UNIVERSITÉ DE SHERBROOKE
Faculté de génie
Département de génie civil

ÉVALUATION DE LA FRAGILITÉ SISMIQUE DES BARRAGES POIDS EN UTILISANT DES FONCTIONS MULTIVARIÉES BASÉES SUR DES MÉTA-MODÈLES

SEISMIC FRAGILITY ASSESSMENT OF CONCRETE GRAVITY DAMS USING META-MODEL-BASED MULTIVARIATE FUNCTIONS

Thèse de doctorat
Specialité: génie civil

Rocio L. SEGURA

Sherbrooke (Québec) Canada

July 2019

JURY MEMBERS

Patrick PAULTRE

Supervisor

Jamie E. PADGETT

Co-supervisor

Jean PROULX

Examiner

Georgios BALOMENOS

Examiner

Jacky MAZARS

Examiner

RÉSUMÉ

Les conséquences de la rupture d'un barrage peuvent être considérables, en termes de pertes humaines, économiques et environnementales. De plus, la sécurité des barrages et des aménagements hydroélectriques est une préoccupation majeure au Québec étant donné que plus de la moitié de la population vit dans des zones potentiellement inondables. Il y a environ 933 grands barrages au Canada et 333 ou un peu plus sont situés au Québec. Parmi eux, beaucoup d'entre eux ont été construits il y a plus de 50 ans. Au cours de cette période, d'importants progrès ont été réalisés dans les méthodes d'évaluation des risques naturels. Bien que la défaillance totale d'un barrage en béton à la suite d'un tremblement de terre soit rare, les tremblements de terre sont une cause majeure de dommages à différents degrés de gravité. Par conséquent, le vieillissement et ses problèmes associés, combiné aux nouvelles méthodes d'estimation des charges sismiques ont entraîné la nécessité de revoir et d'améliorer les méthodes d'analyse sismique pour les barrages. Au cours des dernières décennies, les outils probabilistes sont devenus de plus en plus populaires pour l'évaluation sismique des barrages. Cependant, de telles méthodes nécessitent souvent un grand nombre d'analyses dynamiques non linéaires de modèles complexes par éléments finis. Par conséquent, le compromis entre l'exactitude du modèle numérique et la quantité de calcul rend ce type d'analyse non viable. Toutefois, l'évaluation sismique des barrages peut être améliorée en incluant les incertitudes liées aux paramètres sismiques et aux paramètres de modélisation et accélérée en réduisant l'importante quantité de temps de calcul avec l'utilisation de techniques d'apprentissage automatique pour développer des modèles de substitution ou des méta-modèles qui serviront prédire la réponse du barrage.

L'objectif principal de la recherche est de mettre au point une méthode d'évaluation de la sécurité sismique des structures de type barrage-poids grâce à une analyse de fragilité, effectuée avec la mise en œuvre de méta-modèles et en identifiant correctement le scénario sismique susceptible d'être présent sur le site du barrage. La méthodologie est appliquée à un barrage-poids situé dans l'est du Canada, dont la fragilité est évaluée par comparaison avec les études antérieures et directives de sécurité actuelles. On observe que la procédure plus précise présentée ici pour choisir les accélérogrammes produit des estimations moins conservatrices de la fragilité pour le barrage. Nous avons trouvé que la surface de réponse polynomiale de 4ème ordre est le méta-modèle le plus performant, et elle a été utilisée pour générer des fonctions de fragilité multivariées pour tenir compte de la variation des paramètres les plus critiques du modèle influençant la fragilité sismique du barrage. A partir de l'analyse de ces modèles, des recommandations pratiques de conception ont pu être formulées et il a été observé que le paramètre de modélisation affectant le plus l'analyse de la fragilité est la cohésion entre le béton et la roche.

Mots-clés : barrages-poids, aléas sismiques, analyse de fragilité, méta-modèle, courbes de fragilité, surfaces de fragilité, paramètres de modélisation

ABSTRACT

The consequences of dam failure can be substantial, in terms of casualties, economic and environmental damage. Moreover, the safety of dams and hydroelectric developments is a major concern in Quebec given that over half the population lives in potential flood zone. There are about 933 large dams in Canada and 333 or slightly more are located in Quebec. Among them, many have been built more than 50 years ago. During that time, important advances in the methodologies for evaluating the natural hazards have been made. Although total failure of a concrete dam following an earthquake is rare, earthquakes are a major cause of damage at different degrees of severity. Consequently, the combination ageing and its associated problems with new methods for estimating seismic loads, have resulted in the need to review and upgrade the methods of seismic analysis for dams. In recent decades, probabilistic-based tools, have become increasingly popular for the seismic assessment of dams. However, such methods often require a large number of nonlinear dynamic analysis of complex finite element models. As a result the trade-off between the accuracy of the numerical model and the computational burden render unviable this type of analysis. However, the seismic assessment of dams can be enhanced by including seismic and modeling parameters uncertainties and expedited by reducing the substantial computational time with the use of machine learning techniques to develop surrogate or meta-models to predict the response of the dam.

The proposed research addresses direct gaps in the seismic performance assessment of dams, while also shedding new light on the use of machine learning to support fragility modeling of complex systems like dams. Accordingly, the main objective of this research was to develop a method for assessing the seismic safety of gravity dam-type structures through a fragility analysis, conducted with the implementation of meta-models and by properly identifying the seismic scenario likely to be present at the dam site. The methodology is applied to a case study gravity dam located in eastern Canada, whose fragility is assessed through comparison with previous studies and current safety guidelines. It is observed that the more accurate procedure presented herein to select ground motions produces less conservative fragility estimates for the case study dam. Likewise, the 4th order polynomial response surface came up as the best performing meta-model, and it was used to generate multivariate fragility functions to account for the most critical model parameter variation influencing the dam seismic fragility. From the analysis of these models, practical design recommendations could be formulated and it was observed that the modelling parameter affecting the fragility analysis the most was the concrete–rock cohesion.

.

Keywords: gravity dams, seismic hazard, fragility analysis, meta-model, fragility curves, fragility surfaces, modeling parameters

To the other two musketeers: all for one and one for all, united we stand divided we fall.-

- Alexandre Dumas

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$$
\begin{aligned}
\mathbf{a} &= \text{constant coefficient vector for PRS meta-model;} \\
a_0 &= \text{bias constant term for RBF meta-model;} \\
a_1, \ldots, a_m &= \text{connection weights between hidden layers for RBF meta-model;} \\
b &= \text{NN bias parameter;} \\
b(\cdot) &= \text{basis functions of MARS meta-model;} \\
c &= \text{constant coefficients for MARS meta-model;} \\
\mathrm{D}_{595} &= \text{earthquake significant duration;} \\
E_{rock} &= \text{foundation modulus of elasticity;} \\
E_s &= \text{concrete modulus of elasticity;} \\
f(\cdot) &= \text{decision tree base models for RFR;} \\
f_0 &= \text{DRF fundamental frequency;} \\
F_{IM_1|IM_2} &= \text{theoretical CDF of seismic IM}_1 \text{ conditioned on IM}_2; \\
\mathrm{FI} &= \text{input Fourier transform;} \\
\mathrm{FMI} &= \text{modified Fourier transform;} \\
\mathrm{FO} &= \text{output Fourier transform;} \\
\mathrm{FT} &= \text{target Fourier transform;} \\
f_{\mathbf{x}}(\mathbf{x}) &= \text{probability density function of a random variable;} \\
g(\cdot) &= \text{surrogate model function;} \\
\mathrm{H} &= \text{block height;} \\
\mathrm{H2} &= \text{historical source model;} \\
\mathrm{HY} &= \text{hybrid source model;} \\
I_a &= \text{Arias intensity;} \\
\mathrm{IH} &= \text{input accelerogram;} \\
K &= \text{water bulk modulus;} \\
k &= \text{positive definite kernel function;} \\
\mathbf{m} &= \text{mean vector;} \\
\mathrm{M} &= \text{moment magnitude;} \\
\mathcal{M}(\cdot) &= \text{random margin;} \\
\mathrm{MIH} &= \text{modified input accelerogram;} \\
\mathrm{MOH} &= \text{deconvolved output accelerogram;} \\
\mathcal{N}(\cdot) &= \text{normal distribution;} \\
n_f &= \text{number of FE simulations;} \\
\mathcal{O}^n &= n^{th} \text{ order of the polynomial response surface;} \\
\mathrm{OH} &= \text{output accelerogram;} \\
p &= \text{binary variable;} \\
p(\mathbf{x}) &= \text{limit state function;} \\
P_e &= \text{probability of exceedance prescribed by the CDA;} \\
\mathrm{P}_f &= \text{probability of an event } (\in (0,1)); \\
\mathrm{P}_{f,unc} &= \text{unconditional probability of exceedance;} \\
P_r &= \text{probability of exceedance for a CDA return periods;} \\
\mathrm{PGA}_V &= \text{vertical peak ground acceleration;} \\
\mathrm{PGV}_{MCE} &= \text{peak ground velocity associated with the MCE;}
\end{aligned}
$$

$$
\begin{aligned}
p_1, \ldots, p_n &= \text{parameters characterizing a limit state;} \\
p_\theta, p_\beta &= \text{regression coefficients;} \\
Q &= \text{hybrid parameter;} \\
R &= \text{source-to-site distance;} \\
R2 &= \text{regional source model;} \\
R^2 &= \text{coefficient of determination;} \\
Sa_H(T^*) &= \text{horizontal spectral acceleration at target period } T^*; \\
Sa_H(T) &= \text{horizontal spectral acceleration corresponding to period } T; \\
Sa_V(T) &= \text{vertical spectral acceleration corresponding to period } T; \\
SaRotD50 &= Sa \text{ 50th percentile directionality ratio;} \\
Sv_H(T_1) &= \text{horizontal spectral velocity corresponding to period } T_1; \\
T &= \text{period of vibration;} \\
T^* &= \text{target period of vibration;} \\
T_1 &= \text{fundamental mode period;} \\
TH &= \text{target accelerogram;} \\
\mathbf{v} &= \text{normally distributed model error vector;} \\
v &= \text{model error;} \\
\hat{V} &= \text{estimated hazard curve;} \\
v_H &= \text{horizontal ground velocity;} \\
VS30 &= \text{shear wave velocity;} \\
\bar{V}(Sa(T_1)) &= \text{complementary CDF of } Sa(T_1); \\
V(Sa(T_1)) &= \text{CDF of } Sa(T_1); \\
v(Sa(T_1)) &= \text{PDF of } Sa(T_1); \\
v_V &= \text{vertical ground velocity;} \\
\mathcal{W}(\cdot) &= \text{Weibull distribution;} \\
w &= \text{NN covariates' weights;} \\
w(\cdot) &= \text{radial function for RBF meta-model;} \\
w_v &= \text{wave length;} \\
\mathbf{X} &= \text{covariate matrix;} \\
\mathbf{x} &= \text{row vector of the covariate matrix;} \\
y &= \text{structural response;} \\
z &= \text{NN hidden variable;} \\
\hat{y} &= \text{predicted structural response;} \\
\alpha_\alpha &= \text{PCE coefficients;} \\
\alpha_0, \alpha_1, \alpha_2 &= \text{hybrid parameter's linear correlation coefficients;} \\
\beta &= \text{standard deviation of the log-normal distribution;} \\
\beta_0, \ldots, \beta_5 &= \text{demand fragility surfaces' regression coefficients;} \\
\beta_\mathcal{W} &= \text{Weibull shape parameter;} \\
\delta_{\max} &= \text{maximum relative sliding at the base;} \\
\varepsilon &= \text{spectral shape;} \\
\zeta &= \text{standard deviation of the damping log-normal PDF;} \\
\Theta &= \text{vector of basis functions for PRS meta-model;} \\
\theta, \beta &= \text{parameters characterizing a cumulative density function;} \\
\vartheta &= \text{logistic regression coefficients;} \\
\hat{\theta}, \hat{\beta} &= \text{analytical formulation of the cumulative density function parameters;}
\end{aligned}
$$

$$\kappa(\cdot) = \text{fit-type regression function;}$$
$$\lambda = \text{mean of the damping log-normal PDF;}$$
$$\lambda_f = \text{demand fragility surfaces' median;}$$
$$\mu = \text{mean of the log-normal distribution;}$$
$$\mu_{\ln Sa} = \text{mean response spectrum;}$$
$$\mu_{\ln Sa(T_i)|\ln Sa(T^*)} = \text{conditional mean response spectrum;}$$
$$\mu_{\mathcal{W}} = \text{Weibull scale parameter;}$$
$$\nu_{rock} = \text{Poisson coefficient of the foundation;}$$
$$\xi = \text{damping ratio;}$$
$$\xi_f = \text{demand fragility surfaces' standard deviation;}$$
$$\rho = \text{linear correlation coefficient between IMs;}$$
$$\rho_{concrete} = \text{concrete density;}$$
$$\rho_{rock} = \text{density of the foundation;}$$
$$\rho_w = \text{water density at time } t;$$
$$\rho_{w,0} = \text{initial water density;}$$
$$\rho(T_i, T^*) = \text{correlation coefficient between a pair of } \varepsilon \text{ at two periods;}$$
$$\boldsymbol{\Sigma} = \text{covariance matrix;}$$
$$\sigma = \text{standard deviation of the response in the dataset;}$$
$$\sigma_c = \text{shear failure stress;}$$
$$\sigma_H = \text{horizontal stress time series;}$$
$$\sigma_{\ln Sa} = \text{standard deviation of the response spectrum;}$$
$$\sigma_{\ln Sa(T_i)|\ln Sa(T^*)} = \text{conditional standard deviation response spectrum;}$$
$$\sigma_n = \text{normal stress;}$$
$$\sigma_s = \text{shear stress;}$$
$$\sigma_t = \text{tensile failure stress;}$$
$$\sigma_V = \text{vertical stress time series;}$$
$$\sigma_v = \text{standard deviation of the normally distributed model error;}$$
$$\Phi(\cdot) = \text{cumulative density function;}$$
$$\Phi_{\mathcal{N}} = \text{standard normal cumulative density function;}$$
$$\Phi_{RBF_i}(\cdot) = \text{nonlinear mapping function for RBF meta-model;}$$
$$\phi = \text{concrete-rock angle of friction;}$$
$$\Psi_\alpha = \text{multivariate orthonormal polynomials;}$$
$$\omega_{eqk} = \text{earthquake angular frequency.}$$
$$\Omega_r = \text{frequency ratio.}$$

LIST OF ACRONYMS

ABFC	adaptive basis function construction
ASCE	American society of civil engineers
BIC	Bayesian information criterion
BTR	boosted regression trees
CDA	Canadian dam association
CDF	cumulative distribution function
CRC	concrete-rock cohesion
CRF	concrete-rock angle of friction
CRGP	Centre de recherche en génie parasismique et en dynamique des structures
CS	conditional spectrum
CV	cross-validation
DOE	design of experiments
DR	drain efficiency
DRF	dam-reservoir-foundation
DS	damage state
EDF	empirical distribution function
EDP	engineering demand parameter
ELF	equivalent lateral force
ENA	eastern North America
EPRI	Electrical Power Research Institute
ERF	earthquake rupture forecast
FEM	finite element model
FEMA	federal emergency management agency
FORM	first-order reliability method
FOSM	first-order second-moment
FSI	fluid-structure interaction
GCIM	generalized conditional intensity measure
GMM	ground motion models
GMTS	ground motion time series
GSC	geological survey of Canada
GP	Gaussian process
HCLPF	high confidence low probability of failure
ICOLD	International Commission on Large Dams
IDA	incremental dynamic analysis
IM	intensity measure
KS	Kolmogorov-Smirnov
LEM	limit equilibrium method
LHS	Latin hypercube simulation
LS	limit state
MARS	multivariate adaptive regressive splines
MCE	maximum considered earthquake
MCS	Monte Carlo simulation

MCMC   Markov chain Monte Carlo
   ML   machine learning
  MLE   maximum likelihood estimator
  MOL   maximum operating level
   MP   model parameters
  MSA   multiple stripe analysis
 NBCC   National building code of Canada
  NGA   Next Generation of Ground-Motion Attenuation Models
   NN   neuronal networks
 PBEE   performance based earthquake engineering
  PCE   polynomial chaos expansions
  PDF   probability density function
 PEER   Pacific Earthquake Engineering Research Center
  PGA   peak ground acceleration
  PGD   peak ground displacement
  PGV   peak ground velocity
  PRS   polynomial response surface
 PSHA   probabilistic seismic hazard analysis
  POE   probability of exceedance
  RBF   radial basis function
  RCC   roller-compacted concrete
  RFR   random forest for regression
 RMAE   relative maximum absolute error
 RMSE   root mean square error
  SCM   seismic coefficient method
   SD   standard deviation
  SHM   structural health monitoring
 SHMC   seismic hazard model for Canada
   SI   spectrum intensity
 SORM   second-order reliability method
 SPRA   seismic probabilistic risk assessment
  SSE   sum of squared errors
  SSI   soil-structure interaction
  SVM   support vector machines
  UHS   uniform hazard spectrum
USACE   United States Army Corps of Engineers
 USBR   United States Bureau of reclamation
  WNA   western North America

CHAPTER 1

INTRODUCTION

1.1 Context and problematic

Dams are one of the most audacious human constructions, considering the social, environmental and economic risks that a work of such magnitude entails. A dam is designed to impound and store water behind it safely throughout its lifetime. Thus it has to maintain its structural integrity in the face of the different hazards and loading conditions that arise during construction, normal operations, and extreme environmental events.

Several dams have experienced rupture (partially or totally) under the effect of exceptional floods, and to a lesser extent due to earthquakes. In the last years the documented damage caused by seismic events [4, 5, 6, 7] has revealed the vulnerability of these structures whose consequences are catastrophic and expensive. In addition, the Canadian Risk and Hazards Network indicates that a significant earthquake is probably Canada's greatest potential natural disaster [8]. The safety of dams and hydroelectric developments is a major concern in Quebec given that over half the population lives in potential flood zone. Thus, there are about 333 large dams in Quebec and around 933 in Canada [9]. Among them, many have been built more than 50 years ago [6]. During that time, important advances in the methodologies for evaluating the natural hazards have been made, causing the review and modification of the design guidelines, in some cases significantly. With the increasing knowledge of seismicity, a growing number of dams fail to meet the revised safety criteria that incorporates new seismic hazard information [6]. Consequently, the combination of ageing and its associated problems with new methods for estimating seismic loads and with the increasing demands of society to ensure a high level of safety has resulted in the need to review and upgrade the methods of seismic analysis for dams.

Furthermore, in Quebec, after the dam failure produced by the flood occurred in Saguenay region in July 1996, a Dam Safety Law was promulgated which is currently being applied. The aforesaid law confers a key role to the engineer and requires from the dam owners or operators to produce documented safety analyses of their works [10]. Consequently, questions have arisen regarding the capacity of existing dams, in their current condition, to withstand extreme environmental events beyond the original design basis during a future service period with an acceptable margin of safety.

Methods for analyzing the structural response of a dam-reservoir-foundation system rely on deterministic or probabilistic approaches. Deterministic methods are often considered too conservatives, or even unsafe in some cases, because of the use of extreme load cases with very low probabilities of occurrence and because of the way of considering different sources of uncertainty [11, 12, 13]. In the deterministic analysis, the way of handling the uncertainties involved in dam performance and safety assessment is by the traditional approach through the use of safety factors. These safety factors are applied separately at different stages, and inconsistently in the processes of analysis, design and assessment, leading to an unknown and unpredictable margin of safety [14]. Procedures that provide a more rational way of assessing the safety of concrete gravity dams are required to establish priorities. Moreover, a probabilistic approach is required to manage the various sources of uncertainty that may impact the dam performance and decisions related thereto [11]. Within this probabilistic framework, a fragility analysis is a promising alternative, particularly suited to study the seismic vulnerability of structures and to estimate the level of damage likely to be caused by seismic events.

Nevertheless, seismic fragility analysis and vulnerability assessment of key infrastructure elements, often requires a large number of non-linear dynamic analyses of complex finite element models (FEM). The substantial computational time may be reduced using machine learning techniques to develop surrogate or meta-models, which are an engineering method used when an outcome of interest cannot be easily directly measured, so a model of the outcome is used instead [15]. Such a challenge is particularly relevant to the case of large scale infrastructures, such as dams, subjected to seismic loads. Accordingly, the main goal of this research is to apply statistical learning techniques to develop a seismic probabilistic demand model or meta-model to predict the seismic response of the dam. The latter will be used to generate fragility functions and recommendations to expedite the safety assessment process. The proposed methodology is applied to a case study gravity dam located in north-eastern Canada. Additionally, the procedure presented herein has the added asset of properly depicting the seismic scenario likely to occur at a specific site enhancing the accuracy of the analysis.

1.2   Project description and objectives

The project is part of one of the research axes of the Centre de recherche en génie para-sismique et en dynamique des structures (CRGP) of the civil engineering department at Sherbrooke University for evaluating seismic vulnerability of structures in Canada and developing tools to do this. Several previous projects performed by CRGP allowed

studying and developing a methodology for using fragility curves, mainly in the area of bridges [16, 17, 18, 19] and more recently in the area of dams [20, 21]. Moreover, the CRGP has vast expertise in the dynamic behaviour and seismic analysis of dams and hydraulic structures.

This study aims to assess, in a probabilistic way, the seismic vulnerability of concrete gravity dams using meta-model-based single-variable and multi-variable fragility functions, including the most relevant model parameters (MP) and seismic intensity measures (IM) affecting the response of the structure.
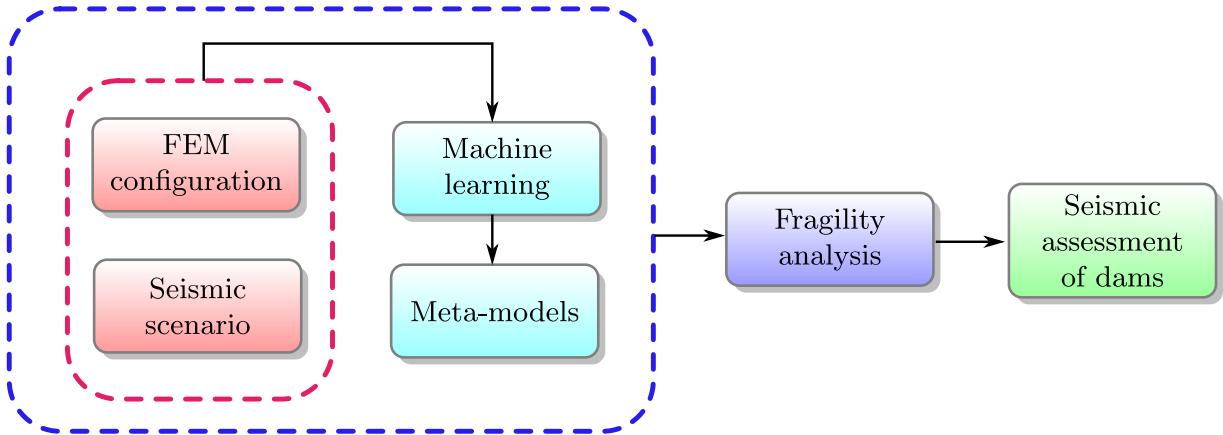
Figure 1.1   Research project main objective diagram

The present research project and the proposed methodology focus on a particular case study gravity dam located in north-eastern Canada. Particulary, this project has the following specific objectives:

1. Implementation of a non-linear finite element model, accounting for fluid-structure interaction (FSI) and soil structure interaction (SSI) to estimate the seismic response of the dam.

2. Characterize the seismic hazard at the dam site by performing a probabilistic seismic hazard analysis and by selecting a representative suite of ground motion time series (GMTS).

3. Identification of the most relevant MPs and seismic IMs affecting the seismic response of the dam through a statistically based screening study.

4. Explore the applicability of meta-models for the seismic assessment of gravity dams and systematically compare and determine the most suitable one.

5. Generate fragility functions from the meta-model output for the dam under study and gain insight on the influence of the model parameters affecting the dam perfor-

mance by explicitly quantify their effect with the generation of multivariate fragility functions.

6. Asses the expected seismic performance of the dam, providing an overview on the importance of the parameters influencing the dam performance to formulate practical design recommendations from the analysis, e.g. appropriate model parameter range to achieve target risk.

1.2.1   Case study - description

As it was aforementioned, the present study is focused on a particular concrete gravity dam in Quebec, Canada. It is the largest gravity dam in the province, with 19 unkeyed monoliths, a maximum crest height of 78 m, and a crest length of 300 m (Figure 1.1). The dam rests on a foundation consisting mainly of anorthosite gabbro and granitic gneiss [22], which corresponds to hard rock ($VS30 > 1500$ m/s). The dam was chosen for its simple and almost symmetric geometry and because of the availability of forced vibration test results used to calibrate the dynamic properties of the numerical model [22]. Moreover, previous studies concerning the seismic fragility of one of the central blocks of the dam [20, 21] are also available, enabling a direct comparison with the results of the method proposed herein.

(a)   (b)

Figure 1.2   Case study dam: (a) aerial view of the dam [1] and (b) cross section

1.3   Originality and contribution

Methods for analysing the structural behaviour of the dam-reservoir-foundation system rely on deterministic or probabilistic methods. Currently in Quebec the safety assessment

of hydraulic structures is mainly deterministic and limited to methods based on safety factors. At present, probabilistic assessment methods, such as fragility functions, are rarely used in the field of dams. This research project will allow the implementation of this tool and develop a methodology for its application in dam-type structures. To implement this type of design, an accurate estimate of the seismic demand of the structural systems is especially important. Such an estimate requires, in turn, a ground motion record selection technique that properly depicts the seismic scenario and adequately propagates the record-to-record variability and uncertainty related to the seismic hazard throughout the fragility analysis [23, 24]. To this end, a probabilistic seismic hazard analysis (PSHA) will be conducted, consistent with the National Building code of Canada (NBCC) 2015, to depict the seismic scenario likely to be present at the dam site. The primary advantage of PSHA over alternative representations of the earthquake threat is that PSHA integrates over all possible earthquake occurrences and ground motions to calculate a combined probability of exceedence that incorporates the relative frequencies of occurrence of different earthquakes and ground-motion characteristics. These results will be used, together with a refined ground motion selection method, the generalized conditional intensity measure (GCIM) approach, to select ground motion time series that match target seismic IMs that highly correlate with the structure's probable damage states. This method proposes, therefore, a more realistic target distribution and its application has produced encouraging results by recent studies in different contexts [25, 26, 27]. Currently, to the best of the author's knowledge, no previous studies have conducted seismic assessment of dams while properly considering the contribution of several IMs. Moreover, there are no studies that have used the GCIM method for this region. As such, this research aims to improve the fragility analysis for concrete gravity dams by using an enhanced procedure to select ground motion records that includes different intensity measures that are relevant to the probable damage states of such structures.

Within this framework, and to better simulate the real dynamic response of the structure, a non-linear FEM of the dam, which fully account for dam-reservoir-foundation interactions will be used. This approach will provide a more realistic simulation of the structural response and material behaviour, which are known to vary with the intensity of dynamic loading. In the same way, the effect of the modeling assumptions will be explored to decide whether this simplifications are considerable when it comes to the dynamic behaviour of the dam. The substantial computation time will be reduced using machine learning techniques to develop surrogate or meta-models, whose basic idea is to avoid the temptation to invest computational budget, but to develop fast mathematical approximations instead. This

research will also highlight the advantages and disadvantages regarding the traditional methodology.

Summarizing, most of prior studies on seismic assessment of concrete gravity dams via machine learning techniques are limited to the consideration of a few meta-models at a time in the context of a single study, simplified 2D finite element models and single-variate fragility functions or engineering demand parameter-seismic intensity measure fragility functions where the methodologies and/or parameterized equations are not provided [28, 29, 30]. Therefore, they do not explore the most suitable meta-models for fragility analysis of this type of structure, nor they explore the influence of the model parameters variation in the seismic fragility analysis. Moreover in any of these studies is discussed the proper definition of the seismic scenario likely to occur at a specific site in a probabilistic manner. To address the aforementioned gaps, this project aims to identify the most viable meta-model for the seismic fragility assessment of gravity dams and to provide an overview on the importance of the parameters influencing the dam performance to be able to formulate design recommendations from the analyses. The major contributions of this research, can be listed as: (i) systematically compare and determine the best performing regression meta-model to predict the base sliding of gravity dams, among 14 different regression techniques, for the first time; (ii) present a methodology to fit parameterized fragility surfaces, as a function of IMs and MPs, from the meta-models; (iii) consider the correlation between the seismic IMs from a probabilistic seismic hazard analysis to generate the samples to characterize the seismic scenario where the meta-model will be evaluated; (iv) gain insight on the influence of the model parameters affecting the dam performance, and explicitly quantify their effect with the generation of multivariate fragility functions and (v) formulate model parameter design recommendations from the analysis, e.g. appropriate parameters range to achieve target risk.

1.4   Structure of the document

In order to meet the objectives of the research project, the content of this thesis is divided into seven chapters. The second chapter presents a review of the literature that includes: (i) the seismic performance of concrete gravity dams, (ii) eastern Canada seismic hazard, (iii) fragility analysis and (iv) meta-models. The probabilistic seismic hazard analysis procedure, ground motion record selection methods, different types of fragility functions, meta-modeling techniques and the few examples available in the field of dams are discussed. The third chapter is dedicated to the modeling and characterization of a concrete gravity dam for fragility analysis, where a methodology for the proper consider-

ation of the uncertainties to assess the seismic vulnerability is proposed. It also presents a summary of the methodology adopted to generate the dataset with which the meta-models were trained. More specifically, this chapter includes the modeling of the case study dam, parameter uncertainties consideration, ground motion selection procedure, design of experiments and meta-model training, optimal regression technique selection and parametric fragility surfaces development. The fourth chapter presents the results of the PSHA to characterize the seismic scenario at the dam site and the development of up-to-date fragility curves for the case study dam using a record selection method based on the generalized conditional intensity measure approach. These fragility functions are then combined with the aforementioned regional hazard data to evaluate the annual risk, which is measured in terms of the unconditional probability of limit state exceedance. The fifth chapter provides insight on viable meta-models for the seismic assessment of gravity dams for use in fragility analysis. A methodology to generate multivariate fragility functions from the meta-models is presented, which offers efficiency while accounting for the most critical model parameter variation influencing the dam seismic fragility. From the analysis of these functions, practical design recommendations were formulated. As it can be seen from Figure 1.3, the modality of the thesis is by articles and two scientific papers were produced in the context of this project. It should be noted that chapters 4 (Part 1) and 5 (Part 2) constitute the articles presented in scientific journals, where the samples generated in Part 1 with the finite element model and the seismic scenario characterization were used in Part 2 to train the meta-model that will emulate the finite element model behaviour. Finally, the results of Part 1 and Part 2 were used to perform the seismic assessment of the dam.
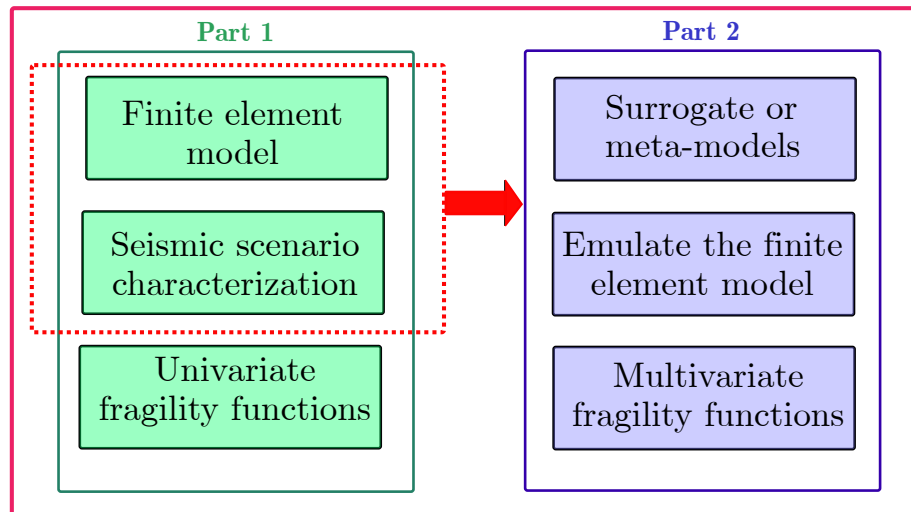
Figure 1.3 Structure of the document - Thesis modality

Chapter six includes complementary verifications and comparisons regarding the seismic fragility of the case study dam. These results, presented in conference papers, are included in this document for completeness. As such, this chapter presents the comparison of the fragility functions developed with the 2010 and 2015 hazard models and with several record selection methods and the estimation of the expected seismic performance under extreme seismic events to assess the impact of these factors in the fragility of the system. Also, the best performing meta-model is used to develop fragility curves and to identify the most influencing model parameter in the seismic response by developing fragility regions. Finally, the last chapter presents the conclusions of the research project and gives guidelines for future projects.

CHAPTER 2

STATE OF THE ART

A concrete gravity dam is a massive structure with an essentially triangular profile that consists of rigid monoliths built side by side. Its longitudinal axis runs straight along the entire dam. The structural stability of a concrete gravity dam is derived mostly from its weight. The primary load affecting concrete gravity dams arises from the hydrostatic pressure.

Certainly, concrete gravity dams are designed to remain operational under normal conditions, to sustain minimal damage under infrequent operative conditions and to prevent total loss of reservoir under extreme events. To avoid the uncontrollable release of water during a strong seismic event, the dam must be able to withstand the ground shaking of even an extreme earthquake given the human fatalities and the economic losses that can result from. It arises then, the importance of the behaviour of the dam under seismic loads. From a deterministic point of view, large storage dams are generally considered safe if they can survive an event with a return period of 10.000 years, i.e., having a 1 % chance of being exceeded in 100 years. It is very difficult to predict what could happen during such a rare event as very few earthquakes of this size have actually affected dams.

Current procedures for concrete gravity dams design are based on static, deterministic analyses. These involve rigid body analysis and static equilibrium for monolith stability and foundation stresses. Where required, seismic loads are approximated by pseudo-static forces obtained using the seismic coefficient method [31]. The governing criteria in concrete gravity dam design are that the dam must be safe from overturning, sliding and bearing failure in the dam or foundation. Under the criteria currently in use, dams are designed to withstand or resist hazards that conceivably could occur in their lifetimes. However, it is not possible to provide "absolute" safety against all hazards. Extreme events, like earthquakes, that exceed the dam's capacity can occur, although they may have very low probabilities of occurrence. Moreover, it is not possible to determine what would happen (damage without full failure) for lower hazards.

Methods for seismic analysis of concrete dams have been improved extensively in the last few decades and growth in the processing power of computers has expedited this improvement. Advanced numerical models of dams have become more feasible and manageable

and constitute the basis of more adequate procedures of designing and assessing. Procedures that provide a more rational way of assessing safety of concrete gravity dams are required to establish priorities. Such procedures should be based on modern principles of structural/geotechnical engineering. Moreover, a probabilistic framework is required to manage the various sources of uncertainty that may impact dam performance and decisions related thereto [32]. Seismic actions have essentially stochastic characteristics, therefore they should be considered as probabilistic loads. Moreover, material properties and structural dimensions have random variations in the spatial domain. Thus, a probabilistic dynamic analysis of the coupled system dam-reservoir-foundation is necessary [33, 34].

2.1 Seismic performance of gravity dams

Historically, few dams have been significantly damaged by earthquakes. On a worldwide basis, only about a dozen dams are known to have failed completely as the result of an earthquake [35, 7, 4]. These dams were primarily tailings or hydraulic fill dams, or relatively old, small, earth-fill embankments of inadequate design. Earthquakes represent multiple hazards with the following features in the case of a storage dam:

– Ground shaking causes vibrations and structural distortions in dams, appurtenant structures and equipment, and their foundations.
– Fault movements in the dam foundation or discontinuities in dam foundation near major faults can be activated, causing structural distortions.
– Fault displacement in the reservoir bottom may cause water waves in the reservoir or loss of free board.
– Rock falls and landslides may cause damage to gates, spillway piers (cracks), retaining walls (overturning), surface powerhouses (cracking and puncturing and distortions), electro-mechanical equipment, pen stocks, masts of transmission lines, etc,.
– Mass movements into the reservoir may cause impulse waves in the reservoir.
– Mass movements blocking rivers and forming landslide dams and lakes whose failure may lead to over topping of run-of-river power plants or the inundation of powerhouses with equipment, and damage downstream.
– Ground movements and settlements due to liquefaction, densification of soil and rock-fill, causing distortions in dams.
– Abutment movements causing sliding and distortions in the dam.

Dams are designed to withstand construction, normal and infrequent operating conditions, and extreme environmental load conditions. The two primary environmental load

conditions affecting concrete gravity dams arise from the differential hydrostatic level that
occurs during hydrologic events and the dynamic loads due to seismological events. Several stages (limit states) of a dam behaviour occur under progressively increasing levels of
excitation. At low levels, the dam-foundation system remains essentially elastic, displacements are small, drains are fully effective, and full control of the reservoir is maintained.
In the elastic range, there are no permanent deformations, and the traditional 2D equilibrium analysis of the dam as a rigid body is sufficient. At the onset of non-linear behaviour,
material cracking occurs, deformations may become permanent, drainage characteristics of
the dam and the operation of gates begin to be affected, and 3D structural actions within
the dam are initiated. At this stage, 2D rigid body analysis of a dam monolith may no
longer provide a good model of structural performance, and a finite element model of the
complete dam may become necessary. Excessive deformations may cause objectionable
cracking and functional disruption. Finally, at ultimate conditions prior to impending
failure, the drains are ineffective due to large deformations, and structural behaviour becomes unstable and unpredictable due to sliding, flotations, or loss of foundation material
bearing capacity. In extreme cases, loss of control of the reservoir may occur [32, 36].

2.1.1 Loads, failure modes and limit states

As it was mentioned, the structural stability of a concrete gravity dam is derived mostly
from its weight. The loads acting on the structure can be summarized in Figure 2.1.
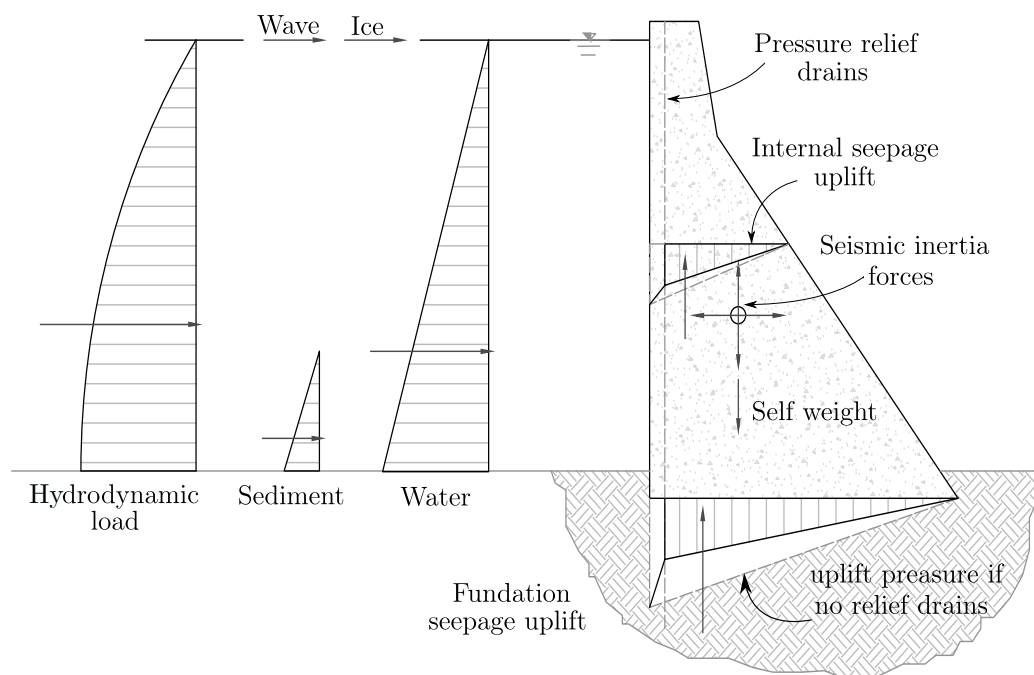
Figure 2.1 Commonly used loads in design practice

Two types of loads solicit a dam: static and dynamic loads. The main static loads are: the weight of the dam, the hydrostatic pressure of the reservoir, uplift-pressure at the foundation, ice thrust and the stresses generated by volumetric changes and thermal gradients. Attention must be focussed on the operational limit states: gravity only (including staged construction), normal reservoir only, gravity plus normal reservoir, gravity plus normal reservoir plus winter/summer temperature; which are of primary importance for dam structures [37]. For their part, the dynamic loads are caused by the action of an earthquake which induces horizontal and vertical inertia forces and reservoir hydrodynamic thrust. Loads due to seismic excitation are one of the most important actions that must be considered in concrete dam design, even in regions of low seismic activity [33, 34]. The relevant limit states of a dam, from the most expected to the least expected [14, 38, 37, 39], can be listed as:

1. Material failure - concrete crushing failure.

2. Excessive stress in the foundation.

3. Excessive deformation of the dam body inducing service limitation for equipments and installations.

4. Seismic hammering between the blocks.

5. Considerable opening of cracking and joints.

6. Sliding at the dam-foundation interface or at the construction joints.

7. Uplifting and overturning.

8. Deflection of the top of dam relative to heel.

When subjected to strong ground motions, gravity dams may be damaged in different ways. Nevertheless, in recent years, typical damage modes in dams after a seismic event have been identified. Based on that seismic damage levels can be established. This is a necessary step before seismic vulnerability analysis can be performed [40]. In their study, Zhong et al. [40] presented the expected typical gravity dams failure modes based on non-linear seismic simulation results. Strictly speaking, classification of seismic damage levels should be established based on documents survey of dams actually damaged by earthquakes in history, or on a large number of simulation of dams subjected to strong earthquakes.

2.1.2   Seismic analysis methods and structural modeling

To assess the safety of a dam, it is necessary to proceed with the analysis of the latter. Two approaches are generally used to model a dam: seismic coefficient method (SCM)

and the finite element method. The SCM method assumes that the dam is infinitely rigid, earthquake forces are treated simply as static forces and are combined with the hydrostatic pressures, uplift, backfill soil pressures, and gravity loads. The analysis is primarily concerned with the rotational and sliding stability of the structure treated as a rigid body. The inertia forces acting on the structure are computed as the product of the structural mass, added-mass of water, and the effects of dynamic soil pressures, times a seismic coefficient. The magnitude of the seismic coefficient is often taken as a fraction of the peak ground acceleration expressed as a decimal fraction of the acceleration of gravity. This method has been widely used in practice because of its simplicity and according to the United States Corps of Engineers (USACE) [41] this method may still be used in the preliminary design and stability analyses. Nevertheless, it is not applicable in the case of a three-dimensional behaviour and it neglects several phenomena which may be important such as the dynamic characteristics of the dam-reservoir-foundation system and the characteristics of the ground motion.

The finite element method is based on the finite element theory to calculate strains and stresses in the dam. This method is much more accurate and provides a better assessment of the behaviour of a dam [42]. The global response of a dam is obtained by first calculating the effect of static loads to which we add the effect of seismic loads. Analysis of seismic loads can be carried out with three methods: pseudo-static method, pseudo-dynamic method and temporal dynamic method.

Pseudo-static method

The equivalent lateral force method (ELF), which is commonly used for the seismic design of buildings, has also been developed for preliminary seismic analysis of gravity dams assuming that the structure response is predominantly in the first mode [43]. In a similar way, the ELF method presents the same limitations as the SCM.

Pseudo-dynamic method

Also called the modal response spectrum analysis procedure is the most basic and truly dynamic method of analysis. In this method, the peak responses of linear elastic structures to earthquake ground motions characterized by response spectra are determined [44]. The number of modes required varies for each analysis, however all modes with significant contribution to the total response of the structure should be included. Usually the number of modes is adequate if the total mass participation of the modes used in the analysis is at least within 90% of the total mass of the structure. Within the pseudo-dynamic methods, the time history modal analysis procedure is similar to the one described for the modal

response spectrum, except that earthquake demands are in the form of acceleration time histories, rather than response spectra. The results are in terms of displacement and stress (or force) time histories. Peak values of various response quantities are extracted from the response histories. Time history modal analysis provides valuable time dependent information that is not available in the modal response spectrum analysis procedure. Especially important is the number of excursions beyond displacement levels where the structure might experience strength degradation (strain softening). This type of analysis is preferred over the ELF method which is limited to a single mode but it is still conditioned to linear elastic responses.

Temporal dynamic method

This type of analysis involves the direct integration of the equations of motion, and therefore is the most powerful method available for evaluating the response of structures to earthquake ground motions [41, 45]. It is a step-by-step numerical integration procedure, which determines stresses (or forces) and displacements for a series of short time increments from the initiation of loading to any desired time. The time increments are generally taken of equal length for computational convenience. The condition of dynamic equilibrium is established at the beginning and end of each time increment. The motion of the system during each time increment is evaluated on the basis of an assumed response mechanism. The advantage of this method is that it can be used for both linear and nonlinear analyses. In the case of nonlinear analyses, structure properties (including nonlinear behaviour) can be modified during each time increment to capture response behaviour appropriate to that deformed state. The application of non-linear analysis to concrete hydraulic structures is limited to cases for which experimental or observational evidence of non-linear behaviour is available and that validity of the numerical models have been demonstrated. These include certain nonlinear behaviour such as joint opening mechanisms, tensile cracking, sliding and rotational stability of blocks isolated by opened joints and cracked sections, and local yielding and cracking.

2.1.3   Modeling considerations

Structural models for dynamic analyses are developed much in the same manner as for static analyses. However, distribution of mass, stiffness and dynamic interaction between the structure and water and between the structure and foundation as well as with the backfill soil should be established accurately. The response of a structure under severe ground shaking may approach or exceed the yield/cracking state. This means that in a linear-elastic dynamic analysis the use of effective stiffness is more appropriate than the

initial elastic stiffness used in the static analysis, and that the damping should be selected consistently with the expected level of deformation and the extent of non-linear behaviour. Furthermore, concrete deterioration and cracking can reduce structural stiffness of an existing structure; thus these effects should be considered in estimation of a representative effective stiffness.

The dynamic interaction with the foundation introduces inertia and flexibility at the base of the model and could provide additional damping mechanisms through material and radiation damping [41]. One of the problems associated with this, is that the seismic signal must be modified within the foundation. Consequently the earthquake to be applied at the base of the foundation is not one from a record in the soil surface, but a modification thereof (deconvolved records). A hydraulic structure also interacts with the impounded, surrounding, or retained water through hydrodynamic pressures at the structure-water interface. This interaction is coupled in the sense that motions of the structure generate hydrodynamic pressures that affect deformations (or motions) of the structure which, in turn, influence the hydrodynamic pressures. In the same way the compressibility of the water and the radiation damping in the reservoir given the finite length in the reservoir modeling are factors that affect the dynamic response of the dam. Since foundation properties, structural properties, and boundary conditions can vary, it is advisable to systematically vary parameters that have a significant effect on structure response until the final results cover a reasonable range of possible responses the structure could experience during the design earthquake [41].

3D vs. 2D modeling

At present, two dimensional analyses of the dam geometries is still the most common approach for the design or evaluation of the gravity dams. Most of the three dimensional analysis methods in the past were developed in reference to arch dams [46]. Practising engineers almost always prefer the two dimensional analysis tools because of the practicality and wide experience in using these tools and evaluating the results from such models. The presence of the construction joints, which is a requirement for conventional concrete dam bodies, justifies the use of two dimensional analyses of concrete gravity dams to some extent, based on the assumption that the monoliths behave independently during seismic events. It may happen that the accuracy of 2D models is not enough and one must resort to the 3D modeling. 3D modeling involves the principle of non-linearity especially when the foundation is modeled. Yilmazturk et al. [47], pointed out the need of three dimensional analyses for the seismic design of dams in narrow canyons in addition to the fact

that 3D model non-linearities are particularly more apparent when the soil shear resistance curve as a function of the normal stress is considered.

Linear vs. non-linear analysis

As previously mentioned, when subjected to strong ground motions, gravity dams may be damaged in different modes. This implies that depending on the response of a structure under severe ground shaking (approach or exceed the yield/cracking state), structural and material properties must be modified accordingly. Concrete deterioration and cracking can reduce the structural stiffness of an existing structure [48]; thus these effects should be considered in the estimation of a representative effective stiffness. Inelastic material models provide more realistic results, as illustrated in Figure 2.2. In the same way, the damping characteristics should be selected consistent with the expected level of deformation and the extent of non-linear behaviour.
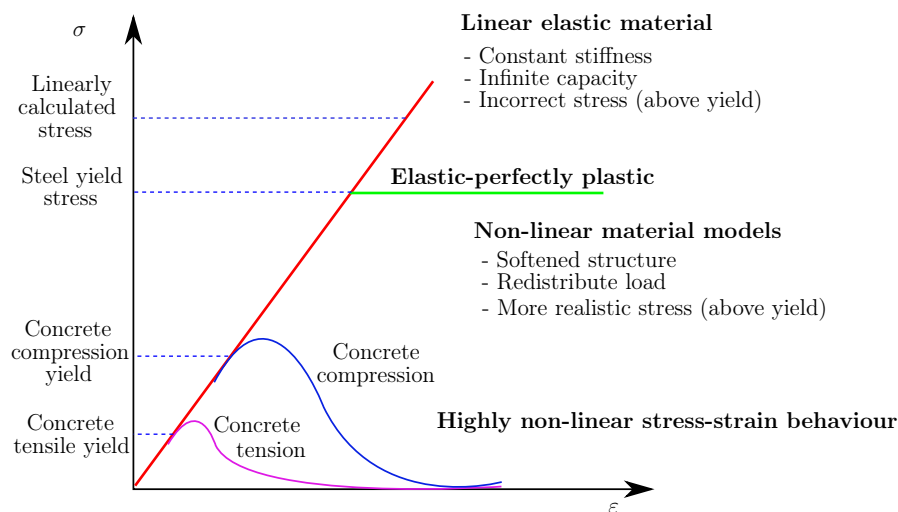
Figure 2.2   Linear vs. non-linear material behaviour

Linear elastic finite element analyses of gravity dams often show very high tensile stress at the upstream heel, especially for two-dimensional models. The reentrant corner goes into tension when the dam deforms downstream. With a fully connected dam-foundation interface, combined with linear elastic materials, the results typically show tension that would exceed that of the concrete by a large margin. Many schemes have been developed to analyse the fictitious high stresses and arrive at a reasonable dam safety decision, but each relies heavily on engineering judgement. Schultz et al. [49] showed, in a reevaluation of a case study dam using non-linear analysis, a more realistic model that better aligned with the past performance. Instead of unrealistic tension stress between the dam and foundation, the resulting stress distribution showed that the highest tensile stress does

not occur at the reentrant corner, but rather about two-thirds height on the upstream face.

