

Probabilistic models for the erosion rate in embankments and reliability analysis of earth dams

Marco Andreini^{1,*}, Paolo Gardoni², Stefano Pagliara¹, Mauro Sassu³

¹*Department of Energy, Systems, Territory and Constructions Engineering (DESTEC), University of Pisa, Largo Lucio Lazzarino, 1 – 56126 Pisa, Italy.*

²*Department of Civil and Environmental Engineering and Director, MAE Center, University of Illinois at Urbana-Champaign, Newmark Civil Engineering Laboratory, 205 North Mathews Ave., Urbana, IL 61801–2352, USA.*

³*Department of Civil, Environmental Engineering and Architecture (DICAAR), University of Cagliari, Via Marengo, 2 – 09123 Cagliari, Italy.*

Abstract

Probabilistic models for the concentrated leak erosion of earthen water retaining structures are presented. The models predict the values of the critical shear stress, the coefficient of erosion and the pipe radius enlargement, starting from other measurable soil properties and the geometrical dimensions of the embankment. The models account for both the non-cohesive and cohesive contributions to the erosion behavior. A Bayesian approach is used for the treatment of the unknown parameters. An importance sampling simulation is adopted to calibrate the models and estimate the posterior distribution of the unknown model parameters using laboratory and in situ experimental data. The new proposed probabilistic model for the pipe radius is then used to develop fragility curves that capture the pipe enlargement as a function of time for a given earth dam.

Keywords

Piping, concentrated leak erosion, earth dams, levees, water retaining structures, Bayesian approach, probabilistic models, fragility.

* Corresponding author. Tel.: +39 050 221 7700; e-mail address: m.andreini@ing.unipi.it (M. Andreini).

1. Introduction

The uncertainties tied to the internal erosion failure of the embankments are still the subject of current study and research activities, since earth dams and levees are widespread throughout the world and have been used since ancient times [1]. Still today, internal erosion phenomena are one of the most common causes of breach mechanism in earth dams [2]; they are ruled by the propensity of the embankment soil to be eroded.

The erodibility of a soil is commonly evaluated considering two main parameters [3]: the critical shear stress τ_c and the coefficient of erosion C_e . The critical shear stress is defined as the tangential stress threshold at which we have the initiation of the erosion mechanism. The coefficient of erosion is a constant that expresses the intensity of the erosion rate.

Jet Erosion Test (JET) [4–6] and Hole Erosion Test (HET) [7–10] are among the most recent experimental test procedures for determination of critical shear stress and the coefficient of erosion, which has proved to be a successful test for a number of soils. However, these tests can be time-consuming and difficult to perform compared to most common geotechnical investigations. Wan and Fell [8] presented regression models to predict C_e based on physical and geotechnical variables for soils showing non-cohesive or cohesive behavior. Andreini et al. [11] presented probabilistic models for the two soil erosion parameters accounting for both non-cohesive and cohesive contributions to the erosion behavior, where the unknown parameters were calibrated using HET laboratory data [7,12,13]. Andreini et al. [11] proposed also fragility curves indicating the conditional probability of erosion initiation, without addressing the evolution of the pipe enlargement in an embankment. Such an evolution requires the definition of a proper analytical formulation of the hole radius increase versus time [10]. For that purpose, a model for pipe flow with erosion analysis was developed in the last decades [15,17-19] on the basis of the equations of two-phase flow with erosion.

This paper revisits the probabilistic models for the critical shear stress τ_c and the coefficient of erosion C_e proposed by Andreini et al. [11]. Specifically, a Bayesian updating of the unknown model parameters is performed through additional in situ JET results. The posterior statistics of the parameters are determined by a Monte Carlo simulation [17], implementing the procedure presented by Gardoni et al. [18–20] in the parallel computation algorithm proposed by Andreini et al. [11] to reduce the calculation time. The updated models of the erosion parameters are then used in the formulation of nested probabilistic models of the pipe radius evolution in an embankment, as shown in the scheme in Figure 1. The unknown parameters of such a model are firstly calibrated using the in situ data from an experimental test, carried out on a scaled physical model of a dam [21]. The parameters for all models are subsequently updated/calibrated on the basis of all the available experimental data.

The safety of earth dams toward internal erosion phenomena is commonly estimated by inspecting the factors that can lead to the initiation of an erosion process, checking the effectiveness of the eventual filters, and assessing whether some detection and repair interventions could be executed. The probability of failure is commonly assessed by means of event tree methods [22], where conditional probability of each event can be estimated by mathematical models, experiments, or expert judgment. The proposed probabilistic models are used to determine the conditional probability (or fragility) of not exceeding a certain value of the pipe radius in a specified time from the initiation of the concentrated leak erosion mechanism. This is needed to calculate the probability of no detection, intervention and repair.

First, this paper presents the general formulation of the probabilistic models and the procedure used to for the Bayesian estimation of the unknown parameters. Then, the results of the model calibration and the posterior statistics of the unknown model parameters are presented and discussed. Finally, the paper presents the estimated conditional probability of no detection, intervention and repair for a typical earth dam.

2. Development of probabilistic models

2.1 General formulation

The probabilistic models presented herein are valid under the assumption that their standard deviation σ is independent of the physical variables (homoskedasticity assumption) and the model error follows the normal distribution (normality assumption). Under these assumptions, a generic univariate model C of a quantity of interest (or a transformation of the quantity of interest) can be expressed in the following general form [18–20]:

$$C(\mathbf{x}, \Theta) = c(\mathbf{x}, \theta, \lambda) + \sigma\varepsilon, \quad (1)$$

where \mathbf{x} represents the physical variables, $\Theta = (\theta, \lambda, \sigma)$ denotes the set of unknown parameters (where θ is the set of model parameters used to best fit the data and λ is the parameter that controls the family of transformations used to satisfy the homoskedasticity and normality assumptions.), and $\sigma\varepsilon$ is the model error (where ε is a normal random variable with zero mean and unit variance and σ is the standard deviation of the model error). In this paper, we use the family of dual power transformations proposed by Yang [23] as follows

$$f_\lambda(\cdot) = \begin{cases} \frac{(\cdot)^\lambda - (\cdot)^{-\lambda}}{2\lambda} & \text{if } \lambda \neq 0, \\ \log(\cdot) & \text{if } \lambda = 0 \end{cases}, \quad (2)$$

which inverse is

$$f_\lambda^{-1}(\cdot) = \begin{cases} \left(\lambda(\cdot) + \sqrt{1 + \lambda^2(\cdot)^2} \right)^{\frac{1}{\lambda}} & \text{if } \lambda \neq 0, \\ e^{(\cdot)} & \text{if } \lambda = 0 \end{cases}. \quad (3)$$

2.2 Bayesian updating

In the Bayesian approach [17,24], the unknown model parameters are estimated through the following updating rule

$$p(\Theta|\mathbf{y}) = \kappa L(\Theta|\mathbf{y})p(\Theta), \quad (4)$$

where $p(\Theta | \mathbf{y})$ is the posterior distribution that reflects the updated state of information about the vector of unknown parameters Θ , $L(\Theta | \mathbf{y})$ is the likelihood function that captures the information from the data vector of n observations $\mathbf{y} = (y_1, y_2, \dots, y_n)$, $p(\Theta)$ is the prior distribution that represents the information available before collecting the data, and $\kappa = \left[\int L(\Theta | \mathbf{y}) p(\Theta) d\Theta \right]^{-1}$ is a normalizing factor. In addition, the updating rule in Eq. (4) can be used to continuously update the current knowledge every time new data become available. For example, if an initial sample of observations \mathbf{y}_1 is originally available, a first application of the updating rule gives

$$p(\Theta | \mathbf{y}_1) = \kappa_1 L(\Theta | \mathbf{y}_1) p(\Theta), \quad (5)$$

where κ_1 is the normalizing factor for the first updating. If a second sample of observations \mathbf{y}_2 , distributed independently from the first one, becomes available, $p(\Theta | \mathbf{y}_1)$ can be updated to account for the new information obtaining

$$\begin{aligned} p(\Theta | \mathbf{y}_1, \mathbf{y}_2) &= \kappa_2 L(\Theta | \mathbf{y}_2) p(\Theta | \mathbf{y}_1) \\ &= \kappa_2 L(\Theta | \mathbf{y}_1) L(\Theta | \mathbf{y}_2) p(\Theta), \end{aligned} \quad (6)$$

in which κ_2 is the normalizing factor for the second updating.

Eq. (6) is of the same form as Eq. (5) (i.e., the posterior distribution of Θ given \mathbf{y}_1 works as the prior distribution for the second sample). Such updating process can be carried out any number of times, obtaining the following general form for n_q sets of samples

$$\begin{aligned} p(\Theta | \mathbf{y}_1, \dots, \mathbf{y}_{n_q}) &= \kappa_{n_q} \left[\prod_{q=1}^{n_q} L(\Theta | \mathbf{y}_q) \right] p(\Theta), \\ &= \kappa_{n_q} L(\Theta | \mathbf{y}_{n_q}) p(\Theta | \mathbf{y}_1, \dots, \mathbf{y}_{n_q-1}) \end{aligned} \quad (7)$$

where $\kappa_{n_q} = \left\{ \int \left[\prod_{q=1}^{n_q} L(\Theta | y_q) \right] p(\Theta) d\Theta \right\}^{-1}$ is the normalizing factor associated to the n_q -th updating.

Computation of these posterior statistics can be challenging, especially when large numbers of parameters are to be determined. Andreini et al. [11] proposed an algorithm using importance sampling, which may be used in parallel computing to reduce the analysis time. The analyses performed and presented in this paper are executed using this algorithm and it was detected that the time to satisfy the termination criterion was proportionally decreasing with the number of processing elements used.

3. Probabilistic models adopted for the critical shear stress and the coefficient of erosion

The adaptation of the probabilistic models proposed by Andreini et al. [11] to the case of multiple Bayesian updating is shown herein. The selected deterministic models for cohesive and non-cohesive soils of the erosion parameters and the behavior combination formula are described. The probabilistic model formulation and the model corrective terms are presented. The description of the likelihood function and the adopted prior distribution for the unknown parameters are finally given.

3.1. Cohesive and Non-cohesive behavior combination formula

Referring to van Ledden [25] and van Rijn [26] who investigated the erosion phenomena in coastal engineering, embankment soils can be treated as a mixture of sand/gravels and mud, determinable as those components having an average diameter respectively higher and lower than $63 \mu\text{m}$.

Rearranging the combination formula proposed by van Ledden [25], Andreini et al. [11] found that the generic erosion parameter ξ (i.e. $\xi = \tau_c$ for the critical shear stress and $\xi = C_e$

for the coefficient of erosion) for sand and mud mixtures can be expressed as a combination of those related to each component ξ_k (i.e. $k = s$ for pure sand and $k = m$ for pure mud). Such a combination can assume the form

$$f_c(\xi) = \sum_{k=s,m} w_{\xi,k} f_c(\xi_k) + \Delta_\xi, \quad (8)$$

and therefore

$$\xi = f_c^{-1} \left[\sum_{k=s,m} w_{\xi,k} f_c(\xi_k) + \Delta_\xi \right], \quad (9)$$

where $f_c(\cdot)$ is a transformation of the generic erosion parameter, $f_c^{-1}(\cdot)$ is its inverse, $w_{\xi,k}$ are the weight functions, and Δ_ξ is a corrective term.

Following Gardoni et al. [18,19], Andreini et al. [11] revisited the combination formula of Van Ledden [25] in a global expression containing two nested probabilistic models having a global set of unknown parameters $\Theta_\xi = (\theta_{\xi,k}, \theta_{\xi,w}, \sigma_\xi)$, in which $\theta_{\xi,k}$ are those of the pure sand and pure mud models, $\theta_{\xi,w}$ are those of the weight functions and σ_ξ is the standard deviation of the model error. The form of such a general expression is

$$C_\xi(\mathbf{x}_k, \mathbf{x}_w, \Theta_\xi, \lambda) = f_\lambda \left\{ f_c^{-1} \left[\mathbf{W}_\xi \cdot \mathbf{F}_c(\mathbf{F}_\lambda^{-1} \mathbf{C}_\xi) + \Delta_\xi \right] \right\} + \sigma_\xi \mathcal{E}_\xi, \quad (10)$$

where \mathbf{C}_ξ is the vector of the probabilistic models $C_{\xi,k} = C_{\xi,k}(\mathbf{x}_k, \theta_{\xi,k}, \lambda)$ related to pure sand and pure mud behaviour $\mathbf{C}_\xi = \mathbf{C}_\xi(\mathbf{x}_k, \theta_{\xi,k}, \lambda) = [C_{\xi,s}(\mathbf{x}_s, \theta_{\xi,s}, \lambda), C_{\xi,m}(\mathbf{x}_m, \theta_{\xi,m}, \lambda)]$, while \mathbf{W}_ξ is the vector of the weight functions $w_{\xi,k} = w_{\xi,k}(\mathbf{x}_w, \theta_{\xi,w})$, that is $\mathbf{W}_\xi = \mathbf{W}_\xi(\mathbf{x}_w, \theta_{\xi,w}) = [w_{\xi,s}(\mathbf{x}_w, \theta_{\xi,w}), w_{\xi,m}(\mathbf{x}_w, \theta_{\xi,w})]$, and $\Delta_\xi = \Delta_\xi(\mathbf{x}_w, \theta_{\xi,w})$ is a corrective function. The two terms $\mathbf{F}_c(\cdot)$ and $\mathbf{F}_\lambda(\cdot)$ in Eq. (10) are the vector-valued functions that respectively apply the scalar functions $f_c(\cdot)$ and $f_\lambda(\cdot)$ to each vector component. In the

mentioned expressions, \mathbf{x}_k are the variables for the sand and the mud components, while \mathbf{x}_w are those for the weight functions.

The model of both sand/gravels and mud components can be expressed as the sum of two contributions

$$C_k(\mathbf{x}_k, \boldsymbol{\theta}_{\xi,k}, \lambda) = \hat{c}_{\xi,k}(\mathbf{x}_k, \lambda) + \gamma_{\xi,k}(\mathbf{x}_k, \boldsymbol{\theta}_{\xi,k}, \lambda), \quad (11)$$

where, the deterministic model terms for both sand and mud component

$\hat{c}_{\xi,k}(\mathbf{x}_k, \lambda) = f_\lambda(\hat{\xi}_k^*(\mathbf{x}_k))$, are obtained by respectively applying the Yang transformation in Eq.

