Development of Probabilistic Dam Breach Model Using Bayesian Inference

S. Peter, A. Siviglia, J. Nagel, S. Marelli, R. M Boes, D. Vetsch, B. Sudret

Development of probabilistic dam breach model using Bayesian inference

S. J. Peter, A. Siviglia, J. Nagel, S. Marelli, R. M. Boes, D. Vetsch, and B. Sudret

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Data Sheet

Development of probabilistic dam breach model using Bayesian inference

S. J. Peter, A. Siviglia, J. Nagel, S. Marelli, R. M. Boes, D. Vetsch, and B. Sudret

June 9, 2018

Abstract

Dam breach models are commonly used to predict outflow hydrographs of potentially failing dams and are key ingredients for evaluating flood risk. In this paper a new dam breach modeling framework is introduced that shall improve the reliability of hydrograph predictions of homogeneous earthen embankment dams. Striving for a small number of parameters, the simplified physics-based model describes the processes of failing embankment dams by breach enlargement, driven by progressive surface erosion. Therein the erosion rate of dam material is modeled by empirical sediment transport formulations. Embedding the model into a Bayesian multilevel framework allows for quantitative analysis of different categories of uncertainties. To this end, data available in literature of observed peak discharge and final breach width of historical dam failures was used to perform model inversion by applying Markov Chain Monte Carlo simulation. Prior knowledge is mainly based on non-informative distribution functions. The resulting posterior distribution shows that the main source of uncertainty is a correlated subset of parameters, consisting of the residual error term and the epistemic term quantifying the breach erosion rate. The prediction intervals of peak discharge and final breach width are congruent with values known from literature. To finally predict the outflow hydrograph for real case applications, an alternative residual model was formulated that assumes perfect data and a perfect model. The fully probabilistic fashion of hydrograph prediction has the potential to improve the adequate risk management of downstream flooding.

1 Introduction

Earthen dams have been built by humans to store water for multiple purposes for millennia (*Schnitter*, 1994). While they are regarded as safe structures, history taught us that they nevertheless may fail. Extensive literature is available reporting historic dam failures (e.g. *Broich*, 1996). To protect people and infrastructure downstream of a potentially failing

1

dam, today's supervising authorities and dam operators take precautionary measures, install emergency warning systems, and set up evacuation plans. To do so, reliable predictions about the amount and timing of released water in case of a dam failure are needed. This information comes from dam break models which provide outflow hydrographs for subsequent flood routing models.

In this paper the focus is on homogeneous and non-cohesive earthfill embankment dams. Processes that belong to the breach formation are of interest here, whereas any analysis concerning probability of failure, breach initiation, or flood propagation downstream is not considered. Herein the main breach formation process is regarded as the enlargement of an initial breach due to slope erosion of dam material leading to increasing breach discharge. Research has put much effort into understanding the complex natural phenomena of breach formation processes, both by field tests (e.g. *Morris and Hassan*, 2005) and laboratory experiments (e.g. *Schmocker and Hager*, 2012; *Frank*, 2016). The gained insights help understand the breach formation process and provide relevant information for model development. However, their application to real-life breach situations is often limited due to simplifications and scale effects (*ASCE/EWRI*, 2011). Regardless of the complexity and high non-linearity of dam breach processes, the necessity of dam breach prediction tools induced the development of a multitude of numerical models over the last decades. The physical properties usually investigated are the outflow hydrograph and the size of the final breach. Dam breach modeling techniques vary strongly in complexity and can be classified according to their level of detail:

1. *Statistically-based* models are purely data-driven and rely on regression analysis of historical dam breach events and make estimations about embankment breach characteristics, e.g. peak discharge, time to failure, or final breach width. The dam breach model thereby can be described either by power functions of governing dam-reservoir quantities (*Froehlich*, 2008) or by applying methods of artificial intelligence (*Amini et al.*, 2011; *Hooshyaripor and Tahershamsi*, 2012). Since no physics are considered in these models, their reliability and explanatory power is generally low.

2. *Simplified physics-based* models take into account the description of selected physical processes, e.g. draw-down of the reservoir and enlargement of the breach over time due to erosion of dam material. Because of their simplicity, many a-priori assumptions have to be made, e.g. initial breach size, breach geometry, flow over the breach, sediment transport, and geomechanical concepts. The computational cost of these models is low and the number of parameters that have to be defined is commonly small, hence allowing for real time applications (*Ma and Fu*, 2012). A comprehensive list of simplified dam breach models can be found in *ASCE/EWRI* (2011).

3. *Detailed physics-based* models describe the embankment breaching process in one, two, or even three dimensions by applying sophisticated numerical approaches. During

the last decades these models became popular due to advances in computer sciences (e.g. *Broich*, 1996; *Wang and Bowles*, 2006; *Faeh*, 2007; *Wu et al.*, 2009; *Volz*, 2013). Nevertheless, describing and parameterizing not only the macroscopic but also the microscopic phenomena of a gradually failing dam is very challenging. In particular, the interaction between water and dam material inside the breach, the short term stability of saturated and compacted soils, and the quantification of high-concentration sediment transport capacity pose difficulties.

Before dam breach models are applied as prediction tools, no matter what type of model, they are often calibrated to a data set. Historical dam breach data are therefore used to feed the models, but data on real-life embankment failures are usually poorly documented for various reasons (*ASCE/EWRI*, 2011). Uncertainties on breach properties, such as average breach width and failure time, have been reported in *Froehlich* (2008), and their relation to outflow hydrograph are investigated in *Ahmadisharaf et al.* (2016). Uncertainties of dam breach models are often quantified in terms of prediction errors, which are minimized during model calibration by selecting appropriate data and fitting procedures. To the best of our knowledge, full quantification of uncertainties in physical dam and reservoir properties and output variables has not been explored in dam breach modeling. In simplified physics-based models systematic sensitivity analysis has been carried out (*Fread*, 1984; *Walder and O'Connor*, 1997; *De Lorenzo and Macchione*, 2014), unlike in detailed models where the computational cost is too high. Surrogate modeling techniques, also known as meta-modeling or response surface modeling, could offer an alternative to this end, such as Kriging (*Santner et al.*, 2003) and polynomial chaos expansion (*Xiu and Karniadakis*, 2002). *Wahl* (2004) assessed the prediction errors of various statistically-based embankment breach models by applying them to a set of 108 dam failures. The resulting prediction intervals are approximately $\pm 1/3$, and $\pm 1/2$ to 1 orders of magnitude for predicted breach width and peak outflow, respectively. In 2013 the International Committee of Large Dams (ICOLD) organized a numerical benchmark where the participants were invited to predict the breach outflow hydrograph of a hypothetical dam failure, i.e. the true outcome was not known (*Zenz and Goldgruber*, 2013). The results demonstrate that the variations between the predicted hydrographs of different dam breach models/modelers are large and can significantly influence the result of hydrodynamic calculations (*Escuder-Bueno et al.*, 2016). The origin of the discrepancies in the resulting hydrographs of the benchmark participants can be mainly attributed to variability of the erodibility of embankment material as a result of different soil types, compaction effort, and water content (*Morris et al.*, 2008). Ultimately, on the basis of a single model the wide range of possible breach outflow hydrographs cannot be quantified reliably enough for prediction purposes.

On these grounds many researchers emphasized the need for a systematic quantification of all types of uncertainties in dam breach models to reliably predict embankment dam

failure processes (*Wahl*, 2004; *Froehlich*, 2008; *ASCE/EWRI*, 2011) and consequently to incorporate into comprehensive risk management systems (*Altinakar et al.*, 2009). This has been achieved e.g. in hydrological sciences by first promoting the use of uncertainty estimation as routine (*Pappenberger and Beven*, 2006), setting up an integrated risk management framework (*Büchele et al.*, 2006), critically discussing the deterministic and probabilistic approaches (*Di Baldassarre et al.*, 2010), and proposing new methodologies within the probabilistic framework (*Alfonso et al.*, 2016).

Generally, uncertainties are categorized as (i) aleatory, that describe the natural variability of a physical process, (ii) epistemic, that describe the lack of knowledge in parametric description of the process, and (iii) global model inadequacy and data uncertainty (*Kennedy and O'Hagan*, 2001). Based on this classification *Nagel and Sudret* (2016) proposed a unified framework to quantify uncertainties on all levels by using Bayesian inverse modeling and a deterministic model as backbone. Related frameworks were applied in different fields, such as 3-D environment modeling (*Balakrishnan et al.*, 2003), sediment entrainment modeling (*Wu and Chen*, 2009), groundwater modeling (*Laloy and Vrugt*, 2012; *Shi et al.*, 2014), rating curve derivation (*Mansanarez et al.*, 2016), or design flood estimation (*Steinbakk et al.*, 2016).

The goal of the present paper is to enhance the reliability of hydrograph predictions by developing a new dam break model in a fully probabilistic manner. The model is herein regarded as a strong approximation of an open system. Accordingly the verification and validation of the truth of such a model is nearly impossible (*Oreskes et al.*, 1994). However, a model can be conditionally confirmed, similar to the original idea of Rev. Thomas Bayes: for given evidence in certain circumstances, find the model, out of a set of feasible models, that explains the evidence best. The model predictions will supposedly be more reliable (i) the more evidence is consulted to conditionally confirm the model, (ii) the higher the evidence quality is, (iii) the more widely accepted physical knowledge is implemented in the model, (iv) the more sophisticated methods are applied to find the best model.

In this vein, a new deterministic dam breach model is proposed that is embedded into a Bayesian multilevel framework (*Nagel and Sudret*, 2016). Therein the vector of model parameters is split into different parameter types according to their uncertainty character. The underlying parameter distributions are defined by Bayesian inference, in which the prior information is specified by empirical knowledge and the reference data is based on measured quantities of dam failures. The underlying population of the data is represented by worldwide and historically failed, man-made, homogeneous embankment dams. The novelty comprises a complete quantification of parametric as well as residual uncertainties in dam breach modeling. The latter consists of both data and structural model errors. The outcome of the present study is not only the final probabilistic dam breach model, but instead providing a framework that allows the integration of further data. Since the

underlying deterministic model is treated as a black-box, it can be replaced in a simple manner for future analysis. As stated by *Morris et al.* (2008), it is not easy to improve the accuracy of dam breach models, that is the degree to which model predictions represent a real dam failure. Nevertheless, by incorporating the information about model uncertainties these predictions become more reliable, i.e. they are more trustworthy in case of not knowing the true outcome of a possible dam failure. The enhanced reliability of predicted breach hydrographs is of major importance regarding the quality of risk quantification induced by dam failures.

This paper is organized as follows: The deterministic dam breach model formulation and its parameters is introduced in Section 2. The outline of the Bayesian multilevel approach is given in Section 3, including successfully embedding the dam breach model into the probabilistic framework. The resulting quantification of uncertainties of all levels is shown in Section 4. Finally critical points are discussed within an examplary model application in Section 5and conclusions are drawn in Section 6.

2 Dam Breach Model

The aim of this section is to present the development of the deterministic dam breach model, which is the backbone of the probabilistic framework proposed in this paper. The model under investigation can be attributed to simplified physics-based dam breach models. Because of prevailing lack of data in dam breach modeling, the number of parameters is kept as small as possible to circumvent the risk of too many degrees of freedom. In addition it is of paramount importance to have a computationally efficient model due to its probabilistic inversion where million evaluations are needed. At the same time the model must be able to reproduce the main dam breach phenomenon. The physical basis of the breach formation process is the water-sediment interaction: the discharging water is the driving force of breach erosion, and vice versa the breach enlargement controls the rate of discharge (e.g. *Singh and Scarlatos*, 1988).

2.1 Formulation of the Physical Processes

Processes that lead to an initial breach are not considered here. In reality, the initial breach is being formed by a multitude of different failure causes, e.g. overtopping, internal erosion, slope instabilities, or foundation problems (*Foster et al.*, 2000). This breach initiation process is much slower than the subsequent breach enlargement and varies strongly between different failure causes. Regardless, discharge rates are still low and do not influence the hydrograph of a failing dam (*Wahl*, 2004; *Morris et al.*, 2008). After developing a sufficiently large breach, the flow rate increases rapidly and the failure cannot be prevented.

This breach formation process is essentially dominated by the mechanics of overtopping
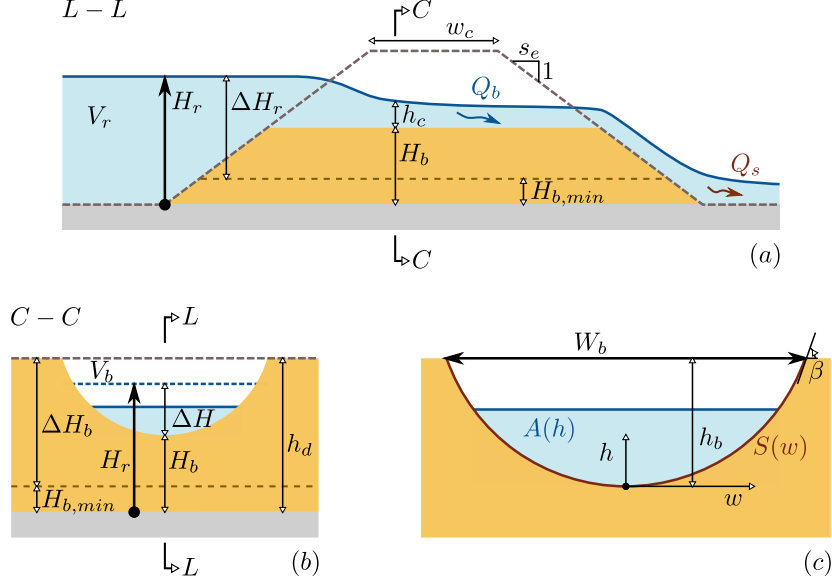
Figure 1: Longitudinal (a) section $L - L$ and transverse (b) control section $C - C$ of the dam-reservoir system, and (c) breach shape definition. Thick arrows indicate the unknown variables H_r and W_b in Eq. (1).