Depending on the type of analysis a suitable numerical method should be selected to calculate the response of the dam. Non-linear finite element analysis is a broad term that encompasses a wide variety of numerical methods designed to capture important structural responses that can not be achieved using a continuous mesh or with elements formulated on a linear stress-strain assumption. The terms "implicit" and "explicit" analyses are commonly used to distinguish between two general approaches for finite element analysis with respect to time history analysis. If the calculation at the current time step t_{n+1} depends only on the results at the preceding time steps (say at time t_n), the method is said to be explicit. If the calculation at time t_{n+1} depends on the results both at the preceding time steps and at the current time t_{n+1}, the method is said to be implicit [44].

Fluid-structure-interaction (FSI) modeling

Earthquake analysis of dams is further simplified if compressibility of water is ignored because then hydrodynamic effects can be modeled with an added mass frequency independent model moving with the dam [50]. When considering the compressibility of water, in addition to the static water pressure, the dam undergoes dynamic forces from the reservoir when the system is subjected to earthquake ground motion. The magnitude of this additional hydrodynamic force is quite significant and may lead to crack initiation and propagation in the dam even under a moderately strong seismic event [51].

Three approaches can be used to represent the hydrodynamic effect of the fluid, the Westergaard or added mass, the Eulerian and the Lagrangian method [52, 41, 51]. For the added mass approach the dynamic effect of the reservoir water is modeled as masses applied at the upstream dam face. The added hydrodynamic mass influences the structure response by lengthening the period of vibration, which in turn changes the response spectrum ordinate and thus the earthquake forces. The added hydrodynamic damping arises from the radiation of pressure waves and, for dams, also from the refraction or absorption of pressure waves at the reservoir bottom. The added damping reduces the amplitude of the structure response especially at the higher modes.

The Eulerian formulation also known as potential or pressure based formulation, allows one degree of freedom per element (pressure, potential displacement or potential velocity) and also offers a significant reduction in the number of the model degrees of freedom. However, interface elements are needed to ensure the compatibility and equilibrium conditions in the fluid-structure interface. These elements allow to connect fluid elements with adjacent

solid elements. Each node of these elements has one potential degree of freedom as well as the displacement degrees of freedom. In the Eulerian approach the displacements are the variables in the structure and the pressures are the variables in the fluid, thus a special purpose computer program is required to obtain the solution of the coupled systems [53, 51].

The Lagrangian formulation is an extension of the classical finite element displacement formulation developed for solids; that is, the degrees of freedom for the fluid are the same as for the solid: the nodal displacements. The shape functions of the solids and fluids elements are the same. However, fluid elements are characterized by a volume elasticity modulus equal to the fluid compressibility modulus and zero shear resistance to simulate an inviscid flow. Compatibility in the fluid-structure interface is done by simply imposing the same normal displacements to the fluid and solid overlapping nodes on the interface. In other words, in the Lagrangian approach, the reference configuration of the continuum, solid or fluid, is the undeformed state, the behaviour of both the fluid and structure is expressed in terms of the displacements [54]. Therefore, compatibility and equilibrium are automatically satisfied at the nodes along the interfaces between the fluid and structure. This makes the Lagrangian displacement-based fluid finite elements very attractive; they can be readily incorporated into a general purpose computer program for structural analysis, since special interface equations are not required.

Earthquake analysis of dams is greatly simplified if compressibility of water is ignored because then hydrodynamic effects can be modeled by Westergaard's concept of a frequency-independent added mass moving with the dam. The significant discrepancies observed regarding more advanced method including water-structure interaction render the simplistic methods inadequate for seismic safety evaluation of existing dams. To evaluate the threshold where the water compressibility can be ignored, it is necessary to consider the frequency response function of a dam. The latter, when plotted in normalized form, depends on parameters characterizing the dam-water system, such as the ratio of water depth to dam height, the hydrodynamic wave reflection coefficient at the reservoir bottom and the frequency ratio, Ω_r, which is the ratio between the fundamental natural vibration frequency of the impounded water and the fundamental natural vibration frequency of the dam alone. For a fixed cross-sectional shape, the frequency ratio Ω_r , is proportional to $1/\sqrt{E_s}$, where E_s is the concrete modulus of elasticity of the dam. Thus Ω_r decreases with increasing E_s, or dam stiffness, and vice versa. However, the frequency response functions are independent of E_s and Ω_r if the reservoir is empty or if water is assumed to be incompressible. A study of frequency response functions for dams conducted by Chopra [50]

showed that with increasing Ω_r , or decreasing E_s, the effects of water compressibility on response become smaller and the response curve approaches the incompressible case. For systems with $\Omega_r = 2$, the effects of water compressibility are insignificant in the response to horizontal ground motion but are still noticeable in the response to vertical ground motion. As least in the response to horizontal ground motion, the effects of water compressibility become insignificant for systems with $\Omega_r > 2$. Interpreting this physically implies that in a very flexible dam the effects of water compressibility become much smaller in the earthquake response of dams with very low values of elastic modulus. Water compressibility is expected to be significant in the response of most gravity dams because the elastic modulus of mass concrete used in dams produces a Ω_r much smaller than 2 [50]. Ignoring water compressibility, which permits modeling of hydrodynamic effects by an added mass of water moving with the dam, would lead to unreliable decisions in seismic safety evaluation of proposed design of new dams and of existing dams [50].

Soil-structure-interaction (SSI) modeling

A SSI model refers to a case where interaction between the structure and its foundation requires special consideration in terms of the ground motion at the base of the structure and the flexible support provided by the soil foundation. Such interaction generally introduces frequency-dependent interacting forces at the structure-foundation interface requiring more elaborate analysis. In practice however, simplified models that include only the flexibility of the foundation and not its inertia and damping are more common [55]. At soil sites the bed rock motion is affected by the local soil conditions as it travels to the ground surface, and the presence of the structure produces a further change to this motion due to kinematic constraints. Additionally, the foundation interacts with the structure by elongating periods of vibration, by providing additional damping mechanisms (through material and radiation damping) and with the presence of rock mass inertia. The material and the radiation damping in the foundation region have the effect of reducing the structural response [56, 57, 58]. Earthquake analysis of dams is greatly simplified if foundation rock is assumed to have no mass, because then only flexibility of the foundation needs to be considered, which can be computed by analysing a bounded-size model of the foundation as part of a standard finite element analysis. By assuming foundation rock to be massless, seismic demands are overestimated considerably in some cases by factors of 2 to 3; this may lead to overly conservative and hence unnecessarily expensive designs for new dams, and to the erroneous conclusion that an existing dam is unsafe, requiring unnecessary retrofit that can be very expensive [50]. Soil-structure-interaction methods and their exhaustive mathematical formulation can be found in the literature [55, 59].

2.2   Seismic risk assessment

Seismic risk describes the potential damages or losses that a region is prone to experience following a seismic event. Methods of analysis and evaluation are nowadays made deterministically. Therefore, they are often considered too conservative because of the use of extreme load cases with very low probabilities of occurrence [11, 14, 60], and because they neglect several sources of uncertainty which affect both the structural capacity and demand. Among them are noteworthy the seismic forces, the material properties and the system configuration. Moreover, deterministic methods give very little useful information on vulnerability and reliability of a structure. To address these issues, probabilistic risk assessment and safety methods are becoming more and more popular [61, 14].

2.2.1   Seismic probabilistic risk assessment - SPRA

Probabilistic risk assessment was originally conceived for using in nuclear industry given the extent of damage that might occur in extreme events, like earthquakes [62, 63]. Over time this method was adapted for different types of structures, such as bridges and dams, and natural hazards like seismic events. The objective of the seismic probabilistic risk assessment (SPRA) is to estimate the probability of occurrence of different seismic events that may affect the structure and to assess the structure response to such earthquakes. The key elements of an SPRA can be reduced to three:

1. Seismic hazard analysis: the selected motion parameter represents the effect of the seismic events at the structure site. The hazard analysis element must consider all the sources of such motion and group them accordingly when estimating occurrence frequencies for each of several magnitudes of ground motion.

2. Seismic fragility evaluation: estimates the conditional probabilities of reaching a specific damage limit state of the structure as a function of the seismic intensity.

3. System analysis: this step is to define the sequence of events that can lead to system damage and estimate damages and losses resulting therefrom.

The framework for modern risk assessment is provided by the theorem of total probability [64]:

$$P_f(\text{Loss} > c) = \sum_s \sum_{\text{LS}} \sum_d P_f(\text{Loss} > c | \text{DS} = d) P_f(\text{DS} = d | \text{LS}) P_f(\text{LS} | \text{IM} = s) P_f(\text{IM} = s)$$

(2.1)

In which IM is the ground motion intensity measure, $P_f(IM = s)$ the probability of reaching a structural limit state LS, given the occurrence of IM$= s$, $P_f(DS = d|LS)$ the probability of damage state DS, given limit state LS, and $P_f(Loss > c|DS = d)$ the probability that the loss exceeds c, given that DS$= d$. The breakdown in this equation clearly identifies the fundamental contributors to the risk assessment: seismology, structural engineering, and structure economics and losses. The term $P_f(DS = d|LS)$ bridges the gap between structural engineering analysis, which assesses limit states in terms of forces and deformations, and loss estimation, which relates damage states (e.g. minor, moderate, severe) to economic losses, expressed as a percentage of replacement cost.

2.2.2   Seismic hazard - Quebec

Although earthquakes occur in all regions of Canada, certain areas have a higher probability of experiencing damaging ground motions caused by earthquakes. Figure 2.3, provides an idea of the likelihood of experiencing strong earthquake shaking at various locations across the country. The potential damage of an earthquake is determined by how the ground moves and how the structures within the affected region are constructed. Expected ground motion can be calculated on the basis of probability, and the expected ground motions are referred to as seismic hazard.
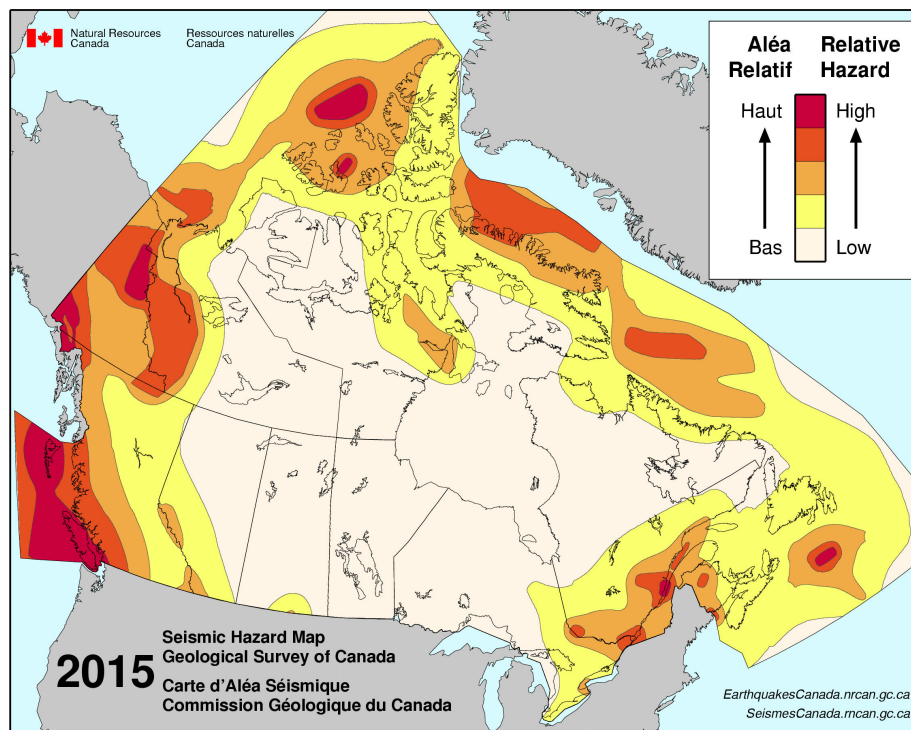
Figure 2.3   Simplified seismic hazard map [2]

Even though eastern Canada is not located in a critical zone, following several major earthquakes in the south-east of Canada between 1663 and 2010 (Figure 2.4), eastern Canada is considered a moderate seismic zone. This seismic activity in the stable interior of the North American Plate, where eastern Canada is located, is believed to be related to the regional stress fields, with the earthquakes concentrated in regions of crustal weakness. Years of instrumental recordings have identified certain clusters of earthquake activity in Eastern Canada where earthquakes occur at depths varying from surface to 30 km. West Quebec, Charlevoix-Kamouraska, Lower St. Lawrence and a small portion of the Northern Appalachians are the clusters located in the Province of Quebec. Each year, approximately 450 earthquakes occur in eastern Canada [8]. Of these ground motion time series four will exceed magnitude 4, thirty will exceed magnitude 3, and about twenty five events will be reported felt. A decade will, on average, include three events greater than magnitude 5. A magnitude 3 event is sufficiently strong to be felt in the immediate area, and a magnitude 5 event is considered the threshold of damage.
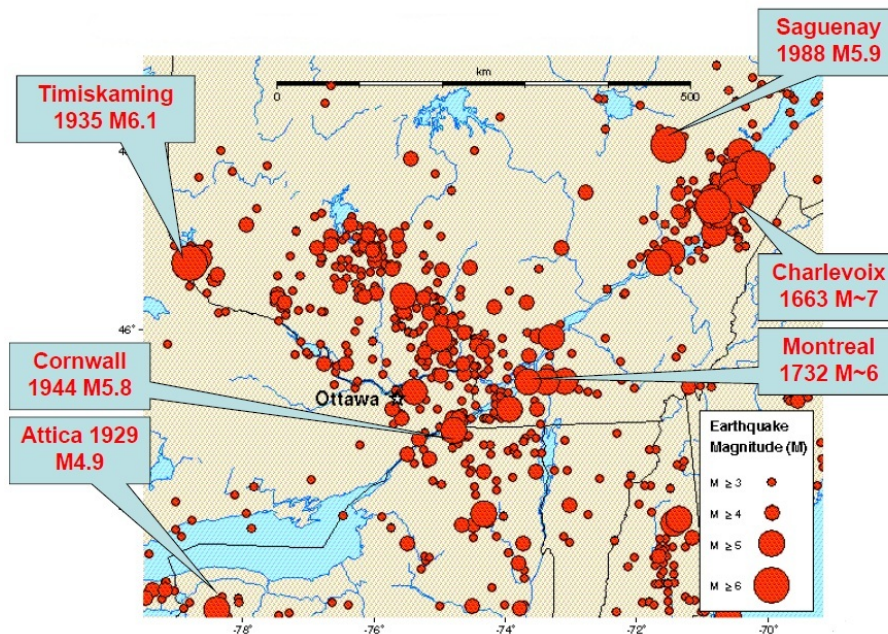
Figure 2.4   Southeastern Canada earthquakes since 1700 [3]

2015 National hazard model update - 5th generation

The requirements for seismic hazard modeling and mapping have changed over the years as scientists understandings of earthquakes and their effects on buildings have evolved and improved. As the knowledge of, and sophistication in, probabilistic seismic hazard modeling techniques have advanced, Canada's national mapping efforts have evolved from

qualitative assessment in 1953, to fully probabilistic for the hazard model accepted for the seismic provisions of the 2015 NBCC [65].

The 2015 national hazard model update (the 5th generation) yields many important advances on its predecessors, including: reconfigured seismic sources and special consideration of large rare eastern earthquakes; the use of a suite of representative backbone ground-motion models, and; explicit definition of crustal fault sources in the Yukon Territory and offshore western margin faults (north of Cascadia) based on GPS and paleoseismic slip rates [66]. The model takes advantage of contemporary scientific knowledge and replaces the 4th generation "robust" combination of alternative source models used for the 2010 seismic hazard model for Canada (SHMC) with a fully probabilistic model.

Regarding the seismic sources zones, for the 4th generation two models were used, distinguished primarily as historical cluster (H2) and regional seismotectonic (R2) models. For the 5th generation model, this framework is preserved in northeastern Canada. In southeastern Canada, an additional type of source - hybrid between H2 and R2 (HY) - is used together with the updated H2 and R2 models. The different physical properties of the crust in eastern and western Canada and the different nature of the earthquake sources in southwestern Canada require the use of separate ground motion models (GMMs), as detailed by Atkinson and Adams [67]. Unlike the 4th generation SHMC which used a single published relation (with rather arbitrary uncertainty bounds) for each region, the 5th Generation model uses representative suites of GMMs. A suite of crustal relations based on the ground-motion values from five appropriate eastern GMMs was used for eastern Canada.

2.2.3 Ground motion record selection method

To assess the seismic vulnerability of structures, an accurate estimate of the seismic demand of the structural systems is especially important. Such an estimate requires, in turn, a ground motion record selection technique that properly depicts the seismic scenario and adequately propagates the record-to-record variability and uncertainty related to the seismic hazard throughout the fragility analysis [23, 24]. Generally, the primary considerations in selecting ground motion recordings make use of appropriate seismological properties. However, these seismological parameters alone might not properly capture the seismic behaviour of the structure, failing to predict of structural demands [68]. With this limitation in mind, a consensus is emerging that it is more productive to consider time series properties rather than seismological parameters when selecting ground motions [69].

One commonly used approach for record selection is to select recorded or simulated ground motions whose response spectra match a target mean response spectrum. Irrespective of the procedure used to obtain a target response spectrum, there are several methods for selecting input ground motions that match a desired spectrum. Frequently, ground motions are selected (sometimes after scaling) to individually deviate the least from the target response spectrum. The deviation can be measured using the sum of squared differences between the response spectrum of the record and the target response spectrum. An alternate approach is to select a ground-motion set, rather than one record at a time, by minimizing the mean spectrum of the selected records from the target response spectrum. In some situations, matching only a target mean response spectrum is not sufficient since the approach ignores the inherent variance that may exist in the response spectrum. When matching a target response spectrum mean and variance, it does not suffice to treat ground motions individually, but rather requires comparisons of the mean and variance of sets of ground motions to the target values [70]. There is generally an intractably large number of possible ground-motion sets, and so identifying the best set is a computationally expensive combinatorial optimization problem. Following the work done by Jayaram et al. [70] and more recently by Baker and Lee [71] the selection algorithm first uses simulation techniques, Monte Carlo simulation (MCS) or Latin Hypercube sampling (LHS), to probabilistically generate multiple response spectra from a distribution parameterized by the target means and variances. For each simulated response spectrum, a ground motion with a similar response spectrum is then selected. Since the simulated response spectra have the desired mean and variance, the response spectra of the selected recorded ground motions will also have the desired mean and variance. A greedy optimization technique then further improves the match between the target and the sample means and variances. This step replaces one previously selected ground motion at a time with a record from the ground-motion database that causes the best improvement in the match between the target and the sample means and variances.

2.2.4   Target seismic conditional distributions

For estimating the seismic demand a typical approach is to use the same records at different seismic intensity levels. This is not realistic since different target properties of the ground motions are expected at each intensity level. Furthermore, in some cases, the ground motion records are selected from a target spectrum such as a uniform hazard spectrum (UHS). Spectral accelerations given by a UHS are obtained by considering the same probability of exceedance for all periods. Thus, it does not represent the spectra of any single seismic event [72] and, due to its inherent conservatism, is often considered unsuit-

able to be used in a probabilistic approach, such as a fragility analysis [21]. Accordingly, in recent years, the implementation of specific target conditional distribution, such as the conditional (mean) spectrum (CS) [73] or the generalized conditional intensity measure approach [74], have been recommended over the frequently employed uniform hazard spectrum due to its capability to support the selection of records that match proper ground motion characteristics for a given intensity measure level [75].

Conditional spectrum - CS

This approach allows to select different sets of ground motion records with appropriate target properties at each seismic IM level. The conditional spectrum describes the expected (mean and variance) response spectrum of a ground motion conditioned such that spectral acceleration matches a target amplitude at a given period, $Sa_H(T^*)$ [73]. The first step to compute the CS is to determine a target spectral acceleration at a conditioning period, T^*. One must then identify a representative earthquake scenario in terms of magnitude (M), distance (R) and ε-value, which is a measure of the difference between the spectral acceleration of a record and the mean of a ground motion prediction equation at the given period. If the target $Sa_H(T^*)$ value is obtained from probabilistic seismic hazard analysis (PSHA), this $M-R-\varepsilon$ scenario can be obtained from disaggregation. Afterwards, a mean response spectrum, $\mu_{\ln Sa}(M, R, T_i)$, and its associated standard deviation, $\sigma_{\ln Sa}(T_i)$, are computed using this scenario and a ground motion model consistent with the PSHA. The conditional mean spectrum in Eq. (2.2) and the conditional standard deviation spectrum in Eq. (2.3), can now be computed according to Baker [73].

$$\mu_{\ln Sa(T_i)|\ln Sa(T^*)} = \mu_{\ln Sa}(M, R, T_i) + \sigma_{\ln Sa}(T_i)\varepsilon(T^*)\rho(T_i, T^*) \qquad (2.2)$$

$$\sigma_{\ln Sa(T_i)|\ln Sa(T^*)} = \sigma_{\ln Sa}(T_i)\sqrt{1 - \rho^2(T_i, T^*)} \qquad (2.3)$$

where $\rho(T_i, T^*)$ is the correlation coefficient between a pair of ε at two periods and $\varepsilon(T^*)$ is the ε-value at the conditioning period. Eq. (2.2)–(2.3) coupled together are defined as the CS.

Although improved with respect to the traditional UHS, one limitation of the ground motion selection method based on spectral acceleration matching is that only the characteristics of ground motion represented in terms of (linear) spectral acceleration are considered, whereas it is acknowledged that the severity of ground motion, in general, depends on its intensity, frequency content, and duration [74]. Therefore, a further refined method,

the generalized conditional intensity measure, was proposed by Bradley [74] for ground motion record selection.

Generalized conditional intensity measure - GCIM

The essence of the GCIM approach is the definition of a multivariate distribution for any set of ground-motion intensity measures (IM_i) conditioned on the occurrence of a specific ground-motion intensity measure (IM_j) [74]. The calculation of the GCIM distributions involves two major steps: (i) determining the probability that if a ground motion was observed with IM_j, it was caused by a particular earthquake scenario; and (ii) given the observed ground motion with IM_j from a particular earthquake scenario, defining the distribution of the other ground motion intensity measures. Moreover, the general variable $\text{IM}|Rup$ (where "$|Rup$" indicates conditioning on a specific earthquake rupture scenario) is assumed to follow a multivariate log-normal distribution, as shown in Eq. (2.4), and the marginal distributions of all of the scalar intensity measures in $\text{IM}_j|Rup$, $\text{IM}_i|Rup$ can be estimated via existing GMMs.

$$f_{\text{IM}_i|Rup,\text{IM}_j} \sim \ln \mathcal{N}(\mu_{\ln \text{IM}_i|Rup,\text{IM}_j}, \sigma^2_{\ln \text{IM}_i|Rup,\text{IM}_j}) \tag{2.4}$$

Within the GCIM framework, any number of ground-motion IMs identified as relevant for a particular seismic response problem can be considered allowing the estimation of conditional distributions based on the full distribution of disaggregation results. This method proposes therefore a more realistic target distribution and its application has produced encouraging results by recent studies in different contexts [25, 26, 27]. Further details regarding the GCIM methodology are given in Section 4.4.1.

2.3   Fragility analysis

As it was mentioned before the codes and procedures for earthquake engineering show a tendency toward performance-based seismic design. For the implementation of this type of design, the capacity and demand of the structural system have to be defined. The capacity and demand models must represent the variability in the test data in a probabilistic manner, as they do with fragility functions.

2.3.1   Fragility curves

A univariate fragility function, as depicted in Figure 2.5, is a conditional probability that gives the likelihood that a structure will meet or exceed a specified level of damage for a

given ground motion intensity measure [61]. This conditional probability is given by:

$$\text{Fragility} = \text{P}_f(\text{LS} \mid \text{IM} = s) \tag{2.5}$$

where LS is the damage limit state, IM the seismic intensity measure and s is the achieved condition for the measured intensity of the seismic event.
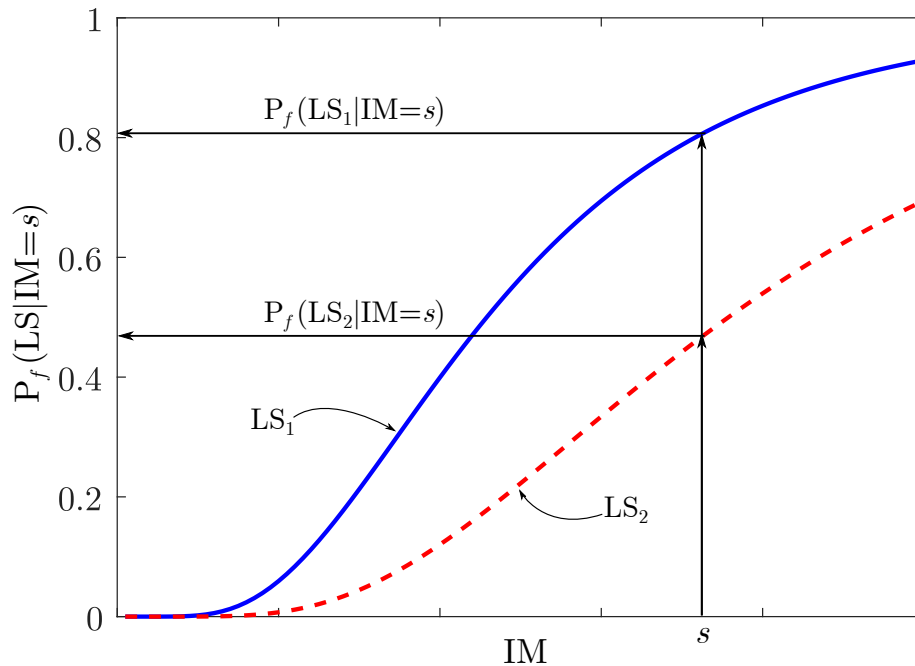
Figure 2.5   Conceptual fragility curve

The shape of a fragility curve describes the uncertainty in the capacity of the system to withstand a load or, alternatively, uncertainty in what load will cause the system to fail. If there is little uncertainty in capacity or demand, the fragility curve will take the form of a step function. For elastic, poorly understood, or complex systems, there is uncertainty in the capacity of the system to withstand a load. In these cases, the fragility curve takes the form of an S-shaped function. The S-shaped fragility curve is appropriate when there is uncertainty in the capacity of the system to withstand a load. The choice of the intensity measure must allow to appropriately reflect the fragility. This choice is not an easy task when developing fragility curves as this measure should help to effectively represent the seismic hazard of the site and the behaviour of the structure under consideration. Moreover, this measure must be practical and easy to implement and use. It has been shown that the spectral acceleration at the fundamental period ($Sa(T_1)$) is the preferred IM for evaluating seismic vulnerability of a particular structure [76, 77, 78]. However, in the case of concrete dams, recent studies [30] have shown that among the motion-dependent scalar

IMs, PGV is the best option while the combined acceleration response spectra, Sa^{1-to-N}, is the most practical and proficient structure-dependent spectral IM.

Approaches for developing fragility curves can be classified into three broad categories: judgemental, empirical and analytical, as further described below. Judgemental approaches are based on expert opinion or engineering judgement, empirical approaches are based on observations and analytical approaches are based on numerical models. Each approach differs in terms of the level of effort required to implement it and the precision that is attached to the results. However, none of the approaches is always best. The choice of what approach to use involves making a trade-off between cost and precision that is appropriate for the application.

Judgemental approaches

Fragility curves that are based on some form of expert opinion are classified as judgemental. When relying on expert opinion, it is important to devise replicable and verifiable procedures to elicit the opinions from experts. Fragility curves based on experts judgement have been developed to assess the seismic vulnerability of the components of hydroelectric facilities in Canada [79]. This study was based on the report ATC-13 from the Applied technology Council for California [80] done with the opinion of 71 experts. The applicability of this study depends on the modifications that must be done to fit the seismic characteristic of eastern Canada. This subjectivity produces that the judgemental approaches are often used as a last resort because of limitations in the availability of observational data and models.

Empirical approaches

Empirical fragility curves are based on observational data documenting the performance of structures under a variety of loads. Observational data tend to be highly specific to their source situations and may be sparse in the domain representing the more extreme events, which may also tend to be the events of most interest. However, the approach is generally limited to situations in which a sufficient quantity of data can be collected, and lack of data is often cited as a barrier to use the approach. Empirical fragility curves are available in the literature for bridges and buildings [76, 81]. Currently, no empirical fragility curve is available in the literature for dams.

Analytical approaches

Analytical fragility curves are based on structural models that characterize the performance limit state of the structure. The performance of the structure is a function of some

vector of "basic" variables, $\mathbf{x}$. These variables determine both the capacity of a structure to withstand a load and the demand placed on the structure. Basic variables include material properties, geometry, or dimensions; they could also include environmental variables (such as temperature or humidity) that might in some way affect the capacity. The limit state equation, also known as the performance function, can be expressed, in terms of safety margin, as the difference between capacity $R(\mathbf{x})$ and demand $S(\mathbf{x})$:

$$Z = p(R, S) = R(\mathbf{x}) - S(\mathbf{x}) \tag{2.6}$$

However, this formulation is still not the most general, since it assumes that there are distinct capacity and demand values. In many structural reliability problems such a distinction is not possible because they cannot be explicitly defined [82]. Thus, the general structural reliability problem in terms of a limit-state function is defined as :

$$p = p(\mathbf{x}) \tag{2.7}$$

The solution space consists of three regions: $p(\mathbf{x}) < 0$ is the failure domain; $p(\mathbf{x}) = 0$ is the limit state surface; and $p(\mathbf{x}) > 0$ is the survival domain.
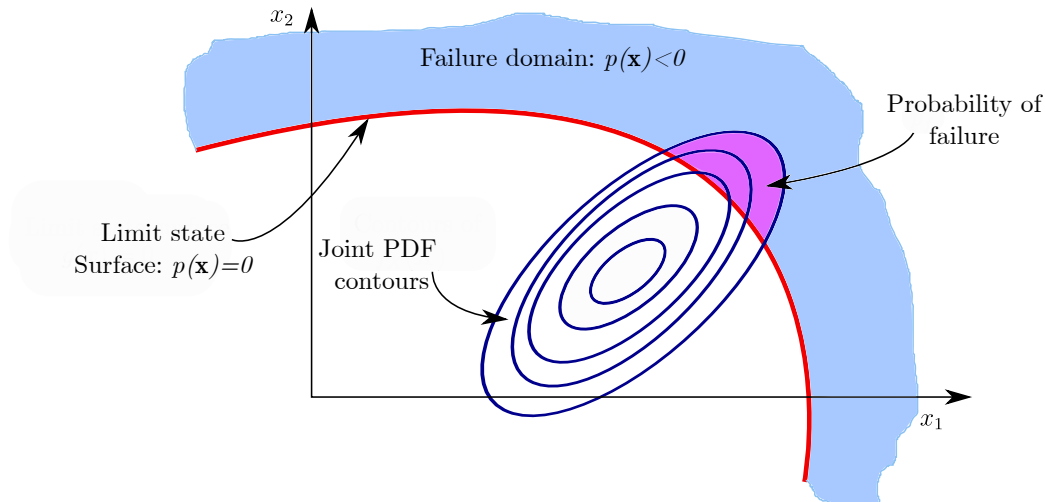
Figure 2.6   Conceptual 2D fragility problem

Basic variables can be either random variables or deterministic variables. The probability of failure is given by integration of a multivariate probability density function (PDF) for the n-dimensional vector of basic random variables, $f_{\mathbf{x}}(\mathbf{x})$, over the failure domain, $p(\mathbf{x}) \leq 0$:

$$\text{Fragility} = \text{P}_f(p(\mathbf{x}) \leq 0) = \int \ldots \int_{p(\mathbf{x}) \leq 0} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \tag{2.8}$$

A fragility curve is constructed by calculating the probability of failure under loads ranging from those at which failure is highly unlikely to those at which failure is almost certain. Analytical approaches can be decomposed into four distinct groups based on whether the limit state function is an explicit function or an implicit function and whether the probability of failure is obtained using analytical solution methods or numerical solution methods. An explicit limit state function is one that could be written explicitly in terms of basic random variables. An implicit limit state function is one that cannot be written in closed form as a function of basic variables, but is implied through a numerical model. Among them, the most popular used in the literature [82] are listed below,

   – Analytical solution methods (explicit).
     -First-order second-moment (FOSM) analysis.
     -First-order reliability method (FORM).
     -Second-order reliability method (SORM).
   – Numerical solution methods (implicit).
     -Monte Carlo simulation.
     -Latin Hypercube simulation.

Numerous studies showed that fragility may be adequately described, using a log-normal cumulative distribution function (CDF) [62, 76, 60]. Consequently the fragility is commonly approximated by a log-normal law with the following form,

$$P_f(s) = \Phi_{\mathcal{N}} \left(\frac{\ln(s/\mu)}{\beta} \right) \tag{2.9}$$

in which $\Phi_{\mathcal{N}}$ is the standard normal CDF, μ is the median capacity (expressed in units that are dimensionally consistent with the demand parameter, s) and β, is the standard deviation. Other CDFs like the Normal, Weibull distribution, etc., might be used [62, 83], as well as any form of Sigmoid function. Regarding the fitting techniques, maximum likelihood estimator (MLE), least-squares standard and least-squares in the linear space can be used to estimate the parameter of the fragility curves.

2.3.2   Fragility curves applied to dams

The implementation of fragility curves in the domain of dams is a relatively new practice. The first study in the field was carried out by Tekie and Ellingwood [36] who illustrated the methodology applied to the Bluestone Dam on the New River in West Virginia, which was designed in the late 1930s. Four limit states, confined to structural failure modes, were considered in the seismic fragility analyses: material failure-concrete (at the neck of the dam), material failure-foundation (at the toe), sliding at the dam-foundation interface,

and deflection of the top of the dam relative to the heel. Non-linear temporal dynamic analysis were performed including the reservoir-dam-foundation interaction effects. A total of nine uncertain parameters were defined in this study. Twelve samples of the model were obtained using the method of LHS and were coupled with twelve historical accelerograms. These samples were analysed for six levels of seismic intensity, by calibrating the acceleration time compared to the spectral acceleration measured at the pitch period, to obtain the fragility curves. This study reveals that in all cases the log-normal model of fragility fit the results of the simulation very well. In the same manner, Mirzahosseinkashani and Ghaemian [84] applied this criteria in order to illustrate seismic fragility curves for Pine Flat Dam. A non-linear analysis of the tallest monolith of the dam was performed with a flexible massless foundation and considering the dam-reservoir-foundation interaction. The maximum PGA that the dam-reservoir-foundation can endure in every horizontal earthquakes was obtained when the system reaches to 5% energy balance error based on nonlinear dynamic time history analysis. Two limit states were considered. The first limit state, was based on the crack length at the base and the second one based on the total area of cracked elements in the body of dam. The log-normal distribution was used for developing the fragility curves. The seismic fragility curves illustrated in this paper demonstrate that the occurrence of structural limit states are probable for massless foundation.

The same methodology proposed by Tekie and Ellingwood was used by Ghanaat et al. [85] to assess the vulnerability of Mühleberg Dam in Switzerland. Thirty samples were obtained with the LHS method and were simulated for eight levels of seismic intensity by calibrating the accelerograms based on the peak acceleration at the rock. Six uncertain parameters were considered. Dynamic non-linear temporal analyses were performed using the computer software SAP2000 and LS-Dyna. A log-normal distribution was used to represent results. The fragility analysis was performed first including all variables and later including only the random uncertainty related to the earthquake. Ghanaat et al. [83] have also studied the seismic vulnerability of a case study dam using LHS and advanced non-linear analysis with structural failure capability considering full dam-water and dam-foundation interaction. Sliding at the dam base and lift joints were identified as two prominent failure modes of the dam. Random and uncertainty input variables influencing the seismic fragility were generated using LHS and randomly selected to develop seismic fragilities for both failure modes. Ten samples were obtained and an incremental dynamic analysis (IDA) was conducted for each sample until reaching the limit state of damage. The probability of failure for each random non-linear trial was calculated as a function of the peak failure acceleration associated with the incipient sliding. The calculated results

were used to obtain the best-fit distribution to the data. A Weibull distribution was used to estimate the fragility curves.

Similarly, the work performed by Lupoi and Callari [38] presents a probabilistic seismic assessment method able to manage the physical complexity of the dam-foundation-reservoir system and the uncertainties regarding structural data and external actions. The methodology was applied to Kasho dam, a concrete gravity dam located in Japan. The seismic response of the structure was estimated from a reduced number of dynamic time history analyses, employing a finite element discretisation of the dam-reservoir-foundation structure. The limit states under consideration were the excessive deformation of the dam, joints cracking recovery and concrete-rock interface slipping. A probabilistic demand model was obtained by calculating the response of the dam with ten historical accelerograms. The fragility curves were obtained via a standard Monte Carlo simulation procedure. No probability was assumed to represent the fragility curves.

Further studies on the subject were performed by Zhong and Lin [40]. With consideration of the persistent uncertainty in ground motion input as well as material properties of concrete, 180 non-linear seismic analyses were performed based on which, typical seismic damage modes were obtained. This research used samples obtained by thirty Monte Carlo simulation and six levels of seismic intensity for determining the seismic vulnerability of a gravity dam. Three parameters were defined as uncertain and five limit states were defined qualitatively as a function of the intensity and cracking magnitude. In this study, no probability distribution was used for estimating fragility.

The most recent studies belongs to Bernier et al. [20, 21], Hariri-Ardebili and Saouma [29], Lallemant et al. [86] and Segura et al. [87]. In the two first studies [20, 21], the methodology is applied to the highest concrete gravity dam in Quebec. The finite element method is used to model a single block of the dam and it takes into account the different interactions between the dam, the reservoir and the foundation. Uncertainties in the ground motions and modeling parameters are included, and a sampling technique is used to propagate these sources of uncertainty. A sensitivity analysis is also performed to determine the model parameters that have a significant influence on the seismic response of the system. The fragility curves are developed using non-linear time-history analysis to evaluate two limit states: sliding at the dam base and at lift-joints. In the fist paper of Bernier et al. [20], the uncertainty related to the spatial variation of angle of friction across these large infrastructure systems is included in the fragility analysis through the incorporation of random fields modeling. However, the results reveal that this additional source of uncertainty has only a slight impact on the fragility of the structural system and can be

neglected. In the second paper of Bernier et al. [21] the same methodology is used together with the CS method to select ground motion times series.

In the study of Hariri-Ardebili and Saouma [29], seismic fragility curves for gravity dams with and without ground motion vertical component are explored. The structural analyses are performed using multiple-record incremental dynamic ones. An optimal intensity measure parameter is also selected among 37 variations, and it is determined that the combined spectral acceleration leads to lowest dispersion. The derived fragility curves using scaled records are compared with those from probabilistic seismic demand analysis (un-scaled records). Results show acceptable consistency between the two methods.

In like manner, the study performed by Lallemant et al. [86] provides a synthesis of the most commonly used methods for fitting fragility curves and highlights some of their significant limitations. More novel methods are described for parametric fragility curve development (generalized linear models and cumulative link models) and non-parametric curves (generalized additive model and Gaussian kernel smoothing). It also proposes methods for treating the uncertainty in intensity measure, a common issue with empirical data.

Finally in the study by Segura et al. [87], a similar model of the case study dam in Bernier et al. [20, 21] is used to develop up-to-date fragility curves for the sliding limit states of gravity dams in Eastern Canada using a record selection method based on the generalized conditional intensity measure approach. These fragility functions are then combined with the recently developed regional hazard data to evaluate the annual risk, which is measured in terms of the unconditional probability of limit state exceedance. It was observed that the more accurate procedure proposed herein produces less conservative fragility estimates for the case study dam.

These studies show that as far as dam safety is concerned the most reliable methods should be implemented. Undoubtedly, one of the best methods of analysing gravity dams is a non-linear dynamic time history analysis [76]. To minimize the cost of the non-linear finite element analyses required to develop the fragility functions, sampling techniques (usually referred to as variance reduction techniques) must be used. Crude MCS can produce most accurate results, but it requires an enormous amount of computational cost when non-linear time history analyses are involved [82]. For most fragility analyses the pure random sampling of MCS may not be necessary, as long as distributions of the input parameters are reproduced accurately. According to the existing studies in the literature, LHS technique has been found to be very useful in reliability problems involving complex

systems. Further and more exhaustive information can be found in the state of the art review of seismic fragility applied to concrete dams by Hariri-Ardebili and Saouma [28].

2.3.3 Fragility surfaces

Earthquake shaking represents complex loading to a structure. It cannot be accurately characterized by a single parameter such as peak ground acceleration [88]. Traditional vulnerability assessment methods develop fragility functions by using a single parameter to relate the level of shaking to the expected damage, which consequently produces a robustness of predictions that is highly dependent on the selected parameter. However, the estimation of the fragility of the system can be potentially improved by increasing the number of parameters; in this way, a more complete description of the properties of ground motions can be obtained [89]. Single parameter demand models and fragility curves suffer from two potential drawbacks: (i) inability to assess the impact of structural model parameter variation on structure performance during earthquakes without costly re-analysis for each new set of parameter combinations and (ii) lack of flexibility to incorporate field instrumentation data from monitoring of existing structures to enable the updating of seismic fragility estimates [90]. Furthermore, the effect of the variation of the material properties in the seismic fragility analysis of structures with complex numerical models, such as dams, is frequently overlooked due to the costly and time-consuming revaluation of the numerical model. Consequently the use of multi-parameter models to predict the response of a certain structure is beginning to be used increasingly. The goal is to identified the role of various strong motion or structural parameters on the induced damage in the structure using numerical calculations and vulnerability analysis. The most influential parameters are then used to build multi-variable fragility functions, in order to reduce some of the uncertainty inherent in the response to seismic loading. Similar to fragility curves, multivariate fragility functions offer the conditional probability of exceeding different limit states given the occurrence of an earthquake of a certain intensity. The only difference is that the specific limit state is characterized with n parameters $p_1, p_2 \ldots p_n$ instead of one parameter, as is the case with fragility curves. Hence the probability of limit state exceedance is conditioned on the resulting set of critical parameters. The fragility function corresponding to the limit state l is defined as follows:

$$F_l(x_1, x_2, \ldots, x_n) = \mathrm{P}_f(\mathrm{LS} > \mathrm{LS}_l | p_1 = x_1, p_2 = x_2, \ldots, p_n = x_n) \qquad (2.10)$$

where LS is the limit state damage index, and LS_l is the value corresponding to the lth limit state. This results in an equation for a fragility surface, as depicted in Figure 2.7,

that offers a more complete and accurate view of the vulnerability of the structure. Such fragility surfaces can be implemented within earthquake risk evaluation tools and they should provide more precise damage estimations. It is expected that this procedure can lead to more accurate planning and retrofitting policies for risk mitigation.
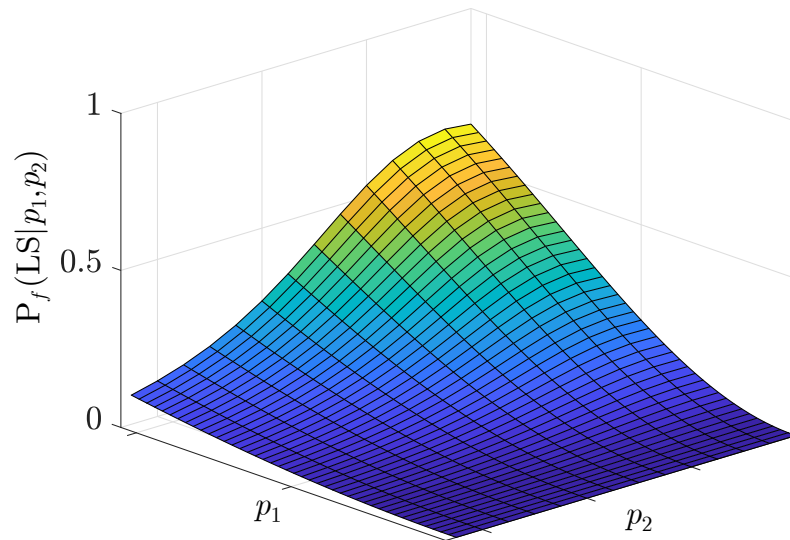
Figure 2.7   Conceptual fragility surface

While fragility curves are usually represented by well-known and readily parameteriz-able probability distributions like the log-normal one, the problem gets more complex for surfaces, where bivariate distributions must be computed. Follows, a description of the fragility surfaces construction methods used in the literature and applied to seismic engineering. As it will be seen, there are only a few studies where fragility surfaces have been developed for dam-type structures, most of the available studies in the literature have been implemented for bridges [91] and buildings [92]. Nevertheless, they can be extrapolated to dams or any other system.

PEER methodology for demand fragility surfaces

The Pacific Earthquake Engineering Research Center (PEER) in their 2011/01 report [91], proposed a methodology to develop demand fragility surfaces characterizing probability of exceeding common engineering demand parameters (EDP) for bridges as a function of the most important and useful seismic IM. The adopted approach was to utilize six constants to fully define the demand fragility surface. The fragility functions at every EDP value can be defined by a median, λ_f, a standard deviation of natural logarithms of the data, ξ_f, and an additional scalar value representing the limiting value of cumulative probability, p_{max}. Linear regression is used to define the medians, standard deviations, and peak probabilities as functions of the EDP. Linear regression requires specification of

a slope and an intercept value, hence two constants each are required for median, standard deviation, and maximum probability, and the entire demand fragility surface can therefore be defined by six constants, as shown in Eq. (2.11)–(2.13),

$$\lambda_f = \beta_0 + \beta_1 \ln(edp) \tag{2.11}$$

$$\xi_f = \beta_2 + \beta_3 \ln(edp) \tag{2.12}$$

$$p_{max} = \beta_4 + \beta_5 \ln(edp) \leq 1 \tag{2.13}$$

where $\beta_0, \dots, \beta_5$ are the linear regression coefficients. By combining Eq. (2.11)–(2.13) with commonly used parametric formulations for fragility curves, such as the log-normal CDF, the probability of exceedance can be calculated using Eq. (2.14) for all EDPs,

$$\mathrm{P}_f(\mathrm{EDP} < edp \mid \mathrm{IM} = im) = 0 \leq \Phi_{\mathcal{N}} \left(\frac{\ln(im) - (\beta_0 + \beta_1 \ln(edp))}{\beta_2 + \beta_3 \ln(edp)} \right) (\beta_4 + \beta_5 \ln(edp)) \leq 1 \tag{2.14}$$

where $\Phi_{\mathcal{N}}$ is the standard normal cumulative distribution function, im is the seismic IM level, and edp is an engineering demand parameter level, defining an specific limit state. Eq. (2.14) provides a convenient means of specifying the entire demand fragility surface using only six variables, where the fitted surfaces agree reasonably well with the discrete data.

Neighbourhood method

This method is meant to be used for fragility surfaces as a function of two seismic IMs. For two chosen strong-motion intensity parameters, a $x-z$ space is built containing points of coordinates (x_i, z_i), x and z representing the two seismic IMs. This space is associated with the following norm, between points $A(x_1, z_1)$ and $B(x_2, z_2)$:

$$d(\mathrm{A,B}) = \sqrt{\left[\frac{\ln(x_1/x_2)}{\ln(x_{max}/x_{min})} \right]^2 + \left[\frac{\ln(z_1/z_2)}{\ln(z_{max}/z_{min})} \right]^2} \tag{2.15}$$

To avoid bias due to differently-scaled parameters this norm is also normalized by the amplitude of each parameter. For each point of the $x-z$ space it can be defined a neighbourhood V of radius d, where the probability to reach or exceed the limit state l

can be evaluated,

$$P_f(\text{LS} > \text{LS}_l) = \frac{N_{V,l}}{N_{V,tot}} \tag{2.16}$$

where $N_{V,tot}$ is the total population in the neighbourhood V and $N_{V,l}$ denotes the number of points for which the damage state reaches or exceeds LS_l.

However, surfaces generated using this original approach present some anomalies (such as local peaks) [92]. This observation highlights the main issue of this approach: in order to define the neighbourhood V of each point, a value of the radius d has to be chosen more or less arbitrarily. Greater values of the radius d help to smooth the surface, but there is a risk of generating a description of the fragility behaviour which might be too blurry (loss in precision). On the contrary, to obtain a sharper image of the structure vulnerability, it is recommended to decrease the value of d (at the expense of regularity). This method has been satisfactory used in the development of fragility surfaces for reinforced concrete buildings [92, 88].

Hybrid parameter method

The hybrid parameter methodology is an improvement of the neighbourhood method, proposed for fragility surfaces as a function of two correlated parameters, as it is the case of seismic IMs [92]. The main idea is to generate an hybrid parameter, $Q(x, z)$, as a combination of the two considered fragility surface's parameters. It is then possible to fit the points by a well-known analytical function, depending on the variable R, hence it is a function of the two variables x and z. If the parameter Q is defined as a linear combination of x and z and a log-normal CDF is considered, then the fragility surface can be determined through Eq. (2.17)–(2.18) as follows,

$$P_f(\text{LS} > \text{LS}_k \mid \text{IM}_1 = x, \text{IM}_1 = z) = \Phi_{\mathcal{N}}\left(\frac{\ln(Q) - \mu_Q}{\beta_Q}\right) \tag{2.17}$$

$$Q(x, z) = \alpha_0 + \alpha_1 \ln x + \alpha_2 \ln z \tag{2.18}$$

where α_0, α_1 and α_2 are linear correlation coefficients, and μ_Q and β_Q are the parameters characterizing the standard normal cumulative distribution function, $\Phi_{\mathcal{N}}$. Besides a linear correlation, other formulations such as the logarithmic distance between points, has produced fair results [92]. The methodology has been employed for generating fragility surfaces for reinforced concrete buildings [92, 93] and for multi-span bridges [94].

Logistic regression

It is a statistical learning model commonly used in a wide-range of applications where predictions of a binary outcome are sought. In our case the binary outcome of interest is the exceedance or not of the prescribed damage level. Logistic regression is the extension of linear regression for classification problems which basic idea is to transform the linear function output into the (0, 1) interval describing the probability $P_f(p = +1 \mid \mathbf{x})$ [95]. The output of a linear model can be transformed in the interval (0, 1) by passing it through a sigmoid function. Let p denote the dependent binary variable (this is the damage indicator variable which becomes 1 if the prescribed damage level is exceeded and 0 otherwise), the fragility, or probability of exceedence conditioned on the seismic IM and a vector variable $\mathbf{x}$ of structural parameters is

$$P_f(p = 1 \mid \mathbf{x}) = \frac{e^{\sum_i \vartheta_i \mathbf{x}_i}}{1 + e^{\sum_i \vartheta_i \mathbf{x}_i}} \tag{2.19}$$

where ϑ_i are the coefficients of the model which must be selected in order to provide the best fit to the available data. These coefficients can be estimated with the maximum likelihood estimation method, with Bayesian inference, Markov chain Monte Carlo (MCMC), etc. Logistic regression has been widely used in the literature for the fragility assessment with fragility curves for bridges [90, 96, 97] and ports [98] and, to a lesser extent, with fragility surfaces [99].