(2) to the dimensionless expressions $\hat{\xi}_k^*(\mathbf{x}_k)$ of sand and mud components of the critical shear

stress $\hat{\tau}_{c,k}^*(\mathbf{x}_k)$ and the coefficient of erosion $\hat{C}_{e,k}^*(\mathbf{x}_k)$. In Eq. (11), $\gamma_{\xi,k}(\mathbf{x}_k, \boldsymbol{\theta}_{\xi,k}, \lambda)$ are the

correction terms for the deterministic models of sand and mud components.

The weight functions and the corrective terms are defined in Table 1 as piecewise functions of

the mud content percentage P_m , whose critical value $P_{m,cr}$ represents the threshold above

which the non-cohesive behavior of the mixture becomes cohesive. In this case, $\mathbf{x}_w = P_m$ and

the set of unknown parameters for the weights $\boldsymbol{\theta}_{\xi,w}$ concerns two terms: $\theta_{\xi,w,1} = P_{m,cr}$ and an

empirical exponent $\theta_{\xi,w,2}$.

3.2 Deterministic models

3.2.1 Non-cohesive soils/sand component

For the deterministic model of the critical shear stress we use the Meyer-Peter & Müller empirical formula [27]

$$\hat{\tau}_{c,s}(\mathbf{x}_s) = 0.047g(\rho_d - \rho_w)d_{s,50} \left(\frac{k_r}{k_s} \right)^{\frac{3}{2}}, \quad (12)$$

where $k_r = 26d_{s,90}^{-1/6}$, $k_s = 21d_{s,50}^{-1/6}$, in which $d_{s,50}$ and $d_{s,90}$ are respectively the diameters corresponding to the 50% and 90% of the only sand and gravel component (grains with a diameter higher than $75 \mu m$).

Dividing the Eq. (12) by the quantity $g(\rho_d - \rho_w)d_{c/s}$, where $d_{c/s} = 5 \mu m$ is the diameter which represents the transition between silt and clay in the fine part of the soils considered, we obtain the following expression of the dimensionless deterministic model of the critical shear stress

$$\hat{\tau}_{c,s}^*(\mathbf{x}_s) = \frac{0.065}{d_{c/s}} \frac{\sqrt{d_{s,50}^5}}{\sqrt{d_{s,90}}}. \quad (13)$$

where the (*) identifies the dimensionless quantity. For the coefficient of erosion, we refer to van Rijn [26], which indicates the following formulation for non-cohesive soils

$$\hat{C}_{e,s}(\mathbf{x}_s) = 0.018 \frac{\rho_d}{\tau_{c,s} D_*^{0.9}} \sqrt[4]{\frac{g d_{s,50} (\rho_d - \rho_w)}{\rho_w}}, \quad (14)$$

where D_* represents the dimensionless grain size

$$D_* = d_{s,50} \sqrt[3]{\frac{g(\rho_d - \rho_w)}{\rho_w \nu^2}}, \quad (15)$$

ν the kinematic viscosity, and $\tau_{c,s}$ is the critical shear stress for pure sand for which we use

Eq. (13) setting $\tau_{c,s} = \hat{\tau}_{c,s}(\mathbf{x}_s)$. A dimensionless form of the Eq. (14) is obtained by multiplying

it by the quantity $\sqrt{g d_{c/s}}$, so that we have

$$\hat{C}_{e,s}^*(\mathbf{x}_s) = \hat{C}_{e,s}(\mathbf{x}_s) \sqrt{g d_{c/s}}. \quad (16)$$

3.2.2 Cohesive soils/mud component

For the deterministic model of the critical shear stress, we consider the formula from Smerdon and Beasley [28], which associates the critical shear stress to the plasticity index (PI) in the following way:

$$\hat{\tau}_{c,m}(\mathbf{x}_m) = 0.167 Pa (\text{PI})^{0.84}. \quad (17)$$

The plasticity index gives the range of the water content where soils exhibit plastic behavior. The dimensionless form of Eq. (17) is obtained as presented above for the Mayer-Peter and Müller formula [27] as follows

$$\hat{\tau}_{c,m}^*(\mathbf{x}_m) = \frac{0.167 Pa (\text{PI})^{0.84}}{g(\rho_d - \rho_w) d_{c/s}}. \quad (18)$$

For the coefficient of erosion we use the formula proposed by Hanson and Simon [29]

$$\hat{C}_{e,m}(\mathbf{x}_m) = \rho_d \frac{2 \times 10^{-7}}{\sqrt{\tau_{c,m}(\mathbf{x}_m) / Pa}}, \quad (19)$$

where $\tau_{c,m}$ is the critical shear stress for pure mud, for which we use the Eq. (18) putting

$$\tau_{c,m} = \hat{\tau}_{c,m}(\mathbf{x}_m).$$

A dimensionless form, as well as seen for the sand component, is obtained as follows

$$\hat{C}_{e,m}^*(\mathbf{x}_m) = \hat{C}_{e,m}(\mathbf{x}_m) \sqrt{g d_{c/s}}. \quad (20)$$

3.3 Models' corrections

Following Gardoni et al. [18–20] and Andreini et al. [11], the initial correction terms

$\gamma_{\xi,k}(\mathbf{x}_k, \boldsymbol{\theta}_{\xi,k}, \lambda)$ are written as the following quadratic combination:

$$\gamma_{\xi,k}(\mathbf{x}_k, \boldsymbol{\theta}_{\xi,k}, \lambda) = [\mathbf{h}_k(\mathbf{x}_k, \lambda)]^T \boldsymbol{\theta}_{\xi,k} [\mathbf{h}_k(\mathbf{x}_k, \lambda)] = \sum_{i=1}^{n_k} \sum_{j \geq i}^{n_k} \theta_{\xi,k,ij} h_{k,i}(\mathbf{x}_k, \lambda) h_{k,j}(\mathbf{x}_k, \lambda), \quad (21)$$

in which the sets of the unknown parameters $\boldsymbol{\theta}_{\xi,k}$ are represented using an upper triangular matrix form, and \mathbf{h}_k is a vector of n_k explanatory functions $h_{k,i}(\mathbf{x}_k)$ obtained by applying the

Yang transformation expressed in Eq. (2) to a set of basis functions $\eta_{k,i}(\mathbf{x}_k)$, that is

$$h_{k,i}(\mathbf{x}_k) = f_\lambda[\eta_{k,i}(\mathbf{x}_k)].$$

3.4 Likelihood function and definition of prior distribution

Considering the generic q -th set of data $(\mathbf{y}_{\xi,q}, \mathbf{x}_{k,q}, \mathbf{x}_{w,q})$, the expression of the l -th residual $r_{\xi,q,l}$ related to the transformed reading $C_{\xi,q,l} = f_\lambda(y_{\xi,q,l})$, for given $(\mathbf{x}_{k,q,l}, \mathbf{x}_{w,q,l})$, assumes the following form [11]:

$$r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda) = C_{\xi,q,l} - f_\lambda \left\{ f_c^{-1} \left[\mathbf{W}_\xi \cdot \mathbf{F}_c \left(\mathbf{F}_\lambda^{-1} \mathbf{C}_\xi(\mathbf{x}_{k,q,l}, \boldsymbol{\theta}_{\xi,k}, \lambda) \right) + \Delta_\xi(\mathbf{x}_{w,q,l}, \boldsymbol{\theta}_{\xi,w}) \right] \right\}, \quad (22)$$

that is $r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda) = \sigma_\xi \varepsilon_{\xi,q,l}$. Under the assumption of statistically independent observations, we obtain the general form of the likelihood function [18,19]

$$L(\boldsymbol{\Theta}_\xi, \lambda | \mathbf{y}_{\xi,q}) \propto \prod_{\text{Equality Data } l} P \left[\sigma_\xi \varepsilon_{\xi,q,l} = r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda) \right] \times \prod_{\text{Upper Bound Data } l} P \left[\sigma_\xi \varepsilon_{\xi,q,l} < r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda) \right]. \quad (23)$$

Since ε_ξ has a standard normal distribution, we can write

$$L(\boldsymbol{\Theta}_\xi, \lambda | \mathbf{y}_{\xi,q}) \propto \prod_{\text{Equality Data } l} \left\{ \frac{1}{\sigma_\xi} \varphi \left[\frac{r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda)}{\sigma_\xi} \right] \right\} \times \prod_{\text{Upper Bound Data } l} \left\{ \Phi \left[\frac{r_{\xi,q,l}(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}, \lambda)}{\sigma_\xi} \right] \right\}, \quad (24)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ respectively represent the standard normal probability density function and its cumulative distribution function.

Considering the availability of n_q different sets of data \mathbf{y}_{ξ,n_q} , the Bayesian updating rule expressed in the Eq. (7) can be written as

$$p(\boldsymbol{\Theta} | \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi,n_q}) = \kappa_{n_q} \left[\prod_{q=1}^{n_q} L(\boldsymbol{\Theta}_\xi | \mathbf{y}_{\xi,q}) \right] p(\boldsymbol{\Theta}_\xi) = \kappa_{n_q} L(\boldsymbol{\Theta}_\xi | \mathbf{y}_{n_q}) p(\boldsymbol{\Theta}_\xi | \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi,n_q-1}), \quad (25)$$

where $\kappa_{\xi, n_q} = \left\{ \int \left[\prod_{q=1}^{n_q} L(\Theta_{\xi} | \mathbf{y}_{\xi, q}) \right] p(\Theta_{\xi}) d\Theta_{\xi} \right\}^{-1}$ is the normalizing factor associated to the n_q -th updating.

4. Formulation of the probabilistic model for pipe enlargement

4.1 Deterministic model based on a simplified mechanical approach

A model for the pipe enlargement was developed on the basis of the integrated reduced Navier-Stokes/Prandtl equations with erosion by Bonelli and Brivois [14]. The enlargement of the hole resulting from the erosion has been described by

$$\frac{dR}{dt} = \begin{cases} \frac{C_e \tau_c}{\rho_d} \left(\frac{\tau_b}{\tau_c} - 1 \right)^p & \text{if } \tau_b > \tau_c, \\ 0 & \text{otherwise} \end{cases}, \quad (26)$$

where R is the radius of the pipe increasing with the time t , ρ_d is the dry density of the soil, τ_b is the shear stress exerted by the flow on the soil and p is a positive empirical exponent equal or lower than 1.

We refer to the case shown in Figure 2, where the transversal section of an earth dam is in the condition of absence of a filter or inability to stop the continuation of the erosion process. Such a section is geometrically defined by the total height H_{dam} , the widths of the crest b_{min} and the base b_{max} , and the gradients of the upstream and downstream slopes c_u and c_d .

As in Bonelli et al. [10], we consider the simplified case of a straight circular pipe of length $b(z, h_w) = (c_u + c_d)(H_{dam} - h_w + z) + b_{min}$ and current radius $R(t)$, occurred at a certain depth z from the free surface, considering a water level h_w from the dam base.

In this content, Bonelli et al. [10] proposed a formulation to obtain a closed form expression of $R(t)$, by means of the simplifying assumptions that constant pressure drop and both the head

and the momentum loss factors, α and β , are stationary during the pipe enlargement. Accordingly, it is possible to show that the initiation of the erosion occurs when the initial radius of the pipe is higher than a critical value R_c depending on the critical shear stress τ_c by means of the relation

$$R_0 > R_c = \frac{2\tau_c b(z, h_w)}{\rho_w g z}. \quad (27)$$

Integrating the Navier/Stokes equations with erosion concerning Eq. (26), Bonelli et al. [10] showed that the radius increase factor $\delta_r = R / R_0$ can be given by

$$\delta_r = \hat{\delta}_r(t) = \begin{cases} \frac{\tau_c}{\tau_{b0}} + \left(1 - \frac{\tau_c}{\tau_{b0}}\right) e^{\frac{t}{t_{er}}} & \text{if } p = 1 \\ \frac{\tau_c}{\tau_{b0}} + \left[\left(1 - \frac{\tau_c}{\tau_{b0}}\right)^{1-p} + (1-p) \frac{t}{t_{er}} \left(1 - \frac{\tau_c}{\tau_{b0}}\right)^{1-p} \right]^{\frac{1}{1-p}} & \text{if } 0 < p < 1 \end{cases}, \quad (28)$$

where τ_{b0} and t_{er} are respectively the initial shear stress applied on the lateral surface of the pipe and the characteristic time, given by the following expressions

$$\tau_{b0} = \frac{R_0 \alpha \rho_w g z}{2b(z, h_w)}, \quad (29)$$

$$t_{er} = \frac{2\rho_d b(z, h_w)}{\beta C_e \alpha \rho_w g z}. \quad (30)$$

Since considering that both factors α and β are stationary is a strong assumption, they are treated together with the coefficient p as unknown parameters in the probabilistic model described in the following.

4.3 Probabilistic model formulation

The probabilistic model C_r of the radius increase factor δ_r can be expressed through the general form

$$C_r(\Theta_\xi, \Theta_{\delta_r}, \lambda) = f_\lambda \left[\delta_r(\mathbf{x}_{\delta_r}, \Theta_\xi, \Theta_{\delta_r}) \right] + \theta_r + \sigma_{\delta_r} \varepsilon_{\delta_r}, \quad (31)$$

where $\Theta_{\delta_r} = (\theta_{\delta_r}, \theta_r, \sigma_{\delta_r})$ is the additional set of parameters of such a model, $\mathbf{x}_{\delta_r} = (\mathbf{x}_k, \mathbf{x}_w, \mathbf{x}_r)$ represents the set of variables in which $\mathbf{x}_r = (z, R_0, h_w, t)$, and $\theta_{\delta_r} = (\theta_\alpha, \theta_\beta, \theta_p)$ are the parameters which account, as mentioned at the end of the previous section, for the uncertainties of the factors α, β and the exponent p . In addition, θ_r is a parameter used to correct for the bias of the approximated formulation of Bonelli et al. [10] in Eq. (28), ε_{δ_r} is a normal random variable with zero mean and unit variance and σ_{δ_r} is the standard deviation of the model error $\sigma_{\delta_r} \varepsilon_{\delta_r}$.

