Sensitivity Analysis

The procedure of input uncertainty characterization was presented for independent and dependent input variables in Section 3.2. Those uncertainties were propagated to model outputs through our implementation of the numerical model of Wu (2016). Statistical descriptors are used to quantify the resulting uncertainties and to characterize the relation between uncertainties on input and output variables. In this work, Sobol indices of total order were used to highlight the most influential input uncertainties. It can be expressed as the ratio between significance descriptors and the global output variance.

Let the random variable $\boldsymbol{X} = (X_1, \ldots, X_{m_1})$ be partitioned into n statistically independent random subsets of variables, so that $\boldsymbol{X}^1 = (X_1, \ldots, X_{m_1})$, $\boldsymbol{X}^2 = (X_{m_1+1}, \ldots, X_{m_2})$, ..., $\boldsymbol{X}^n = (X_{m_1+\ldots+m_{n-1}+1}, \ldots, X_{m_1+\ldots+m_n})$. The significance descriptor related to subset j may be approximated as (Sobol, 2001; Arnst & Ponthot, 2014):

$$s_{\boldsymbol{X}^j} = E\left[V\left(g\left(\boldsymbol{X}^j, \boldsymbol{x}^{\sim j}\right)\right)\right] \approx s_{\boldsymbol{X}^j}^\nu = \frac{1}{2\nu} \sum_{l=1}^\nu \left[g\left(\boldsymbol{x}_l^j, \boldsymbol{x}_l^{\sim j}\right) - g\left(\tilde{\boldsymbol{x}}_l^j, \boldsymbol{x}_l^{\sim j}\right)\right]^2, \tag{25}$$

where $\boldsymbol{X}^{\sim j}$ denotes the random variable that contains all the components of $\boldsymbol{X}$ that are not components of $\boldsymbol{X}^j$, $P_{\boldsymbol{X}^{\sim j}}$ and $P_{\boldsymbol{X}^j}$ the probability distributions of $\boldsymbol{X}^{\sim j}$ and $\boldsymbol{X}^j$, respectively, $\{\boldsymbol{x}_l^j, 1 \le l \le \nu\}$ and $\{\tilde{\boldsymbol{x}}_l^j, 1 \le l \le \nu\}$ two independent ensembles of independent and identically distributed (i.i.d.) samples from $P_{\boldsymbol{X}^j}$ and $\{\boldsymbol{x}_l^{\sim j}, 1 \le l \le \nu\}$ an independent ensemble of i.i.d. samples from $P_{\boldsymbol{X}^{\sim j}}$, and g the model output function. As expressed Equation 25, the significance descriptor associated to $\boldsymbol{X}^j$ represents the mean output variance obtained when the values of all input variables are fixed, except those related to input parameters contained in $\boldsymbol{X}^j$. Sobol indices of total order are finally obtained by dividing this indicator by the total output variance, that is, the variance obtained when all input variables are unknown and vary according to their respective PDF.

In Equation 25, the larger ν the better the approximation of the significance descriptor. Noting that $g\left(\boldsymbol{x}_l^j, \boldsymbol{x}_l^{\sim j}\right)$ and $g\left(\tilde{\boldsymbol{x}}_l^j, \boldsymbol{x}_l^{\sim j}\right)$ correspond to output values of two different runs, 2ν runs are required to evaluate Equation 25.

4. Results and Discussion

This section starts with the comparison of the two sampling methods of dependent inputs introduced in Section 3.2.2 and in Text S1 in Supporting Information S1. Using Method 1, a global sensitivity analysis is then applied to each configuration presented in Section 2.2. The focus is set on two model outputs, namely the peak breach discharge, $Q_{b,\text{peak}}$, and the time to peak, t_{peak}, that is, time between the overtopping initiation and the peak breach discharge occurrence. All information used to characterize the independent input variables uncertainty is gathered in Table S4 in Supporting Information S1.

4.1. Comparison of Dependent Input Samples Generation Methods

Before examining global sensitivity analysis results, we compare Method 1 and Method 2 used to generate samples of dependent input variables (Section 3.2.2). Figure S4 in Supporting Information S1 shows regression curves

Table 3
Comparison of Computation Time Required to Generate Regression Coefficient Samples With Method 1 ($k = \nu = 4,000$) and Method 2 ($k = 1,000$; $\nu = 4,000$)

	Method 1	Method 2			
	Regression procedure = sample generation ($\nu = k = 4,000$)	Regression procedure ($k = 1,000$)	Sample generation ($\nu = 4,000$)	Total time	Computation time reduction
$\{C_a^*, C_b^*, C_c^*, C_d^*\}$	411.9 [s]	103 [s]	0.6 [s]	103.6 [s]	**−75%** (−308.3 [s])
$\{q_{b,a}^*, q_{b,b}^*\}$	5.1 [s]	1.3 [s]	0.2 [s]	1.5 [s]	**−71%** (−3.6 [s])

obtained using random samples of $\{C_a^*, C_b^*, C_c^*, C_d^*\}$ and $\{q_{b,a}^*, q_{b,b}^*\}$ generated using Method 1. All regression curves seem in agreement with experimental point clouds, proving that Method 1 provides meaningful random samples. Conversely, Method 2 provides results that approximate the results obtained by Method 1, but computational cost is reduced. The objective is to determine which method presents the best trade-off between computational efficiency and accuracy. In this analysis, results are computed using Test 10 (reference configuration).

Sobol indices of total order were computed for the maximum breach discharge, $Q_{b,\text{peak}}$, and the related elapsed time since overtopping initiation, t_{peak}. Using Method 1, relative variations in Sobol indices are smaller than 1% when $\nu > 4,000$ (Figure S5 in Supporting Information S1) and the results are thus considered as converged. Figure S6 in Supplement compares those values with results obtained by Method 2 using $k = 500$, $k = 1,000$ and $k = 2,000$. In all cases, results converged for $\nu = 4,000$. Increasing the value of k does not improve accuracy on Sobol indices related to $Q_{b,\text{peak}}$. Regarding t_{peak}, a significant improvement is observed when k rises from 500 to 1,000, with no additional accuracy gain when k reaches 2,000. Therefore, $k = 1,000$ was chosen. The maximum error on Sobol indices is about 5% compared to Method 1 for both model outputs.