2.3.4   Fragility surfaces applied to dams

As it was aforementioned, the estimation of the fragility of the system can be potentially improved by increasing the number of parameters. Noted advantages of these parameterized or multivariate fragility functions include the potential for efficient posterior uncertainty propagation, exploring sensitivities or the influence of design parameter variation, and enabling application across a portfolio of structures. Nevertheless, given the large number of simulations required, the development of fragility surfaces or multivariate fragility functions that leverage complex numerical models, such as dams, can impose high computational burdens. Accordingly, there are not many studies in the literature that have developed fragility surfaces for dams, and the ones that are available do not provide a methodology or a parameterized formulation.

Hariri-Ardebili and Saouma, in their state-of-art paper [28], present an exhaustive review of seismic fragility functions (curves and surfaces) for the vulnerability assessment of dams. In this review, the raw data from Yao et al. [100], where a seismic fragility analysis is

applied to the safety evaluation of a concrete arch dam, is used to generate fragility surfaces as a function of PGA and the EDP joint sliding and joint opening with an EDP-based Bayesian cloud analysis [101]. Yao et al. [100] considered only ground motion record-to-record variability, where the case study is a 305 m high arch dam with 72 % of filled reservoir. The dam originally included 25 vertical contraction joints; however, only 3 are modeled using commercial finite element code with nonlinear contact models, while the concrete is assumed to be linear elastic. Finally, the hydrodynamic pressure is modeled using Westergaard's added mass approach. A total of 18 ground motions are selected and categorized in 3 groups, and no fragility functions are given. However, the seismic fragility surfaces for joint response, developed by Hariri-Ardebili and Saouma [28], based on raw data from Yao et al. [100] show that the probability of exceedance of a specific PGA is higher for joint sliding than for opening (mostly at smaller EDP values).

Hariri-Ardebili and Saouma [30] analyzed the tallest non-overflow monolith of a 122 m high gravity dam. 2D mesh of the dam and foundation was provided while the only source of nonlinearity being the interface between the two. It is noteworthy that beside gravity and hydrostatic loads, the uplift pressure is automatically adjusted in terms of crack length. The aleatory uncertainty included by ground motion record-to-record variability only while the epistemic uncertainty was ignored. Using the algorithm proposed by Jarayam et al. [70], 100 ground motions were selected. Fragility curves and surfaces were derived with the previously mentioned cloud analysis [101], using 70 different IMs and the optimal one was then identified. The results show that the fragility curves as a function of the spectral acceleration at the fundamental period for joint opening and sliding at the dam-rock interface for three different LSs: 2 mm (initiation of opening/sliding), 5 mm (propagation of the opening/sliding), and 8 mm (near collapse condition). In all cases, joint sliding has the highest probability of exceedance. As the LS increases, the differences between those two fragility curves diminishes. In-so-far all the fragility curves are derived for a finite values of the EDP (such as sliding of 3, 5 and 8 mm). However, this was generalized for continuous values of EDPs and a fragility surface was then generated.

Finally, Hariri-Ardebili and Saouma [28], from the data adapted from Hariri-Ardebili and Saouma [102], generated fragility surfaces with the methodology explained before, as a function of the crest horizontal displacement and the LS. Whereas most fragility functions are expressed in terms of IMs (specific to a site and a structure), it is preferable to express them in terms of EDPs. The advantage being twofolds: (i) curve is less site specific, and more generic thus potentially applicable to other similar structures; and (ii) it ties with the PBEE paradigm. The finite element model, similar to a previous one [30], with the

difference that beside joint nonlinearity, concrete nonlinearity is also accounted for through smeared crack model. Two types of ground motion combinations are considered, horizontal only and horizontal and vertical. Furthermore, for each case two loading scenarios are further considered, full and empty reservoirs. Finally, each one of these four cases is subjected to 21 site-specific actual ground motions scaled by 14 seismic intensity levels to perform an incremental dynamic analysis for collapse fragility curves.

2.4   Surrogate or meta-models

Machine Learning (ML) describes a family of methods that allow learning from data, what relationships exist between quantities of interests [95]. The goal of learning relationships between quantities of interests is to gain information about how a system works and to make predictions for unobserved quantities. This new knowledge can then be employed to support decision making. ML techniques are usually divided into two main types supervised and unsupervised [95], the first being the one that will be address in this document. In the predictive or supervised learning approach, the goal is to learn a mapping from inputs $\mathbf{x}$ to outputs y, given a labeled set of input-output pairs $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n_f}$. Here $\mathcal{D}$ is called the training set, and n_f is the number of training examples. In the simplest setting, each training input $\mathbf{x}_i$ is a D-dimensional vector of numbers, representing, say, the characteristics and configurations of the system. These are called features, attributes or covariates. Similarly the form of the output or response variable can in principle be anything, but most methods assume that y_i is a categorical or nominal variable from some finite set, $y_i \in \{1, \ldots, C\}$ (such as collapse or survival of the structure), or that y_i is a real-valued scalar (such as crack length). When y_i is categorical, the problem is known as *classification* or pattern recognition, and when y_i is real-valued, the problem is known as *regression* [103]. Throughout this section, only regression problems will be considered.

Over the last two decades, there has been an explosion in the ability of engineers to build finite-element models to simulate how a complex structure will perform. A surrogate model is an engineering method used when an outcome of interest cannot be easily directly measured, so a model of the outcome is used instead [15]. The basic idea in the surrogate model approach is to avoid the temptation to invest computational budget in answering the question at hand and, instead, invest in developing fast mathematical approximations to the long running computer codes. One way of gaining this desirable increased insight into the problems being studied is via the use of surrogate (or meta) models. Such models seek to provide answers in the gaps between the necessarily limited analysis runs that can be afforded with the available computing power. The simplest, and currently most

common, use of surrogate models is to augment the results coming from a single, expensive simulation code that needs to be run for a range of possible inputs dictated by some design strategy (perhaps a planned series of runs or those suggested by some search process). The basic idea is for the surrogate to act as a "curve fit" to the available data so that results may be predicted without recourse to use of the primary source (the expensive simulation code). The approach is based on the assumption that, once built, the meta-model will be many orders of magnitude faster than the primary source while still being usefully accurate when predicting away from known data points.

To reduce the computational expense, surrogate models have been employed in structural reliability problems to approximate the response of structures with complex finite element models, or to estimate the limit state function using an approximating function. The meta-model can be described by,

$$y = g(\mathbf{x}) + v \tag{2.20}$$

where the surrogate model $g(\mathbf{x})$ statistically predicts the response of the structure, y, for a given set of covariates including intensity measures and model parameters, $\mathbf{x}$, and v is the error due to the lack of fit of the surrogate model. The scientific challenge of surrogate modeling is the generation of a meta-model that is as accurate as possible, using as few simulation evaluations as possible [104]. The process comprises three major steps which may be interleaved iteratively:

- Sample selection (also known as sequential design, optimal experimental design or active learning).
- Construction of the surrogate model and optimizing the model parameters (bias-variance trade-off).
- Appraisal of the accuracy of the surrogate.

The accuracy of the surrogate depends on the number and location of samples (expensive experiments or simulations) in the design space. Various design of experiments (DOE) techniques cater to different sources of errors, in particular errors due to noise in the data or errors due to an improper surrogate model.

2.4.1   Regression techniques

The following subsections provide a brief overview of the different regression techniques for their ability to offer viable meta-models of the seismic response of dams. Only the most relevant features of each technique will be presented given that exhaustive mathematical formulation can be found in the literature. Description of polynomial response

surface (PRS), adaptive basis function construction (ABFC), multivariate adaptive regressive splines (MARS), radial basis functions (RBF), support vector machine for regression (SVMR) and random forest for regression (RFR) techniques and their applications in the literature can be found in section 5.2.2.

Polynomial chaos expansions - PCE

Polynomial chaos expansions have been introduced in the literature on stochastic mechanics in the early 90's by Ghanem and Spanos [105]. In the original setting, a boundary value problem is considered in which some parameters are modeled by random fields. The quantities of interest are the resulting stochastic displacement and stress fields. Thus the use of PCE has been intimately associated with spatial variability and considered as a separate topic with respect to structural reliability for a while.

Polynomial chaos expansions can be considered as an intrinsic representation of a random variable that is defined as a function of the input random vector $\mathbf{x}$. In the context of structural reliability the limit state function leads to define the random margin $\mathcal{M}(\mathbf{x})$. The probability of failure is then defined by $P_f(\mathcal{M}(\mathbf{x}) \leq 0)$. Assuming that this variable has a finite variance and that the input parameters in $\mathbf{x}$ are independent, the following representation holds [106],

$$\mathcal{M}(\mathbf{x}) = g(\mathbf{x}) = \sum a_\alpha \Psi_\alpha(\mathbf{x}) \tag{2.21}$$

where $\Psi_\alpha(\mathbf{x})$ are multivariate orthonormal polynomials in the input variables and a_α are coefficients to be computed. Since the component of $\mathbf{x}$ are independent, the joint PDF is the product of the marginal. Considering the expansion in itself as a meta-model that is suitable for reliability analysis has been originally explored by Sudret and Der Kiureghian [107]. Later on, the use of PCE has started to be progressively used with the emergence of so-called non intrusive methods. More specifically the regression approach has been developed and applied to reliability analysis in Berveiller et al. [108, 109] and Choi et al. [110], among others. Further details and a much more comprehensive formulation can be found in [111, 112, 106].

Gaussian process - GP

Rather than claiming $g(\mathbf{x})$ relates to some specific models, it can represent obliquely, but rigorously, by letting the data "speak" more clearly for themselves [113] with a Gaussian process (GP) approach. This technique is still a form of supervised learning, but the training data are harnessed in a subtler way. As such, GP is a less parametric tool. How-

ever, it is not completely free-form, and if it is unwilling to make even basic assumptions about $g(\mathbf{x})$ then more general techniques should be considered. Gaussian process regression works by describing the prior knowledge of a system's response over its covariate domain using a joint multivariate normal probability density function. Then it employs the properties of the multivariate Normal in order to update the prior using empirical observations [95].

A GP is a generalization of the Gaussian probability distribution. Whereas a probability distribution describes random variables which are scalars or vectors (for multivariate distributions), a stochastic process governs the properties of functions. A GP assumes that $p(g(\mathbf{x}_1), \dots, g(\mathbf{x}_N))$ is jointly Gaussian, with some mean $\mathbf{m} = m(\mathbf{x})$ and covariance $\mathbf{\Sigma}(\mathbf{x}) = k(\mathbf{x}_i, \mathbf{x}_j)$, where k is a positive definite kernel function. The key idea is that if $\mathbf{x}_i$ and $\mathbf{x}_j$ are deemed by the kernel to be similar, then we expect the output of the function at those points to be similar, too. One of the main attractions of the Gaussian process framework is precisely that it unites a sophisticated and consistent view with computational tractability [114]. In supervised learning, some inputs $\mathbf{x}_i$ and some outputs y_i are observed. It is assumed that $y_i = g(\mathbf{x}_i)$, for some unknown function g, possibly corrupted by noise. The optimal approach is to infer a distribution over functions given the data, $p(g \mid \mathbf{X}, y)$, and then to use this to make predictions, $\hat{y}_*$, given new inputs, $\mathbf{x}_*$, as shown in Figure 2.8.
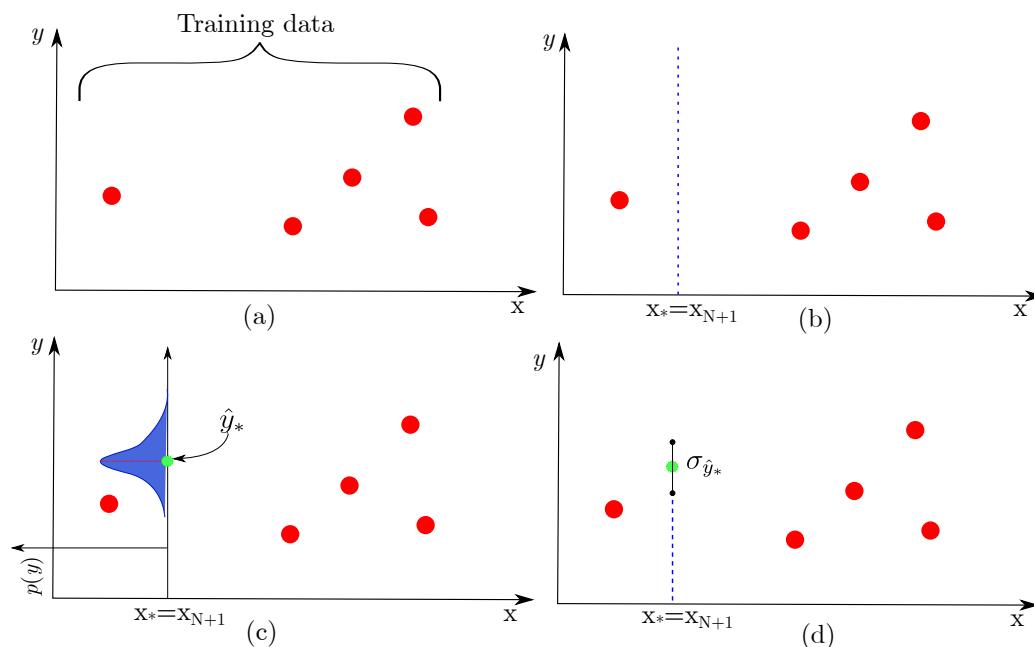
Figure 2.8 Gaussian process model prediction: (a) training data generation, (b) prediction point (c) estimation of y_* as a Gaussian distribution and (d) prediction error bars

Let the prior on the regression function be a GP, denoted by:

$$g(\mathbf{x}_i) \sim GP(m(\mathbf{x}_i), k(\mathbf{x}_i, \mathbf{x}'_i)) \tag{2.22}$$

For any finite set of points, this process defines a joint Gaussian:

$$p(g \mid X) \sim \mathcal{N}(g \mid \mathbf{m}, \mathbf{\Sigma}) \tag{2.23}$$

This GP will be used as a prior for Bayesian inference and one of the primary goals computing the posterior is that it can be used to make predictions for unseen test cases. Let g be the known function values of the training cases, and let g_* be a set of function values corresponding to the test set inputs, $\mathbf{x}_*$.

$$\begin{bmatrix} g \\ g_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m} \\ \mathbf{m}_* \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma} & \mathbf{\Sigma}_* \\ \mathbf{\Sigma}_*^\top & \mathbf{\Sigma}_{**} \end{bmatrix} \right)$$

where $\mathbf{m} = m(\mathbf{x}_i), i = 1, ..., n$ for the training means and analogously for the test means μ_*; for the covariance we use $\mathbf{\Sigma}$ for training set covariances, $\mathbf{\Sigma}_*$ for training-test set covariances and $\mathbf{\Sigma}_{**}$ for test set covariances. Since the values for the training set g are know, the conditional distribution of g_* given g can be expressed as:

$$g_* \mid f \sim \mathcal{N}(\mathbf{m}_* + \mathbf{\Sigma}_*^\top \mathbf{\Sigma}^{-1}(g - \mathbf{m}), \mathbf{\Sigma}_{**} - \mathbf{\Sigma}_*^\top \mathbf{\Sigma}^{-1} \mathbf{\Sigma}_*) \tag{2.24}$$

The predictive performance of GPs depends exclusively on the suitability of the chosen kernel [114, 113, 103, 95]. In other words, in order for the GP techniques to be of value in practice, we must be able to choose between different mean and covariance functions in the light of the data.

Neuronal networks - NN

Only the feedforward architecture for neuronal networks will be addressed in this section for its simplicity. In feedforward NN the information is transferred from the input layer to the output layer by propagating it into layers of hidden variables. The idea is to approximate complex functions by a succession of simple linear combinations of hidden variables organized in layers [95]. Linear regression can be employed to introduce the

concepts surrounding feedforward neural networks, according to Eq.(2.25):

$$z = b + \sum_{i=1}^{n} w_i \mathbf{x}_i \tag{2.25}$$

where the covariates $\mathbf{x}$ are associated with a single observed system response $y = z + v$, where z represents a hidden variable defined by the linear combination of covariate values using weights, w_i, and one bias parameters, b, and v is the model error.

Neural networks became extremely popular in the period from 1980 until the mid 1990' when they became overshadowed by other emerging methods. It is not until the mid 2000' that with the appellation of deep learning, they took again the leading role in the development of machine learning methods. In deep learning, the label *deep* refers to the great number of layers of hidden variables in the model. One limitation is that in order to achieve the exceptional performance it is renown for, deep learning requires large datasets containing from thousands to millions of labeled examples for covariates $\mathbf{x}_i$ and the system responses y_i [95].

2.4.2   Meta-models in the structural engineering domain

The combination of numerical models, probabilistic approaches and machine learning has gained considerable interest in the literature in recent years for engineering design and structural reliability [106, 115, 116]. This combination is justified by the significant randomness that characterizes not only the earthquake excitation but also the structural system itself (e.g., stochastic variations in the material properties, degradation due to aging and temperature fluctuation, etc.)

Two approaches so far are found in the literature. In the first one the basic premise of applying statistical learning techniques to provide meta-models is to replace the "true" seismic demand estimates with computationally efficient approximating functions. After a discrete surrogate model is employed for the failed/safe classification of the structure according to the damage limit states of the fragility analysis. This is the case of the work done by Gosh et al. [90], where a surrogate model demand was developed to approximate the relationship between the seismic response (peak seismic response of a bridge component) and a vector including ground motion intensity measures and parameters of the system. Lastly, logistic regression was employed to develop parametrized fragility models. This methodology is also used by Ataei and Padgett [97] and Balomenos and Padgett [98] in their research for coastal bridges under extreme load conditions caused by hurricanes

and wharf/pier structures in port facilities that are subjected to hurricane-induced storm surge and wave loading, respectively.

In the second approach, an acknowledged alternative for using meta-modeling in fragility analysis is applied. It consist in constructing meta-models for a continuous response parameter of the structure, where the response meta-models are themselves compared to capacity estimates for reliability computation. Typically, polynomial approximation has been employed to find the predictive models of structural behaviour used in structural reliability studies; these parametric models are known as response surfaces. However, this type of model is inherently limited by the properties of the polynomial function and its transformations. Hence, polynomial meta-models may not accurately represent the portion of the design space that is of interest to the engineer. Nevertheless, this approach is still considered owing to its simplicity and the fact that past studies have shown them to be efficient and accurate for assessing seismic performance of other complex structures [90, 117, 96, 118].

Non-polynomial-based meta-models, such as GP, that have the capability of accurately modeling large portions of highly non-linear and non-monotonic design spaces while retaining the highly desired capability of fast execution, are increasingly being used. Not only because of its predictive capabilities but also because an estimate of the accuracy of the GP surrogate model is always available as a natural by-product of the underlying GP mathematics. This method has been used mostly in the structural dynamics domain as an emulator tool [119, 120, 121, 122].

Surrogate modeling techniques within a seismic fragility framework have found recent applications for the safety assessment of buildings and bridges, among other structures [123, 118, 96, 90, 117, 124]. Even though many of these studies considered several seismic intensity measures, model parameters and geometric uncertainties for building the meta-models to predict the response of the structure, most of them do not clearly depict the influence of all the considered parameters in the form of multivariate fragility functions from the respective meta-models.

2.4.3   Surrogate models applied to dams

The research in this field is very limited and there is a gap between all the theoretical aspects and the real world dam engineering applications. Gaspar et al. [125] proposed a probabilistic thermal model to propagate uncertainties on some roller-compacted concrete's (RCC) physical properties. A thermo-chemo-mechanical model was then used to

describe the RCC behavior. Moreover, the global sensitivity analysis was performed to evaluate the impact of random variables. Mata et al. [126] proposed a method based on linear discriminant models for the construction of decision rules for the early detection of developing failure scenarios. They developed a single classification index by combining the physical measured quantities.

An extensive comparison between machine-learning data-based predictive models for monitoring the dam behavior can be found in Salazar et al. [127, 128]. In their study it is discussed and contrasted some of the machine learning based predictive models for dam safety assessment, i.e., random forests, boosted regression trees (BRT), neural network, support vector machine, and multivariate adaptive regression splines. The prediction accuracy in each case was compared with the conventional statistical model and BRT models stood out as the most accurate overall, followed by NN and RFR. It was also observed that the model fit can be improved by removing the records of the first years of dam functioning from the training set.

More recently, Hariri-Ardebili and Pourkamali-Anaraki [129, 130] and Hariri-Ardebili [131] have used machine learning techniques to perform reliability analysis applied to gravity dams against flooding, earthquakes and aging, considering, in some cases, explicit limit state functions and simplified FEM in others. In Hariri-Ardebili and Pourkamali-Anaraki [129], SVM method is adopted and applications for a simplified flood reliability assessment of gravity dams and for nonlinear seismic finite element method based analysis are presented. Up to seventeen random variables are considered in the example and the results of SVM are contrasted with classical reliability analyses techniques. Two examples are studied in this paper. The first one is a gravity dam analyzed analytically based on the limit equilibrium method (LEM) approach. In the second example, the finite element model of a gravity dam-foundation system is analyzed. Similarly, Hariri-Ardebili and Pourkamali-Anaraki [130] presented a simplified reliability framework for gravity dams subjected to flooding, earthquakes, and aging. Response of the dam is analyzed with explicit limit state functions. The probability of failure is directly computed by either classical Monte Carlo simulation or the refined importance sampling technique. Three classification machine learning techniques (i.e., K-nearest neighbor, SVM, and naive Bayes classifier) are adopted for binary classification of the structural results. Results are then generalized for different dam classes (based on the height-to-width ratio), various water levels, earthquake intensity, degradation rate, and cross-correlation between the random variables. Hariri-Ardebili [131], investigated the potential application of DOE in order to develop appropriate response surface meta-models and present an explicit expression

of dam response. More than ten DOE techniques are compared and contrasted and the accuracy of meta-models is evaluated using finite element simulations from crude MCS. The impact of polynomial degree is studied for an actual gravity dam with five random variables. Pine Flat gravity dam is selected as case study.

ML techniques have also found their application in structural health monitoring (SHM) for dams. Nguyen and Goulet [132] proposed an anomaly detection method that combines the existing Bayesian Dynamic Linear Models framework with the Switching Kalman Filter theory. The approach operates in a semi-supervised setup where normal and abnormal state labels are not required to train the model. The potential of the new method is illustrated on the displacement data recorded on a dam in Canada. The results show that the approach succeeded in identifying the anomaly caused by refection work, without triggering any false alarm.

Most of the studies that can be found in the literature on the seismic assessment of dams via machine learning techniques are limited to the consideration of a few meta-models in the context of a single study, simplified finite element models and univariate fragility functions. Consequently, they do not inquire in the most suitable meta-model for fragility analysis of this type of structure, nor they explicitly assess the influence of the variation of the model parameters on the seismic fragility analysis.

CHAPTER 3

METHODOLOGY

The main objective of this chapter is to present the case study dam to illustrate the methodology for modeling and characterization of the uncertainties associated with this type of structure to perform a seismic fragility analysis from the meta-models. Accordingly, this chapter includes: (i) the development and validation of a nonlinear finite element model of a dam-type structure, taking into account the soil-fluid-structure interaction and the characterization of the uncertainty sources and material properties; (ii) the procedure to conduct a PSHA and the ground motion record selection algorithm; (iii) the meta-model generation, training, comparison and validation process and (iv) the fragility point estimate generation and parameterized fragility surfaces fitting.

3.1 Case study dam modeling and characterization

The dam-reservoir-foundation (DRF) system studied was modeled with the finite element method following the recommendations of the United States Bureau of Reclamation [37, 133]. The model take into account the different interactions between the structure, the reservoir and the foundation. The model meshing was carried out using the computer software ANSYS-ICEM [134] while the computer software LS-Dyna [135] was used for modeling and analyzing the system considering as guidelines mostly the work done by Depolo et al. [136], Bernier et al. [20] and Noble [42].

3.1.1 DRF Finite element model

The numerical model must adequately represent dynamic behavior to minimize the epistemic uncertainties related to the modeling assumptions. Figure 3.1 presents the main features of the finite element model. To reduce the computational burden, an explicit integration method with single-point-of-integration hexahedral elements was used to estimate the seismic response of the dam.

Due to the presence of contraction joints, the interaction between the blocks of the dam is not significant enough to justify the realization of a complex 3D model. Thus, only one block is modeled with 3720 elements and a linear elastic material to model the concrete behavior. The non-linearities are introduced later on in the model in the form of contact
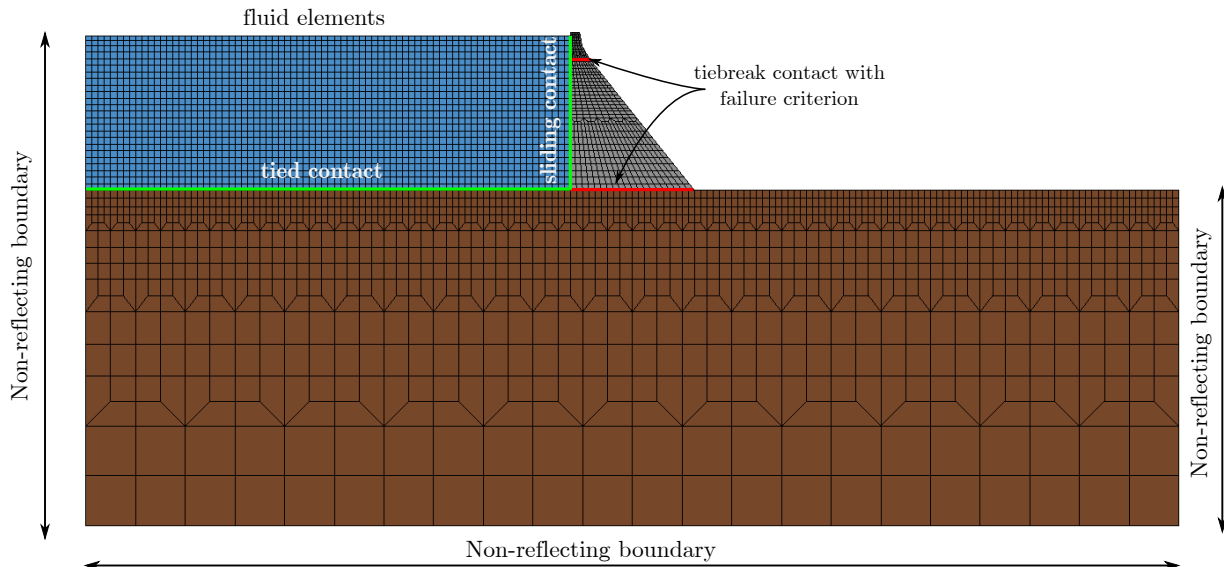
Figure 3.1 Dam-reservoir-foundation LS-Dyna model

surfaces between the dam system components. Some elements such as drainage galleries
are not taken into account since their influences on the global behavior of the structure
are not significant.

Part of the foundation is modeled to account for its inertia, flexibility and damping given
that the structure-foundation interaction elongates the periods of vibration and provides
an additional energy dissipation mechanism. The extent of the foundation is defined
as a function of the block height H (78 m) [37]. Its length in the upstream-downstream
direction is 545 m (6H), and its depth is 165 m (2H). The foundation mesh element sizes are,
in general, controlled by seismic requirements. Consequently, the maximum element size
depends on the foundation material properties and on the frequencies of the structure [37].
There should be no fewer than 10 elements per wavelength, and every attempt should be
made to maintain uniform elements. The wavelength of interest is given by:

$$w_v = \frac{\sqrt{\dfrac{E_{rock}}{2(1 + \nu_{rock})\rho_{rock}}}}{f_0} \tag{3.1}$$

where w_v is the wave length, $E_{rock}, \nu_{rock}, \rho_{rock}$ are the modulus of elasticity, the Pois-
son coefficient and the density of the foundation, respectively, and f_0 is the fundamen-
tal frequency of the structure-reservoir-foundation system. Considering from preliminary
analysis that the system's fundamental frequency is approximately 10 Hz, the maximum
foundation element dimension is fixed at 25 m. To limit the number of elements in the
model, the foundation width is identical to that of the block, and the size of the elements

is progressively increased, as shown in Figure 3.1. The final model of the foundation comprises 9636 elements.

To account for the interaction between the reservoir and the structure, part of the reservoir is modeled with fluid elements having the physical properties of water. The reservoir is modeled with the same length as the foundation, has a constant depth of 76 m and possesses 7488 elements. The compressibility of water is considered in order to adequately model the propagation of pressure waves in the reservoir, whereas the viscous effect is neglected since it had no influence in the fluid-structure interaction. The elements must be sufficiently small (< 15 m) to ensure the adequate transmission of the soil movement through the water [37]. This condition is widely respected in the present model since the maximum dimension of a reservoir element remains less than 3 m. The elements of the reservoir are modeled using a Lagrangian formulation and NULL material in LS-Dyna that has no shear stiffness and no yield strength and behaves in a fluid-like manner. The NULL material is associated with an equation of state, and as recommended by Noble [42], a linear polynomial relation is used, where the pressure P applied to one element is given by:

$$P = K \left(\frac{\rho_w}{\rho_{w,0}} - 1 \right) \tag{3.2}$$

where K is the water bulk modulus equal to 2.18 GPa, ρ_w is the water density at time t during the analysis and $\rho_{w,0}$ is the initial water density.

Boundary conditions and contact interfaces

In this study, as recommended by several studies [136, 37] and detailed below, the loads are applied in two phases: a dynamic relaxation phase for static loads and a dynamic phase for the seismic loads. The boundary conditions are different depending on the loading phase. During the dynamic relaxation phase, a symmetric boundary condition is applied in the bottom and downstream-upstream faces, where the normal displacements are zero to simulate the constraints of the model in the space. In the dynamic phase, symmetric boundary conditions are removed, and the reactions obtained from the dynamic relaxation phase are also applied on the boundary faces for equilibrium. Additionally, non-reflective boundaries are added on the bottom, downstream and upstream faces of the model to take into account the radiation damping and to dissipate the energy trapped in the foundation due to the finite length of the model, i.e., to simulate a semi-infinite behavior. In addition, horizontal and vertical displacements are restricted at the bottom edges of the foundation to prevent rigid body movements of the model during dynamic analyses.

Regarding the contact surfaces between the different components of the system, sliding contact with zero friction was used to model the concrete-reservoir interface. For the reservoir-foundation interface, tied contact was applied, except near the upstream face of the block, where sliding contact with zero friction was used to maintain the reservoir load during the sliding of the block. Preliminary linear analyses, detailed further below, identified the block-foundation interface at the base of the block as high tensile stresses areas and, therefore, where cracking and sliding are likely to occur. Consequently, the model nonlinearity was constrained to this areas only, using tiebreak contact elements with a tension-shear failure criterion. The finite element code assumes the contact surface is broken when the following criterion is satisfied:

$$\left(\frac{|\sigma_n|}{\sigma_t}\right)^2 + \left(\frac{|\sigma_s|}{\sigma_c}\right)^2 \geq 1 \tag{3.3}$$

where σ_n and σ_s are the normal and shear stress, respectively, σ_t is the tensile strength and σ_c is the cohesion. As long as the failure criterion is not reached, the tensile-compressive and shear stresses are transmitted at the contact level. If the normal stress is in compression, the first term of Eq. (3.3) is not considered, and only cohesion is taken into account. In addition, no resistance is mobilized by friction before breaking the contact. Once the failure criterion is reached, only the friction contributes to the resistance of the contact.

Damping and hourglassing Control

Energy dissipation within a complex dam-reservoir-foundation model may occur from a number of mechanisms. One major source is the use of non-reflecting boundary conditions; otherwise, too much energy could be trapped within the foundation domain. Another energy dissipation mechanism comes from the interaction of the dam with the reservoir through the generation of elastic waves in the reservoir. This phenomenon is known as radiation damping, which is considered in the model by also using a non-reflecting boundary condition in the upstream face of the reservoir. Regarding the block structure, to consider structural damping and potential nonlinear behavior, a viscous damping is associated with the concrete material. For the foundation, and as recommended by the USBR [37], a viscous damping of $\xi = 5\,\%$ was also associated with the rock material.

The largest disadvantage of single-point integration elements is the need to control the zero energy modes that can arise, known as hourglassing modes [54]. The hourglass control method used in this study is the stiffness method, where a small elastic stiffness capable of stopping the formation of the anomalous modes but having a negligible effect on the stable global modes is added to the model. The hourglass modes control methods

must be calibrated so that the energy associated with these modes remains less than 10 % of the internal energy of the system and optimally less than 5 % [133]. By setting the hourglassing parameters according to the recommendations of the USBR [133] in LS-Dyna, these parasitic modes are eliminated from the model, and the hourglassing energy criterion is always respected.

Calibration of the numerical model

To evaluate the accuracy of the FEM, the dynamic properties of the DRF system were compared to the results obtained from in-situ dynamic experimental results. Details on dynamic testing and treatment of results are available in Proulx and Paultre [22]. The validation of the dynamic characteristics were based on the fundamental period of the system and global damping. Because of the type of element used to model the reservoir in the LS-Dyna model, a modal analysis cannot be performed; therefore, a free vibration test is simulated to estimate the fundamental period and the damping of the system. Considering gravity loads only, a force is applied at the crest of the block in the upstream-downstream direction and then suddenly withdrawn. The recorded horizontal displacement time series of a node at the crest of the block is then used to estimate the fundamental period of the system through its Fourier spectrum. The global damping is approximated using the logarithmic decrement:

$$\xi = \frac{1}{2\pi m} \ln\left(\frac{u_n}{u_{n+m}}\right) \tag{3.4}$$

where u_n is the displacement measured at a given time and u_{n+m} is the displacement measured m cycles later. The damping was calculated as the average value using the first representative peak (peak n) and m numbers of cycles afterward.

With the material properties presented in Table 3.1, a fundamental frequency of 4.0 Hz is obtained and compared to the frequency of 4.85 Hz measured by Proulx and Paultre [22] during in situ test. The 20 % difference with the experimental value is explained by the fact that only the highest block of the dam is modeled. Further details in the calibration of the FEM can be found in Bernier et al. [20].

Static and dynamic loads

The loads considered in this study are summarized in Table 3.2. Loads resulting from ice thrust, sediment deposit and thermal gradient are neglected, considering they would not affect the overall behavior of the system under seismic actions. Only one loading case is considered to analyze the seismic response of the case study structure, which includes the self-weight of the block, the hydrostatic and hydrodynamic loads exerted by the reservoir

Table 3.1 Material properties - calibrated model

Parameter	Concrete	Foundation	Water
Density (kg/m^3)	2400	2700	1000
Modulus of elasticity (GPa)	32.0	75.0	0.7×10^{-3}
Bulk modulus (GPa)	-	-	2.18
Poisson coefficient	0.2	0.33	0.49995
Damping (%)	1.5	5.0	-

on the block, the uplift pressures at the concrete-rock contact and the horizontal and vertical seismic loads.

Table 3.2 Load conditions

Load	Details
Self-weight	$\rho_{concrete} = 2400$kg/m^3
Reservoir weight	Reservoir level = MOL
Upstream hydrostatic thrust	Reservoir level = MOL
	Concrete-rock contact
Uplift	Reservoir level = MOL
	Drain efficiency: between 0 % and 67 %
Hydrodynamic load	Automatic with fluid elements
	Reservoir level = MOL
Horizontal seismic load	Representative GMTS
Vertical seismic load	Representative GMTS

MOL: Maximum operating level.

The weight of the block is integrated in all the analyses, and it is determined directly by LS-Dyna as well as the reservoir and the hydrostatic load. For all the analyses, the maximum operating level (MOL) of the reservoir is considered, i.e., a constant reservoir elevation of 76 m with respect to the foundation. Even if the foundation is associated with a mass, its self-weight is not taken into account in the application of gravity loads to avoid settlement during the analysis. Concerning the uplift loads, they are only applied at the concrete-rock contact, and they are calculated at each node of the structure base according to the prescribed expressions by the USACE [137]. The dynamic loads considered in the analysis comprise the seismic load in the horizontal (upstream-downstream) and vertical direction. In addition, the model of the reservoir with fluid elements makes it possible to automatically take into account the hydrodynamic thrust due to the seismic shocks. The dynamic effects on the uplift distribution are neglected, assuming that the uplift profile remains constant during seismic loading [137, 37].

Application of static and dynamic loads

The application of gravity and static loads in an explicit analysis can produce unexpected results if the load is applied too quickly because the solution may diverge. To avoid this problem, as previously mentioned, the static loads are applied during a dynamic relaxation phase, and the seismic loads are applied in a dynamic analysis phase. The dynamic relaxation then reduces the nodal velocity at each time step, which is equivalent to a highly damped dynamic analysis [54]. Within the context of this study, the application of the gravity and statics loads takes a total duration of $8\,$s. The self-weight of the dam is applied progressively from 0–$4\,$s, whereas the weight of the reservoir and uplift pressures are applied gradually, after the self-weight, between 4–$8\,$s. At the end of the dynamic relaxation phase, the reaction forces at each node that belong to the foundation faces were recorded to be used in the next phase to maintain a quasi-static state of the model.

Given the presence of an absorbing boundary condition during the dynamic phase, seismic loading cannot be applied at the foundation base in the form of accelerograms [37], and it needs to be applied as a stress time series using the following expressions:

$$\sigma_H = 2\rho_{rock}v_H\sqrt{\frac{E_{rock}}{2\rho_{rock}(1+\nu_{rock})}} \tag{3.5}$$

$$\sigma_V = 2\rho_{rock}v_V\sqrt{\frac{E_{rock}(1-\nu_{rock})}{\rho_{rock}(1+\nu_{rock})(1-2\nu_{rock})}} \tag{3.6}$$

where σ_H is the horizontal stress time series, ρ_{rock} is the foundation density, v_H is the horizontal ground velocity (obtained by integration of the horizontal acceleration), σ_V is the vertical stress time series, v_V is the vertical ground velocity (obtained by integration of the vertical acceleration) and E_{rock} and ν_{rock} are the modulus of elasticity and the Poisson coefficient of the foundation, respectively.

Deconvolution of the ground motions

Given that the foundation mass and inertia effect are considered in the analysis, ground motions need to be deconvolved. Deconvolution methods allow for the adjustment of the amplitude and frequency contents of a seismic ground motion applied at the base of the foundation to achieve the desired target acceleration time series at the structure-foundation interface. Following the method proposed by Sooch and Bagchi [57] and presented in Figure 3.2, the ground motion is initially applied at the base of the foundation (IH), and it is assumed to be the same as the free-field ground acceleration (TH). The acceleration

time series at the top surface (OH) (i.e., block-foundation interface) is then estimated by solving the wave propagation problem using the finite element model. This estimated or reproduced ground motion at a reference point on the block-foundation interface (FO) is then compared to the original free-field ground motion (FT) after transforming both signals into the frequency domain using Fourier analyses. The synthesized and free-field signals at the top of the foundation are then compared in the frequency domain, and a correction factor for each frequency is computed (CF). The modified ground motion (FMI) is then transformed back into the time domain (MIH), and the wave propagation analysis for the foundation system is repeated iteratively with the modified ground motion applied at the base of the foundation. This procedure is repeated until the difference between the spectra of the output (MOH) and the target ground motion (TH) is less than $5\,\%$ over the range of periods of interest $(0.2T_1 - 2T_1)$ and $10\,\%$ elsewhere. As shown in Figure 3.3, the comparison of the acceleration response spectra further confirms the effectiveness of the deconvolution.
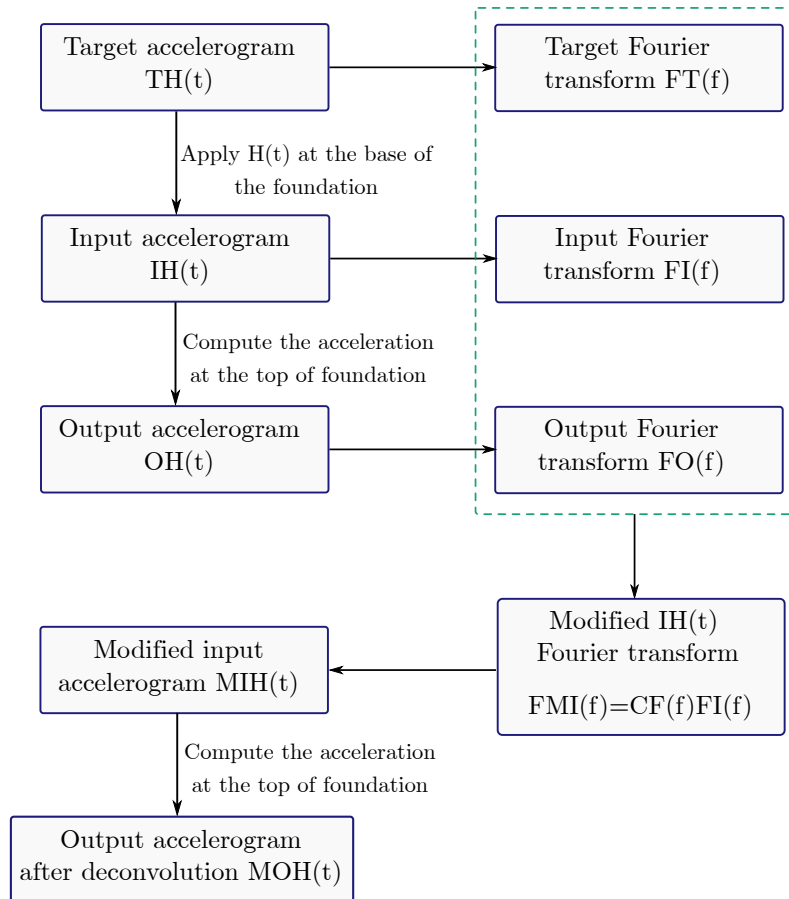
Figure 3.2   Deconvolution methodology
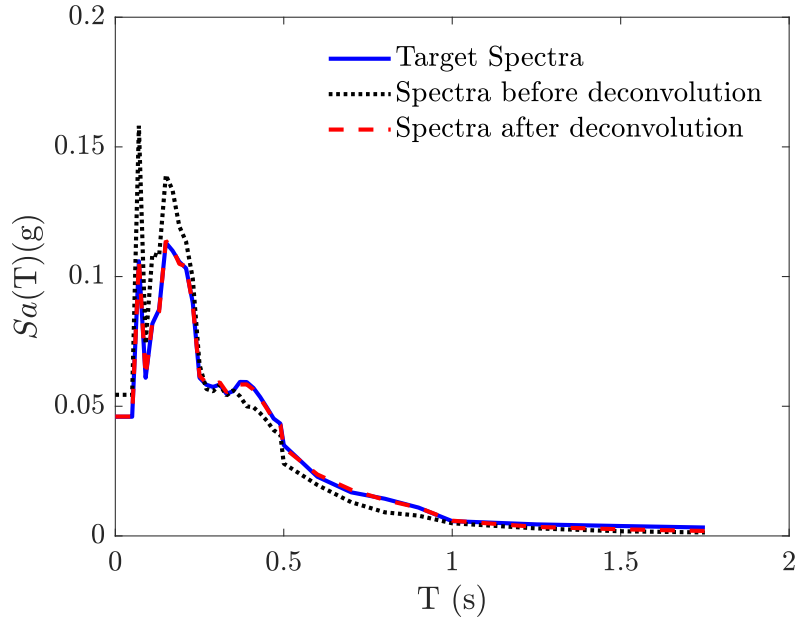
Figure 3.3 Comparison between response spectrum (before and after deconvolution)

3.1.2 Uncertainty modeling

Each parameter of the analysis is defined either as a fixed value or as a random variable associated with a probability density function. The preliminary analyses carried out in the finite element model showed that some parameters (e.g., concrete damping) will have to be defined as random variables to take into account the uncertainty about the estimated value. The lack of knowledge relative to the parameters quantifying the resistance of the structure, loading conditions and seismic input are considered in this study. All the parameters considered as random variables and their associated PDFs are listed in Table 4.3, while all the material properties considered constant throughout the analysis are summarized in Table 3.1, except for concrete damping. The modulus of elasticity of concrete and rock are not considered as uncertain variables since these values were determined during the calibration of the numerical model. All other modeling parameters not listed are kept constant using their best estimate. The 8 uncertain parameters listed in Table 4.3 were sampled using Latin hypercube sampling to generate $n_f = 250$ model samples to train the meta-models and to propagate sources of uncertainty in the fragility analysis. LHS sampling method was chosen because of its ability to efficiently generate representative samples by dividing the range of possible values for each variable into N equiprobable intervals. Since no material investigations are available for the case study dam, the probability distributions of material properties are defined from the literature and empirical data of similar dams. The uniform distribution is used for all parameters,

except damping, since it is the maximum entropy probability distribution when data is limited to an upper and lower bound.

Concrete damping

The concrete damping is difficult to estimate, especially since the possible cracking of the structure under the effect of the seismic load will cause additional energy dissipation. Given that the concrete is associated with an elastic linear behavior in the numerical model, the increase in damping with cracking is not automatically taken into account. The average damping should therefore be increased with respect to the $1.5\,\%$ value that is obtained from an experimental test for relatively small vibrations [22]. As proposed by Ghanaat et al. [85], for the concrete damping, a log-normal distribution with a median of $5\,\%$ and a standard deviation of 0.35 was adopted. Taking into account these values, the total system damping remains under $7\,\%$ [138].

Shear and tensile strength

In the absence of sufficient data, a uniform PDF was also used to describe the parameters characterizing the shear and the tensile strength. For determining the shear and tensile strength parameters at the rock-concrete and concrete-concrete contact and based on the construction plan of the dam, the foundation rock at the dam site was determined to be composed of diorite, gneiss and granite. The maximum and minimum values for the cohesion and the angle of friction were estimated from literature test results for the same type of rock foundation and concrte properties [139, 7].

Drain efficiency

The efficiency of the drainage system is defined by the proportion of the flow that can be captured by this system to reduce the hydraulic load. To account for the uncertainties in the uplift pressures, the drain efficiency is defined as a random variable with a uniform PDF. The USBR [37] recommends limiting the effectiveness of the drains to $67\,\%$. Consequently, the efficiency of the drain system was sampled considering values between 0 and $67\,\%$ to include also the undrained case.

3.2   Characterization of the seismic scenario

Classical PSHA allows calculating the probabilities of exceeding, at least once in a given time span, and at a given site, a set of ground motion parameter levels considering all possible earthquake ruptures defined in a seismic source model. Such a list of probability

values is usually referred to as hazard curve. $P_f(\text{IM} \geq im \mid t)$ indicates the probability that a ground-motion parameter, IM, exceeds at least once in a time span t, a level im. To calculate the frequency of exceedence three major elements are needed:

1. Characterisation of seismic sources.

2. Characterisation of attenuation of ground motion or ground motion models (GMM).

3. Calculation of probabilities

Regarding the definition of the source model, according the Open file 7576 [140] produced by the Geological Survey of Canada (GSC), the 5th generation model comprises four components; one for each quadrant of Canada. The subdivision of the national model into four components was necessary to reduce computation time. These four components each comprise multiple weighted sub-models. The sub-models are implemented with the following weighting:

– The northwestern and southwestern models comprise a single sub-model, weighted at 1.0;
– The northeastern model comprises two sub-models: Historical (H2) weighted at 0.6 and Regional (R2) weighted at 0.4;
– The southeastern model comprises three sub-models: H2 weighted at 0.4, Hybrid (HY) weighted at 0.4 and R2 weighted at 0.2.

Concerning the characterization of GMM, they are provided in the form of look-up tables (based on Atkinson and Adams [67]) to be used with the respective sources. The tables have also been modified to provide hazard values directly in terms of Soil Class C ($VS30 = 450\,\text{m/s}$), which is the reference ground condition for NBCC 2015 [65].

The calculation of probabilities to obtain the hazard curve and the disaggregation results at the dam site were performed with the open source software OpenQuake [141]. Further details with respect on the methodology for the characterization of the seismic scenario are given in section 4.3.1.

3.2.1 Ground motion selection method

The results of the PSHA were included in a ground motion selection method to define a representative suite of GMTS likely to occur at the dam site. Moreover it was intended to include in the selection method all relevant seismic IMs for the fragility analysis. As a result the GCIM procedure [74] was chosen to compute the target distribution of the IMs to select ground motion records. Additional explanation is given in section 4.4.2.

3.2.2   Selection procedure

The major steps of the proposed ground motion selection process are based on the work done by Baker and Lee [71] and they are illustrated in Figure 3.4. The process starts by specifying a target response spectrum. Formally, the target mean values of spectral acceleration and the covariance matrix are specified. Eq. (4.3) and Eq. (4.4) were used to compute the target distribution with the GCIM method. The next step is to statistically simulate realizations of response spectra from the target distribution. This is done by sampling from a multivariate log-normal distribution with the target mean and covariance matrices. Since this simulation step is extremely fast, it is performed multiple times and the set of simulations best matching the target spectrum is utilized for the following steps. Next is the specification of candidate ground motions to select from. Relevant meta-data from a candidate ground motion database is loaded, including horizontal and vertical spectral acceleration values, PGA, PGV and rupture parameters for each ground motion. Once the database meta-data has been loaded, a pre-screening must be performed so that only appropriate ground motions are considered for selection. The pre-screening considers appropriate values of earthquake magnitude, source-to-site distance, and $VS30$. These so-called causal parameters are important to screen in order to assure that the considered time series are reasonably consistent with the conditions of interest in ground motion selection, but they should not be screened so aggressively that an insufficient number of candidate motions remain for the next stage of selection [142]. The following step involves selecting a suite of ground motions from the database that best match the conditional mean and standard deviation of the statistically simulated spectra. The selected suite of motions is evaluated to see whether it is sufficiently close to the target distribution. The maximum percentage mismatch of the mean and standard deviation of the selected motions' spectra, relative to their targets are calculated. If the errors are too large, then a finite number of optimization rounds are performed to further improve the selection. Finally, the greedy optimization involves further optimizing the initial selection if needed. At this stage, the selected set of ground motions are modified by replacing individual ground motions from the set with available motions from the screened database and seeing whether the set is improved in its match to the target response spectrum.

