THE SEISMIC ASSESSMENT OF EXISTING CONCRETE GRAVITY DAMS: FE MODEL UNCERTAINTY QUANTIFICATION AND REDUCTION

Dissertation

submitted to and approved by the

Faculty of Architecture, Civil Engineering and Environmental Sciences
University of Braunschweig – Institute of Technology

and the

Department of Civil and Environmental Engineering
University of Florence

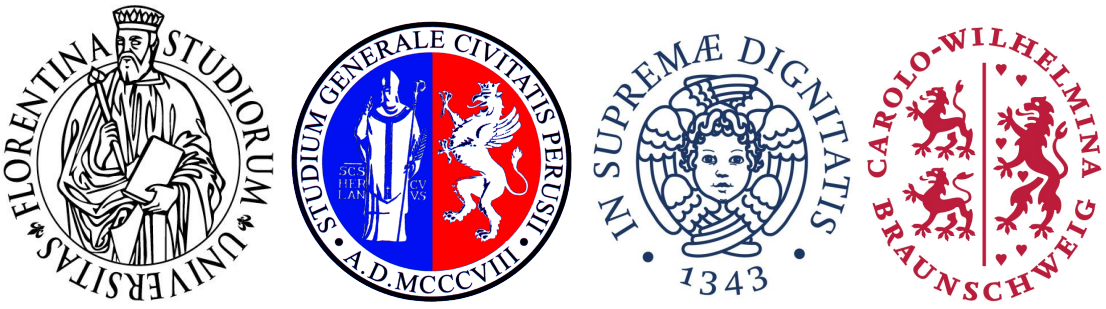
in candidacy for the degree of a
Doktor-Ingenieur (Dr.-Ing.) /
Dottore di Ricerca in Civil and Environmental Engineering*)

by
Giacomo Sevieri
born 27/03/1988
from Lucca, Italy

Submitted on	11 February, 2019
Oral examination on	7th May, 2019
Professorial advisors	Prof. Hermann G. Matthies Prof. Anna De Falco

2019

*) Either the German or the Italian form of the title may be used.



Universities of Florence, Perugia, Pisa, Braunschweig

Department of Civil and Environmental Engineering
International Doctorate in Civil and Environmental Engineering - XXXI cycle

The seismic assessment of existing concrete gravity dams: FE model uncertainty quantification and reduction

PhD candidate: Giacomo Sevieri

Italian Supervisors
Prof. A. De Falco

German Supervisor
Prof. H. G. Matthies

2019

Abstract

Most of the dams around the world were designed before the introduction of seismic regulations and without concerns about their dynamic behaviour. The failure of a large concrete gravity dam might have catastrophic effects and its partial or total collapse can give rise to uncontrolled water releases, putting at risk a large number of human lives, not counting the considerable economic consequences. However, existing concrete gravity dams are an important source of energy impossible to give up to, hence their preservation, also against seismic events, is a fundamental task for our society. Since there are no case histories of concrete gravity dams failed after seismic events, numerical models assume great importance for the evaluation of the seismic performance of such structures. If integrated in a Structural Health Monitoring (SHM) system, numerical models are also a powerful tool to control the structural behaviour of the dam during its regular use, and to detect possible damage after seismic events.

Several different sources of uncertainty are involved in concrete gravity dams modelling, e.g. geometrical uncertainties, epistemic uncertainties related to the material constitutive models, Soil-Structure-Interaction modelling approach. The effects of epistemic uncertainties related to the material constitutive models can be reduced by exploiting all available information about the dam. For this purpose, material test results are fundamental source of information, even though they are not sufficient to properly calibrate a reliable FE model of the dam. Measurements recorded on the dam body, e.g. displacements or ambient vibrations, represent important sources of information to calibrate the parameters of dam twin models used to reproduce a particular Quantity of Interest (QI) of the dam.

In this research work two probabilistic SHM frameworks, defined in the Bayesian setting, for concrete gravity dams are proposed. The first procedure allows calibrating the parameters of a static twin model of the dam by using static displacement records and environmental measurements. In this way, epistemic uncertainties related to the mechanical parameters

of the materials are reduced. The calibrated static twin model is integrated in a SHM system in order to control the dam behaviour during its regular use.
The second proposed SHM framework allows both calibrating the parameters of a dynamic twin model of the dam, by using ambient vibrations thus reducing the effects of the epistemic uncertainties, and controlling the health state of the structure. In particular, the proposed procedure exploits the results of Operational Modal Analysis (OMA) techniques which allows determining the experimental modal characteristics of the system by elaborating ambient vibrations. The use of modal characteristics as reference QI in the updating process allows reducing the computational burden. However, in this context the coherence between experimental and numerical modes must be ensured. For this purpose a modified version of the numerical algorithm Markov Chain Monte Carlo (MCMC) is proposed. Indeed, by considering a reordering criterion within MCMC the numerical modes can be matched to the experimental ones, discarding those with no correlation. The resulting calibrated dynamic twin model of the dam is used to predict the seismic behaviour of the structure through the fragility curves calculation. The reduction of epistemic uncertainties, obtained using the proposed procedure allows improving the reliability of the fragility estimation of the dam. The predictive model of the dam mode shapes is integrated in a SHM system for damage detection purpose. Indeed, by exploiting the sensitivity of mode shapes with respect to damage, the predictive model allows detecting anomalous phenomena.
Finally, a procedure, based on the Optimal Bayesian Experimental Design (OED), which allows designing the SHM system layout maximizing the probability of damage detection, is proposed. Moreover, by exploiting the relationship between damage development and strength parameters of materials constitutive models, these latter can be updated.
Probabilistic procedures are generally expensive in terms of computational cost, which becomes prohibitive when complex FE models are involved. Therefore in this research work, meta models based on the general Polynomial Chaos Expansion (gPCE) technique are widely used to reduce the computational burden.
A study of the main modelling issues in the seismic analysis of concrete gravity dams has been done in order to recognize the main uncertainty sources whose reduction is necessary. In particular, the effects of the choice of the material constitutive models of the SSI modelling approach and of the FSI modelling approach on the seismic analysis result are investigated.
A wide literature review on the seismic fragility assessment of concrete gravity dams and the use of SHM system records for model parameters calibration is presented in order to contextualize the present research work. Moreover, the theoretical background behind the

statistical tools used in this work is introduced in order to provide the reader with the basic knowledge for understanding the present research work.
In the last part of the thesis, two real Italian concrete gravity dams are analysed in order to test the proposed procedures. The numerical results highlight the possibility to apply the procedures developed in this research work to real cases, thus improving both the estimation of the structural fragility and the dam control.

Acknowledgements

I want to thank my mentors Prof. Anna De Falco and Prof. Hermann G. Matthies for all they taught to me during the PhD, and for the possibility they gave me to undertake this journey.
There are plenty of people who I would like to thank, in particular Prof. Massimiliano Lucchesi and Prof. Nicola Zani, who helped me to thoroughly understand constitutive models of materials; Prof. Bojana Rosić, a true friend who supported me in codes development and probability theory; Dr. Marco Andreini for the help in the definition of the probabilistic models; Dr. Giuseppe Chellini for help in understanding the Operational Modal Analysis techniques; Mr. Matteo Mori who besides being a friend also contributed to the development of the numerical models.
Finally, obviously, part of the merit of this work goes to all friends and my family, who have always endured me during these years.

"Probability is the most important concept in modern science, especially as nobody has the slightest notion of what it means".

Bertrand Russell, 1929.

Contents

List of Figures

List of Tables

List of Symbols

$\gamma(\cdot)$	Correction term of the additive probabilistic model
$\mathbb{E}[\bullet]$	Expected value
$\mathbb{P}$	Probability measure
$\mathbf{I}$	Multi-index set
$\mathbf{x}$	Measurable variables of the probabilistic model
$\mathfrak{F}$	σ-algebra
Ω	Space of all events
ω	Event in the space Ω
$\mathbf{\Phi}$	Mode shapes matrix
$\mathbf{\Psi}$	Orthogonal polynomials matrix
$\mathbf{\Theta}$	Vector of all unknown probabilistic model parameters
$\boldsymbol{\lambda}$	Vector of the covariance functions coefficients λ^m
$\boldsymbol{\theta}_\gamma$	Combination coefficients of the explanatory functions
$\boldsymbol{\theta}_{\mathrm{FA}}$	Vector of the Fourier coefficients
$\boldsymbol{\theta}_{\mathrm{gPCE}}$	Vector of gPCE parameters
$\boldsymbol{\theta}_{\mathrm{m}}$	Vector of unknown probabilistic model parameters
$\boldsymbol{\theta}_{\mathrm{s}}$	Vector of unknown material strength parameters
$\boldsymbol{\varpi}$	Seismic action parameters vector

$\boldsymbol{\Xi}$	Vector of the hyperparameters
Ψ_α	Orthogonal polynomials
σ	Standard deviation of the probabilistic model error
Σ_f	Covariance matrix of the error terms of the frequencies predictive model
Σ_ϕ	Covariance matrix of the error terms of the mode shapes predictive model
$\boldsymbol{s}$	Vector of deterministic model samples
$\boldsymbol{v}$	Matrix of polynomial coefficients in the projection equation
$\boldsymbol{W}$	Sensing matrix
$\mathbf{d}$	Vector of the design variables
$\mathbf{f}$	Frequencies vector
$\mathbf{r}$	Residual vector
$\mathbf{w}_d$	Vector of the covariance functions coefficients w_d^m
$\mathbf{y}$	Vector of new observations
δ	Total static displacement of the dam
δ^H	Static displacement of the dam due to the basin level variation
δ^H	Thermal part of the total static displacement of the dam
$L(\cdot)$	Likelihood function
P	Polynomial expansion cardinality
$S_a(\cdot)$	Spectral acceleration
$S_d(\cdot)$	Spectral displacement
$S_v(\cdot)$	Spectral velocity
u	Model response
$U(\cdot)$	Expected utility

$u^{(\alpha)}$	PCE coefficients
u_N	PCE of u
CIA	Cumulative Inelastic Area
CID	Cumulative Inelastic Duration
CLA	Cloud Analysis
DCR	Demand Capacity Ratio
DI	Damage Index
DSDR	Damage Spatial Distribution Ratio
DV	Damage Variable
ETA	Endurance Time Analysis
FE	Finite Elements model
FEA	Finite Element Analysis
FSI	Fluid Structure Interaction
gPCE	general Polynomial Chaos Expansion
HMS	Health Monitoring System
IDA	Incremental Dynamic Analysis
IM	Intensity Measure
LS	Limit State
MAC	Modal Assurance Criterion
MCE	Maximum Credible Earthquake
MSA	Multiple Stripe Analysis
OED	Optimal Bayesian Experimental Design
OED	Optimal Experimental Design

OMA	Operational Modal Analysis
PGA	Peak Ground Acceleration
PGV	Peak Ground Velocity
PMF	Potential Maximum Flood
RV	Random Variable
SHM	Structural Health Monitoring
SPSA	Simultaneous Perturbation Stochastic Approximation algorithm
SSI	Soil Structure Interaction

Chapter 1

Introduction

1.1 General overview of the topic and motivations

Concrete gravity dams are an important resource for our community due to their use for energy production, floods control, industrial and agricultural purposes and much else. However, due to economical and environmental constraints, in developed countries new dams are no longer being built. The largest part of the existing concrete gravity dams have been built before the introduction of seismic regulations or they have been designed without concerns about their dynamic behaviour. The existing ones are ageing and their deterioration levels are reaching critical values. The American Society of Civil Engineers (ASCE) 2013 report card for America's infrastructure (American Society of Civil Engineers 2013), which categorized the hazard related to US dams, classified 14'700 dams as high risk, 12'400 as significant risk, 59'000 low risk and 1300 as undetermined. The number of high-hazard dams is increased by nearly 40% over the past decade (USACE - U.S Army Corps of Engineers 2015). Following the Association of State Dam Safety Officials (ASDSO) report by 2020, the 70% of the US dams will be over 50 years old (ASDSO 2011) and most of them are unlikely to safely withstand current design guidelines for potential maximum floods (PMF) and maximum credible earthquakes (MCE).

As for Europe, and in particular high seismic areas as Italy, some similarities with the American case can be found. Around 500 dams of different typologies are in this country and among them about 90% have been built before the 1980s and mainly designed following static concepts.
Considering only the concrete gravity dams, none of them has ever collapsed after seismic events, neither in Italy or in other parts of the world, as reported by Hall (1988) and Zhang et al. (2016). However, due to the important role of such structures and the catastrophic effects, in terms of human and economic losses, that the collapse of a dam could lead to, the preservation and control of concrete dams are important tasks of our society. Numerical models are powerful tools to understand and to investigate the seismic behaviour of concrete dams, in particular because of the absence of case histories. Moreover, numerical models can be used within a Structural Health Monitoring (SHM) system to predict the dam behaviour, thus improving its control.
Several different sources of uncertainty, which are usually categorized as epistemic and aleatory, influence the numerical model results. Pragmatically, the main difference between epistemic and aleatory uncertainties is the possibility of the analyst to reduce them once new information are available, in the case of espitemic ones, or the impossibility to update the state of knowledge, in the case of aleatory uncertainty, due to their randomness nature (Der Kiureghian and Ditlevsen 2009). In the field of modelling for seismic assessment of existing concrete gravity dams, once a particular deterministic model has been defined, the main source of aleatory uncertainty is the seismic action, while the main epistemic uncertainty sources can be categorized as:

- **Geometric uncertainty**: related to differences between the project and its development.
- **Material constitutive models**: materials behaviour in the deterministic model.
- **Model Parameters**: material constitutive models parameters, boundary conditions, damping etc..

Moreover, the choice of analysis method and verification criteria leads to other uncertainties in the final result. Regarding this point, international codes and scientific community propose different approaches. The scientific community, in recent years, has been developing reliable procedures to assess the structural fragility of concrete gravity dams by using analysis methods generally developed for other kind of structures (Hariri-Ardebili and Saouma 2016c). The second approach is proposed by regulations, that suggest the

use of deterministic biased models which are corrected by safety factors related to specific kind of analyses.
Epistemic uncertainties need to be reduced as much as possible by using all available information on the structure. The observations recorded by the monitoring system are important information on the structural behaviour and they can be used not only to assess the health of the structure but also to calibrate the model parameters (ICOLD 2000). The typology, the quality, as well as the devices position, lead to acquire different kind of information and the possibility to update the state of knowledge about different uncertainties. By introducing the information coming from the SHM into the structural analysis, twin models of the dam can be calibrated. These models can be used to predict the structural behaviour during its regular use or during seismic events, thus improving the structural safety.
Only few applications of SHM to concrete gravity dams can be found in the literature, the majority of which are performed in a deterministic setting.
The main aim of this thesis is the definition of suitable SHM systems for concrete dams by combining structural analysis, statistical tools, static and dynamic records, with the aims to control the structural behaviour and to calibrate twin models of the dam. Regarding this latter purpose, a calibrated twin model of the dam can be used both to improve the reliability of the SHM systems themselves or to improve the estimation of the structural fragility. Indeed in this context, a twin model is a mathematical expression which simulates the structural behaviour, static or dynamic, as a function of some parameters. Finite Element model as well as probabilistic predictive models can be considered as twin, and they can be used in different way as real time structural control or dam behaviour prediction. Statistical tools as Bayesian inference are used to set up a robust framework to reduce the uncertainties. However, probabilistic procedure are computationally expensive and they become prohibitive if applied in context where complex FE models are commonly used, as in dam engineering field. Therefore, in this research work meta models are used to reduce the computational burden, thus speeding up the proposed procedures.
Two real Italian concrete gravity dams are analysed in order to test the proposed procedures. In both examples real measurements recorded on the dam bodies are used to update the state of knowledge. Since there are no information about the behaviour of the damaged dams, it is simulated through high-fidelity models of the structures. In this context, a high-fidelity model is a very detailed numerical model of the dam which allows accurately simulating the structural behaviour.

1.2 Objectives

The aim of this research work is the definition of a suitable SHM framework for static and dynamic control of concrete dams, by combining structural analysis, monitoring systems and statistical tools. The final SHM framework should be able both to assess the structural health of the dam and to reduce the epistemic uncertainties of the twin models. Since the twin models are used to predict the dam behaviour, the reliability of the SHM and the estimation of the structural fragility are improved. The proposed research work, framed in a probabilistic setting, aim to:

- *Static SHM system*: calibrate the parameters of the dam predictive model in order to simulate the structural behaviour during normal operations, and control its health state.
- *Dynamic SHM system*: calibrate the parameters of the dam predictive model in order to simulate the structural behaviour during seismic events, thus improving the estimation of its seismic fragility and controlling its health state during and after seismic events.
- *Devices Optimization for dynamic SHM*: optimise the dynamic SHM system in order to both maximize the probability of damage detection and acquire information for the reduction of epistemic uncertainties.

Probabilistic procedures based on Bayesian inference have been developed using static and dynamic measurements recorded by the monitoring system as source of information.
The application to real cases highlights the effectiveness of the proposed methodologies in reducing the epistemic uncertainties involved in the seismic assessment of existing concrete gravity dams.

1.3 Novelties of the work

The novelties of this research work are the following.

- The development of a procedure within a Bayesian framework, to calibrate the model parameters using displacements recorded during normal operations, environmental data and results of materials tests. The calibrated model can be employed to control the health status of the dam and its daily behaviour. The procedure is applied to a real large Italian dam, as case of study.

- The development of a procedure within a Bayesian framework, in order to calibrate the model parameters using ambient vibrations and results of materials tests. The calibrated model can be used to predict the behaviour of the dam during seismic events. In this way the epistemic uncertainties in the fragility analysis are reduced. The procedure is applied to a real large Italian dam as case of study.
- The development of a procedure based on the Optimal Bayesian Experimental Design (OED), to optimise the monitoring system layout with respect to the damage detection and, to the acquisition of information useful for material strength parameters updating.

In this context, the general Polynomial Chaos Expansion (gPCE) technique is used to reduce the computational burden, which is usually prohibitive when complex FE models are involved.

1.4 Structure of the thesis

This thesis consists in 9 Chapters. The main contents of each Chapter is summarized as follows:

- **Chapter 1**
  A general overview of the topic with the motivations is presented. The basic concepts and the main novelties regarding the seismic assessment of concrete gravity dams developed in this work, are briefly introduced.
- **Chapter 2**
  An overview of the literature contributions on the seismic assessment and Structural Health Monitoring of concrete gravity dams is provided.
- **Chapter 3**
  The main modelling issues in the seismic analysis of concrete gravity dams are shown. This includes a study of the influence of material constitutive models, the soil-structure interaction and the fluid-structure interaction on the model output.
- **Chapter 4**
  The theoretical background of the statistical tools used to develop the present research work is introduced. In particular, the Bayesian inference, the gPCE technique and the Optimal Bayesian Experimental Design are discussed.

- **Chapter 5**
  The proposed static SHM framework, based on data recorded by the monitoring system during normal operations of the dam, is described.

- **Chapter 6**
  The proposed method to calibrate the dynamic model parameters using ambient vibration data is presented.

- **Chapter 7**
  The proposed dynamic SHM framework and the optimization procedure for designing it are presented.

- **Chapter 8**
  The analysis of two case studies are presented.

- **Chapter 9**
  This part collects conclusions of the thesis. The main goals achieved in the thesis are summarised and suggestion for future developments are given.

Chapter 2

Literature background and main contributions

2.1 Introduction

Among the infrastructures being part of the assets of a country, dams are fundamental due to their use for energy production, flood control and agricultural-industrial sustenance. However, they are characterized by a high-risk level against seismic events since the largest part of them have been designed without concerns about their dynamic behaviour. Because of their importance for the community, existing concrete gravity dams must be preserved. In this regard numerical models are a powerful tool since they can be used to control the structural behaviour and to predict it during seismic events. However, epistemic and aleatory uncertainties are involved in this process. Therefore, the aims of this research project are: the development of probabilistic SHM systems which allows reducing the epistemic uncertainties and to control the structural behaviour. This research work can be inserted in the research field of dam engineering and SHM. The present Chapter aims to review the main research contributions related to the seismic assessment of concrete gravity dams and the applications of SHM, static or dynamic, to concrete gravity dams.

With the main objective of framing the problem by identifying the major research trends, the Chapter is organized as follows:
Section 2.2 analyses the main issues related to the structural fragility assessment of existing concrete gravity dams and how in literature those problems are usually faced.
Section 2.3 describes how monitoring systems of existing concrete gravity dams can be used to control the structural behaviour and to detect damage. Regarding the latter point, a review of the most common damage indexes available in literature for concrete gravity dams is presented.
Section 2.4 illustrates the literature gap in the field of seismic assessment of existing concrete gravity dams.

2.2 Seismic fragility analysis of concrete gravity dams

2.2.1 Introduction

One of the main motivations of this work is the improvement of the existing concrete gravity dams control by performing FE twin models which can be used to predict the structural behaviour during its regular use and during seismic events. This last aspect can be collocated in the research line of seismic fragility assessment of existing concrete gravity dams.
The literature review presented in this section focuses the attention on the application of common fragility assessment methodology to the case of concrete gravity dams, with particular emphasis on the choice of the analysis typology, the epistemic and aleatory uncertainties and definition of the limit state functions.

2.2.2 Fundamental concepts

In this section the fundamental idea behind the fragility assessment of structures is presented. The concept of structural fragility can be defined in a non technical way as *"the quality of a structure or a structural component, of being easily broken or damaged"*. The introduction of fragility function in the seismic engineering field is due to Kennedy et al. (1980). The authors defined a fragility function as a probabilistic relationship between frequency of failure of a component of a nuclear power plant and peak ground acceleration in an earthquake. In the field of seismic engineering the fragility function can be defined as a mathematical relation which expresses the probability that some undesirable events

occurs, i.e. the structure or a structural component reaches a limit state (LS), as a function of a measurable characteristic of the earthquake (IM), i.e. peak ground acceleration, magnitude and so on. Alternatively, a fragility function can be defined as the cumulative distribution of the capacity of a structure to resist an undesirable limit state (Porter 2003). The previous definitions justify the use of fragility functions, not only to assess the structural vulnerability, but also to predict the structural behaviour during seismic events. Moreover, fragility functions allow working in a probabilistic setting considering both epistemic and aleatory uncertainties. As discussed in Porter (2003), fragility functions are commonly used to assess the structural vulnerability, as shown in figure 2.1. Once a par-

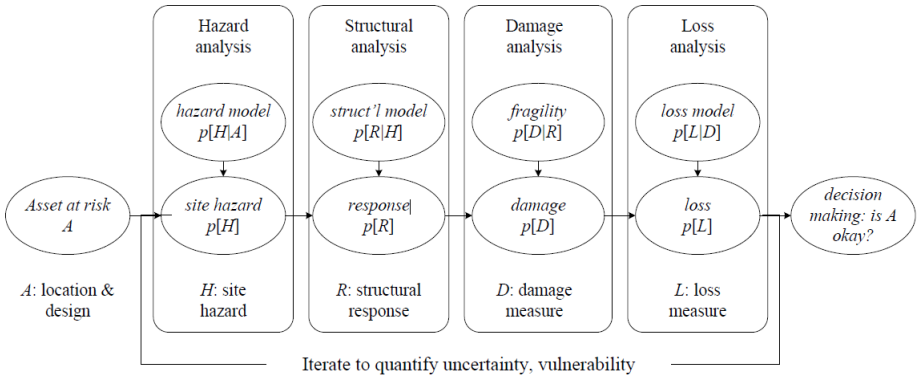


Fig. 2.1: Analytical methods for estimating seismic vulnerability of a single asset (Porter 2003).

ticular LS has been defined as well as the IM, the fragility function F_{LS}, related to the selected LS, can be defined as

$$F_{\mathrm{LS}}\left(im\right) = P\left[\mathrm{LS}|\mathrm{IM} = im\right], \tag{2.1}$$

where im is a particular value if IM. Alternatively, by defining a demand parameter D and the structural capacity C_{LS} associated with the given LS, the fragility function F_d becomes

$$F_d\left(im\right) = P\left[D \geq d|\mathrm{IM} = im\right], \tag{2.2}$$

where d is a particular value of D. In some real applications multiple fragility functions must be defined, since a structure or its components could experience several different limit states, sequentially, simultaneously or in a mutual exclusive way (Porter 2017).
Usually the probability of failure is assumed to be log-normally distributed (it is not the only possible assumption or the best one or the more coherent with the reality, but it is

just the most common one). In this way the Cumulative Density Function (CDF) can be written as

$$F_{\mathrm{LS}}(im) = P[\mathrm{LS}|\mathrm{IM} = im] = \Phi\left[\frac{\ln(im) - \ln(\eta)}{\beta}\right], \tag{2.3}$$

where $\Phi[\cdot]$ is the standard normal CDF, β is the logarithmic standard deviation (also called dispersion) and η is the median of the fragility function.
Fragility functions can be derived based on one of the following approaches (Porter 2017):

- **Empirical**: An empirical fragility function is created by fitting a function to the observation recorded during experimental tests or in the real world. Usually, this kind of fragility functions are built starting from the observations acquired during post-earthquake survey, obviously this fragility functions lack of generality (Muntasir Billah and Shahria Alam 2015).

- **Analytical**: An analytical fragility function is based on the results of numerical models. In this case particular attention must be placed on the choice of the analysis typology and how uncertainties are considered in the model.

- **Heuristic**: A heuristic fragility function is based on expert judgements, a pool of experts guess or judge failure probability as a function of environmental excitation. This approach is useful when empirical observations are very limited (ATC-13 1985).

- **Hybrid**: The above approaches can be combined to build fragility functions, reducing the computational effort but controlling the bias.

in the context of analytical fragility functions the choice of the analysis typology is a fundamental aspect. In particular the analysis typology should be chosen balancing the computational burden and the accuracy of the result.
Following Hariri-Ardebili and Saouma (2016c), the analysis typologies can be categorized in a progressive way, figure 2.2, in every analysis typology the uncertainties play a different role and so they have a different weight.
In the application of the analytical definition of the fragility functions, equation 2.3, the calibration of η and β, by fitting the analytical data point is fundamental. Therefore the choice of the fitting methodology is an important aspect as indicated in Baker (2015).
The combination of aleatory and epistemic uncertainties in the seismic assessment of structure is one of the biggest issue. In this context, the epistemic uncertainties are these related to the parameters of material models, while the aleatory ones are related to the record-to-record variation. The best approach would suggest to use methods able to account for

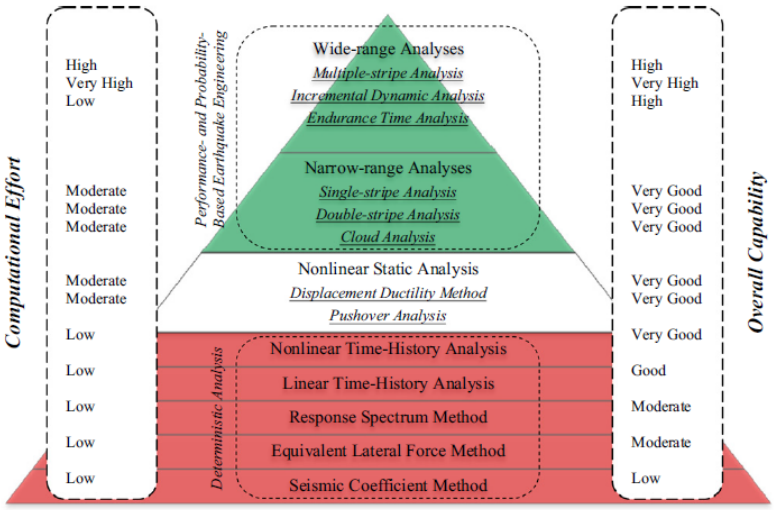


Fig. 2.2: Progressive analysis methodology for concrete dams (Hariri-Ardebili et al. 2014).

both uncertainties, like sampling based methods (Celarec and Dolšek 2013) or spectral approach (Xiu 2010). In this regard, some examples are available in the literature, Celik and Ellingwood (2010) analysed non-ductile reinforced concrete frames by using a sampling based method to incorporate aleatory and epistemic uncertainties, the authors proposed a confidence bounds on the fragilities as a measure of their accuracy. However, these methods lead to a high computational cost, so if complex numerical models are involved these methods become prohibitive. Therefore, simplified methods have been proposed in literature. Usually these methods face this issue either by simplifying the deterministic model or combining the mean values and the variances calculated separately for epistemic and aleatory uncertainties. The combination is related to some measure of the gradient of each random variable, likewise the First Order Reliability Method (FORM) and the First Order Second Moment (FOSM) reliability method (Ditlevsen and Madsen 2007). Several examples are available in literature, Dolšek (2012) considered epistemic uncertainties by using non linear static analyses, which were incorporated with the aleatory ones by using incremental dynamic analysis (IDA) performed on an equivalent single degree of freedom (SDOF) model of the structure. Vamvatsikos and Fragiadakis (2010) compared the fragility curves obtained by using: (a) a Monte Carlo simulation with Latin Hypercube sampling, (b) point estimate and (c) FOSM technique, showing the difference in the final results. In the same work, as well as in Cornell et al. (2002) and Celarec and Dolšek (2013), the two uncertainty sources were combined through the well known square-root-sum-of-square

rule,

$$\beta_{RU} = \sqrt{\beta_R^2 + \beta_U^2}. \tag{2.4}$$

In equation 2.4, β_R is the dispersion related to the aleatory uncertainty, β_U is related to the epistemic uncertainty and the β_{RU} is their combination. Liel et al. (2009) developed two simplified methods, one which reduces the computational burden by approximating the model output through a response surface. Whereas, the second method estimates the gradients related to the epistemic uncertainties with a variational approach similar to the FOSM. This latter is explained in detail since it has been used in this work. The approach called Approximated Second Order Second Moment (ASOSM) method allows calculating the fragility curve considering both epistemic and aleatory uncertainty as a combination of the curves calculated considering only the record-to-record variation or the uncertainty in the model parameters. Once the fragility curve related only to the aleatory uncertainty is calculated, the shifts in terms of mean values and dispersions, due to the epistemic uncertainties, are determined by perturbing every random variables $\pm 1.7\sigma$ $\left(X_i^+ = X_i + 1.7\sigma \quad X_i^- = X_i - 1.7\sigma\right)$. Likewise FOSM, the gradients represent the average slope about the mean

$$\frac{\partial g\left(X\right)}{\partial X_i} = \frac{\Delta \mu_{\ln,\mathrm{Sa,col}}}{\Delta X_i} = \frac{\mu_{\ln,\mathrm{Sa,col}}\left(X_i^+\right) - \mu_{\ln,\mathrm{Sa,col}}\left(X_i^-\right)}{X_i^+ - X_i^-} \tag{2.5}$$

where $g\left(X\right)$ is the collapse capacity and $\mu_{\ln,\mathrm{Sa,col}}$ is the mean value of the log-normal collapse capacity. Once the gradients are defined, the dispersion related the epistemic uncertainties $\sigma^2_{\ln,\mathrm{mod}}$ can be computed,

$$\sigma^2_{\ln,\mathrm{mod}} = \left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{\partial g\left(X\right)}{\partial X_i}\frac{\partial g\left(X\right)}{\partial X_j}\right]\rho_{i,j}\sigma_i\sigma_j\right]. \tag{2.6}$$

In equation 2.6, $\rho_{i,j}$ is the correlation coefficient between the variables i-th and j-th, σ_i and σ_j are the standard deviations of the random variables i-th and j-th. The dispersion $\sigma^2_{\ln,\mathrm{mod}}$ is combined with the one related to the record-to-record variation, following the SRSS rule

$$\sigma^2_{\ln,\mathrm{total}} = \sigma^2_{\ln,\mathrm{mod}} + \sigma^2_{\ln,\mathrm{RTR}}. \tag{2.7}$$

The response asymmetry is given by the parameter Δ^+/Δ^-,

$$\Delta^+/\Delta^- = \frac{\hat{m}^+/\hat{m}}{\hat{m}/\hat{m}^-}, \tag{2.8}$$

where $\hat{m}$ is the median capacity of the model with mean model parameters and

$$\hat{m}^+ = \frac{1}{n}\sum_{i=1}^{n} m_{X_i+\sigma_i} \qquad \hat{m}^- = \frac{1}{n}\sum_{i=1}^{n} m_{X_i-\sigma_i} \tag{2.9}$$

represent the average of the median collapse capacities when the model random variables are perturbed, and n is related to the analysis number, equal to the number of the model random variables. The mean value of the fragility curve which accounts for both aleatory and epistemic uncertainties $\hat{m}_{\text{mod}}$ is obtained by multiplying the mean value of the curve related only to the record-to-record variability with the coefficient Δ^+/Δ^-.
Usually, incremental dynamic analysis with multi-records are used in order to properly consider the record-to-record variation, this analysis is characterised by a high computational burden.
In the next section, the most significant applications on fragility analyses of concrete gravity dams are presented.

2.2.3 Review of the existing applications

One of the first work about probabilistic analysis of concrete gravity dams was presented in Milton De Araiijo and Awruchb (1998). In this paper both the aleatory uncertainty, due to seismic excitation, and the epistemic ones, related to the materials characteristics, were considered and combined by using a Monte Carlo simulation. The seismic excitation was generated as non-stationary stochastic process with two basic random variables: the acceleration amplitude and the phase angles. In particular the acceleration amplitude was modelled as a Gaussian random variable, while the phase angles were considered independent uniform random variables. In this paper the concrete compressive strength was assumed as base random variable to generate, through deterministic relation, the tensile strength, the concrete Young's modulus and the adhesion of the dam-foundation interface. The authors calculated the safety factor against slide at the base, concrete crushing at the toe and tensile cracking at the heel, for 50 simulations. In figure 2.3 the CDF of the safety factor against cracking at the heel is shown assuming corresponding both to a Gaussian random variable and to a log-normal one. Finally, the authors used the Anderson-Darling (AD) test (Anderson and Darling 1954) to determine the most appropriate distribution, discovering that the log-normal one better fits the observations.
Tekie and Ellingwood (2003) analysed the effects of epistemic and aleatory uncertainties in the fragility assessment of the Bluestone dam in West Virginia (US). The probabilistic analysis was done through a Monte Carlo simulation with LHS sampling method in order

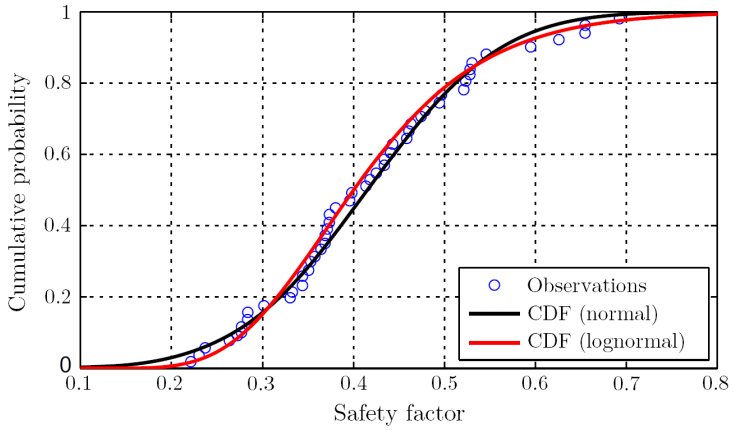


Fig. 2.3: CDF of the safety factor against cracking, adapted from Milton De Araiijo and Awruchb (1998).

to reduce the computational burden. The deterministic numerical model was composed by rigid blocks in which the shear transfer between them occurred by friction. Whereas, the dam-foundation interface was characterized by a perfectly plastic Mohr-Coulomb friction law. The free fields records were deconvolved and applied at the base of the foundation, while the fluid structure interaction was modelled by using the simplified two parameters Darbre's model. In this work, the aleatory uncertainty was the random earthquake, after a seismic hazard analysis 12 records were used. The epistemic uncertainties were:

- *Uniform distribution*: drain effectiveness, grout curtain effectiveness, tail water elevation, effective water elevation, effective uplift area, angle of friction, cohesion, dilation angle of foundation, mass and stiffness proportional to Rayleigh damping.
- *Normal distribution*: concrete compressive strength.

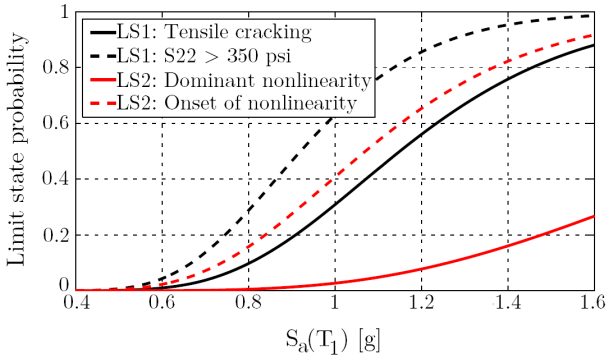
Moreover, four LSs were considered:

- LS1 Cracking of the neck: this LS was achieved when tensile stress at the neck exceeded the concrete tensile strength.
- LS2 Foundation material compressive failure at the toe: this LS was assumed to be achieved when: 1) the plastic strains around the toe were higher than 10^4 or 2) when plastic strains of 10^3 dominated the non-linear behaviour of the foundation.
- LS3 Sliding at the dam-foundation interface: the authors considered that 1) a slippage of 3 mm was the onset of sliding, 2) slippage of 13 mm affected the performance of

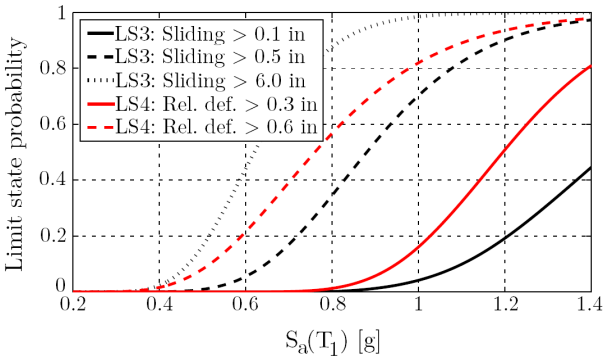
the drainage system and 3) slippage of 152 caused differential movements between adjacent monoliths.

- LS4 Deflection at the crest with respect to the heel: the authors considered two values: 8 and 15 mm which correspond respectively to 0.0014% and 0.0028% of the dam height.

Finally, in figure 2.4 the fragility curves at different LSs are shown, they were calculated as log-normally distributed and it is apparent that LS1 and LS2 are less likely to occur than LS3 and LS4.



(a) LSs 1 and 2.



(b) LSs 3 and 4.

Fig. 2.4: Seismic fragility curves with four LSs, adapted from Tekie and Ellingwood (2003).

Lin and Adams (2007) presented a set of empirical seismic fragility curves for concrete dams, located in eastern and western Canada, based on the work ATC-13 (1985). According to ATC-13 (1985) the curves were derived in terms of damage states, this was possible due to the similar design codes, construction methods and seismic condition between California

and western Canada. Therefore, the fragility curves were derived by fitting the optimal log-normal CDF to discrete points of damage states; figure 2.5 illustrates the relation between a group of fragility curves and a column of damage state.