The computing time required to generate regression samples using Method 1 ($k = \nu = 4,000$) and Method 2 ($k = 1,000$; $\nu = 4,000$) is presented in Table 3. Method 2 is about four times faster than Method 1. Nonetheless, the time required to perform a numerical simulation of Test 10 is about 0.21 s. Since 4,000 runs should be launched for each of the 21 parameters or parameters subsets, the total simulation time reaches about 18,500 s. The computation time saved on the regression coefficient samples generation is thus marginal in the present case. For this reason and to optimize results accuracy, Method 1 has been used to lead the global sensitivity in the following section.

4.2. Application of the Global Sensitivity Analysis

Sobol indices of total order related to the peak breach discharge and the time to peak are presented in Figures 3a and 3b for laboratory-scale configurations and in Figures 4a and 4b for field-scale configurations. In those figures, the y-axis corresponds to the values of the Sobol indices while numbers along the x-axis refer to the tested configurations listed in Table 2. Circles aligned on a same vertical line refer to Sobol indices computed in a similar test configuration. One color is associate to each input variable. Only parameters with a Sobol index larger than 5% for at least one configuration are presented. Vertical dotted lines separate groups of configurations in which the value of a single parameter is varied.

When considering laboratory-scale configurations, uncertainties in seven input parameters appear to be critical for both outputs of interest, namely c_{eff}, A_n, $A_{n'}$, n_{min}, θ_{cr}, λ and h_d. Additionally, uncertainties in the inflow discharge (Q_{in}) play a major role in the determination of $Q_{b,\text{peak}}$, while t_{peak} is also slightly impacted by uncertainties in q_b^* parameters, that is, $q_{b,a}^*$ and $q_{b,b}^*$.

Results appear to be particularly different between configurations with no inflow discharge but an initial water level slightly higher than the initial breach depth (Tests 36 and 40 to 44) and tests with an inflow discharge (all other tests). In the latter case, uncertainties in input parameters affecting erosion, that is, effective shear stress and sediment concentration are almost systematically the most critical for both outputs, as already pointed in many previous works (Alhasan et al., 2016; Kalinina et al., 2020; Peter et al., 2018; Sattar, 2014; Westoby et al., 2015). Those include A_n (parameter in Strickler formula), θ_{cr} (critical Shields parameter) and λ (empirical coefficient
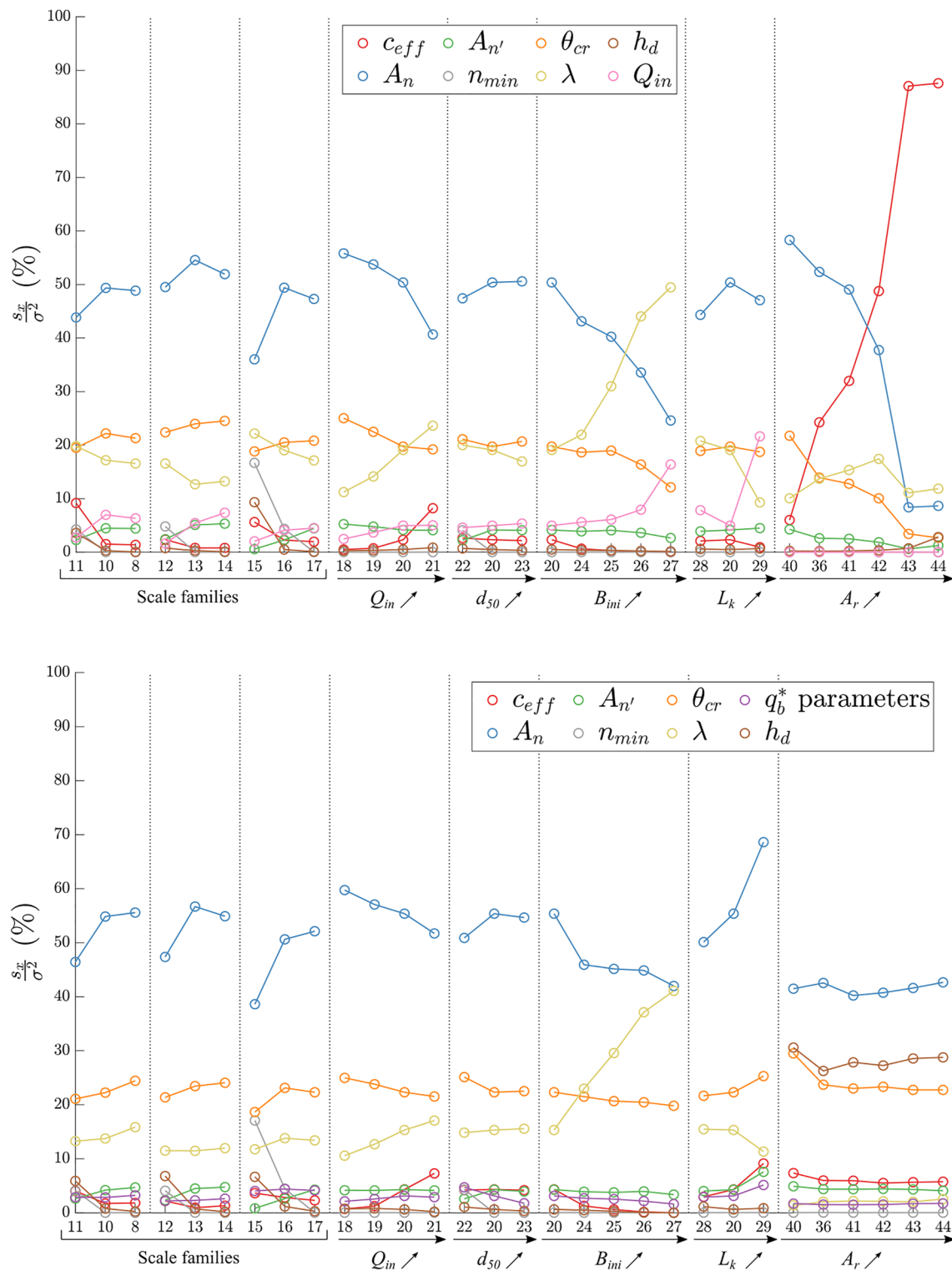
Figure 3. Sobol indices of total order for (a) peak breach discharge, and (b) time to peak in laboratory-scale configurations (numbers = Frank's, 2016 test ID).