With the formulation in Eq. (31), the radius increase factor expression $\delta_r = \delta_r(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \theta_{\delta_r})$ can be obtained by plugging the values of the following functions in Eq. (28):

$$\tau_c = \tau_c(\mathbf{x}_k, \mathbf{x}_w, \Theta_{\tau_c}, \lambda) = g(\rho_d - \rho_w) d_{c/s} f_\lambda^{-1} \left[C_{\tau_c}(\mathbf{x}_k, \mathbf{x}_w, \Theta_{\tau_c}, \lambda) \right], \quad (32)$$

$$C_e = C_e(\mathbf{x}_k, \mathbf{x}_w, \Theta_{C_e}, \lambda) = f_\lambda^{-1} \left[C_{C_e}(\mathbf{x}_k, \mathbf{x}_w, \Theta_{C_e}, \lambda) \right] / \sqrt{g d_{c/s}}, \quad (33)$$

$$t_{er} = t_{er}(z, \theta_\alpha, \theta_\beta, \Theta_{C_e}) = \frac{2\rho_d b(z)}{C_e(\mathbf{x}_k, \mathbf{x}_w, \Theta_{C_e}, \lambda) \theta_\alpha \theta_\beta \rho_w g z}, \quad (34)$$

$$\tau_{b0} = \tau_{b0}(z, h_w, R_0, \theta_\alpha) = \frac{R_0 \theta_\alpha \rho_w g z}{2b(z, h_w)}. \quad (35)$$

Considering a generic set of data $(\mathbf{y}_{\delta_r}, \mathbf{x}_{\delta_r})$, the expression of the l -th residual $r_{\delta_r, l}$ related to the transformed reading $C_{r, l}(y_{\delta_r, l}, \lambda)$ assumes the following form:

$$r_{\delta_r, l}(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \lambda) = C_{r, l}(y_{\delta_r, l}, \lambda) - f_\lambda(\delta_r(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \theta_{\delta_r})) - \theta_r = \sigma_{\delta_r} \varepsilon_{\delta_r}. \quad (36)$$

Under the assumption of statistically independent observations, we obtain the general form of the likelihood function

$$L(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \sigma_{\delta_r}, \lambda) \propto \prod_{\text{Equality Data } l} P \left[\sigma_{\delta_r} \varepsilon_{\delta_r} = r_{\delta_r, l}(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \lambda) \right] \quad (37)$$

and, since ε_{δ_r} responds to a standard normal distribution, we can write

$$L(\Theta_{\xi}, \Theta_{\delta_r}, \lambda) \propto \prod_{\text{Equality Data } l} \left\{ \frac{1}{\sigma_{\delta_r}} \varphi \left[\frac{r_{\delta_r, l}(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \lambda)}{\sigma_{\delta_r}} \right] \right\}. \quad (38)$$

We have two options to determine the posterior statistics of the unknown parameters Θ_{δ_r} :

1. Application of the probabilistic model of the radius increase factor, considering the updated posterior means $\hat{\Theta}_{\tau_c}$ and $\hat{\Theta}_{C_c}$ of the erosion parameters' models as deterministic.
2. Estimation of the descriptors of the joint posterior distribution of the total sets of parameters $(\Theta_{\xi}, \Theta_{\delta_r}, \lambda)$, by means of models' updating using a global likelihood function.

Following the second option, the definition of a joint prior distribution of the parameters $(\Theta_{\xi}, \Theta_{\delta_r})$ is needed. This is discussed in the section 5.3.

5. Calibration of the probabilistic models

5.1 Data sets used to construct the probabilistic models

Andreini et al. [11] presented a first calibration of the formulated probabilistic models for the erosion parameters. They used the data set $\mathbf{y}_{\xi, 1}$ concerning the HETs performed by Wan and Fell [7] on 144 samples of 13 different types of soils. In addition, 11 tests carried out by Benahmed and Bonelli [12] and Benahmed et al. [13] on 100% clay samples were considered. All these tests offered a definite value of the critical shear stress so that they are used as equality data for the calibration of the parameters related to C_{τ_c} . Wan and Fell [7] reported 27 cases where the erosion rate slowed down gradually or negligible erosion was detected. These cases were used in [11] as upper bound data in calibration of the coefficient of erosion model.

The calibration of pipe enlargement models requires data recorded on physical models of embankments subjected to concentrated leak erosion. In this paper, we consider the first of the three experiments carried out by the Hydraulic Engineering Research Unit of the U.S. Bureau of Reclamation [21] on large-scale models of homogeneous embankment dams. Concentrated leak erosion was produced by embedding a pipe with radius R_0 of 20 mm in the embankment and pulling it out through the downstream side to begin each test. After that, the pipe enlargement in terms of downstream width variation versus time was monitored and recorded up to the failure occurred after 13 minutes for loss of ability to support a roof. For the sake of simplicity, we consider in the following that the pipe maintained a cylindrical shape for the entire duration of the experiment and, according with the images of the collapse, the collapse occurred corresponding to a final value of the radius equal to $H_{dam}/3$. Although it represent a strong assumption [10], such a simplification allows us to construct the data set \mathbf{y}_{δ_r} by easily scaling the function shape of the pipe width readings reported by Hanson et al. [21] from an initial value $\delta_r(0) = 1$ to a final one of $\delta_r(13 \text{ min}) = H_{dam} / (3R_0) = 20.33$. Moreover, since two soil samples from such a physical model were subjected to JET [30-32], we use the related results as new data $\mathbf{y}_{\xi,2}$ to update the parameters of the probabilistic models of τ_c and C_e , before the calibration of those related to C_{δ_r} model.

5.2 Erosion parameters' models

Defining $\Sigma_{\xi} = \sigma_{\xi}^2$ as the variance of the erosion parameters' model, it generally assumed its independence from $\boldsymbol{\theta}_{\xi,k}$ and $\boldsymbol{\theta}_{\xi,w}$, so that the prior distribution of the parameters Θ_{ξ} is, therefore, $p(\Theta_{\xi}) \approx p(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w})p(\Sigma_{\xi})$.

Considering a noninformative prior distribution for the parameters $\boldsymbol{\theta}_{\xi,k}$, and referring to Box and Tiao [24], it is possible to show that it is locally uniform, thus we obtain that

$p(\boldsymbol{\theta}_{\xi,k}, \boldsymbol{\theta}_{\xi,w}) \propto p(\boldsymbol{\theta}_{\xi,w})$; while, following Gardoni et al. [19], for $\boldsymbol{\Sigma}_\xi$ we adopt $p(\boldsymbol{\Sigma}_\xi) \propto 1/\sigma_\xi$, so that the prior distribution assumes the following form,

$$p(\boldsymbol{\Theta}_\xi) \propto \frac{p(\boldsymbol{\theta}_{\xi,w})}{\sigma_\xi}. \quad (39)$$

A more extended description of the prior distribution $p(\boldsymbol{\Theta}_\xi)$ is in [11], as well as for a first Bayesian updating using the first set of data $\mathbf{y}_{\xi,1}$, which leads to the posterior distribution $p(\boldsymbol{\Theta}_\xi | \mathbf{y}_{\xi,1})$.

5.3 Pipe enlargement

Since each component of $\boldsymbol{\theta}_{\delta_r}$ has a range in the interval $[0, 1]$, we assume that their prior distributions are Beta distributions $f_{\Theta_{\delta_r,i}}(\theta_{\delta_r,i}) = \theta_{\delta_r,i}^{q_i-1} (1-\theta_{\delta_r,i})^{r_i-1} / B(q_i, r_i)$, in which $B(q_i, r_i)$ is the Beta function of the shape parameters q_i and r_i .

Considering a noninformative prior distribution for θ_p , and referring to Box and Tiao [24], we take it to be locally uniform. The joint probability density function $p(\theta_\alpha, \theta_\beta)$ is obtained by coupling those marginals of the single parameters through the Nataf model [33], which assumes in this case the following form

$$p(\theta_\alpha, \theta_\beta) = f_{\Theta_\alpha}(\theta_\alpha) f_{\Theta_\beta}(\theta_\beta) \frac{\varphi_2(\mathbf{z}, \mathbf{R}')}{\varphi(z_\alpha) \varphi(z_\beta)}, \quad (40)$$

where \mathbf{z} is the Gaussian random vector whose components z_α and z_β can be determined from the related θ_α and θ_β using the equalities $Z_\alpha = \Phi^{-1}[F_{\Theta_\alpha}(\theta_\alpha)]$ and $Z_\beta = \Phi^{-1}[F_{\Theta_\beta}(\theta_\beta)]$, that associates the marginal cumulative distribution functions $F_{\Theta_\alpha}(\cdot)$ and $F_{\Theta_\beta}(\cdot)$, obtained by numerically integrating $f_{\Theta_\alpha}(\cdot)$ and $f_{\Theta_\beta}(\cdot)$, and the standard normal cumulative probability distribution function Φ ($\Phi^{-1}(\cdot)$ denotes its inverse). In Eq. (40), $\varphi_2(\cdot)$ is the bivariate standard

normal probability density function whose correlation matrix \mathbf{R}' is defined by the coefficient $\rho'_{\alpha\beta}$ obtained from the correlation coefficient $\rho_{\alpha\beta}$, between θ_α and θ_β as follows

$$\rho_{\alpha\beta} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\theta_\alpha - \mu_{\Theta_\alpha}}{\sigma_{\Theta_\alpha}} \right) \left(\frac{\theta_\beta - \mu_{\Theta_\beta}}{\sigma_{\Theta_\beta}} \right) \varphi_2(z_\alpha, z_\beta, \rho'_{\alpha\beta}) dz_\alpha dz_\beta, \quad (41)$$

where μ_{Θ_α} , σ_{Θ_α} , μ_{Θ_β} and σ_{Θ_β} are respectively the mean and the standard deviation of the parameters θ_α and θ_β .

Referring to the simplified approach for time to failure and peak flow presented by Bonelli et al. [10], a first approximation of a stationary value of the head loss factor α is about 0.26, and considering the height and the material of the dam physical model described by Hanson et al. [21], an average value of the momentum loss factor β can be assumed to be equal to 0.55. Thus, assuming these values as the means μ_{Θ_α} and μ_{Θ_β} and considering a reasonable value of the coefficient of variation equal to 0.20, we can obtain the Beta distribution shape parameters $q_\alpha = 18.24$, $r_\alpha = 51.91$ and $q_\beta = 10.70$, $r_\beta = 8.76$. Considering herein a reasonable value of $\rho_{\alpha\beta} = 0.50$, the Eq. (41) furnishes an almost identical value of $\rho_{\alpha\beta} = 0.502$.

In this paper, λ is assumed to follow a standard normal distribution. Considering noninformative prior distribution for λ , and referring to Box and Tiao [24], it is assumed to be locally uniform while following Gardoni et al. [19], we assume $p(\Sigma_{\delta_r}) \propto 1/\sigma_{\delta_r}$. We assume the a priori independence between Θ_ξ and $(\theta_{\delta_r}, \theta_r, \lambda)$ as well as the variance of the univariate model $\Sigma_{\delta_r} = \sigma_{\delta_r}^2$. Following this approach, the joint prior distribution can be written as

$$p(\Theta_\xi, \Theta_{\delta_r}, \lambda | \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi,n_\xi}) \propto \frac{\prod_{\xi} p(\Theta_\xi | \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi,n_\xi})}{\sigma_{\delta_r}} \quad (42)$$

which leads, together with Eq. (38), to the following general form of the posterior distribution

$$p\left(\Theta_{\xi}, \Theta_{\delta_r}, \lambda \mid \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi, n_{\xi}}, \mathbf{y}_{\delta_r}\right) \propto \frac{1}{\sigma_{\delta_r}^{n_{\delta_r}+1}} \prod_{i=1}^{n_{\delta_r}} \left\{ \varphi \left[\frac{r_{\delta_{ri}}\left(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \lambda\right)}{\sigma_{\delta_r}} \right] \right\} \quad (43)$$

$$\times p\left(\theta_{\alpha}, \theta_{\beta}\right) \times p\left(\Theta_{\tau_c} \mid \mathbf{y}_{\tau_c,1}, \dots, \mathbf{y}_{\tau_c, n_{\tau_c}}\right) p\left(\Theta_{C_e} \mid \mathbf{y}_{C_e,1}, \dots, \mathbf{y}_{C_e, n_{C_e}}\right)$$

which simplifies, in the case of 2 sets of data $\mathbf{y}_{\xi,1}$ and $\mathbf{y}_{\xi,2}$, to the particular form

$$p\left(\Theta_{\xi}, \Theta_{\delta_r}, \lambda \mid \mathbf{y}_{\xi,1}, \mathbf{y}_{\xi,2}, \mathbf{y}_{\delta_r}\right) \propto \frac{1}{\sigma_{\delta_r}^{n_{\delta_r}+1}} \prod_{i=1}^{n_{\delta_r}} \left\{ \varphi \left[\frac{r_{\delta_{ri}}\left(\Theta_{\xi}, \theta_{\delta_r}, \theta_r, \lambda\right)}{\sigma_{\delta_r}} \right] \right\} \quad (44)$$

$$\times p\left(\theta_{\alpha}, \theta_{\beta}\right) \times p\left(\Theta_{\tau_c} \mid \mathbf{y}_{\tau_c,1}, \mathbf{y}_{\tau_c,2}\right) p\left(\Theta_{C_e} \mid \mathbf{y}_{C_e,1}, \mathbf{y}_{C_e,2}\right)$$

that is adopted in the remainder of this paper.

6. Calibrated probabilistic models

6.1 General remarks on C_{ξ} models' correction

We consider a first candidate explanatory function $h_{k,1}(\mathbf{x}_k) = 1$, in order to capture the potential bias of the deterministic models of each of the two components, that is independent on \mathbf{x}_k and λ . Relating to the non-cohesive part (sand component), we select $\eta_{s,2}(\mathbf{x}_s) = d_{s,50}/d_{c/s}$, $\eta_{s,3}(\mathbf{x}_s) = d_{s,90}/d_{c/s}$ and $\eta_{s,4}(\mathbf{x}_s) = s_{\%}$ (being $s_{\%}$ the total percentage of sand) to capture the potential bias associated to each component of \mathbf{x}_s . For the cohesive part, we select $\eta_{m,2}(\mathbf{x}_m) = \text{PI}$ to correct the bias associated to the deterministic terms $\hat{\xi}_m^*(\mathbf{x}_m)$, $\eta_{m,3}(\mathbf{x}_m) = \rho_d/\rho_{d,\max}$ to account for the influence of the percentage of compaction referred to the maximum dry density $\rho_{d,\max}$ of the mixture, $\eta_{m,4}(\mathbf{x}_m) = \omega$ to capture the possible effect of the water content percentage ω , $\eta_{m,5}(\mathbf{x}_m) = \text{OWC}$ to correct the bias depending on the optimum water content percentage (OWC), and finally $\eta_{m,6}(\mathbf{x}_m) = S$ to account for the degree of saturation S expressed in percentage. The explanatory functions $h_{k,i}(\mathbf{x}_k)$ are obtained by means of the Eq. (2) where the value of λ is determined simultaneously to other unknown parameters

using the maximum likelihood approach [34]. In all the cases related to the C_ξ models, such a criterion gives λ equal to 0, that makes the Yang transformation [23] corresponding to the natural logarithm and its inverse to the exponential function.

