Fluixá-Sanmartín, J.; Altarejos-García, L.; Morales-Torres, A.; Escuder Bueno, I. (2019). Empirical Tool for the Assessment of Annual Overtopping Probabilities of Dams. Journal of Water Resources Planning and Management. 145(1):1-12. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001017

Empirical Tool for the Assessment of Annual Overtopping Probabilities of Dams

Javier Fluixá-Sanmartín[1]*, Luis Altarejos-García[2], Adrián Morales-Torres[3], Ignacio Escuder-Bueno[4]

[1] *M.Sc. in Hydraulic Engineering and Environment, Centre de Recherche sur l'Environnement Alpin (CREALP), Sion, Switzerland.*
[2] *Ph.D. in Civil Engineering, Department of Civil Engineering, Universidad Politécnica de Cartagena (UPCT), Cartagena, Spain.*
[3] *Ph.D. in Civil Engineering, iPresas Risk Analysis, Valencia, Spain.*
[4] *Ph.D. in Civil Engineering, Department of Hydraulic Engineering and Environment. Universitat Politècnica de València (UPV), Valencia, Spain.*
**Corresponding author. E-mail:* javier.fluixa@crealp.vs.ch. Other authors' emails: luis.altarejos@upct.es; adrian.morales@ipresas.com; iescuder@hma.upv.es

Abstract:

This work presents a simple tool for the assessment of maximum overtopping probabilities of dams. The tool is based on empirical relations between the overtopping probability and the basic hydrological and hydraulic characteristics of the dam-reservoir system: the Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}, both weighted with the relative importance of the 1,000-year flood. The surface issued from the tool represents the limit above which no $V^*_F - Q^*_{Cap}$ combination is statistically expected to offer a higher probability. The tool was calibrated using the detailed overtopping models of 342,233 synthetic cases generated from 30 existing dams and then validated against a set of 21 independent cases. The tool is useful when analyzing a portfolio of dams in previous screening phases of dam risk analysis. It aims at identifying overtopping as a relevant failure mode and easily classifying each dam in terms of its overtopping probability. The tool is also a support for the definition and prioritization of corrective measures since it assesses their impact in the overtopping probability reduction.

INTRODUCTION

Large dams and protective dikes and levees are critical infrastructure whose associated risk must be assessed and managed. Risk analysis techniques are useful tools to help owners make decisions in the field of dam safety management. The development and application of these techniques helps the decision-making process by optimizing the existing resources and pointing at the most efficient ways of using them (ANCOLD 2003; ICOLD 2005; SPANCOLD 2012; USACE 2011).

The analysis of failure modes is a key process in risk analysis (FEMA 2015). The goal is to identify, describe and structure the possible ways in which the dam may fail. Techniques such as Failure Modes Effects and Criticality Analysis (FMECA) help systematize this process and ranking failure modes depending on their probability of occurrence. To produce meaningful results using risk analysis it is important to capture all those potential failure modes whose occurrence probability is deemed to be not irrelevant. A review of dam safety information during a first screening stage is necessary to assess whether a potential failure mode should be incorporated into the process or not.

Among the potential failure modes of a dam, overtopping is of particular importance. It has been the main cause of nearly 30% of all dam failures worldwide over the last century (ICOLD 1995; Jandora et al. 2008). Overtopping happens when the water level in the reservoir reaches the dam crest and starts spilling over the downstream face of the dam. Overtopping can be caused by one or a combination of several of the following reasons: floods, spillway gates failure, wind or earthquake triggered waves, dam crest settlement due to earthquakes and/or internal erosion in the case of earth dams, and reservoir operation errors. This work is focused on the overtopping caused by floods, which is the most frequent reason for overtopping (Costa 1985; Wahl 1998).

In this paper it is understood that overtopping is produced when the dam's outlet works are unable to successfully route the incoming flood, so the water level in the reservoir rises until the dam crest is reached and water starts spilling over the dam. This endangers the structure itself, as the flow overrunning on the downstream slope possess a certain erosive potential. The probability of a failure by overtopping is the combination of two components: (i) the probability of occurrence of the triggering event, i.e. the overtopping of the dam; and (ii) the conditional probability of the structural failure once the overtopping begins.

In order to determine whether this is a relevant failure mode that deserves to be considered deeply in a risk analysis, it is necessary to previously assess the probability of occurrence of the overtopping, p_{over}. This probability is defined as the probability that a spill over the dam crest takes place in any year, regardless of its magnitude or duration, and is expressed in inverse time units.

Several approaches of different complexity can be followed to define the probability of overtopping occurrence (Kuo et al. 2007). Typically, studies use dam-specific models based on the routing of inflow hydrographs to compute the maximum reservoir water level. A detailed analysis of overtopping requires comprehensive information on hydrology, hydraulics, characteristics of dam, spillways and reservoir and the use of routing methods (Michailidi and Bacchi 2017). In order to address the uncertainty related to hydrologic and hydraulic factors, the use of probabilistic techniques such as Monte Carlo simulation (Goodarzi et al. 2014; Kwon and Moon 2006) or Latin hypercube sampling (Hsu et al. 2011) is widespread due to their conceptual simplicity and robustness. These techniques provide results with higher accuracy but at the expense of a larger computational cost (Kuo et al. 2007; Sun et al. 2012; Wang and Zhang 2017).

However, it is not always advisable to perform analyses with a high level of detail (SPANCOLD 2012) but to look for a balance between accuracy and effort. Depending on the scope of the study or the stage of the process it may be reasonable to employ a less time and cost demanding methodology. This is especially useful when analyzing a portfolio of dams, and in particular in the selection of the failure modes that have to be incorporated in subsequent stages of the risk analysis. In these cases, the use of screening level techniques such as Screening Portfolio Risk Analyses (SPRA) may be more appropriate. Performed at an initial stage of portfolio risk assessment (Bowles et al. 1998), SPRA aim at identifying relevant failure

modes and classifying each dam in a general and simplified manner, in terms of their potential failure risk (FEMA 2015). SPRA techniques are based primarily on available information. They allow to prioritize additional investigations for the dams at a higher risk. These techniques have been applied in several cases (De Cea et al. 2010; Escuder-Bueno et al. 2008; Jeffrey T. McClenathan and Andy Harkness 2008; USBR and USACE 2011).

The objective of this work is to provide a simple tool for the estimation of dam overtopping probability that does not require the elaboration of *ad hoc* models but that is yet able to provide results on the conservative side and with an adequate level of accuracy. Such a tool should be easy to apply and make use of main characteristics of spill and storage capacity, as well as on hydrological loading, readily available for any dam owner in the dam-reservoir technical information files.

The conceptual approach followed in the paper is as follows. First, a comprehensive, detailed model that is able to accurately estimate the probability of overtopping is described. This detailed overtopping model is of widespread use in the dam engineering field, and has been used by the authors in more than 40 dam safety analyses in the last 15 years (Escuder-Bueno et al. 2016; Morales-Torres et al. 2016; Serrano-Lombillo et al. 2017; Setrakian-Melgonian et al. 2017). This model provides an estimation of the overtopping probabilities that will be used as reference probabilities against which the tool will be tested. Using the information of 30 existing Spanish dam-reservoir systems as raw data, and varying their properties, a large set of more than 5.6×10^5 virtual dam-reservoir systems is generated. Then, using the detailed model, the probability of overtopping for each case is estimated. In the next part, the structure and formulation of the proposed simple model is described and justified, and the set of 5.6×10^5 overtopping probabilities is used to calibrate the parameters of the model. Finally, the calibrated tool is tested using some real dam-reservoir systems as benchmark cases.

DETAILED MODEL FOR THE ASSESSMENT OF ANNUAL OVERTOPPING PROBABILITIES

For the calculation of the annual probability of overtopping of a dam, overtopping probability models were used, henceforth referred as overtopping models. These are based on risk models, which complete description can be found in SPANCOLD (2012). This section describes the structure of the typical detailed model used in this work.

Usually, such modeling is performed using event trees, an exhaustive representation of all the possible chains of events –represented by nodes– that can produce the dam overtopping (Serrano-Lombillo et al. 2011). The tree's branches represent all the possible outcomes of their event of origin and have an assigned probability. Any path between the initiating node and each of the nodes of the tree represents one of the possible outcomes that might result from the original event and can be calculated as the product of the probabilities associated with each branch.

These models can be represented using influence diagrams such as the one presented in **Fig. 1**. Each variable is represented by a node and each relationship by an arrow. In this case, the overtopping model used is composed of 5 nodes: *Floods*, *Previous Reservoir Level*, *Spillways reliability*, *Flood routing* and *Overtopping probability*. These nodes contain information on the characteristics of the dam-reservoir system, the hydrological loads and how they are routed. This information flows from one node to the next until the last one in which the overtopping probabilities are finally calculated. For this work, the *iPresas* software (iPresas 2017) has been used to calculate these probabilities.

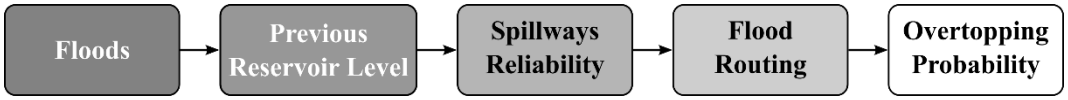

Fig. 1. *Structure of the detailed overtopping model used.*

Node Floods (hydrological loads)

Floods are the initiating event (node) and the probabilities of the emerging branches are defined by the probability linked to the inflow hydrographs, introducing the temporal component to the overtopping probability. These are associated with a given return period (T) or its equivalent annual exceedance probability, (World Meteorological Organization 2008). In this work, the hydrographs of return periods up to 100,000 years have been extracted from the dam-reservoir technical information files and then incorporated to each *Floods* node. In the cases where there is no hydrograph available with such high return periods, it has been necessary to resort to statistical extrapolation through distribution functions of the peak discharges (Chow et al. 2008).