involved in mixing length computation). They dictate how fast the breach expands and are therefore key parameters in the breaching process modeling. Surprisingly, the influence of uncertainties in q_b^* and C_* parameters is mostly negligible for both outputs, that is, influence of parameters related to bed load transport capacity and suspended load concentration, respectively. This can however be understood by considering their low uncertainty
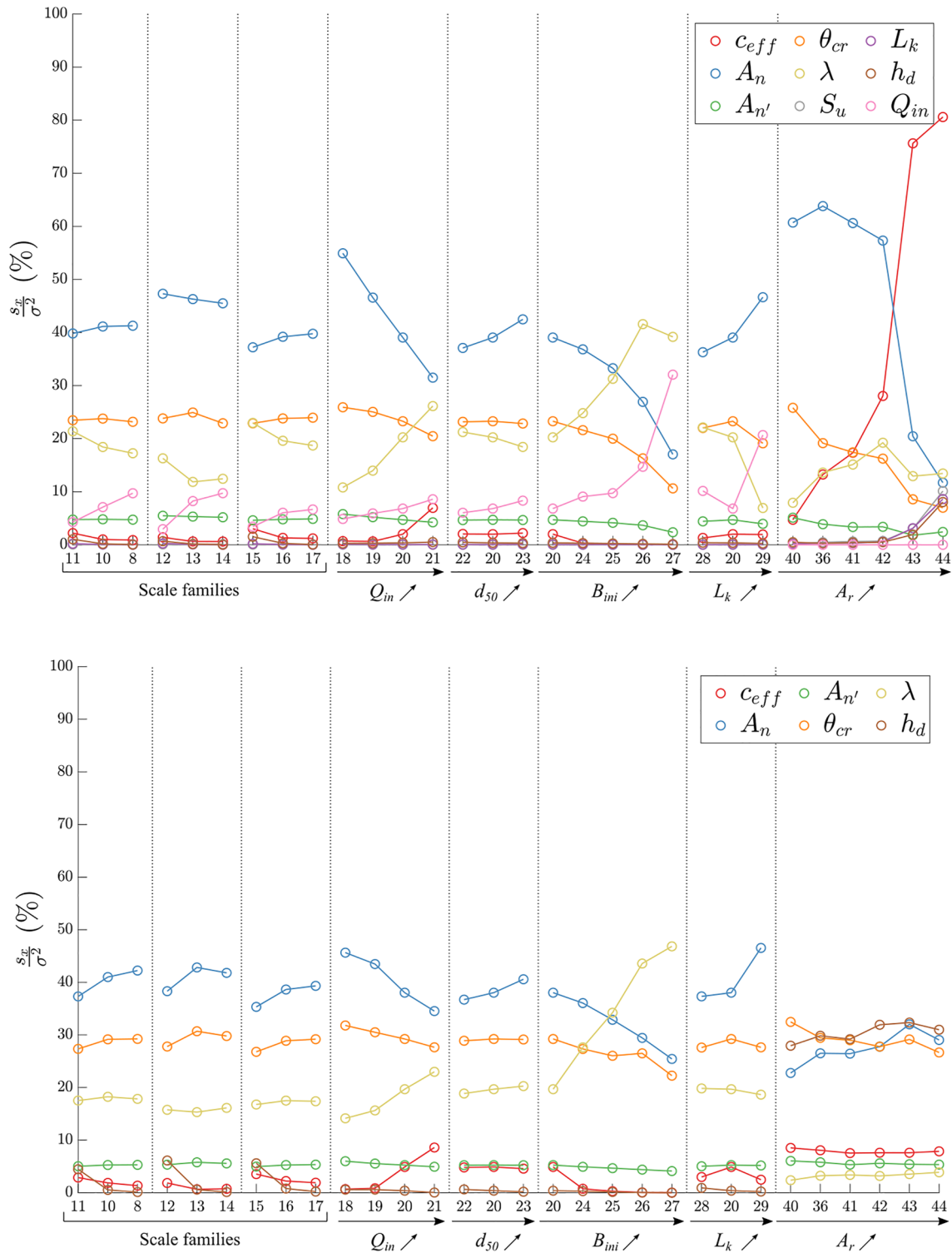
Figure 4. Sobol indices of total order for (a) peak breach discharge, and (b) time to peak in field-scale configurations (numbers = Frank's, 2016 test ID).

resulting from the bootstrapping method presented in Section 3.2.2 for input parameters with dependent PDF. A small uncertainty associated to parameter $A_{n'}$ also explains why it exhibits Sobol indices of negligible values.

When increasing the inflow discharge (Tests 18 to 21), Sobol indices related to bed load concentration (A_n and θ_{cr}) and the ones related to λ exhibit opposite trends for both outputs. As illustrated in Figure S7 in Supporting

Information S1, suspended load corresponds to a larger fraction of the overall sediment concentration when the inflow discharge rises, to the expense of bed load. This induces a smaller relative impact of uncertainties in A_n and θ_{cr}, which leads to a higher relative influence of uncertainties in λ. Similar trends are observed when the initial breach width enlarges (Tests 20 and 24 to 27). The mixing length, L_s, involved in Equation 10 is the product of λ and the water free surface width on the considered reach, B_{ini}. By increasing the initial breach width, B_{ini} increases as well and uncertainties in λ have more impact on L_s, which strongly influences erosion dynamics. In this case, Sobol indices related to λ rise substantially to the expense of other parameters.