(*Singh*, 1996; *ASCE/EWRI*, 2011; *De Lorenzo and Macchione*, 2014)) and according to literature the main physical processes involved are (i) gradual erosion of dam material, (ii) breach enlargement, (iii) increasing breach outflow, and (iv) decreasing water level in the reservoir. These four processes can generally be formulated by a system of two ordinary differential equations (ODE) (see notation in Figure 1)

$$\frac{\mathrm{d}H_r}{\mathrm{d}t} = -Q_b \left(\frac{\mathrm{d}V_r}{\mathrm{d}H_r} \right)^{-1} \tag{1a}$$

$$\frac{\mathrm{d}W_b}{\mathrm{d}t} = Q_s \left(\frac{\mathrm{d}V_b}{\mathrm{d}W_b} \right)^{-1}. \tag{1b}$$

Eq. (1a) represents a continuity equation for conserving the water volume stored in the reservoir $V_r \left[\mathrm{m}^3 \right]$, where the depletion of the reservoir level $H_r \left[\mathrm{m} \right]$ over time $t \left[\mathrm{s} \right]$ is described by the discharge of breach outflow $Q_b \left[\mathrm{m}^3 \, \mathrm{s}^{-1} \right]$ and the rate of change of reservoir volume with respect to reservoir level $\frac{\mathrm{d}V_r}{\mathrm{d}H_r} \left[\mathrm{m}^2 \right]$. The continuity equation Eq. (1b) is conserving the volume of eroded dam material $V_b \left[\mathrm{m}^3 \right]$, where the increasing breach width $W_b \left[\mathrm{m} \right]$ over time is controlled by the discharge of dam material that is transported out of the breach $Q_s \left[\mathrm{m}^3 \, \mathrm{s}^{-1} \right]$ and the rate of change of breach volume with respect to breach width $\frac{\mathrm{d}V_b}{\mathrm{d}W_b} \left[\mathrm{m}^2 \right]$. Similar formulations have been adopted by *Singh and Scarlatos* (1988) and *Macchione* (2008). In the model presented here, processes that belong to the mechanism of progressive surface erosion are considered only, whereas the mechanisms of head cutting (cohesive dam material) and interlocking (coarse dam material) are neglected. Dams showing cohesive or rockfill materials are clearly outside of the application range of this model.
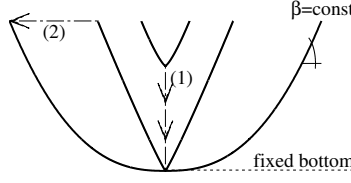
Figure 2: Breach shape development with constant breach side angle β at the top and fixed bottom ($H_{b,min}$): stage of vertical erosion (1), i.e. from initial breach to fixed bottom ($H_b > H_{b,min}$); and stage of lateral widening (2), i.e. no further deepening of the breach ($H_b = H_{b,min}$). In case of low erodibility the stage of vertical erosion will never be exceeded, what can be referred to as partial failure.

As depicted in Figure 2, the breach enlargement is split into two distinct breach development states in time. In a first stage the breach development is dominated by vertical erosion. The breach width W_b is enlarged and breach bottom level H_b [m] is lowered at the same time with constant rate, preserving self-similarity of the breach shape over time (*Pickert et al.*, 2011; *Frank*, 2016). When the breach bottom reaches the foundation of the dam $H_{b,min}$ [m] it is assumed that no further deepening of the breach is possible and lateral widening is controlling the breach enlargement only (e.g. *Coleman et al.*, 2002; *Chinnarasri et al.*, 2004). Thus the development of the breach bottom is described by the following simple ODE:

$$\frac{\mathrm{d}H_b}{\mathrm{d}W_b} = \begin{cases} -\frac{h_b}{W_b} & \text{if} \quad H_b > H_{b,min}, \\ \\ 0 & \text{if} \quad H_b = H_{b,min}, \end{cases} \tag{2}$$

where $h_b = h_d - H_b$ [m] is the breach height and h_d [m] is the dam height (see Figure 1).

Despite reducing the complex physical processes of breach formation into simple continuity equations (1a) and (1b), it has been shown that the physical processes contained in this are sufficient to reproduce the hydrograph of rather detailed models (*Vonwiller et al.*, 2015) or laboratory experiments (*Chinnarasri and Saelim*, 2009). Below further details are given about assumptions that are made in this study to parameterize and quantify the variables of the right-hand side in the system of equations (1).

2.2 Breach Geometry

In this section the simplifications of the breach geometry description are stated. The final goal is to quantify the breach volume change rate $\frac{\mathrm{d}V_b}{\mathrm{d}W_b}$ (see Eq. (1b)), which is dependent on geometrical parameters only.

The longitudinal breach shape is reminiscent of an hourglass shape and the hydraulic control section is defined by the curved weir crest at the inlet of the breach (*Walder et al.*, 2015). In order to lower the complexity of the model and to reduce the number of model parameters, the three dimensional breach geometry is simplified as a prismatic channel. Its

cross-section is regarded as the hydraulic control section (see Eq. (10)). For different existing parameter models the description of the cross-sectional shape is varying, i.e. triangular, trapezoidal, rectangular, or parabolic (*ASCE/EWRI*, 2011). To include all types of shapes here, the breach area is described as a power law of h, i.e. $A \propto h^k$ (see Figure 1c), where k is a shape parameter ranging from $k = 1$ representing a rectangular breach shape to $k = 2$ representing a triangular shape. Strictly speaking, the model requires $k < 1$ and the breach shape converges to a rectangular shape for $k \to 1$. A similar approach describing breach shapes of instantaneous dam breaks is followed by *Pilotti et al.* (2010). For given top width W_b and bottom level H_b the breach side wall is obtained as

$$S\left(w\right) = h_b \left(\frac{2|w|}{W_b}\right)^{\frac{1}{k-1}},\tag{3}$$

where $w \in \left[-\frac{W_b}{2}, \frac{W_b}{2}\right]$ is a control variable running across the breach section, starting at the breach center (see Figure 1c). The length of the breach wall on one side in the interval $[a, b]$ is

$$S_L\left(a, b\right) = \int_a^b \sqrt{1 + S'^2}\ \mathrm{d}w,\tag{4}$$

where

$$S'\left(w\right) = \frac{h_b}{k-1}\left(\frac{2}{W_b}w^{2-k}\right)^{\frac{1}{k-1}}\tag{5}$$

is the breach side slope (for $w \geq 0$). This is not calculated analytically and has to be integrated numerically. Further, for given water level h inside the breach, the water surface width W and the corresponding breach area A are

$$W\left(h\right) = W_b \left(\frac{h}{h_b}\right)^{k-1},\quad A\left(h\right) = \frac{W_b}{kh_b^{k-1}}h^k.\tag{6}$$

The breach volume is calculated by integration of $A\left(h\right)$ along the breach (i.e. across the dam) with w_c being the crest width and s_e the embankment slope of the dam, hence

$$V_b = \frac{W_b h_b}{k}\left(w_c + \frac{2s_e h_b}{k+1}\right).\tag{7}$$

The breach side angle β at the top of the breach (see Figure 1c and Figure 2) represents the short-term critical failure angle of the dam material and is assumed to be constant during the erosion process (*Chinnarasri et al.*, 2004; *Frank*, 2016). This angle can exceed the long-term critical failure angle by far even for non-cohesive soil materials, due to stabilizing effects of the apparent cohesion (*Volz*, 2013; *Volz et al.*, 2017). Setting the breach side slope at the top $S'\left(W_b/2\right) = \tan\beta$ (see Eq. (5)), the shape exponent is then given by

$$k = \frac{2h_b}{W_b \tan\beta} + 1.\tag{8}$$

The breach volume rate of change $\frac{\mathrm{d}V_b}{\mathrm{d}W_b}$ in Eq. (1b) is different for the two breach development states vertical erosion $H_b > H_{b,min}$ and lateral widening $H_b = H_{b,min}$. In the first case the breach depth grows steadily with increasing top width (see Figure 1 and Eq. (2)) and

8

therefore both $\frac{\mathrm{d}h_b}{\mathrm{d}W_b}$ and k are constant. After the breach bottom has reached the dam foundation, the breach depth does not change anymore, but k is now decreasing with enlarged breach (see Eq. (8)). After some algebraic calculations the breach volume rate of change is finally given by

$$\frac{\mathrm{d}V_b}{\mathrm{d}W_b} = \begin{cases} h_b \left[\frac{2}{k}w_c + \frac{6}{k^2(k+1)}s_e h_b\right] & \text{if} \quad H_b > H_{b,min}, \\[2ex] h_b \left[\frac{2k-1}{k^2}w_c + \frac{2(3k^2-1)}{k^2(k+1)^2}s_e h_b\right] & \text{if} \quad H_b = H_{b,min}. \end{cases} \tag{9}$$

2.3 Breach Hydraulics

In this section all hydraulic variables inside the breach are defined, including information about the underlying assumptions. Finally the breach outflow Q_b to be used in Eq. (1a) is estimated.

The hydraulic variables are defined within the prismatic channel where the occurrence of critical flow (Froude number $Fr = 1$) is assumed to hold (see Figure 1a). The effect of the streamline curvature, implicating an energy head loss and $Fr < 1$ (*Walder et al.*, 2015), is not included here. This effect could be incorporated by introducing an additional parameter, e.g. discharge coefficient, or the Froude number. The breach geometry of this control section is assumed to be representative for the entire breach. At the same time it acts only as a hydraulic control section and therefore the location of this transition in flow regime is not relevant for the model (*Singh*, 1996; *Coleman et al.*, 2002; *Macchione*, 2008). The energy head is given by

$$H_e = H_r - H_b = h + \frac{v^2}{2g}, \tag{10}$$

where h is the water depth, v is the average flow velocity within the breach and g is the acceleration due to gravity. The critical flow condition implies the maximum breach discharge for given H_e. Thus the critical water depth and flow velocity are

$$h_c = \frac{2k}{2k+1}H_e, \qquad v_c = \sqrt{g\frac{h_c}{k}}. \tag{11}$$

Consequently the breach discharge can be written as

$$Q_b = A(h_c)\,v_c = \frac{W_b}{h_b^{k-1}}\sqrt{\frac{g}{k^3}}h_c^{k+1/2}. \tag{12}$$

In addition the hydraulic radius of the critical cross-section ($h = h_c$) can be calculated through the following relation

$$r_h = \frac{A(h)}{P_w(h)}, \quad P_w(h) = 2 \cdot S_L\left(0, \frac{W(h)}{2}\right) \tag{13}$$

where $P_w(h)$ is the wetted perimeter along the breach side walls up to level h (see Eq. (4) and Eq. (6)).

2.4 Breach Erosion

In this section the parameters that control the sediment transport rate inside the breach are introduced. Ultimately the sediment transport rate out of the breach Q_s is defined, which is needed in Eq. (1b).

The flow velocity v and the corresponding hydraulic radius r_h are supposed to control the sediment transport within the breach. Empirical sediment transport formulas that quantify bed, suspension or total load are commonly formulated as $q_s \propto (\tau_b - \tau_c)^{c_1} v^{c_2} r_h^{c_3} \left[\mathrm{m^2 \, s^{-1}} \right]$, where $\tau_b \propto r_h J$ is the bottom shear stress for steady flow conditions and τ_c is the critical shear stress for incipient motion. The breach formation process is regarded as intense bedload phenomena, where $\tau_b >> \tau_c$, and the threshold τ_c is neglected therefore. The energy slope J can be defined by applying empirical flow formulas of type $v \propto r_h^{c_4} J^{0.5}$. Exponents $c_{1...4}$ are constants that differ between distinct empirical formulas. Knowing the flow velocity v from Eq. (11) and hydraulic radius r_h from Eq. (13))in the control section of the breach, the transport rate is then

$$q_s = \gamma \cdot v^{\nu} \cdot r_h{}^{\eta} \tag{14}$$

with γ being a global scaling coefficient that acts as a tuning factor to accelerate or hinder the erosion process, and $\nu = 2c_1 + c_2$ and $\eta = c_1 (1 - 2c_4) + c_3$ are exponents that combine the empirical formulation of transport and friction laws. For example *Macchione* (2008) applied the sediment transport equation by Meyer-Peter and Müller ($c_1 = \frac{3}{2}, c_2 = c_3 = 0$) and the friction law by Strickler ($c_4 = \frac{2}{3}$), which yields $\nu = 3$ and $\eta = -0.5$. The two exponents are physically interpreted as: (i) the larger $\nu \, (> 0)$ is, the stronger will be the influence of the hydraulic condition on the sediment transport (high flow velocity leads to more erosion); and (ii) the smaller $\eta \, (< 0)$ is, the stronger will be the influence of breach geometry on the sediment transport (narrow shapes leads to more erosion).

The rate of sediment that is transported out of the breach is finally quantified as

$$Q_s = P_e \cdot q_s, \quad P_e = 2 \cdot S_L \left(w_0, \frac{W(h)}{2} \right), \tag{15}$$

where P_e is the erodible perimeter, defined by the length along the breach side wall S_L bounded by

$$w_0 = \begin{cases} 0 & \text{if} \quad H_b > H_{b,min}, \\[2ex] \frac{2-k}{k} \frac{W(h)}{2} & \text{if} \quad H_b = H_{b,min}, \end{cases} \tag{16}$$

and the transverse location where the water surface intersects with the breach side. Hence in case of vertical erosion the erodible perimeter is equal to the wetted perimeter. During lateral widening the erodible perimeter is dependent on the breach shape: for triangular shape ($k = 2$) the complete wetted perimeter is accessible for sediment transport, for rectangular shape ($k \to 1$) only the vertical part of the breach is erodible, and for all intermediate breach shapes a smooth transition between these two extremes occurs, i.e. the closer the breach

shape is to a rectangle, the smaller will the erodible zone get due to fixed dam foundation in the middle of the breach.

2.5   Reservoir Depletion

The aim of this section is to quantify the reservoir volume change rate $\frac{\mathrm{d}V_r}{\mathrm{d}H_r}$. It is the last missing variable for the characterization of the ODE in Eq. (1).

The reservoir retention curve is parameterized by a power function

$$V_r = V_{r,0} \left(\frac{H_r}{H_{r,0}} \right)^\alpha, \tag{17}$$

where $V_{r,0}$ and $H_{r,0}$ are initial values for the water volume and its associated reservoir level, e.g. the storage volume and full supply level. The exponent α characterizes the reservoir topography, i.e. $\alpha = 1.0$ represents a rectangular basin whereas $\alpha = 4.0$ denotes a basin located in rather mountainous valleys (*Kühne*, 1978). Consequently, the reservoir volume depletion rate can be described as

$$\frac{\mathrm{d}V_r}{\mathrm{d}H_r} = \frac{\alpha V_{r,0}}{H_{r,0}{}^\alpha} H_r{}^{\alpha-1}. \tag{18}$$

2.6   Summary and Numerical Solution of the Parametric Breach Model

To solve the system of ODEs in Eq. (1), the geometrical properties of the dam-reservoir system (h_d, w_c, s_e, β, α) have to be defined. Furthermore the state variables of the system and their initial values have to be set given all model parameters (see Table 1): the drop in reservoir level ΔH_r, released reservoir volume ΔV_r, and final breach height ΔH_b. Detailed information are provided in Appendix A.

In each time step, the depletion rate of the reservoir level $\frac{\mathrm{d}H_r}{\mathrm{d}t}$ and the widening rate $\frac{\mathrm{d}W_b}{\mathrm{d}t}$ are calculated through Eqs. (9), (12), (15), and (18). The system of ODEs is numerically integrated using classic four step Runge-Kutta scheme, initially choosing a large time step. By re-integration with decreased time step, the peak of the resulting breach outflow hydrograph will be closer to its exact value. This iteration is done until a certain relative precision is reached. In this study a value of 1‰ was chosen.