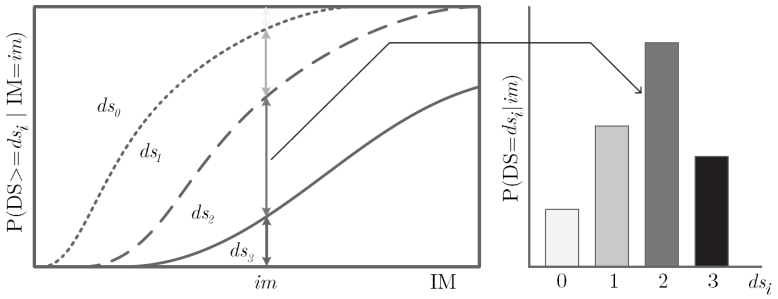


Fig. 2.5: Relationship between continuous fragility curves and DPM, adapted from (Hariri-Ardebili et al. 2014).

Finally, the previous fragility curves (figure 2.5) were modified to consider the different characteristics of seismic motion between California and Canada, just by shifting left or right the original curves. The resulting fragility curves were considered as "standard", the authors proposed to obtain fragility curves for old dams (designed/constructed in the 1900s) by dividing the PGA value by $\sqrt{2}$, and by multiplying by $\sqrt{2}$ to obtain the curves for "new" dams (designed/constructed in the 2000s).
Mirzahosseinkashani and Ghaemian (2009) presented a work where only record-to-record uncertainty was considered, while the epistemic uncertainties were neglected. The authors analysed the Pine Flat dam through a FE model, assuming an elastic massless soil, and a smeared crack model for the concrete. Two LSs have been considered:

- LS1: crack length at the base of the dam.
- LS2: total cracked area on the dam face.

Figure 2.6 shows the resulting fragility curves, the authors wrote that the fragility curves were based on log-normal CDF, while Hariri-Ardebili et al. (2014) revealed this was not possible and so two different log-normal CDFs were determined.
Lupoi and Callari (2012) presented a new probabilistic procedure for the fragility assessment of existing concrete dams which considered separately epistemic and aleatory uncertainties. By defining the external actions $\mathbf{y}$ and the structural uncertainty $\mathbf{x}$, the i-th LS is characterized by the limit state function $g_i(\mathbf{x}, \mathbf{y})$, which can be expressed as

$$g_i(\mathbf{x}, \mathbf{y}) = C_i(\mathbf{x}) - \max_t [D_i(\mathbf{x}, \mathbf{y}, t)], \tag{2.10}$$

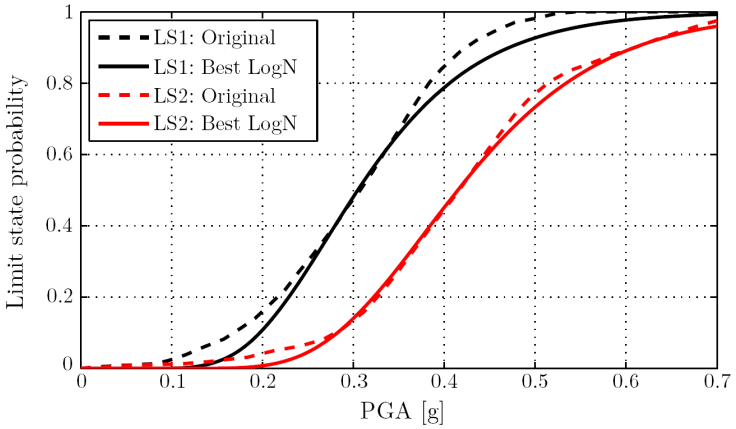


Fig. 2.6: Seismic fragility curves with two LSs, adapted from Mirzahosseinkashani and Ghaemian (2009), by Hariri-Ardebili et al. (2014).

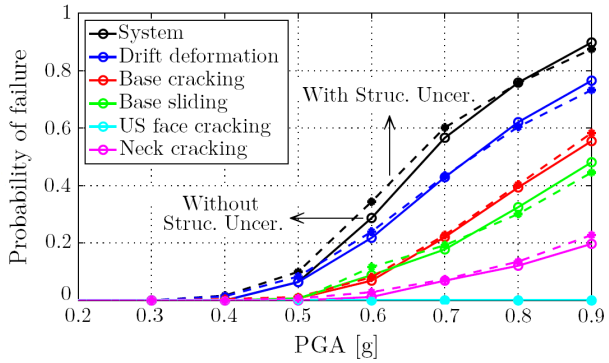
where D_i is the structural demand and C_i is the corresponding capacity. Since the random variables are separated, the computational burden can be reduced by analysing the variability of D_i using few analysis varying the seismic action and keeping constant the value of the epistemic uncertainties. This leads to an approximation, by a linear expansion, of D_i

$$D_i\left(\mathbf{x},\mathbf{y},t\right) \simeq D_i\left(\bar{\mathbf{x}},\mathbf{y},t\right) + \sum_j \left.\frac{\partial D_i\left(\mathbf{x},\mathbf{y},t\right)}{\partial x_j}\right|_{\tilde{x}} \left(x_j - \tilde{x}_j\right), \tag{2.11}$$

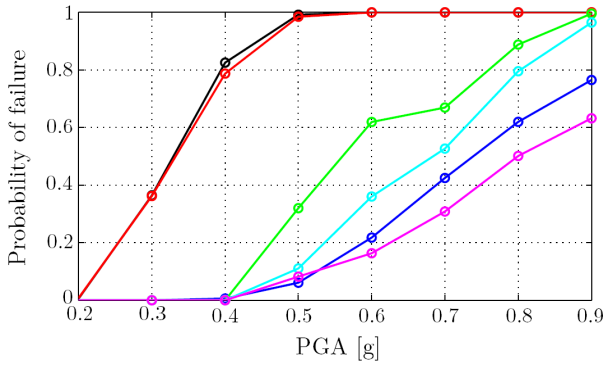
where $\left.\frac{\partial D_i(\mathbf{x},\mathbf{y},t)}{\partial x_j}\right|_{\tilde{x}}$ is the Jacobian. Regarding the capacity C_i, the authors used a deterministic semi-empirical model calibrated with expert judgements. Finally, the system reliability problem was characterized in a general cut-set formulation, and the probability of failure calculated by using a Monte Carlo simulation. The authors applied the proposed procedure to the case of the Kasho gravity dam, using an elastic FE model in which the fluid structure interaction was modelled by using displacement-based fluid elements. Only operational LSs were considered:

- Crest displacements with respect to the heel (without threshold).
- Cracking at the base of the neck, by controlling the tensile stress on selected groups of nodes.
- Base sliding governed by tangential stress.
- Cracking at the upstream face, as in the case of the neck.

Only two model parameters were treated as random variables (epistemic uncertainties): the concrete strength and the soil strength. Whereas, two aleatory uncertainties were considered: the seismic action and the water level. The resulting fragility curves are shown in figure 2.7, the authors did not use any distributional model to fit empirical data, but according to the results series components increases the probability of failure compared to the parallel arrangements. Moreover, the parallel arrangements are more sensitive to the uncertainties since the slopes of the related curves are slightly less than in the case of series components.



(a) Parallel arrangement for components.



(b) Series arrangement for components.

Fig. 2.7: Seismic fragility curves for $H_W = 36$ m, adapted from Lupoi and Callari (2012).

Yao and Jiang (2011) studied an arch dam considering only the record-to-record variation. However, 3 of the 25 vertical contraction joints were modelled with non-linear contacts, while the concrete was assumed to be linear elastic. Westergaard's added masses (Westergaard 1933) were used to model the hydrodynamic pressures, and 18 ground motions divided in 3 groups, defined with regard to the scaled PGAs, were selected. The authors

did not define the LSs and they did not derive the fragility curves.
Hebbouche et al. (2013) used the procedure proposed by Tekie and Ellingwood (2003) in order to consider both epistemic and aleatory uncertainties in the fragility calculation. The authors assumed a linear behaviour for the dam body while the soil was modelled as perfectly plastic material with Mohr-Coulomb yield criteria. The dam-soil interface was governed by Coulomb's friction law. The LHS method was used to sample from the six uncorrelated random variables:

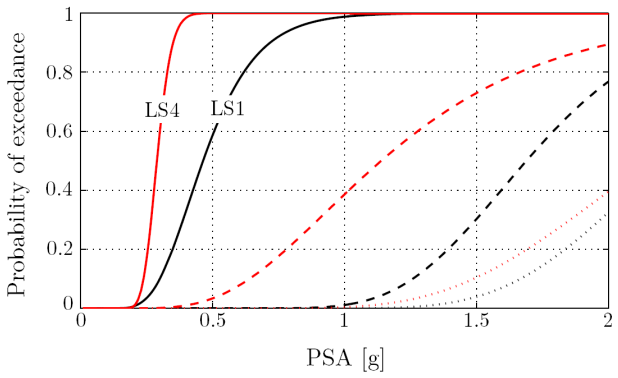
- *Uniform distribution*: friction angle, cohesion, foundation dilatation angle, elasticity modulus of concrete, and Young's modulus of soil.
- *Normal distribution*: concrete compressive strength.

Moreover,four LSs were considered, assuming three different thresholds:

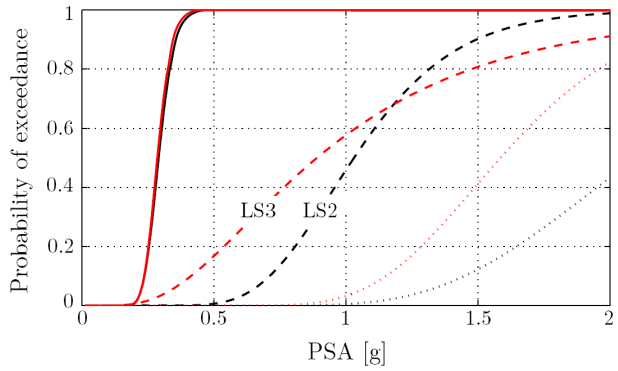
- LS1: Tensile stress at the neck with thresholds: 1.0, 1.5 and 2.0 MPa.
- LS2: Sliding at the dam-foundation interface with thresholds: 5.0, 10.0 and 20.0 mm.
- LS3: Crest relative horizontal displacement to heel, with thresholds: 5.0, 20.0 and 40.0 mm.
- LS4: Compressive stress at the dam heel, with thresholds: 1.0, 3.0 and 6.0 MPa.

Moreover, six near-fault ground motions were used and scaled with regard to the pseudo-spectral acceleration (PSA) from 0.2 to 2 g. The resulting fragility curves are shown in figure 2.8.
Ju and Jung (2015) analysed a weir overflow with steel reinforcements considering the record-to-record variability only. A 2D linear elastic model was performed and the following actions were applied: self-weight, hydrostatic and hydrodynamic pressures, uplift pressure, slit pressure and seismic action. Moreover, 30 far-field and 30 near-field ground motions records were selected and scaled to 7 PGA levels as: 0.1, 0.2, 0.4, 0.6, 0.8, 1.0 and 1.5 g. Five LSs were considered:

- LS1: Compressive stress at the weir and stilling basin, with threshold $0.25 f_c^{weir}$.
- LS2: Tensile stress at the weir body and stilling basin, with threshold $0.42\sqrt{f_c^{weir}}$.
- LS3: Compressive stress at the mass concrete block, with threshold $0.25 f_c^{block}$.
- LS4: Tensile stress at the mass concrete block, with threshold $0.42\sqrt{f_c^{block}}$.



(a) Stress-related LS.



(b) Deformation-related LS.

Fig. 2.8: Seismic fragility curves, adapted from Hebbouche et al. (2013).

- LS5: crest displacement, with threshold 10 mm.

Log-normal CDFs were fitted to empirical points in order to build the fragility curves, which are shown in figure 2.9. Figure 2.9 shows that near-field ground motions lead to higher probability of failure for most of the LSs. Moreover, comparing the fragility curves of the weir and those of the concrete block, the latter show higher probability of failure. Ghanaat and his co-authors studied different concrete gravity dams developing the relative fragility curves, providing important contributions in this area. The authors used sampling approach based on the LHS method assuming both material parameters and seismic input as random variables. In their first work (Ghanaat et al. 2011) the authors analysed the seismic performance of the Mühleberg's gravity dam, including: a weir structure, turbine building and an administration building. An accurate model was performed and non-linearity steamed in the contact surface between dam and foundation. The authors

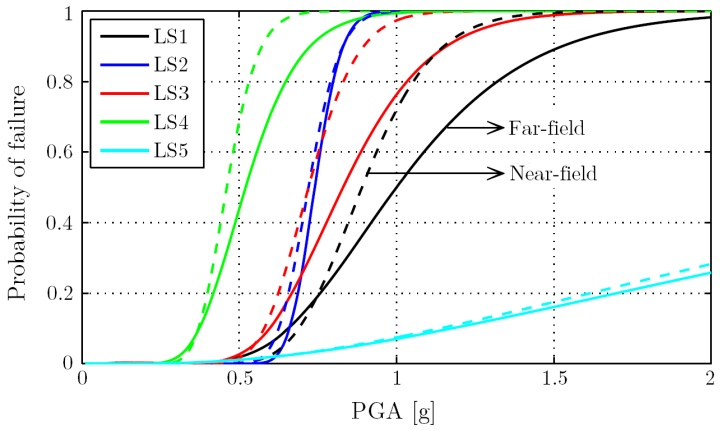


Fig. 2.9: Seismic fragility curves of the weir structure, adapted from Ju and Jung (2015).

considered the following uncertainties:

- *Log-normal distribution*: concrete elastic modulus, concrete damping, rock elastic modulus, rock elastic modulus.
- *Other distribution*: cohesion and friction angle, a logic tree analysis was performed in order to determine the most appropriated distribution.

Through a series of deterministic analyses, the authors discovered that the failure was governed by the turbine building section. Therefore, they performed seismic analyses with 30 three-components ground motion records scaled seven times: two to higher intensities and five to lower intensities. Given the high computational cost, they stopped each analysis when the first convergence failure occurred, i.e. collapse, in this way the concept of possible "resurrection" at higher intensity was discarded. In the end, a log-normal CDF was fitted through the data points using the least-square approach, thus obtaining the fragility curve shown in figure 2.10.
In Ghanaat et al. (2012) a simplified version of the previous method was presented. In this case the authors used only 10 trial analyses to combine epistemic and aleatory uncertainties instead of 30 as in the previous work. Moreover, a different scaling approach and different directional factors were used, and the fragility curve was calculated by considering a Weibull distribution instead of a log-normal one. Therefore, this approach was applied to the tallest non-overflow section of a gravity dam (geometrically similar to the Folsom dam). A 3D model of the dam was performed, again the non-linearity steamed in the dam-foundation interface and in an upper lift joint at the neck. This time the authors considered two

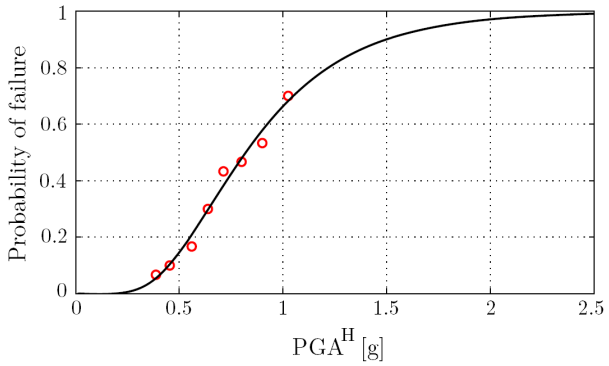


Fig. 2.10: Seismic fragility curve, adapted from Ghanaat et al. (2011).

failure modes, i.e. local and global failure modes. The elastic modulus and the damping of the dam concrete were treated as random variables as well as the rock elastic modulus. Regarding the non linear joints, the tensile strength, friction angle and cohesion were considered random variables. In figure 2.11 the resulting fragility curves are shown for the lower seismic intensities the probability of having sliding at the base joint is higher than the on at the upper lift joint, and vice versa.

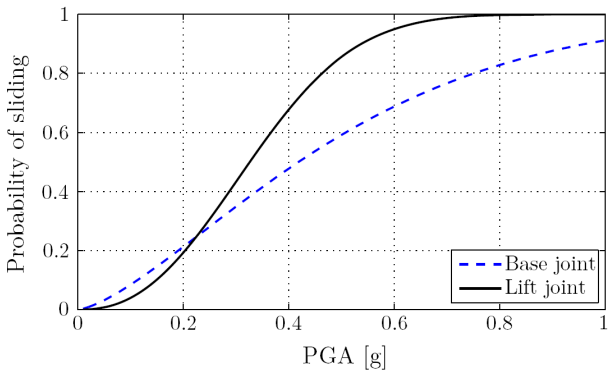


Fig. 2.11: Seismic fragility curve, adapted from Ghanaat et al. (2012).

In Ghanaat et al. (2015) the authors applied a procedure similar to the previous one, but in this case the overflow section had a completely non-linear behaviour and joints elements were applied at the base of the dam. Thirty ground motions were selected and scaled in order to achieve five LSs. Each analysis was performed twice, treating the following parameters as RVs in one case and as deterministic in the second one:

- Concrete: elastic modulus (Log-normal), compressive strength (Log-normal), tensile

strength (Log-normal), maximum aggregate size (Log-normal).

- Dam-foundation joint: rock modulus (Log-normal), tensile strength (Log-normal), cohesion (Triangular), friction angle (Triangular).
- Pier reinforcing steel: yield strength (Triangular), Strain-hardening slope (Log-normal), rapture strain (Log-normal).
- Others: Drain efficiency (Triangular), damping (Log-normal).

The obtained results were fitted through a log-normal CDF as shown in figure 2.12. The figure shows the effects of epistemic uncertainties on the fragility curve, the mean values and the variances of the distributions change significantly.

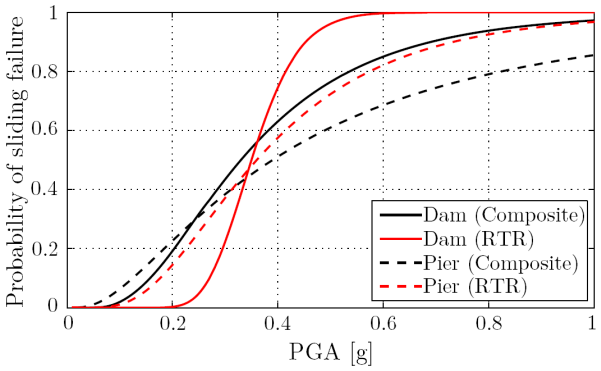


Fig. 2.12: Seismic fragility curve, adapted from Ghanaat et al. (2015).

Kadkhodayan et al. (2015) analysed an arch gravity dam by using a non-linear IDA. However, the non-linearity steamed only in the contraction joints. The percentage of over-stressed area on the dam faces $A_{OS}^{\%}$ (originally proposed by Ghanaat (2004)) was assumed as DI, the relation between DI and IM is shown in figure 2.13. This approach leads to a brittle behaviour, since this damage is not incrementally captured, e.g. for IM$<im_1$ or IM$>im_2$, so this criterion is not optimal for an IDA. Moreover, the authors used a 5^{th} order polynomial expansion for the regression analysis of the relation IM-$A_{OS}^{\%}$, instead the classical cubic polynomial spline or logarithmic model. A 3D FE model was performed, where only the contraction and peripheral joints had non-linear behaviour; a massless elastic soil was assumed. Therefore, nine three-components ground motion records were selected and three potential IMs considered: PGA, PGV and $S_a\,(T_1)$. Three LSs were defined based on the relative damage level. Figure 2.14 shows the resulting fragility curves, the use of a normal distribution leads to a non zero probability of failure for zero IM, that problem

could be alleviated by using a log-normal CDF. Moreover, the LSs defined according to the damage level lead to the intersection of two fragility curves.

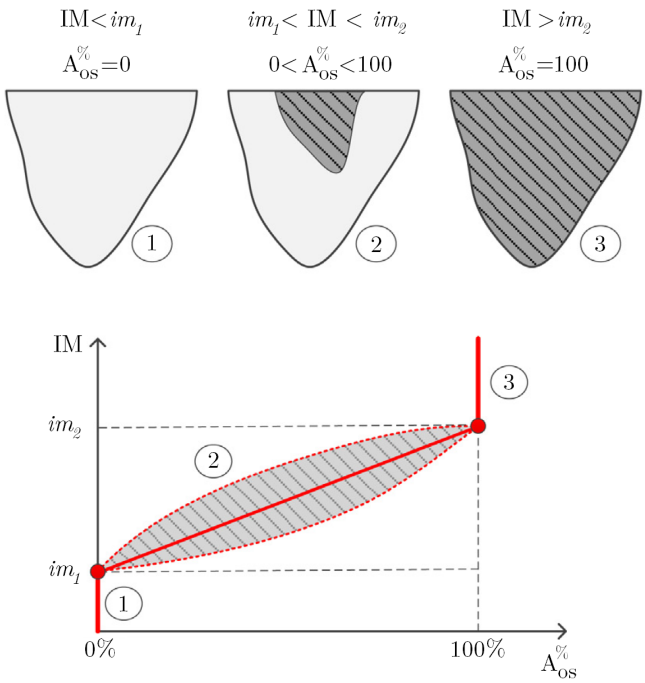
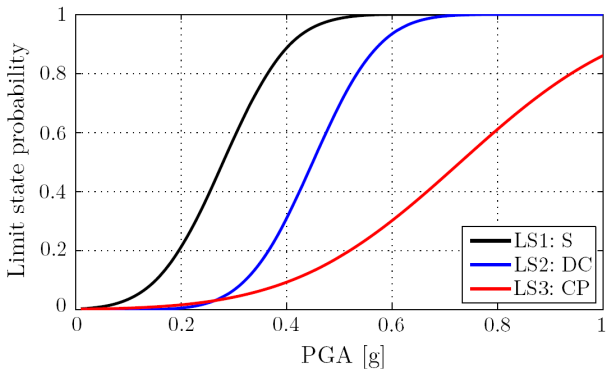


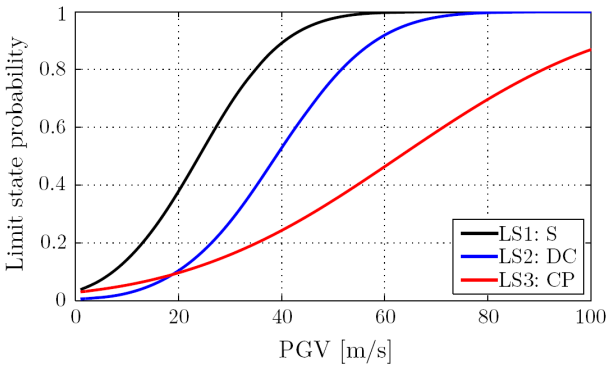
Fig. 2.13: Evolution of the overstressed area on dam face under increasing IM, adapted from (Hariri-Ardebili et al. 2014).

Hariri-Ardebili and Saouma introduced modern fragility curves and surfaces for concrete gravity dams in their publications series. In their publication Hariri-Ardebili et al. (2016b), for the first time Multiple Stripe Analysis (MSA) was used to assess the seismic fragility of concrete gravity dams. A total of nine ground motions were selected and used for three different intensity levels, thus resulting in 27 transient analyses. The authors performed a set of 27 linear elastic analyses and a set of non-linear analyses where the non-linearity was modelled rotating smeared crack and Mohr-Coulomb based joints. However, only the record-to-record variability were considered, neglecting the epistemic uncertainties. The fragility curves were derived considering different DI, in the case of linear analyses:

- Demand Capacity Ratio (DCR): the ratio between the value of the tensile stress and the tensile strength.
- Cumulative Inelastic Duration (CID): which is the total duration of the stress or



(a) PGA-based.



(b) PGV-based.

Fig. 2.14: Seismic fragility curve for dam upstream face, adapted from Kadkhodayan et al. (2015).

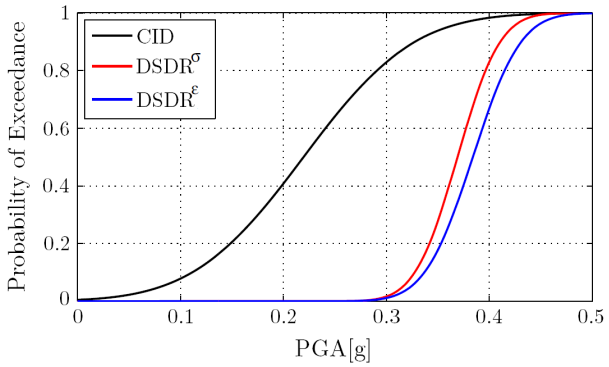
strain excursion beyond a threshold value associated with a DCR.

- Cumulative Inelastic Area (CIA): which is the integral of stress time history for a given DCR.
- Damage Spatial Distribution Ratio (DSDR): ratio of the overstressed region to total dam area for a given DCR.

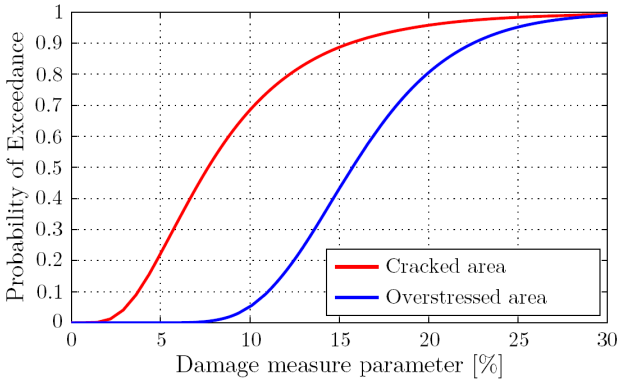
Whereas, in the case of non-linear analyses:

- Joint opening damage index: $\mathrm{DI}^{\mathrm{opening}} = \max\left(\delta^{\mathrm{opening}}\right)$.
- Joint sliding damage index: $\mathrm{DI}^{\mathrm{sliding}} = \max\left(\delta^{\mathrm{sliding}}\right)$.
- Crack-based damage index: $\mathrm{DI}^{\mathrm{cracking}} = \frac{A_c}{A_T}$, i.e the ratio of the cracked to the total area.

An arch dam was analysed, even though the authors proposed criteria for the limit of acceptability for different dams typologies. Figure 2.15a shows the fragility curves in the linear case considering three indices with thresholds CID=0.4 s and DSDR=20 %. The spread of the curves is due to the different spatial coverages since the CID is a local indicator while the DSDR is a local one. In figure 2.15b the results of the non-linear case are shown the linear and the non-linear fragility curves are nearly identical in terms of dispersion, while a shift between them can be observed.



(a) Need for nonlinear analysis.



(b) Linear vs. nonlinear correlation for SIL 3.

Fig. 2.15: Seismic fragility curve for an arch dam, adapted from Hariri-Ardebili et al. (2016b).

Hariri-Ardebili and Saouma (2016b) used for the first time the Cloud Analysis (CLA) for the seismic assessment of concrete gravity dams. The authors analysed the tallest non-overflow monolith of a 122 m high dam by performing a 2D model the only source of non-linearity was the interface between dam and foundation. Uplift pressures were automatically adjusted in terms of crack lengths. Epistemic uncertainties related to the model parameters were neglected but a large set of ground motions, around 100 records,

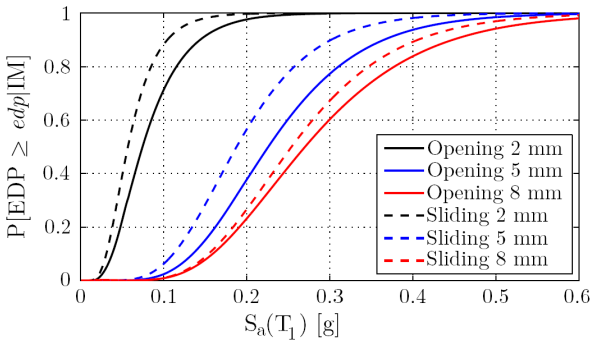
was considered. Moreover, the authors considered 70 different IMs to develop fragility curves and surfaces, in order to compare them and identify the optimal one. The optimal ones were found to be:

- Structure-dependent spectral IMs: i.e. $S_a(T_1)$, $S_v(T_1)$ and $S_d(T_1)$.
- Ground motion-dependent scalar IMs: i.e. PGA and PGV.
- Structure-independent spectral IMs: i.e. API and EPA.

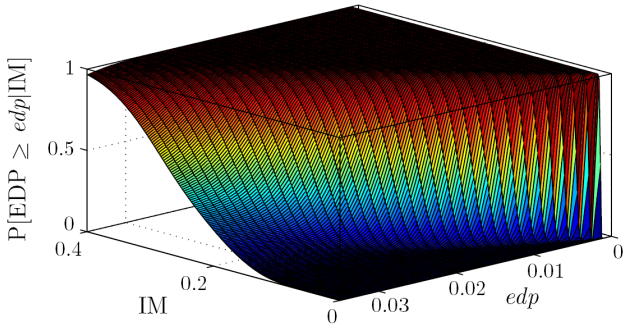
The authors highlighted that the combined acceleration response spectra including the effective mass S_a^{1-to-N} was the best IM. Figure 2.16a shows the fragility curves for joint opening and sliding at dam-foundation interface, considering three LSs: initiation of opening or sliding 2 mm, propagation of opening or sliding 5 mm and near collapse condition 8 mm. For every LS, joint sliding has the highest probability of exceedance. This results can be generalized for continuous values of EDPs obtaining the fragility surface (figure 2.16b), even though it is a challenging task (Gehl et al. 2013).
The first application of IDA to the seismic assessment of concrete gravity dams can be found in Hariri-Ardebili and Saouma (2016a). The authors performed a model similar to the one used in Hariri-Ardebili and Saouma (2016b), but in this case the concrete nonlinearity was accounted also through the use of smeared crack model. Two ground motions combinations were used, horizontal component only or horizontal and vertical components. Two different loading scenarios were assumed, empty and full reservoir. 21 ground motions, scaled by 14 SILs, were applied for each combination of loading scenarios and shaking motion assumptions. Finally, for each of the four cases, a log-normal CDF was fitted through the model outputs using different approaches:

- *Method of moments* (MOM): the CDF parameters are calculated assuming that the functional relationship and the discrete data points have same moments.
- *Sum of squared error* (SSE): the CDF parameters are determined by the minimization of the sum of the squared errors between functional relationship and discrete data points.
- *Maximum likelihood estimation* (MLE): the CDF parameters are determined by the maximization of the likelihood function.

The authors turned out that the best fitting procedure was the simplest one, i.e. MOM, according to the evaluation procedure suggested by Lilliefors (1967). The resulting fragility



(a) Based on joint opening/sliding.



(b) 3D fragility function.

Fig. 2.16: CLA-based seismic fragility curve and surface for a gravity dam, adapted from Hariri-Ardebili and Saouma (2016b).

curves, figure 2.17, shows a higher probability of failure when the vertical components of the ground motion is considered. Moreover, also in the case where the reservoir is considered full, the probability of failure increases.

Hariri-Ardebili and Saouma (2016d) derived the fragility curves for a concrete gravity dam through the Endurance Time Analysis (ETA) method. However, since only one ETA was performed, the aleatory uncertainty related to the record-to-record variation was neglected. A 2D model was developed, and the non-linearity was concentrated in an interface between dam and foundation. The random variables considered in this work were:

- *Interface joint*: tangential stiffness, normal stiffness, tensile strength, cohesion, friction angle, dilatancy angle, specific mode I fracture energy and specific mode II fracture energy.
- *Concrete and rock*: concrete modulus of elasticity and Poisson's ratio, foundation

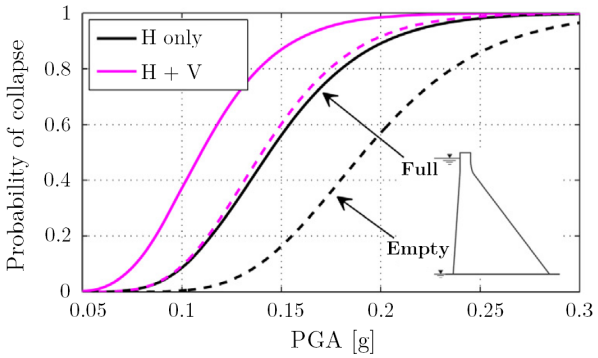
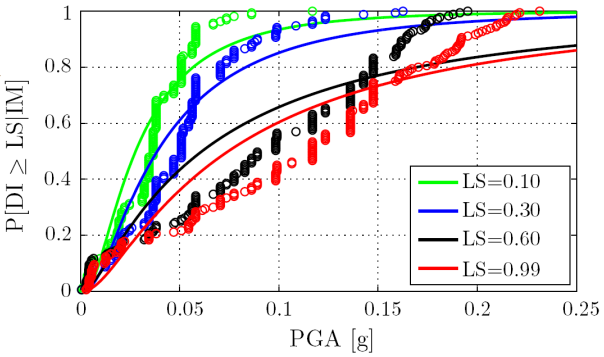


Fig. 2.17: IDA-based seismic fragility curves for a gravity dam, adapted from Hariri-Ardebili and Saouma (2016a).

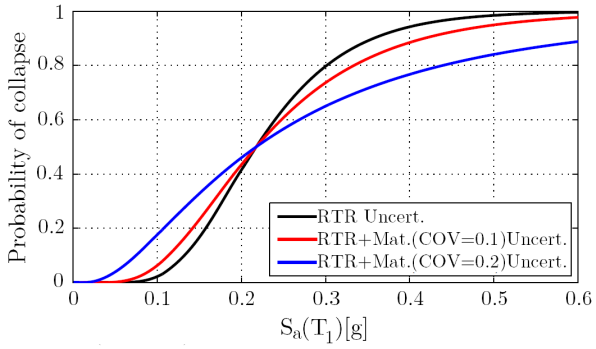
modulus of elasticity.

All RVs were modelled with a truncated normal distribution, with mean values derived from expert judgements and covariances arbitrarily set at 10-20%. Moreover, two probabilistic analyses were performed through Monte Carlo simulations, with LHS sampling, one considering correlated RVs and one without correlation. Finally, log-normal CDFs were fitted through the deterministic results by using the MLE approach in order to build the fragility curve shown in figure 2.18. Figure 2.18(a) shows the fragility curves for the correlated RVs, considering 4 LSs. In this case the DI is the ratio between crack length at the base and the total length. Finally, in figure 2.18(b) the fragility curves built considering both aleatory and epistemic uncertainties are shown, the combination was done by the root sum squared.

Another important contribution of Hariri-Ardebili and Saouma was the seismic analysis of a gravity dam treating the concrete elastic modulus as a random field (Hariri-Ardebili et al. 2018). In this work the Koyna dam was analysed and two cases were compared. In the first case the dam concrete was considered homogeneous, and the elastic parameters were treated as random variables. In the second case the dam concrete was considered heterogeneous and the elastic parameters were treated as random field. A 2D elastic model was used assuming a massless soil and the structural safety was evaluated during a post-process step, in which several performance index were considered: DCR, CID, CIA and DSDR. The authors varied the correlation length of the random field showing the effect on the model output in terms of crest and neck displacement, principal stress at the base and performance index. The results showed that the response of the heterogeneous model and the one of the homogeneous model were very similar, thus the heterogeneity had little



(a) Accounts only for MM uncertainty.



(b) Accounts for both MM and RTR uncertainties.

Fig. 2.18: ETA-based seismic fragility curves with epistemic uncertainty for a gravity dam, adapted from Hariri-Ardebili and Saouma (2016d).

influence with respect to the previous parameters. Increasing the correlation length of the random field, in the heterogeneous case, the COV of the output increased. Finally, the seismic intensity measure and the record-to-record variability were considered. In both cases, the results are highly influenced by the record-to-record variability. In the end the authors did not build the fragility curves, but this paper represent one of the first application of random field to the seismic analysis of concrete gravity dams.

Bernier et al. (2014) and Bernier et al. (2016a) analysed a concrete gravity dam using a 3D finite element model. The non-linearity of the model steamed only in the dam-foundation interface and in a lift-joint placed at the neck of the dam. Both epistemic and aleatory uncertainties were considered, in particular the ground motion was applied also in the vertical direction, by scaling the horizontal component with a random factor between 0.5 and 0.8. Moreover, the parameters of the non-linear interfaces were treated as random variables. In

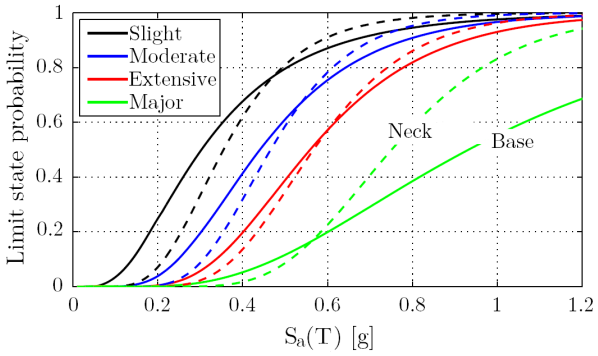
particular the cohesion, tensile strength and friction angle have been assumed uniformly distributed. Whereas, the damping ratio was modelled as log-normal random variable. Two LSs were considered:

- sliding at dam-foundation interface.
- sliding at lift-joint interface.

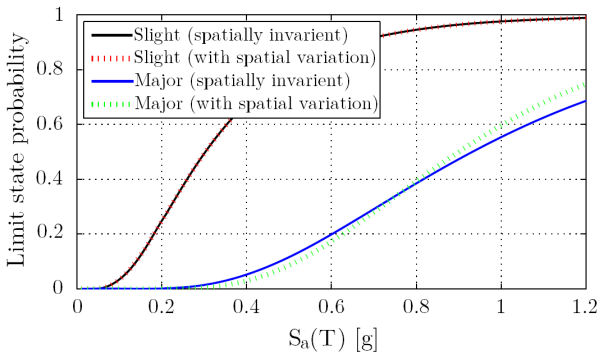
Considering the following damage levels:

- Minor or slight damage: incipient sliding.
- Moderate damage: sliding equal to 25 mm, which leads to a moderate damage of the drain system with an increase of the uplift pressure.
- Severe and extensive damage: sliding equal to 50 mm, which leads to an inefficiency of the drain system.
- Major damage: sliding equal to 150 mm at the base or to 100 mm at the neck, which leads to unacceptable differential movements between adjacent monoliths.

Once the deterministic results were calculated, the authors compared the use of three distributional models, e.g. normal, log-normal and Weibull applying two fitting methods: the SSE and the MLE. Finally, also the effect of the spatial variation of the friction angle in dam-foundation interface, was investigated. This last effect could be linked to construction phenomena or to the heterogeneous nature of the concrete. The resulting fragility curves are shown in figure 2.19. A particularly interesting thing is the effect of the friction angle spatial variation, which affects more the higher damage level than the lower ones, though its effect is very small.
In Bernier et al. (2016b), the authors repeated the previous investigation applying the MSA method and the conditional spectrum (CS) instead the uniform hazard one (UHS). All analyses were repeated using IDA, and a log-normal CDF was used to fit the discrete data points. Finally, the fragility curves for 4 LSs were constructed considering or not the effects of the epistemic uncertainties, as shown in figure 2.20. The median and dispersion parameters of the curves were significantly higher using MSA, than a lower probability of exceedance for the same value of $S_a(T_1)$. Finally, the authors shown that the effects of epistemic uncertainties become less significant when using the CS method.