In Tests 27 (large initial breach width) and 29 (large dike crest width), the relative impact of uncertainties in Q_{in} becomes remarkable in the determination of $Q_{b,peak}$. In Test 27, the initial breach width is particularly large. In this case, additional breach expansion becomes less decisive for the maximum breach discharge value. This trend is illustrated in Figure S8 in Supporting Information S1, which shows a decreasing influence of uncertainties in input parameters when the initial breach width increases, except for the inflow discharge coefficient. In Test 29 (large dam crest width), the breaching initiation phase (i.e., phase of limited overtopping flow and slow erosion) is particularly long due to the extremely large dike crest that must be eroded. The larger the dike crest, the smoother the resulting breach hydrograph, as highlighted by Schmitz et al. (2021) for fluvial dikes. Therefore, the maximum breach discharge becomes less sensitive to variations in input parameters, as highlighted in Figure S8 in Supporting Information S1. Nonetheless, the relation between uncertainties in Q_b and Q_{in} is not altered, and the value of the inflow discharge coefficient becomes more important in the determination of $Q_{b,peak}$. Conversely, t_{peak} is extremely dependent on the duration of the breaching initiation phase, for which influence of erosion related parameters is predominant. For this reason, uncertainties in those parameters keep a dominant role for the determination of t_{peak}.

Although being negligible in most tested configurations, the impact of uncertainties in parameter n_{min} becomes more significant in Test 15 (very small median grain size) because the value of n_{min} is only considered in the computation of Manning's coefficients (Equations 4 and 7) when the median grain size, d_{50}, is particularly small. In this case, n_{min} highly impacts the erosion process by replacing the role of parameter A_n. In the opposite case, its value is not seen at all by the numerical model and its Sobol index becomes zero.

When no inflow discharge is injected (Tests 36 and 40 to 44), observations are quite different. When the reservoir area increases, the water level in the reservoir decreases more slowly, leading to a greater water level above the breach bottom, h_{top}. Through Equation 2, the influence of c_{eff} on Q_b and $Q_{b,peak}$ grows substantially, to the expense of the relative influence of erosion-related parameters. The dam breaching process presents in this case a rather long initiation phase because it is only driven by a small difference between water level in the reservoir and initial breach depth. For this reason, the dam height, h_d, is critical to determine the duration of this initial phase and thus t_{peak}, so do erosion parameters.

Trends observed for field-scale configurations (Figures 4a, 4b and Figure S9 in Supporting Information S1) are extremely similar to the ones presented for laboratory-scale tests. Though, several differences deserve to be highlighted. Due to the increased value of the median grain size in field-scale configurations, parameter n_{min} plays no role anymore. Conversely, Sobol indices related to S_u and L_k slightly rise for $Q_{b,peak}$. This difference is, however, not significant since they hardly rise above the 5% limit value while being just below this threshold in laboratory-scale configurations. The same explanation holds for Sobol indices of q_b^* parameters related to t_{peak}. The small impact of scaling on sensitivity analysis results is confirmed by configurations gathered in the three scale families (Tests 8 and 10 to 17). Besides the punctual significant impact of n_{min} in Test 15 (very small median grain size), results are poorly influenced by limited scale modifications in laboratory and field-scale cases. When analyzing a real-scale dam breaching with the presented numerical model, this observation is of great interest since it suggests that a sensitivity analysis might be applied to a down-scaled embankment to reduce computational cost while leading to similar results.

While sensitivity analysis allows ranking parameters according to their relative impact on specific model outputs, uncertainty quantification assesses the magnitude of the output uncertainties. Both aspects should be combined to properly underline the most critical parameter uncertainties, that is, parameter uncertainties with a critical impact on highly uncertain outputs. Based on the sensitivity analysis led in this study and uncertainty quantification results illustrated in Figures S10, S11 and S12 in Supporting Information S1, this combined approach was implemented in Figure 5 to generate a decision tree. This diagram aims at identifying which parameter uncertainties should be minimized in priority, according to the output uncertainty in the considered configuration. A_n turns to
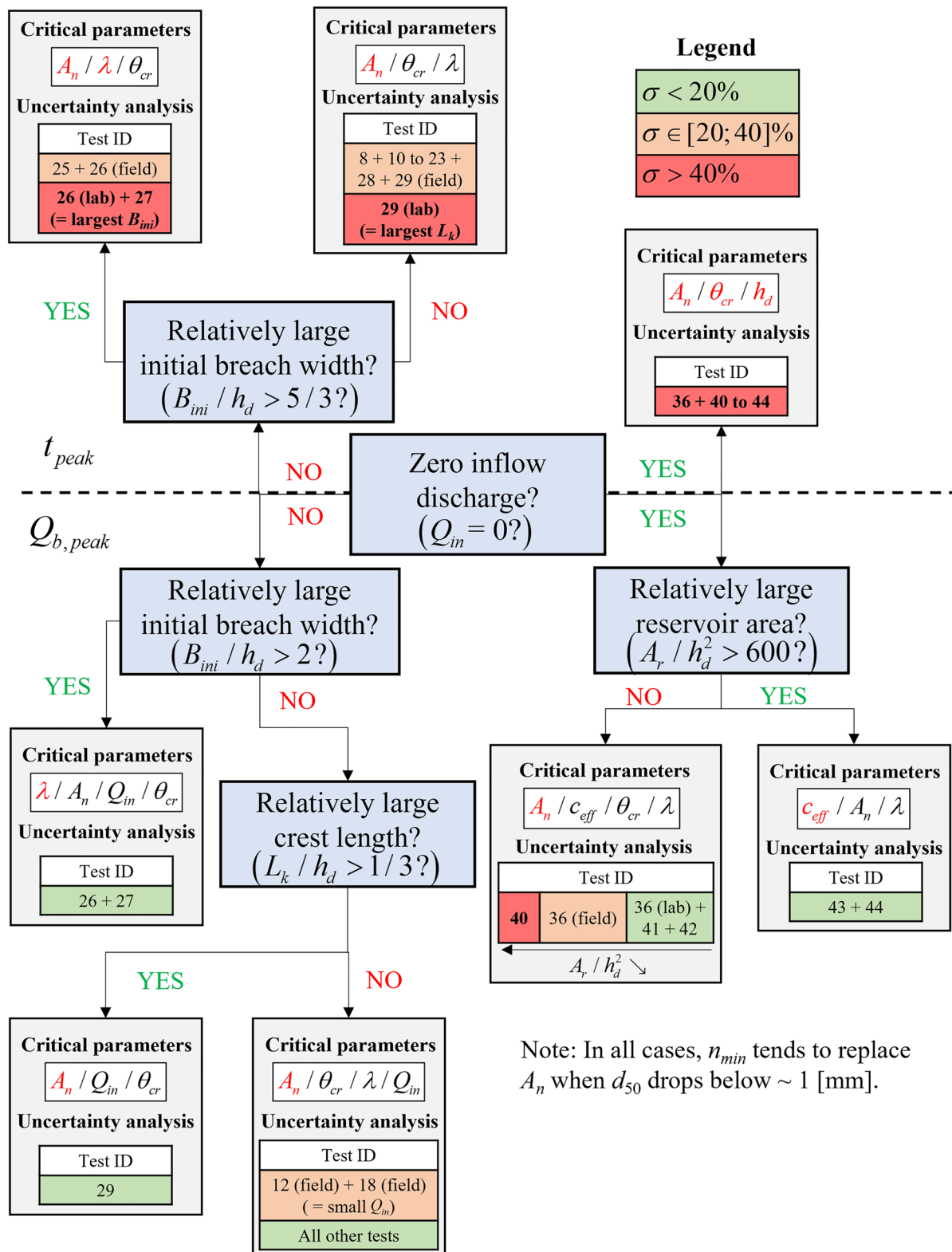
Figure 5. Decision tree indicating most influential parameters (leading parameters are in red) along with output uncertainties magnitude. σ refers to standard deviation values displayed in Figure S12 in Supporting Information S1.