3   Probabilistic Model Calibration

The deterministic dam breach model outlined in the previous section is hereinafter treated as 'black box' where no use of information about the mathematical model implemented by the numerical code is made (*Kennedy and O'Hagan*, 2001). Generally, the calibration procedure described in this section may be applied without fundamental modifications to other dam

Table 1: Parameter vector $\boldsymbol{x}$ of the deterministic dam breach model, categorized by their physical meaning.

description	$\boldsymbol{x}$
dam properties:	
dam height [m]	h_d
crest width [m]	w_c
embankment slope $[-]$	s_e
reservoir properties:	
reservoir basin shape $[-]$	α
reservoir level drop [m]	ΔH_r
released reservoir volume [m^3]	ΔV_r
breach properties:	
breach side angle [°]	β
final breach height [m]	ΔH_b
initial breach depth relative to ΔH_r [-]	r_0
erosion properties:	
exponent for flow velocity $[-]$	ν
exponent for hydraulic radius $[-]$	η
scaling coefficient for transport rate $[-]$	γ

breach models than the one previously introduced, varying in level of details and/or parameterization. However, overly detailed models with many parameters significantly increase the risk of over-parametrization, especially when only scarce or poor-quality data is available. The model of concern is formally described as a function

$$\tilde{\boldsymbol{y}} = \mathcal{M}\left(\boldsymbol{x}\right), \tag{19}$$

where the vector of model parameters $\boldsymbol{x}$ is mapped to the vector of model outputs $\tilde{\boldsymbol{y}}$. The parameters $\boldsymbol{x}$ are listed in Table 1. The model outputs of interest in this study are

$$\tilde{\boldsymbol{y}} = \left(\log_{10}\tilde{Q}_p, \log_{10}\tilde{W}_f\right), \tag{20}$$

with $\tilde{Q}_p$ $\left[\mathrm{m}^3\,\mathrm{s}^{-1}\right]$ being the peak discharge of the outflow hydrograph and $\tilde{W}_f = \bar{W}_b/k$ [m] the average breach width at the end of the breaching process. According to literature, where errors of these quantities are given as orders of magnitude (*Wahl*, 2004; *ASCE/EWRI*, 2011), a logarithmic scale is chosen here.

When applying the proposed dam breach model to predict a hydrograph of an existing but potentially failing embankment dam, some model parameters in $\boldsymbol{x}$ can be easily quantified (e.g. crest width w_c), while others are not known due to lack of data (e.g. basin reservoir shape α) or they are unobservable and act as tuning coefficients (e.g. scaling coefficient of sediment transport γ). Parameters that show a strong epistemic character, e.g. parameters that quantify the breach erosion (see Eq. (14)), will undergo a Bayesian update in the present study. These parameters are defined as vector $\boldsymbol{\phi}$ and called *Quantities of Interest* (*QoI*) hereafter. The goal is to draw conclusions about the epistemic model inputs from real observation data $\langle \boldsymbol{y}_i \rangle$. Henceforward $\langle \boldsymbol{y}_i \rangle$ stands for a sequence $\langle \boldsymbol{y}_i \rangle_{1 \leq i \leq n} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)$ of n experiments, containing Q_p and W_f of historically failed embankment dams (see Table 2). Applying Bayes' rule yields the joint posterior density

$$\pi \left(\boldsymbol{\phi} \mid \langle \boldsymbol{y}_i \rangle \right) = \frac{\mathcal{L} \left(\boldsymbol{\phi}; \langle \boldsymbol{y}_i \rangle \right) \cdot \pi \left(\boldsymbol{\phi} \right)}{C}, \tag{21}$$

where $\mathcal{L} \left(\boldsymbol{\phi}; \langle \boldsymbol{y}_i \rangle \right)$ is the likelihood function, $\pi \left(\boldsymbol{\phi} \right)$ the prior distribution of the model parameters, and C is a normalization constant such that the integral of the posterior distribution $\pi \left(\boldsymbol{\phi} \mid \langle \boldsymbol{y}_i \rangle \right)$ is equal to one.

The unknown parameters in $\boldsymbol{\phi}$ are subsequently quantified by Bayesian inference, that is "the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on observed quantities such as predictions for new observations" (*Gelman*, 2014). This is achieved by performing the following steps of classical Bayesian data analysis:

1. *Setting up a full probabilistic model*, i.e. the previously defined dam breach model is put into a Bayesian multilevel framework to assess different levels of uncertainty according to the underlying physical problem. This includes the specification of prior knowledge and the formulation of the likelihood function.

2. *Conditioning the model on observed data*, i.e. setting up a residual model that describes the forward model discrepancy, quantifying the prior knowledge, defining the likelihood function with given data of historical dam failures, and finally calculating the appropriate posterior distribution of the unobserved quantities of interest.

3. *Evaluating the goodness-of-fit*, i.e. analyzing the posterior distribution and compare the output of the now calibrated model with the observed data.

These three steps are further detailed in the next sections. Further steps would consist of including additional data not considered yet for model conditioning, i.e. performing the three steps again whereas the posterior distribution now represents the new prior knowledge.

3.1 Bayesian Multilevel Framework

The probabilistic shell that is shaped around the deterministic dam breach model can be represented by a Bayesian network (see Figure 3). The Bayesian multilevel approach applied

herein is adopted from *Nagel and Sudret* (2016). It provides "a natural framework for solving complex inverse problems in the presence of natural variability and epistemic uncertainty". The multilevel character of the method at hand is given by the hierarchically composed sub-models, such as the deterministic forward model itself, different categories of parameter uncertainty and/or variability described by the prior model, and prediction errors of the forward model specified in the residual model.
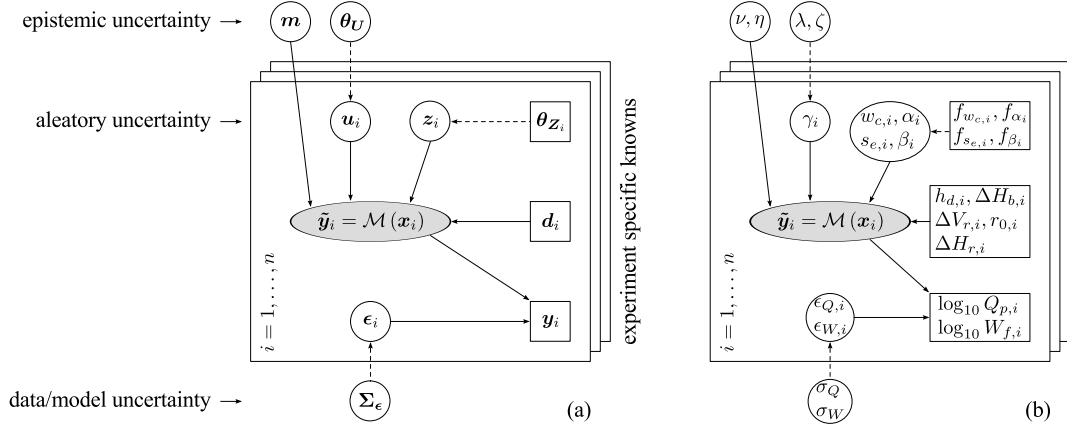

Figure 3: Directed acyclic graph (DAG) representing the probabilistic multilevel modeling approach followed in this study: vertices represent known ($\square$) or unknown ($\circ$) quantities, whereas directed edges symbolize their deterministic ($\rightarrow$) or probabilistic ($\dashrightarrow$) relations (adapted from *Nagel and Sudret* (2016)). On the left side the general formulation is shown (a), whereas on the right side the actual parameters of the dam breach model and their role in the multilevel framework are shown (b).

The parameters in $\boldsymbol{x}$ of the forward model (see Eq. (19)) can be categorized according to their physical meaning (see Table 1). Alternatively they can be classified according to four different categories of parameters (see Figure 4)

$$\boldsymbol{x} = (\boldsymbol{m}, \boldsymbol{u}, \boldsymbol{z}, \boldsymbol{d}) \tag{22}$$

that differ in their (un)certain nature, as done in *Eicher* (2014). This classification is based on general data availability on the one hand, and quality of empirical parameter knowledge on the other hand. The classification proposed in this study is one possible solution. Under different circumstances and/or available expert knowledge the result might be an alternative classification.

The first category is defined as well-known experimental conditions $\boldsymbol{d} = (h_d, \Delta H_r, \Delta V_r, \Delta H_b, r_0)$, that are normally reported for historical dam failure events. These quantities are assumed to be perfectly known for each experiment i and they can be considered as deterministic arguments of the dam breach model (see Table 2). The second category of parameters $\boldsymbol{z} = (w_c, s_e, \alpha, \beta)$ are subject to known and experiment-specific aleatory uncertainty. Their

14

		aleatory uncertainty	
		no fixed value	**yes** sampled from pdf
epistemic uncertainty	**no** known	$\boldsymbol{d}$ $h_d, \Delta H_r, \Delta V_r,$ $\Delta H_b, r_0$	$\boldsymbol{z}$ w_c, s_e, α, β
	yes unknown	$\boldsymbol{m}$ ν, η	$\boldsymbol{u}$ γ

Figure 4: Classification of the dam breach model parameters according to their (un)certain nature: the combinations of aleatory and epistemic uncertainty result in four different classes. The parameters with known values and known probability distribution function (pdf) ($\boldsymbol{d}$, $\boldsymbol{z}$ respectively) are experiment specific whereas the the parameters with unknown values and unknown hyper-parameters ($\boldsymbol{m}$, $\boldsymbol{u}$ respectively) are assumed to be fixed for all experiments (see Figure 3).

true values are not known as for $\boldsymbol{d}$ but follow the probability distributions $f_{\boldsymbol{Z}}(\boldsymbol{z}_i; \boldsymbol{\theta_Z})$. That is, the hyper-parameters $\boldsymbol{\theta_Z}$ are assumed to be well-known, whereas the vector of realizations $\boldsymbol{z}_i$ throughout the experiments $i = 1 \ldots n$ is not known. Parameters belonging to this class are usually not reported for historical dam failures but there is data available from which we can define the population distribution function (*Eicher*, 2014). Furthermore, parameters that are subject to epistemic uncertainty $\boldsymbol{m} = (\nu, \eta)$, i.e. exponents of the transport rate formula in Eq. (14), are treated as constant model parameters, but their true values are not known. Ultimately, $\boldsymbol{u}$ represents unknown parameters that are subject to aleatory uncertainty that itself is unknown and not changing across the experiments. In case of the dam breach model at hand the scaling coefficient γ controlling the velocity of dam material erosion via the sediment transport in Eq. (14) is the only parameter that belongs to this category. Thereby not only the realizations are unknown, but as well the hyperparameters $\boldsymbol{\theta_U}$ describing the probability function $f_{\boldsymbol{U}}(\boldsymbol{u}_i; \boldsymbol{\theta_U})$ from which realizations $\boldsymbol{u}_i$ are drawn. Therein contained is the parametric uncertainty of the erosion process itself. However, the distribution family of $f_{\boldsymbol{U}}$ is assumed to be known.

3.2 Residual Model

Even when assuming perfectly known model parameters $(\boldsymbol{m}, \boldsymbol{u}_i, \boldsymbol{z}_i, \boldsymbol{d}_i)$, predictions $\tilde{\boldsymbol{y}}_i$ are expected to deviate from real observations $\boldsymbol{y}_i$ due to the output imperfection $\boldsymbol{\epsilon}$, that can be regarded as a container of measurement errors, numerical approximations, and general model inadequacies. In that sense observations

$$\boldsymbol{y} = \tilde{\boldsymbol{y}} + \boldsymbol{\epsilon} \tag{23}$$

Table 2: Data and its experiment-specific knowns applied in this study for model inversion, originating from *Wahl* (1998): The underlying population of the data are worldwide, historically failed, homogeneous earthfill embankment dams. Existing and potentially failing dams belonging to this population may be investigated with the model proposed in this study. Experimental conditions, uncertain parameters and observations are listed. Where good data about s_e and w_c are available, the location parameter (mean) is defined based on this information and the scale parameter (standard deviation) is chosen to be much smaller than the location. Otherwise s_e is drawn from $N_{s_e} = N(2.16, 0.66)$ bounded by [1, 10], and w_c from $LN_{w_c} = LN(1.55, 0.51)$, respectively (*Eicher*, 2014). Due to missing information the realization of α and β are sampled from $U_\alpha = U(1, 4)$ and $U_\beta = U(45, 90)$, respectively, representing the physically meaningful range.

Failure		experimental conditions d_i					uncertain parameters z_i				observations y_i	
i	Name	h_d [m]	ΔV_r [10^3 m^3]	ΔH_r [m]	ΔH_b [m]	r_0 [−]	f_{s_e} [m]	f_{w_c} [m]	f_α [−]	f_β [°]	Q_p [m^3 s^{-1}]	W_f [m]
1	Apisapha	34.1	22.2	28.0	31.1	0.2	$N(2.5, 0.01)$	$LN(1.59, 0.01)$	U_α	U_β	6850	93
2	Baldwin Hills	71.0	0.910	12.2	21.3	0.6	$N(1.9, 0.01)$	$LN(2.95, 0.01)$	U_α	U_β	1130	25
3	Butler	7.16[b]	2.38	7.16	7.16 [a]	0.2	N_{s_e}	LN_{w_c}	U_α	U_β	810	63
4	Fred Burr	10.4	0.750	10.2	10.4	0.2	N_{s_e}	LN_{w_c}	U_α	U_β	654	—
5	French Landing	12.2	3.87	8.53	12.2	0.2	$N(2.25, 0.01)$	$LN(0.88, 0.01)$	U_α	U_β	929	27
6	Frenchman Creek	12.5	16.0	10.8	12.5	0.2	$N(2.5, 0.01)$	$LN(1.81, 0.01)$	U_α	U_β	1420	55
7	Hatchtown	19.2	14.8	16.8	18.3	0.2	$N(2.25, 0.01)$	$LN(1.81, 0.01)$	U_α	U_β	3080	151
8	Ireland no. 5	5.18[b]	0.160	3.81	5.18	0.2	N_{s_e}	$LN(0.88, 0.01)$	U_α	U_β	110	14
9	Johnstown	38.1	18.9	24.6	24.6	0.2	$N(1.75, 0.01)$	$LN(1.12, 0.01)$	U_α	U_β	8500	95
10	Lawn Lake	7.9	0.798	6.71	7.62	0.2	$N(1.55, 0.01)$	$LN(0.88, 0.01)$	U_α	U_β	510[c]	22
11	Lily Lake	3.66[b]	0.093	3.35	3.66	0.2	N_{s_e}	LN_{w_c}	U_α	U_β	71	11
12	Little Deer Creek	26.2	1.36	22.9	26.2	0.6	N_{s_e}	$LN(1.81, 0.01)$	U_α	U_β	1330	30
13	Lower Latham	7.01	7.08	5.79	7.01	0.2	N_{s_e}	$LN(1.53, 0.01)$	U_α	U_β	340	79
14	Propsect	4.42[b]	3.54	1.68	4.42	0.2	N_{s_e}	$LN(1.46, 0.01)$	U_α	U_β	116	88
15	Quail Creek	21.3 [b]	30.8	16.7	21.3	0.2	N_{s_e}	LN_{w_c}	U_α	U_β	3110	70

[a] Assumed value: final breach height ΔH_b is equal to drop in reservoir level ΔH_r.
[b] Assumed value: dam height h_d is equal to final breach height ΔH_b.
[c] Value obtained by numerical simulation.

are interpreted as outcomes of a random process and the underlying random generating mechanism, the joint probability distribution of the model parameters, is investigated. Imperfections $\boldsymbol{\epsilon}$ are hereinafter referred to as residuals and are assumed to be realizations of a Gaussian random variable $\boldsymbol{E} \sim f_{\boldsymbol{E}}(\boldsymbol{\epsilon}; \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})$, usually centered around $\boldsymbol{0}$ with variance $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. Errors due to numerical approximations and model inadequacies are collected as structural errors herein. To separate them from imperfections $\boldsymbol{\epsilon}$, more sophisticated residual error models could be used, e.g. representing the structural errors by a functional error term (*Brynjarsdóttir and O'Hagan*, 2014). This analysis is not included in this study because of missing information about dam breach data uncertainties and because it is out of the scope of this study.

