

This article was downloaded by: [Universidad Autónoma del Estado de México], [David-Joaquín Delgado-Hernández]

On: 15 October 2012, At: 17:55

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/nsie20>

### A continuous Bayesian network for earth dams' risk assessment: an application

David-Joaquín Delgado-Hernández<sup>a</sup>, Oswaldo Morales-Nápoles<sup>b c</sup>, David De-León-Escobedo<sup>a</sup> & Juan-Carlos Arteaga-Arcos<sup>a</sup>

<sup>a</sup> Civil Engineering Department, School of Engineering, Universidad Autónoma del Estado de México, Estado de México, Mexico

<sup>b</sup> Netherlands Organization for Applied Scientific Research (TNO), Delft, Netherlands

<sup>c</sup> Faculty of Electrical Engineering Mathematics and Computer Sciences, Delft University of Technology, Delft, Netherlands

Version of record first published: 15 Oct 2012.

To cite this article: David-Joaquín Delgado-Hernández, Oswaldo Morales-Nápoles, David De-León-Escobedo & Juan-Carlos Arteaga-Arcos (2012): A continuous Bayesian network for earth dams' risk assessment: an application, Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance, DOI:10.1080/15732479.2012.731416

To link to this article: <http://dx.doi.org/10.1080/15732479.2012.731416>



PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## A continuous Bayesian network for earth dams' risk assessment: an application

David-Joaquín Delgado-Hernández<sup>a\*</sup>, Oswaldo Morales-Nápoles<sup>b,c</sup>, David De-León-Escobedo<sup>a</sup> and Juan-Carlos Arteaga-Arcos<sup>a</sup>

<sup>a</sup>Civil Engineering Department, School of Engineering, Universidad Autónoma del Estado de México, Estado de México, Mexico;

<sup>b</sup>Netherlands Organization for Applied Scientific Research (TNO), Delft, Netherlands; <sup>c</sup>Faculty of Electrical Engineering Mathematics and Computer Sciences, Delft University of Technology, Delft, Netherlands

(Received 23 December 2011; final version received 11 April 2012; accepted 19 May 2012)

Dams are civil engineering structures to hinder water flows. Criteria such as purpose, size and construction material are useful to categorise them. The latter is used to classify 'earth dams', which tend to have higher risk levels than other types. The failure of a dam leads to significant economic loss and usually to catastrophic impacts. In an effort to comprehensively examine the variables that influence earth dam breaks and describe their interactions, a model has been developed in such a way that it allows to assess risks and to prioritise the allocation of resources for maintenance activities. The research was carried out by systematically reviewing the literature, which led to the choice of Bayesian Networks (BNs) as a tool for assessing risks. Using data from seven case studies in Mexico, a model was built, which helped to rank the dams under study, leading to results comparable with those reported in the literature. While the particular type of BN used and its quantification is presented more extensively in an accompanying paper, the model may be of interest for dam owners, managers, practitioners and academics on their efforts to manage earth dams' risks.

**Keywords:** Bayesian analysis; risk management; dam safety; embankment stability; flood frequency; expert systems

### Introduction

The International Commission of Large Dams (ICOLD, 2008) defines a dam as an artificial obstruction to natural water flows constructed for one or more specific purposes such as accumulating water for farm irrigation, generating electricity, creating artificial lakes for navigation and leisure activities, supplying water to cities or industry, preventing floods, diverting river flows into canals and keeping a reserve of fresh water.

In terms of size, the same source differentiates among two types: small and large. Small dams are structures with < 15 m of height. Large dams, in contrast, are those with 15 m or more from the foundation to the crest or, between 5 and 15 m with a capacity of more than 3 million m<sup>3</sup>. Based on their structure, they can be categorised as: embankment, gravity, arch and buttress dams (Emiroglu, Tuna, & Aislan, 2002).

Regardless of their construction materials, these structures may well fail. A failure is a 'collapse or movement of part of a dam or its foundation, so that the dam cannot retain water. In general, a failure results in the release of large quantities of water, imposing risks on the people or property downstream'

(ICOLD, 1995). A statistical analysis published in the mid-1990s (ICOLD, 1995) revealed that the frequency of failure was higher in embankment dams (85% of the cases studied) than in the other types (10% gravity, 3% arch and 2% buttress). However, Donnelly (2006) stated that simple statistics can be misleading because they have not considered the fact that embankment dams are the most common type of water retaining structures. In this sense, he noted that 2.6% of the concrete buttress failed. This is a percentage greater than that of embankment dams (1.2%), concrete arch (0.7%) and concrete gravity (0.3%). In any case, failure has negative effects and costs.

The impacts of a dam collapse can be enormous, comprising the destruction of private housing, transport and public infrastructure, industrial facilities and agricultural land (Pistrika & Tsakiris, 2007). The losses may also include human harm and serious disruptions in infrastructure operation, leading to significant total economic damages (sometimes reaching hundreds of USD millions). Risk may be viewed from two perspectives: qualitative and quantitative (Modarres, Kaminskiy, & Krivtsov, 1999). The first entails the existence of a danger source (hazard) when there is no protection against it, leading to the possibility of loss. Then, 'risk can be formally defined as the potential of

\*Corresponding author. Email: delgadoh01@yahoo.com

loss (e.g. material, human or environment losses) resulting from exposure to a hazard' (Modarres et al., 1999). From the quantitative viewpoint, the same source defined risk as the following set of triplets ' $R = \langle S_i, P_i, C_i \rangle; i = 1, 2, \dots, n$ ', where  $S_i$  is a state of events that lead to a dangerous situation,  $P_i$  is the likelihood of state  $i$  and  $C_i$  is the consequence of state  $i$ .

In this paper, the quantitative approach will be used to estimate the degree of loss in seven Mexican earth dams, and the discussion will be centred on the costs (human and economic) produced by the eventuality of a failure in any of them. The seven structures under investigation are located in areas exposed to natural hazards, such as seismic activity and excessive rainfall rate. Probable scenarios for the area of interest with respect to seismic activity, rainfall rate and maintenance will be investigated with the model presented, in order to determine the risk levels in monetary terms.

One of the central motivations for carrying out this investigation was the scarcity of systematic research to date, to comprehensively examine the variables that influence dam failures, describe their interactions and combine them into a model that allows to assess risks and to analyse risk reduction alternatives. The literature reports various studies within the dam industry, centred on the analysis of specific failure modes, but very few on mathematical models for dam risk assessment, that make use of continuous Bayesian Networks (BNs). Precisely, the originality of this work is that it utilises this powerful tool to model continuous variables, avoiding the cumbersome work implied when using their discrete counterparts.

Overall, the study aims to develop a quantitative model to assist dam engineers, in particular those in Mexico, on their risk assessment practices. Although the main emphasis of this paper is on the application of the model to the case of seven earth dams in the State of Mexico, in an accompanying paper, the non-parametric approach for continuous BNs and its quantification techniques through expert judgements is discussed (Morales-Napoles, Delgado-Hernandez, De-Leon-Escobedo, & Arteaga-Arcos, 2012).