be the most critical parameter globally. Though, when the median grain size, d_{50}, drops below around 1 mm, the impact of n_{min} becomes predominant and tends to cancel the influence of A_n.

Overall, these results suggest that uncertainties in the estimation of the value of Manning's coefficient, n, considerably influence the model outputs. This aspect is confirmed in Figure S14 in Supporting Information S1, in which relative variations in the calculated bed load transport capacity are given as a function of predefined relative variations in the value of parameter n. It shows that a predefined relative variation in parameter n leads to a larger relative variation in the bed load transport capacity. This is due to the model structure and equations, which involve parameter n at a power above unity, such as in Equation 5 and indirectly in Equation 9. Hence a given degree of uncertainty in n leads to greater degree of uncertainty in the model outcomes. This issue is strongly intertwined with the assumption of uniform flow on the downstream face of the dam (Equation 3). Although widespread (e.g., Cai et al., 2022; Macchione, 2008; Shen et al., 2020), this assumption remains an intrinsic limitation of simplified models, while a spatially distributed (e.g., in 2D) approach would be needed to properly capture the complex flow processes involved in dam breach hydraulics.

To assess the influence on the results of the assumed type of PDF, we have undertaken new computations in which the Beta distributions associated to the most influential inputs (c_{eff}, A_n, Q_{in}, θ_{cr}, λ and h_d, as identified in Figure 5) were replaced by truncated normal distributions. Each truncated normal distribution was defined using the same information as used to generate the corresponding Beta distribution (Section 3.2.1), that is, the mode, the variance, and the variation interval. Figures S12 and S13 in Supporting Information S1 show that this modification has no significant impact on the sensitivity analysis and uncertainty quantification results, that is, the decision tree provided in Figure 5 is not altered.

An alternate way to test the sensitivity of the results to the input uncertainty characterization is to keep the same probability distribution function, but to vary its parametrization, for example, its mean and variance. Bello et al. (2022) led this kind of analysis on a limited number of input variables to which a normal probability distribution function was assigned. They independently varied the mean and variance of each input PDF by ±50% and assessed the impact on the sensitivity analysis results. It appeared that the relative impact of the input uncertainties on the model output remained qualitatively the same, for example, the most critical input uncertainty did not vary. However, the magnitude of the output uncertainties was considerably influenced. Unsurprisingly, it is expected that the larger the variations in the mean, variance, or variation interval extremities of the input probability distribution functions, the greater the impact on the analysis results. Though, an in-depth exploration of combinations of PDF parametrizations gets cumbersome when considering many inputs variables.

The results presented here highlight the need to conduct a global sensitivity analysis and an uncertainty quantification for each specific configuration of interest, especially when the main physical processes represented by the model vary from one configuration to the other (e.g., in the present case, when Q_{in} is equal to or different from zero). This conclusion applies also to numerous other applications in water resources management and environmental sciences, in which the relative weight of various physical processes represented in the model changes with the modeled configuration.

5. Conclusion

This paper highlights the need for performing sensitivity analysis to identify most critical input variable uncertainties in embankment breaching numerical models. It also demonstrated the dependency between sensitivity/uncertainty quantification results and test configuration. Uncertainty propagation was performed using our implementation of the model of Wu (2016), in which 21 sets of input variables were deemed uncertain. The performance of two new sampling methods of dependent inputs were assessed, showing that the direct bootstrap sampling method (Method 1) exhibited the best accuracy, although being more computationally demanding.

Probability distributions were associated to independent input variables. Two new methods were proposed to generate samples of regression coefficients (i.e., dependent input variables) complying with field data. Based on available empirical data, Method 1 generated new data clouds by bootstrapping and computed associated regression coefficients for each of them. This method led to meaningful results but had poor computational performance because a regression is performed for each input set. Method 2 used variable changes and PCA to decouple input variables and allow defining independent marginal probability distributions. Although slightly less accurate, this method reduced dependent inputs sampling procedure computation time by more than 70%

compared to Method 1. However, in the current study this gain was deemed negligible compared to the time required to run simulations using Monte Carlo procedure. Method 1 was finally used to lead the sensitivity analysis presented in this work.

Sobol indices of total order were computed for each input variable in twenty-seven configurations at both laboratory and field scales. It allowed assessing the relative impact of input uncertainties on the variability of two model outputs (peak breach discharge and time to reach the peak). Results were particularly different between configurations with no inflow discharge but an initial water level slightly higher than the initial breach invert level and tests with an inflow discharge. Conversely, the sensitivity of model predictions with respect to uncertainties in input parameters remains similar irrespective of the considered scale (laboratory vs. field scales), that is, there was almost no difference between the Sobol indices obtained for laboratory-scale and for field-scale configurations. Parameter A_n turned to be the most critical parameter globally. Though, when the median grain size dropped below 1 mm, the impact of $n_{\min}$ became predominant and canceled A_n influence. It comes out that uncertainties in Manning's coefficient value have a critical impact on model outputs variability. However, an accurate determination of this coefficient is not possible in practice, as it is a model parameter and not a measured quantity. This issue, together with the assumption of uniform flow on the dam downstream face, point at a limitation of simplified models, which may only be truly overcome by adopting a spatially distributed approach (e.g., 2D model) but at the expense of a considerably higher computational burden.