Once we performed the transformation of the basis functions, a stepwise deletion procedure can be applied in order to simplifying the models, removing the unimportant terms. Following to Gardoni et al. [18,19], the stepwise deletion process, adopted for each set of data $(\mathbf{y}_{\xi,q}, \mathbf{x}_{k,q}, \mathbf{x}_{w,q})$ concerned the following steps:

- 1) Compute the posterior mode of the parameters Θ_ξ and the related approximation of

$$\text{covariance matrix } \Sigma_{\Theta_\xi \Theta_\xi} = - \left\{ \nabla \nabla \ln \left[p(\Theta_\xi) \prod_{q=1}^{n_q} L(\Theta_\xi | \mathbf{y}_{\xi,q}) \right] \right\}^{-1}, \text{ referring to Richards}$$

[35].

- 2) Identify the term $h_{k,i}(\mathbf{x}_k)$ whose coefficient $\theta_{\xi,k}$ has the largest coefficient of variation (C.o.V.). Such a term is the least informative among all the explanatory functions, so one can choose to drop it from the correction term $\gamma_{\xi,k}(\mathbf{x}_k, \theta_{\xi,k}, \lambda = 0)$.
- 3) Assess the reduced model by checking if the value σ_ξ has not increased by an unacceptable amount. If so, accept the reduced model and repeat the step 1 and 2 for further reductions. Otherwise, the performed reduction is not desirable and the model form before such a reduction is as parsimonious as possible.
- 4) Compute the posterior statistics of the parameters Θ_ξ using the parallel computing algorithm proposed by Andreini et al. [11].

It is worth noting that what makes unacceptable the increase of σ_ξ is the level of accuracy and, at the same time, of parsimony desired for the specific problem.

6.2. Critical Shear Stress Model

This section presents the results of the stepwise deletion and the estimation of the posterior statistics of the parameters Θ_{τ_c} . Considering the first set of data, Andreini et al. [11] were left with $h_{s,2}(\mathbf{x}_s)$, $h_{s,3}(\mathbf{x}_s)$, and $h_{s,4}(\mathbf{x}_s)$ for the sand component, and $h_{m,1}(\mathbf{x}_m)$ and $h_{m,5}(\mathbf{x}_m)$ for the cohesive component. Instead, we remain with the entire vector of explanatory functions $\mathbf{h}_s(\mathbf{x}_s)$ and $h_{m,1}(\mathbf{x}_m)$, $h_{m,2}(\mathbf{x}_m)$, $h_{m,4}(\mathbf{x}_m)$ and $h_{m,5}(\mathbf{x}_m)$, adding the second set of data ($\mathbf{y}_{\tau_c,2}$, $\mathbf{x}_{k,2}$, $\mathbf{x}_{w,2}$) and repeating the stepwise deletion procedure. Figure 3 shows the resulting most parsimonious and accurate model. Table 2 reports the posterior statistics of the parameters after the updating, which involves the following expressions of the correction terms:

$$\begin{aligned} \gamma_{\tau_c,s}(\mathbf{x}_s, \Theta_{\tau_c,s}, \lambda = 0) &= \theta_{\tau_c,s,11} + \theta_{\tau_c,s,12} \ln\left(\frac{d_{s,50}}{d_{c/s}}\right) + \theta_{\tau_c,s,13} \ln\left(\frac{d_{s,90}}{d_{c/s}}\right) \\ &+ \theta_{\tau_c,s,24} \ln\left(\frac{d_{s,50}}{d_{c/s}}\right) \ln(s\%) + \theta_{\tau_c,s,34} \ln\left(\frac{d_{s,90}}{d_{c/s}}\right) \ln(s\%) \end{aligned} \quad (45)$$

$$\begin{aligned} \gamma_{\tau_c,m}(\mathbf{x}_m, \Theta_{\tau_c,m}, \lambda = 0) &= \theta_{\tau_c,m,11} + \theta_{\tau_c,m,12} \ln(\text{PI}) + \theta_{\tau_c,m,15} \ln(\text{OWC}) \\ &+ \theta_{\tau_c,m,25} \ln(\text{PI}) \ln(\text{OWC}) + \theta_{\tau_c,m,45} \ln(\omega) \ln(\text{OWC}) \end{aligned} \quad (46)$$

Considering only the first set of data, it seems that the critical percentage of mud $P_{m,cr}$, represented by the parameter $\theta_{\tau_c,w,1}$, be strictly depending on the optimum water content OWC, since their correlation coefficient reported by Andreini et al. [11] is equal to -0.82 . This would imply that the change of the erosion initiation behavior could be expressed in terms of the OWC threshold, but this is not confirmed in the updated model presented here using the second set of data. On the other hand, the posterior mean of $P_{m,cr}$ remains stable around the 30% after the updating, confirming what also observed by Panagiotopoulos et al. [36]. Another observation is that the addition of more erodible materials for the second updating lead to a negative mean ($=-43.21$) of $\theta_{\tau_c,s,11}$, which suggests that the deterministic model $\hat{\tau}_{c,s}(\mathbf{x}_s)$ overestimates the

critical shear stress, independently on \mathbf{x}_s . Furthermore, the large positive values detected for the means of the parameters $\theta_{\tau_c,m,11}$ and $\theta_{\tau_c,m,12}$ indicate that the contribution of the mud component is strongly underestimated by $\hat{\tau}_{c,m}(\mathbf{x}_m)$, whose bias is also due to the absence in its formulation of the OWC and, in the case of the second updating, even of the water content ω .

Figure 4 shows a comparison between the measured and the predicted values of the critical shear stress, applying the deterministic and the reduced probabilistic models, after the first and the second updating, substituting in the models the posterior median values of the parameters.

6.3. Coefficient of Erosion

As for the critical shear stress, the stepwise deletion has been performed to reduce number of elements of Θ_{C_e} , in order to estimate the related posterior statistics of a smaller number of parameters. Considering the first set of data, Andreini et al. [11] found that all of the explanatory functions are needed to correct the bias associated to the variables \mathbf{x}_k . This fact is not confirmed when the second set of readings is added: the stepwise deletion in the updated model (Figure 5) reaches a large variation of σ_{C_e} after 24 steps, allowing to correct the bias of the cohesive component by means of the unique term $\theta_{C_e,m,25}h_{m,2}(\mathbf{x}_m)h_{m,5}(\mathbf{x}_m)$. All of the explanatory functions $h_{s,i}(\mathbf{x}_s)$, except the first one, are instead needed to correct the bias of the deterministic model associated to the non-cohesive component. In this case, the corrective terms can be expressed as

$$\begin{aligned} \gamma_{C_e,s}(\mathbf{x}_s, \Theta_{C_e,s}, \lambda = 0) = & \theta_{C_e,s,14} \ln(s\%) + \theta_{C_e,s,22} \left[\ln\left(\frac{d_{s,50}}{d_{c/s}}\right) \right]^2 + \theta_{C_e,s,23} \ln\left(\frac{d_{s,50}}{d_{c/s}}\right) \ln\left(\frac{d_{s,90}}{d_{c/s}}\right) \\ & + \theta_{C_e,s,24} \ln\left(\frac{d_{s,50}}{d_{c/s}}\right) \ln(s\%) + \theta_{C_e,s,33} \left[\ln\left(\frac{d_{s,90}}{d_{c/s}}\right) \right]^2 + \theta_{C_e,s,34} \ln\left(\frac{d_{s,90}}{d_{c/s}}\right) \ln(s\%) + \theta_{C_e,s,44} \left[\ln(s\%) \right]^2 \end{aligned} \quad (47)$$

$$\gamma_{C_e,m}(\mathbf{x}_m, \boldsymbol{\theta}_{C_e,m}, \lambda = 0) = \theta_{\tau_c,m,25} \ln(\text{PI}) \ln(\text{OWC}). \quad (48)$$

The posterior statistics of the reduced parameters for each updating are reported in Table 3, whereas a comparison between the applications of the deterministic and the probabilistic models is shown in the Figure 6. It is easy to recognize how the corrective terms reduce the bias associated to the deterministic model. The second Bayesian updating of the parameters allows to have a more accurate model for more erodible soils.

Only in the case of the second updating, the mean value of the critical percentage of mud assumes a value out of the range 20-30%. Moreover, the explanatory function associated to the OWC is present in all the forms presented for the correction terms $\gamma_{C_e,m}(\mathbf{x}_m, \boldsymbol{\theta}_{C_e,m}, \lambda = 0)$. Indeed, such a geotechnical property seems to be not negligible to describe the cohesive behavior, both in terms of critical shear stress and coefficient of erosion.

6.4. Pipe enlargement model

The applications of the probabilistic model to the set of data, from the physical model mentioned in Section 5.1, are herein presented. The value of λ is also in this case determined simultaneously to the other parameters using the maximum likelihood approach [34]. As for the C_ξ models, we obtain λ equal to 0. We consider firstly the task of calibrating the parameters

$\boldsymbol{\Theta}_{\delta_r} = (\theta_{\delta_r}, \theta_r, \sigma_{\delta_r})$ substituting the values of the most updated posterior means $\hat{\boldsymbol{\Theta}}_{\tau_c}$ and $\hat{\boldsymbol{\Theta}}_{C_e}$

(Tables 2 and 3) in Eqs. (31-38). In such a case, the Eq. (44) reduces to the form

$$p(\boldsymbol{\Theta}_{\delta_r} | \mathbf{y}_{\delta_r}) \propto \frac{p(\theta_\alpha, \theta_\beta)}{\sigma_{\delta_r}^{n_{\delta_r} + 1}} \prod_{l=1}^{n_{\delta_r}} \left\{ \varphi \left[\frac{r_{\delta_r,l}(\hat{\boldsymbol{\Theta}}_\xi, \theta_{\delta_r}, \theta_r)}{\sigma_{\delta_r}} \right] \right\}, \quad (49)$$

where $p(\theta_\alpha, \theta_\beta)$ is the joint prior distribution of the parameters θ_α and θ_β given by Eq. (40).

Table 4 reports the posterior distribution descriptors of the parameters $\boldsymbol{\Theta}_{\delta_r}$ obtained in this

way and Figure 7 shows a comparison between measured values of the pipe radius and predicted ones by means of the probabilistic model calibrated in this way.

Instead, the problem of the total calibration of the parameters $(\Theta_\xi, \Theta_{\delta_r})$ on the totality of the sets of available data requires the determination of the descriptors of the related joint posterior distribution defined in Eq. (44). Table 5 reports the posterior distribution descriptors obtained by the calibration process. Figure 8 shows a comparison between the measured and predicted values for the critical shear stress, coefficient of erosion, and radius variation obtained by this global parameter calibration.

It can be seen that this global calibration process does not alter the accuracy of the erosion parameter models, while makes the pipe enlargement model overestimating the experimental readings. Moreover, the posterior mean of the parameters θ_α and θ_β presents lower values but still within the ranges given by Bonelli et al. [10] for dams of small dimensions.

7. Fragility estimates

7.1 General Formulation

The probabilistic models presented in this paper can be used to compute the fragility, which can be defined in this content as the probability of attaining or exceeding prescribed limits for a given set of boundary variables. Following Ditlevsen and Madsen [37] and Gardoni [17], we define a limit state function $g_{\delta_r}(\mathbf{x}_{\delta_r}, \mathbf{s}_{\delta_r}, \Theta_\xi, \Theta_{\delta_r})$ as a mathematical model such that the event $\{g_{\delta_r}(\mathbf{x}_{\delta_r}, \mathbf{s}_{\delta_r}, \Theta_\xi, \Theta_{\delta_r}) \leq 0\}$ represents the attainment or exceedance of such a limit. Using the developed capacity models, the limit state function can be formulated as

$$g_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_\xi, \Theta_{\delta_r}) = C_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_\xi, \Theta_{\delta_r}) - D_{\delta_r}(\mathbf{x}_{\delta_r}, \mathbf{s}_{\delta_r}), \quad (50)$$

where $D_{\delta_r}(\mathbf{x}_{\delta_r}, \mathbf{s}_{\delta_r})$ represents the demand models, in which \mathbf{s}_{δ_r} is the vector of demand variables. This way, the fragility can be stated as

$$F_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) = P\left[\left\{g_{\delta_r}(\mathbf{x}_{\delta_r}, \mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) \leq 0\right\} \middle| \mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}\right], \quad (51)$$

which is a function of both the variables and the parameters Θ_{ξ} and Θ_{δ_r} , in order to underline the fact that an estimate depends on how we consider the uncertainties associated to such parameters [19,20].

A point estimate of the fragility can be obtained using a point estimate $\hat{\Theta}_{\xi}$ and $\hat{\Theta}_{\delta_r}$, such as the posterior means or the Maximum Likelihood Estimation (MLE), thus ignoring the uncertainties related to the model parameters. We indicate the corresponding point estimate fragility as

$$\hat{F}_{\delta_r}(\mathbf{s}_{\delta_r}) = F_{\delta_r}(\mathbf{s}_{\delta_r}, \hat{\Theta}_{\xi}, \hat{\Theta}_{\delta_r}). \quad (52)$$

In these cases, all the uncertainties are associated to the models' variables \mathbf{x}_k , \mathbf{x}_w and \mathbf{x}_r , and in the model error term ε_{δ_r} , which is basically aleatory in nature.