(a) Spatially invarient assumption.



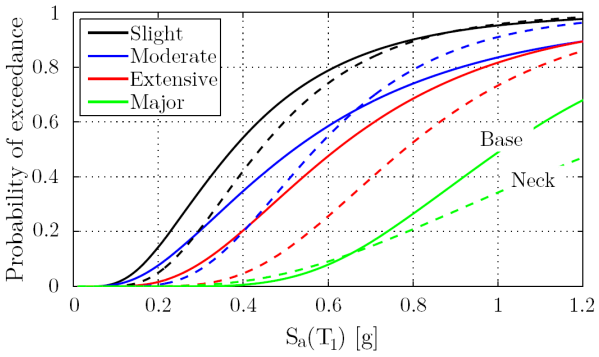
(b) Impact of spatial variation of angle of friction at dam-foundation interface with correlation length equal to the base.

Fig. 2.19: Seismic fragility curves calculated with a 3D FE model, adapted from Bernier et al. (2016a).

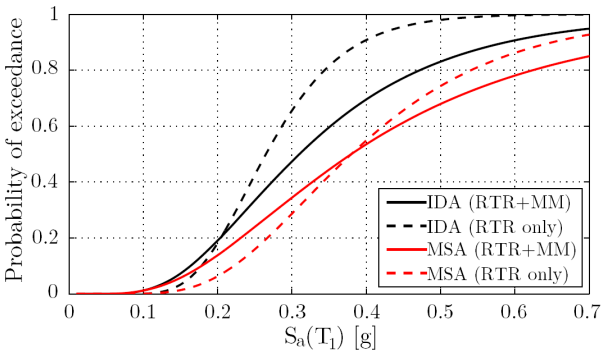
2.3 Structural Health Monitoring and damage indices for concrete dams

2.3.1 Introduction

Structural Health Monitoring (SHM) indicates the process of assessing the state of health of a structure through the analysis of measurements recorded on the structure itself. The complete development of a SHM system involves several different aspects: sensor technology, materials technology, numerical modelling and computational mechanics (Gopalakrishnan et al. 2011). This implies that both hardware components, composed by sensors and the associated instrumentation, and software components, composed by damage modelling



(a) MSA-based fragility curves.



(b) MSA vs. IDA for incipient sliding at the base.

Fig. 2.20: MSA seismic fragility curves, adapted from Bernier et al. (2016b).

and damage detection algorithms, constitute the SHM system. These two aspects strongly interact with each other, so an adequate SHM system can be defined only by considering these aspects together. In this context, the choice of the structural Quantity of Interest (QI), which must be monitored, and the definition of damage levels is fundamental. By simplifying, they respectively represent the *"structural behaviour"* and its *"threshold"*, this latter related to a particular damage state. The former should be both very sensitive to damage development (high value of information embedded in the QI), and easy to measure with a good degree of accuracy. The threshold must be defined in order to reflect a particular damage, and in this context it can be defined with respect to a Limit State (LS) or Damage Index (DI). Moreover, the definition of a threshold should consider the measurement error and the error of the prediction used to determine the structural health state.

SHM can be thought as composed by two steps, the Diagnosis and the Prognosis. On the

one hand, the diagnosis step aims to determine the state of the materials of different parts, and of the full assembly of them constituting the structure as a whole. Indeed, the diagnosis checks the presence of damage, such as cracks, their location and dimension. Moreover, the diagnosis step can be divided in two category: active diagnosis or passive diagnosis. In the passive diagnosis process, passive sensor measurements are used to assess the structural health, while the active diagnosis is based on actuator induced sensor measurements.
On the other hand, the prognosis aims to compute the effect of damage on the structure and its residual life. The probabilistic SHM frameworks proposed in this research work contribute to both the diagnosis and prognosis steps. From the diagnosis point of view, it allows detecting and localizing damage. From the prognosis side, it allows reducing the epistemic uncertainty of the structural model which influences the estimation of the residual life expectancy of the dam.
Structural Health Monitoring is an important aspect in the control process of concrete dams, this is shown by international and national guidelines (ICOLD 2000) developed in the recent years. The traditional procedure to assess the dam integrity is through visual inspection. However, visual inspections have various shortcomings as high manpower demand, insufficient frequency and inaccessibility of critical part of the structure. The consequential lack of information could lead to erroneous evaluation of the structural integrity and hence wrong decisions regarding the maintenance of the dam. For this purpose, static and dynamic monitoring systems can be used to control the structure.
In this section, a review of the most significant research in this area is presented.

2.3.2 Interpretation of the dam structural behaviour from monitoring data

The surveillance of a large concrete gravity dam during normal operations is usually based on the monitoring of displacements caused by the storage level and seasonal thermal variations (ICOLD 2000). Upstream-downstream crest movements of the dam are generally measured by a pendulum system, whether it is direct or inverted. The opening - closing of the joints is mainly measured by removable mechanical strain gauges. At the same time, boundary conditions, such as rainfall, water and air temperatures and storage level are daily gathered, while uplift pressures and water losses through seepages are generally recorded weekly.
Safety monitoring of dam has the dual purpose of highlighting anomalous operations and understanding structural behaviour through the definition of a predictive model of the structure. All the procedures for the model parameters calibration are based on the as-

sumption of linearity of the model and materials. It allows applying the principle of superposition, so that the upstream-downstream displacement δ of a point on a dam can be considered additively composed by the hydrostatic contribution δ^H, the thermal contribution δ^T and a third term δ^K which takes into account unexpected behaviours, creep and other phenomena. δ^H, which mainly depends on the mechanical characteristics of the materials, both concrete and foundation rock, is usually approximated in the literature through a polynomial expansion. Displacements δ^T, which mainly depends on the thermal characteristics of the materials, are simulated by a periodic function in time, similar to that of the average water and air temperatures. Generally, they cannot be neglected because they constitute the greatest part of δ. In this field the two most widespread methods are the observational approach and the functional approximation approach.
On the one hand, the observational approach is based on the constancy of the system response. It makes use of statistical procedures which correlate input variables such as temperature and water level values with the outputs such as crest displacement during normal operations. It aims to find the contribution of external loads to the structure deformation and to identify irreversible components in the structural response. An analytical formulation provides the upstream–downstream crest displacement as the sum of the three different terms defined above regarding temperature change, hydrostatic pressure and any term which takes into account unexpected behaviour. The model's parameters in this case are the coefficients of the three analytical functions. They can be determined through regression procedures based on the least square method, to obtain a good fit between the recorded measurements and their functional representation.
On the other hand, understanding the structural functioning requires a model based on a a-priori scheme which relates input variables and structural response. Upstream-downstream displacements reconstruction through FE models requires identification procedures in order to provide information on the mechanical characteristics of materials. The variables are physical and mechanical parameters such as density and elastic moduli of the materials. In addition, by correlating crest displacements with the opening-closing movements of vertical contraction joints, the actual relationship between adjacent monoliths can be deduced and the reliability of the selected geometrical model can be evaluated. However, displacements reconstruction is a particularly difficult task because of the large number of unknown parameters and because of the complexity of thermal analyses whose results depend on the thermal response of dam concrete. Furthermore, the effects of temperature variations, unlike those of water level variations, do not occur simultaneously with the cause. They occur in time intervals which depend on the thermal characteristics of

concrete, its porosity and saturation level.
To identify the model parameters while avoiding burdensome thermal analyses, it is necessary to separately analyse δ^H by subtracting δ^T from δ, through appropriate regression procedures. For this purpose, it is necessary to set the functions that represent the different displacement quotas.
The advantages of this methodology are implementation easiness and the small number of analyses needed. On the other hand, the result is affected by a high degree of uncertainty and subjectivity mainly due to the selection of the analytical functions describing the output variables.
Dynamic monitoring system are less frequent than the static ones, in this kind of approach accelerations, velocity or displacements are recorded. Dynamic measurements are obtained using either forced vibration tests or ambient vibration data. Therefore, these measurements are elaborated through Experimental Modal Analysis (EMA) techniques (Worden and Tomlinson 2000), or Operational Modal Analysis (OMA) techniques (Brincker and Ventura 2015), in order to determine the modal proprieties of the system, namely the natural frequencies, mode shapes and damping ratio. Forced vibration tests are usually very expensive and and difficult to implement. Whereas, ambient vibration tests are cheap and easy to implement because only the record devices installation is needed. The results of the EMA or OMA are used to calibrate the FE model parameters, in order to reach a good fitting between experimental and numerical behaviour. No applications related to concrete dams available in the literature is set up in a probabilistic framework or aims to set up a suitable dynamic SHM system.

2.3.3 Health monitoring of concrete dams

2.3.3.1 Static monitoring

Bianchi and Bremen (2001) presented procedures to investigate the long-term behaviour of concrete dams and they studied the Ferden and Roggiasca dams in Switzerland. The aim of the study was to detect any changes of the structural behaviour both before the snowfall and after the break period. The authors used environmental measurements to calibrate a deterministic tool (MIC) suitably developed for this work, based on deterministic and probabilistic procedure. Furthermore, the calibrated model was used to predict the dam behaviour in terms of crest displacements over a long period (20 years). Finally, the authors concluded that the presented deterministic tool can be used to control the structural behaviour by continuously comparing the recorded and predicted behaviour.

Behrouz (2002) in his dissertation analysed strain data collected for 17 months from Idukki dam in India. The recorded information were used to calibrate the hydrostatic-season-time (HST) model, and multiple linear regression analysis was used to detect the dependency of maximum strains with the reservoir level, time and environmental temperatures. Therefore, the calibrated model was used to predict the future strains of the dam. In the end the author concluded that HST model was able to accurately predict the future structural behaviour.

Pytharouli and Stiros (2005) analysed horizontal and vertical displacements, extrapolated by geodetic monitoring data, for long term behaviour (30 years) of the Landom dam in Greece. The authors used Discrete Fourier transform and Lomb normalized period-gram methods to understand if both dam deformation and reservoir level variation corresponded to periodic functions of the same period. They calibrated a predictive model based on the previous data, concluding that, for a period greater than 30 years, the structural integrity and deformations were kept at low levels.

Chouinard et al. (2006) applied a multivariate statistical analysis (HST) to the study of Daniel Johnson dam, in order to predict the long term (9 years) behaviour of the structure in terms of irreversible displacements. Moreover, they investigated the dependency between displacements and the environmental measurements through the principal components analysis. The results showed that displacements are mainly due to: thermal effects, water level variation and irreversible phenomena.

De Sortis and Paoliani (2007) compared statistical and structural identification techniques in dam monitoring. The authors analysed the Ancipa dam considering air mean daily temperature, water level variation and crest displacements spanning over 40 years. The authors performed a structural identification in order to calculate the relationship between external loads and dam displacement by using numerical model. The results showed a better accuracy of the structural identification procedure, than the statistical method, in the approximation of the thermal part of the total displacement. Therefore, they concluded that structural identification procedure can successfully used to predict and control the dam behaviour in a long-term period.

Perner and Obernhuber (2010) developed a hybrid model, which combined structural analyses and statistical models, to predict the deformations of the Zillergrund arch dam. Registrations of the deformations recorded in a period of 19 years were used to train the model, and structural models were used to identify the relationship between displacements and environmental measurements, as water level and temperature. The authors concluded that the hybrid model was better than statistical approach since it could be trained with a small

number of observations.
Mata (2011) compared the use of two statistical models, namely multiple linear regression (MLR) and neural network (NN) models to predict the structural behaviour, in terms of crest displacement, of the Alto Robago dam subjected to environmental loads. The two models were calibrated using a 25-year crest displacement, temperature and basin level. The results showed how NN is more flexible and capable to simulate local extreme behaviour due to particular combinations of environmental loads.
Loh et al. (2011) developed a suitable statistical approach to extract trends from long-term structural health monitoring data in order to set an early warning threshold level. The authors analysed the Fei-Tsui dam for which displacements and temperatures were available within a period of 22 years. Several different statistical methods were used and compared, in particular: the singular spectrum analysis with auto regressive model (SSA-AR) and the non-linear principal component analysis (NPCA) using auto-associative neural network method (AANN). The results showed that AANN was able to capture periodic variations (temperature, season) as well as the trend which they supposed to be related only to creep phenomena. Therefore, the authors concluded that AANN is better than SSA-AR model in the long-term prediction of the structural behaviour.
Popescu (2011) analysed the deformations and environmental conditions of the Vidrau dam, with the aim to find how they effect the dam behaviour. The authors also aimed to separate the deformation related to the environmental conditions and the time effects. In order to do that, the authors applied a second order blind source separation (BSS) algorithm to dam monitoring and surveillance data. The results showed that BSS successfully recognized the contribution of each phenomena to the total dam deformation.
Henriques et al. (2012) analysed the Cabril dam, recording 2 days data of temperature, reservoir level and inclination. These data were elaborated with regression in order to understand the relationship between the different effects, and in fact the results showed a correlation between inclination, temperature and basin level variation. Finally, the authors noted that some variations in the dam behaviour cannot be explained by the regression model used in this work.
Demirkaya and Balcilar (2012) compared the use of two statistical models, namely multiple linear regression (MLR) and multiple layer perception (MLP) models, to build a predictive model of the displacements of the Schlegeis arch dam. The two models were trained using water level, air temperature, concrete temperature recorded in 6 points on the dam and displacements, acquired during 8 years. According to the results, MLR showed better performance with regard to the criteria of R^2, than MLP. Finally, results indicated that linear

regression provided the most appropriate solution in of linear problems and acceptable in static dam monitoring.

Mata et al. (2013) analysed the Alto Lindoso dam identifying the effect of the daily variation of air temperature on the structural deformation. The Short Time Fourier Transform (STFT) was used to recognize the influence of air temperature variations on the horizontal displacements, and tracking its variation as function of time which can be used to forecast the future behaviour of the dam. The amplitude of the horizontal displacement was modelled as a function of the air temperature amplitude by a linear regression model. Moreover, the parameters of the statistical model were identified by using the observations recorded from October 2008 to November 2011. The results showed that a better correlation between amplitudes of air temperature and the one of horizontal displacement can be reached when a phase of displacements is regular.

Kao and Loh (2013) used neural networks to build a predictive model of the static behaviour of the dam, which coefficients were trained with the observations acquired by the static monitoring system. The authors studied the Fei-Tsui arch dam comparing three different kinds of Artificial Neural Networks (ANNs) with increasing non-linearities. The authors showed that all of the three ANNs were able to accurately follow the observed behaviour.

Mata et al. (2014) proposed multiple linear regression models for arch dam displacements, whose coefficients were calibrated by regression on the recorded observations. In particular, the authors analysed the use of different models to approximate the thermal part of the displacement, one based on sinusoidal functions and two based on the real recorded temperatures. The analysis of the Alto Lindoso dam proved that the procedure proposed in that work allowed accurately reproducing the static behaviour of the dam.

Su et al. (2016) proposed a Support Vector Machine-based (SVM) model to reproduce the static behaviour of the dam. The authors analysed the most influential parameters to be inserted in the SVM model in order to obtain a good fitting between recorded and predicted dam behaviour. The authors proposed also a procedure for real-time calibration of the SVM model itself.

Kang et al. (2017) proposed a Extreme Learning Machine-based (ELM) predictive model of dam displacements for structural health monitoring purposes. The authors analysed the performance of the proposed model by comparing it with the Multiple Linear Regression model and the Back Propagation neural network approach. The study of the the Fengman dam showed the good agreement between records and predictions calculated with the ELM-based model.

Salazar et al. (2017) presented a predictive model of the dam displacements based on boosted regression trees for damage detection purposes. The authors compared both Causal and Non-Causal models in the study of the Baells dam proving that the second was more efficient in the damage detection. Indeed, they simulated damage of the dam through numerical models, highlighting that the proposed model could be successfully used for real-time control purposes.
Shao et al. (2018) proposed a dam displacement model based on panel data in order to solve the issue of linear relationship between influence factors, which is a feature of statistic models commonly used in dam engineering field. The authors studied a hydro-power station located on Yalong river, showing the good performance of the proposed approach.
Prakash et al. (2018) proposed a framework to build predictive models of the dam static behaviour, e.g. displacements, strains. The use of principal component analysis allowed reducing the model dimension. The coefficients of the predictive models were calibrated by using least squares method, while the threshold of abnormal behaviour was continuously calibrated with the new observations in order to avoid erroneous warnings. The study of an arch dam located in Bulgaria showed the applicability of the proposed procedure.
Su et al. (2018) proposed the combination of Dempster-Shafer Theory of evidence (DST) and Set Pair Theory (SPT) in order to build an approach for health control and decision making in dam engineering field. The proposed method allowed merging multi-source space-time information coming from different parts, e.g. monitoring system and inspection. The study of an hydro-power station built in 1953 proved the effective of the method both to track the structural behaviour and to help decision maker to find the best retrofitting intervention.
Wei et al. (2018a) proposed a predictive model for static dam displacements which considered also lag effects, e.g. reservoir level and rainfall lag effects, and ageing phenomenon of RCC. Firstly, the lag parameters were calculated by using a genetic algorithm. Secondly, a regression analysis was performed in order to calculate the combination parameters. Finally, multilevel-recursive method and regression were applied in order to find time varying parameters. The study of a RCC gravity dam showed that the proposed method leaded to a fitting with observed data better than the one obtained with the statistical method.
Hu et al. (2018b) studied the leakage problem for the Shimantan Reservoir dam, a concrete dam which showed penetrating cracks. The authors showed that systematic field inspections were fundamental in order to understand the causes of leakage problems. Moreover, they proposed a predictive model of dam displacements based on the modification of the Navier-Stokes equation, which was suitable for dam with leakage problems and penetrating

cracks.

Hu et al. (2018a) proposed a method to identify anomalies related to foundation uplift. The elaboration of time series of reservoir levels and uplift pressures with the Dynamic Time Warping (DTW) method and the Local Outlier Factor (LOF) allowed recognising similarities on dependent and independent variables, and to find anomalous causes. The study of the Xixi dam proved the effective of the method.

Dai et al. (2018) studied a RCC gravity dam by applying the statistical model of the displacement and Random Forest Regression (RFR) in order to determine its combination coefficients. The research work showed that the proposed approach reduced the bias of the prediction.

Wei et al. (2018b) proposed a hybrid predictive model of the dam displacements. The statistical model commonly used in dam engineering field to replicate displacements was modified in its hydrostatic component by using FEA. The authors also considered the chaotic residual errors in order to improve the fitting between observations and predictions. The study of a RCC dam showed that the proposed approach leaded to a better fitting than the other methods commonly used in dam engineering field.

Lin et al. (2019) proposed a method to separate dam displacements into two parts: one related to the dam behaviour and one related to the foundation. The idea of partitioned FEM was used to define hybrid equations which allowed separating the two contributions. The observations recorded by the static SHM were used within the procedure in order to obtain the mechanical parameters of the materials. The study of a concrete dam was used to verify the validity of the separation method.

In conclusion, the research works related to static SHM system of concrete dams available in the literature are characterised by the use of different kind of numerical models and statistical or numerical tools. All methods are calibrated using environmental and observations recorded during the regular use of the dam, in order to perform predictive models of the future behaviour. The largest part of the research works aim to improve the fitting between predictions and observations by adding new parameters with no-physical meaning to the predictive models. Usually, the stability of the models, the uncertainties which affect the solution, and error are not discussed, thus making difficult the evaluation of the goodness of the final result.

Focusing the attention only on the fitting between observations and predictions, adding new components to the predictive model could lead to a paradox. Indeed, also the stability of the solution must be checked. A too complex model, with regard to the amount and typology of information, could lead to an *over-fitting* between predictions and observa-

tions. In this way, the predictive model would show good performance if compared with the observations used for training it but a big error otherwise. This concept is also known as Bayesian Ockham Razor (Beck 2010).

2.3.3.2 Dynamic monitoring

The dynamic monitoring of concrete dams can be based on forced vibration tests or ambient vibration. One of the first studies of forced vibration tests on concrete dams can be found in Severn et al. (1980), who studied the Wimbleball buttress dam with the aim to quantify the effects of the water level variation on the structural behaviour. Rotating eccentric mass placed in the upper part of the dam was used to excite the structure. Resonant frequencies in the range of 8.9-9.6 Hz were estimated from resonant peaks. The authors found that an increasing water level caused a fall in resonant frequencies, and they also found that buttress did not act independently from each other for small amplitude of the motion.
Clough et al. (1986) used the results of forced vibration test campaign, executed on the Xiang Hong Dian dam, to calibrate the parameters of a FE model of the structure. The vibration was induced by an eccentric mass shaker, and accelerometers were placed on both the dam crest and the foundation on the downstream face of the dam. 12 fundamental modes, in the range 4-12 Hz, were identified from Frequency Response curve calculated with the test results. Finally, the authors concluded that experimental measurements well fitted analytical predictions.
Cantieni et al. (2004) studied the Norsjo dam, they created and updated a FE model of the structure using the results of forced vibration tests. A servo hydraulic shaker was used as exciter, and the structural behaviour in terms of accelerations was measured in three directions at the dam crest. The least square complex exponential (LSCE) algorithm was used to determine the resonant frequencies in the rage 3.55-12.9 Hz.
Due to economical and practical reasons, usually the execution of forced vibration tests on concrete dams is difficult. Therefore, ambient vibration tests which not require an exciter become a powerful tool to identify the modal properties of the system. One of the first application of ambient vibration test on concrete gravity dam can be found in Brownjohn et al. (1986). The aim of the work was to investigate the feasibility of the installation of a dynamic monitoring system. The authors studied the Contra dam placing accelerometers on the upper part of the structure. Natural frequencies were determined by peaking the peaks on the auto-spectrum of the signals, while the mode shapes were extracted from the output only transfer functions. They identified 8 eight upstream-downstream modes in the

frequency range 1.8-4.2 Hz. Finally, the results showed a direct relationship between the ambient vibration and the electricity generation activity.
Brownjohn (1990) studied the safety and stability of Hermitage Dam, a concrete gravity dam located in Jamaica. Ambient vibrations were used to determine natural frequencies in the range 7-30 Hz, by picking the peaks on the auto-spectrum. The results were used to calibrate the parameter of a FE model of the structure, needed to evaluate the structural safety. The results showed discrepancy between tests field data and mathematical model prediction, but the authors concluded that these discrepancies were too small to invalidate the numerical model.
Loh and Wu (1996) analysed the Fei-Tsui dam, obtaining the dynamic properties of the system by using both ambient vibration and seismic response data. In particular, random decrement method and least-squares method were used to estimate modal parameters of the dam. The results showed a good agreement between the parameters obtained by using ambient vibration and those from seismic excitation.
Kemp (1996) evaluated the structural safety of the Ruskin dam, a concrete gravity dam located in Canada. The dynamic properties of the numerical model were calibrated on the basis of the experimental results obtained by ambient vibration test. Natural frequencies, in the range 6.5-14 Hz for the maximum reservoir level, and 8.5-14.5 Hz for lower level, were obtained from the peaks of average normalized power spectral density.
Daniell and Taylor (1999) used the results of ambient vibration test in order to calibrate the numerical model parameters of the structure. The registrations were acquired by using accelerometers placed in the dam crest, natural frequencies were picked from the peaks of average normalized power spectral densities. The authors were able to identify 6 modes in a frequency range 6.1-11.8 Hz. Finally, numerical and experimental results showed a good agreement, demonstrating that ambient vibration can successfully substitute forced vibration tests.
Darbre et al. (2000) studied the Mauvoisin dam, an arch dam located in Switzerland. The aim of the work was the analysis of the relationship between basin level variation and shift of the natural frequencies. Seven ambient vibration tests between 1995 and 1996 were carried out, and an automated system, which recorded twice daily for six month, was set up. Natural frequencies were obtained by peaking the peaks on normalised power spectral densities of individual accelerations. The results showed that the natural frequencies initially increased with rising the water level. However, after a particular level the natural frequencies started decreasing, this effect was due to two competing phenomena: the increasing mass water which participate to the motion (reduction of the natural frequencies)

and dam stiffening related to closing of vertical construction joints.
Mivehchi et al. (2003) analysed two concrete arch dams in Iran: the Shahid-Rajaee and Saveh dams. Ambient vibration tests were conducted from winter of 1999 to autumn 2000, with the aim to validate the results of numerical models. Partial and rapid opening and closing of the bottom outlet gates of the dam body provided an artificial external excitation during the tests, in order to excite also weak modes. The modal parameters of both dams were extracted by the auto-spectrum of the signals obtaining natural frequencies between 1.46-3.58 Hz for Shahid-Rajaee dam, and 3.91-7.91 Hz for the Saveh dam.
A continuous dynamic monitoring system was installed on the Cabril arch dam after the ambient vibration tests carried out between 200 and 2003 (Oliveira et al. 2004). The results of these tests were used form different purpose:

- Demonstrate that ambient vibration can be used to characterize the dynamic behaviour of arch dams with good precision (Mendes et al. 2004).
- Investigate the influence of contraction joints in the dynamic behaviour of arch dams (Mendes et al. 2007).
- Study the influence of an intake tower on the modal characterization of a dam (Mendes and Oliveira 2009).

The modal characteristics of the system were extracted by using Frequency Domain Decomposition (FDD) and Stochastic Subspace Identification (SSI) techniques. The authors showed how a continuous ambient monitoring system can lead to a good calibration of numerical model of the dam.
Okuma et al. (2008) used the data of long term ambient vibration monitoring system for damage detection purpose. The authors analysed the Hitotsuse dam, the measuring system was set up in order to continuously record at a sampling rate of 200 Hz and to store measurements every 30 minutes. The natural frequencies were determined from the cross spectrum of autoregressive moving average models. The first results showed a good agreement between natural frequencies, estimated with ambient vibration, and those obtained with earthquake observations. Moreover, a strong correlation between the first three modes and the water level was found.
Moyo and Oosthuizen (2010) analysed two arch dams, Roode Elsberg and Kouga dams, performing ambient vibration survey trials, in order to determine the modal characteristics of the two systems, and to use them as baseline measurements for long term dam safety monitoring. Four tests were carried out on Roode Elsberg dam from December 2008 to

April 2010, while only one test was done on the Kouga dam in September 2009. Several operational modal analysis techniques were applied for comparison reasons, as reported in Bukenya et al. (2012b) and Bukenya et al. (2012a). The natural frequencies of Roode Elsberg and Kouga dams were in the range 3-7.70 Hz and 3.72-8.30 Hz, respectively. Regarding the OMA techniques comparison, the results showed that none of them was better than the others.

Ellis et al. (2010) performed ambient vibration tests on the Gem Lake dam, in order to identify the modal properties of the system. Spectral analysis technique, which included the Fast Fourier Transformation and the maximum entropy method, coupled with waterfall plot analyses, were used to elaborate the measurements. Natural frequencies in the range 13.18-27.71 Hz were found. Therefore, a calibrated FE model of the structure was developed in order to deeply understand the dynamic dam behaviour and to develop a suitable field test procedure.

Sevim et al. (2010) analysed the Berke arch dam developing ambient vibration tests, whose results were used to calibrate a numerical model of the structure. Therefore, ambient vibration tests were performed, during which, wind and water pressures were the most important sources of excitation. Eight natural frequencies between 2.74 to 9.66 Hz were found by using enhanced frequency decomposition (EFDD). Finally, the parameters of a FE model of the dam were calibrated in order to fit the experimental results.

Cheng et al. (2015) proposed a health monitoring method for concrete dams based on ambient vibrations in which the Kernel Principle Analysis (KPA) was used to remove the effect of changes in the environmental conditions. In the proposed method, damage was detected by comparing the L_2 norm of the error of the KPCA calculated during normal condition, with the one calculated for the current state of the dam. After the detection, the COMAC indicator was used to localize damage. The authors did not tackle the problem of the mode coherence and they did not define a probabilistic framework.

Hariri-Ardebili et al. (2019) studied the modal behaviour of an arch dam, showing the influence of the uncertainties of the mechanical parameters on the results. The authors proposed a calibration procedure which allowed determining the model parameters once experimental modal characteristics were gathered. However, they did not mention the use of such procedure within a SHM framework.

In the last decade the use of ambient vibration tests on concrete dams has gained great interest, with two main purposes: the calibration of FE models and structural control. These two aspects are strictly related since the structural control, or damage detection, needs reliable numerical model of the dam. Usually model parameters are calibrated by

using minimization techniques in a deterministic setting, neglecting the estimation of the model error, as well as in the static monitoring case. The coherence between recorded and predicted modes is an important aspect of SHM systems based on ambient vibrations, for both model updating and structural control. Due to uncertainties in the predictions or in the records, or to a change of the structural behaviour, experimental and numerical modes can have different order. If they are not reordered coherently, different modes are compared leading to false warning or wrong results in the updating process. Usually, this problem is solved by introducing the concept of *System Modes* as additional variables of the problems (Beck et al. 2001). However, in dam engineering field this is not sufficient. Indeed, the Soil-Structure Interaction (SSI) must be considered in order to perform reliable numerical models. The presence of the SSI leads to a large number of numerical modes with no experimental correlation, thus giving even more importance to the coherence problem. The largest part of the contributions available in the literature which consider this problem did not ensure the mode coherence but they check it a posteriori through the use of synthetic indicator, such as MAC MACEC and so on.

2.3.4 Damage Indices

2.3.4.1 Introduction

The definition of damage indices (DIs) is a fundamental aspect of the structural health monitoring and of the seismic assessment process. A damage index and its threshold express in a synthetic way the damage state of the structural components or of the whole building allowing the structural control. Usually DIs are bounded between 0 and 1, where 0 indicates a healthy state, while 1 indicates a completely damaged structure. In the field of concrete gravity dams several different damage indices have been defined in literature. However, since the main collapse mechanisms of such structures are sliding and rocking along a lift joint, all proposed damage indices are similar. In this section some consideration about damage indices for concrete gravity dams are reported.
Following Hariri-Ardebili and Saouma (2015), a damage index can be defined as a function of damage variables (DV). Any structural response which can be used as indicator of damage is defined as damage variable. If the multiplication is used to express the relationship among damage variables, by defining the the upper bound DV_u^i and the lower one DV_0^i (threshold) of the i-th damage variable DV^i, the damage index can be written as

$$DI = \prod_i \left(\frac{DV^i - DV_0^i}{DV^i - DV_u^i} \right)^{\kappa_i}, \tag{2.12}$$

where κ_i is an exponent related to the change in DI at different stages of DV^i.
As mentioned before, a damage index can be classified in different way:

- *Local or Global*: a local DI is an indicator of damage in a member, portion or limited part of a structure, while a global DI capture the damage state of the whole structure. Global DI can be defined as a weighted summation of local DIs.
- *Single-variable or Multi-variable*: a DI can be defined as a function of only one variable or as a combination of more than one variables.
- *Cumulative or Non cumulative*: a cumulative DI is able to track the accumulation of damage during the time, while a non-cumulative DI expresses the damage state in particular time.
- *Deterministic or Stochastic*: the material variability (epistemic uncertainty) within the dam body can be considered by using a stochastic DI, or neglected by defining a deterministic DI.
- *Structural or Economical*: damage index can be defined in terms of structural parameters or in terms of economical ones.

The definition of a damage index starts from the study of potential failure modes of the structure under control, in the case of concrete gravity dams three potential failure modes are usually considered:

- **Over-stressing**: usually both tensile and compressive stresses are checked.
- **Sliding**: sliding along cracks inside the dam body or planes of weakness as dam-foundation interface or lift joints.
- **Overturning**: this phenomena usually appears with the sliding, the rotation pole lies on the sliding plane.

The potential dam failure is due to a combination of the previous phenomena, but their classification is important for the definition of damage indices. Under severe ground motions a typical gravity dam may suffer tensile cracks, usually located at the base and near the upper downstream face discontinuity (Zhang et al. 2013). Usually, the uppers cracks initiate from the upstream or downstream face and propagate with an angle toward the opposite side of the dam. The main consequence of cracking, if they reach a critical length, is the formation of a weak surface which can lead to sliding or rotational instability. Moreover, the cracking could lead to an increase in uplift pressure, accelerating the instability

phenomena. When cracks due a ground motion start from the upstream surface of the dam they are usually concentrated at the heel. Whereas when the cracks start from the other direction they are usually localised around the slope discontinuity and near to the toe of the dam. Figure 2.21 shows failure mechanisms of concrete gravity dams under increasing ground motions (Léger 2007). Figure 2.21 shows some critical location of cracks:

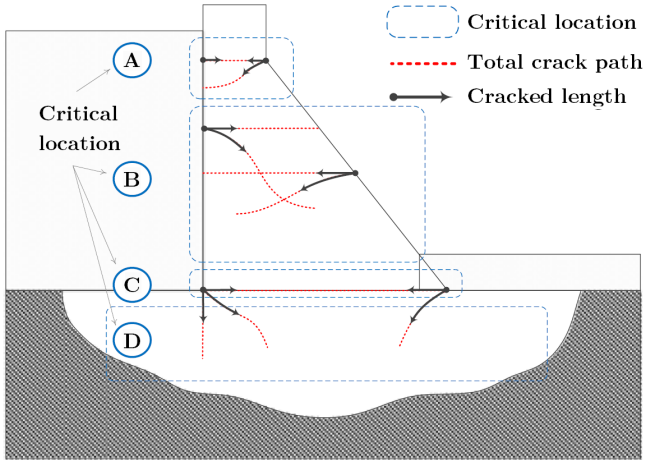


Fig. 2.21: Potential failure modes of a typical concrete gravity dam, critical location and crack paths, adapted from Léger (2007).

- **A**: neck area at the change of the downstream face slope.
- **B**: along lift joints at various elevations.
- **C**: along dam-foundation interface at the toe and heel of the dam.
- **D**: horizontal, vertical or inclined cracking at the foundation in near-field of the dam.

Several applications of DI, defined for other kind of structures and applied to the study of concrete dams, are available in literature. The most commonly used are: cumulative displacement-based indices ("Seismic safety of reinforced concrete members and structures"), cumulative force-based indices ("Reinforced concrete hysteresis model based on the damage concept"), cumulative hysteretic energy-based indices (Cosenza and Manfredi 2000) or a combination of them.

2.3.4.2 Specific damage indices for concrete gravity dams

Interesting examples of DIs specifically defined for concrete dams are those of Hariri-Ardebili and Saouma (2015) and Ansari and Agarwal (2016). In the former case the authors proposed a cumulative multi-variable DI, which handled local, intermediate and global damage states. The damage index was defined in terms of crack length L^c, dissipated energy E_H and maximum drift u_{max}. The dam failure was assumed to be reached when a thorough crack occurs. Therefore, a DI for each of the critical areas shown in figure 2.21 was defined as weighted ratio of cracked length over the estimated local path,

$$DI_i^j = \beta_\Delta \times \frac{L_i^c}{L_i^T}. \tag{2.13}$$

In the previous equation DI_i^j is the damage index related to the i-th crack path in the j-th critical area, L^c and L^T are the cracked and total lengths, respectively, while β_Δ is the controlling coefficient based on u_{max}. Furthermore, ψ_i^j the ratio of dissipated energy along a crack path $(E_H)_i$ with respect to total energy E_H,

$$\psi_i^j = \frac{(E_H)_i}{E_H}, \tag{2.14}$$

is introduced and combined with DI_i^j. Finally, considering n crack paths, the area damage index

$$\bar{DI}^j = \sum_{i=1}^{n} DI_i^j \times \psi_i^j \tag{2.15}$$

can be defined, while the DI for the whole structure is

$$\bar{\bar{DI}} = \sum_{j=1}^{m} \bar{DI}^j \tag{2.16}$$

where m is the total number of critical areas. Furthermore, the authors studied the Koyna dam using 2D and 3D models of the non-overflow section, and considering plastic material models for both soil and dam concrete. They studied the evolution of the proposed DI showing that it can be successfully used to assess the damage state of concrete gravity dams. Special attention must be placed on the definition of L^T, if the crack pre-exists, so the length of L^T is deterministic since it can be measured. Whereas, if the section is partially cracked, L^T can be estimated on the basis of figure 2.21 accounting for the fact that a crack will propagate with a minimum energy dissipation.

Ansari and Agarwal (2016) proposed two global damage indices, one as a combination of crest displacement and cumulative damage energy dissipation DI_d, and another based on factor of safety against sliding DI_{FSS}. The former was defined as

$$DI_d = \begin{cases} 0, & \text{if } d_i \leq d_{y0} \\ \frac{d_{yi}E_{yi} - d_{y0}E_{y0}}{d_{yf}E_{yf} - d_{y0}E_{y0}}, & \text{if } d_{y0} < d_i \leq d_{yf} \\ 1, & \text{if } d_i \geq d_{yf}, \end{cases} \tag{2.17}$$

where d_{y0} is maximum crest displacement at initial yield, d_{yf} is maximum crest displacement at final yield, E_{y0} is cumulative energy dissipated up to initial yield point, E_{yf} is cumulative energy dissipated up to final yield point, d_{yi} is maximum crest displacement during any cycle between initial and final yield points, and E_{yi} is cumulative energy dissipation corresponding to the cycle. The damage index related to the sliding DI_{FSS} was based upon the residual factor of safety against sliding due to cracks occurred during a seismic event. It was defined as

$$DI_{FSS} = \frac{\left[\left(\sum_{i=1}^{n} \left(DI_{FSS_i}\right)_{\text{base/body}}\right) + 0.5\left(DI_{FSS_i}\right)_{\text{neck}}\right]}{n+1}, \tag{2.18}$$

where $DI_{FSS_i} = \sqrt{\left(1 - \frac{FSS_i}{FSS_f}\right)}$ is the partial damage index for any specific crack path, FSS_i is the safety factor against sliding at any intermediate crack, while FSS_f is total safety factor against sliding at no damage state. Finally, the authors analysed a concrete gravity dam by using a 2D model, calculating the fragility curves of the structures for each damage state.

2.4 Literature gap

This Chapter presents the main research works on the seismic assessment of existing concrete gravity dams and the application of SHM in dam engineering. Most of the contributions related to the seismic assessment focus the attention on the application of methodologies developed for other kinds of structures, and the evaluation of the differences due to the use of different LSs. There are no contributions which try to face the problem of the epistemic uncertainties reduction, using real measurements. Some applications regarding the calibration of the model parameters have been presented in section 2.3, using either static or dynamic measurements, but none of these develops a real SHM framework in a

probabilistic setting. As mentioned before, the available works on static SHM focus their attention on the improvement of the fitting between predictions and observations with no concerns about the stability of the solution or the estimation of the error, thus not really implementing a static SHM system for concrete dams. Regarding the application of dynamic SHM for concrete gravity dams, most of them neglect the coherence problem between numerical and experimental modes. Moreover, none of them defines a probabilistic SHM framework. Summing up, the main gaps in the literature are:

- **Seismic fragility assessment of concrete dams**: there are no contributions which try to introduce procedure for epistemic uncertainties reduction.
- **Static SHM**: there is no probabilistic static SHM framework which allows estimating the error of the prediction itself.
- **Dynamic SHM**: there is no probabilistic dynamic SHM framework which allows estimating the error of the prediction itself. None faces the problem of coherence between experimental and numerical modes in dam engineering, with particular attention to the SSI.