To further generalize the findings of this study, the influence of choosing a particular type of PDF (Beta distribution) to characterize input uncertainties was assessed. The calculations were repeated with an alternate type of distribution functions. The PDF associated to the parameter uncertainties influencing most the output variability were defined as truncated normal distributions instead of the initially assumed Beta distributions. It turned out that this modification has no substantial impact on the analysis results. In contrast, extensive modifications in the PDF parametrization (e.g., strong modification of the variation interval or the variance) are expected to have more influence on the analysis results (Reed et al., 2022).

The present work allows identifying parameters whose uncertainty is critical for model outputs, depending on the configuration considered. The relevance of using highly uncertain model parameters which strongly influence the results may be questioned. Also, the reliability of the numerical model may be partly assessed by looking at the magnitude of the output uncertainties related to a given configuration. Finally, the procedure described in this paper may be used to determine whether a modification in the model structure brings a notable improvement in results accuracy but also a reduction of the output uncertainty ranges. The global sensitivity analysis procedure and conclusions presented in this study are not limited to the sole case of dam breaching but are also of relevance to a wide variety of modeling applications in environmental sciences. The global sensitivity analysis procedure presented in Section 3 is easily transferrable to other non-linear computational models containing independent and dependent uncertain scalar input variables. Its use is particularly relevant when the objective is to understand which input uncertainties are the most critical (i.e., factor prioritization) and which ones are negligible (i.e., factor fixing). This information allows simplifying subsequent analyses of the numerical model in the considered configuration as some uncertainties may be discarded.

To further generalize our findings and increase their robustness, similar studies should be performed using different numerical models of different types, additional test configurations and different sensitivity indicators, for example, delta indicator (Borgonovo, 2007). These observations would further support decision-making processes in risk management (Wagener & Pianosi, 2019). When dealing with non-scalar sources of uncertainties, that is, spatially distributed or time-varying uncertain inputs, other more suitable techniques exist to generate relevant input samples (e.g., Baroni & Tarantola, 2014; Lilburne & Tarantola, 2009). With computational demanding models, the Monte Carlo procedure used in this work should be replaced by an optimized sampling method (e.g., Latin Hypercube Sampling). In this case, Method 2 used to generate samples of dependent inputs might lead to significant computational time saving overall.

Data Availability Statement

Data were not used, nor created for this research.

References

Abdedou, A., Soulaïmani, A., & Tchamen, G. W. (2020). Uncertainty propagation of dam break flow using the stochastic non-intrusive B-splines Bézier elements-based method. *Journal of Hydrology*, *590*, 125342. https://doi.org/10.1016/j.jhydrol.2020.125342

Ahmadisharaf, E., Kalyanapu, A. J., Thames, B. A., & Lillywhite, J. (2016). A probabilistic framework for comparison of dam breach parameters and outflow hydrograph generated by different empirical prediction methods. *Environmental Modelling & Software*, *86*, 248–263. https://doi.org/10.1016/j.envsoft.2016.09.022

Alhasan, Z., Duchan, D., & Říha, J. (2016). The probabilistic solution of dike breaching due to overtopping. *Proceedings of the 4th European congress of the International association of hydroenvironment engineering and research, IAHR 2016*. (pp. 505–512).

Arnst, M., & Ponthot, J. (2014). An overview of nonintrusive characterization, propagation, and sensitivity analysis of uncertainties in computational mechanics. *International Journal for Uncertainty Quantification*, *4*(5), 387–421. https://doi.org/10.1615/int.j.uncertaintyquantification.2014006990

ASCE. (2011). Earthen embankment breaching. *Journal of Hydraulic Engineering*, *137*(12), 1549–1564. https://doi.org/10.1061/(asce)hy.1943-7900.0000498

Baroni, G., & Tarantola, S. (2014). A general probabilistic framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. *Environmental Modelling & Software*, *51*, 26–34. https://doi.org/10.1016/j.envsoft.2013.09.022

Bello, D., Alcayaga, H., Caamaño, D., & Pizarro, A. (2022). Influence of dam breach parameter statistical definition on resulting rupture maximum discharge. *Water*, *14*.

Bellos, V., Tsakiris, V. K., Kopsiaftis, G., & Tsakiris, G. (2020). Propagating dam breach parametric uncertainty in a river reach using the HEC-RAS software. *Hydrology*, *7*(4), 1–14. https://doi.org/10.3390/hydrology7040072

Benke, K. K., Lowell, K. E., & Hamilton, A. J. (2008). Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. *Mathematical and Computer Modelling*, *47*(11–12), 1134–1149. https://doi.org/10.1016/j.mcm.2007.05.017

Borgonovo, E. (2006). Measuring uncertainty importance: Investigation and comparison of alternative approaches. *Risk Analysis*, *26*(5), 1349–1361. https://doi.org/10.1111/j.1539-6924.2006.00806.x

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, *92*(6), 771–784. https://doi.org/10.1016/j.ress.2006.04.015

Borgonovo, E., & Plischke, E. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, *248*(3), 869–887. https://doi.org/10.1016/j.ejor.2015.06.032

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Cai, Y., Zhang, X., Xue, R., Wang, M., & Deng, Q. (2022). Numerical simulation of overtopping breach processes caused by failure of landslide dams. *Environmental Fluid Mechanics*, *22*(4), 839–863. https://doi.org/10.1007/s10652-022-09858-1

Cantero-Chinchilla, F. N., Castro-Orgaz, O., & Dey, S. (2019). Prediction of overtopping dike failure: Sediment transport and dynamic granular bed deformation model. *Journal of Hydraulic Engineering*, *145*(6). https://doi.org/10.1061/(asce)hy.1943-7900.0001608

Chen, Z. Y., Ping, Z. Y., Wang, N. X., Yu, S., & Chen, S. J. (2019). An approach to quick and easy evaluation of the dam breach flood. *Science China Technological Sciences*, *62*(10), 1773–1782. https://doi.org/10.1007/s11431-018-9367-4

Damgaard, J. S., Whitehouse, R. J., & Soulsby, R. L. (1997). Bed-load sediment transport on steep longitudinal slopes. *Journal of Hydraulic Engineering*, *123*(12), 1130–1138. https://doi.org/10.1061/(asce)0733-9429(1997)123:12(1130)

Da Veiga, S., Gamboa, F., Iooss, B., & Prieur, C. (2021). *Basics and trends in sensitivity analysis: Theory and practice in R, computational science & engineering* (Vol. CS23). Society for Industrial and Applied Mathematics.