Node Previous Reservoir Level

The water level in the reservoir preceding the arrival of a flood varies and can be determined through water level frequency distribution curves. The variety of possible cases makes it difficult to capture through a single variable the effect of these curves on the overtopping probability.

A homogeneous criterion has to be considered to establish this Previous Reservoir Levels. Consequently, the initial reservoir water level when routing begins will be the Normal Operating Level (NOL) of the reservoir, which is defined as the maximum level reached at the reservoir under normal conditions. Under this assumption the results stay on the conservative side since overtopping probabilities will be overestimated.

Node Spillways reliability

Spillways reliability is of great relevance for dam overtopping as a low reliability increases the probability of reaching high levels in the reservoir, which increases overtopping probability (Fiedler 2010; Kuo et al. 2008). The information that this node includes is the probability that every spillway behaves properly, that is, that in the moment of the arrival of the flood the spillway can be used at its full capacity. It is considered a binary operation condition where each gate of a spillway works perfectly -no failure- or does not work at all -failure- (SPANCOLD 2012). Reliability is simply defined as the probability to be in the no failure state.

The quantitative individual reliability of each outlet has been related to a qualitative description of the condition of the spillway gate system which is based on standard cases used in dam risk analysis as shown in **Table 1**, without being necessary to resort to detailed studies such as fault trees (Escuder-Bueno and González-Pérez 2014).

Table 1. *Standard individual reliabilities of the spillway gates.*

Case	Reliability
Non-gated	100%
New / Very well maintained	95%
Well maintained, some minor problem	85%
Some problem	75%
Unreliable	50%
Not reliable at all / not used	0%

Node Flood routing

Another aspect to consider is the flood routing strategy through the controlled outlets and spillways of the dam. This will determine how the inflows are managed, and consequently the resulting water level in the reservoir. Only the use of spillways is considered here since their capacity and impact on overtopping probabilities is in general much higher than that of the rest of outlets usually employed in the ordinary operation of a dam (USBR and USACE 2011), like bottom outlets. In addition, it has been admitted as homogeneous management criterion for flood routing to make output spills equal to input discharges as much as this is possible (USBR and USACE 2015).

Node Overtopping probability

The last node serves to evaluate overtopping probability. At the end of each branch of the event tree a certain overtopping may or may not have happened. The overtopping probability is calculated as the sum of the probabilities of all those branches where a certain overtopping has actually happened. Neither the overtopping magnitude nor its duration is considered here.

GENERATION OF DAM-RESERVOIR SYSTEMS DATA

In order to have a set of overtopping probabilities covering a wide range of most practical, real-world dam-reservoir systems, the model described above has been used to estimate the overtopping probabilities of a large set of virtual dam-reservoir systems. The objective is to use these probabilities in the definition and calibration process of the simple tool that is searched.

The process starts with 30 real dam-reservoir systems located in Spain. Their main characteristics are presented in **Table S1** of the Supplemental Material. Then this initial set is expanded by changing the key parameters of each system, thus artificially increasing the number of cases. For this purpose, some of the parameters of these 30 dams were systematically modified:

- For those dams with gated spillways, the number of spillways has been increased in -1, +1, +0 and +2, and their length has been multiplied by 0.25, 0.5, 1, 1.5, 2 and 3.
- Different individual reliabilities for these gates: 50%, 75%, 85%, 95% and 100% have been considered.
- For those dams with non-gated spillways, their width has been multiplied by 0.25, 0.5, 1, 1.5, 2 and 3.
- The characteristic relation *Water height vs Volume* of each reservoir has been modified by multiplying their capacity by 0.5, 1, 1.5 and 2.
- The input hydrographs have been modified by multiplying their duration by 0.5, 1, 1.5 and 2, as well as by multiplying their flow rates by 0.5, 1, 1.5 and 2.

The combination of each of these variations for the 30 dams resulted in a total of 561,024 synthetic cases. Even though these cases do not correspond to the real dams, the results are considered valid approximations since the calculation process of overtopping probabilities is logical and consistent.

Once these virtual dam-reservoir systems have been generated, their overtopping probabilities have been calculated using the model described in the previous section. Each dam-reservoir system has been solved obtaining an overtopping probability value, p_{over}. Additionally, the following parameters have been calculated for each problem solved: freeboard storage capacity, V_F; number of existing gates, n, and their reliability, Rel; the release capacity of the gated spillways at the Normal Operating Level, Q_{G_NOL}, at the Design Level, Q_{G_DL}, and at the dam crest, Q_{G_DC}; the corresponding release capacities for non-gated spillways: Q_{NG_NOL}, Q_{NG_DL}, Q_{NG_DC}; and the total volumes and peak flows of the inflow hydrographs (e.g. V_{IN,T_100} represents the volume of the 100-year return period flood, and Q_{P,T_100} is its peak flow). All these parameters are described in detail in the next section. It is worth noting that, in this context, the "Design level" refers to the maximum level reached in the reservoir when the Design Flood is routed, with the following assumptions: (1) initial reservoir level is maximum operating level, and (2) all the spillways and outlets are available and perform without problems.

For the calibration of the tool, it has been decided not to include those cases in which the probability of overtopping is null, since its influence would be masked. Thus, 342,233 synthetic cases are retained and have been used to adjust the simplified model. This amount of cases is large enough to consider the developed tool as statistically significant and robust.

DEFINITION OF THE TOOL FOR THE ASSESSMENT OF ANNUAL OVERTOPPING PROBABILITIES

With the purpose of defining a tool (or simplified model) that estimates the annual probability of overtopping, the characteristics that best capture the dam's behavior against overtopping occurrence have to be identified. These characteristics are the basic components of the formulation of the tool that will be calibrated to adjust to the overtopping probabilities calculated with the detailed overtopping models.

Main variables characterizing the overtopping probability

Hydrological conditions, and particularly flood characteristics, are driving forces in dam overtopping (Bowles et al. 1998). Its occurrence mainly depend on two connected dam-reservoir characteristics: insufficient flood storage capacity and inadequate spillway capacity (Lee and You 2013; USBR 2014). The works of Serrano-Lombillo et al. (2012) and Altarejos-García et al. (2013) highlighted a direct relation between these parameters and the overtopping probability. It is proposed that the tool relies on these interactions and uses the mentioned characteristics of the dam-reservoir system to assess the overtopping probabilities.

On the one hand, the reservoir freeboard volume represents the capacity of the reservoir for the storage of exceeding incoming volumes. It can be defined as the volume available between the previous water level at the beginning of a flood and a characteristic level which, for this particular overtopping analysis, is assumed to be the dam crest (DC). Since the Normal Operating Level (NOL) is the maximum level reached at the reservoir under normal conditions, adopting it as the previous water level leads the results to stay on the conservative side. The freeboard volume V_F has thus been defined as the volume between the NOL and the DC.

On the other hand, the outlet release capacity refers to the maximum discharge that the outlets are able to assure. As stated before, only the spillways are considered. This capacity depends on the water level at the reservoir, on the dimensions of the spillway, and on the regulating gates, if there are any, and their availability. It is of interest, nonetheless, to characterize this capacity at a specific reservoir level. Among the possible levels that could be used, the following three are of special relevance: the NOL, the Design Level (DL), and the DC. In some cases, especially for non-gated spillways, the spillway capacity at the NOL is null and thus should not be adopted as the representative level. Regarding the DL, it may capture the capacity of spillways under normal conditions but not under extreme situations such as overtopping events. Therefore, it has been decided to use the DC to characterize the spillways release capacity Q_{Cap}. A numerical justification of this selection is presented below.

Since some spillways may be controlled by gates, it has been decided to distinguish between the capacity of non-gated spillways, $Q_{Cap,NG}$, and of gated spillways, $Q_{Cap,G}$. The objective is to capture the effect of gates' availability, which is affected by the number of gates, n, controlling the discharge and their individual reliability, Rel. For instance, a spillway with a lower reliability increases the probability of reaching higher reservoir levels, which increases the overtopping probability. In a spillway with N gates there are certain probabilities of having 0, 1, 2… N gates working. If all gates work independently with a given individual reliability, a binomial distribution of parameters n and Rel can be used to compute the probability of each of them working (SPANCOLD 2012). For this work, the mean of the binomial distribution is taken as a single factor to correct individual gated spillways' capacities. The mean of the binomial distribution is expressed as $n \times Rel$, and represents the average number of available gates working. The capacity will worsen as the number of gates and their reliability decrease. Moreover, a dam may be equipped with different types of gated spillways, with different reliabilities and discharge capacities. In the expression of the total capacity of the gated spillways, the correction factor must be separately applied to the different spillway capacities. However, the formulation of the tool does not contemplate combinations of reliabilities and thus for each dam all spillway gates are assumed to have the same reliability.

The total release capacity Q_{Cap} is then expressed as:

$$Q_{Cap} = Q_{Cap,NG} + Rel \cdot \sum_i (Q_{Cap,G\,i} \cdot n_i) \quad (1)$$

where $Q_{Cap,NG}$ is the total capacity of all the non-gated spillways at the DC level; *Rel* is the individual reliability of the spillway gates of the dam; *i* is the index designating the type of gated spillway; $Q_{Cap,G\,i}$ is the individual discharge capacity through one gate of the spillway *i*; n_i is the number of gates of the spillway *i*.