The aim of this work is the definition of a static and a dynamic SHM frameworks, based on the Bayesian inference, able to reduce the epistemic uncertainties in the predictive models of the dam behaviour thus improving both the structural control (Diagnosis step) and the estimation of the seismic fragility of the dam itself (Prognosis step). The calibrated twin models of the structure can be used to control its behaviour during the regular activity or to predict its behaviour during seismic events, thus assessing the structural fragility or controlling the dam behaviour in a more reliable way. In the last part of this work a procedure to design the dynamic monitoring system is proposed. The procedure is based on the Optimal Experimental Design. This technique has been already used to design the optimal monitoring system for concrete dams (Lahmer 2011), with the aim to optimize the damage detection, but in this work the procedure is also used to maximize the information gain with respect to the strength parameters. Therefore, the monitoring system is designed with the double aim of damage detection and strength parameters updating.
The use of the Bayesian inference in SHM is not a novelty (Chatzi 2016) and recently it gains even more attention as shown by the number of papers published about its application to civil structures (Huang et al. 2018). This is mainly due to possibility of transforming an ill-posed problem, as the model calibration, to a well-posed one through the Bayesian inference (Marwala 2010). This can be done by inserting prior information about the parameters. However, the computational cost associated with these kind of procedures is

particularly high, so its application to complex models is still limited. The only available applications of Bayesian updating, in dam engineering field, are related to the calibration of capacity and demand models for the reliability analysis, as reported in Andreini et al. (2016) and Peter et al. (2018). These make use of simplified models in order to reduce the computational burden. In this work the use of meta models allows reducing the computational burden thus making possible the application of Bayesian inference to the SHM of dams.

Chapter 3

Modelling issues in the seismic analysis of concrete gravity dams

3.1 Introduction

Whenever one deals with physical models, regardless their nature, he or she should account for the uncertainties, aleatory or epistemic, which are always involved. Obviously, this is true also for the seismic analysis of existing concrete gravity dams, as discussed in Chapter 2.2. The main source of aleatory uncertainty is the ground motion, while the epistemic uncertainties are mainly related to the model itself, both the assumption needed to build it and the parameters. More specifically, some assumptions are needed in order to define a deterministic model, for instance the choice of materials constitutive models, mesh dimension, interaction and so on. These decisions influence the model output and a better knowledge about the system would leads to reduction of their bias, then this uncertainty has epistemic nature. The most significant sources of uncertainties, related to the choices needed to define the deterministic model in dam engineering field, are:

- *Material constitutive models*: the choice of the material constitutive models is a fundamental step in the deterministic model definition, both for dam concrete and

foundation soil.

- *Soil-Structure Interaction (SSI)*: this is an important aspect which must be faced when big structures are analysed.
- *Fluid-Structure Interaction (FSI)*: in dam analysis, the FSI modelling approach is a fundamental choice which influences all the analyses.

The effects of these three main sources of uncertainties are analysed and discussed in this Chapter.
Once a deterministic model has been defined other epistemic uncertainties arise mainly related to the model parameters. The quantification and reduction of these last uncertainties are the objects of Chapters 5 and 6.

3.2 Material constitutive models

3.2.1 Introduction

The main object of this section is the quantification of the uncertainties due to the choice of the concrete constitutive model and the variation of the strength parameters. Not only the material parameters could be unknown, which is easily understandable, but also the behaviour of the concrete due to its particular characteristics. The so called "dam concrete" is characterized by big size aggregates and an extrapolation from common concrete to dam concrete cannot be directly done, as discussed by Brühwiler and Wittmann (1990). Important efforts have been undertaken from the past in order to study, both theoretically and experimentally, the behaviour of concrete under high loading as reported in Uğurlu (2007) and Wu et al. (2016). During an earthquake, several parts of the dam may experience tensile loading with subsequent crack formation. The safety of these structures is thus controlled by the tensile behaviour of the material (Brühwiler and Wittmann 1990). The upper cracks usually initiate from the upstream or downstream face of the dam and they propagate horizontally, or at an angle towards the opposite face. The consequence of cracking, if extended through the dam section, may be sliding or rotational instability of the separated blocks (Ghanaat 2004) (Zhu and Pekau 2007). The rocking stability of a gravity dam with penetrated cracks was first studied by Saini and Krishna (1973), for the highest monolith of the Koyna Dam. Traditionally, a no-tension stress criterion has been used in the design of concrete dams (Council 1990). However, micro-cracking is always present in concrete and the acceptance of moderate tensile cracking that does not

impair the function of a dam may be a realistic point of view (Council 1990). Several approaches to model the complicated stress-strain behaviour of the concrete are available (Akköse and Şimşek 2010), some of which are based on plasticity models and some others on fracture mechanics. In this context elastic-plastic models (Lee and Fenves 1998a) can be useful to overcome over-stressing problems encountered in the linear analysis of a concrete dam and may predict a more realistic stress distribution in the monoliths during earthquake ground motion. Moreover, plastic regions approximately indicates the critical areas of the structure. In recent years, the non-linear dynamic response of gravity dams under earthquake actions, including cracking of concrete, has attracted more attention of engineers (Hariri-Ardebili et al. 2016a). Stress and crack response of concrete dams may be analysed by means of many non-linear models, commonly applied in most of engineering analysis as reported in Ghrib and Tinawi (1995), Ghaemian and Ghobarah (1999) and Guanglun et al. (2000). Cracking may be modelled by using numerous approaches, which can be classified into two macro-categories: the geometrical approach, that considers the crack a geometrical entity and, if needed, allows updating discretization model with cracks growth; and the non-geometrical one, which only updates the constitutive relationship during the propagation of cracks, the mesh remaining unchanged (Ingraffea 2007). The first one, concerning the discrete cracks, contains two main groups, the linear elastic fracture mechanics (LEFM) and the non-linear fracture mechanics (NLFM). Regarding the latter, there are two basic procedures of modelling cracks commonly used in numerical analysis: the fictitious crack model (FCM), presented by Hillerborg et al. (1976) and the crack band model (CBM), proposed by Bažant and Cedolin (1979) and Bažant and Oh (1983), both of which take the effects of strain softening into account. The FCM overcomes the limitation of LEFM and a non-linear constitutive relation can be introduced in fracture analysis according to the strain-softening mechanism. In this regard, in Pan et al. (2014) a general investigation is presented, in order to evaluate whether the non-linear responses of concrete dams obtained from different fracture modelling approaches are comparable in terms of crack propagation and failure modes.

The second macro-category regards the continuum models and includes smeared cracks and damage mechanics. In this category, two groups may be identified, the constitutive methods and the kinematic ones. The continuum damage model (CDM), belonging to the constitutive methods, offers the possibility of modelling areas where damage causes a multitude of micro-cracks that are not necessarily localized. In particular, in the CDM approach introduced by Rashid (1968), the coalescence of one or more cracks in a volume will result in a deterioration of the stiffness and strength of this volume. The Extended FE Method

(XFEM) approach belongs to the kinematic methods and describes the crack geometry, independently from the background mesh, by enriching the standard displacement-based FE approximation with some pre-knowledge of the physics of crack.
The crack is represented, either in the material constitutive model or in the kinematic model, as an intense localization of strain. In this context, Roth et al. (2015) proposed a crack model that combines the damage mechanics approach and the XFEM in order to predict the propagation of the crack path within the dam section. The CDM (in this case, it comes of a rotating anisotropic damage model) offers the possibility of modelling areas where damage causes a multitude of micro-cracks that are not necessarily localized. It can efficiently predict and continuously adjust crack directions during their evolution. The cohesive XFEM instead, allows a discontinuous displacement field to be well-represented across a localized crack. The use of the CDM allows any initial misprediction of the crack direction to be corrected as a crack grows.
In this section three constitutive models have been considered and compared: an elastic perfectly plastic material with Drucker-Prager yield criteria (Ibrahimbegovic 2009), an elastic plastic damaged model with Lee-Fenves yield criteria (Lee and Fenves 1998b) and a model constituted by non-linear material with damage formulation and a direct limitation of the shear stresses, called "masonry-like" material model (Lucchesi et al. 2017a).

3.2.2 Elastic perfectly plastic material with Drucker-Prager yield criteria

The first considered constitutive model is an elastic perfectly plastic material with Drucker-Prager yield surface (Ibrahimbegovic 2009). The yield function F_y can be defined in terms of the first invariant of the stress tensor I_1 and the second invariant of the deviatoric stress tensor J_2. In this space, the expression of the criteria becomes a line (equation 3.1), as shown in figure 3.2.

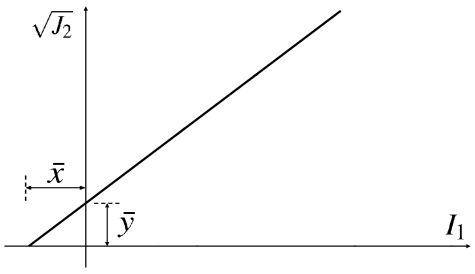


Fig. 3.1: Drucker-Prager failure domain.

$$F(\sigma) = \sqrt{J_2} + \alpha I_1 - k \tag{3.1}$$

The parameters α and k in equation 3.1 define the yield surface and they can be defined w.r.t. the strength parameters of the material,

$$\alpha = \frac{1}{\sqrt{3}} \frac{R_c - R_t}{R_c + R_t} \tag{3.2}$$

$$k = \frac{2}{\sqrt{3}} \frac{R_c R_t}{R_c + R_t}. \tag{3.3}$$

In the equation 3.2 and equation 3.3 R_c and R_t are respectively the compressive and the tensile strength of the material. Finally, the line intersections with the axes, $\bar{x}$ and $\bar{y}$ can be easily determined,

$$\bar{x} = \frac{-\sqrt{3}k}{3\sqrt{3}\alpha} \tag{3.4}$$

$$\bar{y} = \sqrt{3}k. \tag{3.5}$$

In this context, $\bar{x}$ and $\bar{y}$ have been treated as random variables to parametrize the yield criteria. Their distributions can be determined starting from the distributions of the material strengths parameters, which can be defined based on the results of tests or on the data available in literature.

The definition of a plastic constitutive model requires the assumption of a flow rule, which describes the relationship between the increment of the plastic strain tensor $\dot{\epsilon}^p$ and the current state of stress σ. In the field of small plastic strain the direction of the plastic strain increment is defined as

$$\dot{\epsilon}^p = \lambda \frac{\partial Q_p}{\partial \sigma}. \tag{3.6}$$

Where λ is the plastic multiplier, which depends on the current state of stress and load history, and Q_p is the plastic potential. The direction of the plastic strain increment $\dot{\epsilon}^p$ is perpendicular to the surface defined by the plastic potential so the plastic multiplier is determined by the complementary condition,

$$\lambda \geq 0, \quad F_y \leq 0, \quad \lambda F_y = 0. \tag{3.7}$$

When plastic potential and yield function coincide ($Q_p = F_y$) the flow rule is called associated. In this case an associated flow rule has been used.

3.2.3 Elastic-plastic damage model with Lee-Fenves yield criteria

The plastic damage model for cyclic loading of concrete structure, proposed by Lee and Fenves (1998b), is one of the most used constitutive material model for the non-linear analysis of concrete dams. This material constitutive model allows considering both the hardening-softening behaviour and the degradation due to damage evolution.
In this section the most important features of the model are introduced. The evolution equations for hardening variables must be defined for the general multi-axial condition, since 3D FE model has been used in this thesis. Based on Lee and Fenves (1998b) the equivalent plastic strain rates, in tension $\dot{\tilde{\epsilon}}_t^p$ and compression $\dot{\tilde{\epsilon}}_c^p$, are evaluated according to the expressions

$$\begin{aligned} \dot{\tilde{\epsilon}}_t^p &= r\left(\hat{\bar{\boldsymbol{\sigma}}}\right) \hat{\dot{\epsilon}}_{max}^p, \\ \dot{\tilde{\epsilon}}_c^p &= -\left(1 - r\left(\hat{\bar{\boldsymbol{\sigma}}}\right)\right) \hat{\dot{\epsilon}}_{min}^p, \end{aligned} \tag{3.8}$$

where $\hat{\dot{\epsilon}}_{max}^p$ and $\hat{\dot{\epsilon}}_{min}^p$ are respectively the maximum and minimum eigenvalues of the plastic strain rate tensor $\dot{\boldsymbol{\epsilon}}^p$ and

$$r\left(\hat{\bar{\boldsymbol{\sigma}}}\right) = \frac{\sum_{i=1}^{3} \langle \hat{\bar{\sigma}}_i \rangle}{\sum_{i=1}^{3} |\hat{\bar{\sigma}}_i|}, \quad 0 \leq r\left(\hat{\bar{\boldsymbol{\sigma}}}\right) \geq 1 \tag{3.9}$$

is a stress weight factor that is equal to one if the principal stresses $\hat{\bar{\sigma}}_i$ are positive and equal to zero if they are negative. Whether, the eigenvalues of the plastic strain rate tensor $\hat{\dot{\epsilon}}_i^p$ are ordered, then the evolution equation for general multi-axial stress conditions can be expressed in the following matrix form.

$$\dot{\tilde{\boldsymbol{\epsilon}}}^p = \begin{bmatrix} \dot{\tilde{\epsilon}}_t^p \\ \dot{\tilde{\epsilon}}_c^p \end{bmatrix} = \hat{\mathbf{h}}\left(\hat{\bar{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^p\right) \hat{\dot{\boldsymbol{\epsilon}}}^p, \tag{3.10}$$

where

$$\hat{\mathbf{h}}\left(\hat{\bar{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^p\right) = \begin{bmatrix} r\left(\hat{\bar{\boldsymbol{\sigma}}}\right) & 0 & 0 \\ 0 & 0 & -\left(1 - r\left(\hat{\bar{\boldsymbol{\sigma}}}\right)\right) \end{bmatrix}, \tag{3.11}$$

and

$$\hat{\dot{\boldsymbol{\epsilon}}}^p = \begin{bmatrix} \hat{\dot{\epsilon}}_1^p \\ \hat{\dot{\epsilon}}_2^p \\ \hat{\dot{\epsilon}}_3^p \end{bmatrix}. \tag{3.12}$$

The plastic damage concrete model assumes that the elastic stiffness degradation is isotropic and characterised by a single scalar variable d which modifies the initial undamaged elastic stiffness of the material $\mathbf{D}_0^{el}$ obtaining the damaged one $\mathbf{D}^{el}$,

$$\mathbf{D}^{el} = (1-d)\,\mathbf{D}_0^{el}, \quad 0 \leq d \leq 1. \tag{3.13}$$

The scalar degradation variable d must be consistent with the uni-axial monotonic response of the material, and it should capture the complexity associated with the degradation mechanisms under cyclic loading. Therefore, two independents uni-axial damage variable d_t and d_c, function of the plastic strains and field variables, and two restoration functions of the stress state s_t and s_c, which represent the stiffness recovery effect associated with stress reversal, are defined,

$$\begin{aligned} &(1-d) = (1-s_t d_c)\,(1-s_c d_t)\,, \qquad 0 \leq s_t, \quad s_c \leq 1, \\ &s_t = 1 - w_t \hat{\bar{\boldsymbol{\sigma}}}, \quad 0 \leq w_t \leq 1, \\ &s_c = 1 - w_c \left(1 - \hat{\bar{\boldsymbol{\sigma}}}\right), \quad 0 \leq w_c \leq 1. \end{aligned} \tag{3.14}$$

The weight factors w_t and w_c, which are assumed to be material properties, control the recovery of the tensile and compressive stiffness upon load reversal. Since tensile damage is considered, the plastic strains in tension $\tilde{\epsilon}_t^p$ are defined as a function of the damage variable d_t and the cracking strains $\tilde{\epsilon}_t^{ck}$,

$$\tilde{\epsilon}_t^p = \tilde{\epsilon}_t^{ck} - \frac{d_t}{(1-d_t)}\frac{\sigma_t}{E_0}. \tag{3.15}$$

In equation 3.15 σ_t are the tensile stresses and E_0 is the undamaged elastic modulus of the material.

The plastic damage concrete model uses a yield criteria based on the yield function proposed by Lubliner et al. (1989), and incorporates the modifications proposed by the authors to account for different evolution of strength under tension and compression. Therefore, the yield function in terms of effective stresses $\bar{\boldsymbol{\sigma}} = \frac{\boldsymbol{\sigma}}{(1-d)}$ takes the form

$$F\left(\bar{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^p\right) = \frac{1}{1-\alpha}\left[\bar{q} - 3\alpha\bar{p} + \beta\left(\tilde{\boldsymbol{\epsilon}}^p\right)\langle\hat{\bar{\sigma}}_{\max}\rangle - \langle -\hat{\bar{\sigma}}_{\max}\rangle\gamma\right] - \bar{\sigma}_c\left(\tilde{\epsilon}_c^p\right), \tag{3.16}$$

where α and γ are dimensionless material constants,

$$\bar{p} = \frac{1}{3}\bar{\boldsymbol{\sigma}} : \mathbf{I} \tag{3.17}$$

is the effective hydrostatic pressure,

$$\bar{q} = \sqrt{\frac{1}{3}\bar{\mathbf{S}} : \bar{\mathbf{S}}} \tag{3.18}$$

is the Mises equivalent effective stress,

$$\bar{\mathbf{S}} = \bar{p}\mathbf{I} + \bar{\boldsymbol{\sigma}}, \tag{3.19}$$

is the deviatoric part of the effective stress tensor $\bar{\boldsymbol{\sigma}}$ and $\hat{\bar{\sigma}}_{\max}$ is the algebraically maximum eigenvalue of $\bar{\boldsymbol{\sigma}}$. The function $\beta\left(\tilde{\boldsymbol{\epsilon}}^p\right)$ is given as

$$\beta\left(\tilde{\boldsymbol{\epsilon}}^p\right) = \frac{\bar{\sigma}_c\left(\tilde{\epsilon}_c^p\right)}{\bar{\sigma}_t\left(\tilde{\epsilon}_t^p\right)}(1-\alpha) - (1+\alpha), \tag{3.20}$$

where $\bar{\sigma}_t$ and $\bar{\sigma}_c$ are the effective tensile and compressive cohesion stresses, respectively. As mentioned before, α and γ are material constant, the former related to biaxial and uni-axial stresses, while the latter defines the shapes of failure domain in the principal stresses space.
Finally, this material constitutive model assumes non associated flow rule. The Drucker-Prager hyperbolic function is used as plastic potential,

$$Q_p = \sqrt{\left(\varepsilon\sigma_{t0}\tan\left(\psi\right)\right)^2 + \bar{q}^2} - \bar{p}\tan\psi, \tag{3.21}$$

where, ψ is the dilatation angle, σ_{t0} is the uni-axial tensile stress at failure, ε is the eccentricity parameter which defines the rate at which the function approaches the asymptote. The model parametrization has been directly done on the values of the tensile and compressive strengths, assuming that the values of the cohesions were equal to them (Lee and Fenves 1998b). In this way, as indicated in figure 3.2, the domain vertices $\mathbf{P}_1$ and $\mathbf{P}_2$ move along the diagonal keeping constant the shape.

3.2.4 Extended "masonry-like" material

The "masonry-like" model, in its original concept, is a no-tension normal material developed for the study of masonry structures (Lucchesi et al. 2008). In recent years, several authors have contributed to its extension limiting the tensile strength, and bounding the shear stress (Lucchesi et al. 2017b). In this section this material constitutive model is briefly described. Being a normal material model, the constitutive relation is fully described once the tensor of the elastic modulus $\mathbb{C}$ and the space of the admissible tensions

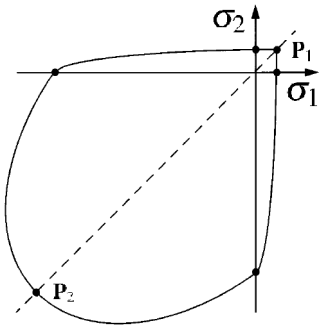


Fig. 3.2: Concrete damage plasticity failure domain.

$\mathcal{K}$ are specified. In particular, $\mathcal{K}$ is a closed and convex subset of Sym, which is the space of all symmetric second order tensors. Once the strain tensor $\mathbf{E}$ has been assigned, the stress tensor $\mathbf{T}$ is obtained by projecting $\mathbb{C} \cdot \mathbf{E}$ on $\mathcal{K}$, with respect to the complementary energy norm. When $\mathcal{K}$ is a polyhedron, an explicit expression of these projections can be determined (Lucchesi et al. 2017b). In this context, the polyhedral shape of $\mathcal{K}$ is defined assuming that the principal stresses belong to the interval $[-R_c, R_t]$, where R_c is the compressive strength and R_t is the tensile strength. The maximum allowed value of the tangential component of the stress τ depends on each plane on the normal component σ,

$$|\tau| = \tau_0 - \sigma \tan(\phi). \tag{3.22}$$

Where τ_0 and ϕ are positive material parameters. Here they have a role analogous to the ones played by the cohesion and the friction angle in the theory of soil plasticity (Chen 2007). Finally, in Lucchesi et al. (2017c) and De Falco et al. (2017) a softening law for the tensile strength, is introduced.
One of the most interesting feature of the "masonry-like" constitutive model is the explicit calculation of the cracking strain, which makes the computational burden very low. Assuming an isotropic behaviour of the material, the relationship among the stress tensor $\mathbf{T}$, the total strain tensor $\mathbf{E}$ and the inelastic part of the strain tensor $\mathbf{E}^a$, is

$$\mathbf{T} = \mathbb{C}(\mathbf{E} - \mathbf{E}^a) = \frac{E}{1+\nu}(\mathbf{E} - \mathbf{E}^a) + \frac{\nu E}{(1+\nu)(1-2\nu)} \mathrm{tr}(\mathbf{E} - \mathbf{E}^a)\mathbf{I}, \tag{3.23}$$

where E is the Young modulus and ν the Poisson's coefficient of the material. Moreover, $\mathbf{T}$, $\mathbf{E}$ and $\mathbf{E}$ are coaxial tensors with common eigenvectors $\mathbf{e}_i$ and corresponding eigenvalues

dented by t_i, e_i and a_i, respectively. Therefore, the constrain of the stresses, previously specified, make a quadrilateral failure domain, shown in figure 3.3, with vertices,

$$\mathbf{T}_1 \equiv (R_t, R_t) \tag{3.24}$$

$$\mathbf{T}_2 \equiv \left(\frac{R_t \left(1 + \sin\left(\phi\right)\right) - 2\tau_0 \cos\left(\phi\right)}{1 - \sin\left(\phi\right)}, R_t \right) \tag{3.25}$$

$$\mathbf{T}_3 \equiv \left(-R_c, \frac{2\tau_0 \cos\left(\phi\right) - R_c \left(1 - \sin\left(\phi\right)\right)}{1 + \sin\left(\phi\right)} \right) \tag{3.26}$$

$$\mathbf{T}_4 \equiv \left(-R_c, -R_c\right). \tag{3.27}$$

In the same way, the elastic strain range in the space of the principal strains is a quadrilateral (figure 3.3), with vertices,

$$\mathbf{E}_1 \equiv \left(\frac{\epsilon_t}{1+\nu}, \frac{\epsilon_t}{1+\nu} \right) \tag{3.28}$$

$$\mathbf{E}_2 \equiv \left(\frac{\epsilon_t \left(\left(1 + \sin\phi\right) - \nu\left(1 - \sin\phi\right)\right) - 2\tau_0 \cos\phi}{\left(1-\nu^2\right)\left(1 - \sin\phi\right)}, \frac{\epsilon_t \left(\left(1 - \sin\phi\right) - \nu\left(1 + \sin\phi\right)\right) + 2\nu\tau_0 \cos\phi}{\left(1-\nu^2\right)\left(1 - \sin\phi\right)} \right) \tag{3.29}$$

$$\mathbf{E}_3 \equiv \left(\frac{\epsilon_c \left(\nu\left(1 - \sin\phi\right) - \left(1 + \sin\phi\right)\right) - 2\nu\tau_0 \cos\phi}{\left(1-\nu^2\right)\left(1 + \sin\phi\right)}, \frac{\epsilon_c \left(\nu\left(1 + \sin\phi\right) - \left(1 - \sin\phi\right)\right) + 2\tau_0 \cos\phi}{\left(1-\nu^2\right)\left(1 + \sin\phi\right)} \right) \tag{3.30}$$

$$\mathbf{E}_4 \equiv \left(\frac{-\epsilon_c}{1+\nu}, \frac{-\epsilon_c}{1+\nu} \right), \tag{3.31}$$

where ϵ_t and ϵ_c are the strain values for which the material strengths in tension and compression, respectively, are reached.

Therefore, since the elastic domains $\mathcal{K}$ and $\mathcal{E}$ are a priori defined, the principal strains space $\mathcal{E}$ can be partitioned. In this way, a general strain state is projected to a segment of the elastic domain in the space $\mathcal{E}$ and the relative stress state is easily determined through the elastic tensor $\mathbb{C}$.

In this application, the parametrization of the failure domain $\mathcal{K}$ has been done by moving the positions of the vertices, as shown in figure 3.4. The positions of the vertices of the polyhedral domain depend on the values of the material strengths parameters, in particular $\mathbf{T}_1$ and $\mathbf{T}_4$ depend on the values of the tensile and compressive strengths, so their distributions may be directly obtained from the distributions of R_t and R_c. The positions of $\mathbf{T}_2$ and $\mathbf{T}_3$ depend not only on the values of R_t and R_c but also on the values of the tangential stress parameters, τ_0 and ϕ. The available information about τ_0 and ϕ

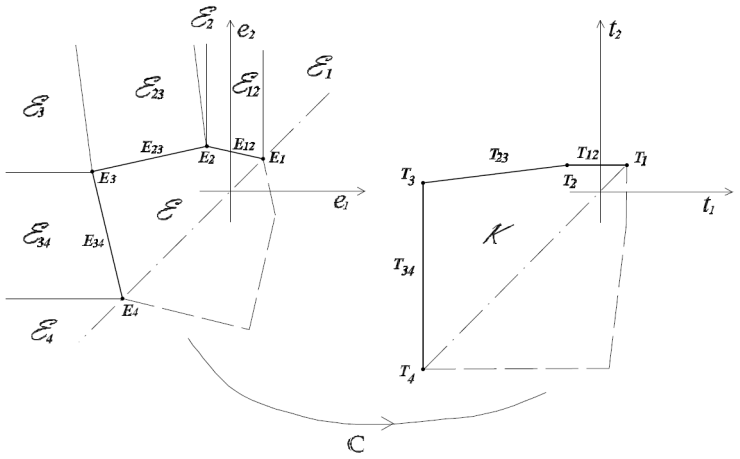


Fig. 3.3: "masonry-like" constitutive model: partition of the strain plain and projection, adapted from Lucchesi et al. (2017b).

are usually very poor, so the segments s and b have been modelled as random variables, using uniform distributions.

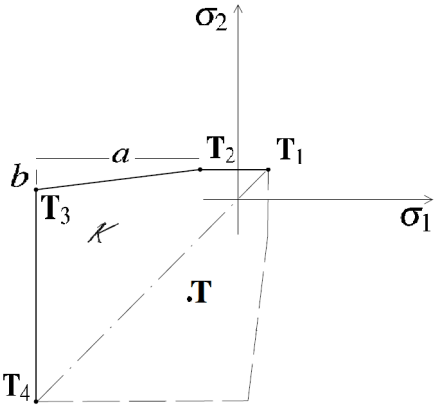


Fig. 3.4: "masonry-like" failure domain.

3.2.5 Comparative analysis

The Koyna Dam, a 103-m-high concrete gravity dam in India, is a benchmark problem which has been widely examined by several researchers in the field of dam modelling. This case is particularly interesting due to the effects registered after the seismic event occurred in the dam area in 1967. Based on a seismic coefficient of 0.05, uniform with the height, the

earthquake forces were expected to cause no tensile stresses. However, the 1967 earthquake caused significant cracking in the dam. The higher monolith of the non-overflow section suffered the worst damage during the seismic event, endangering its stability during future shocks. It is believed that this exaggerated damage resulted from an elevator tower that extended 15 meters above the top of the block and it was therefore subjected to greatly increased inertial forces.

From the computational point of view, the presence of smooth vertical contraction joints enables the use of 2D plane stress models of individual monoliths to predict the seismic behaviour of the structure. Among the many studies of the Koyna dam, the most interesting works are those presented in Gioia et al. (1992), Ghrib and Tinawi (1995) and Roth et al. (2015) in which a static analysis of the dam was performed on cases of reservoir overflow, using fracture mechanics and plasticity-based models. Moreover, Omidi et al. (2013) studied the seismic behaviour of the Koyna dam, varying the damping mechanism in a plastic-damage model. De Falco et al. (2017) compared the results presented by Roth et al. (2015) with those obtained using a simplified discrete-sliding-face model and the extended masonry-like material model with bounded shear stress. In this section the comparative analyses have the main aim to show the influence of the choice of the material constitutive model on the output, and to highlight how the epistemic uncertainties related to the material model parameters, of each of them, lead to a variation of the results.

Starting from the results of material tests, the tensile and the compressive strengths have been modelled as RVs, assuming log-normal distributions as indicated in table 3.1. The other mechanical parameters, needed for the analysis, have been defined according to the literature (Omidi et al. 2013). The General Polynomial Chaos expansion (GPCe) tech-

Table 3.1: Distribution of the concrete strength parameters

Parameter	Mean Value	Standard deviation
R_c (MPa)	25.31	6.33
R_t (MPa)	2.23	0.56

nique, introduced in Chapter 4.4, has been used to speed up the procedure, reducing the analyses number, which in this work is equal to 150. The polynomial coefficients have been determined based on the results of the FE model, which is a plane stress model composed by 658 four-nodes elements and 706 joints. The effect of the soil-structure interaction has been considered by using an appropriate Rayleigh coefficient, and a portion of foundation soil, properly calibrated in terms of stiffness. The effect of the water-structure interaction has been considered by applying the added-masses approach, which allows modelling the

hydrodynamic pressure only by adding lamped masses properly calibrated (Westergaard 1933). The model is subjected to gravity, hydrostatic pressure, of a full reservoir level, and the horizontal ground acceleration components of the 1969 seismic events, while water pressure inside the cracks is neglected.

The results are presented in terms of crest displacements and base shear, highlighting the

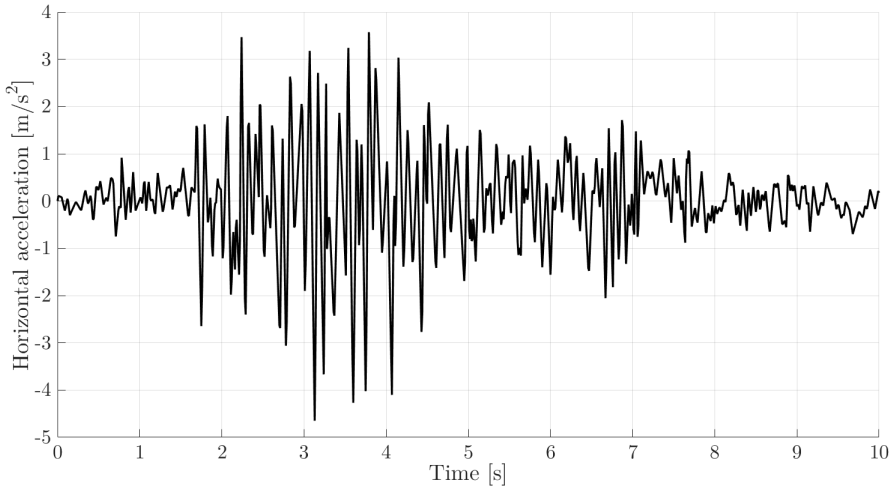


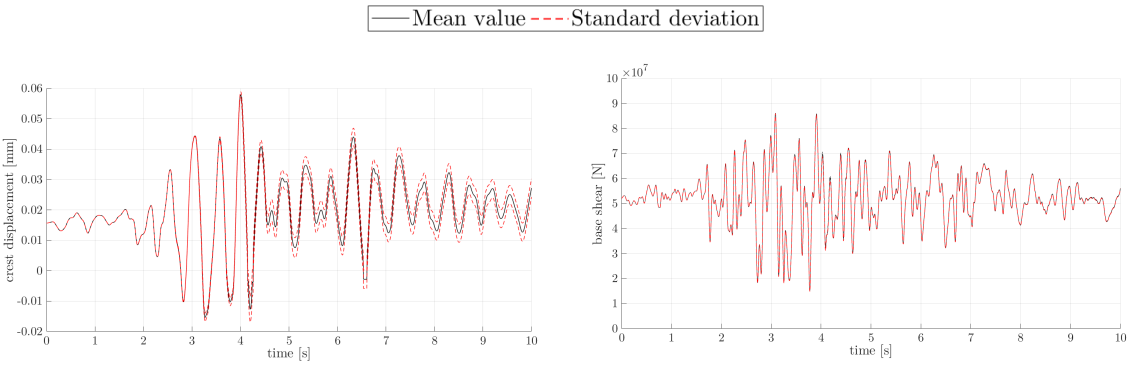
Fig. 3.5: Koyna horizontal ground acceleration.

effect of the strength parameters uncertainty, and comparing the results obtained with the three different constitutive models.

The results of the three models, in terms of crest displacements and base shear, are shown in figures 3.6, 3.7 and 3.8. In all cases, the crest displacements are more sensitive to the strength parameters variation and this fact is underlined by a bigger variance with respect to the base shear.

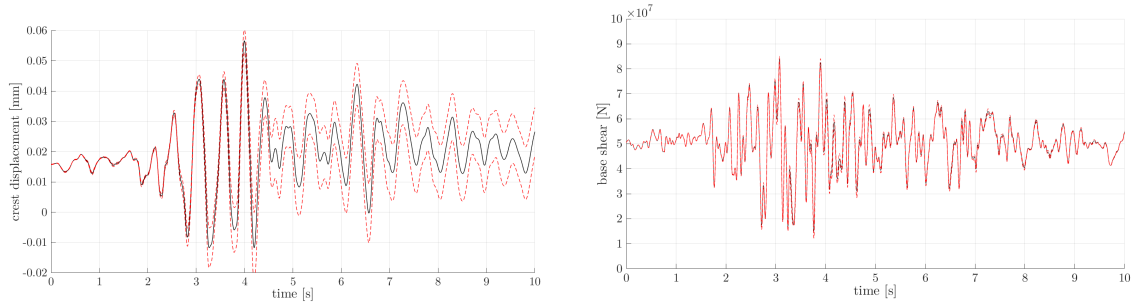
Referring again to figures 3.6, 3.7 and 3.8, and comparing the mean values, several differences can be observed, in particular the displacements are more sensitive to the model choice. The "masonry-like" material shows a different behaviour, if compared to the other two models, both in terms of base shear and displacement. By analysing the collapse mechanisms of the dam the reason of these differences can be found in the direct control of the shear stresses. In fact, this particular constitutive model allows the cracks growing in the proximity of the downstream face due to the direct control of the shear stress.

The maximum values of the relative error of the gPCE in terms of mean values and variances are indicated in table 3.2. The base shear relative errors are smaller than those of the crest displacement and the reason of this fact is that the base shear is usually less



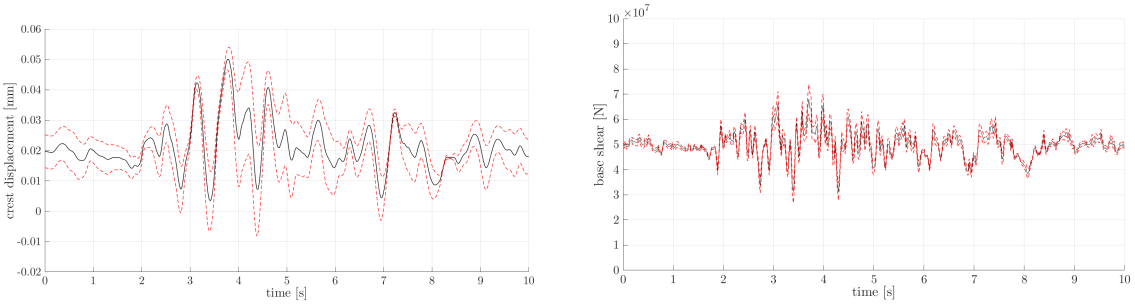
(b) Crest displacement. **(c)** Base shear.

Fig. 3.6: FEA results considering the Elastic-perfectly plastic material model.



(a) Crest displacement. **(b)** Base shear.

Fig. 3.7: FEA results considering the Concrete Damage Plasticity material model.



(a) Crest displacement. **(b)** Base shear.

Fig. 3.8: FEA results considering the "Masonry-like" material model.

influenced by the strength parameters variation, as can be noted in figures 3.6, 3.7 and 3.8.

Table 3.2: Relative errors

Maximum Relative error	mean value	variance
Perfectly plastic material, crest displacement	5.15e-3	0.13
Concrete damage plasticity, crest displacement	3.96e-6	0.10
Extended "masonry-like" material, crest displacement	3.55e-7	0.21
Perfectly plastic material, base shear	1.77e-15	0.33
Concrete damage plasticity, base shear	2.55e-16	0.27
Extended "masonry-like" material, base shear	1.08e-15	0.31

3.2.6 Concluding remarks

In this section the effects on the dam model related to the uncertainty of the material strength parameters are shown, highlighting how they play differently varying the material constitutive model. It is worth noting that the *best model* can be selected only with regard to the available information. Indeed, the Bayesian Ockham Razor (Beck 2010) shows that the model class selection is a trade-off between data-fit and model complexity. This aspect must always be considered when a numerical model is performed.

3.3 Soil-structure interaction (SSI)

3.3.1 Introduction

One of the biggest issues, and source of uncertainty, related to existing dams modelling is the soil-structure interaction (SSI). In this regard, SSI was addressed by many authors who were searching for a reliable simulation of wave propagation in a semi-infinite medium. Nowadays, the Finite Elements method is the most common for a coupled study using both structural and acoustic elements, and it can also simulate the unboundedness of both terrain and reservoir.
In this section, the SSI is considered for existing concrete gravity dams, investigating its effects numerically on 2D FE plane strain models under ground motion. More specifically, the advantages of modelling half-unbounded domains are shown performing frequency response analyses under different boundary conditions and modelling approaches.