3.3   Regression meta-models generation

Following traditional strategies specific for computer simulations [103], in this project meta-models were developed for approximating the seismic response of the dam using the three steps outlined in Figure 5.1. The subsequent sections will detail such steps.
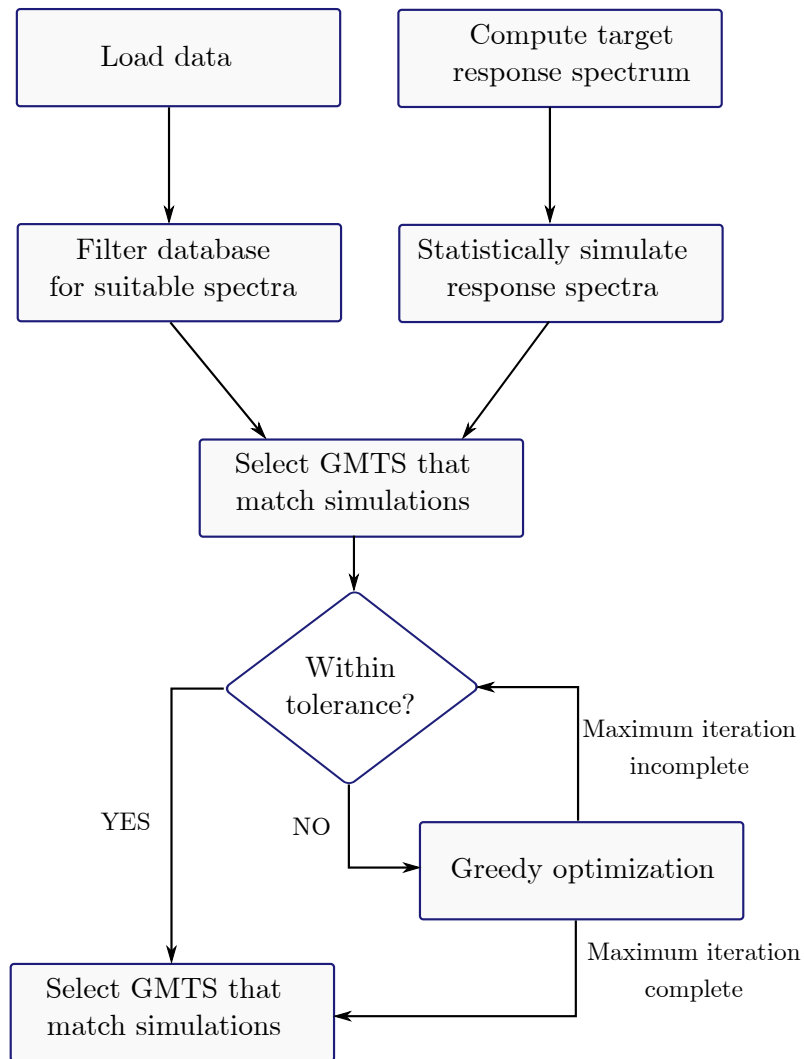
Figure 3.4   Ground motion record selection procedure

3.3.1   Experimental design matrix

Latin Hypercube design of experiments has been widely adopted for deterministic computer experiments [131], wherein a particular combination of dam modeling parameters coupled with a ground motion record will yield the seismic response. The Latin Hypercube design divides the desired range for each element within the parameter vector p into n_f intervals of equal marginal probability $1/n_f$ and then selects a sample once from each interval. The selected n_f samples for the first factor (p_1) are combined with the n_f samples of the second factor (p_2), and subsequent factors $(p_3 \ldots p_m)$ such that it maximizes the minimum distance between the design points. Each row of this Latin Hypercube experimental design matrix, $\mathbf{X}$ is then paired with a suite of ground motions with varying intensity measures. Hence, the dimensions of the original Latin Hypercube experimen-

tal design matrix are $m \times n_f$. Besides the model parameters of Table 4.3, the seismic IMs listed in Table 3.3 were considered in the experimental design matrix. Given that 250 GMTS were selected, the original dimensions of the experimental design matrix were $\mathbf{X}[19 \times 250]$ to which mathematical transformations (natural logarithm, exponential form, etc.) and parameter product combinations were added.

Table 3.3   Seismic IMs considered in the experimental design matrix

Seismic intensity parameters	
$Sa_H(T_1)$	spectral acceleration at the T_1
$Sv_H(T_1)$	spectral velocity at the T_1
PGA	peak ground acceleration
PGV	peak ground velocity
PGD	peak ground displacement
SI	spectrum intensity
ω_{eqk}	earthquake angular frequency
D_{595}	significant duration
I_a	Arias intensity
PGA_V	vertical peak ground acceleration
$Sa_V(T_1)$	vertical spectral acceleration at the T_1

3.3.2   Finite element simulations

For each row, $\mathbf{x}_i$, of the experimental design matrix, $n_f = 250$ finite element simulations were conducted, following the procedure presented in section 3.1.1. The structural response, y_i, considered herein was the maximum relative sliding, $\delta_{\max}$. Hence a response vector $\mathbf{y}[250 \times 1]$ was produced as a result of the FEM simulation, to be used to train different regression techniques.

3.3.3   Meta-model fitting

The regression techniques described in section 5.2.2 were fitted to the training points generated as a result of the finite element simulations considering only the best demand and capacity predictors. In order to select the best predictors, the stepwise regression algorithm in Matlab was used when fitting the meta-models. The algorithm starts with a constant term to predict the response. In the next step, one predictor is added to the model, and the performance of the model is evaluated based on the Bayesian information criterion (BIC). If the model performance improves, the added term is kept; otherwise, it is removed, and this process is repeated until all the proposed predictors are tested. Once all the regression meta-models have been trained, the best performing one is selected in

terms of predictive capacities, based on cross-validation (CV) goodness-of-fit estimates. Further details on the meta-model fitting are given in section 5.3.5 and section 5.2.3. In the following sections, details regarding the implementation of all the considered meta-modeling techniques in this project, are given.

Polynomial response surface

The polynomial response surface methodology in this study is employed to predict the maximum horizontal relative displacement at the base of the dam to quantify the sliding limit state. Traditionally, lower order polynomial (up to second order) are used as meta-models in reliability problems, yet these functions may not be able to capture highly non-linear behaviour. Regarding, higher order polynomials, even if these functions have higher predictive accuracy, they tend to over fit the data. Polynomials up to order 4th were implemented in this study to avoid over fitting issues. The implementation and results of this technique is further detailed in section 5.3.5.

Adaptive basis function construction

The adaptive basis function construction methodology was also employed to predict the maximum horizontal relative displacement at the base of the dam. The software variReg [143] implemented through Matlab was used for this purpose. The performance of this meta-models is shown in Table 5.3 and it can be seen that in terms of local and global goodness-of-fit the model predict with a faire accuracy.

Multivariate adaptive regressive splines

The multivariate adaptive regressive splines methodology were used to replace the output of the numerical model of the dam for the sliding limit state. Two MARS models were trained, one with cubic splines and the other with linear splines. The toolbox for Matlab developed by Jekabsons [144] was used for this purpose. The performance of these meta-models is addressed in Table 5.3 and it can be seen that in terms of local goodness-of-fit the model with linear splines is better while if we consider the global performance indicators the model with cubic splines is more adequate.

Radial basis functions

As it was already mentioned above the RBF method implements basis functions whose response monotonically changes as distance from the central point increases. In the present study, multi-quadratic, thin plate spline and Gaussian are the radial basis functions considered for interpolation. As in the ABFC case, the RBF meta-model was implemented with

the software variReg [145]. The performance of RBF meta-models is shown in Table 5.3. Only 5-CV values are displayed in the table, because the radial basis passes through all the training data, the prediction with the latter is trivial. It can be observed that the RBF training model with multi-quadratic basis functions present the best performance between all the RBF models.

Support vector machine for regression

A support vector machine meta-model for regression was trained using the statistics and machine learning toolbox in Matlab [146]. Four different types of kernels (linear, quadratic, cubic and radial basis functions) were used and their performance is shown in Table 5.3. For the radial basis functions and the cubic kernel, it can be noticed a clear over-fitting of the data given that the meta-models predict with faire accuracy for the training data, but fail to predict for the unseen cases. Besides, almost all the samples in the dataset were used as support vectors reaffirming the over-fitting. Although the local and global performance of the SMVR with linear and quadratic kernels is reasonable, it should be mentioned that approximately 72% of the samples in the dataset were used as support vectors in the multidimensional feature space to find hyperplane that separates all given samples. This is the result of a highly non-linear feature space and could not be appropriate to train this type of models with small datasets (< 300 samples).

Random forest for regression

An assembling technique in terms of random forest regression was performed within 100 decision trees and was also implemented with the statistics and machine learning toolbox in Matlab [146]. The final model adds a sequence of base models and predicts the response by combining the decision trees. Table 5.3 shows that the RFR model overfits the data in the training set, given the lower performance of the model in the cross-validation stage.

3.4   Seismic fragility analysis

Estimation of fragility functions using dynamic structural analysis is an important step in a number of seismic assessment procedures. The first step is the identification of the limit states that are relevant to the system performance, for which a specific fragility function is developed, taking into account a rational assessment of all sources of uncertainty likely to affect it.

3.4.1   Damage states

As is was mentioned in section 2.1.1, when subjected to strong ground motion, gravity dams may be damaged in different ways. In recent years, typical damage modes that could lead to the potential collapse of dams after a seismic event have been identified, and seismic damage levels have been established. Preliminary analyses have identified sliding as the critical failure mode for the case study dam [20], and other failure modes would only occur after sliding has already been observed. As a result, two limit states were considered in this study:

1. Concrete-to-rock sliding at the base of the dam interface.

2. Concrete-to-concrete joint sliding at the neck of the dam interface.

Each limit state was characterized by the sliding damage states presented in Table 4.2.

3.4.2   Single-variate fragility functions

There are a number of procedures for performing non-linear dynamic structural analyses to collect the data for estimating a fragility function. One common approach is incremental dynamic analysis, where a suite of ground motions are repeatedly scaled in order to find the IM level at which each ground motion causes collapse [77]. A second common approach is multiple stripes analysis (MSA), where analysis is performed at a specified set of IM levels, each of which has a unique ground motion set [68]. The latter is the procedure selected in this study to generate fragility functions, curves and surfaces, from non-linear FEM simulations and from the meta-models respectively.

The MSA approach is common when using the GCIM or other approaches to select ground motions representative of a specific site and IM level, because the target properties of the ground motions change at each IM level and thus so do the representative ground motions [74, 75]. Hence, with this approach different ground motions are used at each IM level and the analysis need not be performed up to IM amplitudes where all ground motions cause collapse [68].

Non-linear dynamic analyses were performed for the 30 samples at each intensity level $(Sa_H(T_1) = 0.1\,\text{g}; 0.2\,\text{g}; 0.3\,\text{g}; 0.4\,\text{g}; 0.5\,\text{g}; 0.7\,\text{g}; 0.9\,\text{g})$, resulting in a total of 210 analyses. From each simulation, the maximum relative base and neck displacement is computed. The results from the non-linear analysis were used to calculate 7 fragility point estimates as the number of samples where sliding exceeded the limit state divided by the total number of samples. A well-known mathematical equation was then fitted to these fragility estimates to develop fragility curves for the considered limit states. Several mathemati-

cal formulations, such as the normal, log-normal and Weibull CDF, were fitted to point estimates of the fragility curves. Nevertheless, the log-normal cumulative distribution function, together with the MLE fitting technique, was found to be adequate for depicting the fragility of the considered limit states. Further details on the fragility curves generation are given in section 4.5.3.

3.4.3   Multivariate fragility functions

Given the number of simulations required to generate a fragility surface, instead of non-linear dynamic analysis the results from the meta-models were used to generate the fragility point estimates. Regarding the generation of the samples where the meta-model will be evaluated to predict the dam's response, independence between all the modeling parameters was considered to generate 5×10^5 samples with LHS. The values of the seismic IMs were bounded, as shown in Table 3.4, to efficiently cover the range of values corresponding to return periods from 500–30000 years.

Table 3.4   Seismic IMs range of values

Intensity measures	Range of values
Peak ground velocity (PGV)	0.80–25.0 cm/s
Arias intensity (I_a)	0.0–2.5 m/s
Vertical peak ground acceleration (PGA$_V$)	0.01–0.25 g

To perform the fragility analysis and obtain the fragility point estimates from the meta-models, the methodology depicted in section 5.2.4 and Figure 5.2 was used. The range of each of the parameters of the fragility surface was divided in 100 intervals. As a result, 10^4 fragility point estimates were generated for each limit state, as shown in Figure 3.5.

Two different cases were taken into account to generate fragility surfaces. In the fist case, as it is expected, correlation between the seismic IMs was assumed to generate the samples where the meta-model was evaluated. As a result, fragility surfaces as a function of PGV and each of the MP involved in the prediction of the dam response were generated. In the second case, although less realistic, independence between the seismic IMs was considered to generate the meta-model's samples. Accordingly, fragility surfaces as a function of two seismic IMs were generated to consider more general cases where the seismic scenario is not well defined.
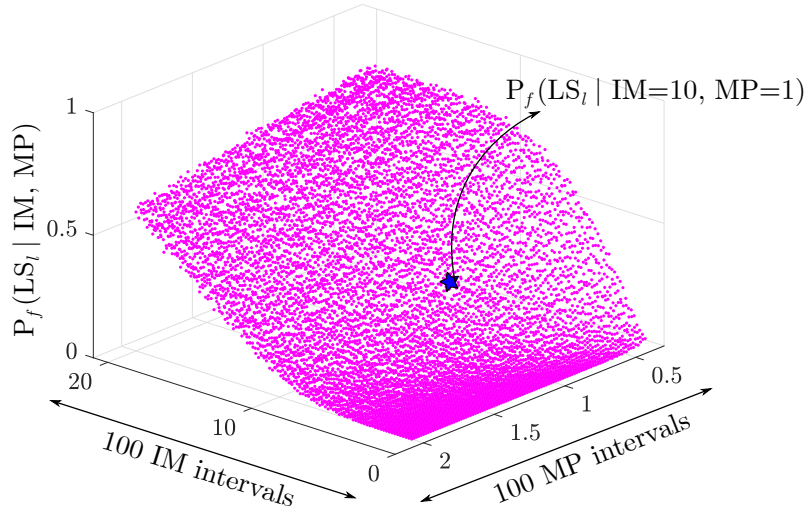
Figure 3.5   Conceptual fragility point estimates from the meta-model MSA

Correlated seismic IMs

For the case study dam, considering the PSHA and the characteristics of the seismic scenario for a specific site, correlation between seismic parameters is to be expected. In order to consider this, the samples of these parameters must be sampled from a multivariate distribution with their respective correlation coefficients, when referring to specific cases. Section 5.3.6, describes the approach that was used to generate point estimates of the fragility surfaces and section 5.2.4 explains the procedure used to create parametric fragility surfaces.

Uncorrelated seismic IMs

In this section, to consider more general cases, the correlation between the seismic IMs was neglected. The fragility point estimates were generated as a function two seismic IMs, IM_1 and IM_2. The uncertainty due to all other parameters involved in the meta-model response is propagated in the analysis by sampling these parameters from their respective PDFs.

For this specific case, the methodology proposed by Baker [68] for the efficient fitting of fragility curves, was extrapolated to fit fragility surfaces. It was assumed that for each limit state, the fragility surface is described by the product of two independent cumulative density functions, as shown in Eq. (3.7):

$$F_s(IM_1, IM_2) = \Phi_{IM_1}(IM_1, \theta_{IM_1}, \beta_{IM_1})\Phi_{IM_2}(IM_2, \theta_{IM_2}, \beta_{IM_2}) \tag{3.7}$$

where $\Phi_{\mathrm{IM}_{1,2}}$ are some form of CDFs and $\theta_{\mathrm{IM}_{1,2}}$ and $\beta_{\mathrm{IM}_{1,2}}$ are the parameters characterizing the respective CDFs. The parameters of each fragility surface were estimated with the MLE method as proposed by Baker [68], when using the MSA for efficient fragility functions' fitting.

3.4.4   Seismic assessment of dams

Expected seismic performance

For structures of special importance, such as dams and nuclear power plants, it is typical to consider the high confidence low probability of failure (HCLPF) criteria. This notion represents the level for which the probability of exceedance of a limit state is sufficiently low to consider that it will not be reached throughout the life of the structure [147] and usually corresponds to a probability of less than 5–6 %. In Canada, the committee of the Canadian Dam Association (CDA) has the vision of being the authoritative source for dam safety practices, and its current design philosophy provides the minimum standards for life safety [31]. Accordingly, to evaluate the seismic performance of the case study dam under extreme limit states, the return period boundaries for the hazard classification provided by the CDA (Table 4.4) for the maximum credible earthquake (MCE), were used. The probabilities of exceedance calculated for the "extreme risk" case [31] were compared with the values proposed by the ASCE 7-16 [148] guidelines for the maximum probability of failure for the MCE. For a risk category III (dam-type structures) and total or partial structural collapse, the maximum probability of failure should be less than 6 %.

In the case of single variate fragility functions, for a given probability of exceedance, prescribed by the safety guidelines, the corresponding seismic IM was extracted from the hazard curve. The fragility estimates of the system for the specified return periods were then calculated from the fragility curves corresponding to the 100 mm and 150 mm damage states for the neck and the base sliding limit states, respectively. Similarly, in the case of the fragility surfaces as a function of IM and MP, to achieve a desired seismic performance, boundaries of model parameters for an adequate performance under extreme limit states can be formulated. As it is the case for fragility curves, the corresponding seismic IM was extracted from the hazard curve and all possible MP values for that IM are established by cutting a slice of the fragility surface along that seismic IM. By ensuring that the probability of exceedance given that an extreme event is in line with the current guidelines for the minimum provisions for life safety, a usable range of MP values can be determined.

Unconditional probability of failure

A design to a specific scenario event or a specific return period ground motion intensity does not address the seismic vulnerability of the resulting structure. Any structure is potentially exposed, throughout its design life, to all the possibilities of the occurrence of ground motion intensities at a given site as characterized by a set of site-specific seismic hazard curves (mean, median, and several specified exceedance percentiles). The use of this complete hazard information is prerequisite to estimating seismic "failure" probabilities. "Failure" is a generic term defining non-performance at a preselected limit state. In order that all pairs of ground motion intensities and associated exceedance probabilities are considered, the site hazard curves are substituted by the mean curve. First, the conditional failure probability curve (fragility curve) is calculated by the convolution of the load and resistance distribution functions. To calculate the unconditional failure probability, each conditional failure probability is multiplied by the corresponding probability density functions of the mean hazard curve, and then summed over all possible values of the hazard curve.

As shown in Figure 4.2, seismic hazard is usually characterized by curves relating the exceedance probability of a ground motion parameter to the intensity level of the ground motion variable. Given the significant uncertainties in ground motion estimates, a proper probabilistic seismic hazard evaluation includes the variability of the selected ground motion parameter at all exceedance probabilities. This uncertainty is characterized in Figure 4.2 (b) and Figure 4.6 (b) by several hazard curves for different source models and earthquake rupture forecast configurations. A hazard curve, $\bar{V}(Sa(T_1))$, is the complementary cumulative distribution function of the selected ground motion parameter, $Sa(T_1)$. The CDF of the hazard, $V(Sa(T_1))$, is simply $[1 - \bar{V}(Sa(T_1))]$. Thus, the PDF of the considered seismic hazard curve is given by,

$$v(Sa(T_1)) = \frac{\mathrm{d}V(Sa(T_1))}{\mathrm{d}Sa(T_1)} = \frac{\mathrm{d}[1 - \bar{V}(Sa(T_1))]}{\mathrm{d}Sa(T_1)} = -\frac{\mathrm{d}\bar{V}(Sa(T_1))}{\mathrm{d}Sa(T_1)} \qquad (3.8)$$

Then, Eq. (3.8) is used together with Eq. (4.9) to calculate the unconditional probability of exceedence. The unconditional probabilities for the most damaging limit states were compared with the annual probabilities of failure found in the literature. MacGregor [149] recommended that the collapse probability of a building-type structure should be on the order of magnitude of 5×10^{-5} and even lower for structures of special importance, such as dams and nuclear power plants. In the same manner, following the risk category framework proposed by the ASCE 7-16 [148], the number of people at risk is between 100 and 1000.

For this range, the FEMA P-1025 [150] guidelines specify a value between 1×10^{-5} and 1×10^{-6} for tolerable risk. The results obtained for the case study dam, are shown in section 4.6.2.

CHAPTER 4

ON THE SEISMIC FRAGILITY ASSESSMENT OF CONCRETE GRAVITY DAMS IN EASTERN CANADA

Avant-propos

Auteurs et affiliation :

R. L. Segura : étudiante au doctorat, Université de Sherbrooke, Faculté de génie, Département de génie civil et de génie du bâtiment.

C. Bernier : assistant de recherche gradué, Rice University, Department of Civil and Environmental Engineering.

R. Monteiro : professeur, Istituto Universitario di Studi Superiori di Pavia.

P. Paultre : professeur, Université de Sherbrooke, Faculté de génie, Département de génie civil et de génie du bâtiment.

Contribution au document : Cet article contribue au document présenté en améliorant la méthodologie de sélection des accélérogrammes pour effectuer une analyse de fragilité sismique, avec l'inclusion des paramètres les plus pertinents qui sont fortement corrélés avec les états limites proposés pour les barrages poids. En outre, une comparaison entre les résultats de l'analyse du modèle probabiliste de risque est effectuée à l'aide des modèles de risque 2010 et 2015 et de leur effet sur les fonctions de fragilité et les fonctions la performance sismique attendue du barrage en cas d'événements extrêmes.

Résumé français : Ces dernières années des méthodes probabilistes, telles que l'analyse de la fragilité, sont devenues des outils fiables pour l'évaluation sismique des structures de type barrage. Ces méthodes nécessitent la sélection d'une série représentative d' accélérogrammes, d'où le besoin d'une méthode de sélection d'accélérogrammes intégrant

tous les paramètres de mouvement du sol pertinents dans l'analyse de la fragilité de ce type de structure. Cet article présente l'élaboration de courbes de fragilité, pour les états limites de glissement de barrages-poids dans l'Est du Canada, actualisées à l'aide d'une méthode de sélection des accélérogrammes fondée sur la méthode de la mesure d'intensité conditionnelle généralisée. Ces courbes de fragilité sont ensuite combinées aux données de risque régionales récemment développées pour évaluer le risque annuel, qui est mesuré en termes de probabilité inconditionnelle de dépassement des états limites. La méthodologie proposée est appliquée à un barrage situé dans le nord-est du Canada, dont la fragilité est évaluée par comparaison avec des études antérieures et les directives de sécurité actuelles. Il a été observé que la méthode, plus précise, proposée dans le document produisait des estimations de la fragilité moins conservatrices pour le barrage étudié.

Abstract:

In recent years, probabilistic methods, such as fragility analysis, have emerged as reliable tools for the seismic assessment of dam-type structures. These methods require the selection of a representative suite of ground motion records, resulting in the need for a ground motion selection method that includes all the relevant ground motion parameters in the fragility analysis of this type of structure. This paper presents the development of up-to-date fragility curves for the sliding limit states of gravity dams in Eastern Canada using a record selection method based on the generalized conditional intensity measure (GCIM) approach. These fragility functions are then combined with the recently developed regional hazard data to evaluate the annual risk, which is measured in terms of the unconditional probability of limit state exceedance. The proposed methodology is applied to a case study dam in north-eastern Canada, whose fragility is assessed through comparison with previous studies and current safety guidelines. It is observed that the more accurate procedure proposed herein produces less conservative fragility estimates for the case study dam.

Keywords: Seismic hazard; dam safety; concrete dams; probabilistic seismic hazard analysis; fragility curves; ground motion record selection.

4.1   Introduction

The consequences of dam failure can be substantial, in terms of both casualties and economic and environmental damage. Therefore, dam safety is given highest priority. With the increasing knowledge of seismicity, a growing number of dams fail to meet the re-

vised safety criteria that incorporate new seismic hazard information. Consequently, the combination ageing and its associated problems with new methods for estimating seismic loads and with the increasing demands of society to ensure a high level of safety has resulted in the need to review and upgrade the methods of seismic analysis for dams. Traditionally, dams were evaluated using deterministic analysis under an extreme event. Nevertheless, in recent decades, probabilistic-based tools, such as fragility functions, have become increasingly popular for the seismic assessment of dams [28, 21, 102, 151, 38, 14]. Fragility analysis, which depicts the uncertainty in the safety margin with respect to specified hazard levels, including the design-basis and review-level events, has proved to be an effective and reliable instrument to support rational risk mitigation decision making and to establish priorities [36, 14]. Based on this, earthquake engineering codes and guidelines have started to shift towards performance-based earthquake engineering (PBEE) design which seeks to improve seismic risk decision-making through assessment and design methods with a strong scientific basis to make informed decisions. To implement this type of design, an accurate estimate of the seismic demand of the structural systems is especially important. Such an estimate requires, in turn, a ground motion record selection technique that properly depicts the seismic scenario and adequately propagates the record-to-record variability and uncertainty related to the seismic hazard throughout the fragility analysis [23, 24]. Generally, the primary considerations in selecting ground motion recordings make use of appropriate seismological properties. However, these seismological parameters alone have proven to be fairly poor predictors of structural demands [68]. With this limitation in mind, a consensus is emerging that it is more productive to consider time series properties rather than seismological parameters when selecting ground motions [69].

Accordingly, in recent years, the implementation of specific target spectra, such as the conditional (mean) spectrum (CS) [73] has been recommended over the frequently employed uniform hazard spectrum (UHS) due to its capability to support the selection of records that match proper ground motion characteristics for a given intensity measure (IM) level [75]. Although improved with respect to the traditional UHS, one limitation of the ground motion selection method based on spectral acceleration matching is that only the characteristics of ground motion represented in terms of (linear) spectral acceleration are considered, whereas it is acknowledged that the severity of ground motion, in general, depends on its intensity, frequency content, and duration [74]. Therefore, a further refined method, the generalized conditional intensity measure (GCIM), was proposed by Bradley [74] for ground motion record selection. Within the GCIM framework, any number of ground-motion IMs identified as relevant for a particular seismic response problem can be considered allowing the estimation of conditional distributions based on the full

distribution of disaggregation results. This method proposes therefore a more realistic target distribution and its application has produced encouraging results by recent studies in different contexts [25, 26, 27].

Currently, no previous studies have conducted seismic assessment of dams while properly considering the contribution of several IMs. Moreover, and referring to the specific context of this study, only a few studies considering the recent CS-based approach for eastern Canada exist [72, 21] and there are no studies that have used the GCIM method for this region. As such, this paper aims to improve the development of fragility curves for concrete gravity dams by using an enhanced procedure to select ground motion records that includes the different intensity measures that are relevant to the sliding limit states of such structures. Specifically, within PBEE of concrete gravity dams, considering their intrinsic structural characteristics, the peak ground velocity (PGV) and the vertical spectral acceleration (Sa_V) have been recognized as important intensity measures to develop probabilistic seismic demand models related to dam base sliding limit states [102, 30]. The GCIM approach, therefore, enables consideration of PGV and Sa_V, in addition to the traditionally employed horizontal spectral acceleration (Sa_H), to improve the selection of the horizontal and vertical components of the records. Such records are then used to seismically assess a concrete gravity dam located in north-eastern Canada.

Fragility curves are developed by performing non-linear dynamic analysis, including fluid-structure-foundation interaction. The produced fragility functions are compared with currently available ones, which were developed with records selected using the standard CS method and an older hazard model. Finally, an expedited safety assessment is performed by computing the probability of occurrence of the ultimate sliding state for the maximum considered earthquake (MCE) and the unconditional probability of occurrence, which are then compared with the current guidelines for the minimum provisions for life safety.

4.2 Case study description and modeling

The present study is focused on a case study concrete gravity dam in Quebec, Canada. It is the largest gravity dam in the province, with 19 unkeyed monoliths, a maximum crest height of 78 m, and a crest length of 300 m (Figure 4.1(a)). The dam rests on a foundation consisting mainly of anorthosite gabbro and granitic gneiss [22], which corresponds to hard rock ($VS30 > 1500$ m/s). The dam was chosen for its simple and almost symmetric geometry and due to the availability of forced vibration test results used to calibrate the dynamic properties of the numerical model [22]. Moreover, previous studies concerning

the seismic fragility of one of the central blocks of the dam [20, 21] are also available, enabling a direct comparison with the results of the method proposed herein.
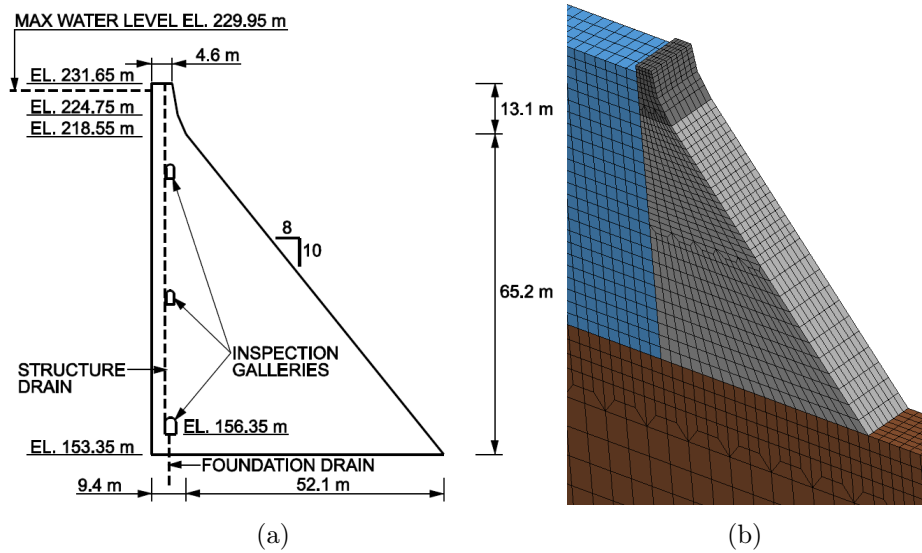
Figure 4.1   (a) Cross section and (b) Finite element model of the tallest mono-lith

The tallest monolith of the dam was selected to represent it and was modeled with computer software LS-Dyna [135], as shown in Figure 4.1 (b), following the recommendations of the United States Bureau of Reclamation (USBR) [37]. An explicit time integration solver [44] was used, within LS-Dyna, due to the contact surfaces and to reduce the computational time. Only one load case combination was considered, which included self-weight, hydrostatic thrust, uplift, hydrodynamic effects and seismic load. The proposed model takes into account the different interactions among the structure, reservoir, and foundation. The reservoir is modeled with compressible fluid elements, whereas the concrete dam and the rock foundation are modeled with linear elastic materials to which a viscous damping is associated. Given that the model should remain stationary after the static loads are applied, two loading phases were considered: (i) a dynamic relaxation phase for static loads, and (ii) a dynamic phase for the seismic loads, each with different boundary conditions. In the first loading phase, a symmetric boundary condition was applied where the normal displacements are zero. For the dynamic phase, non-reflective boundaries were included to prevent artificial amplification of the seismic waves due to the finite length of the foundation and the reservoir. Concerning the contact surfaces between the different components of the system, sliding contact with zero friction was used to model the dam-reservoir interface. For the reservoir-foundation interface, tied contact was applied, except near the upstream face of the dam where sliding contact with zero friction was used to

maintain reservoir load during the sliding of the dam. Preliminary linear analyses identi-
fied the concrete-rock interface at the base of the dam and the concrete-concrete interface
at the crest of the dam as high tensile stresses areas and, therefore, where cracking and
sliding could occur. Consequently, the model non-linearity was constrained to these two
areas only, using tiebreak contact elements with a tension-shear failure criterion. Further
details of the modeling assumptions and the validation of the numerical model can be
found in the study by Bernier et. al [20]. After adjusting the properties of the dam and
the foundation materials, the fundamental period of the dam-reservoir-foundation sys-
tem is 0.25 s, which matches the fundamental period from a previous model of the same
monolith calibrated from in situ forced vibration tests [22].

4.3   Characterization of seismic hazard

A classic probabilistic seismic hazard analysis (PSHA) approach enables the calculation of
the probabilities of exceeding, at least once in a given time span and at a given site, a set
of ground motion parameter levels considering all possible earthquake ruptures defined in
a seismic source model. Such set of probability values is usually referred to as the hazard
curve, and three major steps are required to compute it: (i) characterization of seismic
sources, (ii) characterization of the ground motion models (GMM) and (iii) calculation of
probabilities.

4.3.1   Probabilistic seismic hazard analysis and disaggregation

Although eastern Canada is located in a stable zone, the occurrence of several major
earthquakes in the south-east of the country between 1663 and 2010 has led to the consid-
eration of this area as a moderate seismic zone. The most recent hazard model, on which
the seismic provisions of the 2015 National Building Code of Canada (NBCC) [2] are
based, comprises four components, one for each quadrant (north-west, south-west, north-
east and south-east), to reduce computational burden [140]. In this context, an updated
hazard model for the south-eastern quadrant was considered for this study, using three
sub-models, distinguished primarily as the historical cluster (H2), regional seismotectonic
(R2) and a hybrid approach (HY) between H2 and R2. The adopted weights were 0.4
for both H2 and HY and 0.2 for R2. The empirical ground motion models were provided
in the form of look-up tables (based on Atkinson and Adams [67]) to be used with the
considered sources. To account for the epistemic uncertainty, three weighted GMMs were
used to specify a central (weighted at 0.5), upper (weighted at 0.3) and lower (weighted at
0.2) representative equation set for each region/event type. The tables were also modified

to provide hazard values directly in terms of firm ground site conditions ($VS30 = 450$ m/s), which is the reference ground type for the NBCC 2015. A PSHA was then performed at the dam site using the open-source software OpenQuake [141], which features the aforementioned 2015 NBCC hazard model, as part of a project at the Global Earthquake Model (GEM) Foundation, using information provided by the Geological Survey of Canada (GSC). Deamplification factors consistent with NBCC 2015 were applied to the hazard curve to make it compatible with the hard rock soil conditions at the dam site. Different hazard levels were selected for the conditioning IM ($Sa(T_1) = 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9$ g) to conveniently cover the range of spectral accelerations corresponding to return periods from 700 to 30000 years. The hazard curve was obtained in a discrete manner by calculating the annual rate of exceedance at the hazard levels specified above. Furthermore, the hazard curve was also described analytically to enable seismic safety assessment through analytical or numerical integration, enabling the propagation of uncertainty in the hazard model. Given that at the dam site, the hazard data does not follow a linear relationship in log-log space, a hyperbolic function was used to approximate the hazard curve, as proposed by Bradley et al. [152]. This function provides a good approximation of both linear and non-linear hazard data in the range of adopted intensity levels, whereas it may not extrapolate accurately, and is given in Eq. (4.1),

$$\hat{V} = \alpha \exp \left\{ \beta \left[\ln \left(\frac{Sa(T_1)}{\gamma} \right) \right]^{-1} \right\} \tag{4.1}$$

where $\hat{V}$ is the estimated mean annual frequency of exceedance corresponding to a given $Sa(T_1)$, and α, β and γ are constants characterizing the fitting of the hazard curve data. Figure 4.2(a) compares the mean annual frequency estimates from Eq. (4.1) with the results obtained with OpenQuake, showing close agreement when the constants α, β and γ are, respectively, 5.0×10^5, 2.12×10^2 and 1.87×10^4. To understand and identify the earthquake scenarios with the greatest contribution to the overall seismic hazard, disaggregation was performed for the seven considered hazard levels using the hazard curve corresponding to the branch with the smaller sum of squared errors (SSE) with respect to the mean hazard curve. Figure 4.2(b) shows that the hazard curve corresponding to the hybrid source model (HY) is the closest to the mean hazard curve. Accordingly, Table 4.1 presents the earthquake scenarios resulting from the disaggregation of the HY source model coupled with the central GMM.

The disaggregation results are illustrated in Figure 4.3, where it can be seen (Figure 4.3 (a)) that for low $Sa(T_1)$ values, events of magnitudes between 5.0 and 6.0 and distances between
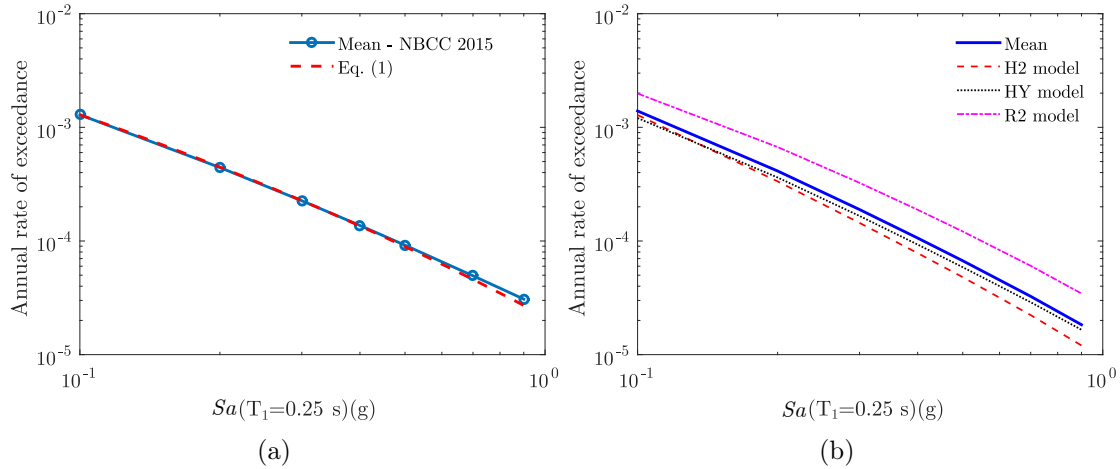
Figure 4.2 (a) Hazard curve at the dam site for $Sa(T_1 = 0.25\,\mathrm{s})$ and (b) Hazard curves according to the different source models

Table 4.1 Earthquake scenarios from the GMM disaggregation at $T_1 = 0.25\,\mathrm{s}$

Intensity level	Annual rate of exceedance	Mean values		
		M	R	ε
0.1 g	1.29E-03	6.01	51.18 km	0.78
0.2 g	4.43E-04	6.18	37.54 km	0.86
0.3 g	2.26E-04	6.31	29.70 km	0.91
0.4 g	1.37E-04	6.43	24.45 km	0.94
0.5 g	9.18E-05	6.52	21.77 km	0.95
0.7 g	4.94E-05	6.61	19.58 km	0.97
0.8 g	3.08E-05	6.82	18.70 km	0.99

15 km and 85 km are predominant at the dam site, while for high $Sa(T_1)$ values, as shown in Figure 4.3(b), contributions to the hazard are mainly from distances between 0 km and 45 km and magnitudes between 6.0 and 7.5.

4.3.2   Comparison of the NBCC 2010 and 2015 Hazard Models

Given the availability of an updated hazard model, the extent to which the PSHA outcome at the dam site, which is essential for the ground motion record selection that follows, is modified by the use of different models was investigated. As the knowledge and sophistication surrounding probabilistic seismic hazard modeling techniques have advanced, Canadian national mapping efforts have evolved from qualitative assessments in 1953 to fully probabilistic hazard models, foreseen by the seismic provisions of the 2015 NBCC. Figure 4.4(a) depicts the hazard curve obtained with the NBCC 2010 and 2015 models.
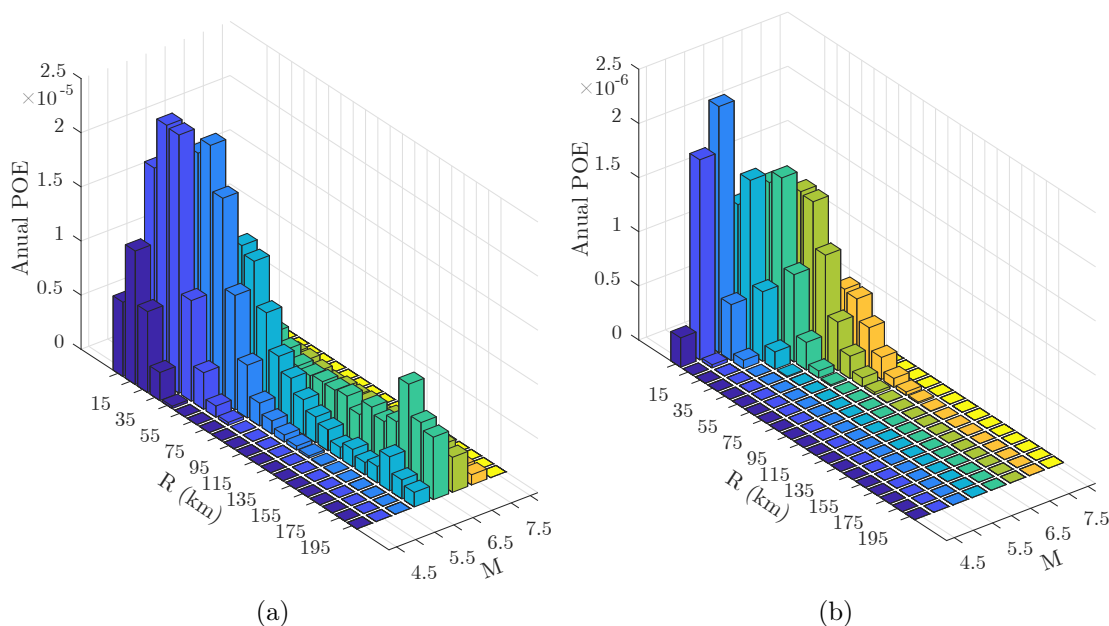
Figure 4.3 Magnitude-distance-annual probability of exceedance (POE) disaggregation at the dam site for: (a) $Sa(T_1 = 0.25\,\text{s}) = 0.1$ g; (b) $Sa(T_1 = 0.25\,\text{s}) = 0.9$ g

The 2015 model provides lower hazard estimates than those obtained with the 2010 model, upon which past fragility curves for gravity dams have been based [21]. This difference can be explained by the fact that the updated hazard model encompasses many important advances compared to its predecessor, such as the consideration of the contributions of different seismic source models, special consideration of large rare eastern earthquakes and the use of representative modern ground-motion models. Additionally, the two hazard models behave differently due to the changing mesh size in the computer software Open-Quake. As shown in Figure 4.4(b), the 2010 model is more stable than the 2015 model when considering a very coarse or refined mesh. Hence, several meshing possibilities were tested to determine the earthquake rupture forecast (ERF) sensibility and to achieve the optimal trade-off between precision and computational time. These improvements are also reflected in the lower disaggregation values of the updated hazard model shown in Figure 4.5.

4.4 Ground motion records selection

To proceed with the evaluation of the vulnerability of the case study dam through the development of fragility functions, a representative set of ground motion time series (GMTS), which properly accounts for the aleatory uncertainty, is necessary. When selecting ground
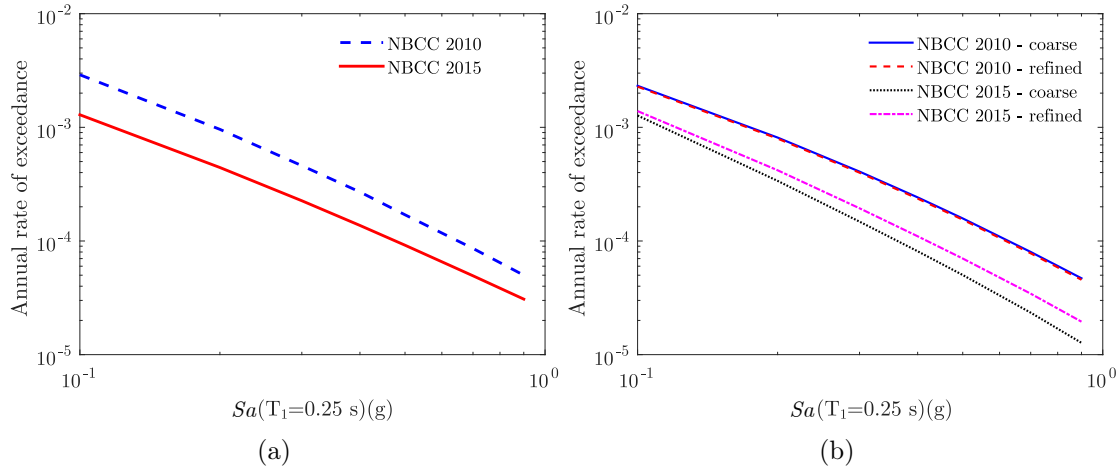
Figure 4.4   (a) Hazard curve at the dam site for $Sa(T_1 = 0.25\,\text{s})$ and (b) Hazard model - ERF sensitivity
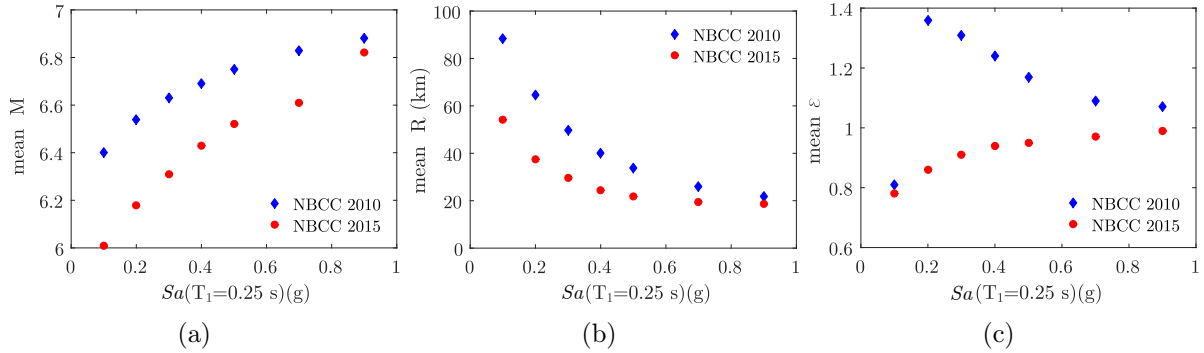
Figure 4.5   Comparison of the mean disaggregation values according to the NBCC 2010 and 2015 hazard models: (a) Magnitude, (b) Distance and (c) Epsilon

motions, the established goal was to obtain time series that were consistent with the ground motion amplitudes and properties computed from seismic hazard analysis. Therefore, a record selection approach that considers all the relevant IMs for the response of the structure was adopted. This approach is based on the aforementioned GCIM, a framework in which the distribution of any ground motion IM can be obtained given the occurrence of another specific ground-motion intensity measure.

4.4.1   Generalized Conditional Intensity Measure Approach

The essence of the GCIM approach is the definition of the multivariate distribution of any set of ground-motion intensity measures (IM_i) conditioned on the occurrence of a specific ground-motion intensity measure (IM_j) [74]. The calculation of the GCIM distributions

involves two main steps: (i) determining the probability that if a ground motion was
observed with IM_j, it was caused by a particular earthquake scenario; and (ii) given
the observed ground motion with IM_j from a particular earthquake scenario, defining the
distribution of the other ground motion intensity measures. Moreover, the general variable
$IM|Rup$ (where "$|Rup$" indicates conditioning on a specific earthquake rupture scenario) is
assumed to follow a multivariate log-normal distribution, and the marginal distributions of
all the scalar intensity measures in $IM_j|Rup$, $IM_i|Rup$ can be estimated via existing ground-
motion prediction equations. This means that the correlation coefficient matrix of IM is
sufficient to uniquely specify the multivariate log-normal distribution for $IM|Rup$; hence,
for each IM_i in IM, $IM_i|Rup,IM_j$ has a univariate (conditional) log-normal distribution
that, according to [153], can be expressed by Eq. (4.2).

$$f_{IM_i|Rup,IM_j} \sim \ln\mathcal{N}(\mu_{\ln IM_i|Rup,IM_j}, \sigma^2_{\ln IM_i|Rup,IM_j}) \tag{4.2}$$

Accordingly, the conditional mean and standard deviation are given by Eq. (4.3) and
Eq. (4.4).

$$\mu_{IM_i|Rup,IM_j} = \mu_{\ln IM_i|Rup} + \sigma_{\ln IM_i|Rup}\rho_{\ln IM_i,\ln IM_j|Rup}\varepsilon_{\ln IM_j} \tag{4.3}$$

$$\sigma_{IM_i|Rup,IM_j} = \sigma_{\ln IM_i|Rup}\sqrt{1 - \rho^2_{\ln IM_i,\ln IM_j|Rup}}, \tag{4.4}$$

where $\mu_{IM_i|Rup,IM_j}$ and $\sigma_{IM_i|Rup,IM_j}$ are the mean and standard deviation of $\ln IM_i$ for a
given rupture and are commonly obtained from the ground motion prediction equations;
$\rho_{\ln IM_i,\ln IM_j|Rup}$ is the correlation coefficient between $\ln IM_i$ and $\ln IM_j$ for a given earth-
quake scenario, commonly provided by empirical correlation equations; and the parameter
$\varepsilon_{\ln IM_j}$, representing the number of standard deviations between the target and the mean
predicted spectral amplitudes, is obtained via Eq. (4.5).

$$\varepsilon_{\ln IM_j} = \frac{\ln IM_j - \mu_{\ln IM_j|Rup}}{\sigma_{\ln IM_j|Rup}} \tag{4.5}$$

In this case, the conditioning intensity measure is the spectral acceleration at the fun-
damental period of the dam ($IM_j = Sa(0.25)\,\mathrm{s}$) obtained from PSHA, and the specific
earthquake scenarios were obtained from the disaggregation analyses for different inten-
sity levels (Table 4.1). Due to the limited availability of recorded ground motions in eastern
North America (ENA), empirical ground motion and correlation models are rarely avail-

able for this region. Therefore, equations proposed for western North America (WNA)
were assumed to be valid for Eastern Canada because these models are suitable for shallow
crustal events, which are the most common type of events in ENA.