Alternatively, to account for the epistemic uncertainties in models' parameters, we must consider Θ_{ξ} and Θ_{δ_r} as random variables and, consequently, the predictive estimate of fragility $\tilde{F}_{\delta_r}(\mathbf{s}_{\delta_r})$ can be evaluated as the expected value of the function in the Eq. (51) over the posterior distribution expressed by the Eq. (44), that are:

$$\tilde{F}_{\delta_r}(\mathbf{s}_{\delta_r}) = \int F_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) p(\Theta_{\xi}, \Theta_{\delta_r} | \mathbf{y}_{\xi,1}, \dots, \mathbf{y}_{\xi,n_{\xi}}, \mathbf{y}_{\delta_r}) d\Theta_{\xi} d\Theta_{\delta_r}. \quad (53)$$

Following Gardoni et al. [18,19], the explicit effect of the epistemic uncertainty in the model parameters can be evaluated by constructing confidence bounds on the reliability index

$$\beta_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) = \Phi^{-1}\left[1 - F_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r})\right] \text{ corresponding to the fragility which are written}$$

as

$$\{\Phi[-\tilde{\beta}_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) - \sigma_{\beta_{\delta_r}}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r})], \Phi[-\tilde{\beta}_{\delta_r}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) + \sigma_{\beta_{\delta_r}}(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r})]\}, \quad (54)$$

where

$$\sigma_{\beta_{\delta_r}}^2 \left(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r} \right) \approx \nabla_{\Theta_{[\tau_c, C_e, \delta_r]}} \beta_{\delta_r} \left(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r} \right) \Sigma_{\Theta_{[\tau_c, C_e, \delta_r]} \Theta_{[\tau_c, C_e, \delta_r]}} \nabla_{\Theta_{[\tau_c, C_e, \delta_r]}} \beta_{\delta_r} \left(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r} \right)^T \quad (55)$$

is the variance of the reliability index approximated by the first-order Taylor series expansions around the posterior mean point $\mathbf{M}_{\Theta_{[\tau_c, C_e, \delta_r]}}$ (in which $\nabla_{\Theta_{[\tau_c, C_e, \delta_r]}} \beta_{\delta_r} \left(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r} \right)$ is the gradient row vector of $\beta_{\delta_r} \left(\mathbf{s}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r} \right)$ evaluated at $\mathbf{M}_{\Theta_{[\tau_c, C_e, \delta_r]}}$ and $\Sigma_{\Theta_{[\tau_c, C_e, \delta_r]} \Theta_{[\tau_c, C_e, \delta_r]}}$ is the posterior covariance matrix of $\Theta_{[\tau_c, C_e, \delta_r]} = [\Theta_{\tau_c}, \Theta_{C_e}, \Theta_{\delta_r}]$), and $\tilde{\beta}_{\delta_r} = \Phi^{-1}[1 - \tilde{F}_{\delta_r}(\mathbf{s}_{\delta_r})]$ is the generalized reliability index corresponding to $\tilde{F}_{\delta_r}(\mathbf{s}_{\delta_r})$.

7.2 Application to a typical earth dam

7.2.1 Problem description

As presented in the introduction, the evolution of the concentrated leak erosion phenomenon, which leads to the collapse of a water retaining structure, can be subdivided in four main phases. These phases can be modeled as a series system of events, in which we have to consider, as a further element, the occurrence of any successful action of detection, intervention or repair. Referring to the Figure 9, where the single events are E^I = initiation, E^C = continuation of erosion, E^P = progression to form a pipe, E^D = not detection, intervention and repair and B = Breach mechanism, we can consider that the failure event is given by the occurrence of all the mentioned events and can be evaluated as $(E^I \cap E^C \cap E^P \cap E^D) \cap B$, where \cap denotes the intersection, or in a simpler notation as $(E^I E^C E^P E^D) B$. Thus, considering each couple of events as mutually exclusive, we have that the probability of failure $P_f = P[(E^I E^C E^P E^D) B]$ can be expanded in $P_f = P[B | (E^I E^C E^P E^D)] P(E^I E^C E^P E^D)$ by means of the multiplication rule [38], which leads to the following expression when applied iteratively:

$$P_f = P(B | E^I E^C E^P E^D) P(E^D | E^I E^C E^P) P(E^P | E^I E^C) P(E^C | E^I) P(E^I). \quad (56)$$

In particular, E^D concerns 3 sub-events: E^{OL} = observation of the concentrated leak because it is observable, E^{DT} = leak detected and E^{ul} = interventions fail. In Figure 9, the overlined symbol \bar{E} indicates the complement of the generic event E .

The fragilities' formulation presented in the Section 7.1 can be used to determine the conditioned probabilities $P(\bar{E}^{DT} | E^I E^C E^P E^{OL})$ and $P(E^{ul} | E^I E^C E^P E^{OL} E^{DT})$, as described in the following section. For this purpose, we refer to the general case of an embankment dam without any filter, having $H_{dam} = 15\text{m}$, $b_{min} = 4\text{m}$ and $c_u = c_d = 1.5$, and made of the soil with null PI and the properties indicated in the Table 6.

7.2.2 Detection, intervention and repair

The probability of not detection, intervention and repair is strictly dependent on the total time required for progression of the pipe enlargement and the development of the breach mechanism.

We consider the fragility $\tilde{F}_{\delta_r}(\mathbf{s}_{\delta_r})$, associated to the limit state function $g_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) = C_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) - D_{\delta_r}(R_D, R_0)$, where $D_{\delta_r}(R_D, R_0) = f_{\lambda}(R_D / R_0)$ and R_D can be seen as a control value of the pipe radius.

The application to the case under study is shown in the Figure 10, where the fragility curves are plotted as a function of the time t and the radius R_D , for $z = h_w = H_{dam} = 15\text{ m}$ and $R_0 = 150\text{ mm}$. Such curves have the expression

$$\begin{aligned} \tilde{F}_{\delta_r}(z, h_w, R_0, t, R_D) &= P\left[g_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) \leq 0 \mid z, h_w, R_0, t, R_D \right] = \\ &= P\left[R(t) \leq R_D \mid z, h_w, R_0, t, R_D \right] \end{aligned} \quad (57)$$

and allow us to determine the probability of having a certain remaining time Δt to the breach, given a definite value of the radius variation interval.

As an example, considering that the failure occurs when the pipe achieves an ultimate value of the radius $R_u(t_u) = H_{dam} / 3 = 5\text{ m}$ at time t_u , as in Bonelli et al. [10], the probability

to not achieve the failure at certain time t is given by putting $R_D = R_u$ in Eq. (57) and, for the case under study, it simplifies to

$$\tilde{F}_{\delta_r}(t) = P\left[g_{\delta_r}(\mathbf{x}_{\delta_r}, \Theta_{\xi}, \Theta_{\delta_r}) \leq 0 \mid t, R_D = R_u\right] = P\left[t \leq t_u \mid t, R_D = R_u\right], \quad (58)$$

which is plotted in the Figure 11. Fell et al. [22] reported the values of the probability of not detection and not intervention as a function of the period of time $[t, t + \Delta t]$, to which belongs the failure time t_u . The probability of such events can be calculated by the application of the total probability rule placed in the forms

$$P(\bar{E}^{DT} \mid E^I E^C E^P E^{OL}) = \sum_{\Delta t} P(\bar{E}^{DT} \mid E^I E^C E^P E^{OL} \Delta t) P(t, \Delta t), \quad (59)$$

$$P(E^{ul} \mid E^I E^C E^P E^{OL} E^{DT}) = \sum_{\Delta t} P(E^{ul} \mid E^I E^C E^P E^{OL} E^{DT} \Delta t) P(\Delta t), \quad (60)$$

where $P(t, \Delta t) = \tilde{F}_{\delta_r}(t) - \tilde{F}_{\delta_r}(t + \Delta t)$ is the probability that $t_u \in [t, t + \Delta t]$. In this example, the terms of Eqs. (59) and (60) are reported in Table 7 and we have $P(\bar{E}^{DT} \mid E^I E^C E^P E^{OL}) = 0.471$ and $P(E^{ul} \mid E^I E^C E^P E^{OL} E^{DT}) = 0.673$.

8. Conclusions

Probabilistic models for the critical shear stress, the coefficient of erosion and pipe enlargement for concentrated leak mechanisms are presented. The models are able to predict the influence on the vulnerability to erosion of the physical and geotechnical properties of the soils for water retaining structures. The probabilistic models of the erosion parameters are constructed considering both the non-cohesive and cohesive components. The contribution of each component is modeled using empirical deterministic models, whose bias is corrected by additional terms that also improve the overall accuracy of the predictions. A stepwise deletion process is used to simplify the erosion parameters' model reducing the number of the correction terms, deleting those are least informative while keeping an adequate level of accuracy. The

proposed models are calibrated and updated using test data. The updating is performed using a Bayesian approach.

The updated models of the erosion parameters are then used in the formulation of nested probabilistic model of the pipe radius evolution in an embankment. The unknown parameters of such a model are calibrated using the in-field data from an experimental test carried out on a scaled physical model of a dam. The totality of the parameters of all the models are subsequently calibrated on the basis of all the experimental data.

The proposed probabilistic models are used for the assessment of the probability of not detection and not intervention in case of a concentrated leak erosion mechanism. The results show how transversal cracks with a radius of few centimeters can be the cause of pipe enlargement rates, which can lead to breach mechanism in few days with a significant probability.

Acknowledgements

The Authors acknowledge Rotary International that supported this work through the Global Grant 1411222, and the Istituto Nazionale di Fisica Nucleare (INFN), section of Pisa (Italy), for the support and computing resources.

References

- [1] Richards KS, Reddy KR. Critical appraisal of piping phenomena in earth dams. *Bull Eng Geol Environ* 2007;66:381–402. doi:10.1007/s10064-007-0095-0.
- [2] Foster M, Fell R, Spannagle M. The statistics of embankment dam failures and accidents. *Can Geotech J* 2000;37:1000–24. doi:10.1139/t00-030.
- [3] Foster M, Fry J-J. State of The Art on the Likelihood of Internal Erosion of Dams and Levees by Means of Testing. In: Bonelli S, Nicot F, editors. *Eros. Geomech. Appl. to Dams Levees*, Hoboken, NJ: John Wiley & Sons Inc; 2013, p. 1–85. doi:10.1002/9781118577165.ch1.

- [4] Hanson GJ. Surface erodibility of earthen channels at high stresses. Part II, developing an in-situ device. *Trans ASAE* 1990;30:132–7. doi:10.13031/2013.31306.
- [5] Hanson GJ. Development of a jet index to characterize erosion resistance of soils in earthen spillways. *Trans ASAE* 1991;34:2015–20. doi:10.13031/2013.31831.
- [6] Hanson GJ, Cook KR. Apparatus test procedure and analytical methods to measure soil erodibility in-situ. *Appl Eng Agric* 2004;20:455–62. doi:10.13031/2013.16492.
- [7] Wan CF, Fell R. Investigation of internal erosion and piping of Soils in embankment dams by the slot erosion test and the hole erosion test. Sydney, AUSTRALIA: 2002.
- [8] Wan CF, Fell R. Laboratory tests on the rate of piping erosion of soils in embankment dams. *Geotech Test J* 2004;27:295–303. doi:10.1520/GTJ11903.
- [9] Wan CF, Fell R. Investigation of Rate of Erosion of Soils in Embankment Dams. *J Geotech Geoenvironmental Eng* 2004;130:373. doi:10.1061/(ASCE)1090-0241(2004)130:4(373).
- [10] Bonelli S, Fell R, Benahamed N. Concentrated Leak Erosion. In: Bonelli S, Nicot F, editors. *Eros. Geomech. Appl. to Dams Levees*, Hoboken, NJ: John Wiley & Sons Inc; 2013, p. 271–341. doi:10.1002/9781118577165.ch4.
- [11] Andreini M, Gardoni P, Pagliara S, Sassu M. Probabilistic Models for Erosion Parameters and Reliability Analysis of Earth Dams and Levees. *ASCE-ASME J Risk Uncertain Eng Syst Part A Civ Eng* 2016;2. doi:10.1061/AJRUA6.0000878.
- [12] Benahmed N, Bonelli S. Investigating concentrated leak erosion behaviour of cohesive soils by performing hole erosion tests. *Eur J Environ Civ Eng* 2012;16:43–58. doi:10.1080/19648189.2012.667667.
- [13] Benahmed N, Chevalier C, Bonelli S. Concentrated Leak Erosion. In: Bonelli S, editor. *Eros. Geomaterials*, Hoboken, NJ: John Wiley & Sons Inc; 2013, p. 155–86.

doi:10.1002/9781118561737.ch5.

- [14] Bonelli S, Brivois O. The scaling law in the hole erosion test with a constant pressure drop. *Int J Numer Anal Methods Geomech* 2008;32:1573–95. doi:10.1002/nag.683.
- [15] Bonelli S, Benahamed N. Piping flow erosion in water retaining structures: inferring erosion rates from hole erosion tests and quantifying the failure time. *IECS 2010, 8th ICOLD Eur. Club Symp. Dam Safety-Sustainability a Chang. Environ.*, 2010, p. 467–72.
- [16] Bonelli S, Benahmed N. Piping flow erosion in water retaining structures. *Int J Hydropower Dams* 2011;18:94–8.
- [17] Gardoni P, editor. *Risk and Reliability Analysis: Theory and Applications*. Cham, Switzerland: Springer International Publishing; 2017. doi:10.1007/978-3-319-52425-2.
- [18] Gardoni P, Der Kiureghian A, Mosalam KM. *Probabilistic Models and Fragility Estimates for Bridge Components and Systems Probabilistic Models and Fragility Estimates for Bridge Components and Systems*. vol. 13. Berkeley, CA: 2002.
- [19] Gardoni P, Der Kiureghian A, Mosalam KM. Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations. *J Eng Mech* 2002;128:1024–38. doi:10.1061/(ASCE)0733-9399(2002)128:10(1024).
- [20] Gardoni P, Mosalam KM, Der Kiureghian A. Probabilistic seismic demand models and fragility estimates for RC bridges. *J Earthq Eng* 2003;7:79–106. doi:10.1080/13632460309350474.
- [21] Hanson GJ, Tejral RD, Hunt SL, Temple DM. Internal erosion and impact of erosion resistance. *30th US Soc. Dams Annual Meet. Conf.*, Sacramento: 2010, p. 773–84.
- [22] Fell R, Foster M, Cyganiewicz J, Sills G, Vroman N, Davidson R. *A Unified Methods*

- for Estimating Probabilities of Failure of Embankment Dams by Internal Erosion and Piping Guidance Document Volumn 1 and 2. Sydney, AUSTRALIA: 2009.
- [23] Yang Z. A modified family of power transformations. *Econ Lett* 2006;92:14–9. doi:10.1016/j.econlet.2006.01.011.
- [24] Box GEP, Tiao GC. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons; 1992.
- [25] van Ledden M. *Sand-mud segregation in estuaries and tidal basins*. Delfth University of Technology, 2003.
- [26] van Rijn LC. *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas*. Amsterdam, The Netherlands: Aqua Publications; 1993.
- [27] Meyer-Peter, E. and Müller R. *Formulas for Bed-Load Transport*. Int Assoc Hydraul Struct Researach - Zweite Tagung - Second Meet - Deuxième Réunion 1948.
- [28] Smerdon ET, Beasley RP. *The tractive force theory applied to stability of open channels in cohesive soils*. Columbia, Missouri: 1959.
- [29] Hanson GJ, Simon a. *Erodibility of cohesive streambeds in the loess area of the Midwestern USA*. *Hydrol Process* 2001;15:23–38. doi:10.1002/hyp.149.
- [30] Wahl TL, Erdogan Z. *Erosion Indices of Soils Used in ARS Piping Breach Tests*. Denver: 2008.
- [31] Wahl TL, Regazzoni P-L, Erdogan Z. *Determining Erosion Indices of Cohesive Soils with the Hole Erosion Test and Jet Erosion Test*. Denver, Colorado: 2008.
- [32] Wahl TL. *Relating HET and JET test results to internal erosion field tests*. *Jt. Fed. Interag. Conf. Sediment. Hydrol. Model.*, Las Vegas, NV: 2010, p. 1–12.
- [33] Liu P-L, Der Kiureghian A. *Multivariate distribution models with prescribed marginals*

and covariances. Probabilistic Eng Mech 1986;1:105–12. doi:10.1016/0266-8920(86)90033-0.