3.3.2 Modelling approach

During ground shaking, the dam-reservoir-foundation system must be considered a coupled system. To date models seldom take into account full interaction effects, because of the lack of adequate numerical implementations or computational resources required by three dimensional detailed models. SSI is described to have two main components: kinematic and inertial interaction ("Dynamic soil-structure interaction, John P. Wolf, Prentice-Hall, Englewood Cliffs, N.J., 1985. No. of pages: £466. Price: £61.45"). The former is governed by soil flexibility. In this regard, the massless foundation model proposed by Clough in 1980 (Clough 1980) has been extensively used in seismic analysis of dam-foundation problems. In this model, recorded displacements are imposed at the boundaries of the domain and the input motion reaches instantaneously the base of the dam. Wave velocity in foundation becomes infinite and the structure takes all kinetic energy. These assumptions seem in general unrealistic (Tan and Chopra 1995).
Inertial interaction is generated by elastic waves that develop under dynamic loads, promoting the energy transport through the soil volume. Such a phenomenon, that carries energy away from the structure, is often referred as "radiation damping". So, while in static SSI analysis the simple truncation of the far field with setting of appropriate boundary conditions gives very often good results, in dynamic cases it makes results to be erroneous because of reflection waves.
Recently, SSI for concrete retaining structures has been addressed by many authors, searching for a reliable simulation of wave propagation in a semi-infinite medium, by modelling the far field part of the foundation. Some interesting methods are: Lysmer boundary conditions (Lysmer and Kuhlemeyer 1969), hyper-elements ("A technique for the analysis of the response of dams to earthquakes"), infinite elements (Kim and Yun 2000), (Yun et al. 2000), rational boundary conditions (Feltrin 1997), boundary element method (Yazdchi et al. 1998), scaled boundary element method (Song and Wolf 2000) and high order non-reflecting boundary conditions (Givoli 2004).
Among the previous approach particularly interesting and widely used are the PML, the infinite elements and the non-reflecting boundary conditions. PML is a technique able to absorb incident waves under any angle and frequency, preventing them from returning back to the medium after incidence to the model boundaries (Johnson 2008). The procedure, first introduced by Berenger in 1994 (Berenger 1994), may be applied to different physical problems. It comes to a complex coordinate stretching of the domain to introduce a decay of the oscillation without any reflection in the source domain, simulating a perfectly absorbing material. The rational scaling of PML is expressed by the following function of

the dimensionless coordinate (Berenger 1994)

$$f_r(\zeta) = s\zeta\left(\frac{1}{3p(1-\zeta)+4} - \frac{i}{3p(1-\zeta)}\right), \tag{3.32}$$

where p is the curvature parameter and s the scaling parameter.
Infinite Elements (IEs), used to incorporate unbounded domains into the finite element method, have a formulation similar to those of FE, except for the infinite extent of the element region and shape function in one direction. Infinite elements method is based on a function which maps the global to the local coordinate system,

$$f(\zeta) = \frac{\zeta}{\gamma - \zeta}\delta p, \tag{3.33}$$

where δp is the pole distance, $\gamma = \frac{\delta s + \delta p}{\delta s}$ and δs is the scaled thickness (Zienkiewicz et al. 1983).
Finally, the low reflecting boundary condition is obtained by imposing a mechanical impedance on the foundation boundary of the model, following the equation

$$\mathbf{T}\cdot\mathbf{n} = -\mathbf{D}\frac{d\mathbf{u}}{d\mathbf{u}}, \tag{3.34}$$

where, $\mathbf{u}$ is the displacements vector, $\mathbf{T}$ the stress tensor, $\mathbf{n}$ the unit vector of the boundary tangent plane and $\mathbf{D}$ is the impedance matrix (Holzapfel 2000),

$$\begin{aligned} \mathbf{D} &= \frac{c_p + c_s}{2}\mathbf{I} \\ c_p &= \sqrt{\frac{K + \frac{4}{3}G}{\rho}} \\ c_s &= \sqrt{\frac{G}{\rho}}, \end{aligned} \tag{3.35}$$

where K is the bulk modulus, G the shear modulus and ρ the density.
In order to understand the necessity of considering both solid and fluid domains as unbounded, three different modelling options are explored in this section: the rigid soil, the Perfectly Matched Layer (PML) technique and the massless soil.

3.3.3 Frequency response analysis

A frequency response analysis is first performed in order to evaluate different modelling approaches for the coupled system regarding the fluid part of the model, the soil and its unboundedness. The case study is an Italian concrete gravity dam 65 meters high.
Three models were set up by using COMSOL Multiphysics® (COMSOL Multiphysics 2009) software (figure 3.9). The first reference model simulates the dam on rigid terrain (figure 3.9a), the second model includes the standard massless foundation in a bounded region (figure 3.9b) and the third model accounts for foundation soil as unbounded half-space provided with mass figure 3.9. Regarding dam and soil domains, the standard Solid Mechanics equations are applied. The solid mesh is made of default second-order serendipity elements: 453 for model 1, 3146 for model 2 and 3692 for model 3. The material setting of the three plane-strain models is reported in table 3.3.

As for the fluid subsystem (full reservoir), it has been simulated both by Westergaard

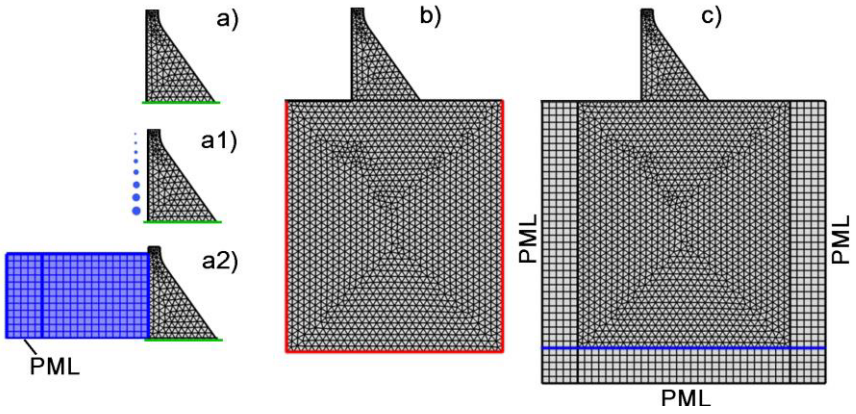


Fig. 3.9: a) model 1: rigid soil; b) model 2: massless soil; c) model 3: infinite terrain model. a1) added mass model; a2) fluid-structure interaction model.

added mass model (Westergaard 1933) and by the Helmoltz equation derived from the full Navier-Stokes equation assuming small vibrations and neglecting viscosity. In this context, the effects of considering coupling and unboundedness of fluid and solid domains are evaluated. PMLs were applied at the bottom and on both sides of the terrain domain and also at the upstream side of the fluid region, in order to simulate the unboundedness of both domains, when explicitly modelled. The PML mesh element sides are directed along the radiation path. In addition, according to Kuhlemeyer and Lysmer (1973), the spatial element size must be smaller than approximately one-eighth of the wavelength associated with highest input frequency.

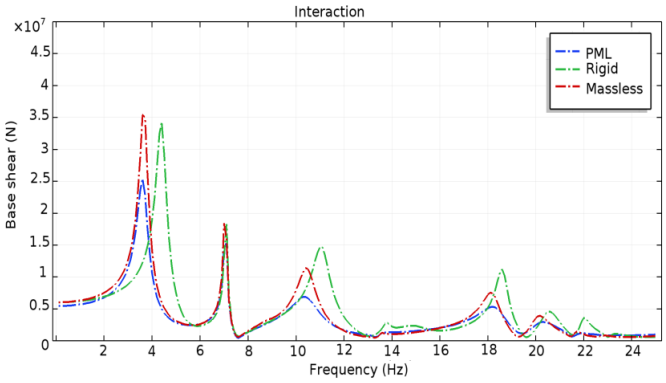
Table 3.3: Materials parameters

	Concrete	Foundation rock
Density ρ $\left[\mathrm{kg/m^3}\right]$	2450	2300
Young modulus E [MPa]	20500	22000
Poisson modulus ν	0.2	0.2
Damping coefficient ξ	0.05	0.05

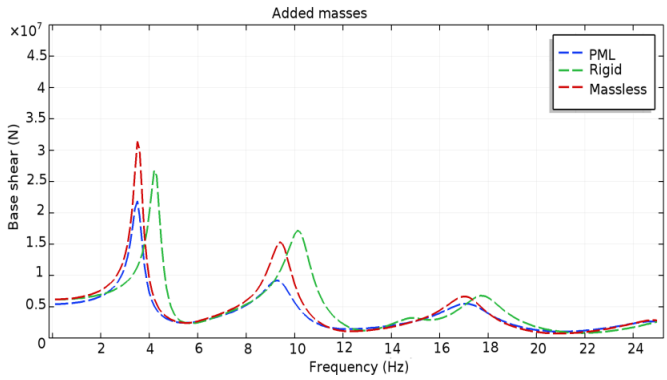
In this case, the frequency sweep ranges between 0 Hz and 25 Hz and the maximum mesh size is 5 m. The analysis has been conducted under both the hypotheses of empty and full reservoir. In this latter case, the basin is simulated using both the simplified Westergaard (Westergaard 1933) added mass model (figure 3.9 a1) and the full acoustic coupling (figure 3.9 a2). When acoustic water domain is modelled, a zero pressure boundary condition is set at the free surface and a rigid boundary condition at the bottom. Base shear has been obtained for a total of nine cases, in order to detect the soil effect.
To make the three soil models comparable to each other, the same horizontal harmonic acceleration has to be provided at the base of the dam. In particular, in models 1 and 2, a horizontal acceleration with peak amplitude of 1 $[\mathrm{m/s^2}]$ is applied to the green and the red contours, respectively.
Unfortunately, the application of PMLs on the exterior boundaries of the terrain model makes them no longer available for a displacement boundary condition. Similarly, by applying the kinematic condition to the interface between the dam and the soil, the outgoing waves are blocked since the field variable u is fixed. This problem is overcome by leaving the displacement unconstrained, a distributed force load was applied at the bottom PML blue interface figure 3.9 and its value was "tuned", frequency by frequency, by a global equation, enforcing an average unit acceleration condition at the dam base.
The results in terms of frequency response for base shear are displayed in all cases of empty reservoir, added masses and full interaction in figure 3.10. The infinite terrain model (blue curves), shows a downward frequency shift with respect to the rigid soil model (green curves) in both cases of Fluid Structure Interaction and added masses. Moreover, it displays a noticeable reduction of the peak response in comparison to the massless case, due to radiation damping. In the case of FSI, the first eigenmode of the reservoir remains well visible, at

$$f = \frac{c}{4h} = \frac{1483\,[\mathrm{m/s}]}{4 \cdot 60\,[\mathrm{m}]} = 6.17\,[\mathrm{Hz}] \tag{3.36}$$

where c is the speed of sound in the water and h is the height of the basin. The reservoir driven resonances can be easily identified, due to their very narrow bandwidth and their frequency location. The fluid pressure and the solid displacement colour maps for models b) and c) in figure 3.9 under a 10 Hz frequency excitation are reported in figure 3.11.
The streamlines represent the acoustic energy flux within the fluid and the mechanical energy flux within the solid that originate from the bottom of the soil domain. A remarkable difference between the two systems can be observed. The massless model displays a circulatory streamlines pattern that becomes curly when frequency increases, without a defined incoming wave front. The infinite soil model, instead, displays a well-defined energy flux direction, as well as a lower amount of energy transmitted to the basin with respect to the former case.



(a) Interaction.



(b) Added masses.

Fig. 3.10: Frequency response curves of the base shear for different approaches.

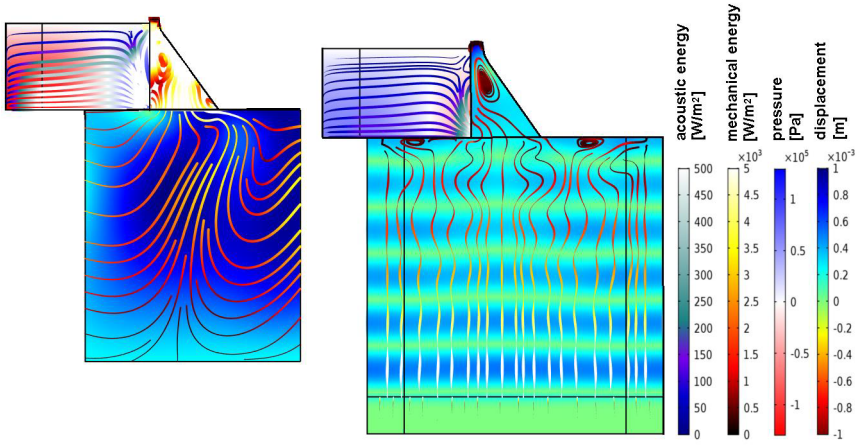


Fig. 3.11: Displacement, pressure and energy flux streamlines plot for massless soil (left) and unbounded soil (right).

3.3.4 Parametric analysis

In order to evaluate the effect of the change in stiffness and density of the soil, a parametric study is performed on a model in the case of empty and full reservoir, allowing a deeper understanding of the terrain contribution. In figure 3.12 response curves in term of base shear are shown for empty reservoir and full reservoir, varying the soil stiffness E_g and density ρ_g. The corresponding parameters of concrete are kept unchanged (table 3.3). The graphs are expressed in function of the logarithm of the ratio between the terrain parameter and the corresponding concrete value in table 3.3. Soil parameters are varied one at a time, on a wide range of values, well beyond a realistic distribution, to emphasize the different effects of each one. It may be deduced that

- For increasing values of the terrain stiffness, both the frequency and the amplitude of the peak response increase.
- For increasing values of the terrain density, the peak response decreases while its frequency location remains unchanged.
- If the terrain has some flexibility, the system's resonant frequency is always lower than the rigid case, regardless of terrain density.
- In case of full reservoir, the peak of the first frequency of the basin is always evident and even more noticeable with increasing density and decreasing stiffness.

As expected, the asymptotic response for very stiff terrain converges to the rigid foundation model (black dashed line). Results obtained from the response analysis in terms of crest acceleration display the same features as the base shear curves.

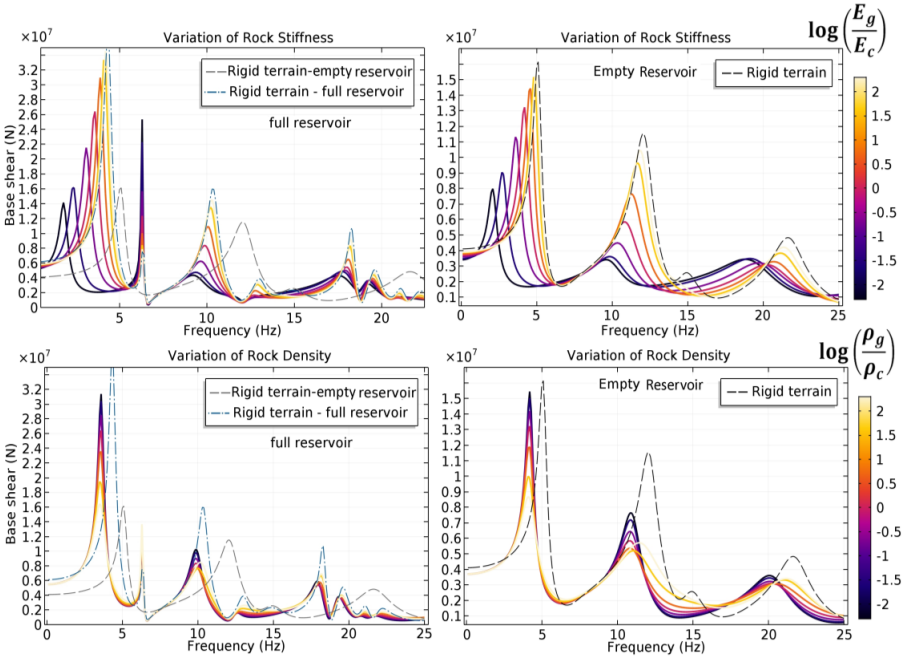


Fig. 3.12: Parametric variation of base shear response curve with soil relative stiffness and density – full reservoir (on the left), empty reservoir (on the right).

3.4 Fluid-structure interaction (FSI)

3.4.1 Introduction

Hydrodynamic forces evaluation on water retaining structures during earthquakes is a challenging issue. In the past, hydrodynamic loads exerted by the reservoir on the upstream face of the dam were obtained from simplified theories, such as Westergaard's (Westergaard 1933) and Zangar's (Zangar 1953), dating back to the first half of the twentieth century. Such theories, initially based on very simplistic hypotheses and successively developed by Chopra (Chopra 1967), are nowadays the base of main international codes and guidelines. The well-known Westergaard simplified theory was developed in 1933 and originated the widespread added masses approach. This useful simplification adopted in highly sophisticated and irregular 3D models could however lead to inaccurate predictions. The possible

differences are estimated in this section.
The physical phenomenon has been investigated using Finite Element models by removing some of the simplifications underlying the aforementioned theories. Hydrodynamic pressures on an Italian large gravity dam have been calculated using different modelling approaches. The analysis has been carried out using two and three-dimensional models, and a parametric variation of boundary conditions has been performed.

3.4.2 Analytical solution

Westergaard's paper (Westergaard 1933) dating back to 1933 is considered as the milestone for the evaluation of hydrodynamic loads on water retaining structures, given its simple yet elegant solution and physical interpretation of the phenomenon. The key assumptions of his solution are the following:

1. small displacements of water particles, neglecting global mass flow;
2. non-viscous fluid;
3. infinite upstream extension of the water reservoir;
4. no surface waves;
5. perfect waves reflection of the reservoir bottom;
6. horizontal and rigid water reservoir bottom surface;
7. rigid, sinusoidal motion of the dam;
8. plane symmetry of the problem;
9. earthquake dominant period higher than dam natural vibration period.

The first three hypotheses allowed obtaining a substantial simplification of the Navier-Stokes equation. These equations, valid for general fluid motion, turn into the more manageable linear wave equation, namely D'Alembert's equation

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t} = 0 \tag{3.37}$$

where p is the pressure and c is the speed of sound in the medium equal to $\sqrt{K/\rho}$, where k is the water bulk modulus and ρ its density. The solution is expressed as particle horizontal

and vertical displacement, provided by Fourier series. The vertical displacements tend to infinity as the excitation period reaches the values:

$$T_n = \frac{4h}{c \cdot n} \tag{3.38}$$

where h is the water height. If that is the case, resonance occurs and the corresponding values of frequency $f_n = 1/T_n$ are defined as reservoir eigenfrequencies. The solution of equation 3.37 can be accomplished both in the time and in the frequency domain, through the well-known Fourier transform.

In the same paper, Westergaard also introduced a simplified expression for the displacement and a physical interpretation of the hydrodynamic forces. Their effect can be estimated by neglecting their dependence on the vibration period and by approximating the pressure distribution by a parabolic expression. If one accepts the trade off, their behaviour is identical to that of a mass distribution $m(y)$ rigidly attached to the upstream face of the dam, according to the following distribution along the height y, measured downwards from the free surface

$$m(y) = \frac{7}{8}\rho\sqrt{hy} \tag{3.39}$$

In a later paper, Chopra (Chopra 1967) extended Westergaard's solution to the whole range of frequencies, by removing hypothesis 9, providing an analogous and more complicated Fourier series solution.

Analytical solutions of the problem did not have any major improvement over the decades, mainly because of the increasing availability of computing resources. Westergaard's solution remained the benchmark and the reference for the international codes up to modern days. In this section the last three hypotheses are removed, and the arising consequences are analysed.

3.4.3 Case study

3.4.3.1 Introduction

The aforementioned investigation is developed by means of a case study regarding an existing Italian dam. The structure is an ordinary concrete gravity dam made up of 27 concrete monoliths, with a maximum height of 65 m, a straight shape in plan, a vertical upstream face and a downstream face with a slope of 0.72. Concrete material parameters are summarized in table 3.4 while water level is supposed at maximum allowed of 60 m.

Table 3.4: Concrete materials parameters.

Density ρ $\left[\mathrm{kg/m^3}\right]$	2450
Young modulus E [MPa]	20550
Poisson modulus ν	0.18
Damping coefficient ξ	0.05

3.4.3.2 FEM analyses of 2D and 3D dam models

2D and 3D finite element models have been performed in COMSOL Multiphysics 5.3 software (COMSOL Multiphysics 2009), as illustrated in the figures 3.13 and 3.14. Each of the two models is studied under different hypotheses

- *Full reservoir, added masses*: hydrodynamic actions are introduced according to Westergaard's simplified expression 3.39 for equivalent masses rigidly connected to the structure. The correct implementation of such masses requires their value to be assigned only to the horizontal degree of freedom of the corresponding mesh element, leaving the vertical contribution to inertial forces unmodified.
- *Full reservoir, fluid-structure interaction*: the fluid domain is explicitly modelled and an acoustic-structural interface is introduced at the dam upstream face.
- *Full reservoir, rigid structure*: this hypothesis is investigated mainly for comparison and interpretation of the previous models.
- *Empty reservoir*: the acoustic part of the model is ignored. This model is used for comparison purposes only, to appreciate the structural effects of the hydrodynamic force.

The 2D model (figure 3.13) is made up of 300 triangular second order Serendipity elements for the solid mechanics part and 624 quadrilateral second order Lagrange elements for the acoustics part, for a total of 841 nodes. The 3D model (figure 3.14) is made up of 8837 tetrahedral second order Serendipity elements for the solid mechanics part and 42384 tetrahedral second order Lagrange elements for the acoustics part, for a total of 17086 nodes.
The two models are analysed by performing a frequency domain study under harmonic base acceleration excitation. Regarding the boundary conditions, on the solid mechanics component, a unit, horizontal acceleration is imposed at the dam base nodes, considering a rigid foundation motion. As for the acoustic component, a zero pressure condition is

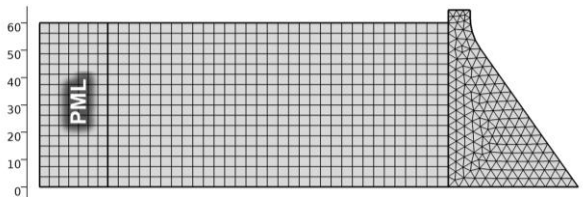


Fig. 3.13: 2D model, Fluid Structure Interaction.

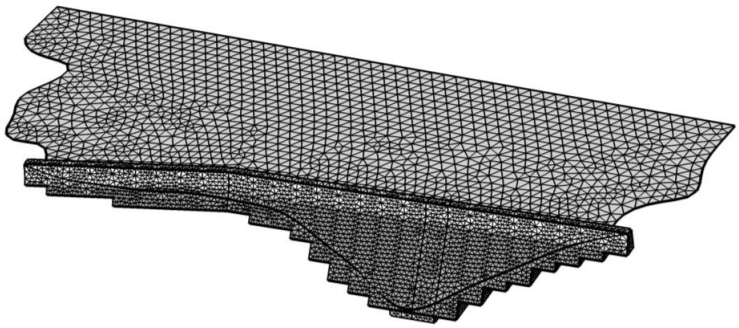


Fig. 3.14: 3D model, Fluid Structure Interaction.

imposed at the water free surface and a rigid wall condition at the bottom of the reservoir. The fluid-structural interaction condition at the interface between the two domains is the following

$$\begin{cases} -n\left(\frac{1}{\rho}\nabla p\right) = -n \cdot u_{tt} \\ F_A = p \cdot n \end{cases} \tag{3.40}$$

where n is the unit vector normal to the interface, F_A is the acoustic force on the structure and u_{tt} is the solid acceleration. The first equation transfers the structural acceleration to the fluid, while the second applies the fluid pressure load on the structure. Such a system of equations expresses a fully coupled problem, where the solid and the acoustic parts have to be solved simultaneously.

The original Westergaard formulation of semi-infinite domain is a reasonable practical assumption of a sufficiently long reservoir. This implies that none of the upstream waves is ever reflected back to the structure during the seismic event. It follows that a non-reflecting boundary condition is necessary in the upstream direction, namely the well-known Sommerfeld radiation condition. This can be accomplished in a finite element model through the Perfectly Matched Layer (PML) (Berenger 1994). It consists in a complex coordinate stretching function $f_r(\zeta)$ of an auxiliary absorbing domain as introduced in section 3.3, equation 3.32. In the present case an auxiliary PML domain has been added on the up-

stream side of the reservoir (figure 3.13) in the FE models.
The 2D model introduces the deformability of the structure and extends the excitation frequency to the whole range of interest for seismic study purposes, thus removing hypotheses 3 and 9. The results are displayed in the form of acoustic pressure magnitude, being a complex quantity in a frequency domain solution. In figure 3.15 the pressure modulus is plotted over the domain.

The hydrodynamic load distribution along the dam upstream face is represented by the

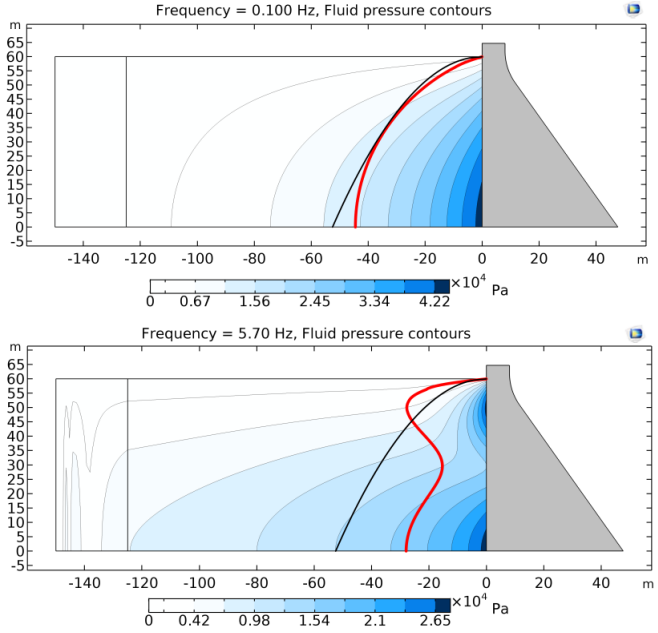


Fig. 3.15: 2D model response for 0.1 Hz (upper image) and 5.7 Hz (lower image). The linear distributed forces on the upstream face are represented in red, in comparison with Westergaard's solution (black curve).

red curves and compared with Westergaard's theory prediction (black curves).
By examining the pressure distributions, it can be clearly seen that Westergaard's solution is reproduced only for frequency values tending to zero (top image). In fact, in this case the pressure distribution on the upstream face follows the theoretical quarter-wave shape, in good agreement with Westergaard's parabolic assumption. Higher excitation frequency values (bottom image) lead to the development of higher mode shapes for the pressure distributions as in the lower image, where a near three-quarter-wave distribution is displayed.
The hydrodynamic load distribution on the dam face is computed also for the 3D model, thus removing hypothesis 8 of planar behaviour. In figure 3.16, in analogy with figure 3.15, the blue surface represents the equivalent distributed load on the upstream face of

the dam deduced by inertial added mass forces, while the red one is the acoustic pressure load obtained from the FE interacting model.

The 3D case displays a significant pressure variation along the cross valley direction,

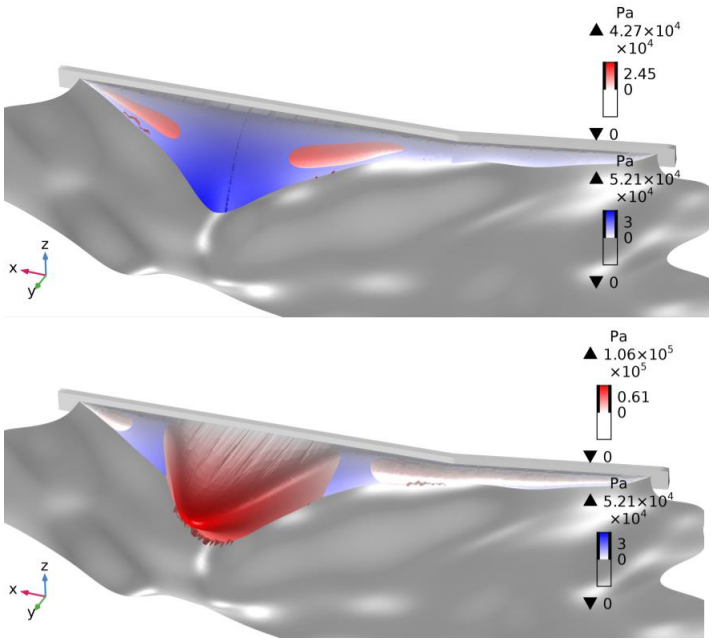


Fig. 3.16: 3D model response for 0.1 Hz (upper image) and 5.7 Hz (lower image): pressures relative to the FS interaction (red surface) are shown, in comparison to the equivalent added masses inertial load (blue surface).

which is impossible to be captured in a 2D modelling context. In addition, a significant phase difference between pressures on each monolith reduces the overall synchronism of hydrodynamic load and thus the amplification of its overall action on the structure.
Model response is evaluated by plotting the amplitude of hydrodynamic resultant forces for the tallest monolith. The plots in figure 3.17 and figure 3.18 show a noticeable difference between the 2D and 3D analyses; the main differences are the following.

- The frequency of the resonance peaks in the 3D case is higher than in the 2D case; a reasonable explanation could be found in the higher degree of constraint given by the adjacent, shorter monoliths, providing additional stiffness to the structure and the corresponding minor water depth.
- A slightly higher intensity of the hydrodynamic resultant force for the 2D model is observed.

- In the 2D case the response for near-zero frequencies of the added masses model tends to the same value as the FSI analysis, matching the theoretical solution. This does not occur in the 3D case, where the difference is still noticeable, being roughly 20% higher. This results may induce to believe that a direct added masses model transposition to three dimensional analysis be arbitrary and lead to errors.

- The added masses simplified model provides a reasonable estimation of the first resonant frequency, while it does not match the FSI analysis for superior modes, providing a much higher amplification.

- In the 2D case the rigid barrier peaks (blue curve) are still visible in the FSI analysis, meaning that specific reservoir dynamic properties are let unmodified by the structural coupling. In fact, the theoretical reservoir eigenfrequencies deduced from equation 3.38 are equal to 6.17 Hz and odd multiples in 2D, which are not affected by the structure. The phenomenon can be explained by noticing that at a certain distance from the dam, a local pressure variation at the upstream face does not influence the far-field response. In fact the latter is governed only by the reservoir depth itself, which is constant in 2D. In the 3D model such effect is less evident.

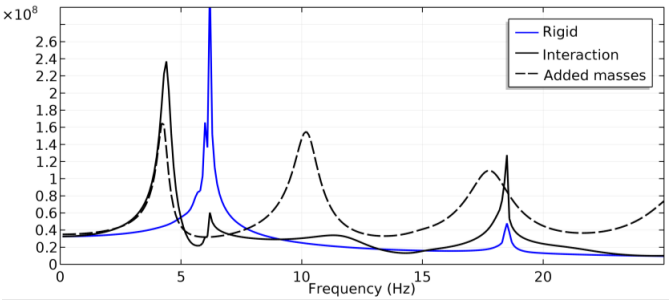


Fig. 3.17: Hydrodynamic force - 2D central monolith model.

In Figures figure 3.19 and figure 3.20 the 3D model pressure contours are plotted for different excitation frequencies. The figures clearly display the variability of pressure distribution along the cross valley direction, even for low excitation frequency values.
Since Westergaard's expression is strictly valid in 2D cases, its extension to 3D can be accomplished in different ways: both by fixing h in equation 3.39 as the overall reservoir maximum depth, or by continuously varying it along the cross valley direction. The first solution leads to major errors in estimating both peak frequencies and response magnitudes. The second solution, adopted in this case, is quite adequate in matching the FSI simulation

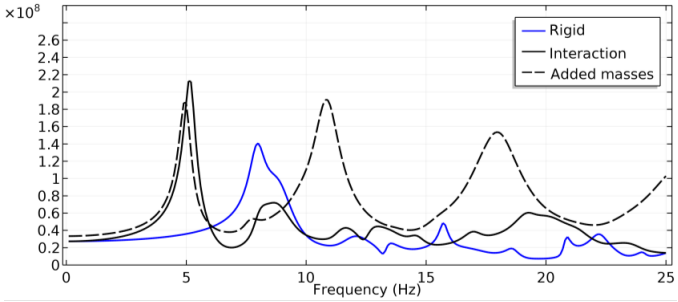


Fig. 3.18: Hydrodynamic force - 3D model, central monolith.

for low frequency ranges. Nevertheless, it remains inappropriate for higher modes, where pressure variations in the cross valley direction are significant due to excitation of superior longitudinal harmonics.

The same figures also plot the streamlines for acoustic intensity, defined as $I = p \cdot v$, with p is the fluid pressure and v the particle velocity. By comparing the figures for a low and a high value of excitation frequency, a remarkable difference is observed, and the energy flux is at least two orders of magnitude greater for high frequencies. This energy is radiated upstream of the reservoir and it is thus subtracted from the dam body. Such effect is equivalent to consider an additional damping, referred to as acoustic radiation damping. The acoustic intensity is integrated over the upstream model boundaries and

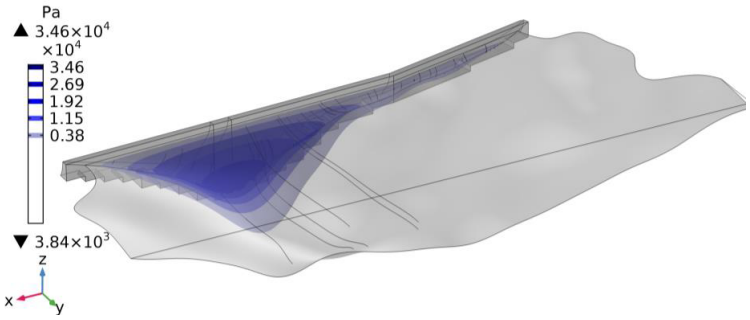


Fig. 3.19: Hydrodynamic pressure contours and acoustic intensity streamlines - 3D model - excitation of 0.1 Hz.

the total radiated power is plotted in figure 3.21 on a semi logarithmic scale for 2D and 3D models. It can be observed that the radiated energy for excitation frequencies below the first eigenfrequency of the reservoir is substantially negligible, compared to that for higher frequencies. The first reservoir eigenfrequency may thus be interpreted as a cut-off limit for energy radiation upstream of the reservoir. It is also shown that the 2D FSI models

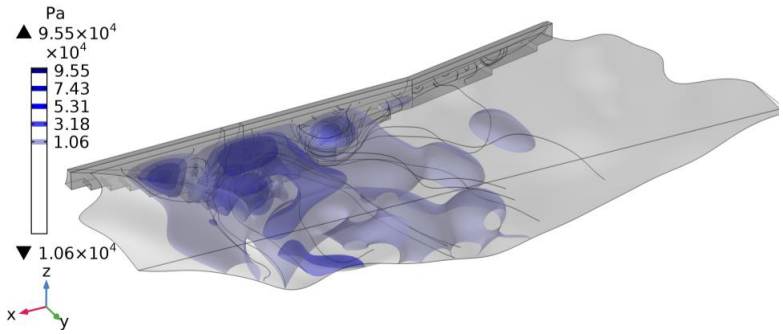


Fig. 3.20: Hydrodynamic pressure contours and acoustic intensity streamlines - 3D model - excitation of 20 Hz.

provides a dissipation throughout the whole range which is orders of magnitude lower than the rigid barrier case. This behaviour is not replicated in the 3D case, where the rigid and FSI values are comparable.

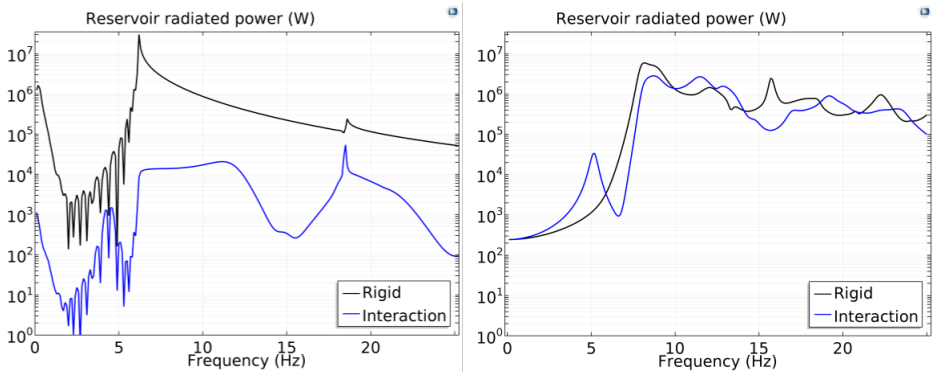


Fig. 3.21: Total reservoir radiated power (semi-logarithmic scale) for 2D (left) and 3D model (right).

The reasons for the different response between 2D and 3D models in terms of peak frequency and resultant force intensity has been investigated through a parametric study involving the shape of the canyon and the angle between lateral dam embankments. Such investigation is performed on two simplified 3D study models.

The first simple FE 3D model simulates the basin between the lateral embankments. It is a cylindrical sector of angle 2α with a deformable dam on the downstream side and an auxiliary PML domain on the upstream side. The mesh of the fluid domain has a maximum of 11346 tetrahedral quadratic Lagrange elements and 3158 nodes. The mesh of the solid domain has 477 tetrahedral second order Serendipity elements with 161 nodes. The cross-valley shape is rectangular and it remains unmodified as α varies. Figure 3.22 displays

the hydrodynamic force, varying the angle 2α in the range $(0, 75)$. One can observe that the hydrodynamic load decreases as α increases. In addition, the peak frequency value is practically unmodified with α. This is reasonable, since the peak frequency depends only on the reservoir depth.

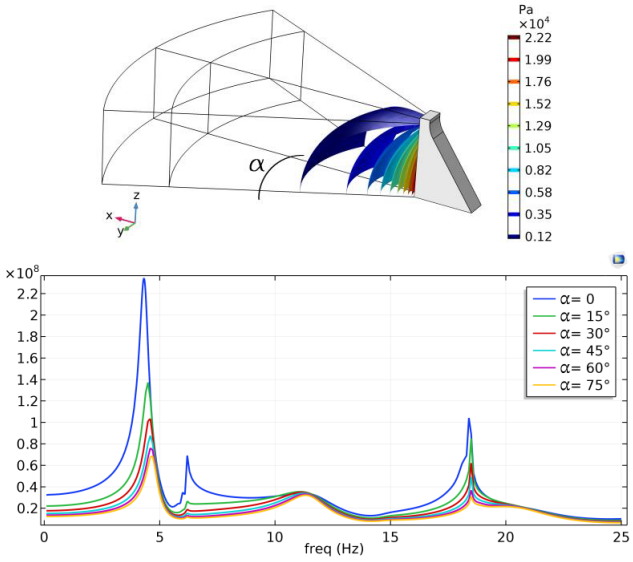


Fig. 3.22: The 3D simplified model with pressure contours for 0.1 Hz frequency (upper figure); Frequency response of the idealized 3D model for different values of α - semi-logarithmic scale (lower figure).

The second simple FE 3D model is a trapezoidal prism, whose lateral walls have slope β. On the downstream side there is a rigid barrier and on the upstream side an auxiliary PML domain is added. The mesh of the fluid domain has at most 2261 tetrahedral quadratic Lagrange elements and 1600 nodes. Figure 3.23 shows the effects of the variation of β on the total hydrodynamic force, while the barrier maintains constant area. The results demonstrate that even a small deviation of the side walls from the vertical ($\beta = 90$) produces a frequency shift to higher values. The peak values tend to infinite, because no sediment absorption is considered.