In order to better capture these situations, the set of cases has been divided into 5 groups, one for each gate reliability. Among the total of cases, 52,021 correspond to the reliability of 100% that includes the non-gated spillways, and 72,553 cases to each of the 50%, 75%, 85% and 95% reliabilities. The non-gated cases are fewer than the rest because this assumption entails more null probabilities and, as stated above, those cases in which the probability of overtopping is null must be removed for the calibration of the tool.

The variables V_F and Q_{Cap} have to be weighted with the relative importance of the hydrologic input load. A given freeboard will not have the same impact on the overtopping probability of a reservoir designed to absorb floods with similar volumes than for flood volumes significantly higher. The same principle applies for a given release capacity, that has to be weighted with the flood peak inflow.

In this study the Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}, are proposed to capture the characteristics of the dam-reservoir system. The V^*_F is defined as the ratio between the freeboard storage capacity V_F and the volume of a representative flood $V_{IN,T}$ with a given return period T [Eq. (2)], while the Q^*_{Cap} is the ratio between the outlet capacity Q_{Cap} and the peak inflow associated to this flood $Q_{P,T}$ [Eq. (3)]. For this work, the flood corresponding to a return period of 1,000 years has been taken as representative for the characterization of the tool, as justified below.

$$V^*_F = \frac{V_F}{V_{IN,T=1,000}} \quad (2)$$

$$Q^*_{Cap} = \frac{Q_{Cap}}{Q_{P,T=1,000}} = \frac{Q_{Cap,NG} + Rel.\sum_i (Q_{Cap,G\,i} \cdot n_i)}{Q_{P,T=1,000}} \quad (3)$$

Key relations between the main characteristics of the dam-reservoir system and the overtopping probability

Fig. 2 shows the projection on the coordinated planes of the 342,233 virtual dam-reservoir systems solved with the detailed overtopping model and grouped by their gates reliability. The values of the annual overtopping probability, plotted in the vertical z-axis, as a function of Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}, are represented. For illustrative purposes, the z-axis is plotted on a logarithmic scale to better appreciate the order of magnitude of its values, and the x-axis (V^*_F) and the y-axis (Q^*_{Cap}) are not represented at the same scale.

Among the 342,233 total cases, a large span of variables are treated, covering the following ranges:

- V^*_F: [0 – 2.1]
- Q^*_{Cap}: [0 – 46]
- p_{over}: [2.64×10^{-15} – 1]

Although the clouds of points obtained indicate a certain dispersion of results, a clear relationship between these variables stands out: the greater the value of the freeboard storage capacity parameter, the smaller the overtopping probability. Similarly, the probability decreases with the release capacity parameter. As expected, the results indicate that for larger freeboards and release capacities, the dam will likely be able to absorb the incoming floods with less risk of overtopping.

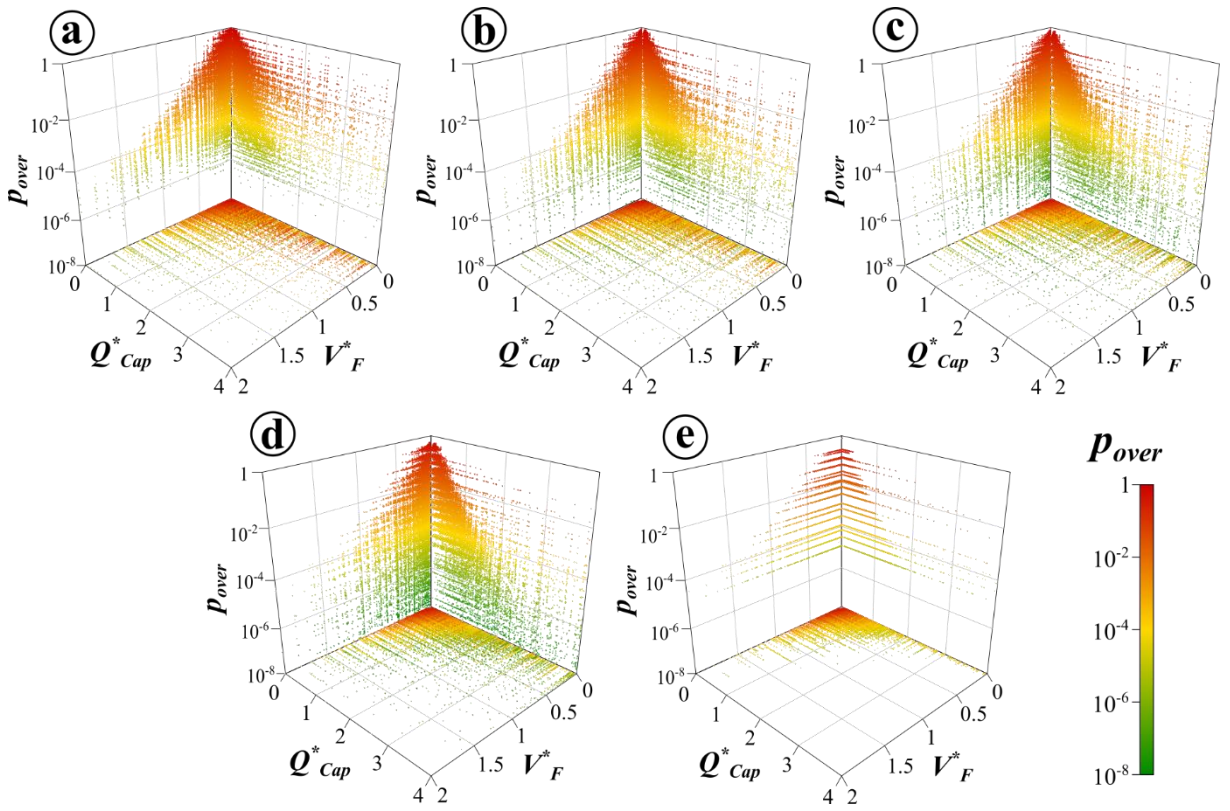

Fig. 2. *Projection of overtopping probability (p_{over}) results with the Unit Storage Capacity (V^*_F) and the Unit Spillway Capacity (Q^*_{Cap}) on the 3 coordinated planes, grouped by individual gates reliability:* ***(a)*** *50%;* ***(b)*** *75%;* ***(c)*** *85%;* ***(d)*** *95%; and* ***(e)*** *100% (non-gated).*

From **Fig. 2**, an upper limit of the overtopping probability clearly exists, above which no more cases substantially emerge. This allows to find a relation between the Unit Storage Capacity, V^*_F, the Unit Spillway Capacity, Q^*_{Cap}, and the maximum annual overtopping probability, $p_{over,max}$, that corresponds to a threshold above which a certain value of p_{over} will likely not be exceeded. This will lead us to stay on the conservative side when estimating the probability. The objective is to find this relation, defined as:

$$p_{over,max} = f\left(V^*_F\,,\ Q^*_{Cap}\right) \tag{4}$$

To overcome the uncertainty related to the dispersion of points, a statistically significant tool that comprehends the variety of cases that can be given by any dam-reservoir system is needed. The expression of the tool [Eq. (4)] should represent the upper envelope that covers the maximum number of points.

For the expression of this three-dimensional surface, the relations between the different variables are analyzed. **Fig. 3** shows an example of the density of points projected on the 3 coordinated planes for a gates reliability of 95% (similar results are obtained for the rest of reliabilities). These densities have been obtained by applying the Kernel Density Estimation technique (Parzen 1962; Rosenblatt 1956). Continuous lines correspond to the intersection of the adjusted surface with the 3 different planes: a) $Q^*_{Cap} = 0$; b) $V^*_F = 0$; c) $p_{over} = 10^{-3}$.

From these plots it can be stated that:

- The envelope curve of the Q^*_{Cap} and V^*_F variables for a given p_{over} appears to be a straight line.
- The logarithm of $p_{over,max}$ can be expressed as the logarithm of the combination of the two other variables, Q^*_{Cap} and V^*_F.
- By definition, the value of $p_{over,max}$ is limited to 1 and its logarithm is limited to 0.
- A cloud of points exceeding any adjusted surface is persistently detected for a range of low V^*_F values [**Fig. 3 (c)**], in particular for the gates reliability of 50%, 75%, 85% and 95%.