4.4.2   Selection algorithm

The purpose of using the GCIM approach to perform the selection of recorded ground
motion time histories is to include the intensity parameters with the greatest influence
on the structural response at the limit states of interest and the consequent seismic risk
assessment. In the case of gravity dam-type structures, PGV was found to be one of the
best-performing ground-motion-dependent scalar IMs to correlate damage [30]. Similarly,
the vertical spectral acceleration Sa_V is also expected to be relevant in heavy structures
of this type and is hence considered in the GCIM to account for the vertical compo-
nent of ground motion in the record selection. As a result, the set of considered IMs
are $\{Sa_H(T), Sa_V(T), PGV\}$, where $Sa_H(T)$ and $Sa_V(T)$ are computed at 20 vibration
periods in the range of $T = \{0.2T_1 - 2T_1\}$, as proposed by [73], leading to a total of 41 IMs
to be considered in addition to the conditioning IM. The following GMMs and correla-
tion models were considered: Atkinson and Adams [67] ($Sa_H(T)$ and PGV), Gulerce and
Abrahamson [154] ($Sa_V(T)$), Baker and Jayaram [155] (ρ_{Sa_H,Sa_H}), Baker and Cornell [156]
(ρ_{Sa_H,Sa_V} and ρ_{Sa_V,Sa_V}), and Bradley [157] ($\rho_{Sa_H,PGV}$). Finally, due to the lack of empirical
correlation models between $Sa_V(T)$ and PGV, and considering that previous studies have
shown that PGV is strongly correlated to moderate period spectral accelerations [153],
the correlation between these two IMs was approximated by Eq. (4.6).

$$\rho_{PGV,Sa_V(T)} \approx \rho_{Sa_H(1.5s),PGV}\,\rho_{Sa_H(1.5s),Sa_V(T)} \qquad (4.6)$$

The GCIM distribution computed with the above-mentioned IMs was then used to simulate
and select 30 ground motion records for each intensity level. Given the lack of recorded
ground motions for the specific case study region, artificial simulated records could have
been employed. However, the former have still been preferred over the latter to include
the added value of working with real ground motions, given that stochastically simulated
spectra might present unrealistic properties [158, 74]. In line with the previously mentioned
limitations in terms of the available GMMs, the records were selected from the PEER
NGA-West2 database [159] due to the limited availability of strong ground motion records
in the PEER NGA-East database [160]. Another reason for not using the NGA-East
database was the lack of the vertical acceleration component in the records, which rendered

the corresponding matching within the selection algorithm impossible. Furthermore, as stated by Jayaram et al. [70], the considered dataset will have a greater effect on the GMMs for the IMs than on the correlation empirical models. As such, the use of empirical GMMs and records from the NGA-West2 database are indeed a limitation of the study in terms of the estimated response of the case study dam, as previously recognized in the study by Bernier et al. [21], given that: (i) hard rock sites in ENA are much stiffer than those in western regions; (ii) when compared to western regions, eastern events tend to generate high amplitudes at low periods; and (iii) ground acceleration attenuates more slowly in ENA, with a distance close to the source at low periods. Therefore, it was guaranteed that the ground motions from the NGA-West2 database matched the GCIM distribution at low periods [161, 162]. Furthermore, a pre-screening of the database in terms of magnitude, distance to the source and soil type was performed to limit the selection of ground motions to those that closely describe the seismic characteristics at the dam site. Specifically, only ground motions with magnitude (M), distance (R) and shear wave velocity (VS30), respectively, of $5.5 \leq M \leq 7.5$, $0\,\mathrm{km} \leq R \leq 200\,\mathrm{km}$ and $550\,\mathrm{m/s} \leq VS30 \leq 2000\,\mathrm{m/s}$ were considered. Alternatively the causal parameter bounds proposed by Tarbali and Bradley [142] could have been used to filter the database for usable spectra but it was preferred, for the sake of simplicity, to consider bounds based on the disaggregation analysis and previous studies [21] to cover all the range of studied IM levels. Despite the fact that the simulations were conducted in a 2D model and that the horizontal components could have been taken individually (obtaining twice the number of records), the geometric mean of the response spectra for the two orthogonal horizontal components of motion ($SaRotD50$) was selected for two main reasons: (i) the PGV values in the NGA-West2 database correspond to the geometric mean of the horizontal component; (ii) empirical ground motion models work with the mean and standard deviation of the logarithm of Sa, which can be represented by a Gaussian distribution. The ground motions were then selected using the greedy optimization algorithm proposed by Jayaram et al. [70] and later improved by Baker and Lee [71]. The statistical Kolmogorov-Smirnov (KS) test was used to compare the target GCIM distribution and the empirical distribution of the selected ground motion set for an IM to evaluate the quality of the selected ground motion sets.

Figure 4.6(a) presents the horizontal response spectra of the selected records for the intensity level of $Sa(T_1) = 0.1\,\mathrm{g}$, and Figure 4.6(b) presents the corresponding vertical response spectra of the selected records. A good match was found for the selected GMTSs that lay within the desired range of values (mean and standard deviation) for the entire range of considered periods and intensity measures. By contrast, Figure 4.7 shows the theoretical

cumulative distribution and the realizations of two sets of records for the vertical spectral acceleration at the fundamental period and PGV selected with the CS and the GCIM approaches. Even if the hazard model and the GMMs are the same, the records selected considering only horizontal spectral acceleration as the IM are less representative of the analytical distributions and hence are less consistent with the seismic scenario at the dam site. Conversely, the distribution of the records selected with $IM = \{Sa_H(T), Sa_V(T), PGV\}$ and the GCIM method is in agreement with the conditional cumulative distributions, as is evident from the empirical distribution function of the realizations between the KS rejection bounds (for a confidence level of 0.05).
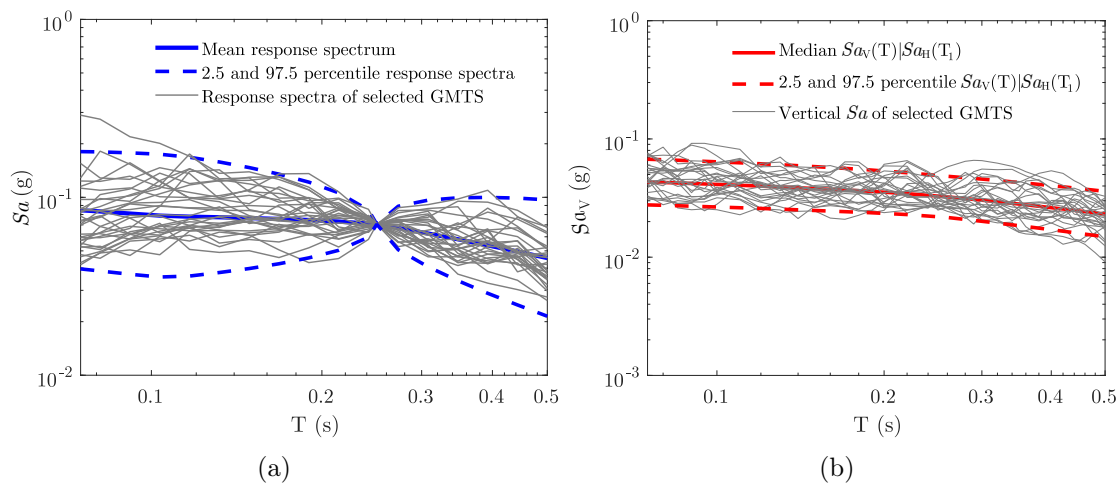
Figure 4.6 Distribution of the 30 selected ground motions at $Sa(T_1 = 0.25\,\mathrm{s}) = 0.1\,\mathrm{g}$ for: (a) horizontal spectral acceleration, (b) vertical spectral acceleration

4.5   Fragility analysis

Seismic assessment through fragility analysis provides a framework to ensure that available information on uncertainties is treated consistently [14]. The first step is the identification of the limit states that are relevant to the system performance, for which a specific fragility curve is developed, taking into account a rational assessment of all sources of uncertainty likely to affect it.

4.5.1   Damage limit states

When subjected to strong ground motion, gravity dams may be damaged in different ways. In recent years, typical damage modes that could lead to the potential collapse of dams after a seismic event have been identified, and seismic damage levels have been
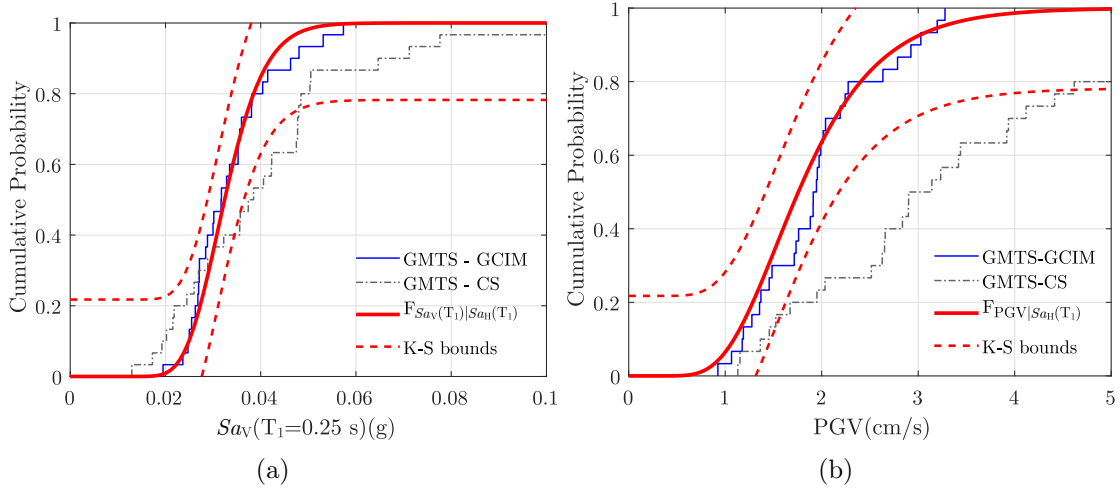
Figure 4.7 Empirical distribution function of the 30 selected GMTSs at $Sa(T_1 = 0.25\,\mathrm{s}) = 0.1\,\mathrm{g}$: (a) $Sa_V(T_1 = 0.25\,\mathrm{s})$, (b) PGV

established. Preliminary analyses have identified sliding as the critical failure mode for the case study dam [20], and other failure modes would only occur after sliding has already been observed. As a result, two limit states were considered in this study: (i) sliding at the concrete-to-rock interface at the base of the dam, and (ii) sliding at the concrete-to-concrete joint at the neck of the dam. Each limit state was characterized by the sliding damage states proposed in Bernier et. al [21], which are presented in Table 4.2.

Table 4.2 Limit states considered for the concrete gravity dam

Sliding damage state	Base	Neck
Slight/minor	5 mm	
Moderate	25 mm	
Extensive	100 mm	
Complete	150 mm	10 mm

4.5.2 Modeling parameter uncertainty

Table 4.3 presents the parameters that were considered as random variables in the numerical analysis of the dam response and for which the uncertainty or likelihood of occurrence was formally included, through their probability distribution function (PDF). All the remaining input parameters were held constant and were represented by their best estimate values. While the seismic response modeling uncertainty has been taken into account in terms of parameter uncertainty, it should be acknowledged that neglecting higher-level uncertainties in the seismic response, can lead to the underestimation of the true model

uncertainty [163]. For the case study dam, due to the limited availability of material investigations, the probability distributions were defined using empirical data from similar dams. The uniform distribution was used for the parameters other than damping, for which a log-normal distribution was adopted, as proposed by Ghanaat et al. [83]. A screening study to assess the effect of each modeling parameter on the response of the dam can be found in Bernier et al. [20]. The uncertainties due to the variables in Table 4.3 were propagated in the fragility analysis using Latin hypercube sampling (LHS) due to its ability to ensure that the set of samples reflects the entire range of the parameters, as demonstrated in past applications to earthquake engineering [164, 165]. Thirty statistically significant samples of the dam model at each intensity level were generated with this procedure.

4.5.3 Fragility curves

A fragility function is defined by the conditional probability that a structure will meet or exceed a specified level of damage for a given ground motion intensity measure, as illustrated by Eq. (4.7):

$$\text{fragility} = \text{P}_f(\text{IM}) = \text{P}_f(\text{LS}|\text{IM} = im), \tag{4.7}$$

where $\text{P}_f(\text{IM})$ is the fragility, LS is the damage limit state, IM is the intensity measure, and im is the achieved condition for the specified IM. In this study, the chosen IM was the spectral acceleration at the fundamental period of the dam-reservoir-foundation system $(Sa(T_1))$. The fragility curves were developed using multiple stripes analysis (MSA) [68], which enables efficient estimation of fragility curves from a small number of dynamic analyses at a limited number of discrete intensity levels. For each of the 7 intensity levels,

Table 4.3   Modeling parameter assumed PDFs

Parameters	PDF	Distribution Parameters	
Concrete-to-rock tensile strength (MPa)	Uniform	L=0.2	U=1.5
Concrete-to-concrete tensile strength (MPa)	Uniform	L=0.3	U=2.0
Concrete-to-rock cohesion (MPa)	Uniform	L=0.3	U=2.0
Concrete-to-concrete cohesion (MPa)	Uniform	L=0.9	U=2.5
Concrete-to-rock angle of friction (°)	Uniform	L=42	U=55
Concrete-to-concrete angle of friction (°)	Uniform	L=42	U=55
Drain efficiency (%)	Uniform	L=0.0	U=66
Concrete damping (%)	Log-normal	λ=-2.99	ζ=0.35

30 samples of the finite element model were generated with LHS and were paired with 30 sets of ground motion records selected via GCIM. The maximum sliding displacements at the base and at the neck interface were computed from the non-linear simulations of the dam-reservoir-foundation system. A log-normal PDF was fit to the sliding displacements at the base and neck for each intensity level; thus, the fragility point estimates were computed as the probability that the displacement at each intensity level exceeds the specific damage state displacement threshold. Finally, to compute the fragility, a log-normal cumulative distribution function (CDF) was fitted to the data points according to Eq. (4.8), in which $\Phi_{\mathcal{N}}$ is the integral probability in the standard normal space, μ is the median capacity (expressed in units that are dimensionally consistent with the demand parameter, im), and β is the logarithmic standard deviation (or dispersion) of the capacity.

$$ P_f(\text{IM}) = \Phi_{\mathcal{N}}\left(\frac{\ln(\text{IM}/\mu)}{\beta}\right) \tag{4.8} $$

Following the recommendations of Baker [68], when using MSA, the maximum likelihood estimation (MLE) method was employed to fit Eq. (4.8) to the fragility point estimates. Figure 4.8 presents the resulting fragility curves.
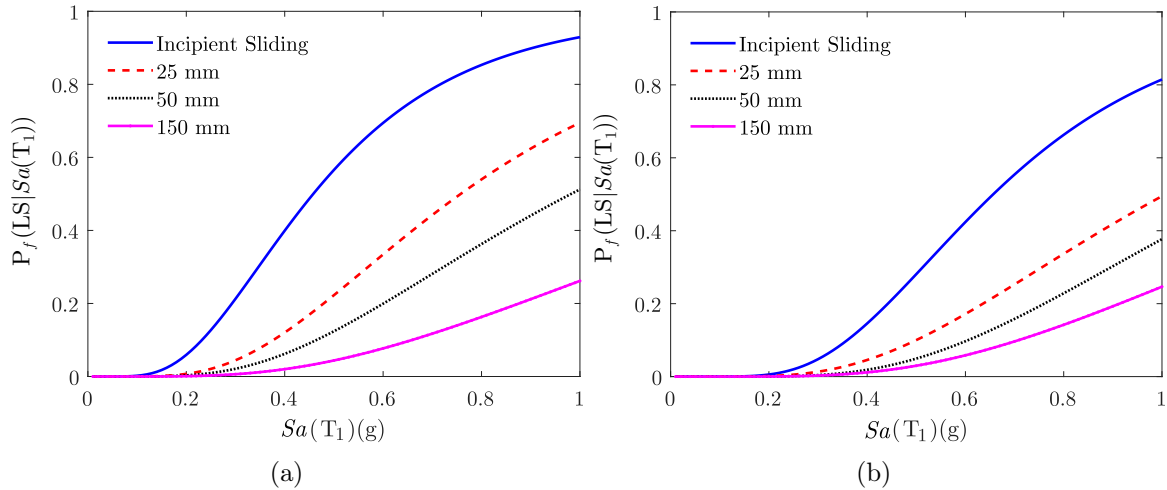
Figure 4.8   Fragility curves for: (a) base sliding, (b) neck sliding

4.5.4   Comparison of the ground motion selection approaches

The superiority of the CS method as a ground motion record selection tool to develop fragility curves compared to the traditional methodology based on the uniform hazard spectrum as the target was demonstrated by Bernier et al. [21] for the seismic assessment of gravity dams. Nevertheless, the ground motion record selection was conducted considering

the horizontal spectral acceleration as the sole IM, whereas the fragility curves developed herein are based on records selected by considering multiple IMs. Therefor, to illustrate the differences, Figure 4.9 directly compares the fragility curves proposed in this study, which were obtained from the updated Canadian hazard model and multiple-IM-based record selection with the GCIM (Method A), with the fragility curves obtained using the 2010 hazard model for Canada and single-IM-based record selection with the CS method (Method B). The fragility estimates of the curves obtained via Method B are conservative with respect to those obtained by Method A, an outcome that is even more evident for the neck sliding limit state. For the base sliding limit state, the median differs by 2 to 20%, whereas for the neck sliding limit state, it differs by 11 to 24 %. These results were expected because it was previously shown in Figure 4.4(a) that the annual rate of exceedance of the 2015 hazard model is lower than that calculated with the 2010 hazard model for the same range of spectral acceleration. Moreover, the representativeness of the selected ground motions with Method B is based on comparison of the mean and standard deviation of the spectral accelerations at various vibration periods, whereas in the proposed approach (Method A), representativeness is based on statistical goodness-of-fit comparisons for each IM of the empirical distribution function (EDF) of the ground motion records and the GCIM distribution. This leads to ground motion record sets that more accurately describe the seismic scenario at the site and the structure in question due to the ability to account for all the relevant IMs. These results also highlights how selecting ground motion records to match a target (elastic) response spectrum can lead to a less precise and overly conservative estimate of the probability of reaching the considered damage limit states.
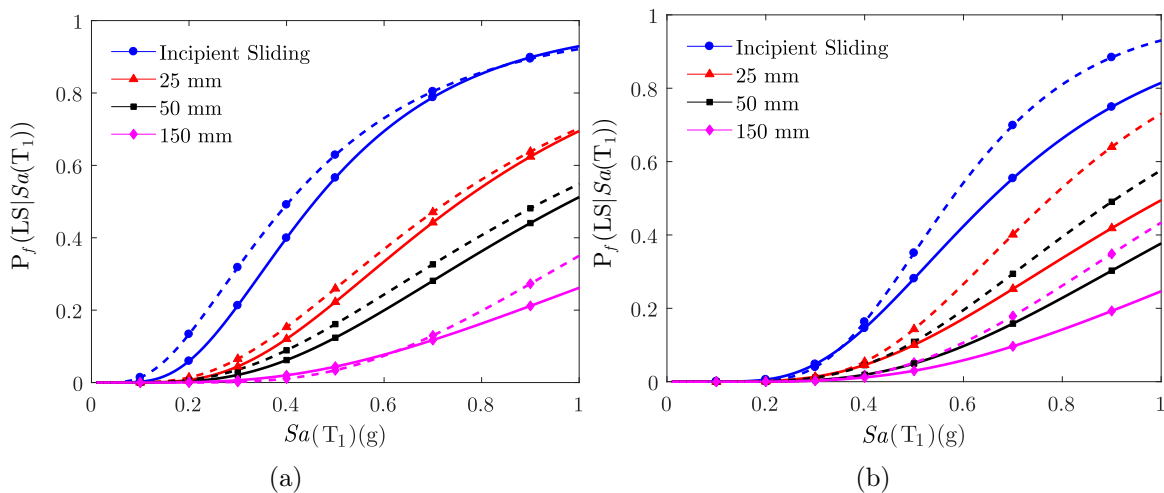
Figure 4.9 Comparison of fragility curves obtained via Method A (solid line) and Method B (dashed line) for: (a) base sliding and (b) neck sliding limit states

4.6   Safety assessment

4.6.1   Expected seismic performance under extreme limit states

Given the high risk associated with concrete gravity dams related to the catastrophic consequences of a potential collapse and uncontrolled release of water, their safety assessment represents one of the most important and complex problems for owners and regulatory organizations. In Canada, the committee of the Canadian Dam Association (CDA) has the vision of being the authoritative source for dam safety practices, and its current design philosophy provides the minimum standards for life safety [31]. Now that the updated fragility functions have been proposed in the previous section, it is important to understand the level of risk (for the considered limit states) to which the case study dam is subject. To evaluate the seismic performance of the case study dam under extreme limit states, the return period boundaries for the hazard classification provided by the CDA (Table 4.4) were used. For a given probability of exceedance (P_e), prescribed by the safety guidelines, the corresponding spectral acceleration was extracted from the hazard curve. The fragility estimates of the system for the specified return periods (P_r) were then calculated from the fragility curves corresponding to the 100 mm and 150 mm damage states for the neck and the base sliding limit states, respectively. The probabilities of exceedance calculated for the "extreme risk" case were compared with the values proposed by the ASCE 7-16 [148] guidelines for the maximum probability of failure for the MCE. For a risk category III (dam-type structures) and total or partial structural collapse, the maximum probability of failure should be less than 6 %. As shown in Table 4.4, for Method A and for both limit states, the probability of exceedance is less than 6 %. By contrast, the results obtained with Method B show that for both limit states, the probability of exceedance is well above the maximum value and is even higher for neck sliding. The main outcome of this short safety assessment exercise is that the outdated hazard model, coupled with a less comprehensive ground motion record selection approach, yields overly conservative estimates, which can have a direct impact in the application of safety guidelines and eventual decision making.

4.6.2   Unconditional Probabilities of Failure

A structure is potentially exposed throughout its lifetime to all the possibilities of the occurrence of ground motion intensities at a given site, as characterized by the site-specific seismic hazard curve. Accordingly, and for completeness, the assessment of the annual risk, measured as the annual probability of exceedance of the specified damage state, was

Table 4.4 Probability of exceedance of the limit states set by the CDA

Hazard	CDA boundaries	Method A		Method B	
		P_r, Base	P_r, Neck	P_r, Base	P_r, Neck
Low	$P_e \geq 1/100$	< 0.01	< 0.01	< 0.01	< 0.01
Significant	$1/1000 < P_e < 1/100$	< 0.01	< 0.01	< 0.01	< 0.01
High	$P_e \geq 1/2475$	< 0.01	< 0.01	< 0.01	< 0.01
Very high	$1/10^4 < P_e < 1/2475$	< 0.01	< 0.01	1.31×10^{-2}	2.11×10^{-2}
Extreme	$P_e \leq 1/10^4*$	3.59×10^{-2}	2.34×10^{-2}	9.73×10^{-2}	1.37×10^{-1}

∗ Maximum considered earthquake (MCE)

also performed. The fragility functions were convolved with the hazard curve resulting from PSHA to calculate the unconditional probabilities of exceedance. Using the hazard function from Eq. (4.1), the unconditional probability of exceedance can be approximated with Eq. (4.9) and used to assess the annual probability of damage, which has been proven to be an accurate approximation of the probability of failure for small periods of time, such as 1 year [147, 166].

$$P_{f,unc} = \int_{Sa(T_1)} [1 - P_f(\text{LS}|\text{IM}, Sa(T_1))] \left| \frac{\mathrm{d}\bar{V}}{\mathrm{d}Sa(T_1)} \right| \mathrm{d}Sa(T_1) \qquad (4.9)$$

$P_f(\text{LS}|\text{IM}, Sa(T_1))$ is given by the fragility function, and $\dfrac{\mathrm{d}\bar{V}}{\mathrm{d}Sa(T_1)}$ is the PDF of the median hazard at the dam site. Table 4.5 presents the risk evaluated as the annual probability of exceeding each damage limit state, and the unconditional probabilities for the most damaging limit states were compared with the annual probabilities of failure found in the literature. MacGregor [149] recommended that the collapse probability of a building-type structure should be on the order of magnitude of 5×10^{-5} and even lower for structures of special importance, such as dams and nuclear power plants. In the same manner, following the risk category framework proposed by the ASCE 7-16, the number of people at risk is between 100 and 1000. For this range, the FEMA P-1025 [150] guidelines specify a value between 1×10^{-5} and 1×10^{-6} for tolerable risk. Table 4.5 shows that for the most severe limit states, the corresponding probabilities of exceedance are within the aforementioned boundaries.

Table 4.5 Annual probability of exceedance

Limit state	$P_{f,unc}$, Base	$P_{f,unc}$, Neck
Incipient sliding	1.32×10^{-4}	5.17×10^{-5}
25 mm	4.35×10^{-5}	1.99×10^{-5}
50 mm	2.47×10^{-5}	1.07×10^{-5}
150 mm (base) - 100 mm (neck)	9.26×10^{-6}	6.63×10^{-6}

4.7 Conclusions

The main objective of this study was to understand and present the level of improvement in the fragility analysis of concrete gravity dams provided by the use of an innovative method for record selection that combines the latest advances in probabilistic seismic hazard analysis and the most influential IMs for the proposed limit states.

The methodology was applied to a case study dam located in north-eastern Canada. PSHA was performed with the open-source software OpenQuake and the updated 2015 hazard model for Canada to characterize the seismic scenario at the dam site. The results of the PSHA at the dam site with the updated hazard model were compared with the results of the previous version to quantify the extent to which the differences between the two models are significant. The updated model showed lower annual probabilities of exceedance for the same site, which can be explained by the consideration of a more extensive database.

The intensity measures identified as relevant for the seismic response of the dam were included in the modified selection algorithm consistent with the GCIM formulation. The sets of selected ground motion records were paired with the 30 samples of numerical models of the dam generated with LHS for each intensity level. The fragility curves for the base and neck sliding limit states of the dam were then generated via multiple stripes analysis. A safety assessment was performed by calculating the expected seismic performance under the extreme limit states proposed by the Canadian national guidelines (CDA) and by determining the unconditional probabilities of damage limit state exceedance. The fragility functions and the results from the safety assessment obtained with the proposed methodology were compared with those resulting from the previous hazard model and using a selection algorithm considering spectral acceleration as the only IM. The comparison showed that the proposed methodology produces less conservative fragility functions which can be related to the more accurate hazard model and to the fact that the selected set of ground motions includes the characteristics of the IM that, in addition to Sa_H, are likely to be present at the dam site. Selecting ground motions by considering only the horizontal spectral acceleration can lead to sets of records whose distributions with respect

to other important IMs are not consistent with the seismic scenario at a specific site. Furthermore, the results of the safety assessment with the present methodology were in line with the minimum safety margins proposed by the current guidelines, whereas the results were above the tolerable limits for the same case study dam when using the previously employed procedure based on a less thorough seismic input definition.

Despite the specificities of this case study, namely, its location in eastern Canada and consideration of only sliding limit states, the conclusions are believed to be useful for the seismic assessment of concrete gravity dams in general. An important limitation of this study, which should be addressed by future developments in the field, is the lack of empirical GMMs and correlation models for ENA and the corresponding impossibility of selecting records from the ENA database.

4.8   Numerical Implementation

The source code for the GCIM distribution and the selection algorithm used for the case study dam are available at https://github.com/SeguraRL/GCIM_gm_eastern_Can. The repository includes meta-data and all the empirical GMMs and correlation models considered for eastern Canada.

4.9   Acknowledgments

CHAPTER 5

META–MODEL–BASED SEISMIC FRAGILITY ANALYSIS OF CONCRETE GRAVITY DAMS

Avant-propos

Auteurs et affiliation :

R. L. Segura : étudiante au doctorat, Université de Sherbrooke, Faculté de génie, Département de génie civil et de génie du bâtiment.

J. E. Padgett : professeur, Rice University, Department of Civil and Environmental Engineering.

P. Paultre : professeur, Université de Sherbrooke, Faculté de génie, Département de génie civil et de génie du bâtiment.

Titre français : Analyse de Fragilité Sismique de Barrages-Poids en Béton Basées Sur des Méta-modèles

Contribution au document : Cet article contribue au document présenté par explorer différentes techniques de méta-modélisation pour sélectionner les plus viables pour l'évaluation sismique de la sûreté des structures de type barrage-poids basées sur des fonctions de fragilité multi-variables, développé par la mise en œuvre de ce méta-modèle. Les résultats de l'analyse probabiliste des aléas sismiques et les accélérogrammes sélectionnées dans l'article constituant le chapitre 4, sont utilisées pour générer les points de formation auxquels les méta-modèles de régression ont été ajustés. De plus, une méthodologie permettant d'ajuster les surfaces de fragilité paramétrées aux estimations des points de fragilité générées avec les méta-modèles est présentée. Enfin, la comparaison de ces fonctions de fragilité avec les directives de sécurité actuelles est présentée, ce qui permet d'établir une gamme de valeurs utilisables des paramètres du modèle.

Résumé français : Des méthodes probabilistes, telles que l'analyse de la fragilité, ont été développées comme alternative prometteuse pour l'évaluation sismique des structures de type barrage. Cependant, compte tenu de la réévaluation coûteuse des simulations de

modèles numériques, l'effet des paramètres de modèle susceptibles d'affecter la fragilité sismique du système est souvent négligé. Reconnaissant le manque d'exploration approfondie de différentes techniques d'apprentissage automatique pour développer des substituts ou des méta-modèles permettant d'approcher efficacement la réponse sismique des barrages, cette étude fournit des informations sur les méta-modèles viables pour l'évaluation sismique des barrages à gravité destinés à l'analyse de la fragilité. La méthodologie proposée pour générer des fonctions de fragilité multivariée est efficace tout en prenant en compte la variation de paramètre de modèle la plus critique influant sur la fragilité sismique du barrage. À partir de l'analyse de ces modèles, des recommandations de conception pratiques peuvent être formulées. La procédure présentée ici est appliquée à un barrage situé dans le nord-est du Canada, dont la fragilité est évaluée par comparaison avec les directives de sécurité en vigueur afin d'établir une plage de valeurs de paramètres de modèle utilisables.

Abstract:

Probabilistic methods, such as fragility analysis, have been developed as a promising alternative for the seismic assessment of dam-type structures. However, given the costly reevaluation of the numerical model simulations, the effect of the model parameters likely to affect the seismic fragility of the system is frequently overlooked. Acknowledging the lack of the thorough exploration of different machine learning techniques to develop surrogates or meta-models that efficiently approximate the seismic response of dams, this study provides insight on viable meta-models for the seismic assessment of gravity dams for use in fragility analysis. The proposed methodology to generate multivariate fragility functions offers efficiency while accounting for the most critical model parameter variation influencing the dam seismic fragility. From the analysis of these models, practical design recommendations can be formulated. The procedure presented herein is applied to a case study dam in northeastern Canada, where the polynomial response surface of order 4 (PRS $\mathcal{O}^4$) came up as the most viable meta-model among those considered. Its fragility is assessed through comparison with the current safety guidelines to establish a range of usable model parameter values in terms of concrete-rock angle of friction, drain efficiency and concrete-rock cohesion.

5.1   Introduction

Methods for the seismic analysis of dams have improved extensively in the last few decades, and the growth of computing power has expedited this improvement. Advanced numerical models have become more feasible and, thus, constitute the basis of improved procedures for design and assessment. Moreover, a probabilistic framework is required to manage the various sources of uncertainty that may impact the system performance and decisions related thereto [32]. Fragility analysis, which depicts the conditional probability that a system reaches a structural limit state, is a central tool in this probabilistic framework. Traditional vulnerability assessment methods develop fragility functions by using a single parameter to relate the level of shaking to the expected damage, which consequently produces a robustness of predictions that is highly dependent on the selected parameter. However, the estimation of the fragility of the system can be potentially improved by increasing the number of parameters; in this way, a more complete description of the properties of ground motions can be obtained [89]. Furthermore, the effect of the variation of the material properties in the seismic fragility analysis of structures with complex numerical models, such as dams, is frequently overlooked due to the costly and time-consuming revaluation of the numerical model. The seismic response and vulnerability assessment of key infrastructure elements often require a large number of nonlinear dynamic analyses of complex finite element models (FEMs). The substantial computational time may be reduced by using machine learning techniques to develop a surrogate or meta-model, which is an engineering method used when an outcome of interest cannot be easily directly measured; thus, a model of the outcome is used instead [15]. In addition, if the outcome of interest comes from non-linear FEM analysis that reflects the dynamic behaviour of the structure under seismic loading, the meta-model will emulate this behaviour. To this end, these algorithms include several features in their mathematical formulation to help capture this highly non-linear behaviour. Among them, the use of higher order sparse polynomials, partitioning the sample space and fitting a series of models and then combines them into an ensemble with an overall better performance and by mapping into high dimensional feature space, between others.

Such a challenge is particularly relevant to the case of large-scale infrastructures such as dams subjected to seismic loads. Thus, the main goal of this study is to explore the applicability of meta-models for the seismic assessment of gravity dams and present a methodology to develop parameterized multivariate fragility functions through the use of the latter. The secondary goal is to explicitly account for the effect of the model parameter variation in the seismic fragility analysis. The proposed methodology has the added asset

of properly depicting the seismic scenario likely to occur at a specific site, enhancing the accuracy of the seismic fragility analysis. The proposed methodology is applied to a case study gravity dam located in northeastern Canada.

5.1.1 Literature review

In the past two decades, univariate fragility functions, which depict the potential for limit state exceedance conditional on a ground motion intensity, have been readily adopted and developed for the seismic assessment of structures, as it can been found in several state-of-the-art reviews [167, 28, 168]. However, given the recognized limitations of such univariate fragility curves, the use of multivariate fragility functions to assess the vulnerability of a given structure or portfolio of structures is beginning to be used progressively [91, 99, 94, 92, 169]. Noted advantages of these parameterized or multivariate fragility functions include the potential for efficient posterior uncertainty propagation, exploring sensitivities or the influence of design parameter variation, and enabling application across a portfolio of structures. Nevertheless, given the large number of simulations required, the development of fragility surfaces or multivariate fragility functions that leverage numerical models can impose high computational burdens.

The combination of numerical models, probabilistic approaches and machine learning has gained considerable interest in the literature in recent years for engineering design and structural reliability [106, 115, 116]. This combination is justified by the significant randomness that characterizes not only the earthquake excitation but also the structural system itself (e.g., stochastic variations in the material properties, degradation due to aging and temperature fluctuation, etc.). Surrogate modeling techniques within a seismic fragility framework have found recent applications for the safety assessment of buildings and bridges, among other structures [123, 118, 96, 90, 117, 124]. Even though many of these studies considered several seismic intensity measures (IMs) and model parameters (MPs) for building the meta-models to predict the response of the structure, most of them do not clearly depict the influence of all the considered parameters in the form of multivariate fragility functions from the respective meta-models. For the specific case of dam-type structures, an extensive comparison between machine-learning data-based predictive models for monitoring the dam behavior can be found in Salazar et al. [127] and Salazar et al. [128]. More recently, Hariri-Ardebili and Pourkamali-Anaraki [129], Hariri-Ardebili and Pourkamali-Anaraki [130], and Hariri-Ardebili [131] have used machine learning techniques to perform reliability analysis applied to gravity dams against flooding, earthquakes

and aging, considering, in some cases, explicit limit state functions and simplified FEM in others.

Most of the prior studies on the seismic assessment of concrete gravity dams via machine learning techniques are limited to the consideration of a single meta-model, simplified finite element models and univariate fragility functions. Therefore, they do not explore the most suitable meta-model for fragility analysis of this type of structure, nor they explore the influence of the variation of the model parameters on the seismic fragility analysis. Moreover, none these studies discuss the proper definition of the seismic scenario likely to occur at a specific site in a probabilistic manner.

5.1.2 Originality and contribution

To address the aforementioned gaps, this paper aims to identify most viable meta-model, from the subset of machine learning methods considered, for the seismic fragility assessment of gravity dams, provide an overview on the importance of the parameters influencing the dam performance and formulate design recommendations from the analysis. The major contributions of this paper, in order of presentation, can be listed as follows: (i) perform a comparative analysis to determine the best performing regression meta-model to predict the base sliding of gravity dams from 6 meta-models with different basis function configurations, yielding a total of 14 different regression techniques, for the first time; (ii) present a methodology to fit parameterized fragility surfaces, as a function of IMs and MPs, from the meta-models; (iii) consider the correlation between the seismic IMs from a probabilistic seismic hazard analysis (PSHA) to generate the samples to characterize the seismic scenario where the meta-model will be evaluated; (iv) gain insight into the influence of the model parameters affecting the dam performance and explicitly quantify their effect with the generation of multivariate fragility functions; and (v) formulate model parameter design recommendations from the analysis, e.g., appropriate range of parameters to achieve target risk.

5.2 Meta-model-based multivariate fragility procedure

Dam seismic assessment is a complex task due to the uniqueness of each of such structures and to the interaction between the different components of the system. Similarly, the seismic response of dam-type structures involves nonlinear dynamic analysis of complex finite element models, often requiring prohibitively high computation times. Machine learning describes a series of methods that allow learning from data, what relationships

exist between the quantities of interest. When performing probabilistic studies related
to structural reliability, it may be interesting to replace the finite element model by a
regression model built on a set of simulated responses, for the purpose of computational
efficiency [95]. Accordingly, the motivation of applying statistical learning techniques to
develop a seismic probabilistic demand model or meta-model is to expedite this safety
assessment process. Such "model of a model" is based on machine learning techniques
to allow the algorithm to learn from the data and because observations are the output
of a simulation, the observation model does not include any observation error. Within
this context, a probabilistic seismic demand model expresses an approximate relationship
between an uncertain seismic response, e.g., a dam's maximum relative base sliding, and a
set of parameters that influence the response. The basic idea is for the surrogate or meta-
model to act as a "curve fit" to the available data so that the results may be predicted
without requiring costly simulation. In the general case, the meta-model can be described
as follows:

$$\underbrace{y_i}_{\text{response}} = \underbrace{g(\mathbf{x}_i)}_{\text{meta-model}} + v, \quad \underbrace{\mathbf{x}_i = \{x_1, x_2, \ldots, x_n\}_i}_{\text{covariates}} \tag{5.1}$$

where the surrogate model $g(\cdot)$ statistically predicts the response of the structure, y_i, for
a given set of intensity measures and model parameters, $\mathbf{x}_i$, and v is the error due to the
lack of fit of the surrogate model.

The meta-models considered herein are within an adaptive scheme, i.e., the functions in
the meta-models can change according to the input data to reduce the burden of manual
selection of several parameters in the meta-model. Therefore, three adaptive meta-models
and an interpolation scheme will be considered. In addition, the performance of two other
statistical learning algorithms based on kernels and decision trees will also be addressed.
Among the different regression techniques, this paper focuses on the following: (1) poly-
nomial response surface models of order 2 to 4 with stepwise regression (PRS $\mathcal{O}^{2\text{--}4}$); (2)
multivariate adaptive regressive splines (MARS) with linear and cubic splines; (3) adap-
tive basis function construction (ABFC); (4) radial basis function (RBF) interpolation
with multiquadratic, thin plate splines and Gaussian basis functions; (5) support vector
machines for regression (SVMR) with linear, quadratic, cubic and radial basis function
kernels; and (6) random forest for regression (RFR).

Within the extent of this study, the first comparative analysis of meta-models in the con-
text of seismic assessment of dams is performed. Using the results from the finite element
simulations, this study develops meta-models for approximating the seismic response of

gravity dam-type structures using the three steps outlined in Figure 5.1. The subsequent sections will detail such steps.

5.2.1 Design of experiments

To minimize the associated cost of running dynamic nonlinear FEM of dams under seismic excitation while analysing an adequate number of loading conditions and structural system configurations, an appropriate experimental design method should be used. Structural and material properties likely to affect the seismic response of the structure should be considered, and their associated ranges should be based on experimental data or values found in the literature. The Latin hypercube sampling (LHS) experimental design method is adopted to generate n_f sample points representing the different configurations of the dam under study. This sampling technique was selected because of its ability to divide the desired range of values for each parameter into n-equiprobable intervals and then select a sample once from each interval.

Step 1: **Experimental design matrix**
- Select the set of parameters defining loading conditions and structural configurations;
- Define range of feasible values;
- Define the number n_f of FE simulations to be performed;
- Generate n_f sample points with LHS in the corresponding space of the selected parameters;

Step 2: **Finite element simulations**
- Conduct n_f dynamic non-linear finite element simulations;
- Obtain the structural response y_i, corresponding to each row $\mathbf{x}_i$, of the experimental design matrix, $\mathbf{X}$;

Step 3: **Fit meta-models**
- Fit different regression techniques $g(\mathbf{X})$ considering only the best demand and capacity predictors;
- Select the best meta-model in terms of predictive capacity based on CV goodness-of-fit estimates;
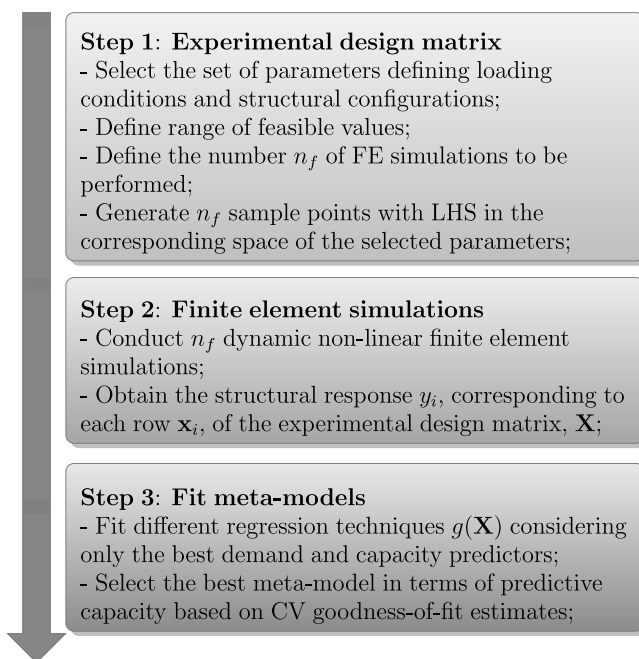
Figure 5.1 Procedure for the generation of meta-models

5.2.2 Meta-modeling techniques

The following subsections provide a brief overview of the different meta-models tested in this study for the ability to offer viable meta-models of the seismic response of dams. Only the most relevant features of each technique will be presented given that exhaustive

mathematical formulation can be found in the literature. Alternatively, relevant details concerning the model fitting and the algorithm settings are provided.

Polynomial response surface - PRS

The polynomial response surface is an m-dimensional surface that predicts desired responses using a computationally efficient closed-form polynomial function developed from a set number of input variables [103]. PRS was implemented together with stepwise regression to select the best explanatory or basis functions. The sparse polynomial response surface can be represented as follows:

$$\hat{y} = \mathbf{a}^{\mathsf{T}}\boldsymbol{\Theta} \tag{5.2}$$

where $\hat{y}$ represents the predicted value with the meta-model, the set of coefficients is represented by a column vector $\mathbf{a}$ and $\boldsymbol{\Theta}$ is a column vector of basis functions. In this study, the response variable was considered normally distributed and the meta-model was trained using the statistics and machine learning toolbox in Matlab [146].

Up to 2nd-order polynomials shall suffice for responses characterized by low curvatures, while 3rd- and 4th-degree polynomials including two factor interactions are more appropriate for significant curvatures [103]. This approach is considered because past studies have shown them to be efficient and accurate for concrete gravity dam seismic performance as well as for meta-models of other complex structures [131, 118].

Adaptive basis function construction - ABFC

ABFC is a sparse polynomial regression model building approach that enables adaptive model building without restrictions on the model's degree, accomplished in polynomial time instead of exponential time, as well as without the requirement to repeat the model building process [145]. The required basis functions are automatically iteratively constructed using heuristic search, specifically for the particular data. The ABFC meta-model can be represented by Eq. (5.2), where the order of the polynomial basis functions are adaptively determined. When working with relatively small datasets, basis functions selection bias and instability should be prevented. To this end, the ensemble of floating adaptive basis function construction (EF-ABFC) method proposed by Jekabsons [145] was used, together with the corrected Akaike's information criterion (AICC) as the penalization criteria for model evaluation. The software variReg [143] implemented through Matlab was used to train this meta-model.

Similar to the PRS, ABFC has also been proven to be a valuable technique for the evaluation of the seismic performance of complex structures [96].

Multivariate adaptive regressive splines - MARS

MARS is a form of regression analysis introduced by Friedman [170]. It is a nonparametric regression technique and can be considered an extension of linear models that automatically take into account nonlinearities and interactions between variables using a tensor product basis of regression splines to represent the multidimensional regression functions. The MARS meta-model can be represented as follows:

$$\hat{y} = \sum_{i=1}^{n} c_i b_i(\mathbf{x}) \tag{5.3}$$

where c_i are constants coefficients and $b_i(\mathbf{x})$ are the basis functions. Due to its adaptive nature, the MARS meta-model partitions the sample space and fits a series of models, each of which has a lower error, and then combines them into an ensemble with an overall better performance [90]. Two MARS models were trained, one with cubic splines and the other with linear splines. The toolbox for Matlab developed by Jekabsons [144] was used for this purpose. The algorithm builds a model in two phases: forward selection and backward deletion. From the backward deletion phase the "best" models of each size are selected and outputted as the final one based on the ones with the lowest generalized cross-validation (GCV) estimates. As suggested by Jekabsons [144], most attention was paid to the maximum number of basis functions included in the model, the maximum degree of interaction between input variables and the GCV penalty per knot. Regarding the maximum number of basis functions, the recommended value for this parameter is about two times the expected number of basis functions in the final model [170] and in the context of this study was limited to 30. The maximum degree of interaction was set to 3 because it was found to be a fair trade off between the meta-model predictive capabilities and the use of computational resources. Concerning the GCV penalty per knot, larger values will lead to fewer knots (i.e., the final model will have fewer basis functions). Simulation studies suggest values in the range of about 2 to 4 [144]. As recommended by Friedman [170], taking into account the maximum number of interactions, these value was set equal to 3.

MARS models are well suited to nonlinear problems, easily interpretable, and have achieved great accuracy in predicting structural response due to their adaptive nature while being computationally efficient for seismic assessment [127, 96, 90].

Radial basis functions - RBF

Radial basis function interpolation uses basis functions whose response monotonically changes as the distance from the central point increases. It was first introduced by Hardy [171] for scattered multivariate data interpolation, using linear combinations of radially symmetric functions based on Euclidian distances or similar metrics to approximate response functions. The RBF can be expressed in the following functional form:

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i \Phi_{RBF_i}[w_i(\mathbf{x})] \tag{5.4}$$

where $\Phi_{RBF_i}[w_i(\mathbf{x})]$ is a nonlinear mapping from the input layer to the hidden layer, a_0 is the bias, and $a_1, ..., a_m$ are the connection weights between the hidden layer and output layer, typically determined using iterative procedures. Multiquadratic, thin plate spline and Gaussian radial basis functions are some of the basis functions typically considered for interpolation [103]. For the Multicuadratic basis functions and the thin plate spline basis functions, the shape parameter was kept constant as $1/n_f$ while for the basis Gaussian basis functions, the smoothing parameter was found using leave one out cross-validation (LOOCV). The RBF meta-model was also implemented with the software variReg [145].

Despite the lack of transparency due to the hidden layer, RBF have been proven to generate excellent approximations to a wide range of structural responses [96, 90, 115] and are thus considered in this study.

Support vector machine for regression - SVMR

Support vector machines are a modern class of statistical learning algorithms with a sparse solution; thus, predictions only depend on a subset of the training data, known as support vectors [103]. This technique was originally designed for binary classification but can be extended to regression. Moreover, it contains all the main features that characterize the maximum margin algorithm: a nonlinear function is learned by a linear learning machine's mapping into high-dimensional kernel-induced feature space. The capacity of the system is controlled by parameters that do not depend on the dimensionality of the feature space. In SVM regression, the input $\mathbf{X}$ is first mapped onto an m-dimensional feature space using a nonlinear mapping, and then a linear model is constructed in this feature the space. The statistics and machine learning toolbox in Matlab [146] was used to train the meta-model with four different types of kernels (linear, quadratic, cubic and radial basis functions). A set of hyper-parameters values must be set before the learning process, which for SVMR includes the soft margin constant (cost function, C), parameters of the kernel function

(width of RBF kernel or degree of a polynomial kernel) and the tolerance width margin, ϵ. Regarding the soft margin constant the value $C = iqr(y)/1.349$, where $iqr(y)$ is the interquartile range of response variable, was used for the RBF kernel and $C = 1$ for all other kernels as recommended by Fan et al. [?]. The width of the RBF kernel was set equal to 1 and, as it was already mentioned, the degree of the polynomial kernel was set equal to 1, 2 and 3. Concerning the tolerance width margin, ϵ was set equal to $iqr(y)/13.49$ as recommended by Fan et al. [?].

SVMR has started to be used progressively for the fragility assessment of bridges [172, 90, 96] and more recently for dams [127, 129], showing satisfactory results.

Random forest for regression - RFR

A random forest is a meta-estimator that fits a number of classifying decision trees on various subsamples of the dataset and uses averaging to improve the predictive accuracy and control overfitting. The random forest model is an additive-type model that makes predictions by combining decisions from a sequence of base models. More formally, we can write this class of models as follows:

$$\hat{y} = \frac{1}{N} \sum_{i=0}^{N} f_i(\mathbf{x}) \tag{5.5}$$

where the final model is the sum of simple base models $f_i(\mathbf{x})$. Here, each base classifier is a simple decision tree. In random forests, all the base models are constructed independently using a different subsample of the data [103]. A regression ensemble was performed within 100 decision trees using Bootstrap aggregation as the ensemble aggregation method due to its ability of reducing over-fitting of the model, handle higher dimensionality very well and a less careful tuning of different hyper-parameters unlike the Boosting ensemble method [173]. Moreover, random forests typically offer a good estimate on prediction accuracy for external data based on the out-of-bag (OOB) accuracy. For bagged decision trees, the maximum number of decision splits was set at $n_f - 1$ and the number of predictors to select at random for each split was one third of the number of predictors. The algorithm was implemented with the statistics and machine learning toolbox in Matlab [146].

RFR meta-models are easy to train and implement, and in contrast to other methods, random forests are not sensitive to outliers. For these reasons, together with the fairly good performance of this technique in the literature [90, 96, 127], RFR are implemented in this study.

5.2.3   Cross-validated goodness-of-fit estimates

The goodness-of-fit estimates depict the discrepancy between the observed values from the FEM simulation and the estimated value with the meta-model in question. The root mean square error (RMSE) provides a measure of global error, quantifying the difference between the responses predicted by the meta-model and actual data and is computed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{n_f} (y_i - \hat{y}_i)^2}{n_f}} \tag{5.6}$$

where y_i are the response values in the dataset, $\hat{y}_i$ are the predicted values and n_f is the total number of points in the dataset (FE simulations). Additionally, the coefficient of determination R^2 can be calculated as follows:

$$R^2 = 1 - \frac{\sum_{n_f} (y_i - \hat{y}_i)^2}{n_f \sigma^2} \tag{5.7}$$

where σ^2 is the variance of the response in the dam response dataset. Similarly, the relative maximum absolute error (RMAE) measures the extent of the local fitting error and is the ratio of the maximum absolute difference between the meta-model and test data responses to the standard deviation of the actual response:

$$\text{RMAE} = \frac{max|y_i - \hat{y}_i|}{\sigma} \tag{5.8}$$

In the present study, the predictive capability of the meta-models will be assessed through 5-fold cross-validation (5-CV) [95]. The dataset is randomly divided into 5 sets, and the meta-model is trained using $k - 1$ sets, with the remaining set used as test data. This procedure is repeated 5 times; thus, 5-CV provides an estimate of the predictive accuracy of the model for unknown data. The average R^2 value resulting from 5-CV will be used along with RMSE and RMAE to compare the different meta-models.

5.2.4   Multivariate fragility analysis

As stated by Ghosh et al. [90], single-parameter demand models suffer from two potential drawbacks: (i) the inability to assess the impact of structural model parameter variation on structure performance during earthquakes without costly re-analysis for each new set of parameter combinations and (ii) the lack of flexibility to incorporate field instrumentation data from monitoring of existing structures to enable the updating of seismic fragility estimates. Consequently, the use of multivariate fragility functions enables the efficient

uncertainty propagation of the random variables and allows for the exploration of the effects of design parameter variation on the vulnerability of the structure. Thus, the goal is to identify the role of the most influential ground motion IMs and MPs on the induced damage in the structure to build multivariate fragility functions from the meta-model output to provide a more complete and accurate view of the vulnerability of the structure.