The unknown error of observed quantities Q_p and W_f and of model inadequacies is taken into account during the calibration procedure, referred to as *residual calibration*. For simplicity uncorrelated and normally distributed errors are assumed with variances

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \begin{bmatrix} \sigma_Q{}^2 & 0 \\ 0 & \sigma_W{}^2 \end{bmatrix}. \tag{24}$$

The standard deviations of both the error in peak discharge $\sigma_Q \left[\log_{10} \mathrm{m}^3/\mathrm{s}\right]$ and final breach width $\sigma_W \left[\log_{10} \mathrm{m}\right]$ are collected in the vector $\boldsymbol{\sigma} = \left(\sigma_Q, \sigma_W\right)$ such that $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \boldsymbol{\sigma}^2 \boldsymbol{I}_2$ and will be quantified through inference analysis.

3.3 Prior Knowledge

In a Bayesian fashion prior or expert knowledge is seen as a subjective degree of belief about the true values of the parameters. Prior parameter knowledge is formulated by means of probability distribution functions. Through Bayesian data analysis this prior knowledge is updated with a set of data resulting in posterior parameter distributions. It is distinguished between (1) structural priors and (2) parametric priors.

The former include information about the prior model of experiment-specific unknowns $\boldsymbol{z}_i$ and $\boldsymbol{u}_i$. It is referred to as prescribed uncertainty because the corresponding variabilities are integrated out in the marginalized formulation of the likelihood function (more details in Section 3.4). Therefore, this type of prior knowledge cannot be improved by additional experiments. The structural priors in this study are formulated as distribution functions $f_{\boldsymbol{Z}}(\boldsymbol{z}_i; \boldsymbol{\theta}_{\boldsymbol{Z}_i})$ to draw samples $\boldsymbol{z}_i$ of the parameter vector $\boldsymbol{z}$ (see Table 2) and

$$f_{\boldsymbol{U}}(\boldsymbol{u}_i; \boldsymbol{\theta}_{\boldsymbol{U}}) = LN(\gamma_i; \lambda, \zeta) \tag{25}$$

being a lognormal distribution to draw samples of the scaling coefficient γ_i. The hyperparameters $\boldsymbol{\theta}_{\boldsymbol{U}} = (\lambda, \zeta)$ are unknowns associated with parametric priors (see below) and composed by the location parameter λ (mean value of the associated normal distribution) and the scale parameter $\zeta > 0$ (standard deviation of the associated normal distribution).

Second, the parametric priors describe the knowledge about the global unknowns $\boldsymbol{m}$, $\boldsymbol{\theta_U}$, and $\boldsymbol{\sigma}$. Model parameters that belong to one of the aforementioned categories are herein referred to as quantities of interest QoI during Bayesian inferential analysis and collected as the vector $\boldsymbol{\phi} = (\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\sigma})$. Their prior formulation will be updated when Bayesian data analysis is applied. In case of $\boldsymbol{m} = (\nu, \eta)$ the expert knowledge falls back on empirical formulas quantifying sediment transport and hydraulic friction (see Appendix B) and are herein defined as bivariate normal distribution

$$\pi_{\boldsymbol{M}}(\boldsymbol{m}) \sim \boldsymbol{N}_2\left(\boldsymbol{\mu_m}, \boldsymbol{\Sigma_m}\right) \tag{26}$$

with $\boldsymbol{\mu_m}$ consisting of mean values $\mu_\nu = 4.0$ and $\mu_\eta = -0.5$, and $\boldsymbol{\Sigma_m}$ being the covariance matrix with $\sigma_\nu = 0.9$, $\sigma_\eta = 0.3$, and $\rho_{\nu\eta} = -0.1$. The hyperparameters $\boldsymbol{\theta_U} = (\lambda, \zeta)$ that quantify the structural prior of the scaling coefficient γ_i cannot be guessed based neither on expert nor on empirical knowledge. This is due to lack of data or even unobservability. Therefore the prior

$$\pi_{\boldsymbol{\Theta_U}}(\boldsymbol{\theta_U}) = \begin{pmatrix} \pi_\Lambda \\ \pi_Z \end{pmatrix} \sim \begin{pmatrix} U\left(-15, 5\right) \\ U\left(0, 2\right) \end{pmatrix} \tag{27}$$

does not contain any valuable information apart from physically feasible and broad enough ranges. The prior knowledge about structural model and data uncertainty constituted in random variable $\boldsymbol{\epsilon}$ is represented by non-informative uniform distributions

$$\pi_{\boldsymbol{E}}(\boldsymbol{\Sigma_\epsilon}) = \begin{pmatrix} \pi_{\sigma_Q} \\ \pi_{\sigma_W} \end{pmatrix} \sim \begin{pmatrix} U\left(0.0, 0.6\right) \\ U\left(0.0, 0.6\right) \end{pmatrix}. \tag{28}$$

Finally, the joint prior distribution is defined as a function of the quantities of interest

$$\pi\left(\boldsymbol{\phi}\right) = \pi\left(\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}\right) = \pi_{\boldsymbol{M}}\left(\boldsymbol{m}\right) \cdot \pi_{\boldsymbol{\Theta_U}}\left(\boldsymbol{\theta_U}\right) \cdot \pi_{\boldsymbol{E}}\left(\boldsymbol{\Sigma_\epsilon}\right) \tag{29}$$

and can be applied in Bayes' formula to evaluate the posterior distribution (see Eq. (21)).

3.4 Definition of the Likelihood Function

The probability of observed data $\boldsymbol{y}_i$ for given model parameters $(\boldsymbol{m}, \boldsymbol{\theta_U})$ and residual variability $\boldsymbol{\Sigma_\epsilon}$ is defined as the likelihood function

$$\mathcal{L}\left(\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}; \langle \boldsymbol{y}_i \rangle\right) = \prod_{i=1}^{n} f\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}\right). \tag{30}$$

The data $\boldsymbol{y}_i$ used in this study are listed in Table 2. The likelihood depends on the experiment specific knowns $(\boldsymbol{\theta_{Z_i}}, \boldsymbol{d}_i)$ and is conditional on the unknowns $(\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon})$. Here a marginalized (or integrated) formulation

$$f\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}\right) = \int_{\mathcal{D}_u} \int_{\mathcal{D}_z} f_{\boldsymbol{E}}\left(\boldsymbol{y}_i - \mathcal{M}\left(\boldsymbol{m}, \boldsymbol{u}_i, \boldsymbol{z}_i, \boldsymbol{d}_i\right); \boldsymbol{\Sigma_\epsilon}\right)$$
$$f_{\boldsymbol{U}\mid\boldsymbol{\Theta_U}}\left(\boldsymbol{u}_i \mid \boldsymbol{\theta_U}\right) f_{\boldsymbol{Z}\mid\boldsymbol{\Theta_Z}}\left(\boldsymbol{z}_i \mid \boldsymbol{\theta_Z}\right) \mathrm{d}\boldsymbol{u}_i \, \mathrm{d}\boldsymbol{z}_i \tag{31}$$

is used, wherein the aleatory uncertainty in the so-called latent variables $(\boldsymbol{u}, \boldsymbol{z})$ is integrated out (*Nagel and Sudret*, 2016). This formulation focuses on global unknown parameters and therefore can be seen as a function of quantities of interest $\boldsymbol{\phi} = (\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon})$. Alternatively the latent variables $(\boldsymbol{u}, \boldsymbol{z})$ can be inferred, avoiding the numerical integration of Eq. 31. However, the knowledge about future realizations of $(\boldsymbol{u}, \boldsymbol{z})$ cannot be improved thereby.

The formal likelihood function $f\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}\right)$ of data set i is numerically approximated by Monte-Carlo integration. Through independently sampling of $\left(\boldsymbol{u}_i^{(k)}, \boldsymbol{z}_i^{(k)}\right)$ from their population distributions and further calculating the vector of residuals $\boldsymbol{\epsilon}_i^{(k)} = \boldsymbol{y}_i - \mathcal{M}\left(\boldsymbol{m}, \boldsymbol{u}_i^{(k)}, \boldsymbol{z}_i^{(k)}, \boldsymbol{d}_i\right)$ the approximation of the formal likelihood function is determined. The distribution function $f_{\boldsymbol{E}}$ of the residuals, described by the covariance matrix $\boldsymbol{\Sigma_\epsilon}$ (see Eq. (24)), is applied directly and the likelihood is therefore estimated as

$$\hat{f}\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_u}, \boldsymbol{\Sigma_\epsilon}\right) = \frac{1}{K} \sum_{k=1}^{K} f_{\boldsymbol{E}}\left(\boldsymbol{\epsilon}_i^{(k)}; \boldsymbol{\Sigma_\epsilon}\right). \tag{32}$$

During the process of likelihood estimation the number of model evaluations K is incrementally increased until a relative precision on the estimated likelihood $\hat{f}\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon}\right)$ of 1% is reached. Usually $K = 10^4$ model evaluations are needed to reach this target.

3.5 Model Inversion

In a deterministic framework inverse modeling connotes finding the parameter set which produces the model output that fits best the observed data. In a Bayesian world inverse modeling implies updating the prior knowledge with observed data to end up with the posterior distribution. In addition to this full uncertainty picture, point estimates of the posterior, like mean or mode, are chosen as characterization of the high dimensional distribution in practice. With $\boldsymbol{\phi} = (\boldsymbol{m}, \boldsymbol{\theta_U}, \boldsymbol{\Sigma_\epsilon})$ being the vector containing the *QoI* of the multilevel framework introduced in Eq. (22), the joint posterior distribution in Eq. (21) is recalled as

$$\pi\left(\boldsymbol{\phi} \mid \langle \boldsymbol{y}_i \rangle\right) \propto \mathcal{L}\left(\boldsymbol{\phi}; \langle \boldsymbol{y}_i \rangle\right) \cdot \pi\left(\boldsymbol{\phi}\right). \tag{33}$$

To perform inferential analysis, random samples are needed from the posterior distribution. Even if the d-dimensional posterior is well defined in the space of the unknown parameters $d = \mathcal{D}_{\boldsymbol{\phi}} = \mathcal{D}_{\boldsymbol{m}} \times \mathcal{D}_{\boldsymbol{\theta_U}} \times \mathcal{D}_{\boldsymbol{\Sigma_\epsilon}}$, drawing samples from it is not trivial nonetheless. The standard approach to circumvent this challenge is the method of MCMC. In this study a differential evolution Markov chain (DE-MC) algorithm is applied (*ter Braak*, 2006) (see Appendix C for more details). Thereby generated samples of the target distribution, i.e. the joint posterior distribution in Eq. (33), can be further processed to finalize inference analysis and interpreting the impact on the breach model.

4 Inference Analysis Results

Statistical inference of the proposed dam breach model will be presented here. First, the methodology of how to interpret the results of inference analysis is introduced. Second, the results of the model inversion are demonstrated with $\boldsymbol{\phi} = \left(\lambda, \zeta, \nu, \eta, \sigma_Q, \sigma_W\right)$. Third, the results are discussed, especially with the focus on real case applications, where results of additional analysis is presented.

4.1 Methodology of Interpretation

MCMC simulation was run with the choice of population size $N = 12$, satisfying $N = 2*d$ for unimodal target distribution (*ter Braak*, 2006). Details concerning the MCMC performance are given in Appendix D. The total number of model evaluations during MCMC simulation was approximately $I \cdot N \cdot n \cdot K \approx 5 \cdot 10^9$. An efficient implementation of deterministic dam breach model and highly parallel execution yields 10^4 model evaluations per second, resulting in a total run time of roughly 150h.

Since the initial motivation of this study was to improve the reliability of dam breach model predictions, the goodness-of-fit (GoF) of the presented parameter model is discussed. For this purpose the model is run using the same input data as used for model inversion, including the structural priors and the updated parametric prior, i.e. the posterior of *QoI*. The most probable parameter set $\hat{\boldsymbol{\phi}}$ of the d-dimensional joint posterior is taken to run the dam breach model. Hence, the variability $\tilde{\boldsymbol{y}}$ originates purely from aleatory uncertainty. Goodness-of-fit evaluations are done by visually comparing the predicted quantities $\tilde{\boldsymbol{y}}$ with the observed data $\boldsymbol{y}$ in form of violin plots (*Hintze and Nelson*, 1998). Further the GoF statistics

$$\boldsymbol{r} = \tilde{\boldsymbol{y}} - \boldsymbol{y} + \boldsymbol{\epsilon} \tag{34}$$

are processed in a descriptive manner. The expected value $\widehat{\mathrm{E}}\left[\boldsymbol{r}\right]$ is calculated for each data i and afterwards averaged over the whole sequence of n data sets to estimate potential prediction biases. Furthermore, the sequence $\langle\boldsymbol{p}_i\rangle$ with $\boldsymbol{p}_i = \mathrm{Pr}\left[\boldsymbol{r}_i \leq 0\right]$ being the percentile of data i is presented in form of a percentile plot. This way the location of potential prediction biases is visualized. As it is done for existing dam breach models (*Wahl*, 2004), the 95% interval $\widehat{\mathrm{I}}_{95}\left[\boldsymbol{r}\right] \approx 2.0\sqrt{\widehat{\mathrm{Var}}\left[\boldsymbol{r}\right]}$ is estimated to quantify the width of prediction band, where $\widehat{\mathrm{Var}}\left[\boldsymbol{r}\right]$ is the sample variance of GoF statistics $\boldsymbol{r}$. All analysis is done for overall objectives $\boldsymbol{y}$ and separately for both objectives peak discharge Q_p and final breach width W_f. This is because the model inversion is performed including the information of both objectives and any potential bias in predicting solely e.g. Q_p must be prevented.