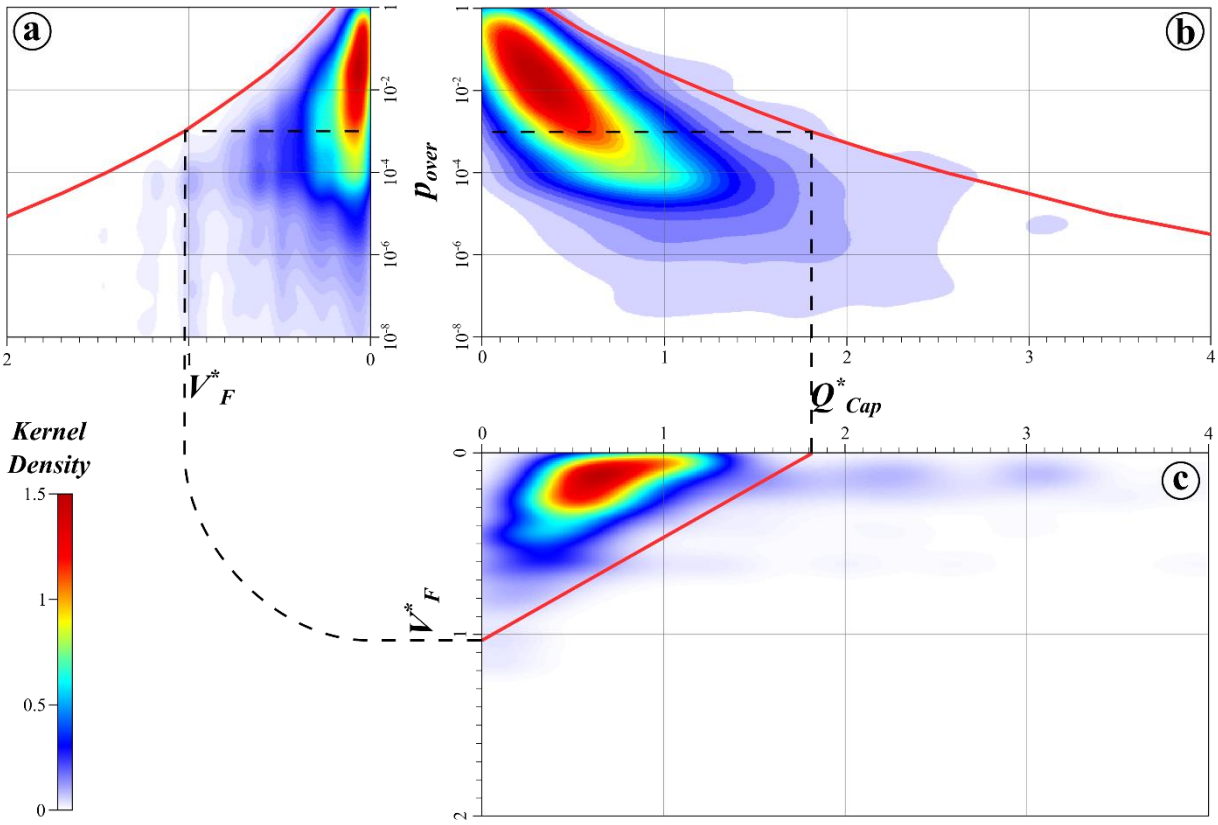

Fig. 3. *Density plot of the projected points on the 3 coordinated planes for a gates reliability of 95%:* ***(a)*** *V^*_F vs p_{over};* ***(b)*** *Q^*_{Cap} vs p_{over};* ***(c)*** *Q^*_{Cap} vs V^*_F for a given p_{over} of 10^{-3}.*

Taking into account the previous considerations, an expression of the Eq. (4) is proposed as follows:

$$log_{10}(p_{over,max}) = min\left(0; a \cdot log_{10}\left(b \cdot {V^*}_F + c \cdot {Q^*}_{Cap} + d\right)\right) \quad (5)$$

Which is equivalent to:

$$p_{over,max} = min\left(1; \left(b \cdot {V^*}_F + c \cdot {Q^*}_{Cap} + d\right)^a\right) \quad (6)$$

where *a*, *b*, *c* and *d* are coefficients that must be adjusted for each gates reliability.

The objective is to adjust these coefficients so the upper envelope covers at least 95% of the calculated points, considered to be statistically significant for this study. However, as stated before, there are points that remain over the adjusted surfaces, corresponding to lower values of the V^*_F parameter. In order to deal with this limitation, it has been decided to establish a critical value of the Unit Storage Capacity, named $V^*_{F,crit}$. Beyond this limit, the tool will be robust enough to capture the maximum expected probability of 95% of the cases with a maximum of 5% of the points exceeding the surface. **Table 2** includes the critical Unit Storage Capacity values, $V^*_{F,crit}$, identified for each gates reliability case. It is worth noting that for the non-gated dams, no Unit Storage Capacity $V^*_{F,crit}$ value is needed to ensure the correct performance of the surface. Taking into account the criteria exposed above, the four coefficients of Eq. (6) have been adjusted on an iterative process for each gates reliability and are presented in **Table 3**.

Eq. (6), along with coefficients of **Table 3**, defines the surface that envelops at least 95% of the cases. An example of resulting surface for the 95% gates reliability is presented in **Fig. 4**, where black and blue points represent the empirical cases that fall over and under the envelope surface, respectively.

Table 2. *Critical Unit Storage Capacity* $V^*_{F,crit}$ *defining the range of application of the tool.*

	Gates reliability				
	50%	75%	85%	95%	100%
Critical Unit Storage Capacity $V^*_{F,\,crit}$	0.30	0.25	0.20	0.15	0 (no limit)

Table 3. *Adjusted coefficients of the tool grouped by gates reliability.*

	Gates reliability				
Coefficients	50%	75%	85%	95%	100%
a	-11.107	-11.056	-11.056	-11.056	-11.056
b	0.839	0.992	1.044	1.042	1.042
c	0.388	0.441	0.493	0.591	0.747
d	0.824	0.802	0.791	0.792	0.789

These results highlight the relative influence of each variable on the expected overtopping probability. They also suggest how a variation of the probability can be achieved through the respective modification of the Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}.

Although the tool has been designed using the $V^*_{F,crit}$ values to capture 95% of the cases under the enveloping surfaces, it is possible to enlarge its scope of application by not considering any limitation. However, this will cause its performance to decrease as a larger number of cases remain above the $p_{over,max}$ surfaces. **Table 4** shows the different performances of the tool under both assumptions: considering and not considering the $V^*_{F,crit}$ limitation. Should the user of the tool decide not to take into account the limitation, it is recommended to treat carefully its results. Performance is here defined as the percentage of points effectively located under the surface issued from the tool.

A summary of the process followed for the definition of the tool is presented in **Fig. 5**.

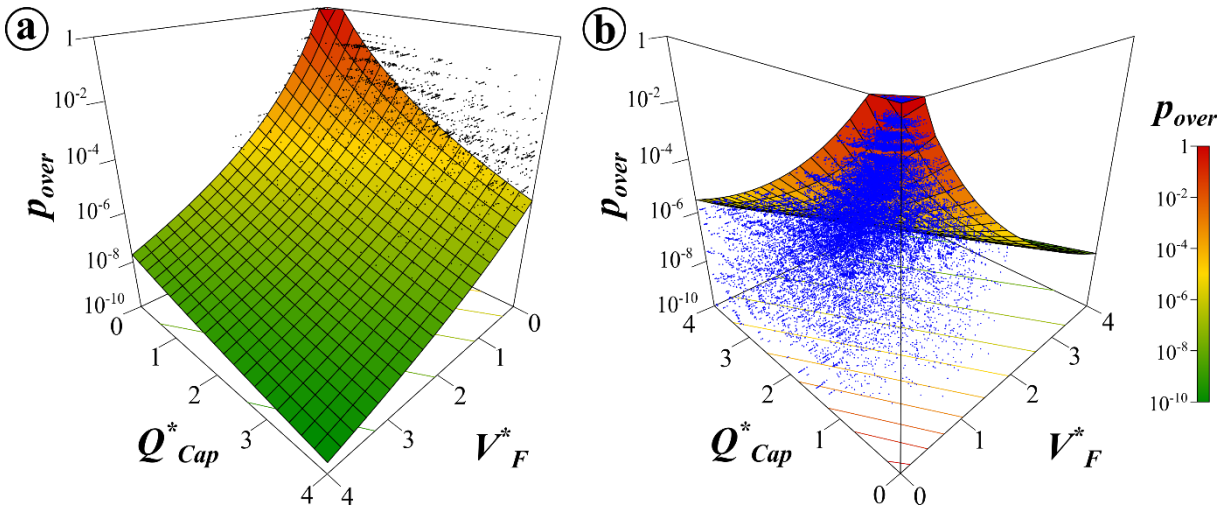

Fig. 4. *Envelope surface and contour lines of equal probability for a gates reliability of 95%. Points represent the empirical cases:* ***(a)*** *located over the envelope surface;* ***(b)*** *located under the envelope surface.*

Table 4. *Performance of the tool considering and not considering the* $V^*_{F,crit}$ *limitation.*

	Gates reliability				
	50%	75%	85%	95%	100%
Considering $V^*_{F,crit}$	95%	95%	95%	95%	97%
Not considering $V^*_{F,crit}$	85%	85%	85%	88%	97%

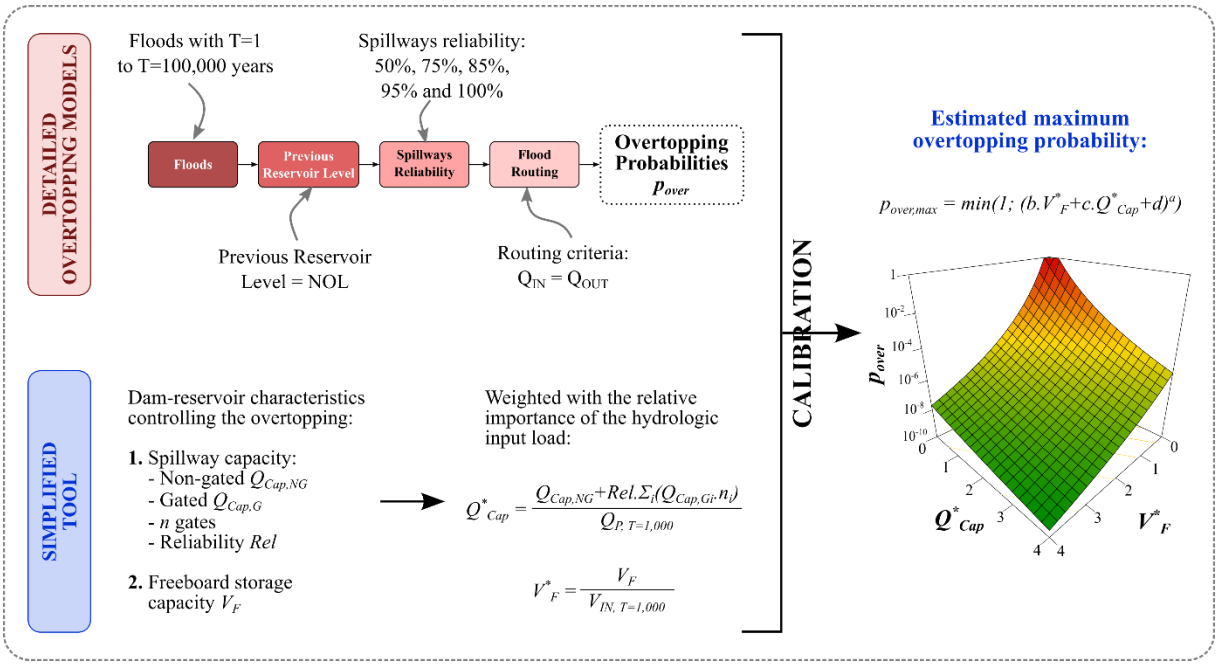

Fig. 5. *Process followed for the definition of the simplified tool based on probabilities issued from detailed models.*

Adequacy of the tool formulation

In the definition of the tool, and more specifically in the choice of the form and the variables of Eq. (6) to be used, a series of assumptions has been taken. It is necessary to justify that these variables explain the behavior of the overtopping probability as well as possible. This is done by comparing how well this probability is explained through these variables using the relation described in Eq. (6) with how well it would be explained using other variables or relations.