Similar to fragility curves, multivariate fragility functions offer the conditional probability of exceeding different limit states given the occurrence of an earthquake of a certain intensity. The only difference is that the specific limit state is characterized with n parameters $p_1, p_2 \ldots p_n$ instead of one parameter, as is the case with fragility curves. Hence the probability of limit state exceedance is conditioned on the resulting set of critical parameters. The fragility function corresponding to the limit state l is defined as follows:

$$F_l(x_1, x_2, \ldots, x_n) = P_f(\text{LS} > \text{LS}_l | p_1 = x_1, p_2 = x_2, \ldots, p_n = x_n) \tag{5.9}$$

where LS is the limit state damage index, and LS_l is the value corresponding to the lth limit state.

Sample generation and fragility point estimates

This paper adopts a sampling strategy for generating point estimates of the fragility functions that draws upon the approach presented in multiple stripe analysis (MSA). However, in this study, both the ground motion and model intensity ranges of parameters are stratified, and rather than conduct nonlinear dynamic analyses, the meta-models are used for approximating the seismic response.

To this end, the selected IMs and MPs to generate the fragility surface are divided in N and M intensity levels, respectively, and samples are generated as shown in Figure 5.2. While keeping one parameter constant, the other is varied among the different levels, and its response is approximated with the meta-model. The fragility point estimate is calculated as the number of samples with a specific IM and MP intensity level that exceed a determined limit state over the total number of samples generated with those specific IM and MP.

Parametric fragility surfaces

While fragility curves are usually represented by well-known and readily parameterizable probability distributions such as the log-normal one, the problem is more complex for surfaces, where bivariate distributions must be computed. To fit an analytical function to the fragility point estimates, within this MSA approach, the methodology proposed by
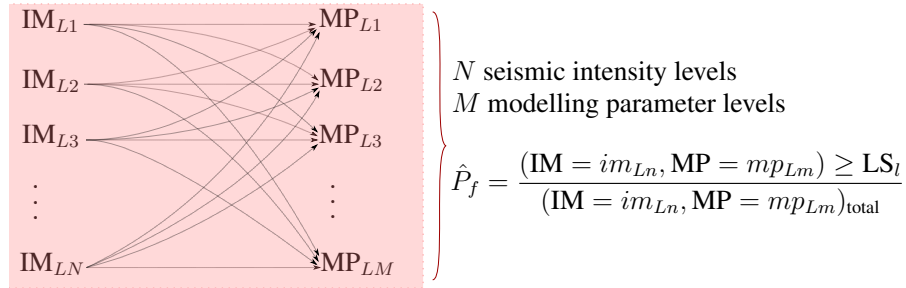
Figure 5.2 Multiple stripe analysis for meta-model fragility point estimate generation

Baker [68] to generate fragility curves and the methodology proposed by Brandenberg et al. [91] to generate fragility surfaces are combined for the first time. This joint procedure is further modified to make it suitable for the generation of fragility surfaces from the meta-model results, as a function of seismic IMs and model parameters, as shown in Figure 5.3. The steps involved in the construction of the parametric fragility surfaces can be summarized as follows:

1. For each limit state, l, fit fragility curves (F_c) according to Eq. (5.10), as a function of the seismic IM, for each MP intensity level, mp_k.

$$F_c(\text{IM}, \text{MP} = mp_k) = \Phi_l(\text{IM}, \theta_k, \beta_k) \tag{5.10}$$

where Φ_l is a cumulative density function (CDF), and θ_k and β_k are the parameters characterizing the associated CDF.

2. Plot the values of parameters θ_k and β_k for each limit state and for each value of mp_k, and fit a functional form to the discrete data:

$$\hat{\theta} = \kappa(p_\theta, \text{MP}) \tag{5.11}$$

$$\hat{\beta} = \kappa(p_\beta, \text{MP}) \tag{5.12}$$

where $\hat{\theta}$ and $\hat{\beta}$ are the analytical expressions of the parameters characterizing the CDF as a function of MP, and p_θ and p_β are the regression coefficients and $\kappa(\cdot)$ is the fit-type function (polynomial, exponential, etc.) selected for each MP.

3. Substitute Eq. (5.11)–(5.12) into Eq. (5.10) to obtain the analytical expression of the fragility surface (F_s) as a function of IM and MP for each limit state, as shown

in Eq. (5.13).

$$F_s(\text{IM}, \text{MP}) = \Phi_l(\text{IM}, \hat{\theta}, \hat{\beta}) = \Phi_l\left(\text{IM}, \kappa(p_\theta, \text{MP}), \kappa(p_\beta, \text{MP})\right) \qquad (5.13)$$

Three different CDFs, (i) normal, (ii) log-normal and (iii) Weibull, were tested, and the one with the best performance in each case was selected. The parameters of Eq. (5.10) were estimated with the maximum likelihood estimation (MLE) method and a Newton-Raphson optimization technique, as recommended by Baker [68] when working with the MSA method.

In the following sections, an application of the proposed methodology to assess the seismic vulnerability of a concrete gravity dam case study is presented, where the effect of the different model parameters influencing the seismic response of the dam is explicitly considered in the fragility analysis.
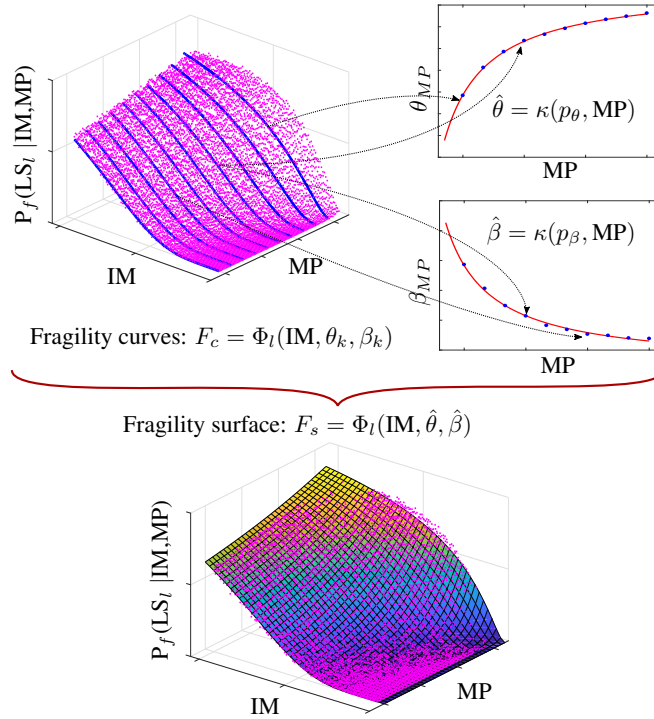
Figure 5.3   Parametric fragility surface construction

5.3   Case study: description and modeling

The proposed methodology in this study is applied to a case study gravity dam in Quebec, Canada. It possesses 19 unkeyed monoliths, a maximum crest height of 78 m, and a crest length of 300 m. The dam was chosen for its simple and almost symmetric geometry,

its well documented dynamic behavior and the availability of forced vibration test results used to calibrate the dynamic properties of the numerical model [22].

5.3.1   Finite element model

The tallest monolith of the dam is selected as representative of the structure and was modeled with the explicit finite element software LS-Dyna [135], as shown in Figure 5.4, following the recommendations of the United States Bureau of Reclamation (USBR) [37]. Only one load case combination was considered, which included self-weight, hydrostatic thrust, uplift, hydrodynamic effects and seismic load. The proposed model takes into account the different interactions among the structure, reservoir, and foundation. The reservoir is modeled with compressible fluid elements, whereas the concrete dam and the rock foundation are modeled with linear elastic materials to which a viscous damping is associated. Preliminary linear analyses identified the concrete-rock interface at the base of the dam and the concrete-concrete interface at the crest of the dam as high areas of tensile stresses and, therefore, where cracking and sliding could occur. Consequently, the model nonlinearity was constrained to these two areas only, using tiebreak contact elements with a tension-shear failure criterion. Further details of the modeling assumptions and the validation of the numerical model can be found in Bernier et al. [20].

Figure 5.4   Finite element model of the case study dam

5.3.2 Model parameter uncertainty

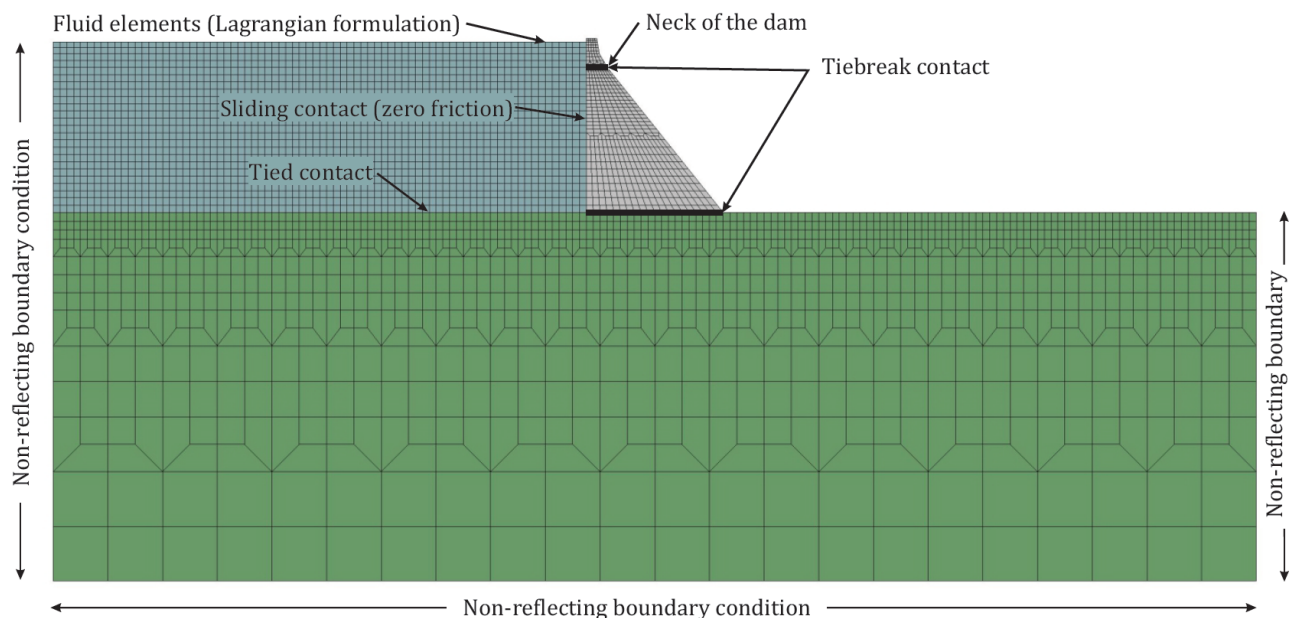Table 5.1 presents the parameters that were considered random variables in the numerical analysis of the dam response and for which the uncertainty was formally included through their probability distribution functions (PDFs). All of the remaining input parameters were kept constant and represented by their best estimate values. For the case study dam, due to limited availability of material investigations, the probability distributions were defined using empirical data of similar dams. The uniform distribution was used for most parameters except for damping, for which a log-normal distribution was adopted as proposed by Ghanaat et al. [83]. By posing the resulting meta-models and fragilities as functions of MPs that have a significant impact on the behavior of the dam, future studies may integrate these models with emerging estimates of these parameters or updated PDFs.

Table 5.1 Parameter distributions considered for statistical design of experiments

Parameters	PDF	Distribution Parameters	
Concrete-to-rock tensile strength (MPa)	Uniform	L = 0.2	U = 1.5
Concrete-to-concrete tensile strength (MPa)	Uniform	L = 0.3	U = 2.0
Concrete-to-rock cohesion (MPa)	Uniform	L = 0.3	U = 2.0
Concrete-to-concrete cohesion (MPa)	Uniform	L = 0.9	U = 2.5
Concrete-to-rock angle of friction (°)	Uniform	L = 42	U = 55
Concrete-to-concrete angle of friction (°)	Uniform	L = 42	U = 55
Drain efficiency (%)	Uniform	L = 0.0	U = 66
Concrete damping (%)	Log-normal	λ = -2.99	ζ = 0.35

5.3.3 Damage limit states

In recent years, typical damage modes that could lead to the potential collapse in dams after a seismic event have been identified, and seismic damage levels can be established. Preliminary analyses have confirmed sliding as the critical failure mode for the case study dam [20], and other failure modes would only occur after sliding had already been observed. As a result, base sliding at the concrete-to-rock interface damage limit states proposed in Tekie and Ellingwood [14] is considered in this study and presented in Table 5.2. The incipient sliding limit state should only cause minor damage since well-dimensioned dams should be able to undergo slight deformations or displacements while remaining stable. The moderate damage limit state can be considered as the onset of nonlinear behavior where material cracking occurs, deformations may become permanent and the drainage system begins to be affected. Displacement greater than 50 mm could cause differential movements between the blocks and potentially lead to the loss of control of the reservoir

and very extensive damage, while displacements greater than 150 mm represent a complete
damage state and high probability of collapse of the structure.

Table 5.2 Limit states considered for
the case study dam

Limit state	Base sliding, $\delta_{\max}$
LS0 - Slight/minor	5 mm
LS1 - Moderate	25 mm
LS2 - Extensive	50 mm
LS3 - Complete	150 mm

5.3.4 Seismic hazard and ground motion selection method

A probabilistic seismic hazard analysis (PSHA) was performed at the dam site with the
computer software OpenQuake [141] to characterize the possible earthquake scenarios at
different intensity levels. The hazard levels were defined in terms of horizontal spectral
acceleration at the fundamental period of the structure ($Sa_H(T_1) = 0.1\,\mathrm{g} : 0.1 : 1.0\,\mathrm{g}$)
to conveniently cover the range of spectral accelerations corresponding to return periods
from 700 to 30000 years.

To proceed with the selection of a representative set of ground motion time series (GMTS),
the generalized conditional intensity measure (GCIM) approach [74] was adopted. The
purpose of using the GCIM approach is to include the most influencing seismic IMs with
respect to the structural response. For the case of gravity dam-type structures, PGV was
found to be one of the best performing ground-motion structure-independent scalar IM
to correlate damage [30]. Similarly, the vertical spectral acceleration Sa_V is also expected
to be relevant in heavy structures of this sort. As a result, the set of considered IMs in
the GCIM are $Sa_H(T)$, $Sa_V(T)$ and PGV, where $Sa_H(T)$ and $Sa_V(T)$ are computed at
20 vibration periods in the range of $T = \{0.2T_1 - 2T_1\}$, as proposed by [73], leading to a
total of 41 IMs to be considered in addition to the conditioning IM, $Sa_H(T_1)$. The GCIM
distribution computed with the abovementioned IMs was then used to simulate and select
250 ground motions. The records were selected from the PEER NGA-West2 database
[159] due to the limited availability of strong ground motion records in the PEER NGA-
East database [160]. Further details on the PSHA and the record selection procedure can
be found in Segura et al. [87].

5.3.5   Most viable meta-model for maximum relative base sliding prediction

By selecting different configurations of the model parameters in Table 5.1, $n_f = 250$ samples of the FEM were generated with LHS and paired with the selected ground motions. The maximum relative sliding at the base $(\delta_{\max})$ was computed from nonlinear simulations, and 14 regression meta-models were fitted to the structural response. An initial pre-screening of the covariates or predictors was made before selecting the algorithm, based on visual inspection of the scatter plots of the possible predictors with respect to the response of the structure and taking into account the parameters affecting the dynamic behaviour of the structure. This initial set of predictors (input) was used to train all the considered meta-models to perform a comparative analysis. In addition to the model parameters listed in Table 5.1, several seismic intensity measures were considered in the starting set of variables, such as spectral acceleration at the fundamental period $(Sa_H(T_1))$, spectral velocity at the fundamental period $(Sv_H(T_1))$, peak displacement, velocity and acceleration (PGD, PGV, PGA), spectrum intensity (SI), earthquake angular frequency (ω_{eqk}), significant duration (D_{595}), Arias intensity (I_a), and peak ground acceleration and spectral acceleration at the fundamental period in the vertical direction $(PGA_V, Sa_V(T_1))$. Given that some algorithms already perform a internal selection of the predictors (with forward and backward iterations), the selected final set of predictors (output) in each meta-model is a sub-set of the initial set of predictors.

Comparison of meta-model predictive capabilities

The performance of the meta-models, namely, their ability to predict the sliding response of the dam, is judged based on the goodness-of-fit estimates shown in Table 5.3. In general, good performance of the adaptive-type meta-models and relatively poor performance of the kernel-based meta-models were observed, which could be explained due to the small training set. Provided that the number of samples considered to train the different algorithms can have en effect on the performance of the latter, Figure 5.6 presents the variation of the prediction performance of the considered meta-modeling techniques with respect to the size of the training set. In general it can been that the prediction accuracy from 5 fold cross-validation improves as the number of training samples increases and that all considered meta-models perform poorly with small training samples (< 50). However, it should be noted that for some meta-models, after a certain number of samples, the goodness-of-fit estimates remain almost constant or with a relatively low rate of improvement. Indeed, for these meta-models, the fact of increasing the size of the training samples beyond a certain

number will not translate in a significant improvement of the predicting capabilities of the algorithm. This is the case for ABFC, PRS $\mathcal{O}^{2-4}$, RFR and, to a lesser extent, MARS and RBF, where beyond 175 training samples the improvement of the predicting accuracy is relatively low. On the contrary, for SVMR with different kernel functions, it can be observed that the performance of the algorithm keeps progressively increasing as the number of samples increases. Moreover, although the local and global performance of SVMR with linear and quadratic kernels is reasonable, it should be mentioned that approximately 72% of the samples in the dataset were used as support vectors in the multidimensional feature space to find hyperplane that separates all given samples. This is the result of a highly non-linear feature space and a larger sample should be use to validate this type of model.

Goodness-of-fit estimates were calculated to evaluate how closely the meta-model's predicted values match the FEM simulation (true) values, considering the whole dataset for training and testing. Despite the fact that some algorithms showed superior performance, it is also known that some of them tend to overfit the data. To this end, the meta-models were trained and validated using 5-CV, and the performance was evaluated by calculating the mean of the goodness-of-fit estimates obtained from each fold. Consequently, by comparing the estimates from the algorithm trained and validated with the whole dataset with the ones from cross-validation, it is possible to identify model overfitting. Figure 5.5 presents the comparison of the goodness-of-fit estimates of the meta-models trained with the entire training set and the average result from 5-CV. It is intended for these indicators to be as close as possible to ensure that the meta-model can predict accurately for the cases that were trained with as well as for the unknown cases, i.e. to prevent overfitting. From Figure 5.5, it can be seen that in terms of RMSE and R^2, the PRS meta-models present similar values for both cases. Similarly, in terms of RMAE, the ABFC and the MARS surrogates, for which the estimates were calculated with the entire training set, are very close to those from 5-CV. On the other hand, the SVMR behavior suggests overfitting of the meta-model, given their large capacity to describe the training dataset but failing to predict beyond it to unseen cases. The RFR meta-model presents fair results regarding the R^2 and RMAE, but the difference between the RMSE for the training data and the average 5-CV reflect some limitations to accurately predict the response of the system. From Figure 5.5 and Table 5.3, the best meta-model in terms of predicting capabilities is the 4th-order PRS (PRS $\mathcal{O}^4$), which is evident from the 5-CV results.

Polynomial response Surface Meta-model

From the meta-model comparison, the selected surrogate is a polynomial response surface of 4th-order (PRS $\mathcal{O}^4$) as a function of three model parameters, three seismic intensity

Table 5.3 Meta-model comparison for base sliding

Meta-model		RMSE	R^2	RMAE	5-CV RMSE	5-CV R^2	5-CV RMAE
PRS[†]	Order: 2	0.367	0.857	1.245	0.381	0.846	0.945
	Order: 3	0.324	0.889	1.156	0.342	0.876	0.923
	Order: 4	0.319	0.907	1.024	0.321	0.887	0.898
ABFC[†]		0.308	0.900	1.149	0.360	0.860	1.040
MARS[†]	Linear	0.281	0.916	0.920	0.408	0.810	-
	Cubic	0.298	0.905	1.040	0.383	0.841	-
RBF[*]	Multiquad	-	-	-	0.355	0.866	1.137
	Thin plate	-	-	-	0.373	0.846	1.165
	Gaussian	-	-	-	0.468	0.767	1.310
SVMR[‡]	Linear	0.361	0.862	1.243	0.379	0.843	0.956
	Quadratic	0.302	0.904	1.184	0.405	0.824	1.136
	Cubic	0.267	0.924	1.056	0.859	0.213	3.582
	RBF	0.194	0.959	0.775	0.651	0.554	1.483
RFR[**]		0.240	0.939	0.987	0.406	0.824	1.113

[†] Adaptive algorithm; [*] Interpolation scheme; [‡] Kernel-based algorithm; [**] Decision trees-based algorithm

measures and their transformations and pairwise products.

$$\boldsymbol{\delta}_{\max} = g(\mathbf{X}) = g(\overbrace{\text{CRF}, \text{DR}, \text{CRC}}^{\text{model parameters}}, \overbrace{\text{PGV}, I_a, \text{PGA}_V}^{\text{seismic IMs}}) + \mathbf{v} \qquad (5.14)$$

$$\mathbf{v} \sim \mathcal{N}(v; 0, \sigma_v^2) \qquad (5.15)$$

where CRF, DR and CRC are the model parameters corresponding to the concrete-rock angle of friction, drain efficiency and concrete-rock cohesion, respectively; PGV is the peak ground velocity; PGA_V is the peak ground acceleration in the vertical direction; and I_a is Arias intensity. A normally distributed model error term $\mathbf{v}$ with zero mean and standard deviation equal to the RMSE is added to the selected surrogate model to reflect the lack of fit [174, 96, 118].

The stepwise regression algorithm in MATLAB was formally implemented together with PRS to improve the task of "manually selecting" the predictors. This first filter was used because the PRS meta-model considers all the covariates present in the dataset. The algorithm starts with a constant term to predict the response. In the next step, one predictor is added to the model, and the performance of the model is evaluated based on the Bayesian information criterion (BIC). If the model performance improves, the added term is kept; otherwise, it is removed, and this process is repeated until all the proposed
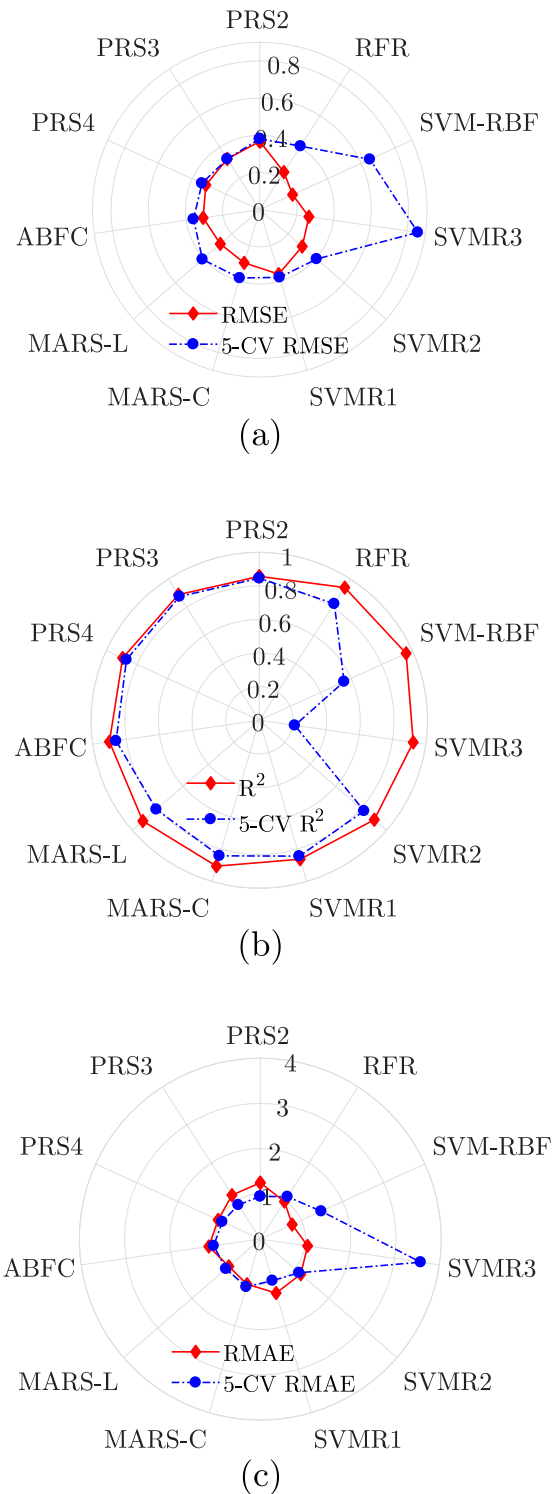
Figure 5.5   Goodness of fit - Training data vs. cross-validation average: (a) RMSE, (b) R^2 and (c) RMAE
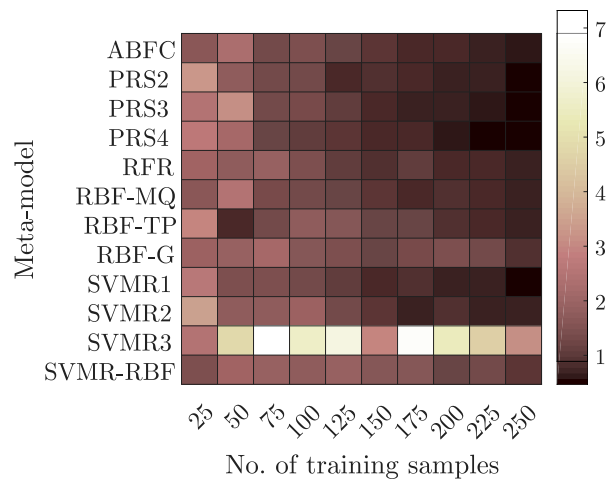
predictors are tested. As a result of the stepwise regression, the best predictors from each 5-CV fold were selected. A polynomial regression model considering the interaction between
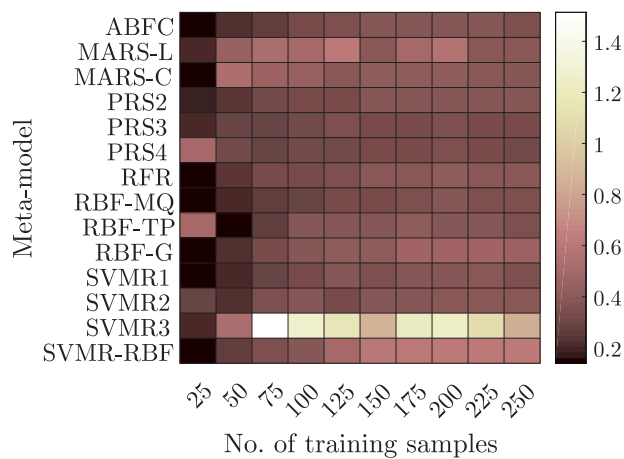
(a)
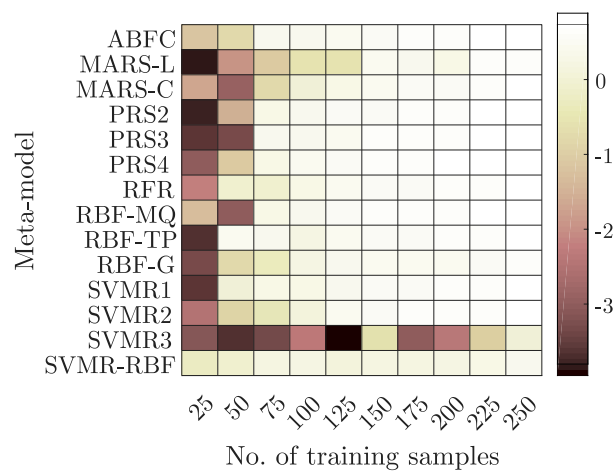
(b)

(c)

Figure 5.6   Prediction capabilities comparison vs. size of the training set: (a) 5-CV RMSE, (b) 5-CV R^2 and (c) 5-CV RMAE

the predictors was fitted to the dataset. In addition to the 5-CV method, which shows the overall accuracy of the model, p-values associated with the explanatory functions were controlled to be smaller than 0.05, assuring that the final terms included in the model are not selected by chance. Figure 5.7 (a) shows that the predicted values obtained with the selected meta-model are in agreement with the simulated dataset, while in Figures 5.7 (b) and (c), it can be seen that the residual normal distribution and the independence between the observation error hypothesis is respected.

(a)   (b)   (c)
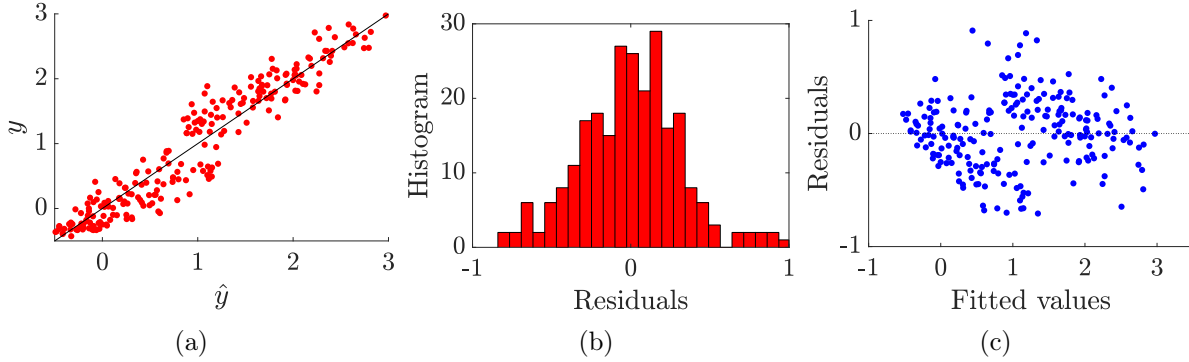
Figure 5.7 PRS $\mathcal{O}^4$: (a) Predicted vs. simulated values, (b) Residuals histogram and (c) Fitted value residuals

5.3.6   Dam sample generation and multivariate fragility functions

Regarding the generation of the samples where the meta-model will be evaluated to predict the dam's response and derive point estimates of fragility, independence among all the MPs was considered to generate 5×10^5 samples with LHS. Nevertheless, it should be noted that for a specific site, the seismic IMs are correlated. For the case study dam and considering the 250 GMTS used to train the meta-models, a linear correlation is to be expected, as seen from Figure 5.8. Figure 5.9 shows the histogram of the logarithm of the seismic IMs that follow an approximate normal distribution. Based on this, the samples of these parameters are taken from a jointly log-normal distribution with their respective correlation coefficients. In addition, to consider a range of values corresponding to return periods from 700 to 30000 years at the dam site, the possible values of PGV, I_a and PGA_V were bounded, respectively, as $0.8\,\mathrm{cm/s} \leq \mathrm{PGV} \leq 25\,\mathrm{cm/s}$, $0.0\,\mathrm{m/s} \leq I_a \leq 2.5\,\mathrm{m/s}$ and $0.01\,\mathrm{g} \leq \mathrm{PGA}_V \leq 0.25\,\mathrm{g}$.

The fragility analysis was performed following the MSA methodology depicted in Figure 5.2. The range of each of the parameters of the fragility surface was divided into 100 intervals. As a result, 10^4 fragility point estimates were generated for each limit state. Moreover, to display how the variation of the model parameters alters the seismic fragility,
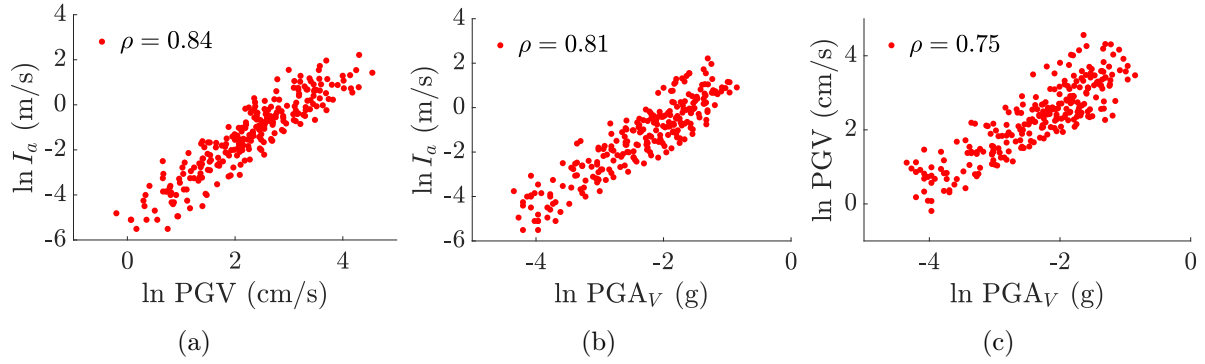
Figure 5.8   Linear correlation between the seismic IMs from the GMTS
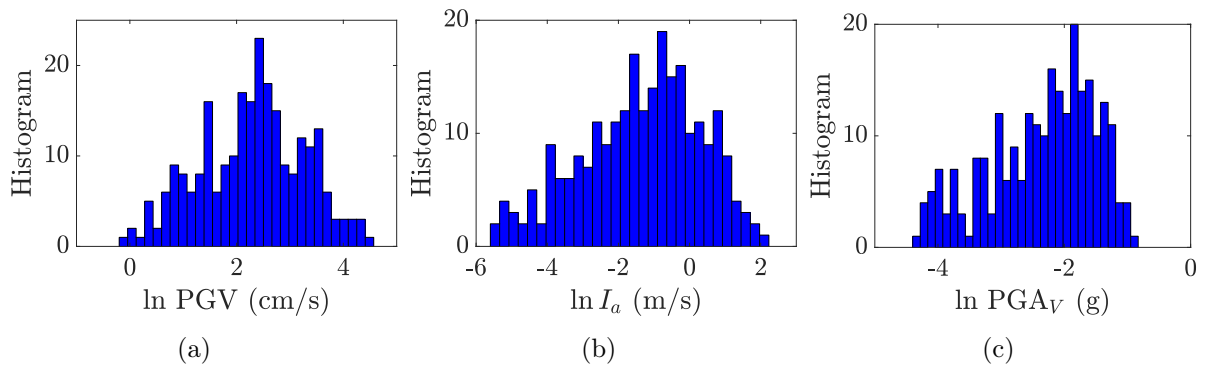
Figure 5.9   Probability distribution of seismic predictors

fragility surfaces as a function of PGV and each of the model parameters considered in the meta-model (CRF, DR and CRC) were generated. It is noteworthy that PGV was selected as the IM to be displayed in the fragility surface not only because of its practicality of database availability and value accessibility but also because, as previously mentioned, it was found to be one of the best performing ground-motion IM to correlate with the proposed damage state [30].

Parametric fragility surfaces for each limit state and for each MP were generated, implementing the methodology explained in Figure 5.3. The resulting fragility surfaces are depicted in Figures 5.10–5.12, and the goodness of fit between the proposed parameterized fragility surfaces and the fragility point estimates is presented in Table 5.4.
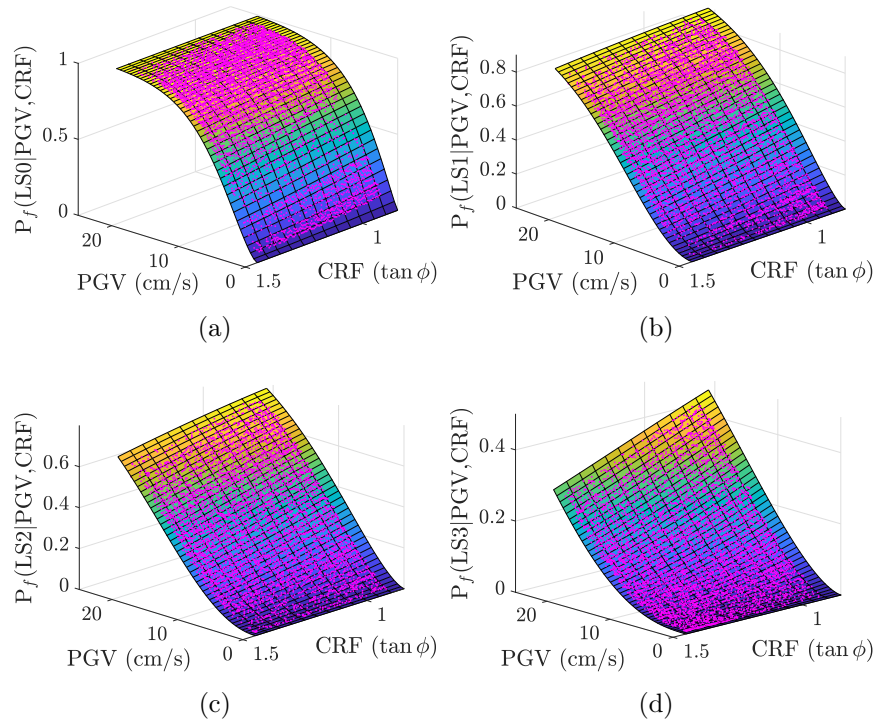
Figure 5.10   Fragility surfaces: F_s(PGV,CRF) for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

Table 5.4   Fragility surfaces' goodness of fit

Parameters	Limit state	R^2	RMSE	RMAE	5-CV R^2	5-CV RMSE	5-CV RMAE
PGV-CRF	LS0	0.997	0.013	0.207	0.991	0.014	0.215
	LS1	0.997	0.015	0.253	0.992	0.016	0.277
	LS2	0.996	0.014	0.314	0.991	0.015	0.364
	LS3	0.991	0.010	0.449	0.991	0.014	0.480
PGV-DR	LS0	0.997	0.013	0.217	0.990	0.014	0.291
	LS1	0.997	0.014	0.263	0.995	0.015	0.293
	LS2	0.996	0.014	0.252	0.994	0.016	0.285
	LS3	0.987	0.016	0.752	0.948	0.018	0.649
PGV-CRC	LS0	0.994	0.023	0.809	0.991	0.024	0.724
	LS1	0.997	0.017	0.288	0.993	0.018	0.295
	LS2	0.995	0.017	0.348	0.992	0.017	0.390
	LS3	0.987	0.015	0.803	0.973	0.016	0.882

Figure 5.11   Fragility surfaces: F_s(PGV, DR) for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

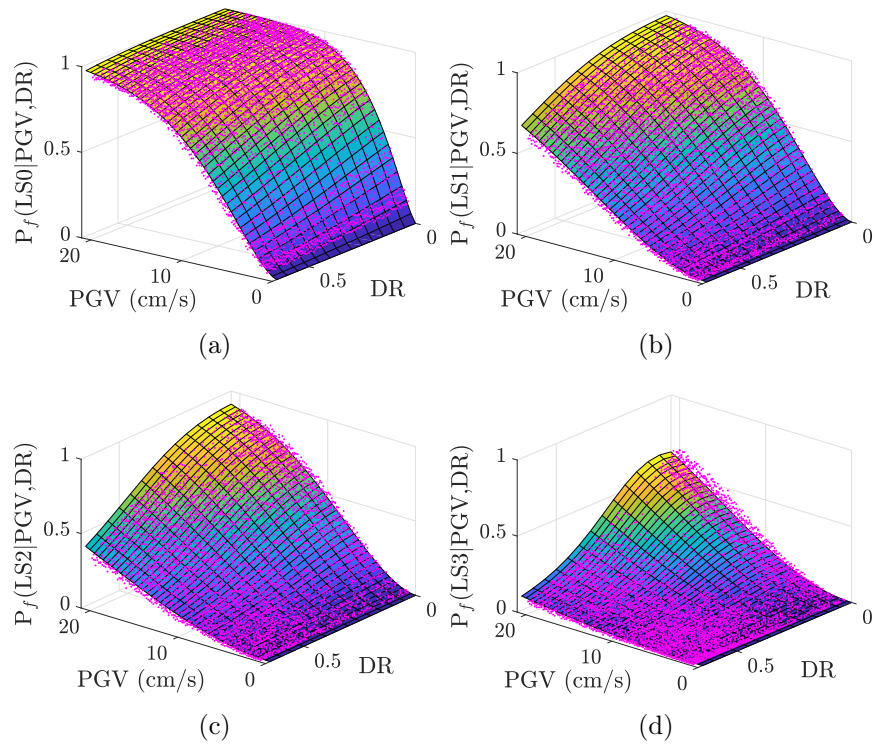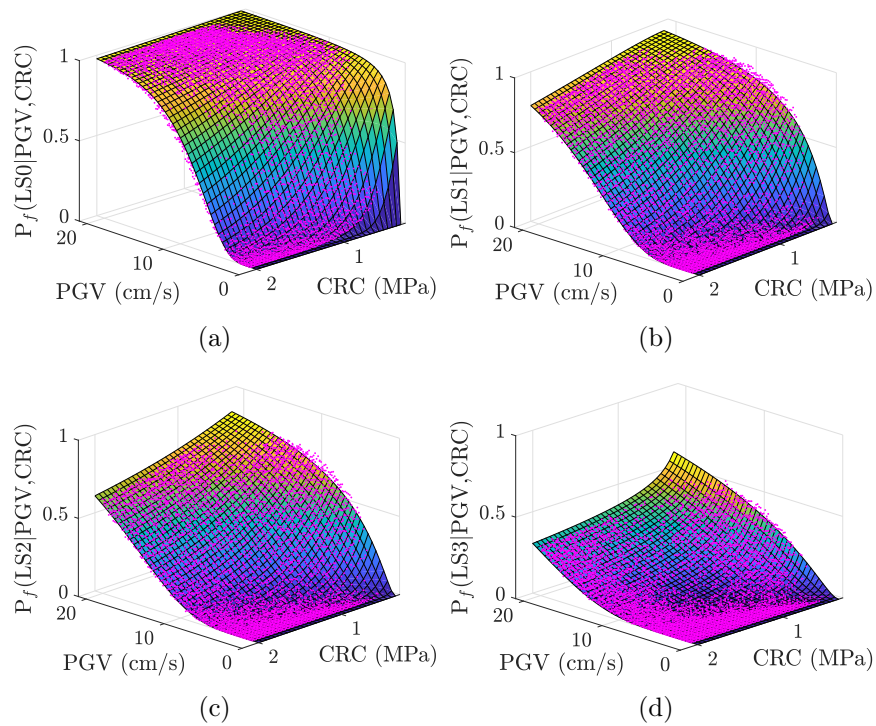**Figure 5.12**   Fragility surfaces: F_s(PGV, CRC) for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

5.3.7   Effect of model parameter variation in the fragility analysis

Concrete-rock angle of friction

Fragility curves as a function of PGV were calculated for each CRF intensity level, according to Eq. (5.10), where Φ_l in this case is the Weibull CDF. The parameters characterizing the CDF were plotted for each CRF value and for each limit state, as shown in Figure 5.13. A linear function was fitted to each parameter whose regression coefficients are shown in Table 5.5. Finally, the parametric fragility surface is depicted by Eq. (5.16),

$$F_s(\text{PGV}, \text{CRF}) = \text{CDF}\mathcal{W}(\text{PGV}, p_{\theta_1}\text{CRF} + p_{\theta_2}, p_{\beta_1}\text{CRF} + p_{\beta_2}) \tag{5.16}$$

where the CDF is also a Weibull distribution to be consistent with the fragility curves. As seen from Figure 5.10 and Table 5.4, the parametric fragility surfaces fit fairly well with the fragility point estimates.

Table 5.5   CRF regression coefficients

Limit state	p_{θ_1}	p_{θ_2}	p_{β_1}	p_{β_2}
LS0	2.619	3.548	0.471	0.799
LS1	4.928	7.753	0.727	0.820
LS2	6.358	10.795	0.920	0.741
LS3	6.836	21.918	1.412	0.439

Drain efficiency

The same procedure described in the above section was used to develop fragility surfaces as a function of PGV and DR; however, a quadratic fit-type function was used instead of a linear one, as seen in Figure 5.14. The parameterized formulation is described by Eq. (5.17);

$$F_s(\text{PGV}, \text{DR}) = \text{CDF}\mathcal{W}(\text{PGV}, p_{\theta_1}\text{DR} + p_{\theta_2}\text{DR}^2 + p_{\theta_3}, p_{\beta_1}\text{DR} + p_{\beta_2}\text{DR}^2 + p_{\beta_3}) \tag{5.17}$$

where $p_{\theta_{1,2,3}}$ and $p_{\beta_{1,2,3}}$ are the linear regression coefficients shown in Table 5.6. Figure 5.11 shows the resulting fragility surfaces.

Concrete-rock cohesion

Similarly, the fragility surfaces as a function of PGV and CRC were generated. In this case, as displayed in Figure 5.15, an exponential function was fitted to each parameter defining the fragility curves, and a log-normal CDF was used for the fragility functions.
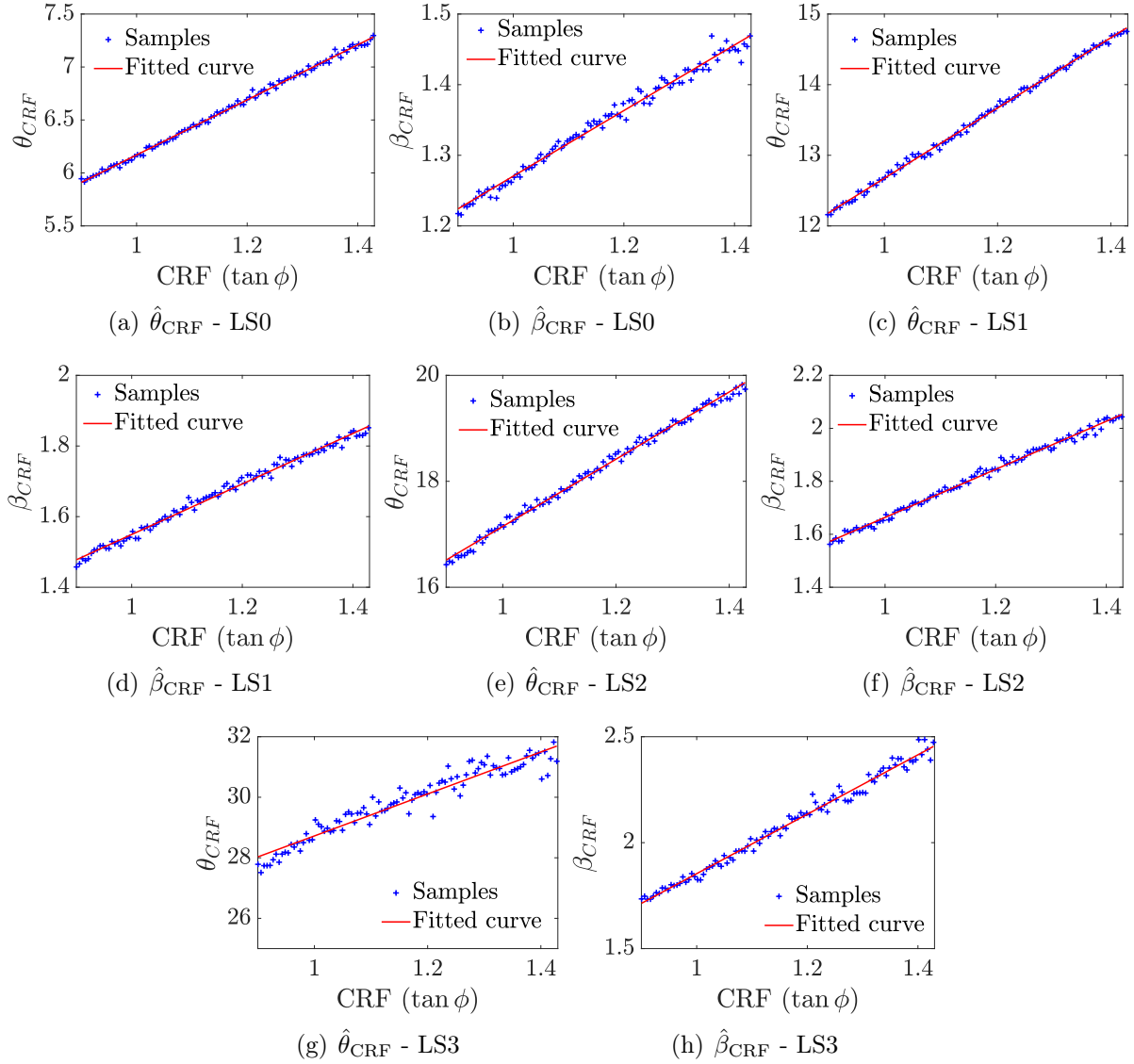
Figure 5.13 CRF fragility curves parameters regression

Table 5.6 DR regression coefficients

Limit state	p_{θ_1}	p_{θ_2}	p_{θ_3}	p_{β_1}	p_{β_2}	p_{β_3}
LS0	4.483	3.964	4.697	-0.027	0.274	1.314
LS1	16.701	5.163	9.621	-0.230	0.393	1.637
LS2	34.515	3.821	12.780	-0.433	0.454	1.782
LS3	123.314	-10.199	20.957	-0.713	0.432	2.033

The parametric fragility surfaces result in the following:

$$F_s(\text{PGV}, \text{CRC}) = \text{CDF} \ln \mathcal{N}(\text{PGV}, p_{\theta_1}\text{CRC}^{p_{\theta_2}} + p_{\theta_3}, p_{\beta_1}\text{CRC}^{p_{\beta_2}} + p_{\beta_3}) \tag{5.18}$$
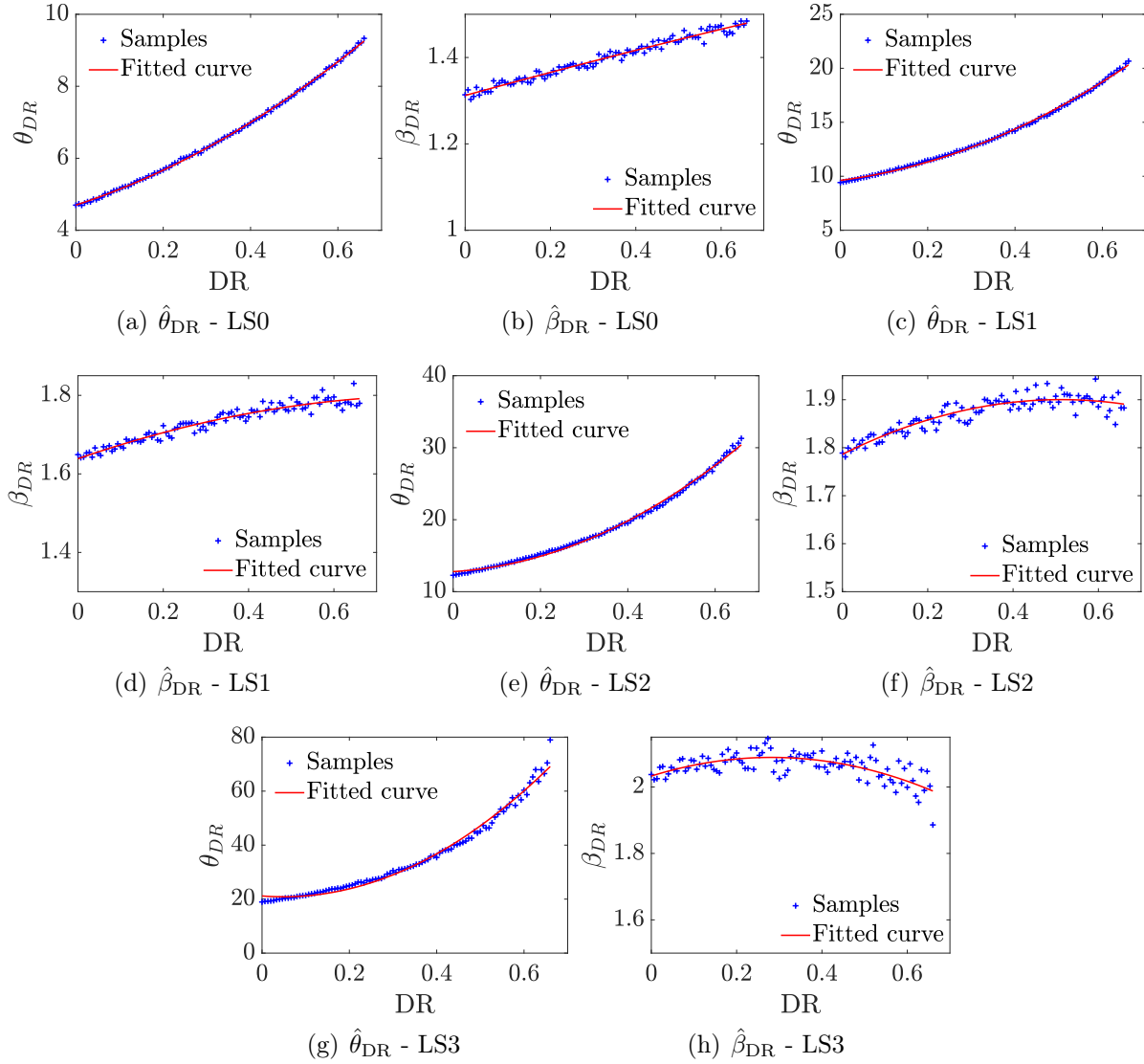
Figure 5.14   DR fragility curve parameter regression

where $p_{\theta_{1,2,3}}$ and $p_{\beta_{1,2,3}}$ are the linear regression coefficients presented in Table 5.7. The fragility surfaces are shown in Figure 5.12.