- [34] Tabandeh A, Gardoni P. Probabilistic capacity models and fragility estimates for RC columns retrofitted with FRP composites. Eng Struct 2014;74:13–22. doi:10.1016/j.engstruct.2014.05.005.
- [35] Richards FSG. A Method of Maximum-likelihood Estimation. J R Stat Soc Ser B 1961;23:469–75.
- [36] Panagiotopoulos I, Voulgaris G, Collins MB. The influence of clay on the threshold of movement of fine sandy beds. Coast Eng 1997;32:19–43. doi:10.1016/S0378-3839(97)00013-6.
- [37] Ditlevsen O, Madsen HO. Structural Reliability Methods. Chichester: John Wiley & Sons Inc; 1996.
- [38] Ang AH-S, Tang WH. Probability Concepts in Engineering Planning and Design: Basic principles. Planning 1975;1:217.

Tables

Table 1: Combination terms for the erosion parameters' models: transformations, weight and corrective functions for the sand and mud components.

Combination terms		τ_c	C_e
$f_c(\cdot)$		Identity	Natural Logarithm
$w_{\xi,s}(\mathbf{x}_w, \boldsymbol{\theta}_{\xi,w})$	if $P_m \leq \theta_{\xi,w,1}$	$(1 + P_m)^{\theta_{\tau_c,w,2}}$	1
	if $P_m > \theta_{\xi,w,1}$	$\frac{(1 + P_m)^{\theta_{\tau_c,w,2}} (1 - P_m)}{1 - \theta_{\tau_c,w,1}}$	$\frac{1 - P_m}{1 - \theta_{C_e,w,1}}$

$w_{\xi,m}(\mathbf{x}_w, \boldsymbol{\theta}_{\xi,w})$	if $P_m \leq \theta_{\xi,w,1}$	0	0
	if $P_m > \theta_{\xi,w,1}$	$\frac{P_m - \theta_{\tau_c,w,1}}{1 - \theta_{\tau_c,w,1}}$	$\frac{P_m - \theta_{C_e,w,1}}{1 - \theta_{C_e,w,1}}$
$\Delta_{\xi}(\mathbf{x}_w, \boldsymbol{\theta}_{\xi,w})$	if $P_m \leq \theta_{\xi,w,1}$	0	$\ln \left[\frac{(1+P_m)^{-\theta_{C_e,w,2}}}{1-P_m} \right]$
	if $P_m > \theta_{\xi,w,1}$	0	$\frac{1-P_m}{1-\theta_{C_e,w,1}} \ln \left[\frac{(1+P_m)^{-\theta_{C_e,w,2}}}{1-P_m} \right]$

Table 2: Posterior Statistics of Parameters in the Critical Shear Stress Model, after the Bayesian update with the second set of data ($\mathbf{y}_{\tau_c,2}, \mathbf{x}_{k,2}, \mathbf{x}_{w,2}$)

	$\theta_{\tau_c,s,11}$	$\theta_{\tau_c,s,12}$	$\theta_{\tau_c,s,13}$	$\theta_{\tau_c,s,24}$	$\theta_{\tau_c,s,34}$	$\theta_{\tau_c,m,11}$	$\theta_{\tau_c,m,12}$	$\theta_{\tau_c,m,15}$	$\theta_{\tau_c,m,25}$	$\theta_{\tau_c,m,45}$	$\theta_{\tau_c,w,1}$	$\theta_{\tau_c,w,2}$	σ_{τ_c}
Mean	-43.208	7.314	5.203	41.133	-23.892	284.615	103.247	-269.508	-34.938	59.178	0.309	1.385	0.861
SD	2.833	0.944	0.694	2.639	1.486	17.756	9.859	15.650	3.370	3.889	0.000	0.055	0.068
Correlation Coefficients													
$\theta_{\tau_c,s,12}$	-0.59												
$\theta_{\tau_c,s,13}$	-0.30	-0.58											
$\theta_{\tau_c,s,24}$	-0.84	0.22	0.63										
$\theta_{\tau_c,s,34}$	0.86	-0.26	-0.59	-1.00									
$\theta_{\tau_c,m,11}$	0.12	-0.25	0.18	-0.02	0.03								
$\theta_{\tau_c,m,12}$	0.19	-0.03	-0.16	-0.22	0.22	0.14							
$\theta_{\tau_c,m,15}$	-0.20	0.19	-0.03	0.15	-0.15	-0.80	-0.71						
$\theta_{\tau_c,m,25}$	-0.20	0.02	0.17	0.22	-0.22	-0.14	-1.00	0.71					
$\theta_{\tau_c,m,45}$	0.22	-0.13	-0.07	-0.20	0.20	0.56	0.90	-0.95	-0.90				
$\theta_{\tau_c,w,1}$	0.08	-0.16	0.12	0.00	0.01	0.48	0.30	-0.52	-0.29	0.45			
$\theta_{\tau_c,w,2}$	-0.31	-0.01	0.30	0.37	-0.36	0.05	-0.14	0.05	0.15	-0.10	-0.04		
σ_{τ_c}	0.29	-0.01	-0.26	-0.33	0.32	-0.01	0.21	-0.13	-0.22	0.18	0.01	-0.47	

Table 3: Posterior statistics of parameters in the coefficient of erosion model, after the Bayesian update with the second set of data ($\mathbf{y}_{C_e,2}, \mathbf{x}_{k,2}, \mathbf{x}_{w,2}$)

	$\theta_{C_e,s,14}$	$\theta_{C_e,s,22}$	$\theta_{C_e,s,23}$	$\theta_{C_e,s,24}$	$\theta_{C_e,s,33}$	$\theta_{C_e,s,34}$	$\theta_{C_e,s,44}$	$\theta_{C_e,m,25}$	$\theta_{C_e,w,1}$	$\theta_{C_e,w,2}$	σ_{C_e}
Mean	-35.321	-36.796	55.873	29.196	-21.351	-18.905	-15.678	0.663	0.490	1.563	1.935
SD	-24.833	-32.135	48.718	23.966	-18.565	-15.791	-12.018	0.750	0.670	1.741	1.827
Correlation Coefficients											
$\theta_{C_e,s,22}$	0.39										
$\theta_{C_e,s,23}$	-0.41	-1.00									
$\theta_{C_e,s,24}$	-0.61	-0.48	0.55								
$\theta_{C_e,s,33}$	0.44	0.98	-0.99	-0.62							
$\theta_{C_e,s,34}$	0.51	0.40	-0.48	-0.94	0.57						
$\theta_{C_e,s,44}$	0.76	0.12	-0.15	-0.38	0.21	0.51					
$\theta_{C_e,m,25}$	-0.28	0.18	-0.22	-0.42	0.24	0.37	-0.48				
$\theta_{C_e,w,1}$	0.20	0.40	-0.44	-0.61	0.48	0.57	0.02	0.69			
$\theta_{C_e,w,2}$	-0.21	-0.88	0.91	0.68	-0.93	-0.64	0.00	-0.54	-0.62		
σ_{C_e}	-0.25	-0.03	0.05	0.18	-0.07	-0.16	-0.18	-0.01	-0.22	0.01	

Tab. 4: Posterior statistics of the parameters in the radius increase model.

	θ_p	θ_α	θ_β	θ_r	σ_{δ_r}
Posterior Mean	0.695	0.252	0.551	-11.627	0.289
Posterior SD	0.01	0.01	0.01	0.20	0.09
Correlation Coefficients					
θ_α	-0.71				
θ_β	-0.87	0.32			
θ_r	-0.86	0.51	0.81		
σ_{δ_r}	0.21	-0.36	-0.10	-0.23	

Tab. 5: Posterior statistics of the parameters after the global calibration using the totality of available data.

	$\theta_{\tau_{c,s},11}$	$\theta_{\tau_{c,s},12}$	$\theta_{\tau_{c,s},13}$	$\theta_{\tau_{c,s},24}$	$\theta_{\tau_{c,s},34}$	$\theta_{\tau_{c,m},11}$	$\theta_{\tau_{c,m},12}$	$\theta_{\tau_{c,m},15}$	$\theta_{\tau_{c,m},25}$	$\theta_{\tau_{c,m},45}$	$\theta_{\tau_{c,w},1}$	$\theta_{\tau_{c,w},2}$	σ_{τ_c}	$\theta_{C_{\rho,s},14}$	$\theta_{C_{\rho,s},22}$	$\theta_{C_{\rho,s},23}$	$\theta_{C_{\rho,s},24}$	$\theta_{C_{\rho,s},33}$	$\theta_{C_{\rho,s},34}$	$\theta_{C_{\rho,s},44}$	$\theta_{C_{\rho,m},25}$	$\theta_{C_{\rho,w},1}$	$\theta_{C_{\rho,w},2}$	σ_{C_ρ}	θ_p	θ_α	θ_β	θ_r	σ_{δ_r}				
Mean	-43.093	7.209	5.246	41.046	-23.843	275.453	98.384	-258.930	-33.285	56.642	0.309	1.397	0.840	-37.848	-38.728	58.906	31.043	-22.556	-20.211	-16.793	0.659	0.488	1.578	1.920	0.670	0.243	0.521	-11.169	0.302				
SD	2.813	0.761	0.541	2.391	1.351	21.422	13.605	23.460	4.605	5.873	0.000	0.027	0.046	3.027	1.850	2.738	1.611	1.030	1.206	1.625	0.047	0.047	0.017	0.136	0.043	0.053	0.109	1.631	0.103				
Correlation coefficients																																	
$\theta_{\tau_{c,s},12}$	-0.66																																
$\theta_{\tau_{c,s},13}$	-0.47	-0.35																															
$\theta_{\tau_{c,s},24}$	-0.85	0.51	0.55																														
$\theta_{\tau_{c,s},34}$	0.85	-0.54	-0.51	-1.00																													
$\theta_{\tau_{c,m},11}$	0.00	0.06	-0.03	0.07	-0.07																												
$\theta_{\tau_{c,m},12}$	0.15	0.05	-0.19	0.03	-0.04	0.55																											
$\theta_{\tau_{c,m},15}$	-0.08	-0.06	0.12	-0.06	0.07	-0.89	-0.87																										
$\theta_{\tau_{c,m},25}$	-0.15	-0.05	0.18	-0.03	0.04	-0.55	-1.00	0.87																									
$\theta_{\tau_{c,m},45}$	0.11	0.06	-0.15	0.05	-0.06	0.78	0.95	-0.98	-0.95																								
$\theta_{\tau_{c,w},1}$	0.05	0.03	-0.04	0.10	-0.10	0.72	0.68	-0.79	-0.67	0.77																							
$\theta_{\tau_{c,w},2}$	0.14	-0.16	-0.01	-0.18	0.18	-0.18	-0.22	0.23	0.23	-0.24	-0.17																						
σ_{τ_c}	0.09	0.08	-0.17	0.00	-0.01	0.03	0.22	-0.14	-0.21	0.17	0.17	-0.04																					
$\theta_{C_{\rho,s},14}$	-0.25	0.12	0.16	0.20	-0.20	0.04	-0.06	0.01	0.06	-0.03	0.03	-0.12	-0.07																				
$\theta_{C_{\rho,s},22}$	-0.06	0.03	0.05	0.08	-0.08	0.21	0.12	-0.19	-0.12	0.17	0.23	-0.12	-0.05	0.47																			
$\theta_{C_{\rho,s},23}$	0.08	-0.03	-0.07	-0.09	0.09	-0.22	-0.11	0.19	0.11	-0.17	-0.23	0.13	0.06	-0.50	-1.00																		
$\theta_{C_{\rho,s},24}$	0.14	-0.08	-0.09	-0.15	0.15	-0.23	-0.09	0.18	0.08	-0.15	-0.12	0.12	0.05	-0.72	-0.53	0.59																	
$\theta_{C_{\rho,s},33}$	-0.11	0.04	0.09	0.11	-0.11	0.23	0.10	-0.19	-0.10	0.16	0.22	-0.12	-0.07	0.55	0.98	-0.99	-0.65																
$\theta_{C_{\rho,s},34}$	-0.15	0.06	0.13	0.15	-0.15	0.24	0.01	-0.15	-0.01	0.10	0.08	-0.07	-0.09	0.63	0.41	-0.49	-0.94	0.58															
$\theta_{C_{\rho,s},44}$	-0.22	0.07	0.19	0.17	-0.16	0.03	-0.18	0.08	0.18	-0.12	-0.03	-0.01	-0.13	0.80	0.22	-0.28	-0.58	0.36	0.70														
$\theta_{C_{\rho,m},25}$	0.09	-0.04	-0.06	-0.06	0.05	0.03	0.09	-0.07	-0.09	0.08	-0.07	0.00	-0.01	-0.12	0.01	-0.03	-0.22	0.04	0.18	-0.29													
$\theta_{C_{\rho,w},1}$	-0.14	-0.09	0.22	0.00	0.01	-0.21	-0.31	0.30	0.31	-0.31	-0.24	0.09	-0.13	0.49	0.25	-0.28	-0.31	0.31	0.28	0.33	0.30												
$\theta_{C_{\rho,w},2}$	0.06	-0.01	-0.08	-0.10	0.09	-0.30	-0.11	0.23	0.10	-0.19	-0.22	0.09	0.06	-0.36	-0.88	0.91	0.68	-0.93	-0.65	-0.23	-0.23	-0.24											
σ_{C_ρ}	0.12	0.07	-0.25	-0.12	0.11	0.00	0.21	-0.11	-0.21	0.15	0.08	-0.12	0.21	-0.11	0.10	-0.08	0.03	0.04	-0.13	-0.23	-0.02	-0.23	-0.04										
θ_p	0.05	0.06	-0.06	0.14	-0.14	0.41	0.40	-0.47	-0.40	0.46	0.36	-0.10	0.13	-0.06	0.06	-0.07	-0.18	0.08	0.21	0.01	-0.05	-0.43	-0.19	-0.08									
θ_α	0.01	0.04	-0.06	0.04	-0.04	-0.11	0.09	0.02	-0.10	0.03	0.01	-0.21	0.11	0.20	0.19	-0.18	-0.10	0.16	-0.01	0.01	0.03	0.16	-0.07	0.34	-0.24								
θ_β	0.03	0.07	-0.11	0.02	-0.03	-0.04	0.05	0.00	-0.05	0.02	-0.01	-0.17	0.09	0.15	0.16	-0.16	-0.12	0.15	0.04	0.02	-0.01	0.05	-0.10	0.37	-0.18	0.62							
θ_r	-0.17	-0.02	0.17	-0.03	0.04	-0.27	-0.36	0.36	0.36	-0.38	-0.28	0.15	-0.20	0.30	0.05	-0.05	-0.06	0.05	0.01	0.19	-0.01	0.43	0.07	-0.20	-0.69	-0.31	-0.38						
σ_{δ_r}	0.21	0.00	-0.28	-0.21	0.20	-0.11	0.18	-0.03	-0.19	0.09	0.01	-0.03	0.18	-0.01	0.07	-0.04	0.04	0.01	-0.16	-0.18	0.07	0.02	0.06	0.46	-0.21	0.31	0.32	-0.02					