Comparing the GoF statistics simplifies the interpretation of the resulting model accuracy and reliability and the implication on the application to predict a potential dam breach hydrograph. Quantitative summaries of this analysis are found in Table 3 and Table 4.
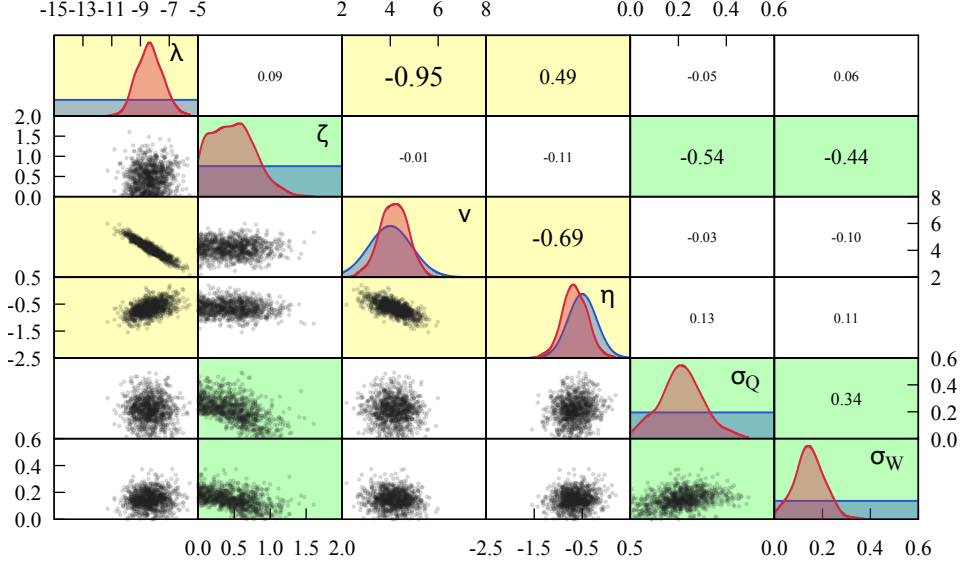
Figure 5: Posterior distribution samples: the diagonals show the marginals, posterior (red) as well as prior (blue) for purposes of comparison; the upper right part indicates the correlation coefficient between parameters; the scatter plots in the lower left part illustrate the actual bivariate distributions between the parameters. The highlighted (λ, ν, η) (yellow) and $(\zeta, \sigma_Q, \sigma_W)$ (green) parameters stand for quantities that show clear correlation patterns among themselves. The yellow subspace seems to be independent of its green counterpart.

4.2 Posterior Distribution

In Figure 5 the posterior distribution is visualized in terms of *iid* samples. The clear correlation pattern between the *QoI* indicates that the scale parameter ζ of the log-normal distribution describing the breach erosion, i.e. quantifying the uncertainty in the global scaling coefficient, is almost independent of other model parameters (λ, ν, η). They show considerable correlation coefficients among themselves, with $\rho_{\lambda, \nu} = -0.95$ being the most distinct correlation close to linear dependency. On the one hand it indicates an over-parametrization, i.e. one of the two parameters could be neglected in the model formulation without losing information or abandoning physics. On the other hand correlations are commonly caused by data that is not informative enough. The prior correlation between the erosion formula exponents $\rho_{\nu, \eta} = -0.1$, from the epistemic point of view, changed to a more noticeable relation in the posterior with $\rho_{\nu, \eta} = -0.69$. Physically speaking, the dependency between highly intense sediment transport and the impact of cross-sectional shape is more accentuated. Looking at the marginal distributions in the diagonals of Figure 5, the inferential information contained in the data used for the Bayesian update (see Table 2) is evident. The non-informative priors π_Λ and π_Z changed to peaky shaped distributions. Otherwise the marginal posteriors of exponents ν and η are not far from their prior $\pi_M(\boldsymbol{m})$ in Eq. (26). For η the mean value
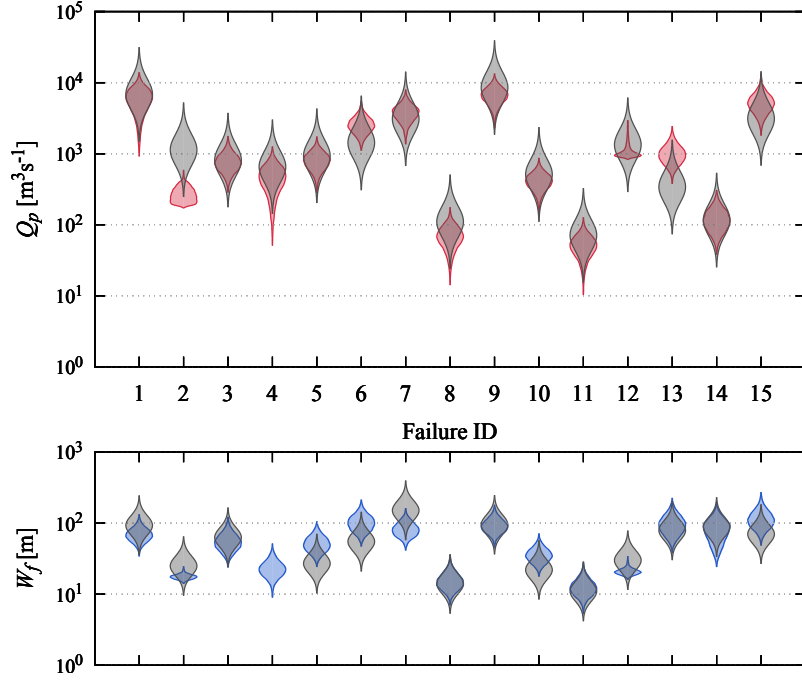
Figure 6: Performance evaluation of model inversion: comparing observed data including calibrated residual uncertainties (gray violins) with modeled data (colored violins). The 15 historic dam failure events used for model inversion (see Table 2) are shown in the upper plot containing peak discharge Q_p data (red) and final breach width W_f data (blue), respectively, in the lower plot.

is slightly shifted to a more negative value, whereas μ_ν did not change significantly, but σ_ν decreased notably. The marginal distribution of the scale parameter ζ shows a wide plateau close to zero. A second correlation pattern is noticed between the parameters $(\zeta, \sigma_Q, \sigma_W)$ which describe uncertainties of the dam breach model. The lower the uncertainty in erosion quantification is, i.e. lower values of ζ, the higher will be the values of structural model and data uncertainties σ_Q and σ_W. This additional pattern seems to be nearly independent of the correlations between parameters (λ, ν, η) that characterize the dam breach erosion but not the associated uncertainties. Estimating the most probable values of the joint posterior distribution for all *QoI*, roughly yields $\left(\hat{\lambda}, \hat{\zeta}, \hat{\nu}, \hat{\eta}, \hat{\sigma_Q}, \hat{\sigma_W}\right) = (-8.4, 0.34, 4.1, -0.61, 0.22, 0.14)$ (see Table 3).

4.3   Goodness of Fit

Running the dam breach model with data used for model inversion and estimated most probable values for unknown parameters, raises the possibility to evaluate the performance of the model inversion. In Figure 6 both objectives Q_p and W_f are displayed. Observed data of peak discharge and final breach width spread over two and one orders of magnitude,
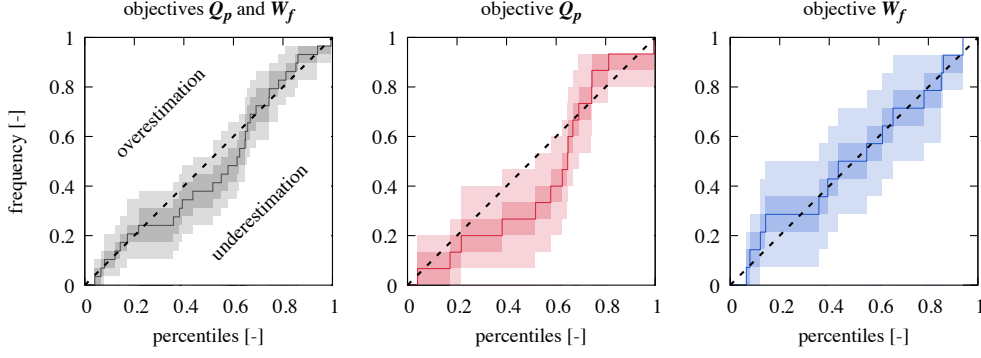
Figure 7: Percentile plot of model inversion performance evaluation, from left to right: summing up both objectives Q_p and W_f (grey); and splitting objectives peak discharge (red) and breach width (blue). Uncertainty bounds were estimated through sampling with replacement, also referred to as bootstrapping (*Efron*, 1979): median value (line), $[25, 75]\,\%$ interval (dark area), and $[5, 95]\,\%$ interval (light area).

respectively, represented by probabilistic means (violins) because of superimposed residual uncertainties σ_Q and σ_W. In many cases an almost complete overlap of the violins based on modeled and observed data is evident. Otherwise large parts of the violin tails show agreement, i.e. the model inversion appears to have succeeded. A quantitative summary of inversion performance is given in Table 4. The average inversion error amounts to $\widehat{\mathrm{E}}\,[\boldsymbol{r}] = -0.03 \pm 0.04$ and implies no biased performance as expected. Analyzing the objectives separately shows small but not significant biases $\widehat{\mathrm{E}}\,[r_Q] = -0.06 \pm 0.06$ and $\widehat{\mathrm{E}}\,[r_W] = 0.01 \pm 0.04$, respectively. This insight is disclosed visually in Figure 7. Final breach width tends to be overestimated, whereas peak discharge predictions tend towards underestimation in the mid percentile range and overestimation in the upper percentile range. Considering both objectives, as done during the process of model inversion, these trends diminish over the full range of percentiles. The average 95% prediction interval is approximately 0.4 order of magnitude when considering both objective quantities. Evaluating only peak discharge yields $\widehat{\mathrm{I}}_{95}\,[r_Q] \approx 0.5$, being within the known range for peak discharge predictions. A clearly smaller interval is observed when looking exclusively at breach width $\widehat{\mathrm{I}}_{95}\,[r_W] \approx 0.35$ compared to the literature (*Wahl*, 2004). Further, the correlation measure between GoF statistics $\widehat{\rho}\,[r_Q, r_W] \approx 0.2$ is fairly small. This indicates that even the simple residual model in Section 3.2 is sufficiently adequate in the case at hand. The results are justified by the assumption of a residual model where only the residual widths of the model output are investigated but no inference about their correlation is allowed for. Considering the variances of error sources (see Table 4) the inference analysis indicates that about ⅔ of the total variance arises from residual uncertainty $\widehat{Var}\,[\boldsymbol{\epsilon}]$ whereas the remaining part seems to originate from parametric uncertainty $\widehat{Var}\,[\boldsymbol{y} - \tilde{\boldsymbol{y}}]$.

4.4 Alternative of Zero Noise Assumption

The results of the successful inference analysis in the previous Section is not directly applicable for hydrograph predictions of a potentially failing dam. The reasons are twofold. On the one hand the observed data Q_p and W_f used for model calibration is not equivalent to the main output of the computational model, which is the outflow hydrograph used as input for flood wave propagation models. This ambiguity restricts potential model errors to be incorporated into the dam breach model for hydrograph predictions. On the other hand the simple residual model applied in this study (see Eq. (24)) prevents to distinguish between structural model errors measurement errors, both represented as part of the residuals ϵ. Consequently neglecting the residuals when predicting the hydrograph, for reasons mentioned before, will likely lead to an underestimation of the model uncertainties and the reliability of model predictions is at stake.

The correlation pattern observed in the posterior distribution (see Figure 5) suggests that changes in the residual model presumably will not affect the set of model parameters (λ, ν, η), but will have an impact on ζ. This represents a quite particular case, while the parameters are usually affected by a change of the residual model (e.g. *Thyer et al.*, 2009). Due to the modularity concept of the Bayesian multilevel framework (*Nagel and Sudret*, 2015), a change of the residual model does not considerably impinge on the other modules from a methodological and technical point of view. Thus, to quantify the parameter uncertainties for direct and predictive model application, a further simplification in the residual model is constituted.

The previous residual model in Eq. (24) is based on the strong assumptions of additivity, homoscedasticity, and Gaussianity. Now these assumptions are substituted by assuming "perfect" data and model. Hence, all imperfections

$$\|\mathbf{\Sigma}_{\boldsymbol{\epsilon}}\| \to \mathbf{0} \tag{35}$$

are expected to diminish, i.e. their variability is not bounded to the assumptions of additivity, homoscedasticity, and Gaussianity anymore, but will be compensated by the variability of model parameters. This residual model is referred to as *zero noise limit* hereinafter. Accordingly the *QoI* are reduced to $\phi = (\lambda, \zeta, \nu, \eta)$ representing a subset of the residual calibration. Now the *QoI* comprise the hyper-parameters (λ, ζ) of the log-normal distribution for γ_i and the exponents ν and η, quantifying the sediment transport within the breach.

The assumption of $\|\mathbf{\Sigma}_{\boldsymbol{\epsilon}}\| \to \mathbf{0}$, i.e. noise-free measurements and a perfectly accurate forward model, is obviously far from the true representation of ϵ in case of dam breach data, but no information about the level of measurement errors is at hand and potential structural errors will be compensated by parametric uncertainties in ζ. Variability in the data is still present due to varying inputs across several experiments. To incorporate the change of the residual model in Eq. (35) into the framework outlined in Section 3, the estimation of the

Figure 8: Posterior distribution samples, assuming zero noise limit: illustration is similar to Figure 5. The correlation pattern of parameters (λ, ν, η) (yellow) remains, whereas the model parameter ζ is independent.

integrated likelihood formulation (see Eq. (31)) is slightly modified and defined as the kernel density function

$$\hat{f}\left(\boldsymbol{y}_i \mid \boldsymbol{m}, \boldsymbol{\theta_u}\right) = \frac{1}{K} \sum_{k=1}^{K} \mathcal{K}\left(\boldsymbol{\epsilon}_i^{(k)}\right), \tag{36}$$

where $\mathcal{K}\left(\boldsymbol{\epsilon}_i^{(k)}\right)$ are Gaussian kernels and the bandwidth selection is based on the normal distribution approximation (*Silverman*, 1986).

This calibration configuration of the zero noise limit is seen as an alternative to the residual calibration. Reusing the data in Table 2 is legitimate as long as the knowledge inferred from residual calibration is not applied. Therefore the prior distribution defined in Section 3.3 is not changed, except from neglecting $\pi_{\boldsymbol{E}}\left(\boldsymbol{\Sigma_\epsilon}\right)$. The performance of the MCMC simulation can be found in Appendix D.1. The posterior resulting distribution is illustrated in Figure 8. Neither the correlation pattern nor the shape of the marginals of parameters (λ, ν, η) have changed compared to the residual calibration. The most probable parameter values $\left(\hat{\lambda}, \hat{\nu}, \hat{\eta}\right) = (-8.3, 4.2, -0.67)$ show negligible differences (see Table 3). A major difference is observed in the marginal distribution of the scale parameter ζ. Now a distinct peak is noticed around $\hat{\zeta} = 0.83$, suggesting that the uncertainty in the scaling coefficient increased, compensating for missing residual variability, not only for structural model errors, but for measurement errors as well.