The following section describes the selection process of the seven structures in central Mexico, which were used as a basis for carrying out the research. A review of risk assessment methodologies and their application in the sector, along with their advantages and limitations will then be highlighted. Next, the model constituent elements will be described in detail together with its use in one of the seven Mexican earth dams. Finally, the discussion, comparison of the results with previous failures, and concluding remarks will be presented. This work forms part of

a bigger research project to develop a comprehensive model for assessing risk in various types of dams.

### Dams in the State of Mexico

In order to develop a quantitative model for risk assessment, some dams located in the State of Mexico, a territory in the heart of the country (just next to Mexico City), were chosen. The criteria for such a selection were as follows: (i) height (between 15 and 30 m), (ii) age (more than 30 years old), (iii) construction material (earth dams) and (iv) location with exposure to hazards (e.g. earthquakes and excessive rainfall). These conditions, according to international dam failures statistics (Foster, Fell, & Spannagle, 2000; ICOLD, 1995), have significant influence in collapse events. Therefore, it was believed that the earth dams in the central Mexican province fulfilling these circumstances were special candidates for risk analysis.

Among the 56 dams built in the State of Mexico before 1976 (30 years prior to starting the current study, which began in 2006), seven structures were identified. Their names are: Embajomuy, San Joaquín, José Trinidad Fabela, Dolores, José Antonio Alzate (San Bernabé), Ignacio Ramírez (La Gavia) and El Guarda, details of which can be found in SRH (1976). Their heights range from 15 to 24 m, their ages from 36 to 66 years, their capacities from 0.55 to 52.5 million  $m^3$ , and all of them are exposed to the same hazards due to their close locations. Irrigation, flooding prevention and hydroelectric power generation can be remarked among their main purposes.

After visiting each structure, it became evident that maintenance activities were not frequent in the vast majority of them. Because of its relative position with respect to inhabited communities downstream, the José Antonio Alzate dam is perhaps the most important structure of the ones under study. Figure 1 shows two recent photographs, the first pointing up the embankment's downstream slope, and the latter both the reservoir and the spillway. It should be noted that the water surface close to the embankment's upstream slope is totally covered by aquatic iris. This is mainly because of the discharges of the nearby industrial complex, which also results in odour pollution in the zone.

Regarding its objectives, design values, dimensions and construction materials, the dam was built for power generation and irrigation in the Lerma River, with 27.3 million  $m^3$  of design capacity, 17.2 million  $m^3$  of overflow capacity, 52.5 million  $m^3$  of total reservoir capacity, height of 24 m, length of 282 m and a free board of 1.8 m. With regard to the embankment, it is composed of a combined volume of 168,000  $m^3$  of impervious core (compacted with heavy equipment),

sand transition zones, rockfill shell, and an upstream concrete block between the embankment and the clay foundation (see Figure 2). The chute spillway is located on the left-hand side, and has a maximum capacity of  $254 \text{ m}^3/\text{s}$ , crest length of 75 m and maximum water surface elevation of 1.42 m.

In terms of the other six structures, they all have similar design characteristics to those of the José Antonio Alzate, which can be found in SRH (1976), being the main difference the amount of people living downstream. In fact, for these dams, it is estimated that the existing population in potential risk range from a few dozens to 20,000 people. With regard to infrastructure, it is common for them to have highways, electrical transmission towers and, again, various settlements downstream. In addition, there are corn farms and other lands used for agricultural purposes. Having described the selection process for the seven earth dams of interest, some risk assessment methodologies will now be presented, highlighting their advantages and drawbacks.

#### Literature review on risk assessment methodologies

The literature reports various studies within the dam industry, centred on the use of specific methods to

perform dam risk assessments. Tree events, influence diagrams and Monte Carlo simulation were applied by Stedinger, Heath, and Thompson (1996). Within their framework, they were able to establish the relation between rainfall depths as well as both fatalities and expected damages to property.

Similarly, McCann, Franzini, Kavazanjian, and Shah (1985) summarised the Federal Emergency Management Agency/Stanford approach which is a tool for generating preliminary dam safety evaluations based on quantitative observations of the structure state. Basically, this instrument allows the estimation of the probability of dam failure as a frequency per unit time. Fell, Bowles, Anderson, and Bell (2000) reviewed the methods for estimating dam failure probabilities, and concluded that the most commonly used were: failure modes identification and probability analysis.

Failure modes can be obtained from historical dam collapses, and used for constructing fault trees (Fell et al., 2000). Another option is to employ expert judgment (Johansen, Vick, & Rikartsen, 1997) to generate the distributions of variables for which hard data do not exist. For instance, it has been utilised to estimate the probability of dam failure due to piping (Fell et al., 2000). Formal probabilistic methods have



Figure 1. Recent photographs of the José Antonio Alzate dam (San Bernabé). (a) Downstream slope (embankment). (b) Reservoir and spillway.

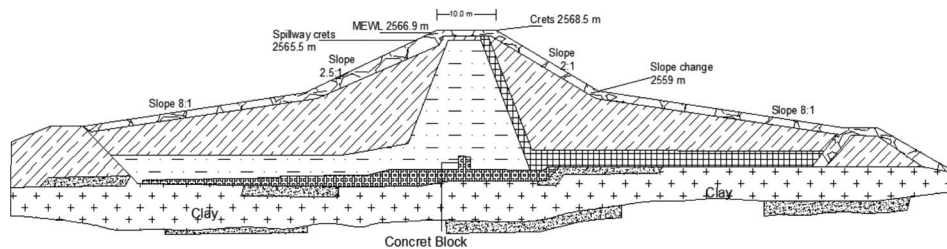


Figure 2. José Antonio Alzate dam (San Bernabé).

also been used in the dam industry for calculating the probability of failure due to loss of global stability (see, for example: Mostyn & Fell, 1997).

Another tool that could be used in the exercise is a BN. Bayesian networks are directed acyclic graphs whose nodes represent random variables and whose arcs (arrows) depict probabilistic relationships amongst them. The direct predecessors (successors) of a node are called parents (children). When all nodes in a BN are discrete (Pearl, 1988) probability tables need to be specified for each node. Marginal distributions are specified for each node without parents, and conditional probability tables are specified for each node given its parents.

One disadvantage of discrete BNs is, however, that the number of parameters to be specified for each node grows exponentially with the number of states of each variable and its parent set. Therefore, the quantification of models that use continuous quantities reduced to discrete BNs, imposes difficulties for the reason stated above. Fault trees may be specified by discrete BNs and hence, its limitations are similar to those described.

In order to avoid this drawback, BNs that can handle continuous nodes have been developed for joint normal variables, which are known as Gaussian BNs (Shachter & Kenley, 1989). In these kinds of models, the arcs in the BN represent the partial regression coefficients applied to the normal units of the transformed variables, and not to the original units. As will be seen later, the variables entering the model are non-normal. When statistical data are scarce, modelling arcs as partial regression coefficients, as explained in the previous sentences, place a heavy burden on expert elicitation processes which is a disadvantage of the approach.

Non-parametric BNs have been developed to cope with some limitations that both discrete and Gaussian BNs present for handling continuous variables (Hanea, Kurowicka, & Cooke, 2006; Kurowicka & Cooke, 2006). In fact, they have been applied in a large-scale model for air transport safety (Ale et al., 2007). Here, nodes represent continuous and/or discrete random variables and arcs represent rank and conditional rank correlations between parents and children.