Table 5.7   CRC regression coefficients

Limit state	p_{θ_1}	p_{θ_2}	p_{θ_3}	p_{β_1}	p_{β_2}	p_{β_3}
LS0	-1.074	-1.079	2.619	0.378	-0.996	0.274
LS1	-0.928	-0.758	3.247	0.306	-1.049	0.310
LS2	-0.510	-0.977	3.177	0.341	-0.989	0.288
LS3	-0.025	-2.931	3.285	0.415	-0.934	0.249

(a) $\hat{\theta}_{\mathrm{CRC}}$ - LS0

(b) $\hat{\beta}_{\mathrm{CRC}}$ - LS0

(c) $\hat{\theta}_{\mathrm{CRC}}$ - LS1

(d) $\hat{\beta}_{\mathrm{CRC}}$ - LS1

(e) $\hat{\theta}_{\mathrm{CRC}}$ - LS2

(f) $\hat{\beta}_{\mathrm{CRC}}$ - LS2

(g) $\hat{\theta}_{\mathrm{CRC}}$ - LS3

(h) $\hat{\beta}_{\mathrm{CRC}}$ - LS3
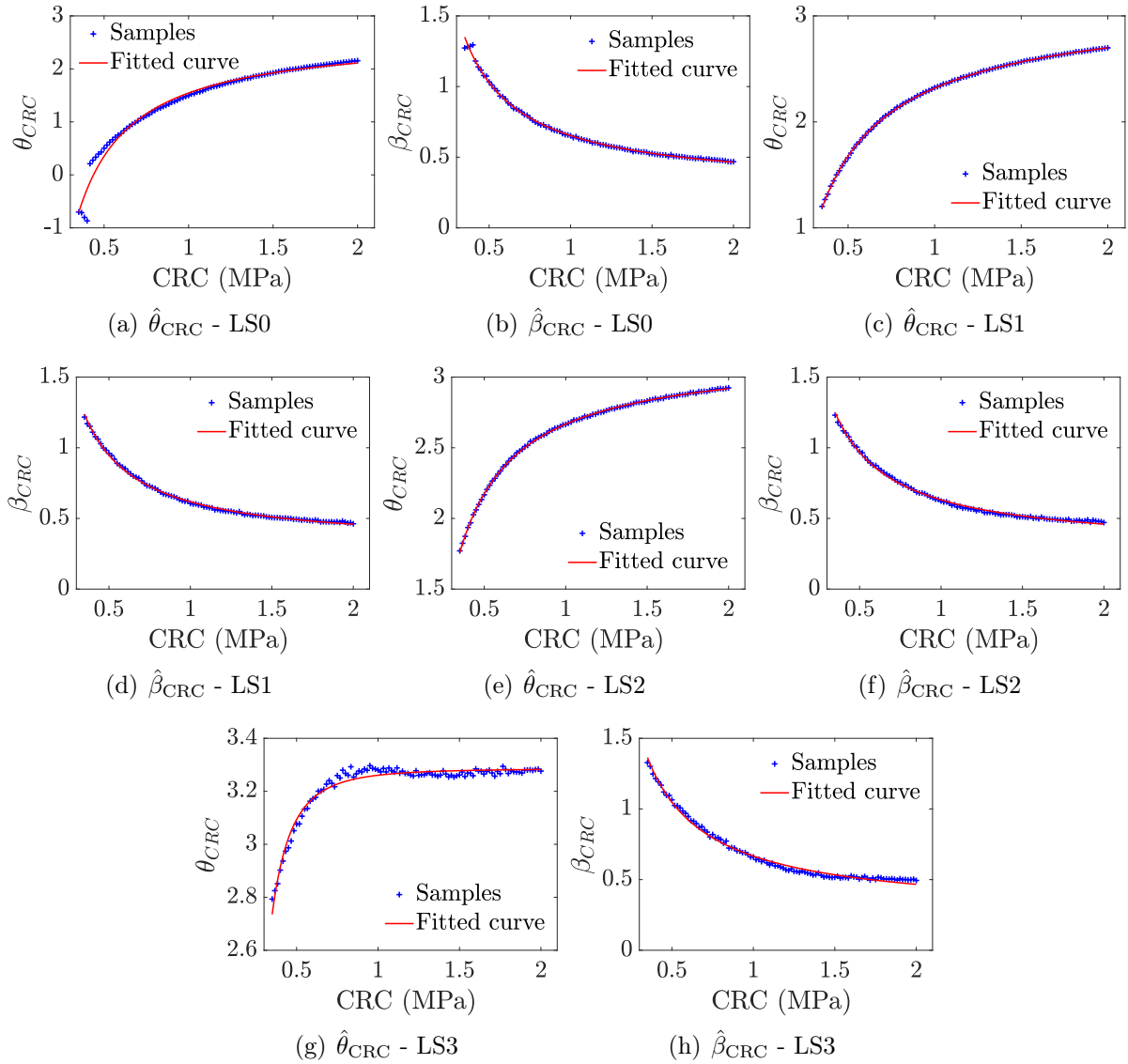
Figure 5.15 CRC fragility curve parameter regression

5.4 Application to seismic assessment of dams

The resulting fragility functions can also be used to assess the seismic performance of the dam by formulating recommendations with respect to the model parameters. To achieve a desired seismic performance, boundaries of model parameters for an adequate performance under extreme limit states can be formulated. The usable range of values is determined by ensuring that the probability of exceedance given that an extreme event is in line with the current guidelines for the minimum provisions for life safety. To establish the admissible values of the model parameters, a return period of 10000 years, which corresponds to the maximum considered earthquake (MCE) [148, 150, 31], was used. The peak ground

velocity associated with the MCE (PGV_{MCE}) was determined from the PSHA and from the fragility surfaces corresponding to the complete damage limit state (LS3), and the fragility curves as a function of the model parameters were extracted for the PGV_{MCE} value.

According to what it is proposed by the ASCE 7-16 [148], for a risk category III (dam-type structures) and total or partial structural collapse, the maximum probability of failure should be less than 6 %. Consequently, and as shown in Figure 5.16, the boundaries of model parameters were established to provide a probability of exceedance less than 6 % for an MCE seismic scenario. As a result, a range of values of $50° \leq CRF \leq 55°$, $0.35 \leq DR \leq 0.66$ and $0.87\,MPa \leq CRC \leq 2.0\,MPa$ should be considered to ensure that the performance of the dam is in line with the minimum values for life safety. Nonetheless, it should be acknowledged that these parameter ranges are derived from the single parameter at a time evaluation, and the joint interaction of the model parameters should be further examined.

(a)   (b)   (c)
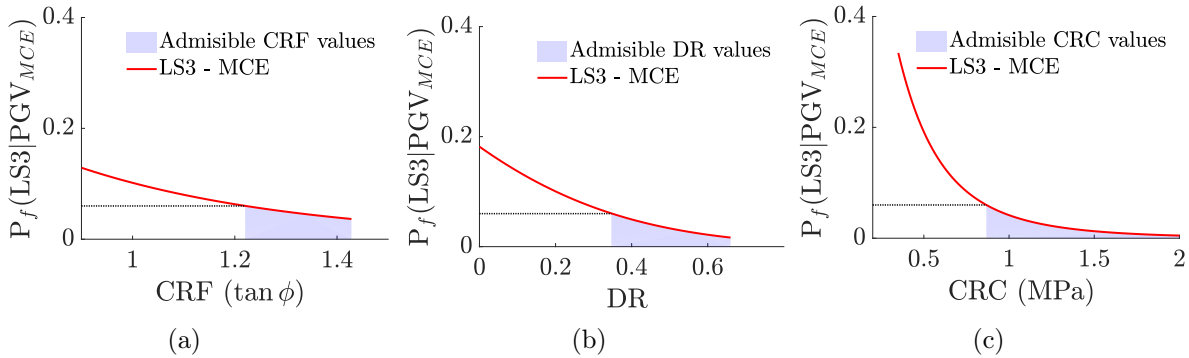
Figure 5.16 Admissible model parameter range of values for MCE events: (a) Concrete-rock angle of friction, (b) Drain efficiency and (c) Concrete-rock cohesion

5.5   Conclusions

The main objective of this study was to identify the most viable meta-model among those considered for the seismic fragility assessment of gravity dams, present a methodology for the development of multivariate fragility functions displaying the effect of the model parameter variation on the dam seismic performance and formulate design recommendations from the analysis.

The methodology was applied to a case study dam located in northeastern Canada. PSHA was performed to characterize the seismic hazard at the dam site and to select 250 ground

motions, consistent with the latter, using the GCIM approach. The sets of selected ground motion records were paired with the 250 samples of numerical models of the dam generated with LHS, representing different material and loading configurations of the system. The dataset used to train the meta-models was generated by performing nonlinear dynamic analysis of these samples, with an FEM that considered the fluid-structure-foundation interaction and by extracting the maximum relative sliding at the base of the dam. Six different types of meta-models with different configurations (basis and interpolation functions) were fitted to the seismic response of the case study dam to predict the sliding limit state at the base of the dam. The 4th-order PRS emerged as the best performing meta-model based on local and global goodness-of-fit estimates from 5-CV, and it was used to generate fragility surfaces as a function of PGV and each of the model parameters. It is observed that the variability of the concrete-rock cohesion model parameter has the most influence on the fragility analysis estimates, followed by the drain efficiency and to a lesser extent, the concrete-rock angle of friction. Finally, a seismic assessment was performed to determine the model parameter boundaries for adequate performance under extreme events, such as the MCE, to respect the minimum safety margins proposed by the current guidelines.

It should be noted that machine learning techniques are indispensable when assessing the vulnerability of structures with computationally expensive FEM such as gravity dams subjected to seismic loading. Similarly, the use of surrogate models allows for the exploration of the impact of different parameters in the fragility without the costly reevaluation of the FEM simulations. Regarding the fragility functions, as evident for the case presented herein and the goodness-of-fit values, the fragility estimates are well depicted by the methodology suggested in this study to fit parametric fragility surfaces.

It is expected that the results of this study can lead to more accurate planning and retrofitting policies to expedite the safety assessment of dams under seismic loads while supporting the decision-making process and to guide the preliminary design of future gravity dams. In future studies, additional insight into the correlation between the parameters defining the model configurations should be made, including other relevant limit states for gravity dams. Additionally, the model parameter variations in the fragility analysis should be further explored to provide parametric fragility functions, including the joint interaction of these parameters using classification meta-modeling techniques.

5.6 Acknowledgments

CHAPTER 6

SEISMIC FRAGILITY ANALYSIS OF DAMS: COMPLEMENTARY RESULTS

This chapter includes additional verifications and comparisons regarding the seismic fragility of the case study dam. These results, presented in conference papers [175, 176], are included in this document for completeness. As such, first an exhaustive comparison between the different record selection procedures considering the NBCC 2010 [177] and 2015 [177] hazard models, and their impact in the development of fragility curves is presented. In the subsequent sections, a polynomial response surface meta-model is used to predict the continuous maximum relative base sliding of the dam to build fragility curves and show the effect of the modeling parameters uncertainty. To do so, fragility regions are developed that represent the fragility curves generated with values corresponding to a 95 % confidence interval of the usable range of the model parameters values; this is done in order to identify the parameter that introduces the most uncertainty in the fragility analysis.

6.1 Ground motion record selection methods

As it was already explained in section 4.3.1, expected ground motion can be calculated on the basis of probability, and the expected ground motions are referred to as seismic hazard. Classical PSHA allows calculating the probabilities of exceeding, at least once in a given time span, and at a given site, of a set of ground motion parameter levels considering all possible earthquake ruptures defined in a seismic source model. Two different source models were considered, the ones provided by the NBCC 2010 and the NBCC 2015. Figure 4.4–4.5 showed not only that the 2015 model provides lower hazard estimates than the 2010 model but also that the 2015 model is more sensitive to the ERF configuration (meshing density). Consequently, it was intended to display the effect of the hazard model in the fragility analysis.

Concerning the different GMTS selection approaches, the CS, GCIM and UHS methods were used to calculate the target distributions, resulting in 4 different cases: (i) GCIM with the NBCC 2015 as presented in section 4.4; (ii) CS with the NBCC 2015; (iii) CS with the NBCC 2010 and (iv) UHS with the NBCC 2010. For cases (i)–(iii) 30 GMTS were selected

from the NGA-WEST2 database for each intensity level ranging from $0.1\,\mathrm{g}$–$0.9\,\mathrm{g}$ covering return periods from 700–30000 years. As such, the correlation equations proposed by Baker and Jayaram [155] for western North America (WNA) were herein assumed for eastern Canada, considering that these models are suitable for shallow crustal events, which are the type of events mostly expected in ENA. To select ground motions with the CS and the NBCC 2015, the exactly same procedure and GMMs as for the GCIM with the NBCC 2015 were used, except that only horizontal spectral accelerations were considered in the target conditional distribution. Similarly, for the CS with the NBCC 2010 the methodology presented in section 3.2.2 was used together with the GMMs proposed by Atkinson and Boore [178]. For case (iv), 20 ground motions from the NGA-WEST2 database with the smallest sum of squared errors with respect to the UHS presented in Figure 6.1 between periods of 0.1 and $1.0\,\mathrm{s}$, were selected.
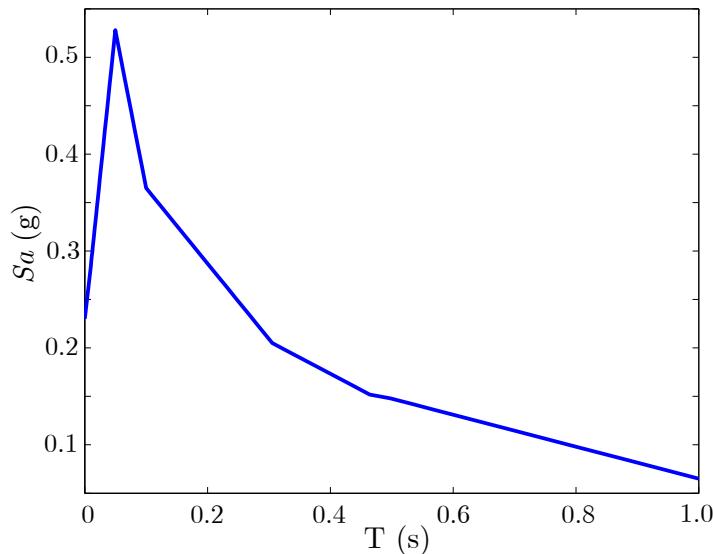
Figure 6.1   NBCC 2010 uniform hazard spectra

6.2   Fragility curves comparison

For each intensity level, 30 non-linear dynamic analyses were performed for cases (i)–(iii) and for case (iv) the values of the fragility curves were extracted from Bernier et al. [21], where the same case study dam is analyzed. The FEM is the one described in section 3.1.1. From each simulation, the maximum relative base and crest displacement are computed. These results were used to calculate 7 fragility point estimates as the number of samples where sliding exceeded the limit state divided by the total number of samples. A log-normal cumulative distribution function was then fitted to these fragility estimates

to develop fragility curves for the considered limit states, together with the MLE fitting technique. The fragility curves developed with the four approaches considered herein are illustrated and compared in Figures 6.2–6.3 for the four levels of damage of each limit state. Tables 6.1–6.2 also presents the parameters of the log-normal distribution for each fragility curve. As it was expected, from Figures 6.2–6.3, as the selection method refinement improves the point estimates of the fragility are less conservative. Under this perspective, the methods can be listed, from the most conservative to the least conservative as: UHS - NBCC 2010, CS - NBCC 2010, CS - NBCC 2015 and GCIM - NBCC 2015. Correspondingly, from tables 6.1–6.2, it can be seen that the mean and dispersion parameters of the curves are significantly higher with the GCIM - NBCC 2015 method, which will result in lower fragility, for a same IM level. Accordingly, for a given IM level, a lower mean response and a lower corresponding dispersion will yield a higher probability of reaching a damage limit state. This highlights how using less refined selection techniques, such as the UHS, and neglecting the spectral shape change of the target ground motions at the different intensity levels can lead to a less accurate estimation of the probability of reaching the considered damage limit states. These results are in line, with the studies related thereto, found in the literature [71, 21, 87].

Similarly, to evaluate the seismic performance of the case study dam under extreme limit states, the return period boundaries for the hazard classification provided by the CDA were used as explained in section 4.6.1. As it was expected in Tables 6.3–6.4, for the cases where the records where selected with the 2015 hazard model (GCIM and CS) and for both limit states, the probability of exceedance is less than 6 %. On the contrary, the results obtained with the 2010 hazard model (CS and UHS) show that for both limit states, the probability of exceedance is well above the maximum value and is even higher for the UHS.

Table 6.1 Comparison of the fragility parameters between the records selection approaches for base sliding

Limit State	GCIM - 2015		CS - 2015		CS - 2010		UHS - 2010	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LS0	0.458	0.531	0.404	0.640	0.390	0.612	0.315	0.483
LS1	0.758	0.545	0.741	0.647	0.700	0.595	0.421	0.323
LS2	0.982	0.584	0.990	0.672	0.921	0.623	0.522	0.311
LS3	1.512	0.648	1.631	0.745	1.203	0.483	0.831	0.390

Figure 6.2 Fragility curves comparison for base sliding damage state for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

Table 6.2 Comparison of the fragility parameters between the records selection approaches for neck sliding

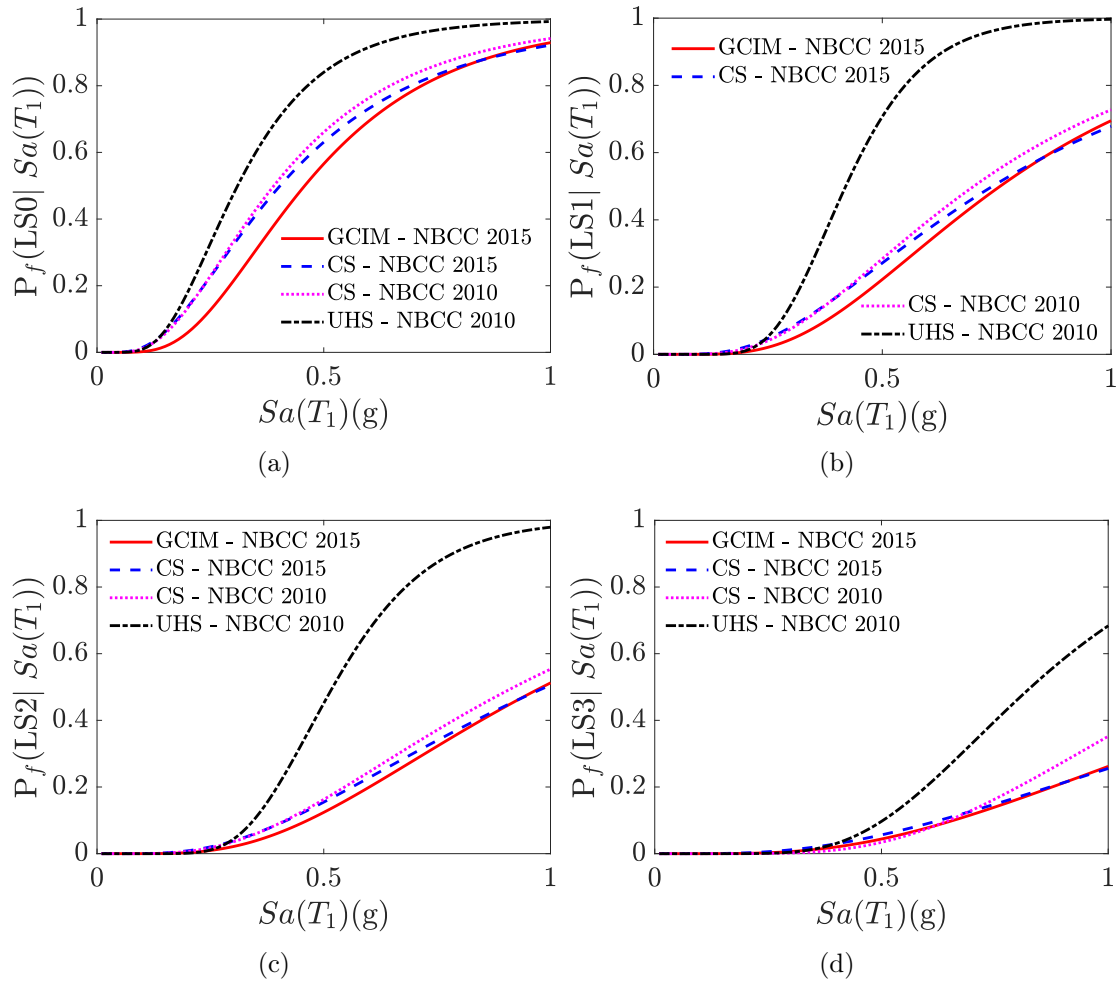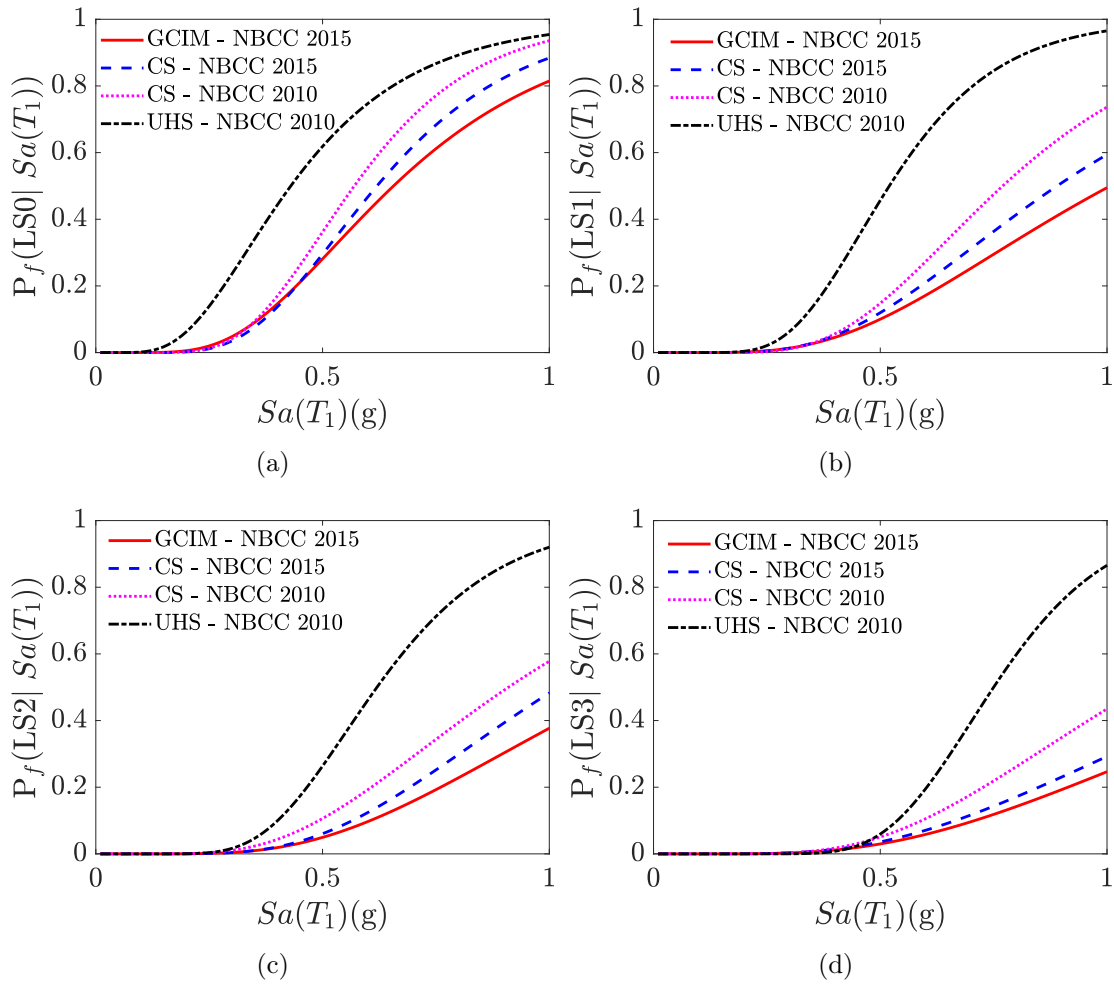| Limit State	GCIM - 2015		CS - 2015		CS - 2010		UHS - 2010	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LS0	0.656	0.469	0.621	0.401	0.573	0.371	0.430	0.501
LS1	1.007	0.547	0.893	0.493	0.771	0.413	0.521	0.363
LS2	1.176	0.518	1.022	0.459	0.912	0.482	0.620	0.342
LS3	1.487	0.578	1.353	0.553	1.080	0.473	0.753	0.261

Figure 6.3 Fragility curves comparison for neck sliding damage state for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

Table 6.3 Comparison of the probability of exceedance of the base sliding limit states set by the CDA

Hazard	CDA boundaries	GCIM - 2015	CS - 2015	CS - 2010	UHS - 2010
		P_r, Neck	P_r, Neck	P_r, Neck	P_r, Neck
Low	$P_e \geq 1/100$	<0.01	<0.01	<0.01	<0.01
Significant	$1/1000 < P_e < 1/100$	<0.01	<0.01	<0.01	<0.01
High	$P_e \geq 1/2475$	<0.01	<0.01	<0.01	<0.01
Very high	$1/10^4 < P_e < 1/2475$	<0.01	<0.01	1.31×10^{-2}	3.66×10^{-2}
Extreme	$P_e \leq 1/10^4$	3.59×10^{-2}	4.76×10^{-2}	9.73×10^{-2}	2.58×10^{-1}

Table 6.4 Comparison of the probability of exceedance of the neck sliding limit states set by the CDA

Hazard	CDA boundaries	GCIM - 2015	CS - 2015	CS - 2010	UHS - 2010
		P_r, Neck	P_r, Neck	P_r, Neck	P_r, Neck
Low	$P_e \geq 1/100$	<0.01	<0.01	<0.01	<0.01
Significant	$1/1000 < P_e < 1/100$	<0.01	<0.01	<0.01	<0.01
High	$P_e \geq 1/2475$	<0.01	<0.01	<0.01	<0.01
Very high	$1/10^4 < P_e < 1/2475$	<0.01	<0.01	2.11×10^{-2}	1.05×10^{-2}
Extreme	$P_e \leq 1/10^4$	2.34×10^{-2}	2.82×10^{-2}	1.37×10^{-1}	2.76×10^{-1}

It is also noted that the fragility curves obtained with the GCIM - NBCC 2015 are close to those obtained with the CS - NBCC 2015. However, it should be taken into account that even if the source models and the GMMs are the same, the records selected considering only horizontal spectral acceleration as the IMs are less representative of the seismic scenario at the dam site and neglect the analytical distribution of other IMs that are relevant for the dam's seismic response and for the considered damage states (Figure 4.7). In fact, this highlights that the effort in better characterizing the seismic hazard and selecting appropriate ground motion records pays off more than simplifying the seismic load input leading more conservative estimates which can have a direct impact in the application of safety guidelines and eventual decision making.

6.3   Fragility functions from meta-models

6.3.1   Fragility curves

Fragility curves were also generated for the case study dam as a function of PGV, as presented in Figure 6.4. The analytical fragility curves were developed using the multiple stripes analysis [71]. To compute the parametric fragility function, a Weibull cumulative

distribution function was fitted to the data points according to Eq. (6.1), in which $\beta_{\mathcal{W}} > 0$ is the shape parameter and $\mu_{\mathcal{W}} > 0$ is the scale parameter.

$$P_f(\text{PGV}) = 1 - \exp[-(\text{PGV}/\mu_{\mathcal{W}})^{\beta_{\mathcal{W}}}] \tag{6.1}$$

The fragility curves were generated in the same manner as the fragility surfaces in section 3.4.3 for the correlated seismic IMs, but in this case the uncertainty in the modeling parameters, was propagated in the analysis by considering a range of possible values as shown in Table 4.3.
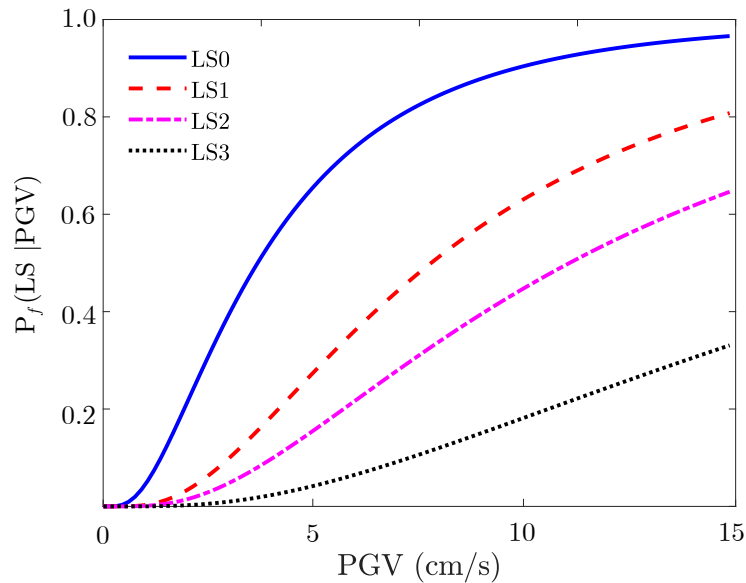
Figure 6.4   Fragility curves from the meta-model

In the same manner, and to explicitly account for the variability in the model parameters, Figure 6.5 shows the fragility curves for each limit state in red, and the shaded areas represent the fragility curves generated with the values corresponding to the 95 % confidence interval of the usable range of the model parameters values. A Weibull CDF was used in all cases except for the concrete-rock cohesion, where a log-normal CDF provided better goodness-of-fit estimates. It can be seen that the concrete-rock cohesion variability introduces the most uncertainty in the fragility analysis, followed by the drain efficiency and to a lesser extent the concrete-rock angle of friction.

6.3.2   Fragility surfaces: uncorrelated IMs

As it was stated in section 3.4.3 for the uncorrelated cases, to consider more general cases, fragility surfaces were generated where the correlation between the seismic IMs was neglected. The fragility point estimates were generated as a function of PGV and I_a, the
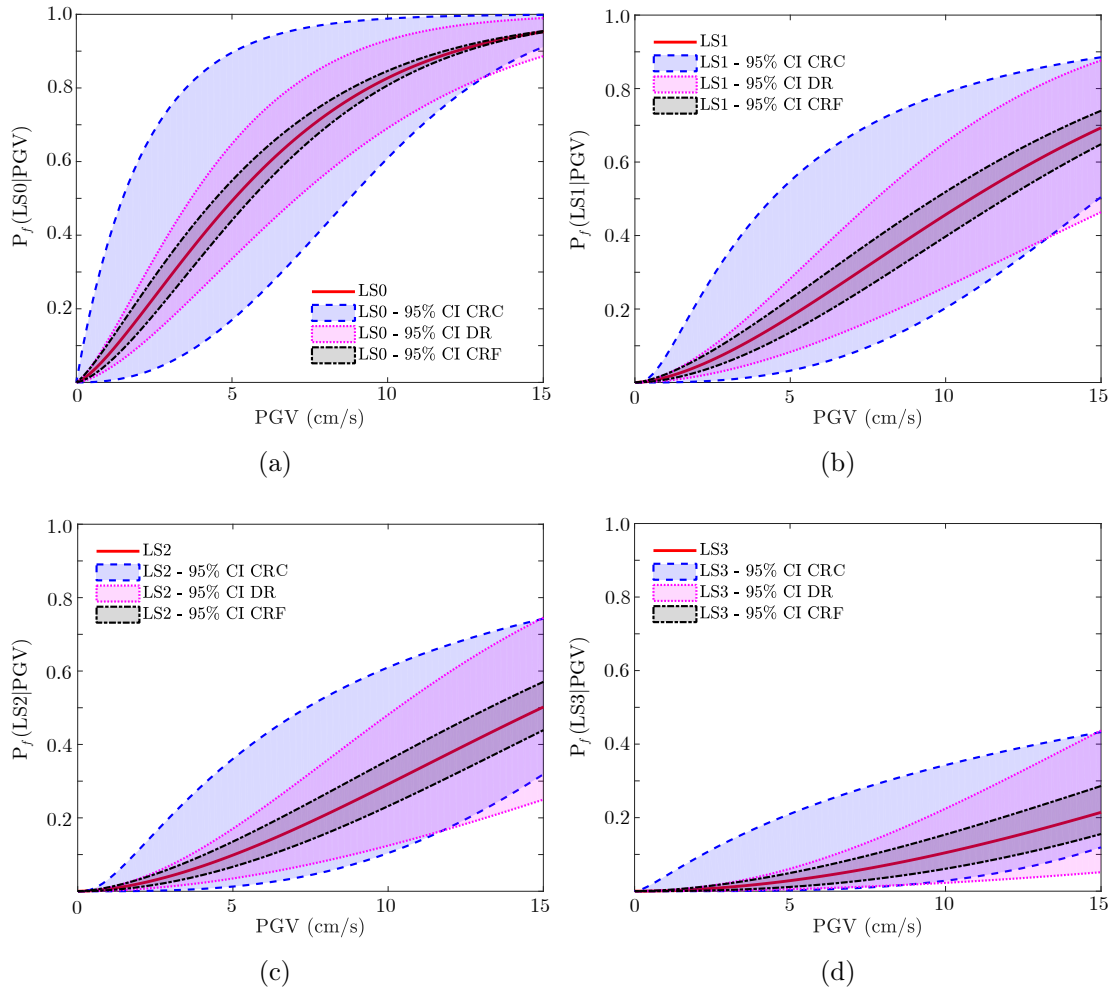
Figure 6.5   Fragility curves from meta-models for base sliding and 95 % confidence interval of the modeling parameters values for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

two most influential seismic IMs. The uncertainty due to the modeling parameters and PGA_V is propagated in the analysis by sampling these parameters with LHS according to their respective range of usable values (Table 4.3 and Table 3.4).

It was assumed that for each limit state in Table 4.2, the fragility surface is described by the product of two cumulative density functions, as shown in Eq. (6.2):

$$F_s(\mathrm{PGV}, I_a) = \Phi_{\mathrm{PGV}}(\mathrm{PGV}, \theta_{\mathrm{PGV}}, \beta_{\mathrm{PGV}})\Phi_{Ia}(I_a, \theta_{Ia}, \beta_{Ia}) \qquad (6.2)$$

where after testing several distributions, Φ_{PGV} and Φ_{Ia} are normal CDFs and θ_{PGV}, β_{PGV} and θ_{Ia}, β_{Ia} are the parameters characterizing each respective CDFs.

The parameters of each fragility surface were estimated with the MLE method and can be found in Table 6.5. As it can be seen in Figures 6.6 and from Table 6.6, the parametric fragility surfaces fit well the fragility estimates calculated with the meta-model and describe well the behaviour of the structure with increasing levels of the selected seismic intensity measures. Nevertheless, it should be noted that using two (usually) correlated seismic IMs to build the fragility surface, can lead to the consideration of seismic scenarios very unlikely to happen at the site of the structure during its lifetime.
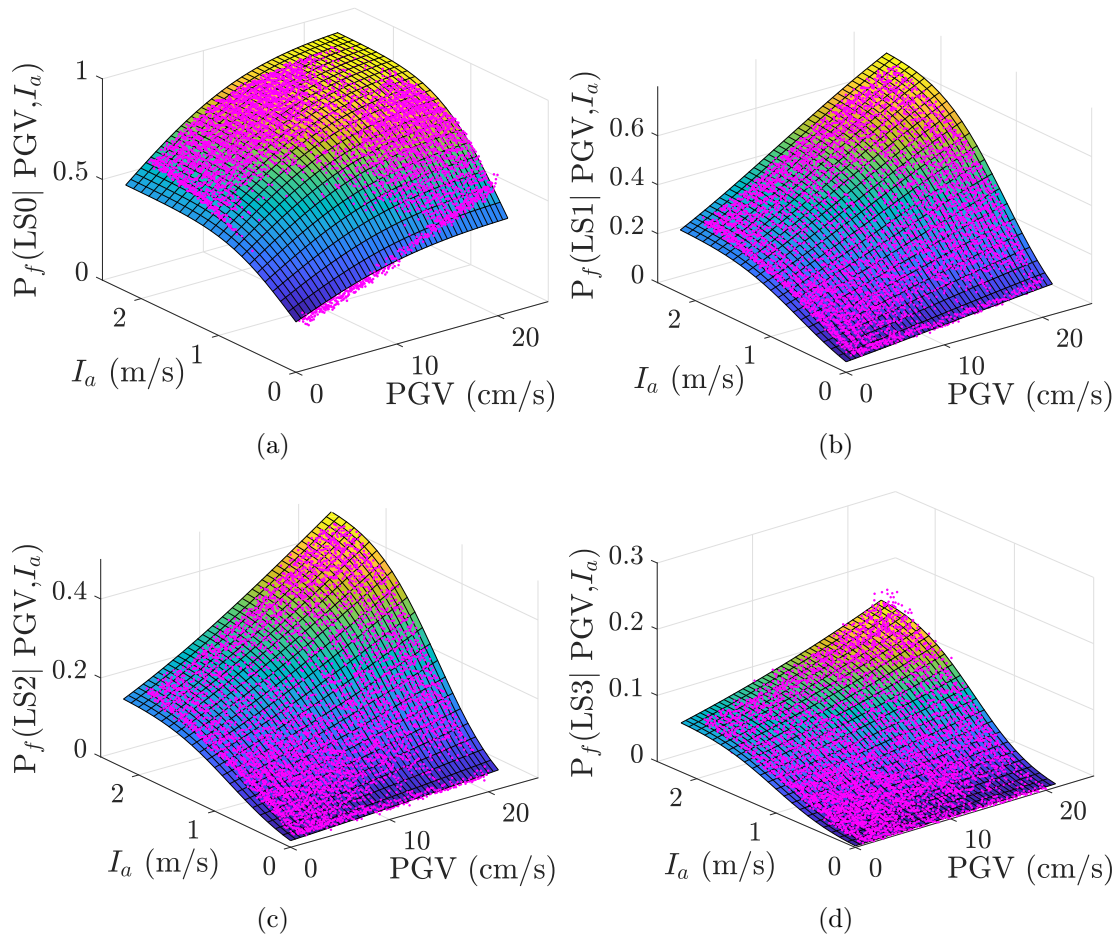
Figure 6.6   Fragility surfaces for (a) LS0, (b) LS1, (c) LS2 and (d) LS3

Table 6.5 PGV-I_a fragility surfaces parameters

Limit state	θ_{PGV}	β_{PGV}	θ_{Ia}	β_{Ia}
LS0	-0.643	9.340	0.039	0.643
LS1	10.189	15.954	0.623	0.641
LS2	20.606	22.146	0.804	0.614
LS3	61.760	42.983	1.024	0.554

Table 6.6 PGV-I_a fragility surfaces goodness-of-fit

Limit state	R^2	RMSE	RMAE
LS0	0.947	0.043	1.184
LS1	0.988	0.018	0.484
LS2	0.984	0.014	0.564
LS3	0.961	0.008	1.238

CHAPTER 7

CONCLUSION

7.1 Summary and conclusions

The consequences of dam failure can be substantial, in terms of both casualties and economic and environmental damage. Thus it has to maintain its structural integrity in the face of the different hazards and loading conditions that arise during construction, normal operations, and extreme environmental events. Therefore, dam safety is given highest priority. Moreover, the safety of dams and hydroelectric developments is a major concern in Quebec given that over half the population lives in potential flood zone. There are about 933 large dams in Canada and 333 or slightly more than a third are located in Quebec. Among them, many have been built more than 50 years ago. During that time, important advances in the methodologies for evaluating the natural hazards have been made, causing the review and modification of the design guidelines, in some cases significantly. Many existing dams fail to meet these revised safety criteria and structural rehabilitation to achieve newly revised criteria may be costly and difficult. Although the occurrence of total failure of a concrete dam following an earthquake is rare, earthquakes are, on the other hand, a major cause of damage to these structures at different degrees of severity.

Consequently, the combination ageing and its associated problems with new methods for estimating seismic loads and with the increasing demands of society to ensure a high level of safety, has resulted in the need to review and upgrade the methods of seismic analysis for dams. Traditionally, dams were evaluated using deterministic analysis under an extreme event. Nevertheless, in recent decades, probabilistic-based tools, such as fragility functions, have become increasingly popular for the seismic assessment of dams, although its main drawback is the requirement of a large number of nonlinear dynamic analysis of complex finite element models, rendering the task computationally expensive.

To address the aforementioned gaps, the main objective of the research project was to develop a method for assessing the seismic safety of gravity dam-type structures based on multi-variable fragility functions, developed by the implementation of meta-models and by properly identifying the seismic scenario likely to be present at the site. The proposed methodology was applied to a case study dam in Quebec which presents a well documented dynamic behaviour. Given the presence of contraction joints, it is assumed

that the blocks of the dam behave independently, a detailed 2D finite element model which fully account for dam-reservoir-foundation interactions and a non-linear analysis was performed. The model meshing was carried out using the computer software ANSYS-ICEM while the computer software LS-Dyna was used for modeling and analyzing the system's response. To evaluate the accuracy of the FEM, the dynamic properties of the DRF system were compared to the results obtained from in-situ dynamic experimental results. The validation of the dynamic characteristics were based on the fundamental period of the system and global damping. This study was divided in two parts, as it can be seen from the two journal articles included in the document.

In the first stage of this study, the main objective was to understand and present the level of improvement in the fragility analysis of concrete gravity dams provided by the use of an innovative method for record selection that combines the latest advances in probabilistic seismic hazard analysis and the most influential IMs for the proposed limit states. To this end, a PSHA was performed with the open-source software OpenQuake and the updated 2015 hazard model for Canada to characterize the seismic scenario at the dam site. The results of the PSHA with the updated hazard model were compared with the results of the previous hazard model to quantify to which extent the differences between the two models are significant. The updated model showed lower annual probabilities of exceedance for the same site, which can be explained by the consideration of a more extensive database.

The results of the PSHA were used to select ground motions consistent with the CS and the GCIM. The intensity measures identified as relevant for the seismic response of the dam were included in the modified selection algorithm consistent with the GCIM formulation. The sets of selected ground motion records were paired with the 30 samples of numerical models of the dam generated with LHS for each intensity level. The fragility curves for the base and neck sliding limit states of the dam were then generated via multiple stripes analysis. A safety assessment was performed by calculating the expected seismic performance under the extreme limit states proposed by the Canadian dam association guidelines and by determining the unconditional probabilities of limit state exceedance. The fragility functions and the results from the safety assessment obtained with the proposed methodology were compared with those resulting from the 2010 hazard model and using the UHS and the CS as target distributions. The comparison showed that the more refined methodology with the GCIM method and the 2015 hazard model produces less conservative fragility functions, followed by the 2015-CS, 2010-CS and finally the 2010-UHS. These results can be related to the more accurate hazard model and to the fact that the selected set of ground motions includes the characteristics of the IM that, in addition

to Sa_H, are likely to be present at the dam site. Selecting ground motions by considering only the horizontal spectral acceleration can lead to sets of records whose distributions with respect to other important IMs are not consistent with the seismic scenario at a specific site. Furthermore, the results of the safety assessment with the present methodology were in line with the minimum safety margins proposed by the current guidelines, whereas the results were above the tolerable limits for the same case study dam when using the previously employed procedure based on a less thorough seismic input definition.

Concerning the second part of this study, it was intended to use the PSHA results and the selected records with the method that provided the best results in the previous analysis to train surrogate models for the fragility seismic assessment of gravity dams. Therefore, the main goal of the second part of this research project was to identify the most viable meta-model for the seismic fragility assessment of gravity dams, present a methodology for the development of multivariate fragility functions displaying the effect of the model parameter variation on the dam seismic performance and formulate design recommendations from the analysis.

The seismic hazard characterization at the dam site was conducted with a PSHA with the 2015 hazard model and 250 ground motions were selected, consistent with the latter, using the GCIM approach. The sets of selected ground motion records were paired with the 250 samples of numerical models of the dam generated with LHS, representing different material and loading configurations of the system. The dataset used to train the meta-models was generated by performing non-linear dynamic analysis of these samples, with the FEM above-described and by extracting the maximum relative sliding at the base of the dam. Six different types of meta-models with different configurations (basis and interpolation functions) were fitted to the seismic response of the case study dam to predict the sliding limit state at the base of the dam. The 4th order PRS emerged as the best performing meta-model based on local and global goodness-of-fit estimates from 5-CV, and it was used to generate fragility curves as a function of PGV and fragility surfaces as a function of PGV and each of the model parameters, considering the correlation between the seismic IMs for a specific site and by sampling from a multivariate log-normal distribution.

To explicitly account for the variability of the modeling parameters in the fragility analysis, fragility curves were generated by propagating the uncertainty of these parameters and by considering the extreme values (95 % confidence interval) of the usable range of values. It was observed that the variability of the concrete-rock cohesion model parameter affects the most the fragility analysis estimates, followed by the drain efficiency and to a lesser extent, the concrete-rock angle of friction. Finally, from the fragility surfaces a

seismic assessment was performed to determine the model parameter boundaries for adequate performance under extreme events, such as the MCE, to respect the minimum safety margins proposed by the current guidelines. Consequently, the boundaries of model parameters were established to provide a probability of exceedance less than 6% for an MCE seismic scenario. As a result, a range of values of $50° \leq \text{CRF} \leq 55°$, $0.35 \leq \text{DR} \leq 0.66$ and $0.87\,\text{MPa} \leq \text{CRC} \leq 2.0\,\text{MPa}$ should be considered to ensure that the performance of the dam is in line with the minimum values for life safety. Nonetheless, it should be acknowledged that these parameter ranges are derived from the single parameter at a time evaluation, and the joint interaction of the model parameters should be further examined.

Comparatively, and to consider more general cases, the correlation between the seismic IMs was also neglected. The fragility point estimates were generated as a function of PGV and I_a, the two most influential seismic IMs. The uncertainty due to the modeling parameters and PGA_V was propagated in the analysis by sampling these parameters with LHS according to their respective range of usable values. It was assumed that for each limit state, the fragility surface is described by the product of two independent normal cumulative density functions. It was observed that the parametric fragility surfaces fit well the fragility estimates calculated with the meta-model and describe well the behaviour of the structure with increasing levels of the selected seismic intensity measures. Nevertheless, using two (usually) correlated seismic IMs to build the fragility surface, can lead to the consideration of seismic scenarios very unlikely to happen at the site of the structure during its lifetime.

Machine learning techniques are indispensable when assessing the vulnerability of structures with computationally expensive FEM such as gravity dams subjected to seismic loading. Similarly, the use of surrogate models allows for the exploration of the impact of different parameters in the fragility without the costly reevaluation of the FEM simulations. Regarding the fragility functions, as evident for the case presented herein and the goodness-of-fit values, the fragility estimates are well depicted by the methodology suggested in this study to fit parametric fragility surfaces. It is expected that the results of this study can lead to more accurate planning and retrofitting policies to expedite the safety assessment of dams under seismic loads while supporting the decision-making process and to guide the preliminary design of future gravity dams.

7.2 Future work recommendations

Despite the specificities of this case study, namely, its location in eastern Canada and consideration of only sliding limit states, the conclusions are believed to be useful for the seismic assessment of concrete gravity dams in general. An important limitation of this study, which should be addressed by future developments in the field, is the lack of empirical GMMs and correlation models for ENA and the corresponding impossibility of selecting records from the ENA database.

In future studies, additional insight into the correlation between the parameters defining the model configurations should be made, including other relevant limit states for gravity dams. Additionally, the model parameter variations in the fragility analysis should be further explored to provide parametric fragility functions, including the joint interaction of these parameters using classification meta-modeling techniques.