Table 6: Random variables for the soil of the example dam.

Parameter	Mean	C.o.V.	Distribution Model
$d_{s,50} [mm]$	0.35	0.20	Lognormal
$d_{s,90} [mm]$	1.10	0.20	Lognormal
Sand Content	0.75	0.20	Beta [0, 1]
OWC	0.15	0.30	Beta [0, 1]
Max Dry Density [Mg/m ³]	1.90	0.15	Lognormal
POC	0.95	0.15	Beta [0, 1]
ω	0.15	0.20	Beta [0, 1]
DOS	0.65	0.20	Beta [0, 1]
Mud Content	0.20	0.20	Beta [0, 1]

Table 7: Probability of not detection and not intervention according to Fell et al. [22], for the example dam in case of monthly inspections and/or monitoring and no public nearby.

Interval of time from Concentrated leak initiation to breach mechanism $[t, t + \Delta t]$	Probability of not detection $P(\bar{E}^{DT} E^I E^C E^P E^{OL} \Delta t)$	Probability of not intervention $P(E^{ul} E^I E^C E^P E^{OL} E^{DT} \Delta t)$	$P(t, \Delta t) = \tilde{F}_{\delta_r}(t) - \tilde{F}_{\delta_r}(t + \Delta t)$
< 3 hours	0.999	0.99	$1 - \tilde{F}_{\delta_r}(3 \text{ hours}) = 0.013$
3 hours to 12 hours	0.99	0.95	$\tilde{F}_{\delta_r}(3 \text{ hours}) - \tilde{F}_{\delta_r}(12 \text{ hours}) = 0.219$
12 hours to 24 hours	0.95	0.9	$\tilde{F}_{\delta_r}(12 \text{ hours}) - \tilde{F}_{\delta_r}(1 \text{ day}) = 0.154$
1 day to 2 days	0.9	0.8	$\tilde{F}_{\delta_r}(1 \text{ day}) - \tilde{F}_{\delta_r}(2 \text{ days}) = 0.050$
2 days to 7 days	0.6	0.7	$\tilde{F}_{\delta_r}(1 \text{ week}) - \tilde{F}_{\delta_r}(2 \text{ days}) = 0.009$
> 1 week	0.1	0.6	$\tilde{F}_{\delta_r}(1 \text{ week}) = 0.445$

Figures

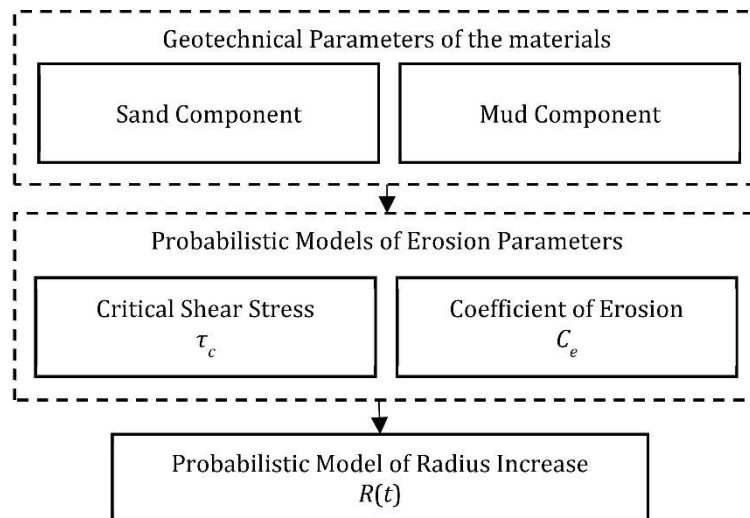


Fig. 1: Concepts for formulation of the probabilistic model for erosion parameters and pipe enlargement

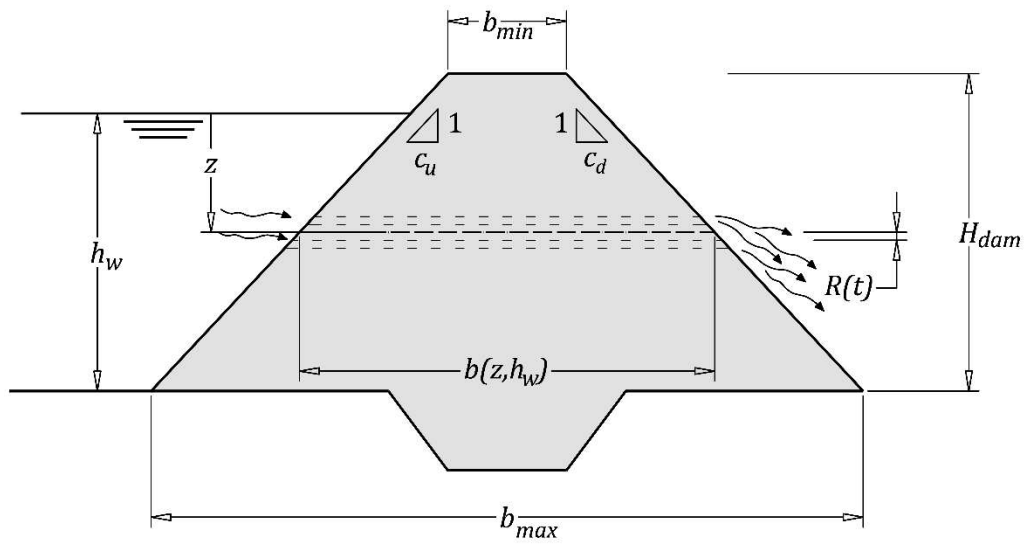


Fig. 2: Scheme of the example dam.

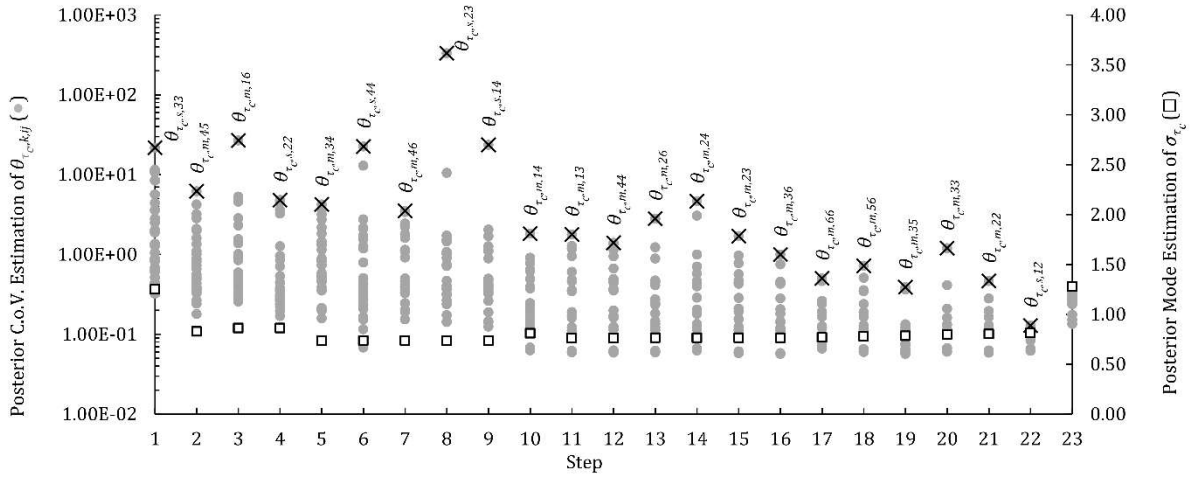


Fig. 3: Stepwise deletion process of the critical shear stress model

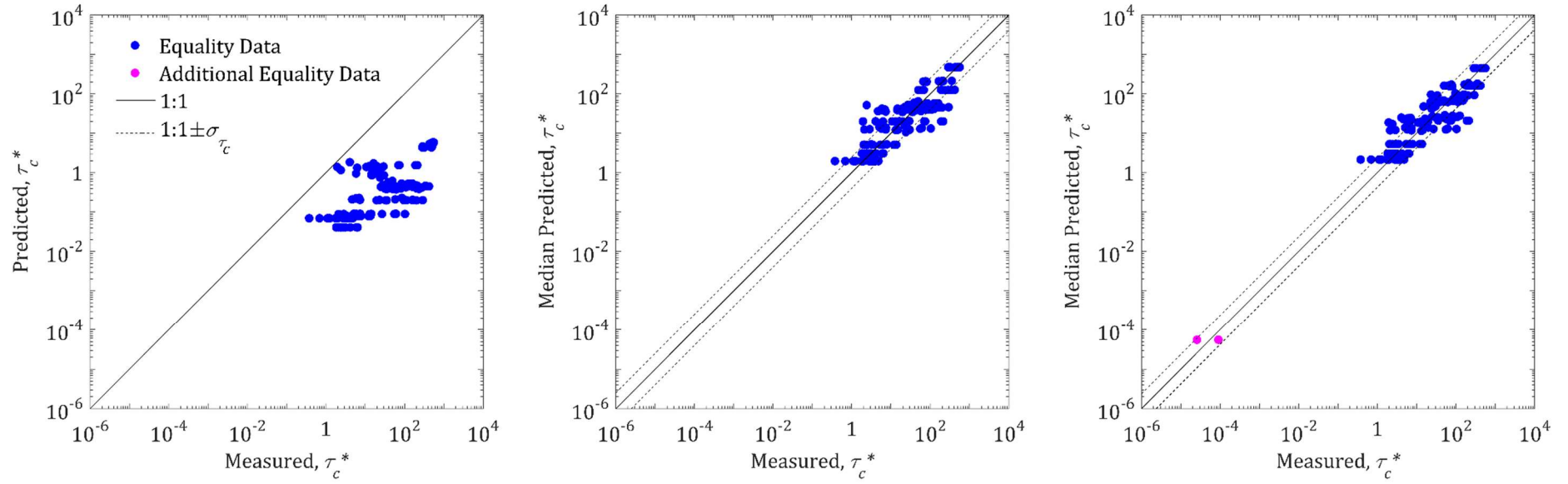


Fig. 4: Comparison between measured and predicted the critical shear stress based on deterministic (left), probabilistic model by Andreini et al. [11] (center) and the updated one with data from two additional JETs [30–32] (right).