The specification of the graph, the marginal distributions for each node, and the rank and conditional rank correlations attached to the edges of the graph, together with a copula for which zero correlation entails independence, are sufficient for the construction of the joint distribution. Once such a distribution is available, it may be updated whenever evidence is available to the analysts. Alongside, techniques for the elicitation of marginal distributions from experts have been developed before (Cooke,

1991). Additionally, methods for the elicitation of dependence information in the form of conditional and unconditional rank correlations have been proposed and used elsewhere (Morales, Kurowicka, & Roelen, 2008).

A similar technique as the one presented in Morales et al. (2008) was used in this study and is presented in the accompanying paper (Morales-Napoles et al., in press). It consists of constructing the dependence structure on the basis of asking experts about ratios of unconditional rank correlations. Because of its advantages, it was then decided to use non-parametric BNs to build the model presented here. Whenever data were not available, the previously described techniques for expert elicitation were used. In the next section, the model variables will be presented.

### Model variables

Once the literature on earth dam failures was reviewed, 10 variables were identified as relevant for this study (Foster et al., 2000; ICOLD, 1995). They are summarised in Table 1, along with their operational descriptions, the source of data used to calculate their marginal distributions and the measurement units selected for quantifying them. Note that some of the available statistical data consisted only of short records (e.g. earthquakes – 8 years; rainfall rate – 37 years), which may not give a precise assessment of the hazard, but it was felt that this information would enable important insights and results to be obtained in the risk assessment exercise. It is worth noting here that human costs were quantified by the consulted experts in current USD, because alternative methodologies to estimate them were considered to be time-consuming and costly. This allowed to combine both human and economic costs in a variable called ‘total costs’. These total costs refer to the concept ‘life-cycle costs’, in order to include all potential costs and consequences of the undesirable dam behaviour within the dam lifetime.

There are more factors than those presented in Table 1 that can influence a dam’s failure, e.g. cracking, static loading, structure problems and soil liquefaction on foundation to name a few (Moser, 2001). However, they have not been considered in this research, because international statistics have shown that earth dam collapses are caused, generally, by those already summarised in the Table 1 (Foster et al., 2000; ICOLD, 1995). Moreover, the consulted experts ratified that cracking problems were unlikely in the seven Mexican structures due to the features of their construction materials, and that aspects such as liquefaction had been considered in their design. Also environmental costs were left out intentionally, mainly

Table 1. Model variables and their descriptions.

Variable name	Operational description	Source of data	Units
(1) Seismic frequency	It is related to the distribution of earthquakes $\geq 5.5$ per year, in Richter magnitude scale, between 2000 and 2008 for the locations of interest	Mexican National Seismographic System	(# earthquakes $\geq 5.5$ )/year
(2) Rainfall rate	It refers to the average value of the seven-basin-five-days moving averages per year	'ERIC' Mexican database from 1961 to 1998	mm/day/year
(3) Maintenance	It is the number of years between maintenance activities, which would lead the dam to an as good as new condition	Expert judgment	# years between maintenance activities
(4) Loss of global stability (safety factor)	It is the distribution of the factors of safety (resisting moment/causing moment), for each of the seven dams, based on their design geometrical features. The 'Swedish method' is used for calculating such factors	Original design data (SRH, 1976)	Unit less
(5) Overtopping	It is the water level, from the crest, during an event in which such a level may increase beyond the total embankment height	Expert judgment	mm
(6) Piping	It is the distribution of water flowing through the embankment that causes its internal erosion (apart from the spillway and outlet pipe torrents)	Expert judgment	lt/sec
(7) Breaching	It refers to the distribution of the average breach width, i.e. the mean of both superior and inferior breach widths, due to embankment's crest erosion. Calculated with the methods reported in Wahl (1998)	Original design data (SRH, 1976)	m
(8) Flooding	It is the average water level, per day, in the downstream flooded area, during a dam failure event per year	Expert judgment	mm/day/year
(9) Economic cost	It refers to both public and private total costs, due to all possible damages in infrastructures (e.g. schools, hospitals, bridges, roads, transport systems), fields (e.g. farms, crops), housing, supply, commercial and entertainment centers, caused by a flooding, consequence of a dam failure. It includes	Expert judgment	\$ (USD)/year

(continued)

Table 1. (Continued).

Variable name	Operational description	Source of data	Units
(10) Human cost	<p>disruption costs, and is measured in current USD/year</p> <p>Since human life is invaluable, estimating this kind of cost is a difficult task, in such a way, it is mainly related to the cost due to the payment of pensions and compensations, and to both public and private total costs, over a time period equivalent to the maximum human remaining life span, due to all possible damages, health and life losses, caused by a flooding, consequence of a dam failure. It is measured in current USD/year</p>	Expert judgment	\$(USD)/year

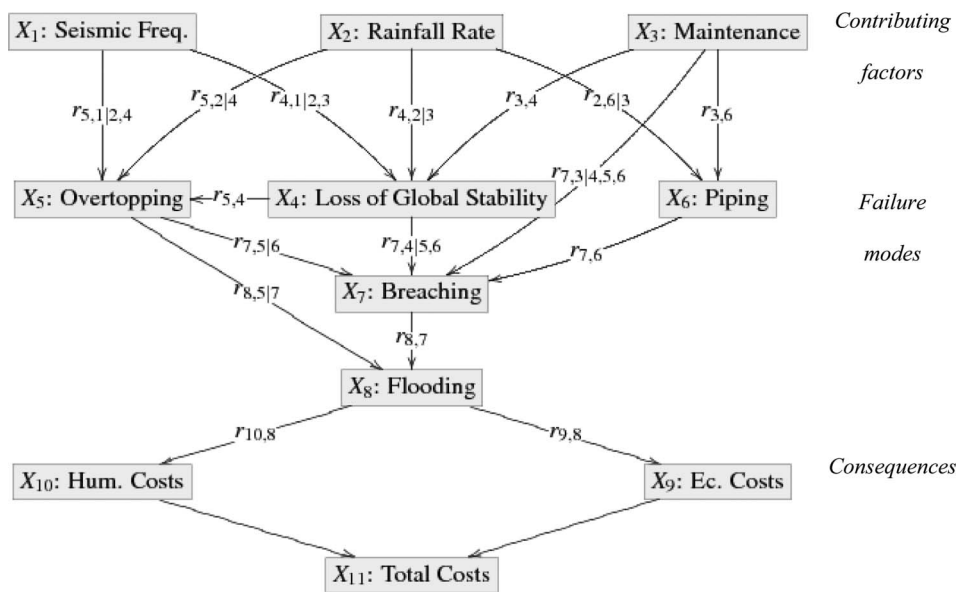


Figure 3. Model for earth dam's risk assessment.

due to the difficulty of measuring them. Nevertheless, future research should explore this aspect in detail.