Da Veiga, S., Wahl, F., & Gamboa, F. (2009). Local polynomial estimation for sensitivity analysis on models with correlated inputs. *Technometrics*, *51*(4), 452–463. https://doi.org/10.1198/tech.2009.08124

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge university press.

De Lorenzo, G., & Macchione, F. (2014). Formulas for the peak discharge from breached Earthfill dams. *Journal of Hydraulic Engineering*, *140*(1), 56–67. https://doi.org/10.1061/(asce)hy.1943-7900.0000796

Fisher, R. A. (1912). On an absolute criterion for fitting frequency curves. *Messenger of Mathematics*, *41*, 155–156.

Frank, P.-J. (2016). *Hydraulics of spatial dike breaches. Ph.D. thesis*. ETH Zurich.

Froehlich, D. C. (2008). Embankment dam breach parameters and their uncertainties. *Journal of Hydraulic Engineering*, *134*(12), 1708–1721. https://doi.org/10.1061/(asce)0733-9429(2008)134:12(1708)

Froehlich, D. C., & Goodell, C. R. (2012). Breach of duty (not): Evaluating the uncertainty of dam-breach flood predictions. In *Proceedings of world environmental and water resources congress 2012: Crossing boundaries*.

Goeury, C., Bacchi, V., Zaoui, F., Bacchi, S., Pavan, S., & El Kadi Abderrezzak, K. (2022). Uncertainty assessment of flood hazard due to levee breaching. *Water*, *14*(23), 3815. https://doi.org/10.3390/w14233815

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, *48*(8). https://doi.org/10.1029/2011wr011044

Hall, J. W., Boyce, S. A., Wang, Y., Dawson, R. J., Tarantola, S., & Saltelli, A. (2009). Sensitivity analysis for hydraulic models. *Journal of Hydraulic Engineering*, *135*(11), 959–969. https://doi.org/10.1061/(asce)hy.1943-7900.0000098

Hall, J. W., Tarantola, S., Bates, P. D., & Horritt, M. S. (2005). Distributed sensitivity analysis of flood inundation model calibration. *Journal of Hydraulic Engineering*, *131*(2), 117–126. https://doi.org/10.1061/(asce)0733-9429(2005)131:2(117)

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, *81*(1), 23–69. https://doi.org/10.1016/s0951-8320(03)00058-9

Helton, J. C., Johnson, J. D., Sallaberry, C. J., & Storlie, C. B. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, *91*(10–11), 1175–1209. https://doi.org/10.1016/j.ress.2005.11.017

Hou, T., Nuyens, D., Roels, S., & Janssen, H. (2019). Quasi-Monte Carlo based uncertainty analysis: Sampling efficiency and error estimation in engineering applications. In *Reliability engineering and system safety 191*.

Iooss, B., & Lemaître, P. (2015). *A review on global sensitivity analysis methods*. Operations Research/ Computer Science Interfaces Series.

Jacques, J., Lavergne, C., & Devictor, N. (2006). Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering & System Safety*, *91*(10–11), 1126–1134. https://doi.org/10.1016/j.ress.2005.11.047

Janssen, H. (2013). Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence. *Reliability Engineering & System Safety*, *109*, 123–132. https://doi.org/10.1016/j.ress.2012.08.003

Jin, R., Chen, W., & Simpson, T. W. (2001). Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization*, *23*, 1–13. https://doi.org/10.1007/s00158-001-0160-4

Kalinina, A., Spada, M., Vetsch, D. F., Marelli, S., Whealton, C., Burgherr, P., & Sudret, B. (2020). Metamodeling for uncertainty quantification of a flood wave model for concrete dam breaks. *Energies*, *13*(14), 3685. https://doi.org/10.3390/en13143685

Kiureghian, A. D., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020

Lee, K. (2019). Simulation of dam-breach outflow hydrographs using water level variations. *Water Resources Management*, *33*(11), 3781–3797. https://doi.org/10.1007/s11269-019-02341-5

Li, G., Rabitz, H., Yelvington, P. E., Oluwole, O. O., Bacon, F., Kolb, C. E., & Schoendorf, J. (2010). Global sensitivity analysis for systems with independent and/or correlated inputs. *Journal of Physical Chemistry A*, *114*(19), 6022–6032. https://doi.org/10.1021/jp9096919

Li, Y., Chen, A., Wen, L., Bu, P., & Li, K. (2020). Numerical simulation of non-cohesive homogeneous dam breaching due to overtopping considering the seepage effect. *European Journal of Environmental and Civil Engineering*.

Lilburne, L., & Tarantola, S. (2009). Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, *23*(2), 151–168. https://doi.org/10.1080/13658810802094995

Macchione, F. (2008). Model for predicting floods due to earthen dam breaching. I: Formulation and evaluation. *Journal of Hydraulic Engineering*, *134*(12), 1688–1696. https://doi.org/10.1061/(asce)0733-9429(2008)134:12(1688)

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

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341. https://doi.org/10.1080/01621459.1949.10483310

Mokhtari, A., & Frey, H. C. (2005). Sensitivity analysis of a two-dimensional probabilistic risk assessment model using analysis of variance. *Risk Analysis*, *25*(6), 1511–1529. https://doi.org/10.1111/j.1539-6924.2005.00679.x

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.1080/00401706.1991.10484804

Onda, S., Hosoda, T., Jaćimović, N. M., & Kimura, I. (2019). Numerical modelling of simultaneous overtopping and seepage flows with application to dike breaching. *Journal of Hydraulic Research*, *57*(1), 13–25. https://doi.org/10.1080/00221686.2018.1442882

Pappenberger, F., & Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, *42*(5). https://doi.org/10.1029/2005wr004820

Peter, S. J., Siviglia, A., Nagel, J., Marelli, S., Boes, R. M., Vetsch, D., & Sudret, B. (2018). Development of probabilistic dam breach model using Bayesian inference. *Water Resources Research*, *54*(7), 4376–4400. https://doi.org/10.1029/2017wr021176

Pheulpin, L., Bacchi, V., & Bertrand, N. (2020). *Comparison between two hydraulic models (1D and 2D) of the Garonne River: Application to uncertainty propagations and sensitivity analyses of levee breach parameters*. Springer Water.