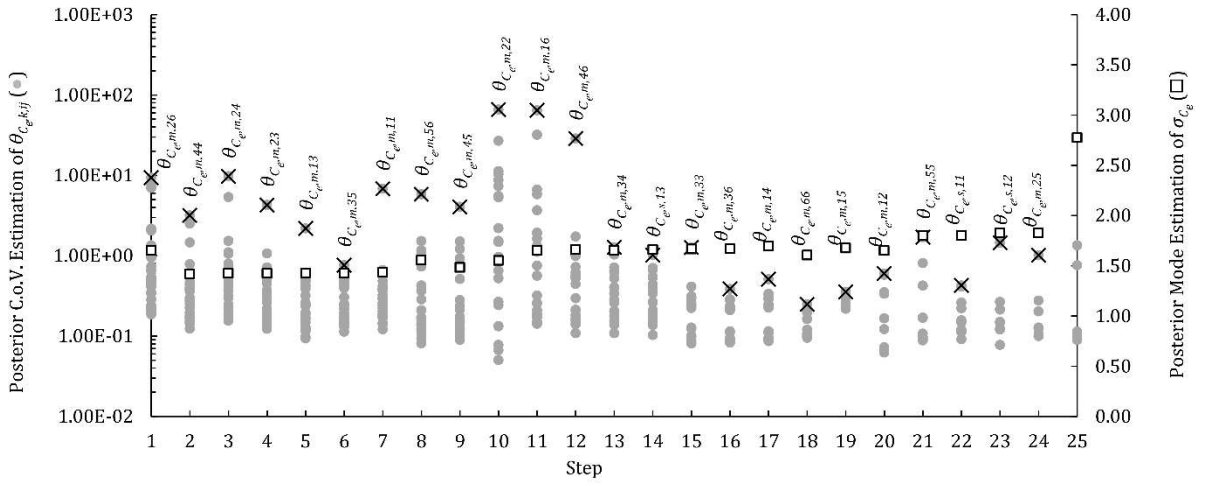


Fig. 5: Stepwise deletion process of the coefficient of erosion model

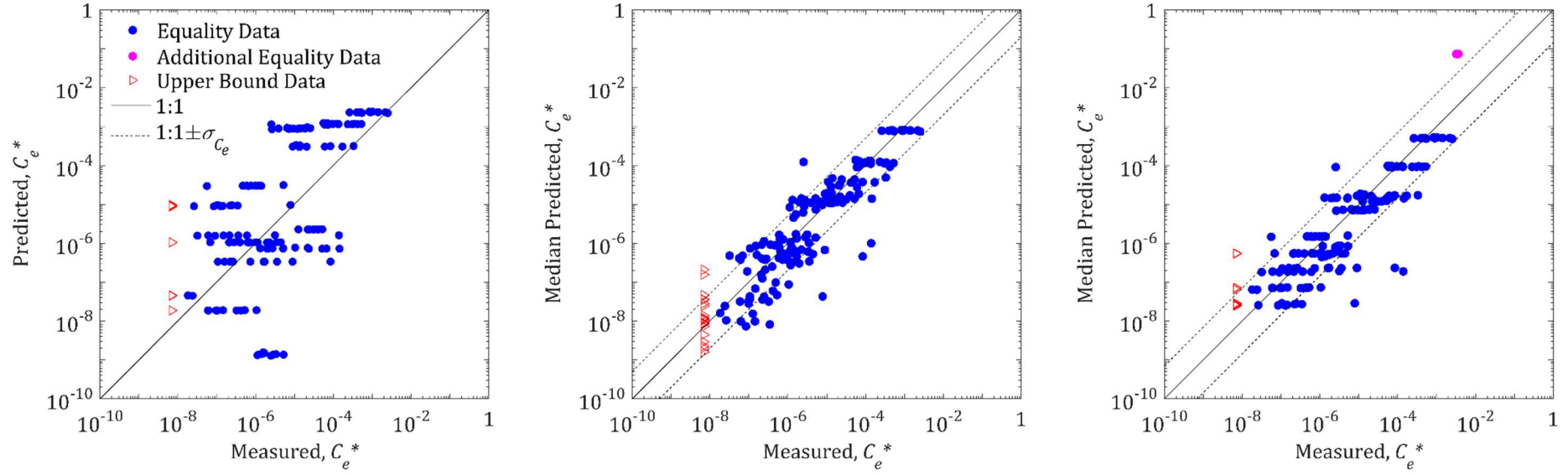


Fig. 6: Comparison between measured and predicted the coefficient of erosion based on deterministic (left), probabilistic model by Andreini et al. [11] (center) and the updated one with data from two additional JETs [30–32] (right).

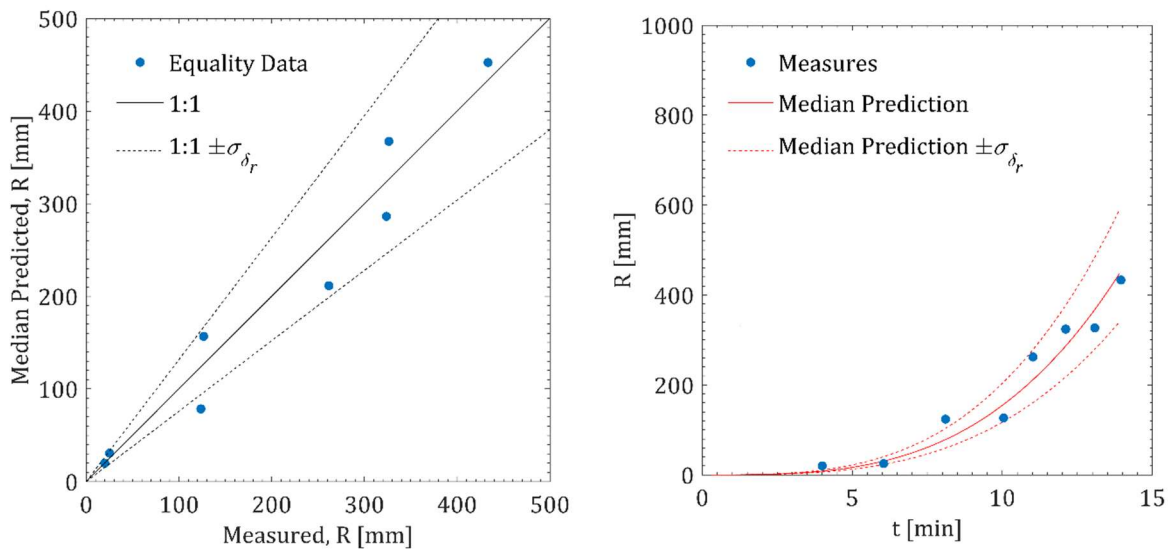


Fig. 7: Comparison between measured values of the pipe radius and predicted ones by means of the calibrated probabilistic model.

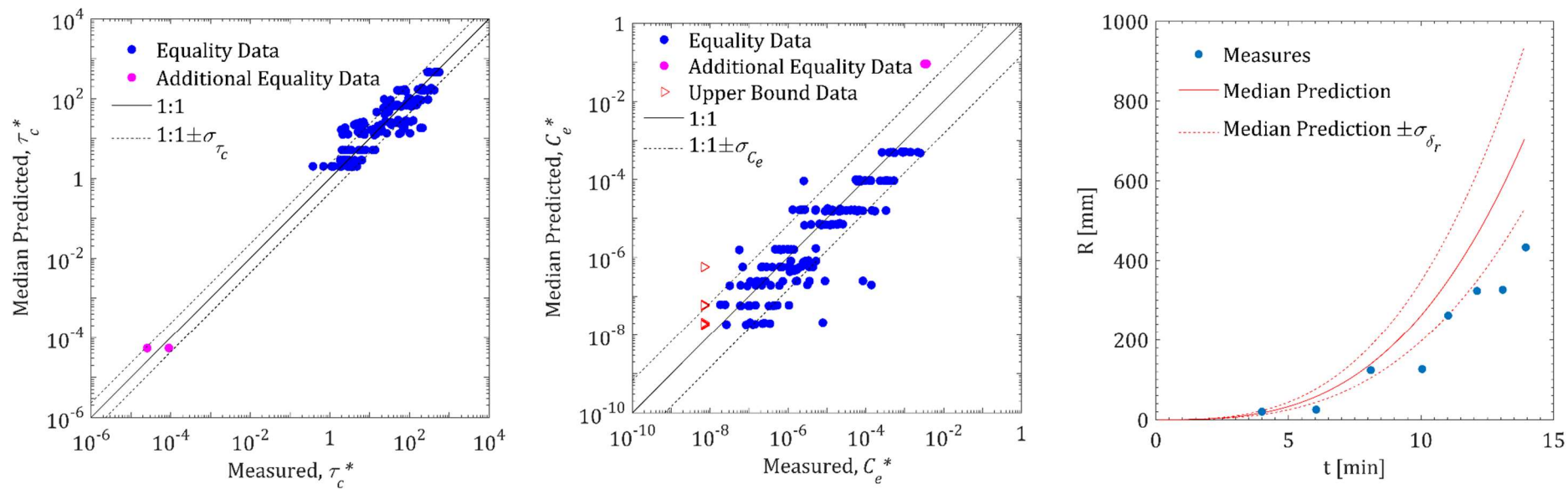


Fig. 8: Comparison between measured and predicted critical shear stress (left), coefficient of erosion (center) and pipe radius increase versus time (right) after the global parameter calibration using the totality of available data.

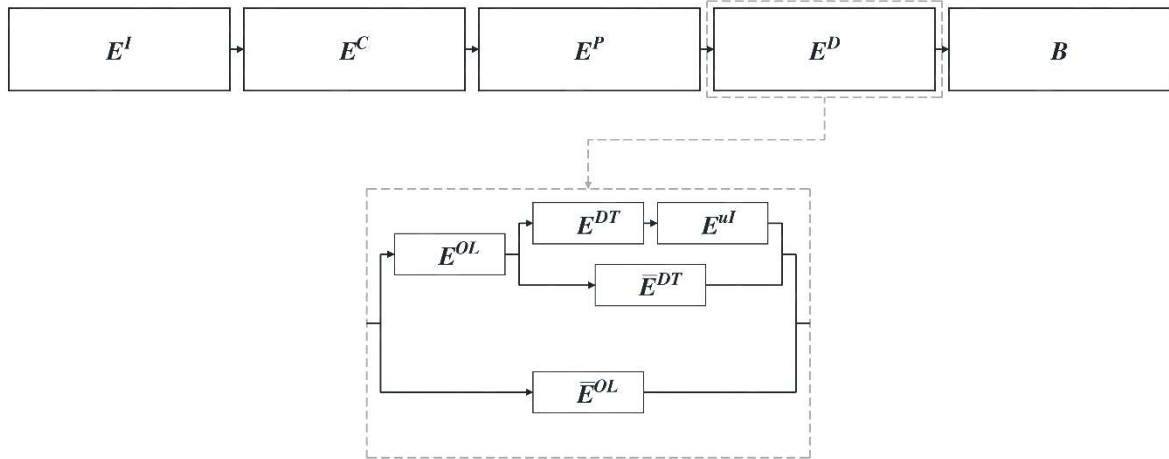


Fig. 9: Scheme of the concentrated leak erosion process: series system concerning the initiation, the continuation of erosion, the progression to form a pipe, no detection, intervention, and repair, and the initiation of the breach mechanism.

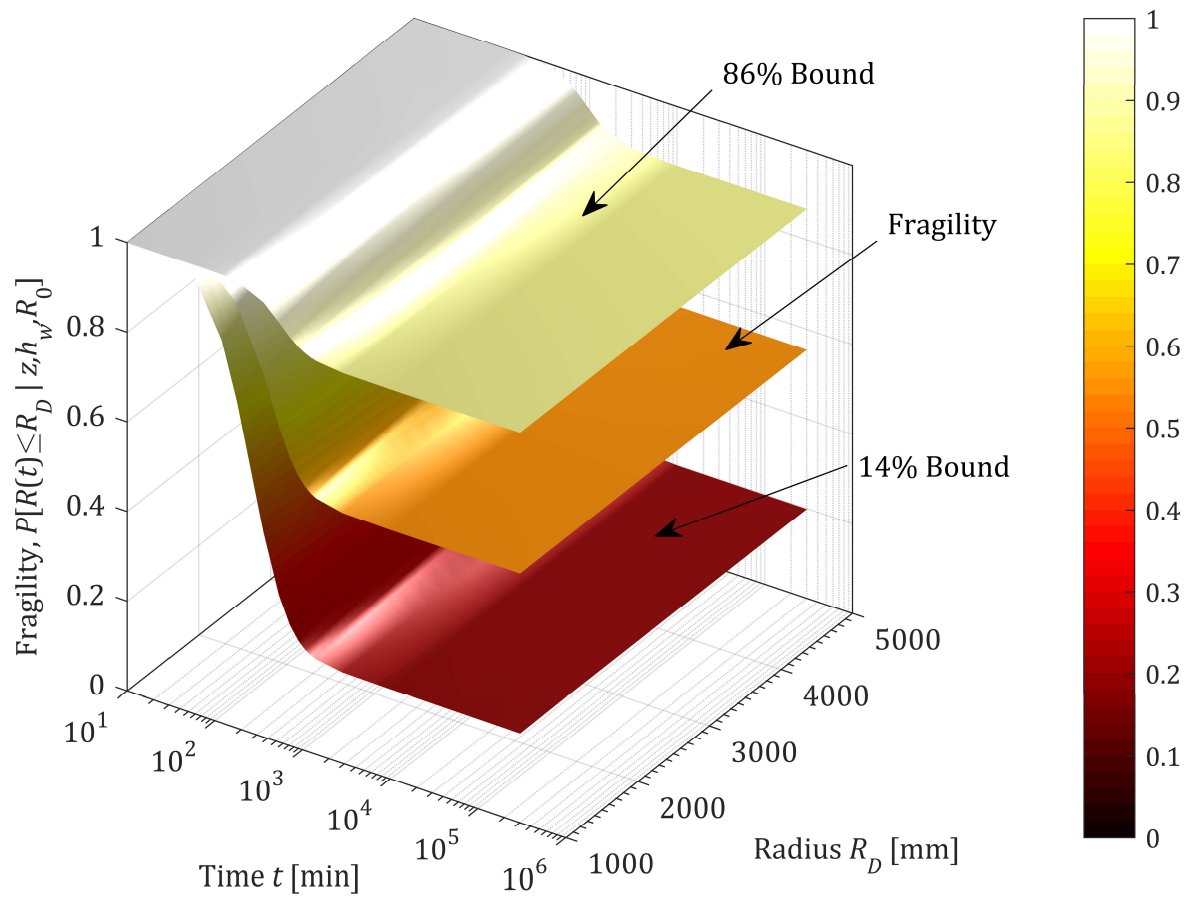


Fig. 10: Fragility estimate for the pipe radius at a certain time from the initiation of the concentrated leak erosion mechanism.

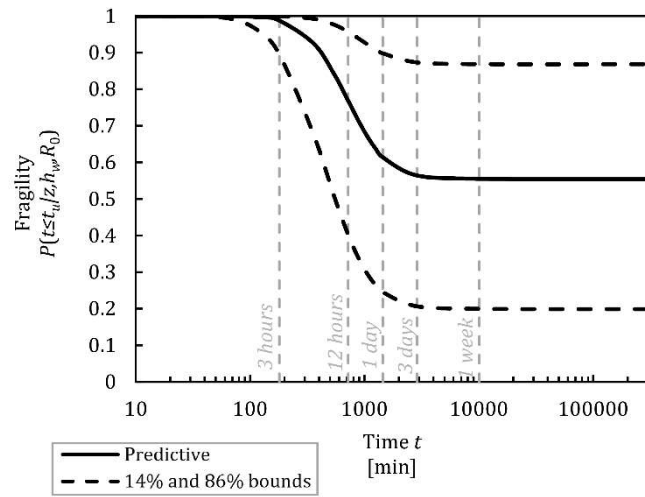


Fig. 11: Fragility estimate of the time to failure for the example dam.