Having identified and analysed the variables, they were grouped into three categories: (i) contributing factors (seismic frequency, rainfall rate and maintenance), (ii) failure modes (loss of global stability, piping, overtopping and breaching) and (iii)

consequences (flooding, human and economic costs). Such a classification helped to generate the model. Figure 3 shows a scheme of the proposed model, which includes the variables recognised. The figure was taken from UNINET, an uncertainty analysis software package for managing both discrete and continuous distributions, using BNs (Hanea, 2008). As can be

seen, the nodes are linked by arrows that represent rank and conditional rank correlations between variables. The rank correlation between two variables is the product moment correlation applied to the distribution function of each variable. The conditional rank correlation is the same as a rank correlation, but measured in the conditional distribution.<sup>1</sup>

For example, looking at the values of Table 2: rainfall rate (2) has a direct impact not only on overtopping ( $r_{5,2|4}$ ) (5), but also on the loss of global stability ( $r_{4,2|3}$ ) (4) and piping ( $r_{2,6|3}$ ) (6). Observe that all these three influences are modelled as conditional rank correlations. All conditional variables are shown next to the small vertical line included in the correlation sub-index. Thus,  $r_{4,2|3}$  means ‘correlation between the loss of global stability (4) and rainfall rate (2), given the presence (or absence) of annual maintenance’ (3). Similarly, flooding (8) has implications for economic (9) and human costs (10). In this case, both are represented as unconditional rank correlations as  $r_{9,8}$  and  $r_{10,8}$ , respectively.

The arcs between both human and total costs, and economic and total costs lack rank correlations, because the total costs are simply the sum of human and economic expenses and hence the relationship is functional (Graham, 2001). It should be recognised that there might be more interactions among the nodes in the previous figure. However, they have not been expressed in the model because (correlations being small) their exclusion does not affect the patterns of dependence between the main variables.

The quantification of rank and conditional rank correlations was carried out by means of structured expert judgments. Details are provided in the accompanying paper. With reference to the one dimensional margins of each variable, the following paragraphs describe the process employed to produce two of them, i.e. seismic frequency (based on statistical data) and maintenance (based on expert judgment). Note, however, that similar steps were followed to build the other eight distributions.

**Seismic frequency**

To generate the one dimensional marginal distribution for the seven structures, and considering that they all are located in a zone where seismic activity is frequent (DDF, 1988), data from the Mexican National Seismographic System limited to the 2000–2008 period were utilised, mainly because no access to a more extended record was feasible. Basically, the database contains the magnitude of all those movements registered in the stations positioned in the area of interest, regardless of their intensity. Hence, the number of earthquakes  $\geq 5.5$  per year, in Richter

Table 2. Statistical values for the variables’ cumulative distributions.

Variable	Units	Mean	Standard deviation	5th percentile	50th percentile	95th percentile
(1) Seismic frequency	(# earthquakes $\geq 5.5$ )/year	6.6790	2.1036	3.6325	6.4714	10.492
(2) Rainfall rate	mm/day/year	5.6852	3.6329	0.4700	5.4300	12.2500
(3) Maintenance	# years between maintenance activities	15.1320	13.9780	1.3577	11.2710	40.9880
(4) Loss of global stability (safety factor)	Unitless	1.6569	0.2582	1.2030	1.6790	2.0424
(5) Overtopping	mm	-3806.6000	2617.0000	-7949.0000	-3528.0000	-74.7100
(6) Piping	lt/seg	19.7710	64.4150	2.7305 E-5	0.2028	94.9730
(7) Breaching	m	8.2943	14.1710	9.4861 E-6	1.3213	44.1310
(8) Flooding	mm/day/year	1831.4000	4517.3000	0.0159	299.3500	6403.2000
(9) Economic cost	\$ (USD)/year	1.4097 E7	2.7059 E7	3045.0000	3.3491 E6	5.3009 E7
(10) Human cost	\$ (USD)/year	2.9576 E7	3.4151 E7	4.8851 E5	2.0371 E7	8.6550 E7



magnitude scale, was used to construct the distribution presented in Figure 4. This marginal distribution is essentially discrete, however, a continuous distribution to approximate it has been proposed. The fit is carried out by finding the parameters of a number of distributions which minimise the sum of squared difference with respect to the empirical distribution. Then, the one with the smallest sum of squared difference is chosen.

As the figure illustrates, in any one year of the eight analysed, at least one and no more than 21 earthquakes  $\geq 5.5$  have been registered in the region. In fact, the mean for the probability density distribution was about 6.7 and the standard deviation was 2.1. Moreover, the 5th, 50th and 95th percentiles of the distribution are reported. In this case, the 5th percentile was 3.63 which means that in only 5% of the years, there are less than 3.63 earthquakes  $\geq 5.5$  in the area. Similarly, it is expected that in 50% of the time, there are less than 6.47 movements above that magnitude and in 95%, they will not exceed 10.49 events. This is not surprising since the zone of attention is known within the country as one of the most dynamic, in terms of seismic activity.

### Maintenance

To create the probability distribution for the Mexican dams, the classical method for expert judgment was used (Cooke, 1991). It is important to state that the

question used for generating the information in this case was:

What is the current (correct but unknown) number of years between maintenance activities, that would lead any of the seven dams of interest, to an as good as new condition?. To express the uncertainty associated with the response, please provide the 5th, 50th, and 95th percentiles of your estimate.

Figure 5 presents the results of the investigation. This distribution (as all others obtained with the classical method for expert judgment) uses the equal weight combination, i.e. each expert's opinion has the same weight. For an overview of the classical method, see Section 'Variables of interest' in the accompanying paper. For details about the method, see the references therein (Morales-Napoles et al., in press). The data from four Mexican specialists were processed in Excalibur software. The mean, standard deviation, 5th, 50th and 95th percentiles of this variable are shown in Table 2 along with the correspondent values for the other nine variables that were under investigation. Now that all variables and their respective distribution values have been summarised, the application of the complete model will be discussed.

### Discussion of the results

Making use of UNINET, the model was manipulated. It should be noted that its final version can predict, or