Although the coefficient of determination R^2 is widely used in statistics to assess how a dependent variable is explained by other independent variables using linear regression models (Rawlings et al. 2001), it is not an optimal choice in a nonlinear regime (Kvålseth 1985; Spiess and Neumeyer 2010). That is because the total sum-of-squares is not equal to the regression sum-of-squares plus the residual sum-of-squares in non-linear regression, which is the case of the present tool. In this work, the Akaike Information Criterion or AIC (Akaike 1974, 1978) has been used to estimate the goodness of fit of the tool relative to each of the other alternatives. The alternative with the lowest AIC is preferred.

Selection of the dam crest level for characterizing the outlet release capacity

The first assumption in the formulation of the tool to be checked is the selection of the level at which the spillway's release capacity has to be defined. In this work, the dam crest level, DC, is used to characterize the spillways capacity Q_{Cap} (*Alternative 0*). This alternative has to be compared to the use of the Design Level (*Alternative 1A*) and of the Normal Operating Level (*Alternative 1B*). **Fig. 6** shows the AIC values when using the different levels, for each gates reliability. Best results are obtained with the dam crest level since the AIC values are lower for *Alternative 0*.

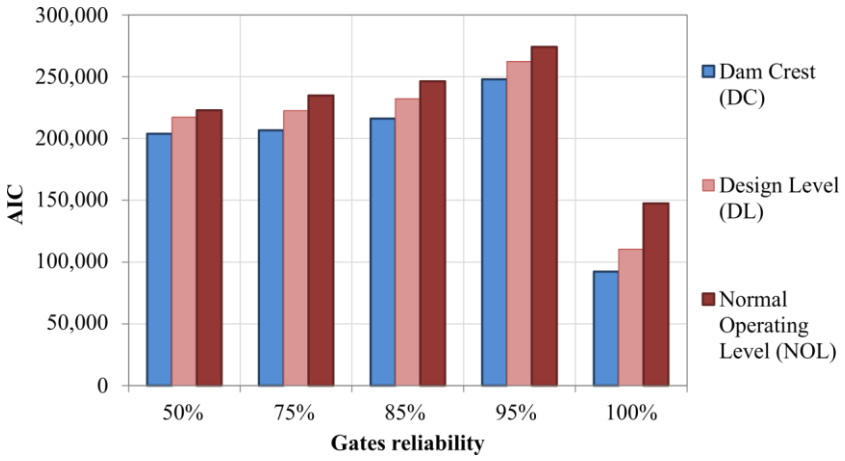

Fig. 6. *Comparison of AIC values considering the Dam Crest (Alternative 0), the Design Level (Alternative 1A) and the Normal Operating Level (Alternative 1B) for the characterization of the spillways capacity, per gates reliability.*

Inclusion of number of gates and their reliability

The next assumption is the inclusion of a factor to correct gated spillways' capacity, expressed as $n \times Rel$ (*Alternative 0*). The *Alternative 2* to be compared to is the use of the Unit Spillway Capacity (Q^*_{Cap}) without taking into account this factor, which is:

$$Q^*_{Cap\_Alt2} = \frac{Q_{Cap}}{Q_{P,T=1'000}} = \frac{Q_{Cap,NG} + \sum_i (Q_{Cap,G\,i} \cdot n_i)}{Q_{P,T=1'000}} \quad (7)$$

Fig. 7 shows the AIC values when using the factor of correction (*Alternative 0*) and not using it (*Alternative 2*), for each gates reliability. The model improves when considering the proposed factor since the AIC values are higher for *Alternative 2*, which justifies its use for the tool.

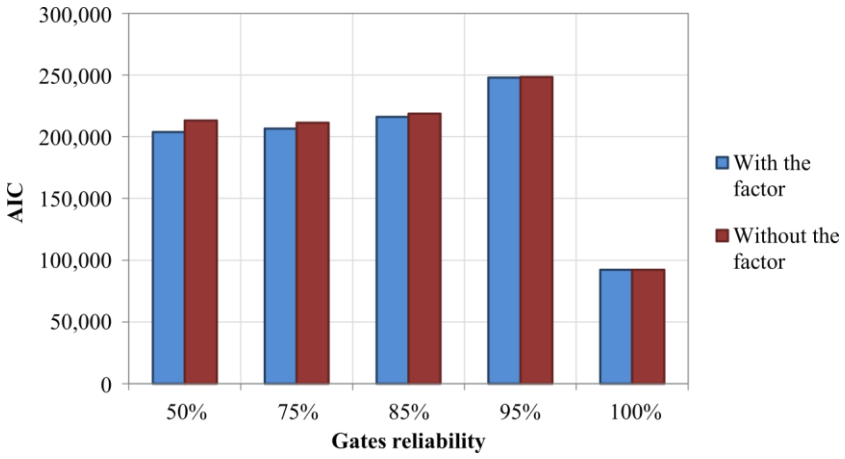

Fig. 7. *Comparison of AIC values considering the factor of correction (Alternative 0) and not considering it (Alternative 2), per gates reliability.*

Selection of the 1,000-year flood as representative for the characterization of the tool

Among the different return periods of floods considered, the tool, and particularly the Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}, is based on the 1,000-year flood as shown in Eq. (2) and Eq. (3). To validate the suitability of such assumption, it has been tested if another available return period in the range from T=2 to T=100,000 years offers a better indicator AIC (*Alternative 3*). For visualization purposes, the graph in **Fig. 8** represents the difference of AIC values of the models using each available return period

and using the T=1,000 as a baseline: $\Delta AIC = AIC_T - AIC_{T=1,000}$. Hence, negative values indicate a more suitable choice. Except for some isolated cases, in general the use of the T=1,000 flood offers better results than the rest of the return periods. Moreover, the choice of the 1,000-year flood has the advantage of being available in most of the hydrological studies for dams.

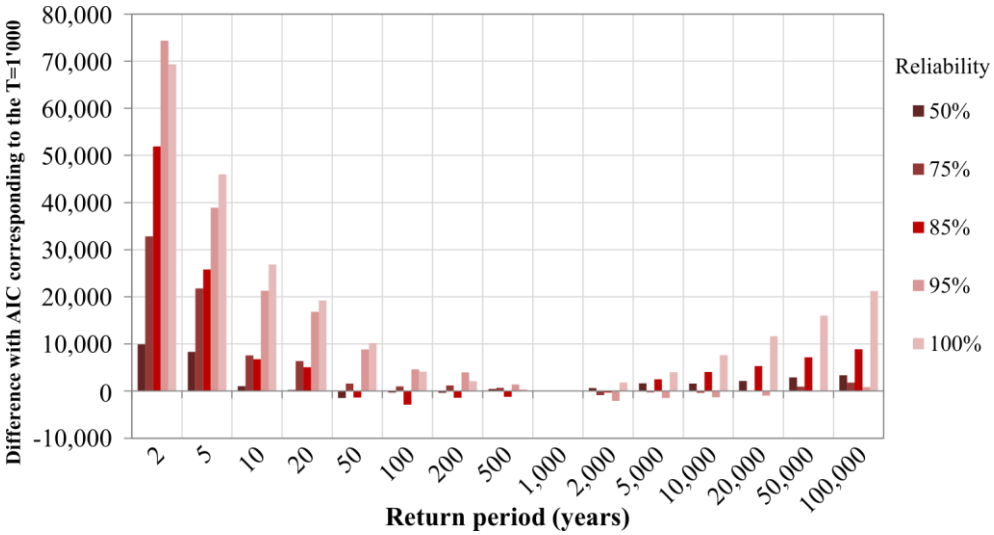

Fig. 8. *Comparison of AIC values between using a given return period T (Alternative 3) and T=1,000 (Alternative 0), per gates reliability.*

Selection of a non-linear model for characterizing the tool's equation

The equation is expressed as a non-linear relation between the two independent variables (V^*_F and Q^*_{Cap}) and the dependent one (p_{over}), leading to a curved surface as shown in **Fig. 4**. It has been tested if a linear model, equivalent to a three-dimensional plane, would be preferable (*Alternative 4*). **Fig. 9** presents the AIC values of the quality of both the non-linear and the linear models. It is clear that a non-linear model is more adapted for capturing the effect of the two independent variables on the overtopping probability.