An attempt to obtain a quick estimation of reservoir natural frequency is performed by defining an equivalent depth $\hat{h}$. By averaging the reservoir depths at each monolith using the corresponding added mass as weight, one obtains

$$\hat{h} = \frac{\sum_i h_i \cdot m_{\mathrm{add}_i}}{m_{\mathrm{add}_i}} \tag{3.41}$$

where h_i is the generic monolith height and m_{add_i} is the added mass calculated for h_i and width b, as illustrated in figure 3.24. By doing so, the equivalent depth is found to be 50.75 m, which corresponds to a first frequency of 7.78 Hz, close to the FEM analysis result in the case of rigid barrier. This simple calculation provides an insight on the nature of the phenomenon. The combined effect of these two parameters varying within a realistic range confirms the differences between the 2D model and the more accurate 3D model.

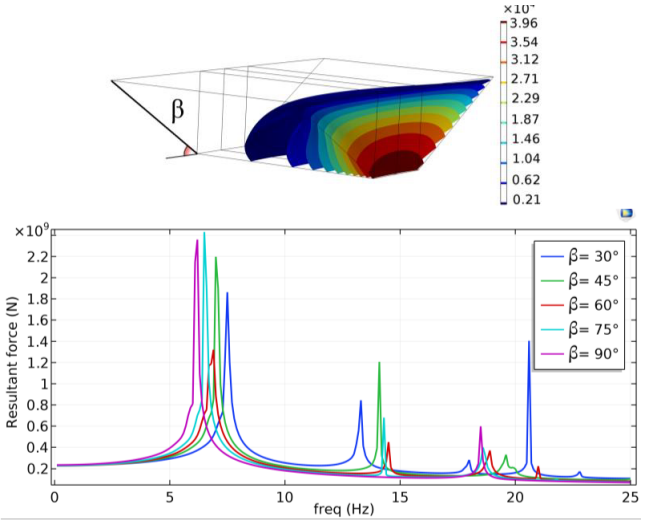


Fig. 3.23: Pressure contours for 0.1 Hz frequency (upper figure), Frequency response of the idealized 3D model for different values of β (linear scale) (lower figure).

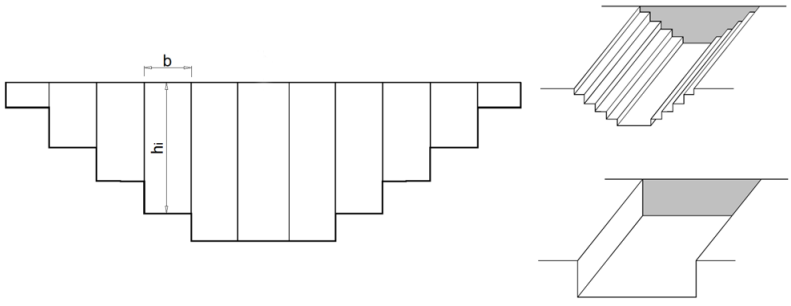


Fig. 3.24: Equivalent reservoir depth definition.

As a last step, the global effects of hydrodynamic loads on the structural response are evaluated in term of base shear force. Results are displayed in figure 3.25, for different modelling approaches, including the case of empty reservoir for reference purposes. It can be noted that, for very low frequencies, the hydrodynamic action increases the response

of the dam only by nearly 30% in both 2D and 3D, while at the resonance condition the shear is nearly doubled. By comparing the different approaches, it can be observed that:

- the first peak frequency of the model of interaction is very close to the one of added masses, while higher frequency peaks differ greatly;
- in the 2D case the rigid barrier model displays a resonance at a higher frequency;
- the 3D model does not display an explicit reservoir resonance peak in the case of rigid barrier.

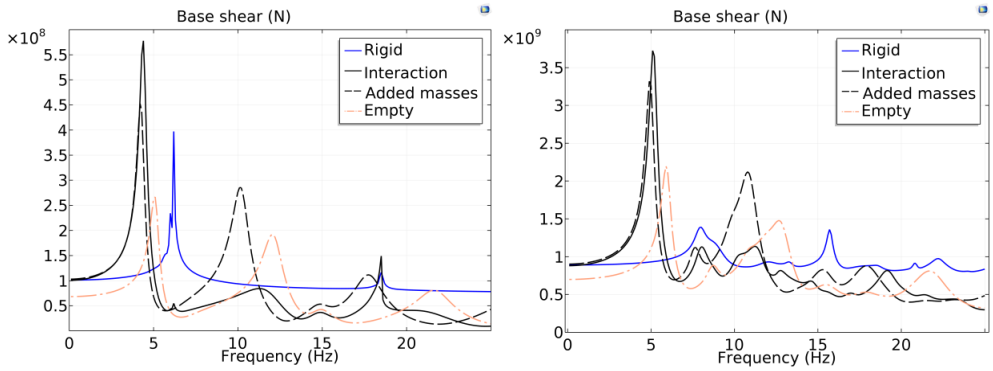


Fig. 3.25: Base shear for 2D model for central monolith (left) and 3D model for the entire dam (right).

3.4.4 Concluding remarks

The comparison between different approaches shows noticeable differences on the overall hydrodynamic response if the excitation frequency is higher than the first mode, both for a simple 2D model of the tallest monolith and the full dam 3D model of a case study. As a result of the subsequent parametric study, some differences between 2D and 3D have been observed. More specifically, in the 3D model the increase of the angle between lateral embankments induces a drastic reduction on the overall response amplitude, while the slope of reservoir walls modifies its frequency content. Such behaviour cannot be reproduced by a 2D plane model. Finally, it can be concluded that a realistic 3D geometry of the reservoir makes the fluid behaviour highly complex and, in 3D, the added mass model may produce even more different results from the full FSI modelling. Therefore, a high fidelity model can be only performed by modelling as precisely as possible the geometry of the system: dam, soil and basin.

3.5 Conclusion

In this Chapter, the main sources of uncertainty involved in the seismic analysis of concrete gravity dams have been introduced and analysed. Some basic assumptions are needed in order to define a deterministic model of the dam; these are mainly related to the constitutive materials models, the Fluid-Structure Interaction and the Soil-Structure Interaction. Once a deterministic model is defined other epistemic uncertainties mainly related to the model parameters arises. One of the aims of this research work is the development of procedures able to reduce the epistemic uncertainties by using all available information about the dam. Particularly interesting is the use of static and dynamic measurements directly recorded on the dam body, because they allows the calibration of predictive models which can be integrated in the monitoring system.
The development of probabilistic procedure requires the use of statistical tools, as for instance the definition of probabilistic models, Bayesian inference, proxy models. Therefore, Chapter 4 introduces all of these tools, providing a review of theoretical background.
For the sake of simplicity, every Chapter which introduces one of the three procedures, developed to reduce the epistemic uncertainties, discusses also the nature of the reference measures used in the updating process and the development of the probabilistic model.
The calibrated predictive models, obtained with the proposed procedures, are indicated as twin model of the dam, because they reproduce the structural behaviour with respect to a particular Quantity of Interest (QI). The calibrated dam twin model are used to predict the structural behaviour during its regular use or during seismic events.

Chapter 4

Statistical analysis

4.1 Introduction

The development of SHM framework suitable for concrete dams, which allows both reducing the uncertainties and controlling the structural behaviour, requires the use of statistical tools. In particular, the procedures proposed in this research work are set up in a Bayesian framework, which is one of the topics introduced in this Chapter. However, the first topic introduced in this work is the development of probabilistic models whose definition is fundamental in order to correctly develop an updating procedure.
Moreover, since probabilistic procedures are computationally expensive due to the large number of solutions of the deterministic model required in the proposed procedures the computational burden is highly reduced by using meta models which approximate the solution of the FE Analysis (FEA). The general Polynomial Chaos Expansion (gPCE) technique is used, so it is described in this Chapter.
In the last part of this thesis a procedure to design a SHM system able to acquire information both to detect damage and update the strength parameters of the materials is introduced. This last procedure is set up in a Bayesian framework by using the Optimal Bayesian Experimental Design, introduced in the last section of this Chapter.

Therefore the Chapter is organized as follows:
Section 4.2 describes the fundamental concepts of probabilistic models.
Section 4.3 introduces the concept of the Bayesian inference.
Section 4.4 describes the general polynomial chaos expansion technique, which has been widely used in this work both to solve the forward problem and to build meta model of the FE analysis outputs.
Section 4.5 introduces the main idea behind the Optimal Bayesian Experimental Design, which has been used to develop the procedure for the devices layout optimization of the structural monitoring system.

4.2 Probabilistic model

4.2.1 Introduction

In structural engineering deterministic biased models are usually used to predict the behaviour of a building or that of its structural components. Together with these models some "calibrated" coefficients are used, in order to obtain a safe estimation of the structural behaviour, but they do not explicitly consider the uncertainties involved in it. Deterministic models like the one previously described are diffusely used to design new buildings and for the assessment of existing structures. However, in structural reliability unbiased predictive models, which directly consider the main uncertainty sources, are needed. Since, the aim of this research work is the development of procedures to calibrate the parameters of concrete gravity dams models, probabilistic models assume great importance.
In this section, the probabilistic models used in this work are described, rather than developing new probabilistic models existing ones have been used extending their use to real applications. In particular, following Gardoni et al. (2002a), corrected additive models have been used and described in this section. Differently from others probabilistic models, in the considered ones correction terms which allows one to properly account for the inherent bias and uncertainties are added. In its general version a correction term is a combination of a set of basis explanatory functions, which allows modelling the role of each uncertainty source. In Chapters 5 and 6 the probabilistic models introduced in this section are specialized for the dam case.

4.2.2 Probabilistic additive model

In the context of this work a *"model"* is a mathematical expression which relates one or more QI, i.e. deformations, displacements structural components capacities and so on, to a set of measurable variables $\mathbf{x}$ related to a particular class of deterministic model $\mathcal{C}$, and other random parameters collected in $\boldsymbol{\Theta}$. The main purpose of the probabilistic model is to provide a means for predicting the quantities of interest for given deterministic, or random values, of the variables collected in $\mathbf{x}$ and $\boldsymbol{\Theta}$. The model is called uni-variate when only one QI is predicted, or multi-variate if more than one QI are considered.
For the sake of simplicity the uni-variate model is first introduced, then the idea is generalized to obtain the multi-variate one. The uni-variate model can be expressed in its general form as

$$C = C\left(\mathbf{x}, \boldsymbol{\Theta}\right), \tag{4.1}$$

where C is the selected QI while the function $C\left(\mathbf{x}, \boldsymbol{\Theta}\right)$ can have a general form, involving integrals, derivative and so on, but it should be derived from mechanical principles. As mentioned before, following Gardoni et al. (2002a) the adopted additive probabilistic model is corrected by a correction term $\gamma\left(\mathbf{x}, \boldsymbol{\theta}_\gamma\right)$ which is a function of some explanatory functions, explained before, and combination coefficients collected in $\boldsymbol{\theta}_\gamma$. Therefore, equation 4.1 can be written as

$$C\left(\mathbf{x}, \boldsymbol{\Theta}\right) = \hat{c}\left(\mathbf{x}, \boldsymbol{\theta}_\mathrm{m}\right) + \gamma\left(\mathbf{x}, \boldsymbol{\theta}_\gamma\right) + \sigma\epsilon, \tag{4.2}$$

in this context $\boldsymbol{\Theta} = (\boldsymbol{\theta}_\mathrm{m}, \boldsymbol{\theta}_\gamma, \sigma)$, where σ is the global error standard deviation of the probabilistic model, $\hat{c}\left(\mathbf{x}, \boldsymbol{\theta}_\mathrm{m}\right)$ is the selected deterministic model, $\boldsymbol{\theta}_\mathrm{m}$ is the vector of unknown model parameters and ϵ is a normal random variable with zero mean and unite variance. Therefore, for given $\mathbf{x}$, $\boldsymbol{\theta}_\mathrm{m}$, $\boldsymbol{\theta}_\gamma$ and σ, the variance of the probabilistic model is $\mathrm{Var}\left[C\left(\mathbf{x}, \boldsymbol{\Theta}\right)\right] = \sigma^2$.
The additive model correction form is valid under the following assumptions:

- *Homoskedasticity assumption*: the model standard deviation σ is independent of $\mathbf{x}$,
- *Normality assumption*: the model error is normally distributed,
- *Additive assumption*: the error can be added to the probabilistic model.

Since these three assumptions are usually not satisfied, a variance stabilizing transformation is used to approximately satisfy them within the range of data. Therefore, given the parameters vector $\boldsymbol{\lambda}$, it can be used to define the variance stabilizing transformation within

a family of possible transformations. For this purpose, the dual power transformation functions, proposed by Yang (2006), has been used in this work,

$$f_\lambda(\cdot) = \begin{cases} \frac{(\cdot)^\lambda - (\cdot)^{-\lambda}}{2\lambda}, & \text{if } \lambda \neq 0. \\ \ln(\cdot), & \text{if } \lambda = 0. \end{cases} \tag{4.3}$$

The inverse functions are

$$f_\lambda^{-1}(\cdot) = \begin{cases} \left(\lambda(\cdot) + \sqrt{1 + \lambda^2(\cdot)^2}\right)^{\frac{1}{\lambda}}, & \text{if } \lambda \neq 0. \\ e^{(\cdot)}, & \text{if } \lambda = 0. \end{cases} \tag{4.4}$$

Such functions show properties similar to the well-known Box and Cox (1964) power transformation, but without the long-standing truncation problem. In this case, $\lambda = \boldsymbol{\lambda}$.
The correction term $\gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma)$ corrects the bias of the deterministic model $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_\mathrm{m})$, but since the deterministic model is usually approximated, the true form of $\gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma)$ is unknown. Therefore, the sources of bias can be explored by using a set of p explanatory basis functions $\mathrm{h}_i(\mathbf{x})$, with $i = 1, ..., p$, whose combination express the bias correction term,

$$\gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma) = \sum_{i=1}^{p} \theta_i \mathrm{h}_i(\mathbf{x}). \tag{4.5}$$

Moreover, by studying the posterior distributions of the coefficients θ_i, the most significant explanatory functions can be identified.
Finally, the previous consideration can be extended in order to define the multi-variate form of the uni-variate probabilistic model. Let is considered a q-dimensional multi-variate additive corrected probabilistic model, to predict q quantity of interest, equation 4.2 can be extended as

$$C_k(\mathbf{x}, \boldsymbol{\theta}_\mathrm{m}, \boldsymbol{\theta}_{\gamma,k}, \boldsymbol{\Sigma}) = \hat{c}_k(\mathbf{x}, \boldsymbol{\theta}_\mathrm{m}) + \gamma_k(\mathbf{x}, \boldsymbol{\theta}_{\gamma,k}) + \sigma_k \epsilon_k, \qquad k = 1, ..., q. \tag{4.6}$$

and the correction term becomes

$$\gamma_k(\mathbf{x}, \boldsymbol{\theta}_{\gamma,k}) = \sum_{i=1}^{p_k} \theta_{ki} \mathrm{h}_{ki}(\mathbf{x}), \qquad k = 1, ..., q. \tag{4.7}$$

With the exception of $\boldsymbol{\Sigma}$, all terms in the above equations have the same meanings as in the case of the uni-variate probabilistic model. In this context, $\boldsymbol{\Sigma}$ is the covariance matrix

of the variables $\sigma_k \epsilon_k$, and its (k, l) element is $\rho_{kl}\sigma_k\sigma_l$ where ρ_{kl} is the correlation coefficient between ϵ_k and ϵ_l. Considering the symmetry of $\mathbf{\Sigma}$, it includes q unknown variances σ_k^2, with $k = 1, ..., q$, and $q\,(q-1)\,/2$ unknown correlation coefficients ρ_{kl}, with $k = 1, ..., q-1$ and $l = k+1, ..., q$. Therefore, the number of the parameters, objects of the updating procedure, considerably increase.

4.2.3 Uncertainty in the probabilistic model

The nature of uncertainties, and then how to treat them, is an open issue among scientist and philosophers since statistic has become a fundamental part of modern science (Lindley 2000). However, the scientific community agrees that the nature and the character of the uncertainty depends on the context in which they are analysed.
Engineering problems are solved in the context of model universe, and in this field the uncertainties are usually categorized as aleatory or epistemic (Der Kiureghian and Ditlevsen 2009). The term aleatory derives from the Latin *alea* which means the rolling of the dice, the aleatory uncertainty is one that is presumed to be the intrinsic randomness of a phenomenon. The word epistemic derives from the Greek $\epsilon\pi\iota\sigma\tau\eta\mu\eta$ (episteme), which means knowledge. Thus an epistemic uncertainty is one that is presumed as being caused by lack of knowledge. Their general definition expresses how their concept in practice applications is vague. However, in the context of physical modelling, these two definitions assume a meaning and greater concreteness. Pragmatically, in this context, epistemic uncertainties can be reduced once new information are gathered, while aleatory ones cannot be reduced. This aspect involves the subjectivity of the analyst, who will decide which of the uncertainties of his model are random and which epistemic, depending on the problem, his own experience, and the available information. An interesting example of this idea is reported in Faber (2005). Thinking about the concrete strength, if an existing building is considered, it can be defined as epistemic uncertainty since new specimens can be tested, improving the knowledge about them, thus reducing the uncertainty. On the other hand, the uncertainty in the concrete strength should be categorized as aleatory, if a future building is considered. In the words of the author, the character of the aleatory uncertainty "transforms" into epistemic uncertainty as the building is realized.
In the context of the present probabilistic model, two sets of parameters have been defined: $\mathbf{x}$ and $\mathbf{\Theta} = (\boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\theta}_{\gamma}, \sigma)$. The set $\mathbf{x}$ represents the known parameters, while $\boldsymbol{\theta}_{\mathrm{m}}$ is the set of the unknown parameters of the deterministic model, which are treated as random. Once new observations are available the knowledge about $\boldsymbol{\theta}_{\mathrm{m}}$ is updated these parameters are considered epistemic uncertainties. The other parameters collected in $\mathbf{\Theta}$ are added to the

deterministic model in order to improve the fitting between model outputs and observations, so the nature of these parameters is epistemic for definition and they can be updated as well.
Finally some of the most common uncertainties related to structural engineering probabilistic models are described below.

- *Model uncertainty*: as mentioned before mathematical models are used to described a physical phenomenon of the real world, but models are approximated. For instance, when a structure is analysed, a particular constitutive model is chosen in order to describe the material behaviour, or a seismic input which derives from a study where assumptions and approximations are involved. Therefore, some modelling hypothesis have to be assumed and this leads to an approximation. Since every mathematical model is an idealization of the real world, then it is uncertain and affected by an implicit error. The only way to assess and reduce this uncertainty is by comparing the model predictions with real world observations. Considering the notation used in this work, this uncertainty arise in the approximation of the deterministic model $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$, and two main sources can be recognized: error in the model form, e.g. linear approximation of the deterministic model and missing variables, or $\mathbf{x}$ and $\boldsymbol{\theta}_{\mathrm{m}}$ contain only a subset of the variables which influences the QI. In this context, the error in the model can be partially compensated by adding the correction term $\gamma(\mathbf{x}, \boldsymbol{\theta}_{\gamma})$, while the error term $\sigma\epsilon$ represents the influence of the missing variables, as well as the remaining model form error. Moreover, the influence of missing variables is completely random, that component of ϵ can be defined aleatory, while the component representing the inexact model form has an epistemic nature. As discussed before, assuming an aleatory or epistemic nature of ϵ is a difficult task.

- *Measurements error*: uncertainty arises from errors inherent in laboratory or field measurements. This error is related to both the observation of the QI C and the model parameters $\mathbf{x}$ which cannot be directly measured. This error can be modelled by adding error terms to the assumed true QI $C_i = \hat{C}_i + e_{C_i}$ and to the assumed true parameters set $\mathbf{x}_i = \hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}_i}$, of the i-th observation. In particular, $\hat{C}_i$ is the i-th measured QI and $\hat{\mathbf{x}}_i$ is the i-th measured set of model parameters, while e_{C_i} and $\mathbf{e}_{\mathbf{x}_i}$ are the respective measurements errors. These errors can be reduced by using more efficient devices, so their related uncertainty is epistemic in nature. Moreover, the mean values of the errors represent the biases in the measurements, while their variances represent the uncertainties inherent the measurements. Usually,

in engineering problems these two errors are considered statistically independent and they are modelled as normally distributed random variables.

- *Statistical uncertainty*: this uncertainty arises from the sparseness of collected data, and increasing the number of observations it can be reduced, so it can be indicated as epistemic uncertainty. This error regards not only the values of the statistics, which define the assumed distribution, but also the choice of the distribution itself. This uncertainty increases by moving toward the tails of the distributions.

Concluding this section, the epistemic part of the presented uncertainty sources can be reduced once new observations are available, and it can be done by using the Bayesian inference, described in the section 4.3.

4.3 Bayesian inference

4.3.1 Introduction

In the field of mathematical modelling the determination of parameters values which allows one to achieve the best fitting between output and recorded behaviour of the structure is called inverse problem. This name is in contrast to the forward problem, whose main aim is the prediction of the model output given certain inputs parameters. Since the available information are usually not sufficient to uniquely determine the model parameters, the inverse problem is an ill-posed problem. The inverse problem can be solved in a deterministic setting by minimizing an objective function. Being an optimization problem it requires a regularization procedure (Vogel 1987). In a probabilist setting the inverse problem becomes a well-posed problem by adding new information regarding the prior knowledge about the model parameters.
In the context of data analysis based on probabilistic models, three principal approaches are available: frequentist, Bayesian and likelihood.
The frequentist approach is based on the frequentis interpretation of the probability, namely the result of a particular experiment can be seen as one of an infinite sequence of possible repetitions of the same experiment. Therefore, the probability distribution of the observed data conditional to the unknown parameters is obtained by sampling directly from a particular model, the likelihood.
In the Bayesian approach the sampling technique and the prior distributions of the model parameters are needed, the posterior knowledge about the parameters is a combination of the prior knowledge and the likelihood function, which contains the new information about

the system.
Finally, in the likelihood, or Fisherian, approach the inference is based only on the likelihood function, which is sampled as in the Bayesian approach but without needing the prior distributions.
Differently to the frequentis approach in the Bayesian and Fisherian ones the experiment is implicitly assumed to be unrepeatable.
In this work the Bayesian approach is widely used to updated the prior knowledge on the model parameters by using the measurements recorded on the structure. In this part of the thesis the Bayesian approach is presented.

4.3.2 Synopsis of Bayesian inference

The well known Bayes' theorem allows updating the prior knowledge about the system prior distributions once new data are collected and incorporated within the likelihood function, getting a new state of knowledge described through the posterior distributions. The Bayes' theorem is usually defined in terms of conditional probabilities and in text books is generally reported in the Laplace form. Considering the probability space defined by the triplet $(\Omega, \mathfrak{F}, \mathbb{P})$, where Ω is the space of all events, $\mathfrak{F}$ is the σ-algebra and $\mathbb{P}$ the probability measure. Therefore, by assuming $\boldsymbol{\Theta}$ as the vector of the unknown model parameters and $\mathbb{P}(\boldsymbol{\Theta})$ as the prior knowledge about them, once new observation $\mathbf{y} = (y_1, ..., y_k)^T$ are available the prior knowledge can be updated getting the posterior one

$$\mathbb{P}(\boldsymbol{\Theta}|\mathbf{y}) = \frac{\mathbb{P}(\mathbf{y}|\boldsymbol{\Theta})}{\mathbb{P}(\mathbf{y})}\mathbb{P}(\boldsymbol{\Theta}) \quad \text{if } \mathbb{P}(\mathbf{y}) > 0, \tag{4.8}$$

where $\mathbb{P}(\mathbf{y}|\boldsymbol{\Theta})$ is the likelihood and $\mathbb{P}(\mathbf{y})$ is the evidence. As mentioned before equation 4.8 is usually defined in terms of conditional probabilities, by defining $p(\boldsymbol{\Theta})$ as prior distribution, $p(\boldsymbol{\Theta}|\mathbf{y})$ as posterior, $p(\mathbf{y}|\boldsymbol{\Theta})$ as likelihood and $p(\mathbf{y})$ as evidence, the Bayes' theorem can be written as (Box and Tiao 1992)

$$p(\boldsymbol{\Theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\Theta})\, p(\boldsymbol{\Theta})}{p(\mathbf{y})} \quad \text{if } p(\mathbf{y}) > 0. \tag{4.9}$$

Moreover, following Fisher (1922) equation 4.9 can be rewritten introducing the likelihood function $L(\boldsymbol{\Theta}|\mathbf{y})$ as

$$p(\boldsymbol{\Theta}|\mathbf{y}) = \kappa L(\boldsymbol{\Theta}|\mathbf{y})\, p(\boldsymbol{\Theta}), \tag{4.10}$$

where $\kappa = \left[\int L(\boldsymbol{\Theta}|\mathbf{y})\, p(\boldsymbol{\Theta})\, d\boldsymbol{\Theta}\right]^{-1}$ is the normalizing factor which ensures that the posterior distribution integrates or sums to one. Moreover, every time new observations become available the state of knowledge can be updated. Assuming that $\mathbf{y}_1$ is the first set of available information, the posterior distribution can be calculated as,

$$p(\boldsymbol{\Theta}|\mathbf{y}_1) \propto p(\boldsymbol{\Theta})\, L(\boldsymbol{\Theta}|\mathbf{y}_1) \tag{4.11}$$

Once new observation $\mathbf{y}_2$ are recorded, they can be introduced in the updating procedure,

$$\begin{aligned} p(\boldsymbol{\Theta}|\mathbf{y}_1, \mathbf{y}_2) &\propto p(\boldsymbol{\Theta})\, L(\boldsymbol{\Theta}|\mathbf{y}_1)\, L(\boldsymbol{\Theta}|\mathbf{y}_2) \\ p(\boldsymbol{\Theta}|\mathbf{y}_1, \mathbf{y}_2) &\propto p(\boldsymbol{\Theta}|\mathbf{y}_1)\, L(\boldsymbol{\Theta}|\mathbf{y}_2) \end{aligned} \tag{4.12}$$

The same updating process can be repeated for m sets of independent observations, with m as big as one wants

$$p(\boldsymbol{\Theta}|\mathbf{y}_1, ..., \mathbf{y}_m) \propto p(\boldsymbol{\Theta}|\mathbf{y}_1, ..., \mathbf{y}_{m-1})\, L(\boldsymbol{\Theta}|\mathbf{y}_m)\,. \tag{4.13}$$

Equation 4.13 can be seen as a learning process where the prior knowledge about $\boldsymbol{\Theta}$ is updated every time that new observations are available.
Turning back to equation 4.8, in case of continuous random variables could happen that the observation set $\mathbf{y}$ has a $p(\mathbf{y}) = 0$ leading to an indeterminate form, namely $0/0$, of the Bayes' rule, so some form of limiting procedures is needed. Several different procedures for doing that are available, as for example the so-called *Borel-Kolmogorov* paradox.
Following Matthies et al. (2016), in the special case where $\mathbf{y}$ and $\boldsymbol{\Theta}$ have a joint pdf $p(\mathbf{y}, \boldsymbol{\Theta})$, equation 4.8 can be reformulated in terms of conditional expectation overcoming the indeterminate form. In the context of this work, $\mathbf{y}$ is essentially a function of $\boldsymbol{\Theta}$, through ϵ, so a conditional pdf between random variables and the observation exists. The conditional expectation is defined as

$$\mathbb{E}(\Psi|\mathbf{y}) := \int_{\Omega} \Psi(\boldsymbol{\Theta})\, p(\boldsymbol{\Theta}|\mathbf{y})\,, \tag{4.14}$$

where Ψ is any measurable function of $\boldsymbol{\Theta}$. Another way to define conditional expectation is through orthogonal projection. Considering the sub-σ-algebra $\mathfrak{B} \subset \mathfrak{F}$ and assuming that random variables have finite variance, so $\mathcal{S}_{\mathfrak{B}} := \mathrm{L}_2(\Omega, \mathfrak{B}, \mathbb{P})$ is a closed sub-space of $\mathcal{S} := \mathrm{L}_2(\Omega, \mathfrak{F}, \mathbb{P})$, and hence a continuos orthogonal projection $\mathcal{P}_{\mathfrak{B}} : \mathcal{S} \to \mathcal{S}_{\mathfrak{B}}$ (Bobrowski 2005). The conditional expectation of a RV $r \in \mathcal{S}$ with respect to $\mathfrak{B}$ is an orthogonal

projection

$$\mathbb{E}\left[r|\mathfrak{B}\right] := \mathcal{P}_{\mathfrak{B}}\left(r\right) \in \mathcal{S}_{\mathfrak{B}}. \tag{4.15}$$

As reported in Bobrowski (2005), the unconditional expectation $\mathbb{E}\left[\cdot\right]$ can be seen as a conditional expectation with respect to the minimal σ-algebra $\mathfrak{B} = \{\emptyset, \omega\}$. Since conditional expectation is an orthogonal projection it minimises the squared error

$$\mathbb{E}\left[\left|r - \mathbb{E}\left[r|\mathfrak{B}\right]\right|^2\right] = \min\left\{\mathbb{E}\left[\left|r - \tilde{r}\right|^2\right] \quad : \quad \tilde{r} \in \mathcal{S}_{\mathfrak{B}}\right\}, \tag{4.16}$$

from which the orthogonality relation can be derived

$$\forall \tilde{r} \in \mathcal{S}_{\mathfrak{B}} : \mathbb{E}\left[\tilde{r}\left(r - \mathbb{E}\left[r|\mathfrak{B}\right]\right)\right] = 0, \tag{4.17}$$

and one has a form of Pythagora's theorem

$$\mathbb{E}\left[\left|r\right|^2\right] = \mathbb{E}\left[\left|r - \mathbb{E}\left[r|\mathfrak{B}\right]\right|^2\right] + \mathbb{E}\left[\left|\mathbb{E}\left[r|\mathfrak{B}\right]\right|^2\right] \tag{4.18}$$

Therefore, the conditional expectation is a form of minimum mean square error (MMSE) estimator. Given the conditional expectation the conditional probability is completely defined, e.g. for $A \subset \Omega, A \in \mathfrak{B}$ by

$$\mathbb{P}\left(A|\mathfrak{B}\right) := \mathbb{E}\left[\chi_A|\mathfrak{B}\right], \tag{4.19}$$

where χ_A in a unit random variable if $\omega \in A$ and vanishes otherwise, so having $\mathbb{E}\left[\cdot|\mathfrak{B}\right]$ allows knowing everything about the conditional probability, then the conditional or posterior density is not needed, overcoming the previously described paradox.

4.3.3 Prior distributions

Historically, the selection of the prior distribution which really reflects the state of prior knowledge was one of the biggest issue related to the Bayesian inference. By citing Box and Tiao (1992), in scientific problems we would like the data *"to speak for themselves"*, so scientists should avoid the use of prior which distorts or dominates *"what data are trying to say"*. This aspect gave rise to several disputes, and the Bayes approach was often criticized as subjective and too sensitive to the choice of the prior. For these reasons the necessity to construct a right prior distribution, when little is known a priori, became a fundamental aspect, so much that Bayes himself tried to give some indications about how to face this problem.

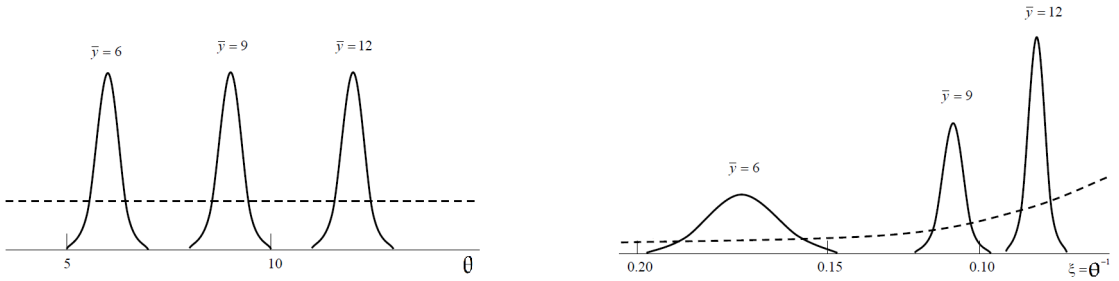
The Bayes' idea was to use uniform distribution as prior in the case of lack of knowledge, this is usually referred to as "Bayes's postulate". Anyway, the inconsistency of this idea can be easily showed by thinking about the parametrization of the random variable. Indeed, if a locally uniform distribution of $\boldsymbol{\Theta}$ is used, the distribution of a transformation of $\boldsymbol{\Theta}$, e.g., $\boldsymbol{\Theta}^{-1}$ and $\ln(\boldsymbol{\Theta})$ would not be locally uniform.
Moreover, if a moderately large number of new observations are available, the influence of a drastic change in the prior leads to small variation in the posterior. The procedures developed in this work are based on the possibility to record an increasing number of data over time, leading to a small influence of the prior. Finally, as remarked by Box and Tiao (1992) we can never be in a complete state of ignorance and the statement *"knowing little a priori"* has only a relative meaning if compared with the information content of an experiment.
These reasons highlight the importance of the prior distribution, in particular when little is known about some parameters, then non-informative prior distributions are explored in this section since in this work they have been widely used. An exhaustive description of the non-informative prior distribution can be found in Box and Tiao (1992), in particular the authors start from very simple cases to complex ones in order to highlight the influence of a prior on the final result.
Particularly interesting are the cases where $\mathbf{y} = (y_1, ..., y_n)^T$ is assumed to be sampled from a normal distribution $\mathcal{N} \sim \left(\theta, \sigma^2\right)$ with in the first case only θ unknown and in the second one only the standard deviation σ considered unknown.
In the former case, the authors show a fundamental concept for which in the case that only the mean value is unknown, the new observation should change only the position of the likelihood. Therefore, if we are in a state of prior ignorance the corresponding distribution should give no information about that position, it means that a non informative prior must be locally uniform, as in figure 4.1a. However, if the object of the statistical analysis is a transformation of θ, e.g. $\xi = \theta^{-1}$, for which the new observation $\mathbf{y}$ changes not only the position of the likelihood but also the shape. As shown in figure 4.1b, the corresponding prior for ξ is not uniform. Therefore, it is possible to assume that θ can be expressed in terms of a metric $\phi(\theta)$ so that the corresponding likelihood is *data translated*. This means that the likelihood for $\phi(\theta)$ is completely determined a priori, except for its location, which depends on $\mathbf{y}$, and the prior distribution is non informative if $\phi(\theta)$ is locally uniform. In the case where $\phi(\theta) = \theta$ a non informative prior is locally uniform in θ itself if

$$p(\theta|\sigma) \propto \text{constant}, \tag{4.20}$$

it is shown in figure 4.2a. Since,

$$p(\xi|\sigma) = p(\theta|\sigma)\,|\frac{d\theta}{d\xi}| = p(\theta|\sigma)\,\theta^2 \propto \xi^{-2} \tag{4.21}$$

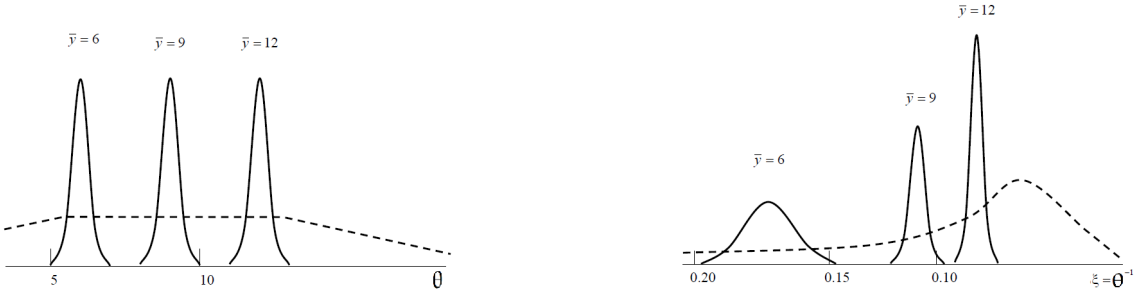
the corresponding non informative prior for ξ is not uniform but proportional to θ^2, as sown in figure 4.2b. In general, the non informative prior should be locally uniform in $\phi(\theta)$ and the corresponding non informative prior in θ is proportional to $|d\phi/d\theta|$, assuming a one to one transformation. In this regard is very important to understand which is the metric for which the prior must be uniform.



(a) The normal mean θ.

(b) Reciprocal of the normal mean $\xi = \theta^{-1}$.

Fig. 4.1: Non informative prior distributions (dotted line) and standardised likelihood curves (solid lines) for the normal mean θ and $\xi = \theta^{-1}$.



(a) The normal mean θ.

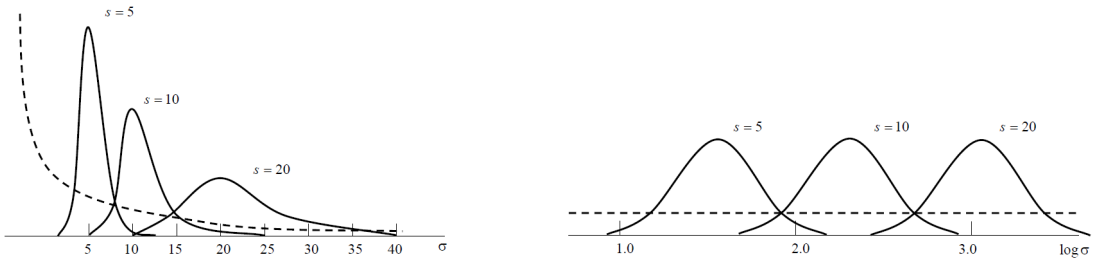
(b) Reciprocal of the normal mean $\xi = \theta^{-1}$.

Fig. 4.2: Non informative prior distributions (dotted line) and standardised likelihood curves (solid lines) for the normal mean θ and $\xi = \theta^{-1}$ seen over a wider range of parameter values.

This principle can be extended to the latter case, in which θ is supposed to be known while σ is unknown, in this case the likelihood in the original metric σ are clearly not data

translated as shown in figure 4.3a. However, if $\ln(\sigma)$ is considered as metric, the likelihood are data translated, as shown in figure 4.3b, so a non informative prior should be locally uniform with regard to $\ln(\sigma)$. Following the previous consideration the non informative prior distribution with respect to σ can be defined as,

$$p(\sigma|\theta) \propto \left|\frac{d\log(\sigma)}{d\sigma}\right| = \sigma^{-1}. \tag{4.22}$$



(a) Normal standard deviation σ.

(b) log of normal standard deviation, $\log(\sigma)$.

Fig. 4.3: Non informative prior distributions (dotted line) and standardised likelihood curves (solid lines) for the normal standard deviation σ and $\log(\sigma)$.

Therefore, a general rule can be derived from the above discussion: given a one to one transformation between θ and $\phi(\theta)$ a prior distribution of θ locally proportional to $|d\phi/d\theta|$ is non informative for θ if the likelihood curve is data translated in terms of ϕ. This means that only the location of the likelihood $L(\phi|\mathbf{y})$ changes with the new observations. A data translated likelihood can be mathematically expressed by the following relation,

$$L(\theta|\mathbf{y}) = g[\phi(\theta) - f(\mathbf{y})], \tag{4.23}$$

where $g(\cdot)$ is a known function independent of $\mathbf{y}$, while $f(\mathbf{y})$ is a function of the new observations.