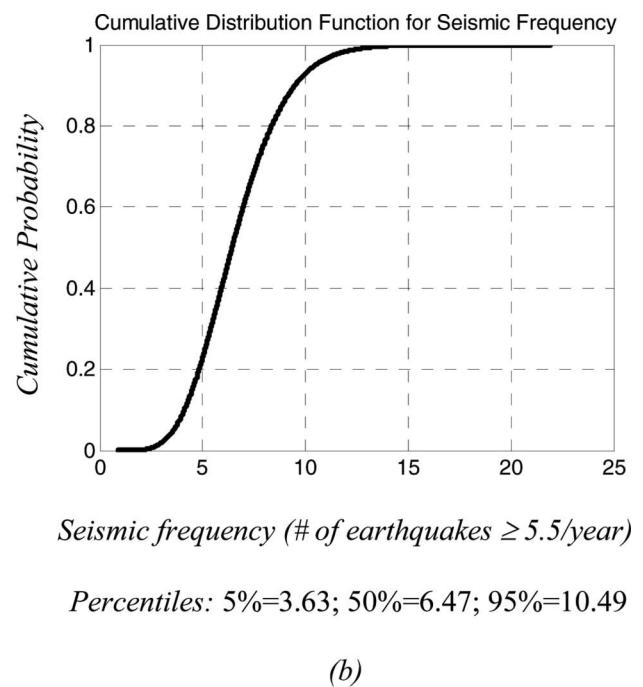
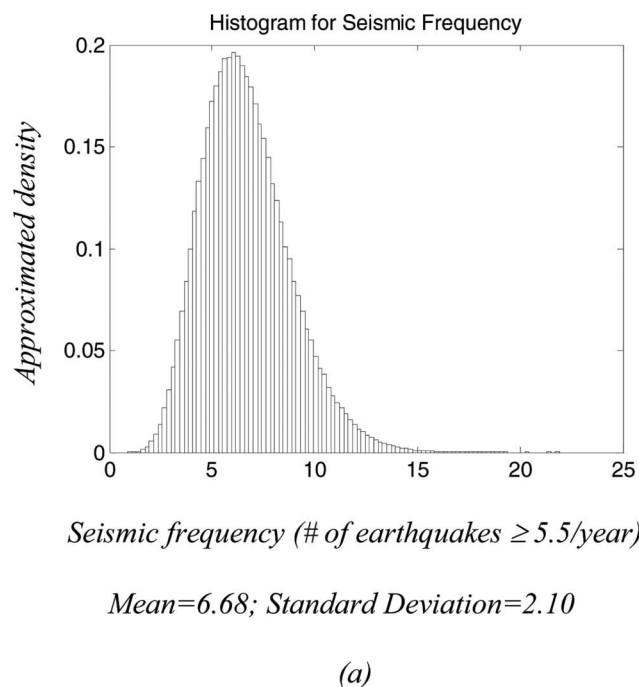
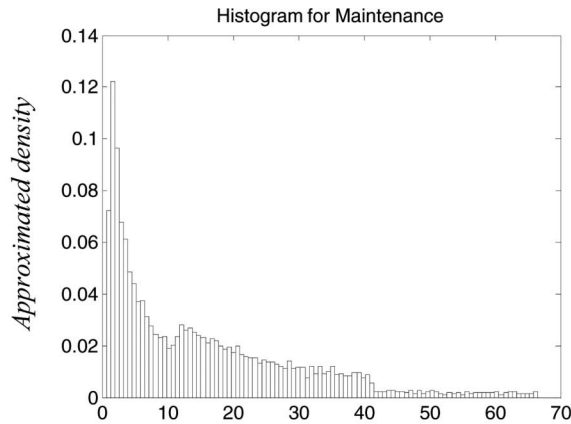


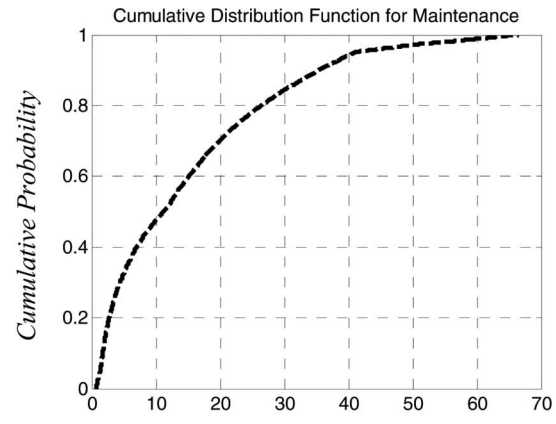
Figure 4. Seismic frequency probability density and cumulative distribution for the seven dams under study.



Number of years between maintenance activities, which would lead the dam to an as good as new condition

Mean=15.13; Standard Deviation=13.98

(a)



Number of years between maintenance activities, which would lead the dam to an as good as new condition

Percentiles: 5%=1.36; 50%=11.27; 95%=40.99;

(b)

Figure 5. Maintenance probability density and cumulative distribution for the seven dams under study.

diagnose, the performance in any of the seven Mexican structures, as it was built considering data from all of them simultaneously. Although the model was also intended to better understand the failure mechanisms of the structures under analysis (e.g. the way in which seismic loading may produce loss of global stability), such mechanisms are not comprehensively explained in the paper. Therefore, further analysis is required to determine, for instance, whether more warning should be available for a flood breach than for an overtopping event. To limit the explanation, its use will only be illustrated here, with information related to the José Antonio Alzate dam, i.e. the one previously presented in Figure 1.

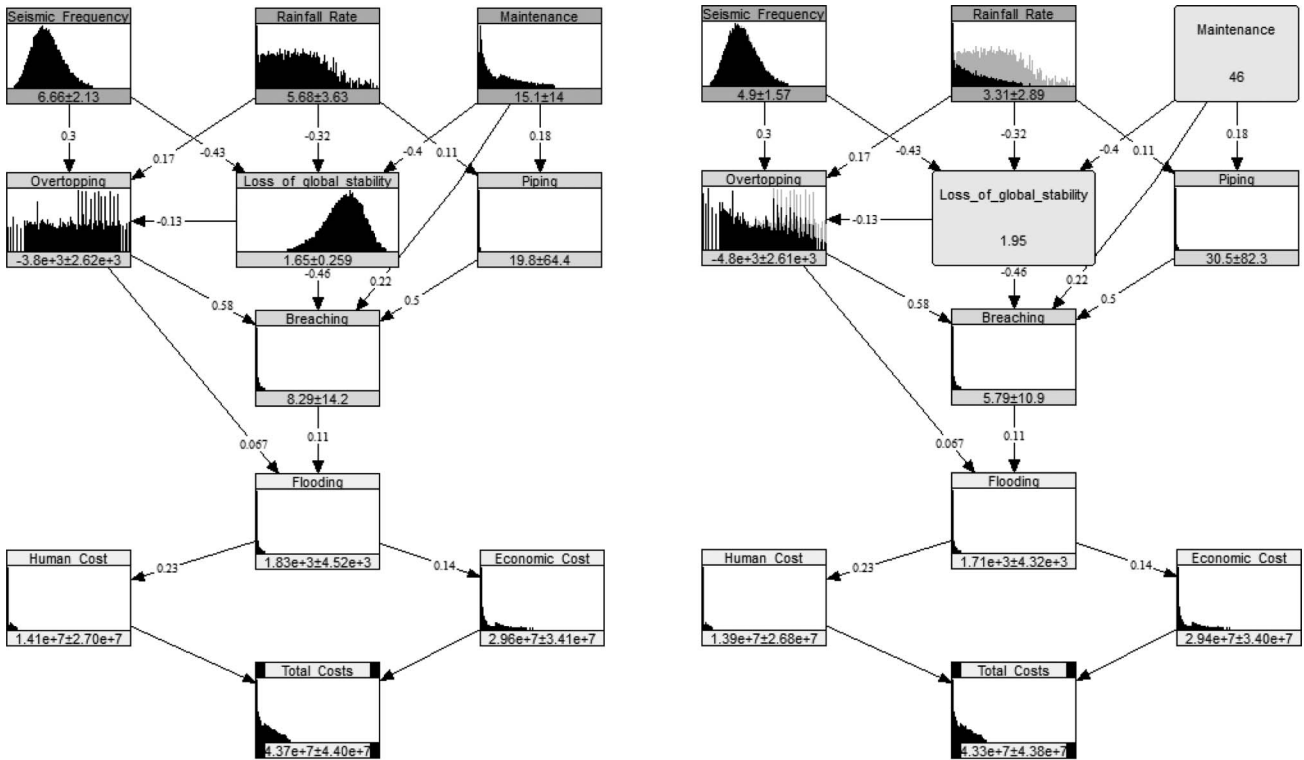
Because of its geometry and year of construction, there are two variables that can immediately be fixed for the dam under analysis. Firstly, the loss of global stability can be quantified by means of the safety factor of the structure, i.e. 1.95, which is to be used as the 'observed value' in the model (see cumulative distribution values of the loss of global stability in Table 2, 1.95 is between the 50th and 95th percentiles). Secondly, the age of the dam is 46 years, which can be associated with the number of years between maintenance activities. Assuming that there has not been any conservation actions since its final construction year (see the distribution values for maintenance in Table 2, 46 is above the 95th percentile), the 'maintenance observed value' can then be established as 46. Figure 6a presents the original model, and Figure 6b presents its adaptation to the dam of interest

with both values, i.e. loss of global stability and maintenance.