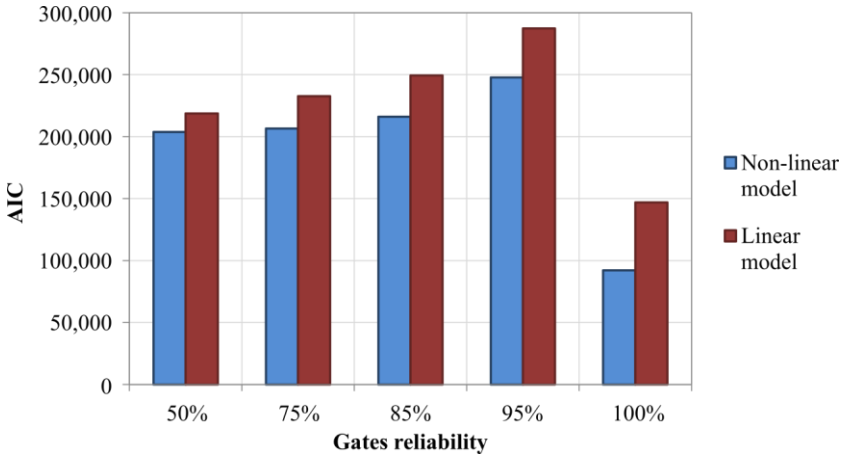

Fig. 9. *Comparison of AIC values using a non-linear (Alternative 0) and a linear (Alternative 4) model, per gates reliability.*

TOOL VALIDATION WITH PRACTICAL CASES

In order to validate its performance, the simple tool developed has been applied to several real world cases that are not part of the set of dams used for the calibration of the tool.

Description of the dam-reservoir systems

The simplified model has been applied to 13 dam-reservoir systems located in different geographical areas: 6 of them from Mediterranean river basins, 6 from Caribbean river basins and 1 from an Atlantic river basin in South America. These cases cover a wide range of spillway capacities, freeboard storage capacity and magnitude of the inflow hydrographs. The main characteristics of the dams are presented in **Table 5**. It is worth noting that in the dams with gated spillways, four reliabilities have been considered: 50%, 75%, 85%, and 95%.

Table 5. *Main characteristics of the 13 dams analyzed.*

Name	Hydrological loads		Freeboard volume	Spillway capacities (m^3/s)					
	$V_{IN,1000}$	$Q_{P,1000}$	V_F	Non-gated	Gated				
	(hm^3)	(m^3/s)	(hm^3)	$Q_{Cap,NG}$	$Q_{Cap,G1}$	$gates_1$	$Q_{Cap,G2}$	$gates_2$	Rel.
Dam 1	233	4241	586	8253	—	—	—	—	—
Dam 2	143	3542	73	10765	—	—	—	—	—
Dam 3	110	2756	126	5653	—	—	—	—	—
Dam 4	153	3201	40	12191	—	—	—	—	—
Dam 5	153	3201	40	5134	—	—	—	—	—
Dam 6	277	6066	99	—	1921	5	—	—	50-95%
Dam 7	2570	6132	1861	—	1255	1	—	—	50-95%
Dam 8	22	600	9	—	212	3	234	3	50-95%
Dam 9	0.9	105	0.05	92	—	—	—	—	—
Dam 10	0.6	25	0.3	277	—	—	—	—	—
Dam 11	4.3	100	0.3	—	91	2	—	—	50-95%
Dam 12	2.6	233	1	283	—	—	—	—	—
Dam 13	1096	7499	216	7649	—	—	—	—	—

Results and discussion

The tool has been applied to the 25 cases issued from these 13 dams, taking into account the different gate reliability considered. The resulting estimated maximum overtopping probabilities $p_{over,max}$ of each dam are presented in **Table 6**.

To contrast these probabilities, detailed overtopping models have been elaborated with the *iPresas* software (iPresas 2017) following the indications described above. **Table 6** also contains the overtopping probabilities p_{over} obtained. It is worth stressing that in 5 cases the calculated probability is null. This fact does not mean in an absolute way that there is never an overtopping in these dams, but instead, that overtopping does not happen within the range of considered floods (i.e., with return periods of up to 100,000 years) and for any of the combinations of the rest of conditions. Moreover, 4 cases are outside the range of application since their V^*_F is below the critical value corresponding to their gates reliabilities (**Table 2**): Dam 11.a, Dam 11.b, Dam 11.c and Dam 11.d. These cases are excluded from the analysis and thus only 21 cases remain. The graphical comparison between maximum overtopping probabilities from the tool and the probabilities calculated with the detailed models is presented in **Fig. 10 (a)**. In this graphic, null probabilities are plotted at the edge of the grid. The coverage of the Q^*_{Cap} and V^*_F values of the 25 validation set cases is shown in **Fig. S1** of the Supplemental Material.

Results show a good performance of the simplified tool. Maximum overtopping probabilities calculated with the simplified tool are greater than the overtopping probabilities calculated with the detailed models except for one case. This case represents the 4.8% of the total cases tested, which concurs with the definition of the tool.

Table 6. *Comparison between overtopping probabilities resulting from the tool and the detailed model (bolded cases are not applicable and thus not analyzed).*

Dam	Simplified tool				Detailed model
	Parameters			Maximum overtopping probability ($p_{over,max}$)	Overtopping probability (p_{over})
	Unit Storage Capacity (V^*_F)	Unit Spillway Capacity (Q^*_{Cap})	Reliability case		
1	2.51	1.95	100%	2.5×10^{-8}	0.0 [a]
2	0.51	3.04	100%	7.2×10^{-7}	0.0 [a]
3	1.16	2.05	100%	8.9×10^{-7}	0.0 [a]
4	0.26	2.50	100%	2.69×10^{-7}	0.0 [a]
5	0.26	1.60	100%	1.2×10^{-4}	1.0×10^{-4}
6.a	0.36	0.79	50%	1.9×10^{-2}	3.7×10^{-2}
6.b	0.36	1.19	75%	3.2×10^{-3}	1.7×10^{-3}
6.c	0.36	1.35	85%	1.3×10^{-3}	3.0×10^{-4}
6.d	0.36	1.50	95%	3.5×10^{-4}	1.1×10^{-4}
7.a	0.72	0.10	50%	1.4×10^{-2}	9.2×10^{-4}
7.b	0.72	0.15	75%	6.1×10^{-3}	5.8×10^{-4}
7.c	0.72	0.17	85%	4.4×10^{-3}	4.5×10^{-4}
7.d	0.72	0.19	95%	3.6×10^{-3}	3.2×10^{-4}
8.a	0.42	1.11	50%	5.2×10^{-3}	1.9×10^{-3}
8.b	0.42	1.67	75%	6.1×10^{-4}	9.6×10^{-5}
8.c	0.42	1.89	85%	2.0×10^{-4}	1.5×10^{-5}
8.d	0.42	2.12	95%	4.4×10^{-5}	4.6×10^{-7}
9	0.06	0.88	100%	1.1×10^{-2}	1.0×10^{-3}
10	0.54	10.87	100%	1.6×10^{-11}	0.0 [a]
11.a	**0.06**	**0.90**	**50%**	$\mathbf{1.0\times10^{-1}}$	$\mathbf{3.4\times10^{-3}}$
11.b	**0.06**	**1.35**	**75%**	$\mathbf{1.5\times10^{-2}}$	$\mathbf{1.4\times10^{-3}}$
11.c	**0.06**	**1.54**	**85%**	$\mathbf{5.0\times10^{-3}}$	$\mathbf{8.1\times10^{-4}}$
11.d	**0.06**	**1.72**	**95%**	$\mathbf{9.7\times10^{-4}}$	$\mathbf{4.0\times10^{-4}}$
12	0.39	1.21	100%	2.8×10^{-4}	1.0×10^{-4}
13	0.20	1.02	100%	2.0×10^{-3}	3.2×10^{-4}

[a] No overtopping events produced for any flood with T<100,000 years.

One of the main applications of this tool in the context of dam safety management is the classification of groups of dams according to their overtopping risk level. In order to verify the suitability of the tool for this purpose, a classification of the 21 dam cases based on the probabilities calculated with detailed models and with the simplified tool has been defined. Dams are ranked from 1, with higher overtopping probability, to 21, with lower probability. The relationship between their positions is presented graphically in **Fig. 10 (b)**.

In order to assess the existing correlation between both classifications, a rank correlation analysis has been performed. For this purpose, the Spearman's rank correlation coefficient ρ is applied to the series of positions of the dams, and evaluates to $\rho = 0.945$ with a p-value < 0.005 (using the t-distribution). This result reveals a good correlation, not obtained by pure chance, at a 99.95% confidence interval: both models capture similarly the safety level of each dam. Therefore, through the use of the tool it is possible to quickly and easily estimate the overtopping risk level of each dam in a group of dams and thus prioritize the adaptation measures to be implemented.