In the previous two examples, related to the normal distribution $\mathcal{N}(\theta, \sigma^2)$, the definition of the metric ϕ was simple ($\phi(\theta) = \theta$ in the former case, and $\phi(\sigma) = \ln(\sigma)$ in the latter one). Usually, in real cases, the definition of a ϕ that allows the likelihood to be expressed as in equation 4.23 is not so easy, or even possible. In these cases an approximated data-translated likelihood can be built. Assuming that $\mathbf{y}$ is sampled by a distribution $p(\mathbf{y}|\theta)$, following Johnson (1967) under some regularity conditions on $p(\mathbf{y}|\theta)$, if the number of

samples is sufficiently large, the likelihood is function of θ, it is approximately normal, and it remains normal under approximate one to one transformation of θ. Moreover, the logarithm of the likelihood is approximately quadratic

$$l(\theta|\mathbf{y}) = \ln[L(\theta|\mathbf{y})] = \ln\prod_{i=1}^{n} p(y_i|\theta) \approx l\left(\hat{\theta}|\mathbf{y}\right) - \frac{n}{2}\left(\theta-\hat{\theta}\right)^2\left(-\frac{1}{n}\frac{\partial^2 l}{\partial\theta^2}\right)_{\hat{\theta}}, \tag{4.24}$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ, so that

$$\left(\frac{\partial l}{\partial\theta}\right)_{\hat{\theta}} = 0. \tag{4.25}$$

Moreover, the function

$$J\left(\hat{\theta}\right) = \left(-\frac{1}{n}\frac{\partial^2 l}{\partial\theta^2}\right)_{\hat{\theta}} \tag{4.26}$$

is positive for a given n and it is only a function of $\hat{\theta}$. Moreover, the logarithm of a normal density function $p(y)$ assumes the following form,

$$\ln p(y) \propto -\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2} \tag{4.27}$$

it is completely described by the location parameter μ and σ. By comparing equation 4.24 and equation 4.27 the standard deviation of the likelihood is approximately $n^{-1/2}J^{-1/2}\left(\hat{\theta}\right)$. Moreover, assuming that $\phi(\theta)$ is one to one transformation the following relation can be derived,

$$J\left(\hat{\phi}\right) = \left(-\frac{1}{n}\frac{\partial^2 l}{\partial\phi^2}\right)_{\hat{\phi}} = \left(-\frac{1}{n}\frac{\partial^2 l}{\partial\theta^2}\right)_{\hat{\theta}}\left(\frac{d\theta}{d\phi}\right)^2_{\hat{\theta}} = J\left(\hat{\theta}\right)\left(\frac{d\theta}{d\phi}\right)^2_{\hat{\theta}}, \tag{4.28}$$

by choosing $\phi\left(\hat{\theta}\right)$ such that

$$\left|\frac{d\theta}{d\phi}\right|_{\hat{\theta}} \propto J^{-\frac{1}{2}}\left(\hat{\theta}\right), \tag{4.29}$$

$J\left(\hat{\theta}\right)$ will be a constant independent of $\hat{\phi}$ and data translated in ϕ. This means that the metric for which the non informative prior has to be locally uniform can be derived as

$$\frac{d\phi}{d\theta} \propto J^{1/2}(\theta) \quad \text{or} \quad \phi \propto \int^{\theta} J^{1/2}(t)\,dt \tag{4.30}$$

where, the integral is improper evaluated in θ, the corresponding prior for θ is

$$p(\theta) \propto \left|\frac{d\phi}{d\theta}\right| \propto J^{1/2}(\theta). \tag{4.31}$$

At this point the previous conclusions can be extended to more general cases, in fact usually the quantity in equation 4.25 is a function of all data $\mathbf{y}$ and $p(\mathbf{y}|\theta)$ is not of the form $p(\mathbf{y}|\theta) = h(\mathbf{y}) w(\theta) \exp[c(\theta) u(\mathbf{y})]$, so by noting that given θ,

$$-\frac{1}{n}\frac{\partial^2 l}{\partial \theta^2} = -\frac{1}{n}\sum_{u=1}^{n}\frac{\partial^2 \ln p(y_u|\theta)}{\partial \theta^2} \tag{4.32}$$

is the average value of n function of the components of $\mathbf{y}$. Moreover, assuming θ_0 as true value of θ, and so $\mathbf{y}$ are sampled from $p(\mathbf{y}|\theta_0)$, for $n \to \infty$ the average in equation 4.32 converge to the expected value, assuming that it exists,

$$\mathbb{E}_{\mathbf{y}|\theta_0}\left[-\frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2}\right] = -\int \frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2} p(\mathbf{y}|\theta_0)\, dy = a(\theta|\theta_0), \tag{4.33}$$

and $\hat{\theta}$ converges to θ_0. Therefore, approximately,

$$\left(-\frac{1}{n}\frac{\partial^2 l}{\partial \theta^2}\right)_{\hat{\theta}} \approx a\left(\hat{\theta}, \theta_0\right) \approx a\left(\hat{\theta}, \hat{\theta}\right) = \boldsymbol{J}\left(\hat{\theta}\right), \tag{4.34}$$

where $\boldsymbol{J} = a(\theta, \theta)$ is

$$\boldsymbol{J}(\theta) = \mathbb{E}_{\mathbf{y}|\theta}\left[-\frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2}\right]. \tag{4.35}$$

Using this expression to approximate 4.25,the metric $\phi(\theta)$ for which a locally uniform prior is approximative non informative, such that

$$\frac{d\phi}{d\theta} \propto \boldsymbol{J}^{1/2}(\theta) \quad \text{or} \quad \phi \propto \int^{\theta} \boldsymbol{J}^{1/2}(t)\, dt \tag{4.36}$$

and the corresponding non informative prior for θ is

$$p(\theta) \propto \boldsymbol{J}^{1/2}(\theta). \tag{4.37}$$

The quantity $\boldsymbol{J}(\theta)$ in equation 4.35 is known as Fisher's measure of information about θ in a single information y, as explained in Fisher (1922)and Fisher (1925), which can be

defined in a more general way,

$$\boldsymbol{J}_n(\theta) = \mathbb{E}_{\mathbf{y}|\theta}\left[-\frac{\partial^2 l}{\partial \theta^2}\right], \tag{4.38}$$

with expectation respect to $p(\mathbf{y}|\theta)$. Moreover, when $\mathbf{y}$ is randomly sampled $\boldsymbol{J}_n(\theta) = n\boldsymbol{J}(\theta)$ and equation 4.37 can be expressed by the Jeffrey's rule, for which the prior distribution for a single parameter θ is approximately non informative if defined proportional to the square root of the Fisher's information measure.
Finally, the previous result can be extended to the case of multiple parameters, that final part is particularity important since that kind of non informative prior has been used in the applications presented in the following Chapters. Considering a vector of observation $\mathbf{y}$ with dimension q having a multi-normal distribution

$$\mathbf{y}|\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma} \sim \mathcal{N}[\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}], \tag{4.39}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}) = [\mu_1(\boldsymbol{\theta}_1), ..., \mu_q(\boldsymbol{\theta}_q)]$ is a vector of function of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, .., \boldsymbol{\theta}_q)$, where $\boldsymbol{\theta}_k = (\theta_{ki}, i = 1, ..., p_k)$, $k = 1, ..., q$ and $\boldsymbol{\Sigma}$ is a $q \times q$ covariance matrix. Therefore, assuming that $\boldsymbol{\Theta} = (\boldsymbol{\theta}, \boldsymbol{\Sigma})$ with $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ approximately independent,

$$p(\boldsymbol{\Theta}) \approx p(\boldsymbol{\theta}) p(\boldsymbol{\Sigma}), \tag{4.40}$$

and the parametrization of $\boldsymbol{\theta}$ is such that it is appropriate consider $\boldsymbol{\theta}$ locally uniform

$$p(\boldsymbol{\theta}) = \text{constant}, \tag{4.41}$$

the non informative prior distribution of the $q(q+1)/2$ matrix $\boldsymbol{\Sigma}$ elements can be derived from the previous consideration. Therefore, following Gelman et al. (2004) the non informative multi-variate Jeffrey's prior density is

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(q+1)/2} \tag{4.42}$$

or equivalently

$$p(\boldsymbol{\Sigma}) \propto |\mathbf{R}|^{-(q+1)/2} \prod_{i=1}^{q} \frac{1}{\sigma_i}, \tag{4.43}$$

where σ_i^2 is the variance, $\mathbf{R}$ is the $q \times q$ correlation matrix and $|\cdot|$ denotes the determinant. These results are valid for multi-parameter models with multi-normal observable vector, if

other distributions are considered non informative prior can be derived by extending the Jeffrey's rule (Box and Tiao 1992).

4.3.4 Likelihood function

The likelihood function expresses the relationship between prior and posterior knowledge, embedding the new information, then it depends on the type and form of the observations. In this part of the thesis only the construction of a likelihood function for a set of n q-variate observations without censored data is considered. Assuming that for given $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_q)$ and covariance matrix $\boldsymbol{\Sigma}$, the vector error is

$$\mathbf{e}_i = \begin{bmatrix} \sigma_1 \epsilon_{1i} \\ . \\ . \\ . \\ \sigma_k \epsilon_{ki} \\ . \\ . \\ . \\ \sigma_q \epsilon_{qi} \end{bmatrix} = \begin{bmatrix} y_{1i} - \mu_{1i}(\boldsymbol{\theta}_1) \\ . \\ . \\ . \\ y_{ki} - \mu_{ki}(\boldsymbol{\theta}_k) \\ . \\ . \\ . \\ y_{qi} - \mu_{qi}(\boldsymbol{\theta}_q) \end{bmatrix}, \quad i = 1, ..., n \tag{4.44}$$

distribute as the q-variate normal $\mathcal{N}_q(\mathbf{0}, \boldsymbol{\Sigma})$, then the n q-variate observations are independent. Furthermore, the joint distribution of the n errors vectors $\mathbf{e} = (\mathbf{e}_1, ..., \mathbf{e}_n)^T$ with $\mathbf{e}_i = (e_{i1}, ..., e_{iq})^T$, is

$$\begin{aligned} p(\mathbf{e}|\boldsymbol{\theta}, \boldsymbol{\Sigma}) &= \prod_{i=1}^{n} p(\mathbf{e}_i|\boldsymbol{\theta}, \boldsymbol{\Sigma}) \quad -\infty < e_{ki} < \infty, \quad k = 1, ..., q, \quad i = 1, ..., n \\ &= (2\pi)^{-qn/2} |\boldsymbol{\Sigma}|^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \mathbf{e}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_i \right) \end{aligned} \tag{4.45}$$

The previous equation can be simplified by introducing the $q \times q$ matrix $\mathbf{S}(\boldsymbol{\theta})$ defined as

$$\mathbf{S}(\boldsymbol{\theta}) = [S_{kl}(\boldsymbol{\theta}_k, \boldsymbol{\theta}_l)] \tag{4.46}$$

with

$$S_{kl}\left(\boldsymbol{\theta}_k, \boldsymbol{\theta}_l\right) = \sum_{i=1}^{n} e_{ki} e_{li}, \tag{4.47}$$

and finally, by using equation 4.45 the likelihood becomes

$$L\left(\boldsymbol{\theta}, \boldsymbol{\Sigma} | \mathbf{y}\right) \propto p\left(\mathbf{e} | \boldsymbol{\theta}, \boldsymbol{\Sigma}\right) \propto |\boldsymbol{\Sigma}|^{-n/2} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\right) \mathbf{S}\left(\boldsymbol{\theta}\right)\right]. \tag{4.48}$$

The definition of the likelihood functions used in this work are based on the previous equation, in every application equation 4.48 has been specialized.

4.3.5 Posterior distribution

Once the prior distribution (equation 4.42) and the likelihood function (equation 4.48) have been defined, the posterior distribution for the parameters $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ of the multi-variate normal model can be calculated as

$$L\left(\boldsymbol{\theta}, \boldsymbol{\Sigma} | \mathbf{y}\right) \propto |\boldsymbol{\Sigma}|^{-(n+q+1)/2} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\right) \mathbf{S}\left(\boldsymbol{\theta}\right)\right], \tag{4.49}$$

where each parameter of $\boldsymbol{\theta}$ can vary from $-\infty$ to ∞. Box and Tiao (1992) show that for $n \geq q$ the marginal posterior distribution of $\boldsymbol{\theta}$ is given by

$$p\left(\boldsymbol{\theta} | \mathbf{y}\right) \propto |\mathbf{S}\left(\boldsymbol{\theta}\right)|^{-n/2}, \tag{4.50}$$

and in particular this equation is valid even when the expectation functions are not linear in the parameters space. Moreover, assuming a uni-variate probabilistic model and making the assumption of linearity of the expectation function in the unknown parameters $\boldsymbol{\theta}$, it can be written as

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta} + \sigma\boldsymbol{\epsilon}, \tag{4.51}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations, $\mathbf{H}$ is a $n \times k$ matrix of known regressors, $\boldsymbol{\epsilon}$ is a $n \times 1$ vector of normally distributed random variables with zero mean and unit variance and σ is the standard deviation of the probabilistic model error. As shown in Box and Tiao (1992), if $n > k$, the posterior distribution of $\left(\boldsymbol{\theta}, \sigma^2\right)$ can be written as

$$p\left(\boldsymbol{\theta}, \sigma^2 | \mathbf{y}\right) \propto p\left(\boldsymbol{\theta}, \sigma^2\right) p\left(s^2 | \sigma^2\right) p\left(\hat{\boldsymbol{\theta}} | \boldsymbol{\theta}, \sigma^2\right), \tag{4.52}$$

where

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{D} \\
s^2 &= \frac{1}{\eta}\left(\mathbf{y}-\hat{\mathbf{y}}\right)^T\left(\mathbf{y}-\hat{\mathbf{y}}\right) \\
\eta &= n-k \\
\hat{\mathbf{y}} &= \mathbf{H}\hat{\boldsymbol{\theta}}.
\end{aligned} \tag{4.53}$$

Moreover, assuming a non informative prior with $\boldsymbol{\theta}$ and $\ln(\sigma)$ approximately independent and locally uniform, i.e.

$$p\left(\boldsymbol{\theta},\sigma^2\right) = p\left(\boldsymbol{\theta}\right)p\left(\sigma^2\right) \propto \sigma^{-2} \tag{4.54}$$

the joint posterior in equation 4.52 can be rewritten as

$$p\left(\boldsymbol{\theta},\sigma^2|\mathbf{y}\right) \propto p\left(\sigma^2|s^2\right)p\left(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}},\sigma^2\right). \tag{4.55}$$

Furthermore, under the normality assumption of $\boldsymbol{\epsilon}$, the marginal posterior of σ^2 is $\eta s^2\chi_\eta^{-2}$ and the marginal posterior distribution of $\boldsymbol{\theta}$ is the following multi-variate t distribution $t_k\left[\hat{\boldsymbol{\theta}}, s^2\left(\mathbf{H}^T\mathbf{H}\right)^{-1},\eta\right]$,

$$\begin{aligned}
p\left(\boldsymbol{\theta}|\mathbf{y}\right) &= \frac{\Gamma\left(\frac{\eta+k}{2}\right)|\mathbf{H}^T\mathbf{H}|^{1/2}s^{-k}}{\left[\Gamma\left(\frac{1}{2}\right)\right]^k\Gamma\left(\frac{\eta}{2}\right)\left(\sqrt{\eta}\right)^k}\left[1+\frac{\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T\mathbf{H}^T\mathbf{H}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T}{\eta s^2}\right]^{-(\eta+k)/2} \\
&-\infty < \theta_i < \infty, \quad i=1,...,k.
\end{aligned} \tag{4.56}$$

In the previous equation $\hat{\boldsymbol{\theta}}$ is the mode and the mean of $\boldsymbol{\theta}$ and its covariance matrix is $\eta s^2\left(\mathbf{H}^T\mathbf{H}\right)^{-1}/(\eta-2)$, and the mean and variance of σ^2 are respectively

$$\begin{aligned}
&\eta s^2/(\eta-2) \\
&2\eta^2 s^4/\left[(\eta-2)^2(\eta-4)\right].
\end{aligned} \tag{4.57}$$

4.3.6 Computation of posterior statistics using Monte Carlo Markov Chain (MCMC)

The closed form solution presented in the previous section is not usually applicable to real problems, as those discussed in this work. Several different reasons could lead to the impossibility to determine the posterior distribution of $\boldsymbol{\Theta}$ using a closed form solution, for

instance the presence of censored data, an error distribution different to the normal one, integrations on manifolds and so on. However, assuming that the posterior statistics of $\boldsymbol{\Theta}$ exist, their calculation could be generally very difficult, as well as the normalizing constant κ which it implies manifold integration over the Bayesian kernel $L(\boldsymbol{\Theta})\, p(\boldsymbol{\Theta})$. These reasons justify the use of different approaches to calculate the posterior distribution, in this work the Markov Chain Monte Carlo (MCMC) technique, in the Metropolis-Hastings formulation, has been used (Gamerman and Lopes 2006a). The choice of this algorithm is mainly due to its versatility which allows one to apply it to a large number of different problems. On the other hand, MCMC is a sampling method so it is computationally expensive and its application, when FE models with a large number of Degrees of Freedom (DoF) are involved, could become prohibitive. Therefore, in this work the computational burden has been significantly reduced by using surrogate models. In this section MCMC technique in the Metropolis-Hastings version is briefly described, referring to specialized texts for a detailed study, for instance Gamerman and Lopes (2006a) and Gelman et al. (2004).
The Monte Carlo Markov Chain approach, in its general formulation, is powerful and versatile because can draw samples from any target probability density function π, for the uncertain parameters $\boldsymbol{\Theta}$, requiring only that this density can be calculated at $\boldsymbol{\Theta}$. The algorithm makes use of a proposal density function $q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_i)$, which depends on the current state of the chain $\boldsymbol{\Theta}_i$, to generate new proposed parameter samples $\boldsymbol{\Theta}^*$. The proposal $\boldsymbol{\Theta}^*$ can be accepted as the next state of the chain $\boldsymbol{\Theta}_{i+1} = \boldsymbol{\Theta}^*$, with acceptance probability $\alpha(\boldsymbol{\Theta}_i, \boldsymbol{\Theta}^*)$, or rejected otherwise with probability $1-\alpha$. The specification of the probability α allows generating a Markov chain with desired target density π. In algorithm 1 MCMC is shown.
Due to practical reasons, a symmetric distribution is chosen as proposal Gamerman and Lopes (2006b).

$$q(\boldsymbol{\Theta}_i|\boldsymbol{\Theta}_{i-1}) = q(\boldsymbol{\Theta}_{i-1}|\boldsymbol{\Theta}_i) \tag{4.58}$$

As explained in Tierney (1994), if the regularity conditions of aperiodicity, irreducibility, and positive recurrence of the Markov chain (Tweedie 1975) are satisfied, the distribution of the sampling sequence $\boldsymbol{\Theta}_i$ converges to the target posterior distribution regardless of the starting point. Nevertheless, diagnostics convergences criteria are applied because of checking of the compliance with the previous conditions could be very difficult in real problems.
Convergence diagnostics is used to determine whether the samples generated by MCMC are representative of the underlying target distribution. In this research work, the diagnostics metric proposed by Brooks and Gelman (1998) has been employed. The convergence

Algorithm 1 Metropolis-Hastings algorithm

procedure MHA $(p(\boldsymbol{\Theta}), q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_i), N)$
  draw initial value $\boldsymbol{\Theta}_0$ from the prior $p(\boldsymbol{\Theta})$
  for $i = 1$ to N **do**
    draw $\boldsymbol{\Theta}^*$ from the proposal distribution $q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_{i-1})$
    solve the deterministic model for $\boldsymbol{\Theta}^*$
    compute the residual vector $\mathbf{r}_i$
    evaluate the probability of acceptance
    $\alpha = \min\left\{1, \frac{\pi(\boldsymbol{\Theta}^*)q(\boldsymbol{\Theta}_{i-1}|\boldsymbol{\Theta}^*)}{\pi(\boldsymbol{\Theta}_{i-1})q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_{i-1})}\right\}$
    accept the next step with probability α
    $\boldsymbol{\Theta}_i = \boldsymbol{\Theta}^*$
    or rejected with probability $1 - \alpha$
    $\boldsymbol{\Theta}_i = \boldsymbol{\Theta}_{i-1}$
  end for
  end procedure

of Markov Chain simulation has been reached when inferences for quantities of interest do not depend on the starting point. Monitoring convergence is obtained by comparing several inferences performed with different starting points. The diagnostics metric is based on the calculation of the Multivariate Potential Scale Reduction Factor (MPSRF) $\hat{R}^p$ in the multivariate case. In general, MPSRF is defined as the ratio between total variance and within-sequence variance, and represents the upper bound of the maximum of the univariate Potential Scale Reduction Factor (PSRF) statistics $\hat{R}$ among $\boldsymbol{\Theta}$ variables. When the convergence is reached, the between-sequence variance should be negligible obtaining $\hat{R}^p = 1$. Usually, $\hat{R}^p = 1.1$ is considered as acceptable, but when the dimension of the problem increases, a convergence criterion $\hat{R}^p = 1.5$ is allowed. Therefore, $\hat{R}^p$ can be defined as

$$\hat{R}^p = \max_{\mathbf{a}} \frac{\mathbf{a}^T \hat{\mathbf{V}} \mathbf{a}}{\mathbf{a}^T \mathbf{W} \mathbf{a}} \tag{4.59}$$

where $\hat{\mathbf{V}}$ is the total variance extended to the multivariate case, $\mathbf{W}$ is the within-sequence variance extended to the multivariate case, and $\mathbf{a}$ is a vector used to achieve the maximum value of the ratio in equation 4.59.

4.4 general Polynomial Chaos Expansion (gPCE)

4.4.1 Functional approximation

The main aim of this work is the development of procedures for the updating of concrete gravity dam model parameters by using measurements recorded on the structure. In this context FE models characterized by a high number of DoF assume great importance to simulate the structural behaviour. Moreover, a large number of analyses must be performed in order to solve the inverse problem thus making the use of these kinds of procedures prohibitive in real cases. Therefore, surrogate models have been used in order to reduce the computational burden, reproducing the model output. Indeed, surrogate models can be used within the updating procedure instead of the real numerical models. In particular, the general Polynomial Chaos Expansion (gPCE) techniques (Xiu 2010) has been widely used in this work, because it also allows one both to solve the forward problem in a straightforward manner and to easily perform a sensitivity analysis.
The gPCE is a technique which propagates the uncertainties due to some unknown parameters, through a deterministic model. Moreover, gPCE allows building a proxy model, called response surface, which can be used in the updating procedure to reproduce the model response, and to reduce the computational burden. The uncertain structural response $u(\omega)$ can be described in a probabilistic space defined by the triplet $(\Omega, \mathfrak{F}, \mathbb{P})$: where Ω is the space of all events, $\mathfrak{F}$ is the σ-algebra and $\mathbb{P}$ the probability measure. Assuming that $u(\omega)$ is smooth enough to be represented in terms of some simple random variables $\boldsymbol{\theta}(\omega)$ (e.g. Gaussians, uniform, etc.) corresponding to the Askey scheme (Xiu and Karniadakis 2002), via the PCE (Wiener 1938), the structural response can be approximated by $u_P(\boldsymbol{\theta}(\omega))$ defined as

$$u(\boldsymbol{\theta}(\omega)) \approx u_P(\boldsymbol{\theta}(\omega)) = \sum_{\alpha \in \mathbf{I}} u^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}(\omega)). \tag{4.60}$$

The random variables collected in the vector $\boldsymbol{\theta}(\omega)$ represent the unknown model parameters. In the previous equation $\Psi_\alpha(\boldsymbol{\theta}(\omega))$ represents the multivariate orthogonal polynomials with degree up to P and with finite multi-index set $\mathbf{I}$, and $u^{(\alpha)}$ are the polynomial coefficients.
Given N samples of $u(\omega)$ the equation 4.60 can be rewritten in a more practical way,

$$u(\boldsymbol{\theta}(\omega_i)) \approx u_N(\boldsymbol{\theta}(\omega_i)) = \sum_{\alpha \in \mathbf{I}} u^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}(\omega_i)) \qquad i = 1, ..., N. \tag{4.61}$$

Denoting P the cardinality of the polynomial expansion, $\boldsymbol{s} := [u(\omega_i)] \in \mathbb{R}^N$, $\boldsymbol{\Psi} := [\Psi_\alpha(\boldsymbol{\theta}(\omega_i))] \in \mathbb{R}^{NxP}$ and $\boldsymbol{v} := \left[u^{(\alpha)}\right] \in \mathbb{R}^P$, equation 4.61 can be written in matrix-vector form

$$\boldsymbol{s} = \boldsymbol{\Psi}\boldsymbol{v}, \tag{4.62}$$

which is equivalent to the more robust projected version

$$\boldsymbol{d} := \boldsymbol{\Psi}^T\boldsymbol{u} = \boldsymbol{\Psi}^T\boldsymbol{\Psi}\boldsymbol{v} =: \boldsymbol{W}\boldsymbol{v}. \tag{4.63}$$

Finally, the choice of the polynomial family depends on the probability distributions of the random variables, as accurately described in Xiu (2010). Moreover, the choice of the maximum polynomial degree and the analyses number must be done balancing the accuracy of the result and the computational burden.

4.4.2 gPCE statistics

As mentioned before, one of the most important features of the gPCE is the easy calculation of the output statistics, by exploiting the orthogonality condition. Given a general polynomial expansion with degree P,

$$Q_P(\theta) = a_P\theta^P + a_{P-1}\theta^{P-1} + ... + a_1\theta + a_0, \quad a_P \neq 0, \tag{4.64}$$

the system of polynomials $\{Q_P(\theta), \quad P \in \mathbb{N}\}$ is orthogonal with respect to a real positive measure η if the following relation is satisfied,

$$\int_S Q_n(\theta)\, Q_m(\theta)\, d\alpha(\theta) = \gamma_n\delta_{mn}, \ m,n \in \mathbb{N}, \tag{4.65}$$

where S is the support of α, δ_{mn} is the Kronecker delta function, and γ_n are positive constant, usually called normalization constants. Regarding γ_n, if it is unitary the polynomial system is called orthonormal, and it is noteworthy that the system can be orthonormalized by defining $\tilde{Q}_n(\theta) = Q_n(\theta)/\sqrt{\gamma_n}$. Therefore, exploiting the orthogonality condition, the statistics can be easily derived, as in the case of the expected value

$$\mathbb{E}[u(\boldsymbol{\theta})] \approx \mathbb{E}[u_P(\boldsymbol{\theta})] = \int \sum_{\alpha\in\mathbf{I}} u^{(\alpha)}\Psi_\alpha(\boldsymbol{\theta})\, dF_{\boldsymbol{\theta}}(\theta) = u^0, \tag{4.66}$$

and the variance,

$$\mathrm{Var}\left(u\left(\boldsymbol{\theta}\right)\right) = \mathbb{E}\left[\left(u\left(\boldsymbol{\theta}\right) - \mathbb{E}\left[u\left(\boldsymbol{\theta}\right)\right]\right)^2\right] \approx \sum_{\alpha \in \mathbf{I}} \left(u^{(\alpha)}\right)^2 \mathbb{E}\left[\Psi_\alpha^2\left(\boldsymbol{\theta}\right)\right]. \tag{4.67}$$

In conclusion, using the gPCE, once the coefficients have been calculated, the model response statistics can be calculated in few simple algebraic steps.

4.4.3 Coefficients calculation via sparse Bayesian approximation

The system in equation 4.63 is usually undetermined, in particular when expensive solvers are involved, so a regularization criteria is needed in order to determine the polynomial coefficients collected in $\boldsymbol{v}$. The problem can be tackled following several different approaches (Ghanem and Spanos 1991), as the regularized least square

$$\boldsymbol{v} = \arg\min\left(\frac{1}{2}||\,W\,\boldsymbol{v} - \boldsymbol{d}||_2^2 + \frac{\lambda}{2}||\boldsymbol{v}||_2^2\right), \tag{4.68}$$

and the basis pursuit denoising

$$\boldsymbol{v} = \arg\min\left(\frac{1}{2}||\,W\,\boldsymbol{v} - \boldsymbol{d}||_2^2 + \frac{\lambda}{2}||\boldsymbol{v}||_1\right) \tag{4.69}$$

methods, also known respectively as l_2 and l_1 minimisation procedures. These are composed by the squared error part used to enforce closeness of $\boldsymbol{v}$ to the data and the regularization term enforcing the smoothness of $\boldsymbol{v}$. The coefficient λ is used to balance the previous two terms, which also represent the noise variance, and it is known to be an uneducated guess whose optimal value is difficult to find. In the case where $\lambda > 0$, l_1 is preferable since it promotes the sparsity of the solution. On the other hand, from the computational point of view l_2 is easier to solve since the solution $\boldsymbol{v}$ is linear in the data $\boldsymbol{b}$. Therefore, from the numerical point of view, several different approaches are available as discussed by Lorenz et al. (2015). In this work the procedure proposed by Rosić and Matthies (2017) has been used to calculate the coefficients $\boldsymbol{v}$. This approach faces the coefficients calculation problem in a Bayesian setting, considering the sparsity of the solution particularly efficient for non-linear problem as the elastic-plastic one. Since the calculation of the polynomial coefficients is an ill-posed problem, the main idea behind this approach is to interpret the samples as measurement data. Therefore, it can be regularized in a Bayesian setting. In particular, using the objective function indicated in equation 4.69, the sensing matrix $\boldsymbol{W}$ must satisfy the restricted isometry property (Tillmann and Pfetsch 2014) in order to

recover the sparse solution. To ensure that, the principle of random projection (Donoho 2006) is used in order to project $\boldsymbol{d}$, of equation 4.63, onto a basis that consist of random linear combination of basis function in $\boldsymbol{\Psi}$. Therefore, the equation 4.63 can be rewritten as

$$\boldsymbol{b} := \boldsymbol{W}\,\boldsymbol{d} = \boldsymbol{W}\boldsymbol{\Psi}^T\boldsymbol{\Psi}\,\boldsymbol{v} = \boldsymbol{A}\,\boldsymbol{v} \tag{4.70}$$

where $\boldsymbol{W}$ denote the carefully chosen random sensing matrix. In a probabilistic interpretation of equation 4.63 the unknown coefficients are priorly modelled as independent random variables $\boldsymbol{v}_f$, defined in the probabilistic space $(\Omega_\zeta, \mathfrak{F}_\zeta, \mathbb{P}_\zeta)$. Moreover, their joint probability density function is

$$p_f\left(\boldsymbol{v}\right) = \prod_{\alpha\in\mathbf{I}} p\left(v^{(\alpha)}\right), \tag{4.71}$$

where $p\left(v^{(\alpha)}\right)$ are the pdfs of the individual parameters. Therefore, equation 4.70 becomes uncertain and is described by prediction

$$\boldsymbol{b}_f\left(\zeta\right) = \boldsymbol{A}\,\boldsymbol{v}_f\left(\zeta\right). \tag{4.72}$$

Given the new data, namely the samples, the coefficients pdf can be updated following the Bayes's rule

$$\pi\left(\boldsymbol{v}|\boldsymbol{b}\right) \sim p\left(\boldsymbol{b}|\boldsymbol{v}\right) p_f\left(\boldsymbol{v}\right) \tag{4.73}$$

where $\pi\left(\boldsymbol{v}|\boldsymbol{b}\right)$ is the posterior density and $p\left(\boldsymbol{b}|\boldsymbol{v}\right)$ is the likelihood. Therefore, assuming a normally distributed prior

$$p_f\left(\boldsymbol{v}\right) \sim \exp\left(-\frac{1}{2}||\boldsymbol{v}||_2^2\right) \tag{4.74}$$

as well as the likelihood, then the posterior pdf assumes the form

$$\pi\left(\boldsymbol{v}|\boldsymbol{b}\right) \sim \exp\left(-\frac{1}{2}||\boldsymbol{A}\,\boldsymbol{v}-\boldsymbol{b}||_2^2\right)\exp\left(-\frac{1}{2}||\boldsymbol{v}||_2^2\right). \tag{4.75}$$

The maximum aposteriori (MAP) estimate of $\pi\left(\boldsymbol{v}|\boldsymbol{b}\right)$ is the minimiser of the objective function in equation 4.68. Following the same assumptions, and only taking the prior to follow the Laplace distribution

$$p_f\left(\boldsymbol{v}\right) \sim \exp\left(-\frac{1}{2}||\boldsymbol{v}||_1\right) \tag{4.76}$$

the posterior MAP estimate is the minimiser of the objective function given in equation 4.69. As in the deterministic case the computation of the posterior distribution by using the Gaussian prior is easier than in the Laplace case, since in the former case the prior and likelihood are conjugated the minimization of l_1 requires a sampling based approach as Monte Carlo Markov Chain algorithm. This problem can be avoided using a hierarchical prior distribution able to mimic the Laplace behaviour, but easier to evaluate. Therefore, following Tipping (2001), the polynomial coefficients can be considered normally distributed

$$p\left(\boldsymbol{v}|\boldsymbol{w}\right) = \prod_{\alpha} \mathcal{N} \sim \left(0, w_{\alpha}^{-1}\right) \tag{4.77}$$

with zero mean and the precision (inverse variance) $\boldsymbol{w}$ distributed according to the Gamma distribution. Therefore, the posterior distribution becomes

$$p\left(\boldsymbol{v}, \boldsymbol{w}|\boldsymbol{y}\right) \propto p\left(\boldsymbol{y}|\boldsymbol{v}, \boldsymbol{w}\right) p\left(\boldsymbol{v}|\boldsymbol{w}\right) p\left(\boldsymbol{w}\right) \tag{4.78}$$

which cannot be computed directly, so it is rewritten as

$$p\left(\boldsymbol{v}, \boldsymbol{w}|\boldsymbol{y}\right) = p\left(\boldsymbol{v}|\boldsymbol{y}, \boldsymbol{w}\right) p\left(\boldsymbol{w}|\boldsymbol{y}\right). \tag{4.79}$$

In equation 4.79 the term $p\left(\boldsymbol{v}|\boldsymbol{y}, \boldsymbol{w}\right)$ follows the normal posterior distribution, while the second term is approximated by the delta function at its modes. This kind of approach requires the marginalization of the hyper-parameters, which has been avoided in the work of Rosić and Matthies (2017) by introducing a filter as briefly described below. The updating problem can be written in the more fundamental Kolmogorov's approach, then the conditional expectation can be considered as a projection onto the subspace generated by the σ-algebra of data

$$\mathbb{E}\left(\boldsymbol{v}|\boldsymbol{b}\right) = P_{\sigma(\boldsymbol{b})}\boldsymbol{v}. \tag{4.80}$$

Considering an orthogonal projection, the conditional exception matches the minimum mean square estimate (Matthies et al. 2016)

$$\min \mathbb{E}\left(||\boldsymbol{v} - \mathbb{E}\left(\boldsymbol{v}|\sigma\left(\boldsymbol{b}\right)\right)||^2\right) \tag{4.81}$$

which implies an orthogonal decomposition

$$\boldsymbol{v} = P_{\sigma(\boldsymbol{b})}\boldsymbol{v} + \left(I - P_{\sigma(\boldsymbol{b})}\right)\boldsymbol{v}. \tag{4.82}$$

Since the direct computation of $P_{\sigma(\boldsymbol{b})}\boldsymbol{b}$ is difficult, following the Doob-Dynkin lemma an optimal measurable function φ, such that

$$\mathbb{E}\left(\boldsymbol{v}|\boldsymbol{b}\right) = P_{\sigma(\boldsymbol{b})}\boldsymbol{v} = \varphi\left(\boldsymbol{b}\right), \tag{4.83}$$

can be defined. In this way equation 4.82 becomes

$$\boldsymbol{v} = \varphi\left(\boldsymbol{b}\right) + \left(\boldsymbol{v} - \varphi\left(\boldsymbol{b}\right)\right). \tag{4.84}$$

The first part of equation 4.84 is the projection altered by the new data, while the orthogonal part remains unchanged, since it represents the prior knowledge. This leads to the filtering equation

$$\boldsymbol{v}_a\left(\zeta\right) = \boldsymbol{v}_f\left(\zeta\right) + \left(\varphi\left(\boldsymbol{b}\right) - \varphi\left(\boldsymbol{b}_f\left(\zeta\right)\right)\right). \tag{4.85}$$

Assuming a linear filter as optimal, the previous equation reduces to the well known Gauss-Markov-Kalman filter

$$\boldsymbol{v}_a\left(\zeta\right) = \boldsymbol{v}_f\left(\zeta\right) + \boldsymbol{K}\left(\boldsymbol{b} - \boldsymbol{b}_f\left(\zeta\right)\right). \tag{4.86}$$

In this context the factor $\boldsymbol{K}$ is the Kalman gain

$$\boldsymbol{K} = \mathrm{cov}_{\boldsymbol{v}_f,\boldsymbol{b}_f}\left(\mathrm{cov}_{\boldsymbol{b}_f} + \mathrm{cov}_{\epsilon}\right)^{\dagger} \tag{4.87}$$

where $\dagger$ represents the pseudo-inverse, while the covariance functions are defined as

$$\mathrm{cov}_{q,y} := \mathbb{E}\left(\left(q - \mathbb{E}\left(q\right)\right) \otimes \left(y - \mathbb{E}\left(y\right)\right)\right). \tag{4.88}$$

The error ϵ related to the truncation of equation 4.70 is introduced in equation 4.87. The linear filter could be non optimal in non linear case, but a higher polynomial order or the iterative version of equation 4.86 can be used to better account for non-linearity. In the latter case, the non-linear measurement operator $Y\left(\boldsymbol{v}\right)$ in $\boldsymbol{b} := Y\left(\boldsymbol{v}\right) + \epsilon$ is approximated as

$$Y_\lambda\left(\boldsymbol{v}\right) = \boldsymbol{M}\left(\boldsymbol{v} - \check{\boldsymbol{v}}\right) + \boldsymbol{a} = \boldsymbol{M}\widetilde{\boldsymbol{v}} + \boldsymbol{a}, \tag{4.89}$$

where $\widetilde{\boldsymbol{v}} := \boldsymbol{v} - \check{\boldsymbol{v}}$ is the fluctuating part of the random variable $\boldsymbol{v}$ considering $\mathbb{E}\left(\boldsymbol{v}\right) = \check{\boldsymbol{v}}$, and $\boldsymbol{M}$ is the linear measurement matrix. Therefore, an iterative formula can be designed

$$\boldsymbol{v}_a^{(i+1)} = \boldsymbol{v}_f + \boldsymbol{K}_\lambda^{(i)}\left(\boldsymbol{b} - \boldsymbol{a}^{(i)} - \boldsymbol{M}^{(i)}\left(\boldsymbol{v}_f - \hat{\boldsymbol{v}}^{(i)}\right) - \epsilon\right). \tag{4.90}$$

In equation 4.90, $\boldsymbol{M}^{(i)}$ and $\boldsymbol{a}^{(i)}$ can be the exact Jacobian and $\boldsymbol{a} := Y(\mathbb{E}(\boldsymbol{v}_f))$, or the inexact Jacobian and $\boldsymbol{a} := \mathbb{E}(Y(\boldsymbol{v}_f))$ in the case of unbiased estimate. The equation 4.90 becomes particularity interesting if seen in a functional approximation scheme, for which, instead of sampling, the random variables are expanded by the polynomial chaos approximation leading to a deterministic algebraic equation

$$\sum_{\beta\in\mathbf{J}} \boldsymbol{v}_a^{(\beta)}\boldsymbol{\Gamma}_\beta(\zeta) = \sum_{\beta\in\mathbf{J}} \boldsymbol{v}_f^{(\beta)}\boldsymbol{\Gamma}_\beta(\zeta) + \boldsymbol{K}\left(\sum_{\beta\in\mathbf{J}} \boldsymbol{b}^{(\beta)}\boldsymbol{\Gamma}_\beta(\zeta) - \sum_{\beta\in\mathbf{J}} \boldsymbol{b}_f^{(\beta)}\boldsymbol{\Gamma}_\beta(\zeta)\right) \tag{4.91}$$

where $\boldsymbol{\Gamma}$ denotes the polynomial corresponding to the distribution of the polynomial coefficients. The third part in equation 4.91 corresponds to the deterministic measurements, so it is a non-zero mean term. By projecting equation 4.91 onto polynomial basis $\boldsymbol{\Gamma}_b$ one can get

$$\mathbf{v}_a = \mathbf{v}_b + \mathbf{K}(\mathbf{b} - \mathbf{b}_f) \tag{4.92}$$

where $\mathbf{v}_f := \left[\mathbf{v}_f^{(\beta)}\right]_{\beta\in\mathbf{J}} = \left[v_f^{(\alpha,\beta)}\right]_{\alpha\in\mathbf{I},\beta\in\mathbf{J}}$. Finally, the Kalman gain $\mathbf{K}$ can be calculated using the algebraic expression of the covariance matrix

$$\boldsymbol{C}_{vf} = \mathbb{E}_\zeta\left((\hat{\boldsymbol{v}}_f - \bar{\boldsymbol{v}}_f)\otimes(\hat{\boldsymbol{v}}_f - \bar{\boldsymbol{v}}_f)\right) = \sum_{\alpha,\beta\in\mathbf{I}_p} \mathbb{E}_\zeta\left(\boldsymbol{\Gamma}_\alpha\boldsymbol{\Gamma}_\beta\right)\boldsymbol{v}_f^{(\alpha)}\otimes\boldsymbol{v}_f^{(\beta)} - \bar{\boldsymbol{v}}\otimes\bar{\boldsymbol{v}} \tag{4.93}$$

where $\bar{\boldsymbol{v}} := \mathbb{E}_\zeta(\boldsymbol{v}_f)$. Equation 4.93 can be written in a matrix form

$$\boldsymbol{C}_{vf} = \tilde{\boldsymbol{V}}_f\boldsymbol{\Delta}\tilde{\boldsymbol{V}}_f^T \tag{4.94}$$

where $(\boldsymbol{\Delta})_{\alpha\beta} = \mathbb{E}_\zeta(\boldsymbol{\Gamma}_\alpha\boldsymbol{\Gamma}_\beta) = \mathrm{diag}(\alpha!)$ and $\tilde{\boldsymbol{V}}_f$ is equal to $\boldsymbol{v}_f := \left(..., \boldsymbol{v}_f^{(\alpha)}, ...\right)^T$ without the mean part.