When the two abovementioned values are fixed, the probability density distribution of all the variables interrelated with them, are updated. This is the reason why there is a 'shadow' in some nodes, showing the original distribution and the one re-calculated (e.g. rainfall rate and overtopping in Figure 6b). To demonstrate more clearly how the model can be used, suppose that a hypothetical extraordinary rainfall rate of 15 mm/day takes place in the dam zone and, concurrently, it is known that the seismic frequency in the region corresponds to eight earthquakes  $\geq 5.5$  per year. Figure 6c shows the results of the exercise. As can be seen, the anticipated flooding value has increased from 1.71 m/day (Figure 6b) to 2.22 m/day (Figure 6c).

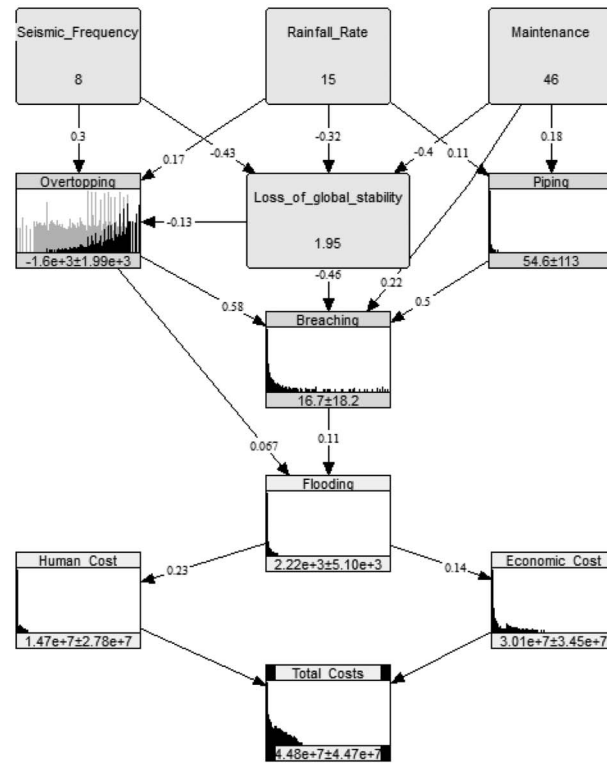
Similarly, the predicted human cost moved from \$ 13.9 to \$ 14.7 USD million, and the economic loss in turn from \$ 29.4 to \$ 30.1 USD million. This means that the intensification of rain, and the presence of earthquakes at the same time, are expected to produce higher levels of water in the potential flood area and, consequently, superior amounts of both human and economic losses. Note that implicitly, risks are compared by the annualised consequences obtained by the model, while hazard exposures, probability of failure and costs have been considered in the BN.

Since one of the main objectives of this research was to determine which of the seven dams required more attention, a risk assessment exercise was carried out. In the event, the impact of a hypothetical



(a)

(b)



(c)

Figure 6. Original and complete model adapted to the case of the José Antonio Alzate dam. (a) Original model (applicable to the seven dams). (b) Model for José Antonio Alzate dam. (c) Complete model adapted to the case of the José Antonio Alzate dam.

overtopping incident of 100 mm was employed to analyse its effects not only in the flood water level downstream, but also in human and economic costs. So, for each of the seven structures, three values were set, i.e. loss of global stability (factor of safety), maintenance (dam age using the assumption above mentioned) and overtopping (100 mm). Again, while this is a hypothetical example, the total costs showed that the two dams demanding more consideration were: 'San Joaquín' and 'José Trinidad Fabela', as they obtained the higher consequence costs i.e. \$46,246,000 USD and \$46,186,000 USD, respectively. In contrast, the two relatively best ranked with the lower total costs, and therefore needing less concern, were: 'José Antonio Alzate' (\$45,574,000 USD) and 'El Guarda' (\$45,525,000 USD). Table 3 summarises the values generated by the model, which helped to prioritise the seven dams based on the amount of expected consequences.

The mean was used in all cases to perform the evaluation, and the total costs for all seven cases were compared against the unconditional value, which refers to the situation where no data about any particular dam are available (see second to last row of Table 3), and the BN has the default values. As it can be observed, the 'worst' dam (San Joaquín) would be expected to have 5.91% more costs than the unconditional scenario. On the other extreme, the 'best' dam (El Guarda) percentage corresponds to 4.2% above the reference value (\$43,663,000 USD). It was no surprise to see that the costs were very similar, as the infrastructure and population features downstream the dams shared common characteristics, a fact ratified not only by the experts but also by the researchers during the visits to the structures.

The exercise was repeated now using an overtopping value of 700 mm, but the difference between the 'best' and the 'worst' dam still remained close to the 1.6% (\$46,246,000/\$45,525,000 with 100 mm versus \$47,023,000/\$46,293,000 with 700 mm) already calculated. Therefore, based on the anticipated consequences for each structure, the dams should be given the priority shown in the last row of Table 3. It is important to note that 'San Joaquín' dam had the lowest factor of safety and, concurrently, it is almost the oldest structure analysed. With regard to 'El Guarda', it has the second best factor of safety and is the youngest of the sample used, so these results were expected. The model could be used in a similar fashion to perform more analysis. In fact, a wide variety of scenarios could be constructed to determine the level of impact of other particular incidents (such as piping or breaching), or a combination of them, in the expected consequences (i.e. total costs). To do so, the analyst should use the model as previously illustrated.

Table 3. Summary of results (mean value).

Dam	Unconditional (UC)	José Antonio Alzate	El Guarda	Dolores	Embajomuy	José Trinidad Fabela	Ignacio Ramirez	San Joaquín
Loss of global stability (factor of safety)	1.65	1.95*	1.92*	1.85*	1.75*	1.37*	1.36*	1.31*
Maintenance (age, in years)	15.1	46*	36*	46*	66*	63*	43*	64*
Overtopping (mm)	-3806	100*	100*	100*	100*	100*	100*	100*
Flooding (mm/day)	6398.8	11,597	11,510	11,816	12,558	13,475	12,772	13,563
Human cost (USD)	\$14,090,000	\$15,083,000	\$15,057,000	\$15,136,000	\$15,293,000	\$15,405,000	\$15,334,000	\$15,437,000
Economic cost (USD)	\$29,573,000	\$30,491,000	\$30,468,000	\$30,539,000	\$30,679,000.00	\$30,781,000	\$30,717,000	\$30,809,000
Total costs (USD)	\$43,663,000	\$45,574,000	\$45,525,000	\$45,675,000	\$45,972,000	\$46,186,000	\$46,051,000	\$46,246,000
Comparison against UC	1	1.0437	1.04264	1.0460	1.0528	1.0577	1.0546	1.0591
Priority	-	6	7	5	4	2	3	1

Note: \*Evidence entered in the model in Figure 6a to compute the conditional means.