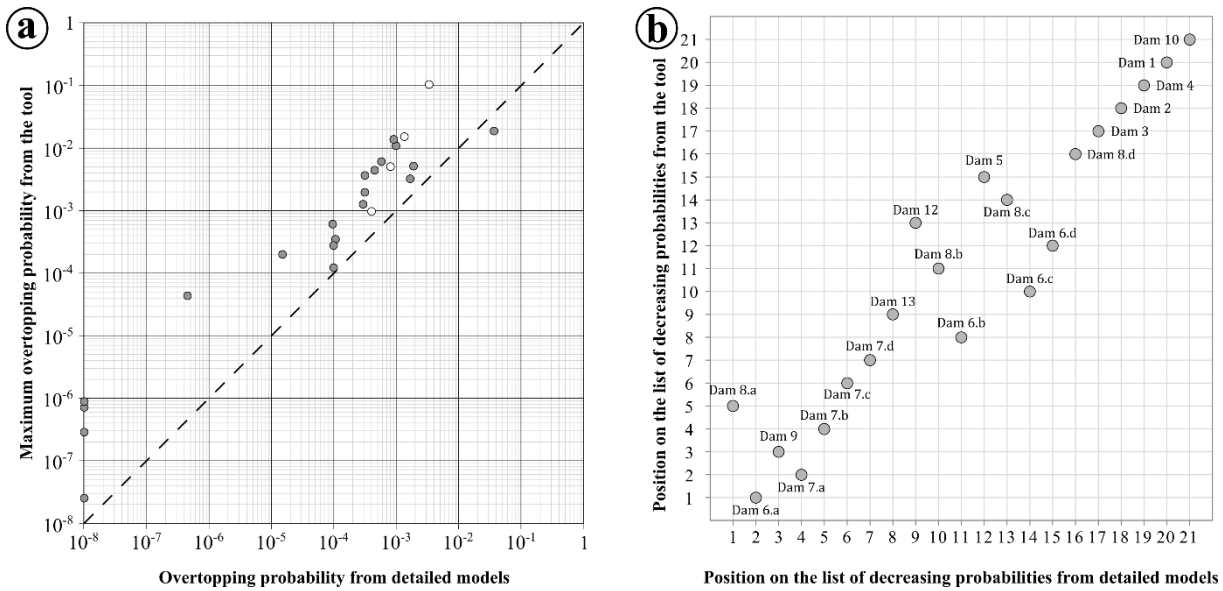

Fig. 10. (a) *Comparison between the overtopping probabilities calculated with the detailed models and with the tool (white points correspond to the 4 cases that are outside the range of application).* ***(b)*** *Comparison of the positions occupied by the 21 studied dams on the decreasingly ordered probabilities calculated by both the detailed models and the simplified tool.*

The tool is also useful for proposing mitigation measures, applicable when the calculated risk of overtopping the dam is not tolerable, and even in some cases when it is tolerable and is to be reduced. Different structural measures (e.g., the augmentation of the spillways' release capacity) and non-structural measures (e.g., the lowering of the pool to increase the freeboard volume) are commonly used (Afshar and Mariño 1990) and can be easily studied by applying the tool. The implementation of any of these measures or a combination of them would result in a modification of the characteristics of the dam-reservoir system: the freeboard storage capacity, the spillway capacity and/or the gates reliability, and consequently in a reduction of the overtopping probability. From the definition of the tool it is possible to determine and compare which measures will have a greater impact in the reduction of the maximum probability of overtopping. This is useful when establishing a classification of corrective measures based on their effectiveness and estimated cost, which gives an idea of their efficiency. For example, consider three adaptation measures envisaged for Dam 6.c, which maximum overtopping probability $p_{over,max}$=1.3×10^{-3}/year (V^*_F=0.36; Q^*_{Cap}=1.35) and overtopping probability p_{over}=3.0×10^{-4}/year. The measures consist of:

- **Measure 1**: the construction of a continuous parapet with height of 1.5 m, which increases the freeboard up to 115.3 hm^3, leading to a Unit Storage Capacity V^*_F= 0.42. The new maximum overtopping probability $p_{over,max}$ is 8.8×10^{-4}/year, while the new overtopping probability p_{over} is 1.2×10^{-4}/year.
- **Measure 2**: the increase of the spillway capacity up to 3000 m^3/s through each gate, calculated at the dam crest level, which leads to a Unit Spillway Capacity Q^*_{Cap}=2.10. The new maximum overtopping probability $p_{over,max}$ is 1.6×10^{-4}/year, while the new overtopping probability p_{over} is 2.0×10^{-5}/year.
- **Measure 3**: fixing of spillway's gates, thus increasing their reliability from 85% to 95%. The new maximum overtopping probability $p_{over,max}$ is 3.5×10^{-4}/year, while the new overtopping probability p_{over} is 3.4×10^{-5}/year.

Therefore, the application of the tool shows that Measure 2 is expected to have a greater reduction of the maximum overtopping probability $p_{over,max}$, followed by Measure 3 and finally Measure 1. This prioritization corresponds with the one based on the overtopping probability p_{over}.

CONCLUSIONS

This work presents a simple tool for the assessment of maximum overtopping probabilities of dams based on a few parameters of the dam-reservoir systems. The tool does not require the elaboration of *ad hoc* overtopping models but is yet able to provide results on the conservative side and with an adequate level of accuracy.

The tool developed is based on empirical relations between the overtopping probability of occurrence and the basic hydrological and hydraulic characteristics of the dam-reservoir system: Unit Storage Capacity, V^*_F, and the Unit Spillway Capacity, Q^*_{Cap}, both weighted with the relative importance of the 1,000-year flood. This information is readily available for any dam owner, generally included in the dam-reservoir technical information files. The surface issued from the tool's formulation represents the limit above which no combination of V^*_F and Q^*_{Cap} is statistically expected to offer a higher overtopping probability.

The probabilities issued from the detailed overtopping models of 30 existing dam-reservoir systems were used to calibrate the tool. This initial set of study cases was expanded by changing the key parameters of each dam, thus artificially increasing the number of cases. A total of 342,233 synthetic cases were generated and used to formulate and calibrate the tool. The coefficients of the tool were adjusted so the upper envelope offered maximum overtopping probabilities $p_{over,max}$ covering the calculated overtopping probabilities in at least 95% of the cases. In order to better capture the diversity of dams, the tool was formulated for 5 different gate reliability cases: 50%, 75%, 85%, 95% and 100% including non-gated dams.

Finally, the tool was successfully applied to 21 independent cases and validated with more complex risk models. Results showed a good performance of the tool and a good correlation with the overtopping probabilities calculated with the detailed models.

This tool is especially useful for dam engineers when analyzing a portfolio of dams in preliminary screening phases of risk analysis. It aims at identifying the overtopping as a relevant failure mode and classifying each dam in a simplified manner in terms of their overtopping probability. This classification would facilitate the allocation of resources for the risk assessment of each dam emphasizing those with more compromised safety levels, thus helping to optimize further efforts and saving the time of developing a detailed risk model for each one. The tool is also a support for the definition and prioritization of corrective measures. It assesses the impact of these measures in the reduction of the overtopping probability.

REFERENCES

Afshar, A., and Mariño, M. A. (1990). "Optimizing Spillway Capacity with Uncertainty in Flood Estimator." *Journal of Water Resources Planning and Management*, 116(1), 71–84.

Akaike, H. (1974). "A new look at the statistical model identification." *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Akaike, H. (1978). "On the Likelihood of a Time Series Model." *The Statistician*, 27(3/4), 217.

Altarejos-García, L., Fluixá-Sanmartín, J., Escuder-Bueno, I., and Serrano-Lombillo, A. (2013). "Towards a simple tool for estimating dam overtopping probability in the context of risk analysis." *Proceedings of the 9th ICOLD European Club Symposium*, Venice, Italy.

ANCOLD. (2003). *Guidelines on Risk Assessment*. Australian National Committee on Large Dams.

Bowles, D. S., Anderson, L. R., Glover, T. F., and Chauhan, S. S. (1998). "Portfolio risk assessment: a tool for dam safety risk management." *Proceedings of the 1998 USCOLD Annual Lectur*, Buffalo, New York.

Chow, V. T., Maidment, D. R., and Mays, L. W. (2008). *Applied hydrology*. McGraw-Hill series in water resources and environmental engineering, McGraw-Hill, New York.

Costa, J. E. (1985). *Floods from Dam Failures*. U.S. Geological Survey Open-File Report, Denver, Colorado, 54.

De Cea, J. C., García, S., and Escuder-Bueno, I. (2010). "Análisis de riesgos de algunas presas de titularidad estatal mediante la técnica de screening." *IX Jornadas Españolas de Presas*, Valladolid, Spain.

Escuder-Bueno, I., and González-Pérez, J. (2014). *Metodología para la evaluación del riesgo hidrológico de presas y priorización de medidas correctoras*. Colegio de Ingeniero de Caminos, Canales y Puertos, Madrid, Spain.

Escuder-Bueno, I., de Membrillera, M. G., Moreno, P., Pérez, O., Ardiles, L., and Jenaro, E. (2008). "First risk-based screening on a Spanish portfolio of 20 dams owned by the Duero River Authority." *Hydrovision*, Sacramento (USA).

Escuder-Bueno, I., Morales-Torres, A., and Castillo-Rodriguez, J. T. (2016). "Use of quantitative risk results to inform dam safety governance: Practical cases in Europe." *36th Annual USSD Conference 2016*, Denver, Colorado, USA.

FEMA. (2015). *Federal Guidelines for Dam Safety Risk Management*. FEMA, Federal Emergency Management Agency.

Fiedler, W. R. (2010). "Performance of Spillway Structures Using Hoover Dam Spillways As a Benchmark." American Society of Civil Engineers, 267–287.

Goodarzi, E., Shui, L. T., and Ziaei, M. (2014). "Risk and uncertainty analysis for dam overtopping – Case study: The Doroudzan Dam, Iran." *Journal of Hydro-environment Research*, 8(1), 50–61.

Hsu, Y.-C., Tung, Y.-K., and Kuo, J.-T. (2011). "Evaluation of dam overtopping probability induced by flood and wind." *Stochastic Environmental Research and Risk Assessment*, 25(1), 35–49.

ICOLD. (1995). *Dam Failures - Statistical Analysis*. Bulletin, International Commission on Large Dams.

ICOLD. (2005). *Risk assessment in dam safety management. A reconnaissance of benefits, methods and current applications*. Bulletin, International Commission on Large Dams.

iPresas. (2017). *iPresas calc. User guide. iPresas risk analysis*.

Jandora, J., Říha, J., and Vysoké učení technické v Brně. (2008). *The failure of embankment dams due to overtopping*. VUTIUM, Brno.

Jeffrey T. McClenathan, P. E., and Andy Harkness, P. E. (2008). "Screening update for portfolio risk analysis for U.S. Army Corps of Engineer dams." *The sustainability of experience - investing in the human factor*, United States Society on Dams, 337–343.