Table 4 shows the average expected residual is approximately zero. The average 95% prediction interval is approximately half order of magnitude, slightly more than in case residual calibration. $\widehat{\mathrm{I}}_{95}\left[r_W\right] \approx 0.4$ is significantly larger, suffering most from the increased

25

Table 3: Most probable parameter value estimations of the joint posterior distribution including the estimation errors measured through bootstrapping.

| QoI | residual model | |
	residual calibration	zero noise limit
$\hat{\lambda}$	-8.37 ± 0.06	-8.25 ± 0.02
$\hat{\zeta}$	0.340 ± 0.010	0.833 ± 0.003
$\hat{\nu}$	4.12 ± 0.04	4.17 ± 0.01
$\hat{\eta}$	-0.610 ± 0.003	-0.669 ± 0.006
$\hat{\sigma}_Q$	0.220 ± 0.003	
$\hat{\sigma}_W$	0.139 ± 0.002	

correlation measure $\widehat{\rho}\left[r_Q, r_W\right] \approx 0.8$. The GoF statistics r_Q and r_W are strongly correlated and therefore the residuals $\boldsymbol{r}$ do not show properties of white noise. These correlated errors are attributed to either global model inadequacies and/or data inconsistencies, caused by the simplification of the zero noise assumption.

4.5 Comparison of Residual Models

Comparing the posterior distributions in Figure 5 and Figure 8 clearly illustrates, that on the one hand the characterization of the sediment transport formula within the breach (see Eq. (14)) is independent of the residual model approach. The correlation pattern among the parameters (λ, ν, η) as well as the set of their most probable values is not dependent on the choice of the residual model. This can be interpreted as a plausibility check for the resulting posteriors. Taking the median value $e^{\hat{\lambda}}$ of the log-normal distribution describing the stochastic scaling coefficient γ and writing the transport formula in a deterministic way, one finds $q_s \approx 0.00025 \cdot v^{4.15} \cdot r_{hy}^{-0.65}$. The exponents $\nu = 4.15$ and $\eta = -0.65$ are not much different from prior knowledge (see Eq. (26)). On the other hand, parameters that contain information about the model uncertainties $(\zeta, \sigma_Q, \sigma_W)$ are strongly dependent on the residual model. When assuming uncorrelated residuals, the model and data uncertainty is represented by $\sigma_Q = 0.22$ and $\sigma_W = 0.14$ and parametric uncertainty describing the scaling coefficient γ is quantified by the QoI $\zeta = 0.34$. For zero noise limit assumption all residual uncertainty is covered by parametric uncertainty of the breach erosion process, consequently ζ is increased from 0.34 to 0.83.

The differences between the two residual models is evident by performing a variance decomposition of model residuals (see Table 4). When assuming diminishing residuals during inversion, the overall output variance is increased from 0.052 to 0.057. No increase is observed in the variance of peak discharge residuals (originally 70% assigned to the data un-

Table 4: Error statistics of the model inversion performance evaluation: Estimated mean error, 95% confidence interval, correlation coefficient between objective quantities Q_p and W_f, and variances of different error sources.

residual model	output quantity	$\widehat{\mathrm{E}}\left[\boldsymbol{r}\right]$	$\widehat{\mathrm{I}}_{95}\left[\boldsymbol{r}\right]$	$\widehat{\rho}\left[r_Q, r_W\right]$	$\widehat{Var}\left[\boldsymbol{r}\right]$	$\widehat{Var}\left[\boldsymbol{y}-\tilde{\boldsymbol{y}}\right]$	$\widehat{Var}\left[\boldsymbol{\epsilon}\right]$
					(expressed in 10^{-2})		
residual calibration	Q_p	-0.06 ± 0.06	0.50 ± 0.01		7.1 ± 0.3	2.2 ± 0.3	4.0
	W_f	0.01 ± 0.04	0.35 ± 0.01		3.2 ± 0.2	1.2 ± 0.2	1.7
	Q_p, W_f	-0.03 ± 0.04	0.44 ± 0.02	0.20 ± 0.03	5.2 ± 0.3	1.7 ± 0.2	2.9
zero noise limit	Q_p	-0.03 ± 0.06	0.52 ± 0.03		7.1 ± 0.8	7.1 ± 0.8	0.0
	W_f	0.05 ± 0.04	0.41 ± 0.02		4.2 ± 0.3	4.2 ± 0.3	0.0
	Q_p, W_f	0.01 ± 0.04	0.47 ± 0.02	0.84 ± 0.04	5.7 ± 0.5	5.7 ± 0.5	0.0

certainty parameter σ_Q). A more prominent increase in variance is detected for final breach width prediction errors, namely, from 0.032 to 0.042 (primarily 60% covered by σ_W). This is explained by the strong correlation of residuals r_Q and r_W for the zero noise assumptions, where model outputs $\tilde{\boldsymbol{y}} = \left(\tilde{Q}_p, \tilde{W}_f\right)$ are correlated due to the physical breach formation processes implemented in the dam breach model. Breaking up this tie during residual calibration shows, that the global uncertainty in Q_p is responsible of overestimating the error in W_f predictions. Furthermore, the residual calibration suggests that more than half of the global uncertainties are associated with residual variability. To assign these global uncertainties to model inadequacies and measurement errors, more sophisticated error modeling would be needed that is beyond the scope of this study, but being a topic for future work. While pure measurement errors are typically closer to white noise, the model inadequacies can be modeled by a functional error term, often represented by Gaussian processes and dependent on the classes of known model parameters $\boldsymbol{d}$ and $\boldsymbol{z}$ (e.g. *Kennedy and O'Hagan*, 2001; *Brynjarsdóttir and O'Hagan*, 2014).

To qualitatively compare the two calibrated dam breach models, differing in the underlying residual model, and selecting the right one is known under Bayesian model selection (*Schöniger et al.*, 2014). This has not been accomplished in this study due to the computational cost of evaluating the evidence C (see Eq. (21)). The motivation to perform the non-standard alternative of zero noise limit (see Section 4.4) is crucial when facing a real case application where reliable hydrograph predictions are requested.

5   Discussion of Model Application

The probabilistic modeling framework and the according calibration results have been discussed in the previous section. Here the applicability of the resulting probabilistic dam breach model is discussed by applying the proposed model in a test case.
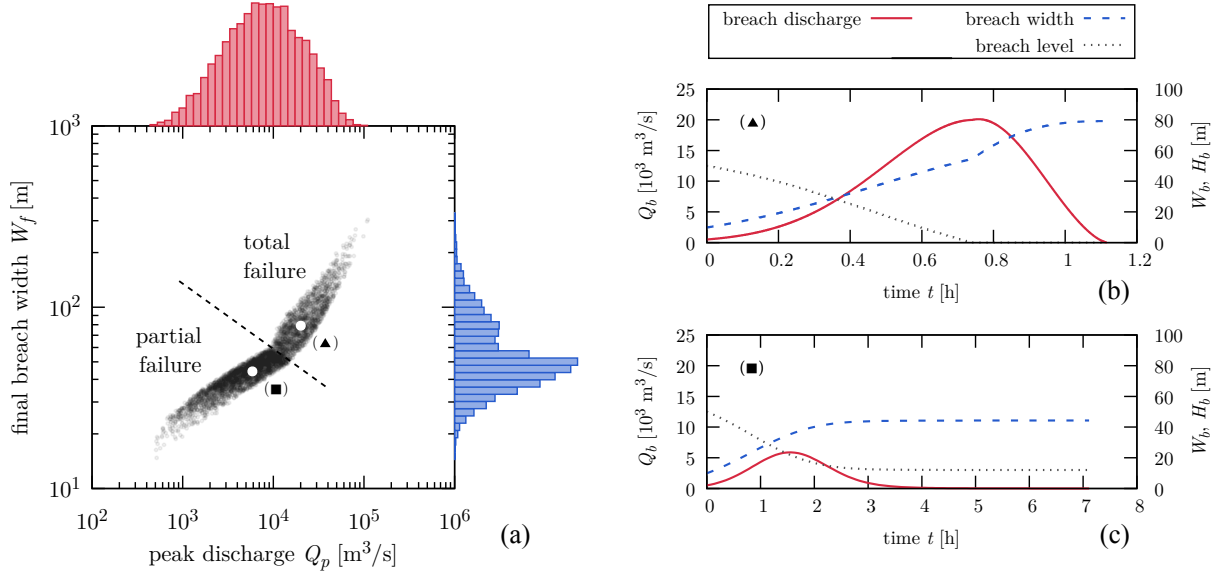
Figure 9: Application example of proposed probabilistic dam breach model (input data taken from *Zenz and Goldgruber* (2013)). Dam breach hydrograph predictions of hypothetical embankment dam by modeling physical breach formation processes: 5000 model evaluations and resulting peak discharge Q_p and final breach width W_f and their marginal distributions (a); and explicitly showing time series of two different samples, indicated with (■) and (▲), for breach outflow, breach width, and breach bottom level ((b) and (c)).

The problem statement of a hypothetical dam failure originates from a numerical benchmark organized by the International Commission on Large Dams *ICOLD* that took place in Graz 2013 (*Zenz and Goldgruber*, 2013). The parameter definition is given as follows (compare with Table 1). The embankment dam is characterized by dam height $h_d = 61\,\mathrm{m}$, crest width $w_c = 24\,\mathrm{m}$, and embankment slope $s_e = 3$. The reservoir is described by a drop in reservoir level $\Delta H_r = h_d = 61\,\mathrm{m}$ and according release of water volume $\Delta V_r = 38\,276\,344\,\mathrm{m}^3$, and the reservoir basin shape parameter approximately follows a uniform distribution $\alpha \sim U\,(2.5, 3.2)$ (estimated from retention curve). The breach side angle at the top is assumed to be uniformly distributed $\beta \sim U\,(50, 85)$, the maximal breach height is $\Delta H_b = 61\,\mathrm{m}$, and the initial breach depth is assumed to be $Y_{b,0} = 50\,\mathrm{m}$, and therefore $r_0 = 0.82$. The erosion of dam material is quantified by the transport formula in Eq. (14). For reasons of applicability and reliability discussed in Section 4.4 the calibration results under the assumption of zero noise are considered here. Making use of the most probable values as point estimates (see Table 3), the applied transport formula yields $q_s \approx \gamma \cdot v^{4.2} \cdot r_{hy}^{-0.67}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ where $\gamma \sim LN\,(-8.3, 0.83)$.

Running a Monte-Carlo simulation of 5000 model evaluations leads to 5000 different progressive breach formation processes. Here a Latin Hypercube Sampling (LHS) was chosen

as sampling strategy for the purpose of faster convergence of the MC simulation. In Figure 9 the resulting distribution of the hydrograph's peak discharge and the final breach width are shown. Two distinct behaviors can be observed in Figure 9a: parameter sets with high erosion rates (large γ parameter realization) lead to total failure of the dam, indicated by final breach W_f being larger than dam height $h_d = 61\,\mathrm{m}$, whereas parameter sets with low erosion rates (small γ parameter realization) lead to a partial failure only. In this exemplary application the driving force of the reservoir water does not seem to be sufficient to erode the large body of dam material in many parameter combinations. Two representative hydrographs for complete and partial failure are shown in Figure 9. In case of total failure the progression is obvious and peak breach outflow is reached within less than 1 h, nearly at the same time when the breach bottom hits the fixed bottom, i.e. transition from vertical erosion to lateral widening. In case of partial failure the gradual erosion is not fast enough to initiate a progressive failure mechanism and the breach bottom never reaches the fixed bottom. Here the duration of the failure process is much longer (a few hours).

The performance of the breach model proposed in this study is based on simple physical assumptions. Thus the model behavior is assumed to be close to physical processes of real dam failures, including the related probabilities of observing one or the other formation process. The resulting peak discharge distribution (see Figure 9a) is therefore much wider than the predicted peak discharges of the benchmark participants (ranging from 10 to $40 \times 10^3\,\mathrm{m^3\,s^{-1}}$). Neglecting the possibility of partial failure in this case might lead to a significant overestimation of the predicted hydrograph. In this case, the consequent flood wave calculation would result in too intensive hydraulic impact, hence the according quantification of the dam break risk is biased.

6 Conclusions

Real dam break risk analysis requires for reliable breach models capable to predict possible breach outflow hydrographs. The challenge to predict a worst case scenario, that still is regarded as a physically feasible event, is of particular interest for engineering applications. By contrast, seeking for more accurate breach models often loses sight of reliably predicting less probable events. In this study the lack of knowledge in describing the progressive process accurately is compensated by the development of a probabilistic dam breach model framework where uncertainties on different levels are quantified. The focus is on homogeneous and non-cohesive earthen embankment dams whose potential failures need to be analyzed for the purpose of further flood risk assessment.

Uncertainties in the quantification of dam material erosion is used to tune the newly proposed dam breach model in a Bayesian fashion. Prior information about dam breach model parameters together with data from historical dam failure events are used to perform Bayesian inference. Model inversion is accomplished by sampling from the posterior distribu-

tion through DE-MCMC simulations. The convergence to the posterior distribution and the consequent statistical inference of the multilevel model was successful using the information contained in a data set of 15 real dam failure events. The very same data has been used to carry out deterministic model calibration of simplified physics-based dam breach models comparable to the model developed in the study at hand (e.g. *Macchione*, 2008). By re-using the same data the focus is on the proposed framework itself instead of possible changes in prediction quality due to different data.

Observed data of peak discharge Q_p and final breach width W_f are used as output quantities. Output imperfections, consisting of structural model errors and measurement errors, are defined as residuals and their prior knowledge is represented by independent Gaussians for each output quantity. Inferring for parametric and residual uncertainties results in a posterior distribution with a distinct correlation pattern. Model parameters that characterize the sediment transport rate inside the breach are independent of residual uncertainty. Moreover, they are not related to the hyper-parameter describing the aleatory uncertainty of the global scaling coefficient γ. The characterization of the average sediment transport can be best described by $q_s \approx 0.00025 \cdot v^{4.15} \cdot r_{hy}^{-0.65} \, \mathrm{m^2 \, s^{-1}}$. Comparing this formulation with the theoretical values from the transport law by Meyer-Peter & Müller and the friction law by Manning ($\nu = 3.0$ and $\eta = -0.5$), the transport within the breach can be attributed to total-load-like ($\nu > 3.0$) and more sensitive to cross-sectional geometries ($\eta < -0.5$). However, the variability of q_s, described by the hyper-parameter ζ, is dependent on the variability of the residuals. The latter are correlated among themselves in addition.

Assessing the goodness-of-fit (GoF) statistics of the performed model inversion yields a 95% confidence interval for model predictions around half orders of magnitude for the model output peak discharge Q_p, falling in the range of values reported in literature for other dam breach prediction tools (*Wahl*, 2004; *Froehlich*, 2016). The model output final breach width W_f shows a 95%confidence interval around third orders of magnitude. The variability of the GoF statistics is explained by roughly 60% from the residual variability and 40% covered by the different parametric uncertainties.