### Comparison with documented failures

It is worth comparing at this point, the model results with previous and documented earth dam failures. None of the events analysed, as far as the authors are concerned, has been registered in the State of Mexico. Nevertheless, there was a recent collapse in Hawaii, on 14 March 2006 (ASDSO, 2007), in which the ‘Kaloko’ earth dam, a 14 m of height structure built in 1890 with a capacity of 1.5 million m<sup>3</sup>, failed because of heavy rains causing seven fatalities and the following damages and related costs:

- (1) 33 USD million for disaster aid.
- (2) 20 USD million for repairing damaged infrastructure (Kuhio highway).

Thus, in total, \$53 USD million was spent to improve the situation, a quantity analogous to that obtained in this work. For instance, from Table 3, the mean values for human and economic costs, in the ‘San Joaquín’ case (age: 64 years, total capacity: 1 million m<sup>3</sup>, height: 20 m) were \$15.4 USD million and \$30.8 USD million in that order. The sum of these two

amounts equals \$46.2 USD million, a quantity comparable to that registered in the documented case. Although there were more costs associated with the Hawaiian tragic event, they have not been considered here for the comparison, because most of them were related to environmental costs (e.g. compensating sugar growers), which were not taken into account, as already stated, in the model.

The Kaloko dam’s case is useful because the magnitude of the population and infrastructure downstream are very similar to those present in the structures located in the State of Mexico, i.e. agricultural fields, communication roads, electrical infrastructure and spread houses. Leaving aside the Kaloko example, related cases have been reported by the Association of State Dam Safety Officials (ASDSO, 2007), which has created a database that contains information about dam failure incidents in the US from 1874 to 2007. While not all the events have occurred in dams such as the ones studied here, the costs reported by ASDSO can also be considered equivalent, ranging from a few thousand USD to \$400 USD million. Table 4 summarises some of the failures

Table 4. Sample of documented earth dam failures in the US and Mexico.

Dam	Country	Construction year	Failure year	Height (m)	Failure mode	Costs (million USD)
South Fork, Pennsylvania	USA	1838–1853	31 May 1889	22	Piping and breaching	> 17
Austin, Pennsylvania	USA	1909	September 1911	15	Overtopping	N/A
Castlewood, Colorado	USA	1890	3 August 1933	21	Overtopping	1.7
Baldwin Hills, California	USA	1947–1951	14 December 1963	20	Breaching	11
Buffalo Creek, Virginia	USA	1945–1947	26 February 1972	14	Overtopping	> 50
Kelly Barnes, Georgia	USA	1899	6 November 1977	12	N/A	30
Lawn Lake and Cascade Lake, Estes Park, Colorado	USA	N/A	15 July 1982	8 y 11	N/A	25
Evans and Lockwood, Fayetteville, North Carolina	USA	N/A	15 September 1989	10 y 10	N/A	10
Silver Lake and Tourist Park, Marquette, Michigan	USA	N/A	13 May 2003	10 y 8	N/A	102
Big Bay Lake, Purvis, Mississippi	USA	N/A	12 March 2004	17	Piping	N/A
Upper Jones Tract, Stockton, California	USA	N/A	3 June 2004	15	N/A	> 90
Kaloko, Kauai, Hawaii	USA	1890	14 March 2006	30	Overtopping	> 67
Needwood, Gaithersburg, Maryland	USA	1946	29 June 2006	21	Rainfall excess	N/A
El conejo y La llave, Guanajuato	Mexico	N/A	1975	N/A	Overtopping	N/A
Laguna I, Necaxa, Puebla	Mexico	N/A	1969	N/A	Piping and breaching	N/A
Laguna II, Veracruz	Mexico	N/A	1972	N/A	Piping and breaching	N/A
La Escondida, Tamaulipas	Mexico	N/A	1972	N/A	Breaching	N/A
Pescaditos, Oaxaca	Mexico	N/A	1953	N/A	Sliding	N/A
El Estribón, Jalisco	Mexico	N/A	1970	N/A	Sliding	N/A
Santa Ana, Hidalgo	Mexico	N/A	1952	N/A	Piping	N/A
Abelardo L. Rodríguez, Sonora	Mexico	N/A	N/A	N/A	Sliding	N/A
El Azúcar, Tamaulipas	Mexico	N/A	N/A	N/A	Sliding	N/A
Unión-Calera, Guerrero	Mexico	N/A	N/A	N/A	Sliding	N/A
Necaxa, Puebla	Mexico	N/A	1909	N/A	Other causes	N/A
Tlalpujahua, State of Mexico	Mexico	N/A	1940	N/A	Other causes	N/A
Cacaloapan, Puebla	Mexico	N/A	1952	N/A	Other causes	N/A

encountered in the US, and some in Mexico (Delgado-Hernandez et al., 2009). When the information was Not Available, an 'N/A' was used.

Again, the consequence costs found are comparable to those produced with the model, and the failure variables have been present in Mexico. Unfortunately, the information about Mexican failures is limited as there seems not to be systematic records of the events. Nevertheless, these results show that the model provides a rational method to assess risk and to quantify it with subjective probabilities, combined with more objective marginal probabilities by means of a BN. This is an important outcome for the involved Mexican dam owners, and it is hoped that they can make better decisions when the time to allocate maintenance budget for their structures come. In short, they can prioritise each construction and generate a preference list of dams requiring maintenance, based on the results generated with the model's simulation of various scenarios.

### Conclusions

This document has dealt with earth dams and their failure modes, emphasising risk assessment in a group of seven dams within the State of Mexico. After reviewing the risk literature, it became evident that the combination of BNs and expert judgment could lead to the development of a model for earth dams risk evaluation. The resultant model incorporates seismic frequency, rainfall rate, maintenance, overtopping, loss of global stability, piping, breaching, flooding, human and economic costs. Having quantified all 10 variables, by means of either statistical historical data or systematic expert judgment, their probability distributions were established. In fact, all of them were represented by continuous rather than discrete distributions, which means that this is probably the first model in the Mexican dam risk assessment literature employing continuous BNs.

Some scenarios were simulated with the model and the results were comparable to those reported in previous documented dam failures. Essentially, the seven Mexican dams could be initially prioritised so as to help decision makers in the resources allocation based on hard data. Therefore, it is strongly believed that the methodology utilised to build the model can be applied to carry out similar exercises in different locations. While the key objectives of this research have been achieved, there were a number of limitations associated with the work, which could be addressed in further research. Firstly, the available records for quantifying the variables were normally scarce. However, using the experts' judgement has proven to be an important aspect for collecting information. With

reference to the comprehensiveness and validity of the model in the dam community, more variables and data should necessarily be investigated in the future. In spite of this, it is strongly believed that the potential contributing factors and dam failure modes relevant to the context of the State of Mexico have been considered in the model.