More specifically, the results of this research could be complemented by additional studies on the following subjects:

- Definition of performance criteria and damage limit states. The main drawback of the methodology used is the lack of data to define the levels of damage. Experimental research must be conducted to characterize the damage states which are usually based on expert judgment, leaving much room for subjectivity. Studies on the concrete modulus of elasticity non-linearities and definition of cracking performance limit states for gravity dams should be given priority.
- Consideration of a 3D model to assess the seismic vulnerability of dams. Two-dimensional analysis of dam geometries is still the most common approach for the design or evaluation of the gravity dams. Nevertheless, it may happen that the accuracy of 2D models is not enough to depict, in a realistic manner, the seismic behaviour and must resort to the 3D modelling whose dynamic properties are known to vary significantly from those found for a two-dimensional model.
- Include the geometric uncertainty in the development of meta-models for fragility analysis. Besides the model parameters and the seismic IMs, efforts should be done to include the geometric uncertainty in the analysis. Geometric uncertainty refers to the geometric characteristic of the dam (monolith's height, base length and width). A geometric variable, such as base-to-height ration of the blocks, should be included to develop meta-models. Given that the case study dam is very regular and quasi-symmetric, it would be possible to generate a meta-model from 3D FEM simulations and another from 2D FEM of each of the blocks of the dam including this geometric

uncertainty and compare the two approaches. Moreover, given that 2D FEM models are widely used for gravity dams, it is believed that by including geometric parameters to the covariate matrix, the resulting meta-model will be useful in predicting the response of a 2D block cross-section gravity dams' portfolio.

– Study the applicability of a two-layered meta-model scheme to efficiently predict the response of the dam. It has been observed that for a given seismic intensity level, and for a given modeling and geometric parameter configuration, the dam system response is not significant enough (base or neck displacement < 3mm or crack length < 2.5mm), which render the task of fitting a regression meta-model very inaccurate in the presence of several values close to zero. Thus, a classification meta-model could be fitted first to determine whether the response of the dam is significant or not, and then for those values which are above the specified limit, a regression meta-model will be fitted to predict the continuous response of the dam.

– Study the effect of neglecting the correlation between the parameters defining the model configurations, such as concrete-rock cohesion and concrete-rock angle of friction.

– Explore the joint model parameter variations impact in the fragility analysis to provide parametric fragility functions, including simultaneous effect of these parameters using classification meta-modeling techniques.

7.3   Sommaire et conclusions

L'occurence d'une rupture de barrage peut avoir des conséquences considérables, telles des pertes en vies humaines, des répercussions économiques et des impacts environnementaux. La sécurité des barrages est d'intérêt public et un souci permanent à toutes les phases du projet, aussi bien de la construction, que de l'exploitation. L'intégrité structurale d'un barrage doit être maintenue face aux différents risques et aux conditions de chargement qui surviennent durant la construction, l'exploitation normale, et les événements environnementaux extrêmes. Au Québec, en particulier, la sécurité des barrages et des aménagements hydroélectriques est une préoccupation majeure étant donné que plus de la moitié de la population vit dans une zone potentiellement inondable. On dénombre environ 933 grands barrages au Canada dont plus du tiers (333) sont situés au Québec. Parmi ces derniers, beaucoup ont été construits depuis plus de 50 ans. Entre-temps, d'importants progrès ont été réalisés dans les méthodes d'évaluation des risques naturels, entrainant la révision et la modification des règles de conception de manière significative dans certains cas. De nombreux barrages existants ne satisfont pas à ces critères de sécu-

rité révisés, et la réhabilitation de la structure pour satisfaire aux nouveaux critères peut s'avérer onéreuse et difficile à réaliser. Bien que la rupture totale d'un barrage en béton à la suite d'un seisme soit rare, les tremblements de terre sont, tout de même, une cause majeure d'endommagement de ces structures à différents niveaux de gravité.

En conséquence, la conjonction de la vétusté des structures, des problèmes associés aux nouvelles méthodes d'estimation des charges sismiques et des exigences croissantes de la société pour assurer un haut niveau de sécurité, a fait ressortir la nécessité de reviser et d'améliorer les méthodes d'analyse sismique des barrages. Traditionnellement, les barrages étaient évalués à l'aide d'une analyse déterministe pour un événement extrême. Toutefois, au cours des dernières décennies, les outils probabilistes, tels que les fonctions de fragilité, sont devenus de plus en plus populaires pour l'évaluation sismique des barrages, malgré que leur principal inconvénient soit l'exigence d'un grand nombre d'analyses dynamiques non linéaires de modèles d'éléments finis complexes, rendant le coût de calcul élevé.

Pour combler les lacunes susmentionnées, l'objectif principal du projet de recherche était d'élaborer un plan d'action pour la mise en œuvre de l'évaluation de la sécurité sismique de structures de type barrage-poids avec des fonctions de fragilité multivariées, développées par l'implémentation de méta-modèles et l'identification fidèle de scénarios sismiques susceptibles de se produire sur le site. La méthodologie proposée a été appliquée à un barrage au Québec qui présente un comportement dynamique bien documenté. Un modèle complet d'éléments finis 2D qui tient compte des interactions barrage-réservoir-fondation (BRF) couplé à une analyse non linéaire a été développé. Le maillage du modèle a été réalisé à l'aide du logiciel ANSYS-ICEM tandis que le solveur LS-Dyna a été utilisé pour modéliser la réponse du système. Pour évaluer la précision du modèle, les propriétés dynamiques du système BRF ont été comparées aux résultats expérimentaux dynamiques obtenus in situ. La validation des caractéristiques dynamiques étaient basées sur la période fondamentale du système et l'amortissement global. Cette étude a été scindée en deux parties, comme on peut le voir dans les deux articles de revue inclus dans le document.

Dans la première phase de l'étude, l'objectif principal était de comprendre et de présenter le niveau d'amélioration de l'analyse de fragilité des barrages-poids en béton grâce à l'utilisation d'une méthode innovante de sélection d'accélérogrammes qui combine les récentes avancées en matière d'analyse probabiliste d'aléas sismiques et de mesures d'intensités (MI) les plus influentes pour les états limites proposés. À cette fin, une analyse probabiliste d'aléas sismiques a été réalisée à l'aide du logiciel libre OpenQuake et du modèle d'aléas sismiques mis à jour en 2015 pour le Canada pour caractériser le scénario sismique sur le site du barrage. Les résultats de l'analyse probabiliste des aléas sismiques (APAS)

sur le site du barrage avec le modèle de risque mis à jour a été comparé aux résultats de la version précédente afin de quantifier l'importance des différences entre les deux modèles. Le modèle actualisé montrait des probabilités annuelles de dépassement plus faibles pour le même site, ce qui peut s'expliquer par la considération d'une base de données plus complète.

Les résultats de l'APAS ont été utilisés pour sélectionner les séismes conformes aux spectres conditionnels et à la GCIM. Les mesures d'intensités jugées pertinentes pour la réponse sismique du barrage ont été incluses dans l'algorithme de sélection modifié conforme à la formulation de la GCIM. Les ensembles de tremblements de terre sélectionnés ont été appariés avec les 30 échantillons de modèles numériques du barrage générés par la Latin Hypercube Sampling pour chaque niveau d'intensité. Les courbes de fragilité des états limites de glissement à la base et aux joints de reprise du barrage ont ensuite été générés au moyen d'une analyse à bandes multiples. Une évaluation de la sécurité a été effectuée en calculant la performance sismique attendue sous les états limites extrêmes proposés par les règles de conception de l'Association canadienne des barrages et en déterminant les probabilités inconditionnelles de dépassement des états limites. Les fonctions de fragilité et les résultats de l'évaluation de la sécurité obtenues avec la méthodologie proposée ont été comparées à celles résultant du modéle d'aléa sismique 2010 en utilisant les distributions UHS et CS comme distributions cibles. La comparaison a montré que la méthodologie la plus raffinée avec la méthode GCIM et le modéle d'aléa sismique 2015 produit des résultats moins conservateurs suivis par les modèles 2015-CS, 2010-CS et enfin 2010-UHS. Ces résultats peuvent être lié au modéle d'aléa sismique plus précis et au fait que les accélérogrammes choisis incluent les caractéristiques des mesures d'intensités sismiques qui, en plus de Sa_H, sont susceptibles d'être présents sur le site du barrage. Sélectionner les accélérogrammes en ne prenant en compte que les accélérations spectrales horizontales peut conduire à des ensembles d'enregistrements dont la distribution par rapport à d'autres MI importantes ne sont pas compatibles avec le scénario sismique d'un site particulier. Par ailleurs, les résultats de l'évaluation de la sécurité avec la méthodologie actuelle étaient conformes aux marges de sécurité minimales proposées par les recommandations actuelles, alors que les résultats étaient supérieurs aux limites tolérables pour le même barrage lorsque la procédure utilisée précédemment était basée sur une définition moins approfondie de l'apport sismique.

En ce qui concerne la deuxième partie de l'étude, il était prévu d'utiliser les résultats de l'APAS et les données sélectionnées avec la méthode ayant donné les meilleurs résultats pour la formation des modèles de substitution pour l'évaluation sismique de la fragilité

des barrages poids. Par conséquent, l'objectif principal de la deuxième partie de cette recherche était d'identifier le méta-modèle le plus viable pour l'évaluation de la fragilité sismique des barrages-poids, de présenter une méthodologie pour le développement de fonctions de fragilité multivariées exhibant l'effet de la variation des paramètres du modèle sur la performance sismique du barrage, et de formuler des recommandations de conception à partir de l'analyse.

La caractérisation des aléas sismiques sur le site du barrage a été effectuée à l'aide d'une APAS avec le modéle d'alea sismique 2015, et 250 accélérogrammes ont été sélectionnés, conformément à ce dernier, en utilisant l'approche GCIM. Les ensembles d'accélérogrammes sélectionnés ont été jumelés avec les 250 échantillons de modèles numériques du barrage générés avec le LHS, ce qui représente différentes configurations de matériaux et de chargement du système. L'ensemble des données utilisées pour former les méta-modèles a été généré à partir d'analyses dynamiques non-linéaires de ces échantillons, à l'aide de la MEF et en extrayant le glissement relatif maximum à la base du barrage. Six différents sortes de méta-modèles avec différentes configurations (fonctions de base et d'interpolation) ont été ajustées à la réponse sismique du barrage à l'étude pour prédire l'état limite de glissement à la base du barrage. La surface de réponse polynomiale du 4e ordre s'est révélée être la meilleure pour l'exécution d'un méta-modèle basé sur des estimations locales et globales de la qualité de l'ajustement à partir de 5-CV, et il a été utilisé pour générer des courbes de fragilité en fonction du PGV et des surfaces de fragilité en fonction du PGV et de chacun des paramètres du modèle, en considérant la corrélation entre les MI sismiques pour un site spécifique et par échantillonnage à partir d'une distribution log-normale multivariée.

Pour tenir compte explicitement de la variabilité des paramètres de modélisation dans l'analyse de fragilité, des courbes de fragilité ont été générées en répandant l'incertitude de ces paramètres et en considérant les valeurs extrêmes (95 % d'intervalle de confiance) de la plage des valeurs utilisables. Il a été observé que le paramètre de modélisation qui influe le plus sur l'analyse de fragilité est la cohésion béton-rocher, suivi de l'efficacité de drainage et, dans une moindre mesure, de l'angle de frottement entre le béton et le rocher. Enfin, à partir des surfaces de fragilité, une évaluation sismique a été réalisée pour déterminer les limites des paramètres du modèle afin d'obtenir un rendement adéquat dans des conditions extrêmes des événements tels que le MCE, pour respecter les marges de sécurité minimales proposées par les recommandations actuelles. Par conséquent, les limites des paramètres du modèle ont été établies pour fournir une probabilité de dépassement inférieure à 6 % pour un scénario sismique MCE. Ainsi, une fourchette de valeurs de $50° \leq \text{CRF} \leq 55°$, $0.35 \leq \text{DR} \leq 0.66$ and $0.87\,\text{MPa} \leq \text{CRC} \leq 2.0\,\text{MPa}$ devrait être prise en compte pour

s'assurer que le rendement du barrage soit conforme aux exigences des valeurs minimales pour la sécurité des personnes. Néanmoins, il faut reconnaître que ces plages de paramètres sont dérivées du paramètre individuel au momment de l'évaluation et que l'interaction conjointe des paramètres du modèle devrait faire l'objet d'un examen plus approfondi.

Comparativement, et pour examiner des cas plus généraux, la corrélation entre les MI sismiques a également été négligée. Les estimations ponctuelles de fragilité ont été générées en fonction du PGV et de I_a, le deux MI sismiques les plus influents. L'incertitude due aux paramètres de modélisation et à PGA_V a été distribuée dans l'analyse en échantillonnant ces paramètres avec le LHS selon leurs valeurs respectives utilisables. On a supposé que, pour chaque état limite, la surface de fragilité est définie par le produit de deux fonctions normales de densité cumulative. Il a été observé que les surfaces de fragilité paramétriques correspondent bien aux estimations de fragilité calculées à l'aide du métamodèle et décrivent bien le comportement de la structure avec des niveaux croissants de mesures d'intensités sismiques choisies. Néanmoins, l'utilisation (habituelle) de deux IMs sismiques corrélés pour construire la surface de fragilité peut mener à considérer des scénarios sismiques très peu susceptibles de se produire à l'emplacement de l'ouvrage pendant sa durée de vie.

Il est à noter que les techniques d'apprentissage machine sont indispensables lors de l'évaluation de la vulnérabilité des structures dont le coût de calcul par la MEF est élevé, comme les barrages gravitaires soumis à des charges sismiques. De même, l'utilisation de modèles de substitution permet d'explorer l'impact de différents paramètres dans la fragilité sans la réévaluation coûteuse des simulations des MEF. En ce qui concerne les fonctions de fragilité, comme cela a été démontré dans le cas présenté ici et dans ceux des valeurs de la qualité de l'ajustement, les estimations de la fragilité sont bien illustrées par la méthodologie suggérée dans cette étude pour ajuster les surfaces de fragilité paramétriques. On s'attend à ce que les résultats de cette étude puissent mener à des politiques de planification et de modernisation plus précises afin d'accélérer l'évaluation de la sécurité des barrages sous des charges sismiques tout en appuyant la prise de décisions, et pour guider la conception préliminaire des futurs barrages gravitaires.

7.4 Recommandations pour les travaux futurs

Malgré les particularités de cette étude de cas, à savoir son emplacement dans l'est du Canada et la prise en compte uniquement des états limites de glissment, les conclusions sont jugées utiles pour l'examen de l'application de l'évaluation sismique des barrages gravitaires en béton en général. Une limite importante de cette étude, qui devrait être prise en compte dans les développements futurs dans ce domaine, est l'absence de MGM empiriques et d'outils d'évaluation de l'impact, de modèles de corrélation pour l'ENA et l'impossibilité de sélectionner des enregistrements dans base de données de l'ENA.

Dans les études à venir, un aperçu supplémentaire de la corrélation entre les paramètres qui définissent l'approche des configurations des modèles devraient être établies, y compris les autres états limites pertinents pour les barrages gravitaires. En outre, les variations des paramètres du modèle dans l'analyse de fragilité devraient être examinées plus en détail pour fournir des fonctions paramétriques de fragilité, y compris l'interaction conjointe de ces paramètres en utilisant des techniques de méta-modélisation de classification.

Plus précisément, les résultats de cette recherche pourraient être complétés par d'autres études sur les sujets suivants :

- Définition des critères de performance et des états limites de dommages. Le principal inconvénient de la méthodologie utilisée est le manque de données pour définir les niveaux de dommages. La recherche expérimentale doit être effectuée pour caractériser les états d'endommagement qui sont habituellement basés sur des données d'experts en laissant beaucoup de place à la subjectivité. Les études sur le module d'élasticité du béton, les non-linéarités et la définition des états limites de résistance à la fissuration des barrages devrait être prioritaires.

- Examen d'un modèle 3D pour évaluer la vulnérabilité sismique des barrages. En deux dimensions, l'analyse de la géométrie des barrages est encore l'approche la plus courante pour la conception ou l'évaluation des barrages gravitaires. Néanmoins, il peut arriver que la précision des modèles 2D ne soit pas suffisante pour représenter de manière réaliste le comportement sismique et on doit recourir à une modélisation tridimensionnelle dont on sait que les propriétés dynamiques varient considérablement par rapport à celles d'un modèle bidimensionnel.

- Inclusion de l'incertitude géométrique dans l'élaboration d'analyse de fragilité des méta-modèles. Outre les paramètres du modèle et les GI sismiques, des efforts devraient être faits pour inclure l'incertitude géométrique dans l'analyse. L'incertitude géométrique fait référence aux caractéristiques géométriques du barrage (hauteur du

monolithe, longueur et largeur de la base). Une variable géomérique telle le rapport de la base à la hauteur des blocs, doit être considérée pour développer des méta-modèles. Étant donné que le barrage étudié est très régulier et quasi-symétrique, l'idée est de générer un méta-modèle à partir de simulations de MEF 3D et un autre à partir de simulations de MEF 2D de chacun des blocs du barrage, y compris cette incertitude géométrique, et comparer les deux approches. De plus, étant donné que les modèles d'EF 2D du barrage sont encore les plus courants pour les barrages-poids, on croit qu'en incluant les données géométriques à la matrice de covariables, le méta-modèle résultant sera utile pour prédire le paramètre de réponse d'un ensemble de barrages gravitaires de section transversale de bloc 2D.

– Étude de l'applicabilité d'un méta-modèle à deux niveaux pour prédire efficacement la réponse du barrage. Ceci est dû au fait que l'on a observé que pour une inten-sité sismique donnée et pour une modélisation et une configuration de paramètres géométriques données, la réponse du système de barrage n'est pas assez significa-tive (glissement de la base et aux joints de reprise < 3 mm ou longueur de fissure $< 2{,}5$ mm), ce qui rend l'ajustement d'un méta-modèle de régression très imprécis en présence de plusieurs valeurs proches de zéro. Ainsi, un méta-modèle de classi-fication pourrait être ajusté d'abord pour déterminer si la réponse du barrage est importante ou non, et ensuite pour les valeurs qui sont supérieures à la limite spé-cifiée, un méta-modèle de régression sera ajusté pour prédire la réponse continue du barrage.

– Étude de l'effet de négliger la corrélation entre les paramètres définissant le modèle tels que la cohésion béton-rocher et l'angle de frottement béton-rocher.

– Exploration des variations des paramètres du modèle dans l'analyse de fragilité afin de fournir des fonctions de fragilité paramétriques, incluant l'interaction conjointe de ces paramètres à l'aide de techniques de méta-modélisation de la classification.

LIST OF REFERENCES

[1] Hydro-Québec. Centrales hydroĺectriques (au 11 novembre 2014). http://www.hydroquebec.com/production/centrale-hydroelectrique.htm, 2014.

[2] Natural Resources Canada (NRC). Simplified seismic hazard map for canada. http://www.earthquakescanada.nrcan.gc.ca/hazard-alea/simphaz-eng.php, 2015. visited on 2016-02-08.

[3] J. Adams. Seismic hazard maps for the national building code of canada - past, present and future. Technical report, 2012.

[4] United States Society of Dams, U.S.S.D. *Observed Performance of Dams During Earthquakes*, volume III. 2014.

[5] M. Wieland. Seismic design of major components - International Water Power. http://www.waterpowermagazine.com/features/featureseismic-design-of-major-components/, 2013.

[6] C. Marche. *Barrages, crues de rupture et protection civile*. Presses internationales Polytechnique, 2 edition, 2008.

[7] A.J. Schleiss and H. Pougatsch. *Les Barrages, du project a la mise en service*. Presses polytechniques et universitaires romandes, 1 edition, 2011.

[8] Canadian Seismic Research Network (CSRN). Canadian seismic research network. http://csrn.mcgill.ca/main.html, 2009. Accessed: 2016-01-27.

[9] Canadian Dam Association (CDA). Dams in canada. https://www.cda.ca/EN/Dams_in_Canada/. Accessed: 2019-03-08.

[10] Hydro-Québec. Loi sur la sécurité des barrages. http://www2.publicationsduquebec.gouv.qc.ca, 2001.

[11] L. B. Lave, D. Resendiz-Carrillo, and F. C. McMichael. Safety goals for high-hazard dams: are dams too safe? *Water Resources Research*, 26:1383–1391, 1990.

[12] M.T. Schultz, B.P. Gouldby, J.D. Simm, and J.L. Wibowo. Beyond the factor of safety: Developing fragility curves to characterize system reliability. Technical report ERDC SR-10-01, U.S. Army Corps of Engineers, 2010.

[13] P.B. Tekie and B.R. Ellingwood. Fragility analysis of concrete gravity dams. *Journal of infrastructure systems*, 127(8):41–48, 2002.

[14] P.B. Tekie and B.R. Ellingwood. Seismic fragility assessment of concrete gravity dams. *Earthquake Engineering & Structural Dynamics*, 32(14):2221–2240, 2003.

[15] A.I.J. Forrester, A. Sóbester, and A.J. Keane. *Engineering Design via Surrogate Modelling: A practical Guide*. Wiley, 2008.

[16] D. Tavares, J.E. Padgett, and P. Paultre. Fragility curves of typical as-built highway bridges in eastern Canada. *Engineering Structures*, 40(7):107–118, 2012.

[17] D. Tavares, Ruiz-Suescun Juliana, P. Paultre, and J.E. Padgett. Seismic Fragility of a Highway Bridge in Quebec. *Journal of Bridge Engineering*, 18(11):1131–1139, 2013.

[18] G. Siqueira, S. Adamou, P. Paultre, and J.E. Padgett. Fragility curves for isolated bridges in eastern Canada using experimental results. *Engineering Structures*, 74(9):311–324, 2014.

[19] G. Siqueira, D. Tavares, P. Paultre, and J.E. Padgett. Performance evaluation of natural rubber seismic isolators as a retrofit measure for typical multi-span concrete bridges in eastern Canada. *Engineering Structures*, 74(9):300–310, 2014.

[20] C. Bernier, J.E. Padgett, J. Proulx, and P. Paultre. Seismic fragility of concrete gravity dams with modeling parameter uncertainty and spacial variation. *Journal of Structural Engineering*, 142:05015002, 2016.

[21] C. Bernier, R. Monteiro, and P. Paultre. Using the Conditonal Spectrum Method for Improved Fragility Assessment of Concrete Gravity Dams in Eastern Canada. *Earthquake Spectra*, 32:1449–1468, 2016.

[22] J. Proulx and P. Paultre. Experimental and numerical investigation of dam-reservoir-foundation interaction for a large gravity dam. *Canadian Journal of Civil Engineering*, 24(1):90–105, 1997.

[23] J.E. Padgett and R. DesRoches. Sensitivity of Seismic Response and Fragility to Parameter Uncertainty. *Journal of Structural Engineering*, 133(12):1710–1718, 2007.

[24] C. B. Haselton, J. W. Baker, Y. Bozorgnia, C. A. Goulet, E. Kalkan, N. Luco, T. Shantz, N. Shome, J. P. Stewart, P. Tothong, J. Watson-Lamprey, , and F. Zareian. Evaluation of Ground Motion Selection and Modification Methods: Predicting Median Interstory Drift Response of Buildings. Technical Report PEER 2009/01, Pacific Earthquake Engineering Research Center, 2009.

[25] B.A. Bradley. The seismic demand hazard and importance of the conditioning intensity measure. *Earthquake Engineering & Structural Dynamics*, 41:1417–1437, 2012.

[26] K. Tarbali and B.A. Bradley. Scenario-based ground-motion selection using the generalized conditional intensity measure (GCIM) approach. *Tenth U.S. National Conference on Earthquake Engineering Frontiers of Earthquake Engineering*, 2014.

[27] L. Sousa, V. Silva, M. Marques, and H. Crowley. On the treatment of uncertainties in the development of fragility functions for earthquake loss estimation of building portfolios. *Earthquake Engineering & Structural Dynamics*, 44:657–675, 2016.

[28] M.A. Hariri-Ardebili and V. Saouma. Seismic fragility analysis of concrete dams: A state-of-the-art review. *Engineering Structures*, 128:374–399, 2016.

[29] M.A. Hariri-Ardebili and Saouma V. Collapse Fragility Curves for Concrete Dams: A Comprehensive Study. *Journal of Structural Engineering*, page 04016075, 2016.

[30] M.A. Hariri-Ardebili and V. Saouma. Probabilistic seismic demand model and optimal intensity measure for concrete dams. *Structural Safety*, 59:67–85, 2016.

[31] Dam safety recommendations Section 6.1, 6.2 and 6.3, revised in 2013. In *Dam safety Guidelines*. 2007.

[32] B.R. Ellingwood and P.B. Tekie. Fragility Analysis of Concrete Gravity Dams. *Journal of Infrastructure Systems*, 7:41–48, 2001.

[33] M. Araujo and A.M. Awruchb. Probabilistic finite element analysis of concrete gravity dams. *Advances in Engineering Software*, 29(2):97–104, 1998.

[34] Y. Bilici, A. Bayraktar, K. Soyluk, K. Haciefendioglu, S. Ates, and S. Adanur. Stochastic dynamic response of dam-reservoir-foundation systems to spatially varying earthquake ground motions. *Soil Dynamics and Earthquake Engineering*, 29:444–458, 2009.

[35] International Commission on Large Dams, I.C.O.L.D. Dam safety and earthquakes. *International Water Power and Dam Construction*, (August):12–14, 2010.

[36] P.B. Tekie and B.R. Ellingwood. Fragility Analysis of Concrete Gravity Dams. *USACE Engineering Manual, ERDC/ITL TR-02-6*, 32, 2002.

[37] Barbara Mills-Bria, Roman Koltuniuk, and Phoebe Percell. State-of-Practice for the Nonlinear Analysis of Concrete Dams 2013. Technical report, U.S Department of the interior Bureau of Reclamation, 2013.

[38] A. Lupoi and C. Callari. A probabilistic method for the seismic assessment of existing concrete gravity dams. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, pages 37–41, 2013.

[39] M.A. Hariri-Ardebili and V. Saouma. Quantitative failure metric for gravity dams. *Earthquake Engineering & Structural Dynamics*, 44:461–480, 2015.

[40] H. Zhong and G. Lin. Seismic Vulnerability Analysis of a Gravity Dam Based on Typical Failure Modes. *15th World Conference on Earthquake Engineering*, 2012.

[41] Earthquake Design and Evaluation of Concrete Hydraulic Structures. Technical report EM 1110-2-6053, U.S. Army Corps of Engineer, 2007.

[42] C.R. Noble. *Finite Element Techniques for Realistically Simulating the Seismic Response of Concrete Dams*. PhD thesis, University of California, 2007.

[43] G.L. Fenves and A.K. Chopra. *Simplified analysis for earthquake resistant design of concrete gravity dams*. Earthquake Engineering Research Center, University of California, Berkeley, 1986.

[44] P. Paultre. *Dynamics of Structures*. John Wiley & Sons, Inc, 2011.

[45] International Commission on Large Dams, I.C.O.L.D. Selecting Seismic Parameters for Large Dams. *Bulletin 72 - 2010 Revision*, 2010.

[46] F. García, J.J. Aznárez, H. Cifuentes, F. Medina, and O. Maeso. Influence of reservoir geometry and conditions on the seismic response of arch dams. *Soil Dynamics and Earthquake Engineering*, 67:264–272, 2014.

[47] S.M. Yilmazturk, Y. Arici, and B. Binici. Seismic assessment of a monolithic RCC gravity dam including three dimensional dam-foundation-reservoir interaction. *Engineering Structures*, 100:137–148, 2015.

[48] P. Paultre. *Structures en béton armé*. Presses internationales Polytechnique, 2011.

[49] M. Schultz, P. Huynh, and V. Cvijanovic. Dams and Extreme Events - Reducing Risk of Aging Infrastructure under Extreme Loading Conditions On the Cover. *34th Annual USSD Conference*, 2014.

[50] A.K. Chopra. Earthquake Analysis of Concrete Dams: Factors to be Considered. *Tenth U.S. National Conference on Earthquake Engineering, Anchorage, Alaska*, 2014.

[51] Tahar B.A., M. Belharizi, A. Laulusa, and A. Bekkouche. Fluid - structure interaction of Brezina arch dam: 3D modal analysis. *Engineering Structures*, 84:19–28, 2015.

[52] G. Hou, J. Wang, and A. Layton. Numerical Methods for Fluid-Structure Interaction - A Review. *Communications in Computational Physics*, 2:337–377, 2012.

[53] B. Miquel. *Nouvelles techniques pratiques pour la modélisation du comportement dynamique des systemes eau-structure*. PhD thesis, École Polytechnique de Montréal, 2012.

[54] J.O. Hallquist. *LS-DYNA Theory Manual*. Number March. Livermore Software Technology Corporation, 2006.

[55] P. Leger and M. Boughoufalah. Earthquake input mechanisms for time-domain analysis of dam-foundation systems. *Engineering structures*, 11:10, 1989.

[56] G.S. Sooch and A. Bagchi. Effect of Seismic wave scattering on the Response of Dam-Reservoir-Foundation Systems. *15th World Conference on Earthquake Engineering, Lisboa*, 2012.

[57] G.S. Sooch and A. Bagchi. A New Iterative Procedure for Deconvolution of Seismic Ground Motion in Dam-Reservoir-Foundation Systems. *Journal of Applied Mathematics*, 2014.

[58] V. Patil Swapnal. Effect of Soil Structure Interaction on Gravity Dam. *International Journal of Science, Engineering and Technology Research*, 4:1046–1053, 2015.

[59] L. Menglin, W. Huaifeng, C. Xi, and Zhai Yongmei. Structure-soil-structure interaction: Literature review. *Soil Dynamics and Earthquake Engineering*, 31:1724–1731, 2011.

[60] P.B. Tekie. *Fragility Analysis of Concrete Gravity Dams*. PhD thesis, Johns Hopkins University, 2001.

[61] M.T. Schultz, B.P. Gouldby, J.D. Simm, and J.L. Wibowo. Beyond the Factor of Safety : Developing Fragility Curves to Characterize System Reliability Geotechnical and Structures Laboratory Beyond the Factor of Safety : Developing Fragility Curves to Characterize System Reliability. Technical Report ERDC SR-10-1, U.S. Army Corps of Engineers, 2010.

[62] B.R. Ellingwood. Validation studies of seismic PRAs. *Nuclear Engineering and Design*, 123(2-3):189–196, 1990.

[63] Seismic probabilistic risk assessment implementation guide. Technical Report Final report 3002000709, Electrical Power Research Institute (EPRI), 2013.

[64] B.R. Ellingwood, O.C. Celik, and K. Kinali. Fragility assessment of building structural systems in Mid-America. *Earthquake Engineering & Structural Dynamics*, 41(11):1549–1568, 2007.

[65] National Research Council of Canada (NRCC). *National Building Code of Canada 2015*. 2010.

[66] T.I. Allen, J. Adams, and S. Halchuk. The seismic hazard model for Canada: Past, present and future. *Tenth Pacific Conference on Earthquake Engineering, Sydney, Australia*, 2015.

[67] G. Atkinson and J. Adams. Ground motion prediction equations for application to the 2015 Canadian national seismic hazard maps. *Canadian Journal of Civil Engineering*, 40:988–998, 2013.

[68] J.W. Baker. Efficient analytical fragility function fitting using dynamic structural analysis. *Earthquake Spectra*, 31:579–599, 2015.

[69] J.J. Boomer. Using real earthquake accelerograms for dynamic analysis of nuclear facilities: defining spectral targets, selecting records and adjusting for consistency. *Transactions, SMiRT 21, New Delhi, India*, 2011.

[70] N. Jayaram, T. Lin, and J.W. Baker. A Computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variancem. *Earthquake Spectra*, 27:797–815, 2011.

[71] J.W. Baker and C. Lee. An Improved Algorithm for Selecting Ground Motions to Match a Conditional Spectrum. *Journal of Earthquake Engineering*, 22:708–723, 2016.

[72] P. Daneshvar, N. Bouaanani, and A. Godia. On computation of conditional mean spectrum in Eastern Canada. *Journal of Seismology*, 19:443–467, 2015.

[73] J.W. Baker. Conditional mean spectrum: Tool for ground-motion selection. *Journal of Structural Engineering*, 137:322–331, 2011.

[74] B.A. Bradley. A generalized conditional intensity measure approach and holistic ground-motion selection. *Earthquake Engineering and Structural Dynamics*, 39:1321–1342, 2010.

[75] T. Lin, S. Harmsen, J.W. Baker, and N. Luco. Conditional Spectrum Computation Incorporating Multiple Causal Earthquakes and Ground Motion Prediction Models. *Bulletin of the Seismological Society of America*, 103, 2012.

[76] M. Shinozuka, J. Lee, and T. Naganuma. Statistical Analysis of Fragility Curves. *Journal of Engineering Mechanics*, 126:1224–1231, 2000.

[77] D. Vamvatsikos and A.C. Cornell. Incremental dynamic analysis. *Earthquake Engineering and Structural Dynamics*, 31(3):491–514, 2002.

[78] K. Mackie and B. Stojadinovic. Seismic Demands for Performance-Based Design of Bridges. Technical Report PEER 2003/16, Pacific Earthquake Engineering Research Center, 2003.

[79] L. Lin and J. Adams. Lessons for the fragility of Canadian hydropower components under seismic loading. *In Proceedings of the Ninth Canadian Conference on Earthquake Engineering, Ottawa, Canada*, pages 1762–1771.

[80] C. Rojahn and R.L. Sharpe. Earthquake damage evaluation data for California. Technical Report ATC-13, Applied Technology Council (ATC), 1985.

[81] G.H. Siqueira. *Évaluation de la vulnérabilité sismique des ponts routiers au Québec réhabilités avec l'utilisation d'isolateurs en caoutchouc naturel.* PhD thesis, Université de Sherbrooke, 2013.

[82] Armen Der Kiureghian. First and second-order reliability methods. In *Engineering Design Reliability Handbook*, chapter 14. CRC Press, 2004.

[83] Y. Ghanaat, R.C. Patev, and A.K. Chudgar. Seismic fragility analysis of concrete gravity dams. *15th World Conference on Earthquake Engineering, Lisboa*, 2012.

[84] S. Mirzahosseinkashani and M. Ghaemian. Seismic fragility assessment of concrete gravity dams. *29th Annual USSD Conference Nashville, Tennessee*, pages 173–181, 2009.

[85] Y. Ghanaat, P.S. Hashimoto, O. Zuchuat, and R.P. Kennedy. Seismic fragility of Mühleberg dam using nonlinear analysis with Latin Hypercube Simulation. *31st Annual USSD Conference San Diego, California*, 2011.

[86] D. Lallemant, A. Kiremidjian, and H. Burton. Statistical procedures for developing earthquake damage fragility curves. *Earthquake Engineering & Structural Dynamics*, 44:1373–1389, 2015.

[87] R.L. Segura, C. Bernier, R. Monteiro, and P. Paultre. On the seismic fragility assessment of concrete gravity dams in eastern Canada. *Earthquake Spectra*, 35(1):211–231, 2018.

[88] D. M. Seyedi, P. Gehl, J. Douglas, L. Davenne, N. Mezher, and S. Ghavamian. Development of seismic fragility surfaces for reinforced concrete buildings by means of nonlinear time-history analysis. *Earthquake Engineering and Structural Dynamics*, 39(1):91–108, 2010.

[89] M. Alembagheri. Investigating Efficiency of Vector-Valued Intensity Measures in Seismic Demand Assessment of Concrete Dams. *Advances in Civil Engineering*, 2018:12, 2018.

[90] J. Ghosh, J.E. Padgett, and L. Dueñas Osorio. Surrogate modeling and failure surface visualization for efficient seismic vulnerability assessment of highway bridges. *Probabilistic Engineering Mechanics*, 34:189–199, 2013.

[91] S.J. Brandenberg, J. Zhang, P. Kashighandi, Y. Huo, and M. Zao. Demand Fragility Surfaces for Bridges in Liquefied and Laterally Spreading Ground. Technical Report PEER 2011/01, 2011.

[92] P. Gehl, D. Seyedi, J. Douglas, and M. Khiar. Introduction of Fragility Surfaces for a More Accurate Modeling of the Seismic Vulnerability of Reinforced Concrete Structures. *Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, Greece*, (n228), 2009.

[93] P. Gehl, D. M. Seyedi, and J. Douglas. Vector-valued fragility functions for seismic risk evaluation. *Bulletin of Earthquake Engineering*, 11:365–384, 2013.

[94] Y. Pan, A. K. Agrawal, M. Ghosn, and S. Alampalli. Seismic Fragility of Multispan Simply Supported Steel Highway Bridges in New York State. II: Fragility Analysis, Fragility Curves, and Fragility Surfaces. *Journal of Bridge Engineering*, 15:462–472, 2010.

[95] J. A. Goulet. *Probabilistic Machine Learning for Civil Engineers*. Polytechnique, Montreal, 2018.

[96] S. Kameshwar and J.E. Padgett. Multi-hazard risk assessment of highway bridges subjected to earthquake and hurricane hazards. *Engineering Structures*, 78:154–166, 2014.

[97] N. Ataei and J.E. Padgett. Fragility surrogate models for coastal bridges in hurricane prone zones. *Engineering Structures*, 103:203–213, 2015.

[98] G.E. Balomenos and J.E. Padgett. Fragility Analysis of Pile-Supported Wharves and Piers Exposed to Storm Surge and Waves. *journal of waterway port coastal and ocean engineering*, 144(2):04017046, 2017.

[99] P.S. Koutsourelakis. Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework. *Probabilistic Engineering Mechanics*, 25(1):49–60, 2010.

[100] X. W. Yao, A. S. Elnashai, and J. Q. Jiang. Analytical Seismic Fragility Analysis of Concrete Arch Dams. *15th World Conference on Earthquake Engineering, Lisboa*, 2012.

[101] F. Jalayer, R. De Risi, and G. Manfredi. Bayesian Cloud Analysis: efficient structural fragility assessment using linear regression. *Bull Earthquake Eng*, 13:1183–1203, 2015.

[102] M.A. Hariri-Ardebili and V. Saouma. Collapse Fragility Curves for Concrete Dams: Comprehensive Study. *Journal of Structural Engineering*, page 04016075, 2016.

[103] Kevin P. Murphy. *Machine Learning A Probabilistic Perspective*. The MIT Press, 2012.

[104] A.I. Khuri and S. Mukhopadhyay. Response surface methodology. *Computational Statistics*, 2(2):128–149, 2010.

[105] R. Ghanem and P. Spanos. *Stochastic finite elements - A spectral approach.* Springer Verlag, (Reedited by Dover Publications, 2003), 1991.

[106] B. Sudret. Meta-models for structural reliability and uncertainty quantification. *Proc. 5th Asian-Pacific Symp. Struct. Reliab. (APSSRA 2012), Singapore*, 2012.

[107] B. Sudret and A. Der Kiureghian. Comparison of finite element reliability methods. *Probabilistic Engineering Mechanics*, 17:337–348, 2002.

[108] M. Berveiller, B. Sudret, and M. Lemaire. Presentation of two methods for computing the response coefficients in stochastic finite element analysis. *9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, Albuquerque, USA.*, 2004.

[109] M. Berveiller, B. Sudret, and M. Lemaire. Stochastic finite elements: a non intrusive approach by regression. *European Journal of computational mechanics*, 15:81–92, 2006.

[110] S. Choi, R. Grandhi, and R. Canfiel. Structural reliability under non-Gaussian stochastic behavior. *Computers & structures*, 82:1113–1121, 2004.

[111] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering and system safety*, 93:964–979, 2008.

[112] G. Blatman and B. Sudret. An adaptive algorithm to build up sparse polyno- mial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25:183–197, 2010.

[113] M. Ebden. Gaussian Processes for Regression: A Quick Introduction. Technical report, Cornell University, 2008.

[114] C.E. Rasmussen. *Gaussian Processes for Machine Learning.* The MIT Press, 2006.

[115] G. Wang and S. Shan. Review of Metamodeling Techniques in Support of Engineering Design Optimization. *Journal of Mechanical Design*, 129:370, 2007.

[116] W. Simpson, J. Peplinski, P.N. Koch, and J.K. Allen. Metamodels For Computer-Based Engineering Design: Survey and Recommenations. *Research in engineering design*, 17:129–150, 2001.

[117] J. Seo and D.G. Linzell. Use of response surface metamodels to generate system level fragilities for existing curved steel bridges. *Engineering Structures*, 52:642–653, 2013.

[118] E.M. Sichani, J.E. Padgett, and V. Bisadi. Probabilistic seismic analysis of concrete dry cask structures. *Earthquake engineering and structural dynamics*, 73:87–98, 2017.

[119] D. Beeson, G. Wiggs, L. Wang, and S. Akkaram. Gaussian process meta-models for efficient probabilistic design in complex engineering design spaces. *ASME 2005 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, 2005.

[120] F.A. DiazDelaO and S. Adhikari. Structural dynamic analysis using Gaussian process emulators. *International Journal for Computer - Aided Engineering and Software*, 27:580–605, 2010.

[121] Z. Xia and J. Tang. Characterization of structural dynamics with uncertainty by using Gaussian Processes. *ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, 2011.

[122] H. Wan and W. Ren. Parameter Selection in Finite-Element-Model Updating by Global Sensitivity Analysis Using Gaussian Process Metamodel. *Journal of Structural Engineering*, 141:04014164.

[123] S. Mangalathu and J. Jeon. Stripe-Based Fragility Analysis of Concrete Bridge Classes Using Machine Learning Techniques. *arXiv preprint*, (arXiv:1807.09761):32, 2018.

[124] N. Buratti, B. Ferracuti, and M. Savoia. Response Surface with random factors for seismic fragility of reinforced concrete frames. *Structural Safety*, 32(1):42–51, 2010.

[125] A. Gaspar, F. Lopez-Caballero, A. Modaressi-Farahmand-Razavi, and A. Gomes-Correia. Methodology for a probabilistic analysis of an RCC gravity dam construction. Modelling of temperature, hydration degree and ageing degree fields. *Engineering Structures*, 65:99–110, 2014.

[126] J. Mata, N. Leitão Schclar, A. Tavares de Castro, and J. Sá da Costa. Construction of decision rules for early detection of a developing concrete arch dam failure scenario. A discriminant approach. *Computers & structures*, 142:45–53, 2014.

[127] F. Salazar, Toledo M.A., T. Oñate, and Morán R. An empirical comparison of machine learning techniques for dam behaviour modelling. *Structural Safety*, 56:9–17, 2015.

[128] F. Salazar, Morán R., Toledo M.A., and T. Oñate. Data-based models for the prediction of dam behaviour: A review and some methodogical considerations. *Archives of Computational Methods in Engineering*, 24:1–21, 2015.

[129] M.A. Hariri-Ardebili and F. Pourkamali-Anaraki. Support vector machine based reliability analysis of concrete dams. *Soil dynamics and earthquake engineering*, 104:276–295, 2018.

[130] M.A. Hariri-Ardebili and F. Pourkamali-Anaraki. Simplified reliability analysis of multi hazard risk in gravity dams via machine learning techniques. *Archives of civil and mechanical engineering*, 18:592–610, 2018.

[131] M.A. Hariri-Ardebili. MCS-based response surface metamodels and optimal design of experiments for gravity dams. *Structure and infraestructure engineering*.

[132] L. H. Nguyen and J.-A. Goulet. Anomaly detection with the Switching Kalman Filter for structural health monitoring. *Structural Control and Health Monitoring*, 25:e2136, 2018.

[133] Barbara Mills-Bria, Anastasia Johnston, and Phoebe Percell. Nonlinear Analysis Research. Technical report, U.S Department of the interior Bureau of Reclamation, 2008.

[134] ANSYS Inc. *ICEM CDF, Release 16.0, software.* Canonsburg, PA, 2014.

[135] Livermore Software Technology Corporation. *LS-Dyna R9.1.0, software.* Livermore, CA, 2013.

[136] D. Depolo, E. Kennedy, T. Walker, and R. Tom. Analysis and Design of Large-Scale Civil Works Structures Using LS-DYNA. *11th International LS-Dyna Users Conference*, 2010.

[137] Gravity Dam Design. Technical Report EM 1110-2-2200, U.S. Corps of Engineer, 1995.

[138] H.K. Gupta and B.K. Rastogi. *Dams and eartquakes.* Elsevier Science, 1 edition, 1976.

[139] Uplift pressures, shear strengths, and tensile strengths for stability analysis of concrete gravity dams. Technical Report TR-100345, Electrical Power Research Institute (EPRI), 1992.

[140] S. Halchuk, T. Allen, J. Adams, , and G. Rogers. Fifth Generation Seismic Hazard Model Input Files as Proposed to Produce Values for the 2015 National Building Code of Canada. Technical Report Open File 7575, Geological Survey of Canada, 2014.

[141] Global Earthquake Model Foundation GEM. *OpenQuake Engine 2.8, software.* 2015.

[142] K. Tarbali and B.A. Bradley. The effect of causal parameter bounds in PSHA-based ground motion selection. *Earthquake Engineering & Structural Dynamics*, 45:1515–1535, 2016.

[143] Jekabsons G. *VariReg: A software tool for regression modelling using various modelling methods, 2010, available at http://www.cs.rtu.lv/jekabsons/.* 2010.

[144] Jekabsons G. *ARESLab: Adaptive Regression Splines toolbox for Matlab/Octave, 2016, available at http://www.cs.rtu.lv/jekabsons/.* 2016.

[145] G. Jekabsons. Adaptive basis function construction: an approach for adaptive building of sparse polynomial regression models. In Y. Zhang, editor, *Machine learning*, pages 127–155. IntechOpen, 2010.

[146] Matlab optimization toolbox, R2017a, Version 11.1.

[147] A. Hadjian. Earthquake reliability of structures. *13th World Conference on Earthquake Engineering, Vancouver, Canada*, 2004.

[148] Minimum Design Loads for Buildings and Other Structures. Technical Report ASCE/SEI Standard 7-16, American Society of Civil Engineers (ASCE), 2016.

[149] J.G. MacGregor. Safety limit states design for reinforced concrete. *Canadian journal of civil engineering*, 3:484–513, 1976.

[150] Federal Guidelines for Dam Safety Risk Management. Technical Report P-1025. Federal Guidelines for Dam Safety Risk Management, Federal emergency management agency (FEMA), 2015.

[151] Y. Ghanaat, R.C. Patev, and A.K. Chudgar. Seismic Fragility for Risk Assessment of Concrete Gravity Dams. *35th USSD Annual Meeting and Conference, Louisville*, 2015.

[152] B.A. Bradley, R.P. Dhakal, M. Cubrinovski, J.B. Mander, and G.A. MacRae. Improved seismic hazard modelwith application to probabilistic seismic demand analysis. *Earthquake Eng. & Structural Dynamics*, 36:2211–2225, 2007.

[153] B.A. Bradley. A ground motion selection algorithm based on the generalized conditional intensity measure approach. *Soil Dynamics and Earthquake Engineering*, 40:48–61, 2012.

[154] Z. Gulerce and N.A. Abrahamson. Site-Specific Design Spectra for Vertical Ground Motion. *Earthquake Spectra*, 27:1023–1047, 2011.

[155] J.W. Baker and N. Jayaram. Correlation of spectral acceleration values from NGA ground motion models. *Earthquake Spectra*, 24:299–317, 2008.

[156] J.W. Baker and C.A. Cornell. Correlation of response spectral values for multicomponent ground motions. *Bulletin of the Seismological Society of America*, 96:215–227, 2006.

[157] B.A. Bradley. Empirical correlations between peak ground velocity and spectrumbased intensity measures. *Earthquake Spectra*, 28:17–35, 2012.

[158] J.J. Boomer, P.J. Stafford, and S. Akkar. Current empirical ground-motion prediction equations for Europe and their application to Eurocode 8. *Bulletin of Earthquake Engineering*, 8(1):5, 2009.

[159] T.D. Ancheta, R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, K.E. Chiou, B.S.J. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. PEER NGA-West2 Database. Technical Report PEER 2013/03, Pacific Earthquake Engineering Research Center, 2013.

[160] C.A. Goulet, T. Kishida, T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. PEER NGA-East Database. Technical Report PEER 2014/17, Pacific Earthquake Engineering Research Center, 2014.

[161] J.W. Baker and C.A. Cornell. Spectral shape, epsilon and record selection. *Earthquake Eng. & Structural Dynamics*, 35:1077–1095, 2006.

[162] N. Luco and P. Bazzurro. Does amplitude scaling of ground motions records result in biased nonlinear structural drift responses? *Earthquake Eng. & Structural Dynamics*, 36:1813–1835, 2007.

[163] B.A. Bradley. A critical examination of seismic response uncertainty analysis in earthquake engineering. *Earthquake Engineering & Structural Dynamics*, 42:1717–1729, 2013.

[164] R. Monteiro. Sampling based numerical seismic assessment of continuous span RC bridges. *Engineering Structures*, 118:407–420, 2016.

[165] C. Zelaschi, R. Monteiro, and R. Pinho. Parametric Characterization of RC Bridges for Seismic Assessment Purposes. *Structures*, 7:14–24, 2016.

[166] E. Miranda, L. Eads, H. Davalos, and D. Lignos. Assessment of the probability of collapse of structures during earthquakes. *16th World Conference on Earthquake Engineering. Santiago Chile*, 2017.

[167] S. Ghosh, S. Ghosh, and S. Chakraborty. Seismic fragility analysis in the probabilistic performance-based earthquake engineering framework: an overview. *International Journal of Advances in Engineering Sciences*, 2017.

[168] A.H.M. Muntasir Billah and M. Shahria Alam. Seismic fragility assessment of highway bridges: a state-of-the-art review. *Structure and Infrastructure Engineering*, 11:804–832, 2015.

[169] M. Grigoriu and E. Mostafa. Fragility surfaces as a measure of seismic performance. *The 7th US National Conference on Earthquake Engineering, Boston*, 2002.

[170] J.H. Friedman. Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1):1–67, 1991.

[171] R.L. Hardy. Multiquadric Equations of Topology and Other Irregular Surfaces. *Journal of Geophysical Research*, 76(8):1905–1916, 1971.

[172] Navid Ataei. *Vulnerability Assessment of Coastal Bridges Subjected to Hurricane Events*. PhD thesis, Rice University, 2013.

[173] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Number 57 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, USA, 1993.

[174] G. Barnston, Anthony. Correspondence among the Correlation, RMSE, and Heidke Verification Measures; Refinement of the Heidke Score. *Notes and Correspondence, Climate Analysis Center*, 1992.

[175] R.L. Segura, R. Monteiro, and P. Paultre. Improved seismic fragility assessment of concrete gravity dams in eastern Canada. *Canadian Dam Association 2018 Annual Conference*, Québec, QC, Canada, 2018.

[176] R.L. Segura, J.E. Padgett, and P. Paultre. Polynomial Response Surface-Based Seismic Fragility Assessment of Concrete Gravity Dams. *12th Canadian conference on earthquake engineering*, Québec, QC, Canada, 2018.

[177] National Research Council of Canada (NRCC). *National Building Code of Canada 2010.* 2010.

[178] G.M. Atkinson and D.M. Boore. Ground-Motion Relations for Eastern North America. *Bulletin of the Seismological Society of America*, 85:17–30, 1995.