The quantities Q_p and W_f representing the observed data are not conforming with the final model output in real model applications. It is the main motivation of the proposed probabilistic dam breach model to predict the time series of breach outflow (breach hydrograph) needed as upper boundary condition in state-of-the-art flood routing models. Assuming non-zero residuals and consequently incorporating reversely the estimated residuals of Q_p into predicted hydrographs is impractical. The clear correlation pattern in the posterior distribution and the hierarchical character of the Bayesian multilevel framework allows for an alternative formulation of the residual model, where residual uncertainties are assumed to diminish. This *zero noise limit* represents a subspace of the full problem. The inference results in this unusual setup show, that the characterization of q_s does not change, but uncertainty

due to both structural model errors and measurement errors is compensated by parametric uncertainty of the sediment transport rate quantification. As a result the uncertainty of model predictions is overestimated, whereas it is probably underestimated in the case of full residual calibration. On these grounds the zero noise assumption is still favorable when applying the proposed modeling framework to a real case, when the computational model predicts full hydrographs instead of maximum flows only. To avoid the non-standard residual model *zero noise limit*, two additional ingredients are needed in the proposed modeling framework: (i) improved error modeling, allowing for disentangling structural model errors and measurement errors; and (ii) development of a physically based technique to reversely apply uncertainties of Q_p and W_f to predicted hydrographs. In fact the community is encouraged to quantify and/or reduce the uncertainties contained in the observed data and thereby improving the precision of the model substantially (*Morris et al.*, 2008; *ASCE/EWRI*, 2011).

The probabilistic dam breach model was successfully applied to a test case, which consists of a hypothetical embankment dam (*Zenz and Goldgruber*, 2013). It clearly demonstrated the benefits of the proposed modeling framework, when taking into account the parametric uncertainties that were quantified in this study. The simplified model, which allows for physical interpretation of the results, combined with the motivation of improving prediction reliability, enabled by enhanced methodologies, provides an adequate compromise between model accuracy and model reliability.

Improving the formulation of the deterministic and simplified dam breach model, such as introducing additional physical parameters (e.g. discharge coefficient), requires to re-assess all parametric uncertainties by inference analysis. To strengthen and/or critically assess the proposed dam breach model framework, the same inferential analysis can be performed subsequently by inferring from an additional data set of observed data, where the prior is defined by the posterior of this study, as done in *Peter* (2017). Since the validity of similar dam breach model formulations have been proven (e.g. *Chinnarasri and Saelim*, 2009; *Capart*, 2013; *Vonwiller et al.*, 2015), observed data from real dam failure events is regarded to provide its maximum value within the proposed calibration procedure to further improve the empirical knowledge of model configurations. The gathering of dam failure data containing qualitatively and quantitatively enough information remains a main challenge.

Eventually, to deal with a probabilistic breach outflow hydrograph, the need for fast and accurate flood wave calculation tools is of paramount importance. Recent advances in computer sciences open a new door to fill this gap (*Kalyanapu et al.*, 2011; *Lacasta et al.*, 2014; *Reguly et al.*, 2015). The possibility of running two dimensional flood wave calculation as Monte-Carlo simulation with random hydrograph is real (*Peter*, 2017). The information contained in the resulting probabilistic flood maps is abundant. Finally, there is a strong need for engineering design aids and standard procedures for embankment breach analysis and hazard management that quantitatively incorporate uncertainties.

A Dam breach model initial conditions

Fixed breach bottom $H_{b,min}$, initial breach level $H_{b,0}$, initial reservoir level $H_{r,0}$ and corresponding reservoir volume $V_{r,0}$ are calculated through (see Figure 1)

$$H_{b,min} = h_d - \Delta H_b \tag{37a}$$

$$H_{b,0} = H_{b,min} + (1 - r_0)\,\Delta H_r \tag{37b}$$

$$H_{r,0} = H_{b,min} + \Delta H_r \tag{37c}$$

$$V_{r,0} = \Delta V_r \frac{H_{r,0}{}^\alpha}{H_{r,0}{}^\alpha - H_{b,min}{}^\alpha}. \tag{37d}$$

Furthermore, having the breach side angle β, that is a proxy for the embankment material property, and the initial breach level $H_{b,0}$ is not sufficient. An initial value of the breach top width W_b is needed in addition. In case of triangular breach shape ($\beta = 45°$) the initial breach width $W_{b,0} = 2h_b$ is unique and the according breach discharge

$$Q_{b,ref} = \sqrt{\frac{512}{3125} g H_e{}^5} \tag{38}$$

is taken as reference value. However, $W_{b,0}$ is not unique for all other choices of β. Thus, the restriction is set, that the initial breach discharge $Q_b\left(\beta, W_{b,0}\right)$ is equal to initial discharge of a triangular breach in Eq. (38)) (*Franca and Almeida*, 2005). When assuming that the reservoir is filled to capacity ($H_{r,0} \,\hat{=}\, h_d$), the initial breach width can then be approximated as

$$W_{b,0}\left(\beta\right) \approx \frac{16}{25}\left(5 - \frac{\beta}{24}\right)\left(h_d - H_{b,0}\right). \tag{39}$$

B Empirical transport and friction formulas

The exponents ν and η in Eq. (14) vary strongly depending on the field of application. Existing transport, respectively friction quantification tools, are based on laboratory and/or field test data that is further processed to result in simple empirical formulas. Numerous examples of such formulas can be found in the literature. According to *Wu* (2008) and *Machiels et al.* (2011) lower and upper bounds for the exponents $c_{1...4}$ (see Section 2.4) are assumed in this study as

$$c_1 \in [1.0, 2.2] \tag{40a}$$

$$c_2 \in [0.0, 2.0] \tag{40b}$$

$$c_3 \in [-0.6, 0.0] \tag{40c}$$

$$c_4 \in [0.4, 0.72]. \tag{40d}$$

Uniformly sampling from the above ranges leads to an empirical approximation of possible model parameter combinations $\nu = 2c_1 + c_2$ and $\eta = c_1\left(1 - 2c_4\right) + c_3$.

C   Markov Chain Monte Carlo (MCMC) algorithm

The MCMC algorithm applied in this study (*ter Braak*, 2006) runs N Markov chains in parallel and has been proven to ensure both detailed balance and ergodicity. Modifications with improved efficiency for high dimensional posterior are available in *Vrugt et al.* (2009) or *Laloy and Vrugt* (2012). The heart of this MCMC algorithm is the absence of a predefined and fixed proposal distribution, but new parameter configurations are instead proposed in an adaptive manner for j'th chain as

$$\boldsymbol{\phi}_p = \boldsymbol{\phi}_j + \delta \left(\boldsymbol{\phi}_{r2} - \boldsymbol{\phi}_{r1} \right) + \boldsymbol{e}, \tag{41}$$

with $\boldsymbol{\phi}_j$ being the actual parameters of chain j, $r1$ and $r2$ are randomly selected chains different from j, $\delta = 2.38/\sqrt{2d}$ is a scaling factor for the jumping width to ensure efficient acceptance rates. To ensure ergodicity of the Markov Chain, $\boldsymbol{e}$ is drawn from a symmetric distribution with a small variance compared to that of the target, but with unbounded support. Initial states are generated by sampling from the prior distribution $\pi(\boldsymbol{\phi})$. In addition, classic Metropolis algorithm is applied, i.e. accepting proposed parameters $\boldsymbol{\phi}_p$ with probability

$$a = \min \left(1, \frac{\pi \left(\boldsymbol{\phi}_p \mid \langle \boldsymbol{y}_i \rangle \right)}{\pi \left(\boldsymbol{\phi}_j \mid \langle \boldsymbol{y}_i \rangle \right)} \right). \tag{42}$$

D   MCMC performance

MCMC simulations run in this study showed time series with acceptance rates between 0.25 and 0.30 over the total number of iterations I, ensuring good mixing of the random walk and therefore converging optimally to the stationary posterior distribution (*Roberts and Rosenthal*, 2001). Starting from initial states, samples of I_b iterations are rejected until the stationary target distribution was reached, also known as burn-in period. To monitor convergence of MCMC to the target distribution, the diagnostic *potential scale reduction factor PSRF* was applied (*Brooks and Gelman*, 1998), in which the variances (covariances in case of multivariate analysis) within each MCMC-chain and between the N-chains are compared. Convergence is assumed to be reached if $PSRF < 1.1$ and the effective number of iterations is $I_{eff} = I - I_b$. Samples of I_{eff} show strong autocorrelation because of the Markov Chain generation mechanism. The number of effective samples, that is samples showing properties of *iid* (independent and identically distributed) random samples, is estimated through

$$\hat{N}_{eff} = \frac{N \cdot I_{eff}}{1 + 2 \sum_{l=1}^{L} \hat{\rho}_l}, \tag{43}$$

where $\hat{\rho}_l$ is the sample autocorrelation with lag l and maximum lag L is the first odd integer for which $\hat{\rho}_{L+1} + \hat{\rho}_{L+2} < 0$ (*Gelman*, 2014). MCMC simulation runs until $\hat{N}_{eff} > 1000$ for each *QoI*. The denominator in Eq. (43) is applied as thinning lag t_l to get rid of the
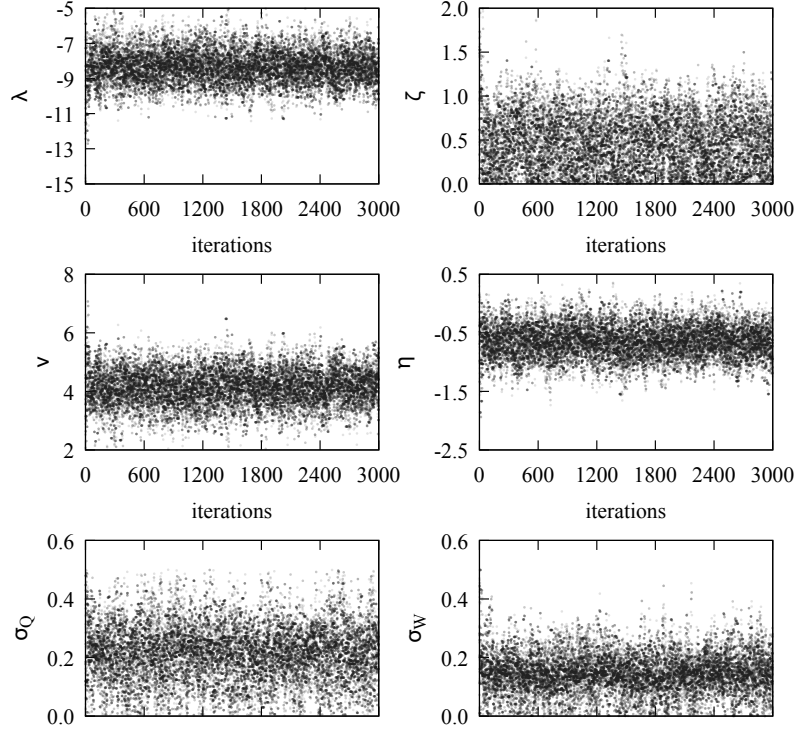
Figure 10: Time Series of the Markov Chain Monte Carlo simulation (differential evolution algorithm with 12 parallel chains (*ter Braak*, 2006)) for the quantities of interest $\phi = \left(\lambda, \zeta, \nu, \eta, \sigma_Q, \sigma_W\right)$.

autocorrelation. Finally, the remaining samples represent independent samples of the joint posterior distribution that will undergo further analysis.

D.1 Residual Calibration

Running the MCMC simulation with $N = 2d = 12$ parallel chains results in the time series given in Figure 10 obtained using a total of $I = 3000$ iterations. The time series plot does not show any pattern of bad mixing what is congruent with a mean acceptance rate over all chains of 0.29. The convergence to a stationary distribution is illustrated by the potential scale reduction factor *PSRF* in Figure 11. After 300 iterations the chains are already close to their stationary distribution. The last quantity to reach the limit of *PSRF* < 1.1 is the multivariate diagnostic and σ_Q at $I_b = 800$ iterations. After having rejected iterations of the burn-in period, the remaining 2200 iterations are thinned with $t_l = 26$ (see Eq. (43)) to finally obtain approximately 1000 *iid* posterior samples put together from the N parallel chains.
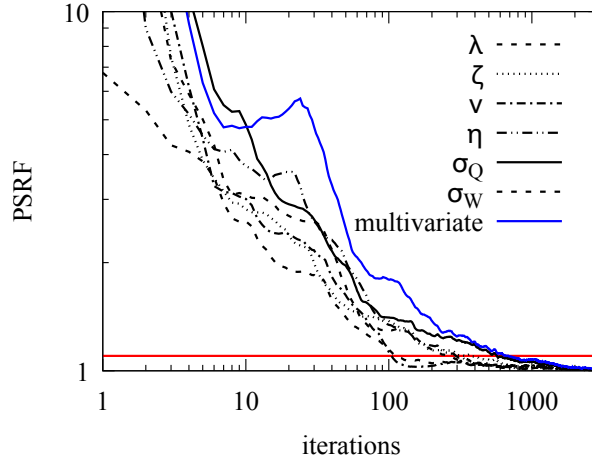
Figure 11: Convergence diagnostic after *Brooks and Gelman* (1998) for the four *QoI* and their multivariate version: after $I_b = 800$ iterations all parallel chains in MCMC simulation seem to have reached the stationary target distribution (potential scale reduction factor $PSRF < 1.1$).

D.2 Zero Noise Limit

Compared to the residual calibration the number of chains $N = 12$ was not changed. The mean acceptance rate of 0.28 is still within the known range of optimal convergence behavior. Due to smaller parameter space $d = 4$ convergence is slightly faster and the burn-in period $I_b = 600$ is shorter accordingly. Likewise the autocorrelation has decreased and a thinning lag $l_t = 19$ was applied to gain *iid* samples. The MCMC simulation was run with $I = 2200$ iterations to end up with approximately 1000 posterior samples.

Acknowledgments

References

Ahmadisharaf, E., A. J. Kalyanapu, B. A. Thames, and J. Lillywhite (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, doi:10.1016/j.envsoft.2016.09.022.

Alfonso, L., M. M. Mukolwe, and G. Di Baldassarre (2016), Probabilistic flood maps to support decision-making: Mapping the value of information, *Water Resources Research*, *52*(2), 1026–1043, doi:10.1002/2015WR017378.

Altinakar, M. S., E. E. Matheu, and M. Z. McGrath (2009), New generation modeling and decision support tools for studying impacts of dam failures, *Tech. rep.*, Association of State Dam Safety Officials (ASDSO).

Amini, A. B., V. Nourani, and H. Hakimzadeh (2011), Application of artificial intelligence tools to estimate peak outflow from earth dam breach, *International Journal of Earth Sciences and Engineering*, *4*(6), 243–246.

ASCE/EWRI (2011), Earthen embankment breaching, *Journal of Hydraulic Engineering*, *137*(12), 1549–1564, doi:10.1061/(ASCE)HY.1943-7900.0000498, ASCE/EWRI Task Committee on Dam/Levee Breaching.