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

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. https://doi.org/10.1016/j.envsoft.2016.02.008

Pianosi, F., & Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, *67*, 1–11. https://doi.org/10.1016/j.envsoft.2015.01.004

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Kevin Tucker, P. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, *41*, 1–28. https://doi.org/10.1016/j.paerosci.2005.02.001

Rakovec, O., Hill, M. C., Clark, M. P., Weerts, A. H., Teuling, A. J., & Uijlenhoet, R. (2014). Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models. *Water Resources Research*, *50*(1), 409–426. https://doi.org/10.1002/2013wr014063

Reed, P. M., Hadjimichael, A., Malek, K., Karimi, T., Vernon, C. R., Srikrishnan, V., et al. (2022). *Addressing uncertainty in multisector dynamics research*. Zenodo.

Rosenblueth, E. (1975). Point estimates for probability moments. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(10), 3812–3814. https://doi.org/10.1073/pnas.72.10.3812

Rossi, R. J. (2018). *Mathematical statistics: An introduction to likelihood based inference*. John Wiley & Sons.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., et al. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, *114*, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, *181*(2), 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Saltelli, A., Annoni, P., & D'Hombres, B. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, *25*(6), 1508–1517. https://doi.org/10.1016/j.sbspro.2010.05.133

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). *Global sensitivity analysis: The primer*. Wiley.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., & others (2004). *Sensitivity analysis in practice: A guide to assessing scientific models*. Wiley Online Library.

Sattar, A. M. A. (2014). Gene expression models for prediction of dam breach parameters. *Journal of Hydroinformatics*, *16*(3), 550–571. https://doi.org/10.2166/hydro.2013.084

Schmitz, V., Erpicum, S., El Kadi Abderrezzak, K., Rifai, I., Archambeau, P., Pirotton, M., & Dewals, B. (2021). Overtopping-induced failure of non–cohesive homogeneous fluvial dikes: Effect of dike geometry on breach discharge and widening. *Water Resources Research*, *57*, e2021WR029660. https://doi.org/10.1029/2021WR029660

Shen, G., Sheng, J., Xiang, Y., Zhong, Q., & Yang, D. (2020). Numerical modeling of overtopping-induced breach of landslide dams. *Natural Hazards Review*, *21*(2). https://doi.org/10.1061/(asce)nh.1527-6996.0000362

Shimizu, Y., Nelson, J., Arnez Ferrel, K., Asahi, K., Giri, S., Inoue, T., et al. (2020). Advances in computational morphodynamics using the International River Interface Cooperative (iRIC) software. *Earth Surface Processes and Landforms*, *45*(1), 11–37. https://doi.org/10.1002/esp.4653

Sobol, I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, *55*(1–3), 271–280. https://doi.org/10.1016/s0378-4754(00)00270-6

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiments*, *4*, 407–414.

Spear, R. C., & Hornberger, G. M. (1980). Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, *14*(1), 43–49. https://doi.org/10.1016/0043-1354(80)90040-8

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, *93*(7), 964–979. https://doi.org/10.1016/j.ress.2007.04.002

Tsai, C. W., Yeh, J., & Huang, C. (2019). Development of probabilistic inundation mapping for dam failure induced floods. *Stochastic Environmental Research and Risk Assessment*, *33*(1), 91–110. https://doi.org/10.1007/s00477-018-1636-8

Vorogushyn, S., Apel, H., & Merz, B. (2011). The impact of the uncertainty of dike breach development time on flood hazard. *Physics and Chemistry of the Earth*, *36*(7–8), 319–323. https://doi.org/10.1016/j.pce.2011.01.005

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

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & von Krauss Krayer, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17. https://doi.org/10.1076/iaij.4.1.5.16466

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., Reynolds, J. M., Hassan, M. A. A. M., & Lowe, A. (2015). Numerical modelling of glacial lake outburst floods using physically based dam-breach models. *Earth Surface Dynamics*, *3*(1), 171–199. https://doi.org/10.5194/esurf-3-171-2015

Wu, W. (2013). Simplified physically based model of Earthen embankment breaching. *Journal of Hydraulic Engineering*, *139*(8), 837–851. https://doi.org/10.1061/(asce)hy.1943-7900.0000741

Wu, W. (2016). *Introduction to DL Breach–A simplified physically based dam/levee breach model*. Clarkson University.

Wu, W., & Wang, S. S. (2006). Formulas for sediment porosity and settling velocity. *Journal of Hydraulic Engineering*, *132*(8), 858–862. https://doi.org/10.1061/(asce)0733-9429(2006)132:8(858)

Wu, W., Wang, S. S., & Jia, Y. (2000). Nonuniform sediment transport in alluvial rivers. *Journal of Hydraulic Research*, *38*(6), 427–434. https://doi.org/10.1080/00221680009498296

Zhang, R. J. (1961). *River dynamics*. Industry Press.

Zhong, Q., Chen, S., & Deng, Z. (2017). Numerical model for homogeneous cohesive dam breaching due to overtopping failure. *Journal of Mountain Science*, *14*(3), 571–580. https://doi.org/10.1007/s11629-016-3907-5

Zhong, Q., Wang, L., Chen, S., Chen, Z., Shan, Y., Zhang, Q., et al. (2021). Breaches of embankment and landslide dams - State of the art review. *Earth-Science Reviews*, *216*, 103597. https://doi.org/10.1016/j.earscirev.2021.103597