Overall, this research has demonstrated that the use of continuous probability distributions, generated through either systematic expert judgment or statistical data, in Mexican dams' risk assessment is not only feasible but also beneficial. This work forms part of a bigger project aimed at developing a more comprehensive model applicable to different types of dams in the country and more infrastructures such as bridges.

### Acknowledgements

The authors would like to thank all four experts who took part in this study. The Autonomous University of the State of Mexico (UAEMex) and the State of Mexico's Council for Science and Technology (COMECyT) and the National Council for Science and Technology (CONACyT projects: 158225 & 145462), partially provided financial support to carry out this research. The second author was also partially financed by the TU Delft and TNO. The authors would also like to acknowledge students David Carlos Pérez-Flores, Benjamín Pérez-Pliego and José Emmanuel Rivero-Santana for their participation in the project.

### Note

1. More details are provided in the accompanying paper by Morales-Napoles et al. (in press) and references therein.

### References

- Ale, B.J.M., Bellamy, L.J., Boom, R., Van Der Cooper, J., Cooke, R.M., Goossens, L.H.J., Hale, A.R., Kurowicka, D., Morales, O., Roelen, A.L.C., & Spouge, J. (2007). Further developments of a causal model for air transport safety (CATS): Building the mathematical heart. Proceedings of ESREL, Stavanger, Norway, 25–27 June (pp. 1431–1439).
- ASDSO. (2007). *US Dam failures and incidents*. Lexington, KY: Association of State Dam Safety Officials. Retrieved from <http://www.damsafety.org>.
- Cooke, R., (1991). *Experts in uncertainty*. New York, NY: Oxford University Press.
- DDF. (1988). *Studies of seismic activity in the valley of Mexico* (In Spanish: *Estudios sobre sismicidad en el valle de México*). México: Secretaría General de Obras, Departamento del Distrito Federal.
- Delgado-Hernandez, D.J., Pérez, P.B., Morales, N.O., Pérez, F.D., Rivero, S. J., & De León, E.D. (2009). Failure of earth dams due to slope landslides and breaching, originated by earthquakes: the case of the State of Mexico. (In Spanish: *Falla de presas de tierra por deslizamiento de taludes y formación de brechas originadas por sismo: un caso en el Estado de México*). Paper presented at the XVII National Congress on Seismic Engineering, Mexican Society for Seismic Engineering, (Congreso Nacional de Ingeniería Sísmica, Sociedad Mexicana de Ingeniería Sísmica), Puebla, Puebla, México, 11–14 November.

- Donnelly, R. (2006). Safe and secure: Risk-based techniques for dam safety. *International Water Power and Dam Construction*, Retrieved from <http://www.waterpower-magazine.com/story.asp?storyCode=2040340>.
- Emiroglu, M.E., Tuna, A., & Aislan, A. (2002). Development of an expert system for selection of dam type on alluvium foundations. *Engineering with Computers*, 18(1), 24–37.
- Fell, R., Bowles, D.S., Anderson, L.R., & Bell, G. (2000). The status of methods for estimation of the probability of failure of dams for use in quantitative risk assessment. *Proceedings of the 20th congress, international commission on large dams*, Beijing, China, 19–22 September.
- Foster, M., Fell, R., & Spannagle, M. (2000). The statistics of embankment dam failures and accidents. *Canadian Geotechnical Journal*, 37, 1000–1024.
- Graham, W. (2001). Human and economic consequences of dam failure. *The National Dam Safety Program Research Needs Workshop: embankment dam failure analysis*, Department of Homeland Security, Federal Emergency Management Agency, Oklahoma City, 26–28 June (pp. 32–45).
- Hanea, A. (2008). *Algorithms for non-parametric Bayesian belief nets* (PhD thesis). Delft University of Technology, The Netherlands.
- Hanea, A., Kurowicka, D., & Cooke, R.M. (2006). Hybrid method for quantifying and analyzing Bayesian belief nets. *Quality and Reliability Engineering International*, 22, 709–729.
- ICOLD. (1995). *Dam failure statistical analysis* (Bulletin 99). Paris, France: International Commission on Large Dams (ICOLD).
- ICOLD. (2008). *International Commission of Large Dams, ICOLD*. Retrieved from <http://www.icold-cigb.net/>.
- Johansen, P.M., Vick, S.G., & Rikartsen, C. (1997). Risk analysis for three Norwegian dams. In E. Broch, D.K. Lysne, N. Flatb., & E. Helland-Hansen (Eds.), *Hydropower 97* (pp. 431–442). Rotterdam: Balkema.
- Kurowicka, D., & Cooke, R.M. (2006). *Uncertainty analysis with high dimensional dependence modelling*. Series in Probability and Statistics. Chichester, UK: Wiley.
- McCann, M.W., Franzini, J.B., Kavazanjian, E., & Shah, H.C. (1985). *Preliminary safety evaluation of existing dams* (Vols 1–2, Report No. 69). Stanford, CA: John A. Blume Earthquake Engineering Center, Stanford University.
- Modarres, M., Kaminskiy, M., & Krivtsov, V. (1999). *Reliability engineering and risk analysis*. Boca Raton, FL: CRC Press.
- Morales, O., Kurowicka, D., & Roelen, A. (2008). Eliciting conditional and unconditional rank correlations from conditional probabilities. *Reliability Engineering & System Safety*, 93, 699–710.
- Morales-Napoles, O., Delgado-Hernandez, D.J., De-Leon-Escobedo, D., & Arteaga-Arcos, J.C. (2012). A continuous Bayesian network for earth dams risk assessment: Methodology and quantification. *Structure and Infrastructure Engineering*.
- Moser, D. (2001). Dam safety decisions: Current USACE practice. *The National Dam Safety Program Research Needs Workshop: embankment dam failure analysis*, Department of Homeland Security, Federal Emergency Management Agency, Oklahoma City, 26–28 June.
- Mostyn, G.R., & Fell, R. (1997). Quantitative and semi quantitative estimation of the probability of landsliding. In D. Cruden & R. Fell (Eds.), *Landslide risk assessment* (pp. 297–316). Rotterdam: Balkema.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufmann.
- Pistrika, A., & Tsakiris, G. (2007). Flood risk assessment: A methodological framework. *Water resources management: New approaches and technologies*, European Water Resources Association, Chania, Crete, Greece, 14–16 June (pp. 1–10).
- Shachter, R.D., & Kenley, C.R. (1989). Gaussian influence diagrams. *Management Science*, 35, 527–550.
- SRH. (1976). *Dams built in Mexico* (In Spanish: *Presas Construidas en México*). México: Secretaría de Recursos Hidráulicos.
- Stedinger, J.R., Heath, D.C., & Thompson, K. (1996) *Risk analysis for dam safety evaluation: Hydrologic risk*. U.S. Army Corps of Engineers USACE) (IWR Report 96-R-13). Water Resources Support Center, Institute of Water Resources Alexandria, VI.
- Wahl, T.L. (1998). *Prediction of embankment dam breach parameters* (Report DSO-98-004). Lexington, KY: Dam Safety Office.