Balakrishnan, S., A. Roy, M. G. Ierapetritou, G. P. Flach, and P. G. Georgopoulus (2003), Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical bayesian framework incorporating the stochastic response surface method, *Water Resources Research*, *39*(12), 1350–1362, doi:10.1029/2002WR001810.

Broich, K. (1996), Computergestützte Analyse des Dammerosionsbruchs, Ph.D. thesis, Universität der Bundeswehr München, Germany, [in German].

Brooks, S. P., and A. Gelman (1998), General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics*, *7*(4), 434–455, doi:10.1080/10618600.1998.10474787.

Brynjarsdóttir, J., and A. O'Hagan (2014), Learning about physical parameters: the importance of model discrepancy, *Inverse Problems*, *30*(11), 114,007–114,031, doi:10.1088/0266-5611/30/11/114007.

Büchele, B., H. Keibrich, A. Kron, A. Thieken, J. Ihringer, P. Oberle, B. Merz, and F. Nestmann (2006), Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks, *Natural Hazards and Earth System Sciences*, *6*(4), 485–503, doi:10.5194/nhess-6-485-2006.

Capart, H. (2013), Analytical solutions for gradual dam breaching and downstream river flooding, *Water Resources Research*, *49*(4), 1968–1987, doi:10.1002/wrcr.20167.

Chinnarasri, C., and A. Saelim (2009), Semi-parametric model of an overtopped dike breach, *Water Management*, *162*(4), 289–296, doi:10.1680/wama.2009.00047.

Chinnarasri, C., S. Jirakitlerd, and S. Wongwises (2004), Embankment dam breach and its outflow characteristics, *Civil Engineering and Environmental Systems*, *21*, 247–264, doi:10.1080/10286600412331328622.

Coleman, S. E., D. P. Andrews, and M. G. Webby (2002), Overtopping breaching of non-cohesive homogeneous embankments, *Journal of Hydraulic Engineering*, *128*(9), 829–838, doi:10.1061/(ASCE)0733-9429(2002)128:9(829).

De Lorenzo, G., and F. Macchione (2014), Formulas for the peak discharge from breached earthfill dams, *Journal of Hydraulic Engineering*, *140*(1), 56–67, doi:10.1061/(ASCE)HY.1943-7900.0000796.

Di Baldassarre, G., G. Schumann, P. D. Bates, J. E. Freer, and K. J. Beven (2010), Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches, *Hydrological Sciences Journal*, *55*(3), 364–376, doi:10.1080/02626661003683389.

Efron, B. (1979), Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*, *7*(1), 1–26, doi:10.1214/aos/1176344552.

Eicher, A. (2014), Bayesian multilevel model calibration of a simplified dam breach model, Master's thesis, ETH Zurich, Switzerland.

Escuder-Bueno, I., G. Mazzà, A. Morales-Torres, and J. T. Castillo-Rodríguez (2016), Computational aspects of dam risk analysis: Findings and challenges, *Engineering*, *2*(3), 319–324, doi:10.1016/J.ENG.2016.03.005.

Faeh, R. (2007), Numerical modeling of breach erosion of river embankments, *Journal of Hydraulic Engineering*, *133*(9), 1000–1009, doi:10.1061/ASCE0733-94292007133:91000.

Foster, M., R. Fell, and M. Spannagle (2000), The statistics of embankment dam failures and accidents, *Canadian Geotechnical Journal*, *37*(5), 1000–1024, doi:10.1139/t00-030.

Franca, M. J., and A. B. Almeida (2005), Modeling of washout of dams, *Journal of Hydraulic Reserach*, *43*(4), 439–444, doi:10.1080/00221680509500141.

Frank, P.-J. (2016), Hydraulics of spatial dike breaches, Ph.D. thesis, ETH Zurich, Switzerland, doi:10.3929/ethz-a-010803310.

Fread, D. L. (1984), A breach erosion model for earthen dams, in *Proceedings of Specialty Conference on Delineation of Landslides, Flash Flood, and Debris Flow Hazards in Utah*, Utah State University, Logan, Utah.

Froehlich, D. C. (2008), Embankment dam breach parameters and their uncertainties, *Journal of Hydraulic Engineering*, *134*(12), 1708–1721, doi:10.1061/ASCE0733-94292008134: 121708.

Froehlich, D. C. (2016), Predicting peak discharge from gradually breached embankment dam, *Journal of Hydrologic Engineering*, *21*(11), doi:10.1061/(ASCE)HE.1943-5584. 0001424.

Gelman, A. (2014), *Bayesian Data Analysis*, Texts in Statistical Science, 3. ed., Chapman & Hall/CRC.

Hintze, J. L., and R. D. Nelson (1998), Violin plots: A box plot-density trace synergism, *The American Statistician*, *52*(2), 181–184, doi:10.1080/00031305.1998.10480559.

Hooshyaripor, F., and A. Tahershamsi (2012), Comparing the performance of neural networks for predicting peak outflow from breached embankments when back propagation algorithms meet evolutionary algorithms, *International Journal of Hydraulic Engineering*, *1*(6), 55–67, doi:10.5923/j.ijhe.20120106.01.

Kalyanapu, A. J., S. Shankar, E. R. Pardyjak, D. R. Judi, and S. J. Burian (2011), Assessment of gpu computational enhancement to a 2d flood model, *Environmental Modelling & Software*, *26*, 1009–1016, doi:10.1016/j.envsoft.2011.02.014.

Kennedy, M. C., and A. O'Hagan (2001), Bayesian calibration of computer models, *Journal of the Royal Statistical Society: Series B*, *63*(3), 425–464, doi:10.1111/1467-9868.00294.

Kühne, A. (1978), Charakteristische Kenngrössen Schweizerischer Speicherseen, *Geographica Helvetica*, *33*(4), 191–199, [in German].

Lacasta, A., M. Morales-Hernández, J. Murillo, and P. García-Navarro, P.avarro (2014), An optimized gpu implementation of a 2d free surface simulation model on unstructured meshes, *Advances in Engineering Software*, *78*, 1–15, doi:10.1016/j.advengsoft.2014.08. 007.

Laloy, E., and J. A. Vrugt (2012), High-dimensional posterior exploration of hydrologic models using multiple-try dream(zs) and high-performance computing, *Water Resources Research*, *48*(1), doi:10.1029/2011WR010608.

Ma, H., and X. Fu (2012), Real time prediction approach for floods caused by failure of natural dams due to overtoping, *Advances in Water Resources*, *35*, 10–19, doi:10.1016/j. advwatres.2011.08.013.

Macchione, F. (2008), Model for predicting floods due to earthen dam breaching. i: Formulation and evaluation, *Journal of Hydraulic Engineering*, *134*(12), 1688–1696, doi: 10.1061/ASCE0733-94292008134:121688.

Machiels, O., S. Erpicum, B. Dewals, and M. Pirotton (2011), Theoretical and numerical analysis of the influence of the bottom friction formulation in free surfce flow modelling, *Water SA*, *37*(2), 221–228, doi:10.4314/wsa.v37i2.65867.

Mansanarez, V., J. Le Coz, B. Renard, M. Lang, G. Pierrefeu, and P. Vauchel (2016), Bayesian analysis of stage-fall-discharge rating curves and their uncertainties, *Water Resources Research*, *52*(9), 7424–7443, doi:10.1002/2016WR018916.

Morris, M. W., and M. Hassan (2005), Conclusions and recommendations from impact project wp5: Combined risk and uncertainty, *Tech. rep.*, European Commission.

Morris, M. W., G. Hanson, and M. Hassan (2008), Improving the accuracy of breach modelling: why are we not progressing faster?, *Journal of Flood Risk Management*, *1*(3), 150–161, doi:10.1111/j.1753-318X.2008.00017.x.

Nagel, J. B., and B. Sudret (2015), Bayesian multilevel model calibration for inverse problems under uncertainty with perfect data, *Journal of Aerospace Information Systems*, *12*(1), 97–113, doi:10.2514/1.I010264.

Nagel, J. B., and B. Sudret (2016), A unified framework for multilevel uncertainty quantification in bayesian inverse problems, *Probabilistic Engineering Mechanics*, *43*, 68–84, doi:10.1016/j.probengmech.2015.09.007.

Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994), Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, *263*(5147), 641–646.

Pappenberger, F., and K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resources Research*, *42*(5), doi:10.1029/2005WR004820.

Peter, S. J. (2017), Dam break analysis under uncertainty, Ph.D. thesis, ETH Zurich, Switzerland, doi:10.3929/ethz-b-000209879.

Pickert, G., V. Weitbrecht, and A. Bieberstein (2011), Breaching of overtopped river embankments controlled by apparent cohesion, *Journal of Hydraulic Reserach*, *49*(2), 143–156, doi:10.1080/00221686.2011.552468.

Pilotti, M., M. Tomirotti, G. Valerio, and B. Bacchi (2010), Simplified method for the characterization of the hydrograph following a sudden partial dam break, *Journal of Hydraulic Engineering*, *136*(10), 693–704, doi:10.1061/ASCEHY.1943-7900.0000231.

Reguly, I. Z., E. László, G. R. Mudalige, and M. B. Giles (2015), Vectorizing unstructured mesh computations for many-core architectures, *Concurrency and Computation: Practice and Experience*, *28*(2), 557–577, doi:10.1002/cpe.3621.

Roberts, G. O., and J. S. Rosenthal (2001), Optimal scaling for various metropolis-hastings algorithms, *Statistical Science*, *16*(4), 351–367, doi:10.1214/ss/1015346320.

Santner, T. J., B. J. Williams, and W. I. Notz (2003), *The Design and Analysis of Computer Experiments*, Springer, New York, doi:https://doi.org/10.1007/978-1-4757-3799-8.

Schmocker, L., and W. H. Hager (2012), Plane dike-breach due to overtopping: effects of sediment, dike height and discharge, *Journal of Hydraulic Reserach*, *50*(6), 576–586, doi: 10.1080/00221686.2012.713034.

Schnitter, N. J. (1994), *A History of Dams - The useful pyramids*, Aa Balkema, Rotterdam.

Schöniger, A., T. Wöhling, L. Samaniego, and W. Nowak (2014), Model selection on solid ground: Rigorous comparison of nine ways to evaluate bayesian model evidence, *Water Resources Research*, *50*(12), doi:10.1002/2014WR016062.

Shi, X., M. Ye, G. P. Curtis, G. L. Miller, P. D. Meyer, M. Kohler, S. Yabusaki, and J. Wu (2014), Assessment of parametric uncertainty for groundwater reactive transport modeling, *Water Resources Research*, *50*(5), 4416–4439, doi:10.1002/2013WR013755.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, *Monographs on Statistics & Applied Probability*, vol. 26, Chapman & Hall/CRC.

Singh, V. P. (1996), *Dam Breach Modeling Technology*, Kluwer Academic Publishers, doi: 10.1007/978-94-015-8747-1.

Singh, V. P., and P. D. Scarlatos (1988), Analysis of gradual earth-dam failure, *Journal of Hydraulic Engineering*, *114*(1), 21–42, doi:10.1061/(ASCE)0733-9429(1988)114:1(21).

Steinbakk, G. H., T. L. Thorarinsdottir, T. Reitan, L. Schlichting, S. Hølleland, and K. Engeland (2016), Propagation of rating curve uncertainty in design flood estimation, *Water Resources Research*, *52*(9), 6897–6915, doi:10.1002/2015WR018516.

ter Braak, C. J. F. (2006), A markov chain monte carlo version of the genetic algorithm differential evolution: easy bayesian computing for real parameter spaces, *Statistics and Computing*, *16*(3), 239–249, doi:10.1007/s11222-006-8769-1.

Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using bayesian total error analysis, *Water Resources Research*, *45*(12), doi:10.1029/2008WR006825.

Volz, C. (2013), Numerical simulation of embankment breaching due to overtopping, Ph.D. thesis, ETH Zurich, Switzerland, doi:10.3929/ethz-a-009967594.

Volz, C., P.-J. Frank, D. F. Vetsch, W. H. Hager, and R. M. Boes (2017), Numerical embankment breach modelling including seepage flow effects, *Journal of Hydraulic Reserach*, doi:10.1080/00221686.2016.1276104.

Vonwiller, L., D. F. Vetsch, S. J. Peter, and R. M. Boes (2015), Methode zur Beurteilung des maximalen Breschenabflusses bei progressivem Bruch homogener Erdschüttdämme an kleinen Stauhaltungen, *Wasser Energie Luft*, *107*(1), 37–43, [in German].

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Hidgon (2009), Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *International Journal of Nonlinear Sciences & Numerical Simulation*, *10*(3), 273–290, doi:10.1515/IJNSNS.2009.10.3.273.

Wahl, T. L. (1998), Prediction of embankment dam breach parameters - a literature review and needs assessment, *Tech. rep.*, U.S. Departement of the Interior, Bureau of Reclamation, Denver.

Wahl, T. L. (2004), Uncertainty of preedictions of embankment dam breach parameters, *Journal of Hydraulic Engineering*, *130*(5), 389–397, doi:10.1061/ASCE0733-94292004130:5389.

Walder, J. S., and J. E. O'Connor (1997), Methods for predicting peak discharge of floods caused by failure of natural and constructed dams, *Water Resources Research*, *33*(10), 2337–2348, doi:10.1029/97WR01616.

Walder, J. S., R. M. Iverson, J. W. Godt, M. Logan, and S. A. Solovitz (2015), Controls on the breach geometry and flood hydrograph during overtopping of noncohesive earthen dams, *Water Resources Research*, *51*(8), 6701–6724, doi:10.1002/2014WR016620.

Wang, Z., and D. S. Bowles (2006), Three-dimensional non-cohesive earthen dam breach model. part 1: Theory and methodology, *Water Resources*, *29*(10), 1528–1545, doi:10.1016/j.advwatres.2005.11.009.

Wu, F.-C., and C. C. Chen (2009), Bayesian updating of parameters for a sediment entrainment model via markov chain monte carlo, *Journal of Hydraulic Engineering*, *135*(1), 22–37, doi:10.1061/(ASCE)0733-9429(2009)135:1(22).

Wu, W. (2008), *Computational River Dynamics*, Taylor & Francis Group, London, doi:10.4324/9780203938485-2.

Wu, W., Y. Kang, and S. S. Y. Wang (2009), An earthen embankment breach model, in *33rd: IAHR Congress: Water Engineering for a Sustainable Environment*, pp. 3920–3926, International Association of Hydraulic Engineering and Research.

Xiu, D., and G. E. Karniadakis (2002), The wiener–askey polynomial chaos for stochastic differential equations read more: http://epubs.siam.org/doi/abs/10.1137/s1064827501387826?journalcode=sjoce3, *SIAM Journal on Scientific Computing*, *24*(2), 619–644, doi:https://doi.org/10.1137/S1064827501387826.

Zenz, G., and M. Goldgruber (Eds.) (2013), *12th International benchmark workshop on numerical analysis of dams*, ICOLD, Austrian National Commission on Large Dams (ATCOLD).