Kuo, J.-T., Hsu, Y.-C., Tung, Y.-K., Yeh, K.-C., and Wu, J.-D. (2008). "Dam overtopping risk assessment considering inspection program." *Stochastic Environmental Research and Risk Assessment*, 22(3), 303–313.

Kuo, J.-T., Yen, B.-C., Hsu, Y.-C., and Lin, H.-F. (2007). "Risk Analysis for Dam Overtopping—Feitsui Reservoir as a Case Study." *Journal of Hydraulic Engineering*, 133(8), 955–963.

Kvålseth, T. O. (1985). "Cautionary Note about R^2." *The American Statistician*, 39(4), 279–285.

Kwon, H.-H., and Moon, Y.-I. (2006). "Improvement of overtopping risk evaluations using probabilistic concepts for existing dams." *Stochastic Environmental Research and Risk Assessment*, 20(4), 223–237.

Lee, B.-S., and You, G. J.-Y. (2013). "An assessment of long-term overtopping risk and optimal termination time of dam under climate change." *Journal of Environmental Management*, 121, 57–71.

Michailidi, E. M., and Bacchi, B. (2017). "Dealing with uncertainty in the probability of overtopping of a flood mitigation dam." *Hydrology and Earth System Sciences*, 21(5), 2497–2507.

Morales-Torres, A., Serrano-Lombillo, A., Escuder-Bueno, I., and Altarejos-García, L. (2016). "The suitability of risk reduction indicators to inform dam safety management." *Structure and Infrastructure Engineering*, 1–12.

Parzen, E. (1962). "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics*, 33(3), 1065–1076.

Rawlings, J. O., Pantula, S. G., and Dickey, D. A. (2001). *Applied regression analysis: a research tool*. Springer texts in statistics, Springer, New York, NY.

Rosenblatt, M. (1956). "Remarks on Some Nonparametric Estimates of a Density Function." *The Annals of Mathematical Statistics*, 27(3), 832–837.

Serrano-Lombillo, A., Escuder-Bueno, I., de Membrillera-Ortuño, M. G., and Altarejos-García, L. (2011). "Methodology for the Calculation of Annualized Incremental Risks in Systems of Dams: Risk Calculation for Systems of Dams." *Risk Analysis*, 31(6), 1000–1015.

Serrano-Lombillo, A., Fluixá-Sanmartín, J., and Espert-Canet, V. J. (2012). "Flood routing studies in risk analysis." *Risk Analysis, Dam Safety, Dam Security and Critical Infrastructure Management*, Ignacio Escuder-Bueno, Enrique Matheu, Luis Altarejos-García, and Jesica T. Castillo-Rodríguez. Leiden: CRC Press, Valencia, Spain, 99–105.

Serrano-Lombillo, A., Morales-Torres, A., Escuder-Bueno, I., and Altarejos-García, L. (2017). "A new risk reduction indicator for dam safety management combining efficiency and equity principles." *Structure and Infrastructure Engineering*, 13(9), 1157–1166.

Setrakian-Melgonian, M., Escuder-Bueno, I., Castillo-Rodriguez, J. T., Morales-Torres, A., and simarro-Rey, D. (2017). "Quantitative Risk Analysis to inform safety investments in Jaime Ozores Dam (Spain)." International Commission of Large Dams (ICOLD), Prague, Czech Republic.

SPANCOLD. (2012). *Risk Analysis as Applied to Dam Safety. Technical Guide on Operation of Dams and Reservoirs*. Professional Association of Civil Engineers. Spanish National Committe on Large Dams, Madrid.

Spiess, A.-N., and Neumeyer, N. (2010). "An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach." *BMC Pharmacology*, 10(1), 6.

Sun, Y., Chang, H., Miao, Z., and Zhong, D. (2012). "Solution method of overtopping risk model for earth dams." *Safety Science*, 50(9), 1906–1912.

USACE. (2011). *Safety of dams - Policy and procedures*. Engineering and design, U.S. Army Corps of Engineers, Washington, DC.

USBR. (2014). "Chapter 3: General Spillway Design Considerations." *Appurtenant Structures for Dams (Spillways and Outlet Works)*, Design Standards, U.S. Bureau of Reclamation.

USBR, and USACE. (2011). *Dam Safety Risk Analysis Best Practices Training Manual*. U.S. Bureau of Reclamation, U.S. Army Corps of Engineers, Denver, Colorado.

USBR, and USACE. (2015). *Best Practices in Dam And Levee Safety Risk Analysis*. United States Bureau of Reclamation and United States Army Corps of Engineers, Denver.

Wahl, T. (1998). *Prediction of Embankment Dam Breach Parameters. A Literature Review and Needs Assessment.* Dam Safety Research Report, USBR, 61.

Wang, F., and Zhang, Q.-L. (2017). "Systemic Estimation of Dam Overtopping Probability: Bayesian Networks Approach." *Journal of Infrastructure Systems*, 23(2), 04016037.

World Meteorological Organization. (2008). *Guide to hydrological practices*. World Meteorological Organization, Geneva.

SUPPLEMENTAL MATERIAL

Table S1. *Main characteristics of the 30 dams used for the calibration of the tool.*

Dam	Characteristics				Hydrological loads		Freeboard volume	Spillway capacities at dam crest (m^3/s)		
	Type	H	L	Capacity at NOL	$V_{IN,1000}$	$Q_{P,1000}$	V_F	Non-gated	Gated	
		(m)	(m)	(hm^3)	(hm^3)	(m^3/s)	(hm^3)	$Q_{Cap,NG}$	$Q_{Cap,G}$	Gates
1	Gravity	47.2	267	22.38	1.85	77	1.04	—	119	1
2	Gravity	57.0	200	66.42	21.03	830	1.71	—	261	2
3	Gravity	41.7	335	18.60	33.86	1,474	2.77	—	606	2
4	Gravity	77.7	270	95.00	8.69	189	4.56	—	237	1
5	Gravity	31.0	464	72.0	51.84	914	9.68	—	159	5
6	Gravity	35.6	112	51.30	55.03	824	6.19	—	325	2
7	Gravity	74.0	534	44.13	55.63	510	24.31	254	33	1
8	Gravity	60.0	517	496.00	127.11	4,080	41.30	—	504	5
9	Gravity	37.6	196	22.30	48.43	4,928	6.06	679	562	3
10	Gravity	75.5	160	69.79	24.18	517	5.85	—	235	2
11	Gravity	48.0	500	247.23	42.24	1,126	21.50	—	291	3
12	Gravity	31.0	198	1.00	9.88	375	0.89	671	—	—
13	Gravity	40.5	311	24.00	15.03	485	2.23	—	405	3
14	Gravity	41.5	206	19.96	2.86	212	2.12	—	138	2
15	Gravity	96.2	229	308.00	51.20	824	25.23	—	876	2
16	Gravity	40.4	425	248.78	43.28	1,843	26.63	—	465	3
17	Gravity	77.6	255	317.83	33.24	993	25.86	—	298	3
18	Arch	36.5	130	10.26	5.47	291	0.91	163	—	—
19	Arch	67.0	454	56.60	53.00	1557	15.64	1,029	146	2
20	Arch	75.5	420	110.00	63.14	5,265	33.16	1,650	—	—
21	Arch	100.6	337	641.35	126.15	2,069	60.95	456	510	2
22	Arch	48.0	248	7.41	20.32	585	2.17	632	85	2
23	Arch	8.3	25	0.27	0.84	43	0.07	97	122	1
24	Weir	13.0	87	2.00	9.46	402	2.85	—	275	3
25	Weir	14.0	156	5.50	77.09	3,656	1.74	—	358	8
26	Weir	16.0	147	6.00	54.19	1,740	1.60	—	400	8
27	Weir	14.4	159	2.35	17.56	330	1.13	—	264	8
28	Embankm.	16.06	263	0.46	0.16	39	0.14	49	—	—
29	Embankm.	15.7	113	14.00	0.42	53	0.97	322	—	—
30	Embankm.	64.0	460	75.00	3.06	265	16.62	410	—	—

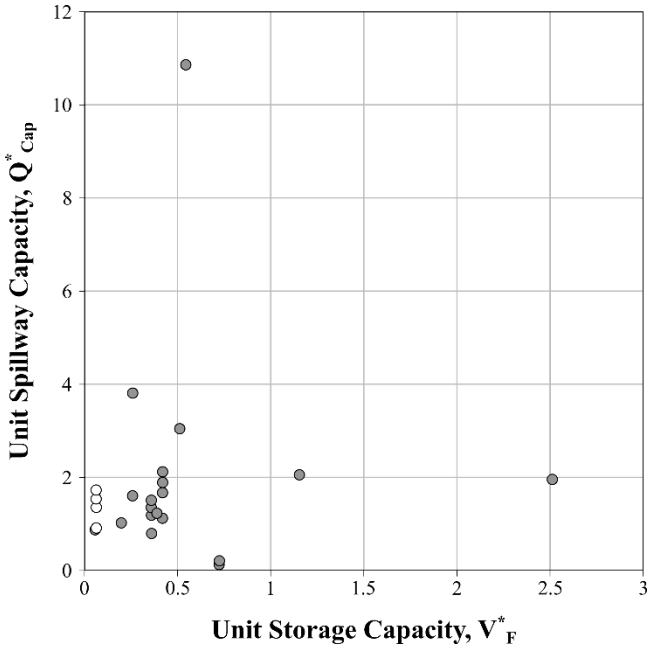

Fig. S1. *Coverage of the Q^*_{Cap} and V^*_F values of the 25 validation set cases (white points correspond to the 4 cases that are outside the range of application).*