Performing only the l_2 minimization 4.92 does not lead to the sparsity of the solution, so a constraint has to be added

$$\begin{aligned} \min \quad & \mathbb{E}\left(||\boldsymbol{v} - \mathbb{E}(\boldsymbol{v}|\sigma(\boldsymbol{b}))||^2\right) \\ \text{such that} \quad & ||\mathbb{E}(\boldsymbol{v}|\sigma(\boldsymbol{b}))||_1 \leq \epsilon \end{aligned} \tag{4.95}$$

Finally, equation 4.95 is non-linear, and its sub-gradient (equation 4.81) can be rewritten as the pseudo-measurement equation

$$\boldsymbol{Z}(\boldsymbol{v}) := \boldsymbol{H}(\boldsymbol{v})\,\boldsymbol{v} - \epsilon = 0, \tag{4.96}$$

where $\boldsymbol{H}(\boldsymbol{v}) := \text{sign}(\boldsymbol{v})$ and ϵ is the tolerance with the covariance $\boldsymbol{C}_\epsilon$ chosen as regularization parameter. Finally, the algorithm consists of a sequential estimation in which the first update is obtained by using real measurements and and equation 4.86, thus by using the pseudo-one and the iterative formula 4.90.

4.4.4 Sensitivity analysis

In the modelling field, whether they are FE or probabilistic models, the comprehension of the solution variation varying the input parameters is a crucial aspect. In problems which involve RVs, the sensitivity aspect becomes even more important. In this context, the Sensitivity Analysis (SA) is fundamental, because it allows studying how the model output is affected by different uncertainty sources. Therefore, SA is used to determine the most contributing input variables to an output behaviour, as the non-influential inputs, or ascertain some interaction effects within the model (Saltelli et al. 2008). Moreover, the objectives of a sensitivity analysis are numerous:

- identify and prioritize the most influential inputs,
- identify non-influential parameters, this aspect assumes great importance if seen in a probabilistic setting because those parameters can be fixed to nominal values,
- map the output behaviour varying the input parameters and, if necessary, by focusing the SA only on a specific region of the model parameters space,
- calibrate some model inputs using available information (inverse problem).

Historically, the first developed SA methods were the local approaches, namely deterministic method based on the perturbation of model inputs. Usually the perturbation occurs around the mean values of the input parameters, and the sensitivity of the model is assessed by calculating the partial derivatives of the outputs in specific points of the solution domain. Local methods are affected by several drawbacks, in particular the linearity and normality assumptions and the local variation. Therefore, a new class of methods was developed in a statistical framework (Iooss and Lemaître 2014) in order to overcome the previous problems. In contrast to the local methods, this new approach was named "global sensitivity analysis", because it considers the whole variation range of the model parameters. Several different global methods have been developed in the recent 30 years. In this Chapter only the functional decomposition of the variance is briefly described since it has been used in this work, while a wide description of the methods is reported in Saltelli et al.

(2008).
In the general case of a non-linear, non-monotonic model, the decomposition of the output variance is still defined and it can be used for the SA. Defining $f(\cdot)$ a square-integrable function,defined on the unit hypercube $[0,1]^d$, it can be represented as a sum of elementary function (Hoeffding 1948),

$$f(\boldsymbol{\theta}) = f_0 + \sum_{i=1}^{d} d_i(\theta_i) + \sum_{i<j}^{d} f_{i,j}(\theta_i, \theta_j) + ... + f_{1,2,...,d}(\boldsymbol{\theta}). \tag{4.97}$$

This expansion is unique under conditions (Sobol 1993),

$$\int_0^1 f_{i_1,...,i_s}(\theta_{i_1}, ..., \theta_{i_s})\, d\theta_{i_k} = 0, \quad 1 \le k \le s, \quad \{i_1, ..., i_s\} \subseteq \{1, ..., d\}, \tag{4.98}$$

this implies that f_0 is a constant. In the SA framework $f(\cdot)$ is the deterministic solver, $\boldsymbol{\theta} = (\theta_1, ..., \theta_d)^T$ is the vector of the mutually independent random variables and $u = f(\boldsymbol{\theta})$ is the model outputs. Therefore, a functional approximation, usually called ANOVA (Efron and Stein 1981), can be defined

$$\mathrm{Var}(u) = \sum_{i=1}^{d} D_i(u) + \sum_{i<j}^{d} D_{i,j}(u) + ... + D_{1,2,...,d}(u), \tag{4.99}$$

where $D_i(u) = \mathrm{Var}[\mathbb{E}(u|\theta_i)]$, $D_{i,j}(u) = \mathrm{Var}[\mathbb{E}(u|\theta_i, \theta_j)] - D_i(u) - D_j(u)$ and so on. Therefore, the variance based sensitivity indices, also called Sobol' indices (Sobol 1993), are defined as

$$S_i = \frac{D_i(u)}{\mathrm{Var}(u)}, \quad S_{ij} = \frac{D_{ij}(u)}{\mathrm{Var}(u)}, \quad ... \tag{4.100}$$

The Sobol' indices indicate the share of variance of u that is due to a given input or a combination of inputs. The number of indices grows exponentially with the dimension d, there will be $2^d - 1$ indices. Therefore, usually, for practical reasons indices of order higher than two are not estimated. Other important measures to asses the model sensitivity are the so-called "*total indices*" or "*total effects*", introduced by Homma and Saltelli (1996),

$$S_{T_i} = S_i + \sum_{i<j} S_{ij} + \sum_{j\neq i, k\neq i, j<k} S_{ijk} + ... = \sum_{l \in \#i} S_l \tag{4.101}$$

where $\#i$ is the subset of $\{1, ..., d\}$ including i. Usually, when d is large only the main effects and the total ones are calculated, giving a good information on the model sensitivity.

Usually, to estimate Sobol's indices sampling based methods are used. These methods are particularly versatile and they can be used to approach every kind of problem. However, sampling methods become very expensive from the computational viewpoint, in particular when complex numerical models are involved. By using gPCE, the calculation of Sobol' indices becomes very easy. Following Sudret (2008) the PCE Sobol decomposition can be defined as

$$\begin{aligned} u_P(\boldsymbol{\theta}) = u^0 + \sum_{i=1}^{d} \sum_{\alpha \in \mathbf{I}_1} u^{(\alpha)} \Psi_\alpha(\theta_i) + \sum_{1 \le i_1 < i_2 \le d} \sum_{\alpha \in \mathbf{I}_{i_1,i_2}} u^{(\alpha)} \Psi_\alpha(\theta_{i_1}, \theta_{i_2}) + \\ ... + \sum_{1 \le i_1 < ... < i_s \le d} \sum_{\alpha \in \mathbf{I}_{i_1,...,i_s}} u^{(\alpha)} \Psi_\alpha(\theta_{i_1}, ..., \theta_{i_s}) + ... + \sum_{\alpha \in \mathbf{I}_{1,2,...,d}} u^{(\alpha)} \Psi_\alpha(\theta_1, ..., \theta_d). \end{aligned} \tag{4.102}$$

For seek of clarity, in the above equation the dependence of each polynomial to the each subset of input parameters has been highlighted. Since every term of the form $\sum_{\alpha \in \mathbf{I}_{i_1,...,i_s}} u^{(\alpha)} \Psi_\alpha(\theta_{i_1}, ..., \theta_{i_s})$ is a polynomial function depending, on all input parameters, then the summands in the Sobol' decomposition of $u_P(\boldsymbol{\theta})$ (see 4.97) is

$$u_{i_1,...,i_s}(\theta_{i_1}, ..., \theta_{i_s}) = \sum_{\alpha \in \mathbf{I}_{i_1,...,i_s}} u^{(\alpha)} \Psi_\alpha(\theta_{i_1}, ..., \theta_{i_s}). \tag{4.103}$$

Due to the uniqueness of the representation in equation 4.97 the equation 4.102 is the Sobol' decomposition of the gPCE.
In the end, the derivation of the sensitivity indices is very simple: once the PC expansion of the model is available, the polynomial coefficients are gathered according to the dependency of each basis polynomial, square-summed and normalized,

$$S_{i_1,...,i_s} = \frac{\sum_{\alpha \in \mathbf{I}_{i_1,....i_s}} \left(u^{(\alpha)}\right)^2 \mathbb{E}\left[\Psi_\alpha^2\right]}{\mathrm{Var}\left[\sum_{j=0}^{P-1} u^j \Psi_j(\boldsymbol{\theta})\right]}. \tag{4.104}$$

In the same way also the total sensitivity indices can be calculated, by having a given integer sequence $\mathbf{I}_{j_1,...,j_t} = \{(i_1, ..., i_s), (j_1, ..., j_t) \subset (i_1, ..., i_s)\}$, the total indices are

$$S_{T_{j_1,...,j_t}} = \sum_{(i_1,...,i_s) \in \mathbf{I}_{(j_1,...,j_t)}} S_{i_1,...,i_s}. \tag{4.105}$$

In conclusion, once the PCE of the model has been done, the Sobol' indices can be calculated without additional cost.

4.5 Optimal Experimental Design (OED)

4.5.1 Introduction

The aim of the last part of the thesis is the development of a procedure both to design the structural health monitoring system and to record information useful to update the material strength parameters. For this purpose, the Optimal Experimental Design (OED) has been used, since it can be defined in a Bayesian setting. In this section of the thesis the main concepts behind the OED are introduced.

4.5.2 Experimental design formulation

The experimental design relies on the construction of a design criterion, or objective function, that reflects how relevant or valuable an experiment is expected to be. Therefore, the first fundamental step is the definition of the experiment goal. The objective function must be defined in order to embody this request, then in a Bayesian setting it could be related to the probabilistic model or to the posterior distribution.
Once the experimental goal has been decided, the design criterion must be defined. In this context, following Huan and Marzouk (2013) the design criteria has been defined in a Bayesian setting. By defining the design variable $\mathbf{d} \in \mathbb{R}^{n_d}$ which represents the experimental condition and considering $\boldsymbol{\theta}$, the vector of unknown parameters and $\mathbf{y}$ the vector of the new observations, then equation 4.9 can be rewritten as

$$p\left(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}\right) = \frac{p\left(\mathbf{y}|\boldsymbol{\theta},\mathbf{d}\right)p\left(\boldsymbol{\theta}|\mathbf{d}\right)}{p\left(\mathbf{y}|\mathbf{d}\right)}. \tag{4.106}$$

All the probability densities are a function of the design variable $\mathbf{d}$. Usually, the prior knowledge expressed through $p\left(\boldsymbol{\theta}|\mathbf{d}\right)$ is assumed independent to $\mathbf{d}$, leading to the approximation $p\left(\boldsymbol{\theta}|\mathbf{d}\right) = p\left(\boldsymbol{\theta}\right)$. Following Lindley (1956) and Lindley (1972), an objective function for experimental design should have the following form,

$$\begin{aligned} U\left(\mathbf{d}\right) &= \int_{\mathcal{Y}}\int_{\Omega} u\left(\mathbf{d},\mathbf{y},\boldsymbol{\theta}\right)p\left(\boldsymbol{\theta},\mathbf{y}|\mathbf{d}\right)d\boldsymbol{\theta}d\mathbf{y} \\ &= \int_{\mathcal{Y}}\int_{\Omega} u\left(\mathbf{d},\mathbf{y},\boldsymbol{\theta}\right)p\left(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}\right)p\left(\mathbf{y}|\mathbf{d}\right)d\boldsymbol{\theta}d\mathbf{y}, \end{aligned} \tag{4.107}$$

where $U\left(\mathbf{d}\right)$ is the expected utility, $u\left(\mathbf{d},\mathbf{y},\boldsymbol{\theta}\right)$ is the utility function, Ω is the support of $\boldsymbol{\theta}$ and $\mathcal{Y}$ is the support of $p\left(\mathbf{y}|\mathbf{d}\right)$. The utility function expresses the usefulness of an experiment characterised by the the condition $\mathbf{d}$ given $\boldsymbol{\theta}$ and $\mathbf{y}$. Since the precise value of $\boldsymbol{\theta}$ is

unknown as the result of an experiment $\mathbf{y}$ before performing it, the expectation of u over the joint distribution of $\boldsymbol{\theta}$ and $\mathbf{y}$ is used.
Following Lindley (1956), the relative entropy from the posterior to the prior, or the Kullback-Leibler divergence, is used as utility function. Considering two generic distribution A and B, then KL divergence is define as

$$D_{KL}\left(A||B\right) = \int_{\Omega} p_A\left(\boldsymbol{\theta}\right) \ln\left[\frac{p_A\left(\boldsymbol{\theta}\right)}{p_B\left(\boldsymbol{\theta}\right)}\right] d\boldsymbol{\theta} = \mathbb{E}_A\left[\ln\frac{p_A\left(\boldsymbol{\theta}\right)}{p_B\left(\boldsymbol{\theta}\right)}\right]. \tag{4.108}$$

This quantity is non-negative, non-symmetric and reflects the difference in information carried out by two distributions (Cover and Thomas 2005). Equation 4.108 can be specialised for the inference problem writing the KL divergence from the posterior to the prior,

$$u\left(\mathbf{d},\mathbf{y},\boldsymbol{\theta}\right) \equiv D_{KL}\left(p_{\boldsymbol{\theta}}\left(\cdot|\mathbf{y},\mathbf{d}\right)||p_{\boldsymbol{\theta}}\left(\cdot\right)\right) = \int_{\Omega} p\left(\tilde{\boldsymbol{\theta}}|\mathbf{y},\mathbf{d}\right) \ln\left[\frac{p\left(\tilde{\boldsymbol{\theta}}|\mathbf{y},\mathbf{d}\right)}{p\left(\tilde{\boldsymbol{\theta}}\right)}\right] d\tilde{\boldsymbol{\theta}} = u\left(\mathbf{d},\mathbf{y}\right). \tag{4.109}$$

In the previous equation $\tilde{\boldsymbol{\theta}}$ is a dummy variable representing the parameters, thus given the internal integration over the parameter space equation 4.109 is not a function of $\boldsymbol{\theta}$. Finally, by substituting equation 4.109 in equation 4.107, and by using $\boldsymbol{\theta}$ instead of $\tilde{\boldsymbol{\theta}}$ the final form of expected utility U in a Bayesian setting, and called expected information gain in $\boldsymbol{\theta}$, can be written as

$$\begin{aligned} U\left(\mathbf{d}\right) &= \int_{\mathcal{Y}}\int_{\Omega} p\left(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}\right) \ln\left[\frac{p\left(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}\right)}{p\left(\boldsymbol{\theta}\right)}\right] d\boldsymbol{\theta} p\left(\mathbf{y}|\mathbf{d}\right) d\mathbf{y} \\ &= \mathbb{E}_{\mathbf{y}|\mathbf{d}}\left[D_{KL}\left(p\left(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}\right)||p\left(\boldsymbol{\theta}\right)\right)\right]. \end{aligned} \tag{4.110}$$

A large KL divergence from posterior to prior implies that the data $\mathbf{y}$ decreases entropy in $\boldsymbol{\theta}$ and those data are more informative for parameters inference.
Finally, the expected utility $U\left(\mathbf{d}\right)$ expressed in equation 4.110 must be maximized over the design space $\mathcal{D}$ in order to find an optimal experimental design.

$$\mathbf{d}^* = \arg\max_{\mathbf{d}\in\mathcal{D}} U\left(\mathbf{d}\right). \tag{4.111}$$

4.5.3 Experimental design calculation

Typically, the expected utility in equation 4.110 has no closed form solution, so it must be approximated. The simplest, but very expensive, approach is that one proposed by Huan and Marzouk (2013), which is based on the Monte Carlo sampling. First equation 4.110

must be rewrite by applying the Bayes' theorem to the quantities both inside and outside the logarithm,

$$\begin{aligned} U(\mathbf{d}) &= \int_{\mathcal{Y}} \int_{\Omega} p(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}) \ln\left[\frac{p(\boldsymbol{\theta}|\mathbf{y},\mathbf{d})}{p(\boldsymbol{\theta})}\right] d\boldsymbol{\theta} p(\mathbf{y}|\mathbf{d})\, d\mathbf{y} \\ &= \int_{\mathcal{Y}} \int_{\Omega} \ln\left[\frac{p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d})}{p(\mathbf{y}|\mathbf{d})}\right] p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d})\, p(\boldsymbol{\theta})\, d\boldsymbol{\theta} d\mathbf{y} \\ &= \int_{\mathcal{Y}} \int_{\Omega} \{\ln[p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d})] - \ln[p(\mathbf{y}|\mathbf{d})]\}\, p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d})\, p(\boldsymbol{\theta})\, d\mathbf{x} d\mathbf{y}. \end{aligned} \tag{4.112}$$

Therefore, the integral in the previous equation can be estimated by using the Monte Carlo sampling,

$$U(\mathbf{d}) \approx \frac{1}{n_{\text{out}}} \sum_{i=1}^{n_{\text{out}}} \left\{ \ln\left[p\left(\mathbf{y}^{(i)}|\boldsymbol{\theta}^{(i)},\mathbf{d}\right)\right] - \ln\left[p\left(\mathbf{y}^{(i)}|\mathbf{d}\right)\right] \right\}, \tag{4.113}$$

where $\boldsymbol{\theta}^{(i)}$ are drawn from the prior $p(\boldsymbol{\theta})$, $\mathbf{y}^{(i)}$ are drawn from the conditional distribution $p\left(\mathbf{y}|\boldsymbol{\theta}=\boldsymbol{\theta}^{(i)},\mathbf{d}\right)$ i.e. the likelihood and n_{out} is the sample number in the outer Monte Carlo estimate. The evidence $p\left(\mathbf{y}^{(i)}|\mathbf{d}\right)$ evaluated at $\mathbf{y}^{(i)}$ typically does not have an analytical form, so it can be estimated using another importance sampling estimate:

$$p\left(\mathbf{y}^{(i)}|\mathbf{d}\right) = \int_{\Omega} p\left(\mathbf{y}^{(i)}|\boldsymbol{\theta},\mathbf{d}\right) p(\boldsymbol{\theta})\, d\boldsymbol{\theta} \approx \frac{1}{n_{\text{in}}} \sum_{j=1}^{n_{\text{in}}} p\left(\mathbf{y}^{(i)}|\boldsymbol{\theta}^{(i,j)},\mathbf{d}\right), \tag{4.114}$$

where $\boldsymbol{\theta}^{(\cdot,j)}$ are drawn from the prior $p(\boldsymbol{\theta})$ and n_{in} is the samples number in the inner Monte Carlo estimate. The combination of equation 4.113 and 4.114 allows the calculation of the biased estimator $\hat{U}(\mathrm{d})$ of $U(\mathrm{d})$. Furthermore, the variance of this estimator is proportional to $A(\mathbf{d})/n_{\text{out}} + B(\mathbf{d})/n_{\text{out}}n_{\text{in}}$, where A and B are terms which depend only to the distributions at hand. Whereas the bias of the estimator is proportional to $C(\mathbf{d})/n_{\text{in}}$, then n_{in} controls the bias while n_{out} controls the variance.
Once the expected utility related to a particular value of d has been calculated, according to equation 4.111, $U(\mathbf{d})$ must be maximized, or in other words, the optimal experimental layout d^* must be determined. Several different approaches are available in literature but in recent years the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm has received considerable attention. The SPSA algorithm, proposed by Spall (1998a), uses finite difference estimates of the gradient calculated on the basis of only two random perturbations regardless of the problem's dimension:

$$\mathbf{d}_{k+1} = \mathbf{d}_k - a_k \mathbf{g}_k(\mathbf{d}_k), \tag{4.115}$$

$$\mathbf{g}_k\left(\mathbf{d}_k\right) = \frac{\hat{U}\left(\mathbf{d}_k + c_k\boldsymbol{\Delta}_k\right) - \hat{U}\left(\mathbf{d}_k - c_k\boldsymbol{\Delta}_k\right)}{2c_k} \begin{bmatrix} \Delta_{k,1}^{-1} \\ \Delta_{k,2}^{-1} \\ \cdot \\ \cdot \\ \\ \Delta_{k,n_d}^{-1} \end{bmatrix}, \tag{4.116}$$

where k is the iteration number,

$$a_k = \frac{a}{(A+k+1)^{\alpha}}, \quad c_k = \frac{c}{(k+1)^{\gamma}}, \tag{4.117}$$

and a, A, α, c and γ are algorithm parameters, with recommended values available in literature (Spall 1998b). Moreover, $\boldsymbol{\Delta}_k$ is a random vector whose components are i.i.d. draws from a symmetric distribution with finite inverse moment, in this work Bernoulli distribution has been used. An intuitive justification of the SPSA is that the error in the gradient "average out" over a large number of iteration (Spall 1998a). Moreover, randomness introduced through the noisy objective function $\hat{U}$ and the finite-difference-like perturbations allows for a global convergence property. Finally, constraints in SPSA are introduced by projection: if the current state is infeasible under all possible random perturbations, then it is projected to the nearest point that does satisfy these conditions. In the context of the present work, this last feature is particularly important because it allows inserting geometrical constraints for the definition of the design domain. Since SPSA is a gradient based method, it takes advantage of any regularity of the objective function, requiring only two function evaluations per step to estimate the gradient, and thus reducing the computational burden. However, a high noise level can lead to slow convergence or the stagnation of the algorithm in local minima.
Finally, it is worth of noting that SPSA is an expensive algorithm if complex numerical models are involved. However, the use of surrogate models, as the gPCE, allows reducing the computational burden, and therefore making possible the use of SPSA in this context.

Chapter 5

Probabilistic framework for static SHM of concrete gravity dams

5.1 Overview

In this Chapter a probabilistic framework for static SHM of concrete gravity dams is shown. As introduced in Chapter 1 the SHM frameworks proposed in this research work have two aims: control the structural health (diagnosis step) and reduce the uncertainties in the estimation of the structural fragility (prognosis step). The static SHM system is based on the measurements acquired during the regular use of the dam, i.e. displacements and environmental conditions. In dam engineering field this kind of monitoring system, which is the most widespread, is usually called *static monitoring system* because in this context the deformation of the dam is mainly due to seasonal temperature variations and changes in the level of the reservoir. Since displacements vary slowly, the problem can be approached as static.

Figure 5.1 shows the flowchart of the SHM system proposed in this Chapter. Three steps can be recognised, the first one is related to the acquisition of environmental data and displacements. These observations are used to updated the predictive model of the dam

displacements through the Bayesian inference. In this context, the measurable variables collected into the vector $\mathbf{x}$, introduced in Chapter 4.2, are the environmental conditions recorded by the monitoring system, namely the reservoir levels and the temperatures. The presence of both thermal and mechanical actions requires coupled analyses, which are computationally demanding, in particular when complex FE models are involved. Moreover, probabilistic procedures for inverse problem solution require a large number of analyses. For this reason, meta models are used to build the predictive model of the dam displacements. In particular, the gPCE (Chapter 4.4) is used to approximate the FE model response for the water level variation δ_i^H. Following the same idea, thermal displacements δ_i^T are approximated through Fourier series, in order to avoid expensive thermal analyses (Noorzaei et al. 2006). Following the literature (Chapter 2), the well-known drift term δ_i^K is modelled by using a linear function in time, i.e. $\delta_i^K = \theta_{\mathrm{K}} t$. In this way, the set of unknown parameters $\boldsymbol{\theta}_{\mathrm{m}}$ introduced in Chapter 4.2 can be thought as composed by the gPCE random variables, collected in $\boldsymbol{\theta}_{\mathrm{gPCE}}$, the combination coefficients of the Fourier analysis $\boldsymbol{\theta}_{\mathrm{FA}}$, and the coefficient of the drift term θ_{K}, that is $\boldsymbol{\theta}_{\mathrm{m}} = [\boldsymbol{\theta}_{\mathrm{gPCE}}, \boldsymbol{\theta}_{\mathrm{FA}}, \theta_{\mathrm{K}}]^T$.
The prior distributions of the mechanical parameters are based on the results of materials tests. Whereas, the prior distributions of parameters with no physical meaning and for which little is known a priori are defined by using non-informative prior distributions, according to Chapter 4.3.3. Once calibrated, the proposed predictive model for dam displacements can be used to control the health state of the structure by comparing the predictions with the observations. This last aspect is called *Control step*.
It is worth noting that the calibrated twin model of the dam can be used also to assess the structural fragility against extreme events. Indeed, the reduction of the effects of the epistemic uncertainty leads to a better estimation of the structural fragility.
The architecture of the proposed SHM allows building a reliable predictive model of the dam displacements which does not need FE coupled analyses. Therefore, the predictive model proposed in this Chapter is characterised by high processing speed and thus it can be used to control in real time the dam behaviour. In this regard, the definition of the threshold beyond which the dam behaviour is considered abnormal is a crucial aspect. The estimated global error standard deviation σ can be used to set a threshold, once a study to assess the correlation between damage and static displacement variation has been performed.
Every time new observations are available the state of knowledge can be updated, improving the reliability of the procedure and the validity of the predictive model of static displacements. Therefore, due to the large amount of data, possible approximations, in

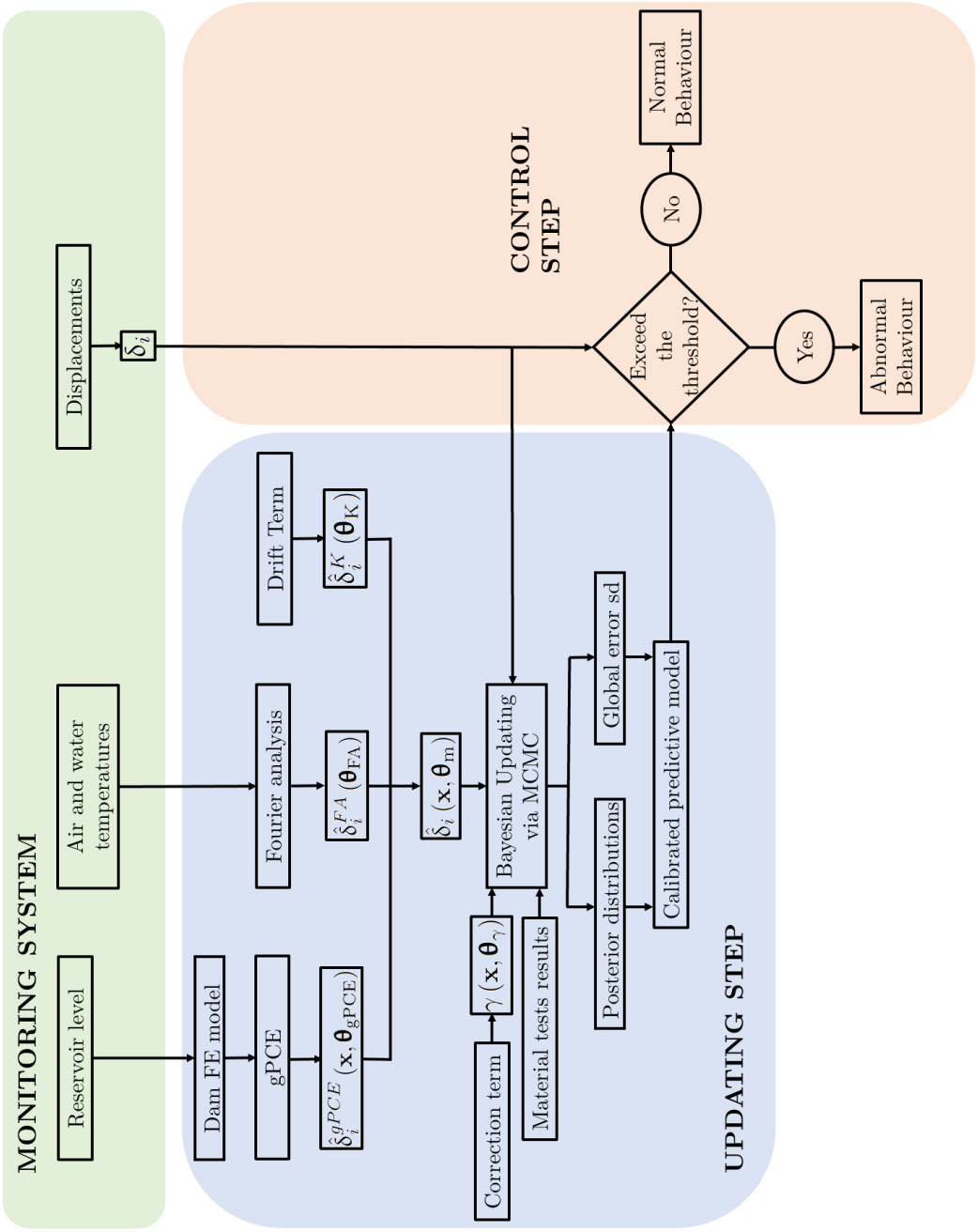


Fig. 5.1: Static SHM system for concrete gravity dams.

the definition of the predictive model are mitigated.
Differently from the static SHM system discussed in Chapter 2, the one proposed in this work introduces the following novelties:

- The static SHM is defined in the Bayesian setting.
- It allows estimating the global error standard deviation which can be used to set a threshold.
- The predictive model has high computational speed because it is a hybrid combination of meta models, so it does not need FE analyses.
- The Bayesian inference allows continuously improving the reliability of the prediction.
- The posterior distributions of the mechanical parameters of the materials can be used to reduced the effect of epistemic uncertainties in the fragility assessment of the dam.

In the following the components of the proposed static SHM framework are discussed in detail.

5.2 Definition of the proxy model for hydrostatic displacements through general Polynomial Chaos Expansion

In this work the gPCE has been used to approximate the displacements of the dam due to the basin level variation calculated through a FE model of the structure. Materials during the normal operation of a dam may be considered linear elastic. The mechanical behaviour of elastic material is thoroughly described by the elastic constitutive matrix $\mathbb{C}$ (Timoshenko and Goodier 1986), whose components may be treated as random variables in a Bayesian updating procedure in the most general case of anisotropic material. Modelling material as anisotropic is convenient only when it is strictly necessary and measurements which are able to infer all the components of the constitutive matrix $\mathbb{C}$ are available. In the case of dams, it would be worthwhile to update the parameters of an orthotropic material by using a multivariate probabilistic model only if displacements in upstream-downstream direction and in the cross-valley direction are recorded.
Conversely, when only upstream-downstream measurements are available, an isotropic ma-

terial is the best choice. For an elastic isotropic material, matrix $\mathbb{C}$ is as follows

$$\mathbb{C} = \begin{bmatrix} K+4G/3 & K-2G/3 & K-2G/3 & 0 & 0 & 0 \\ K-2G/3 & K+4G/3 & K-2G/3 & 0 & 0 & 0 \\ K-2G/3 & K-2G/3 & K+4G/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix}. \tag{5.1}$$

Uncertainties parametrization for elastic materials done by selecting the bulk modulus K and the shear modulus G as random variables is a particularly convenient choice, since they are statistically independent and the components of the tensor $\mathbb{C}$ are linear with respect to K and G. Therefore, in this context the bulk and the shear modulus of the materials are respectively collected in $\boldsymbol{\theta}_{\mathrm{gPCE},K}$ and $\boldsymbol{\theta}_{\mathrm{gPCE},G}$, then $\boldsymbol{\theta}_{\mathrm{gPCE}} = \left[\boldsymbol{\theta}^T_{\mathrm{gPCE},K}, \boldsymbol{\theta}^T_{\mathrm{gPCE},G}\right]^T$.

5.3 Definition of the proxy model for thermal displacements through Fourier analysis

The response of the structure caused by the basin level variation is related to the values of materials mechanical parameters. Whereas, the response of the structure due to thermal effects is related to thermal materials parameters and it is relatively unaffected by mechanical parameters variation. Anyway, thermal displacements are the largest part of total displacements, so they cannot be neglected. Since thermal displacements are characterized by a periodic trend over the years (Léger and Seydou 2009), they can be approximated through a Fourier analysis (FA). Using the Fourier series and defining the vector of the combination coefficients $\boldsymbol{\theta}_{\mathrm{FA}} = \left[\boldsymbol{\theta}^T_{\mathrm{FA},a}, \boldsymbol{\theta}^T_{\mathrm{FA},b}\right]^T$, the target function $s\left(t, \boldsymbol{\theta}_{\mathrm{FA}}\right)$ is approximated by $s_M\left(t, \boldsymbol{\theta}_{\mathrm{FA}}\right)$, which is the sum of simple sine waves (Davis 2016). Usually, it is expressed as

$$s_M\left(t, \boldsymbol{\theta}_{\mathrm{FA}}\right) = \sum_{m=1}^{M} \left[\theta_{\mathrm{FA},a_m} \cos\left(2\pi m t \omega\right) + \theta_{\mathrm{FA},b_m} \sin\left(2\pi m t \omega\right)\right], \tag{5.2}$$

where the combination coefficients θ_{FA,a_m} and θ_{FA,b_m} are respectively collected in the vectors $\boldsymbol{\theta}_{\mathrm{FA},a}$ and $\boldsymbol{\theta}_{\mathrm{FA},b}$. Moreover, t is the time and ω is the frequency of the wave. All combination coefficients are treated as random variables in the same way as the mechanical

parameters, while the number of harmonics is chosen in order to reach the best fitting between recorded and simulated behaviour. In this context, air and water temperatures are used only to set the frequencies of the waves, which are usually close to $\frac{1}{365}$ if the time is expressed in days.

5.4 Definition of the probabilistic model for total displacements

In this section, the probabilistic model response $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$ (equation 4.2), written for the case of concrete gravity dam displacements is defined. In this context, the i-th simulated displacement of a point on the dam can be written as

$$\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = \hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}}) + \hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\mathrm{FA}}) + \hat{\delta}_i^{K}(\theta_{\mathrm{K}}), \tag{5.3}$$

where $\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$ is the i-th simulated total displacement, $\hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}})$ is the part related to the basin level variation approximated through the gPCE, $\hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\mathrm{FA}})$ is the thermal term approximated by the Fourier analysis and $\hat{\delta}_i^{K}(\theta_{\mathrm{K}})$ is the drift term.
As reported in Chapter 4.2, a transformation of the reference quantity is needed to satisfy the assumption at the base of the additive model. Specifically, the normal logarithmic function is used to define $C(\mathbf{x}, \boldsymbol{\Theta})$ and $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$. Therefore, translation l of the reference system is needed to obtain positive values of displacements,

$$\hat{c}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = \ln\left(\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) + l\right). \tag{5.4}$$

Finally, the proposed probabilistic model for the static displacements of a concrete gravity dam can be determined by combining the equations 4.2, 5.3 and 5.4, obtaining

$$\begin{aligned} &\ln\left(\delta_i(\mathbf{x}, \boldsymbol{\Theta}) + l\right) = \ln\left(\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) + l\right) + \gamma(\mathbf{x}, \boldsymbol{\theta}_{\gamma}) + \sigma\epsilon = \\ &= \ln\left(\hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}}) + \hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\mathrm{FA}}) + \hat{\delta}_i^{K}(\theta_{\mathrm{K}}) + l\right) + \gamma(\mathbf{x}, \boldsymbol{\theta}_{\gamma}) + \sigma\epsilon. \end{aligned} \tag{5.5}$$

Once the parameters collected in $\boldsymbol{\Theta}$ have been calibrated, the equation 5.5 can be directly used to predict the dam behaviour.

5.5 Model correction

Several kinds of uncertainties are involved in this model. The greatest uncertainty source is the time interval between cause and effect for the different measured quantities, i.e. temperature variation and thermal displacement. Other phenomena can be considered, such as the unknown behaviour of vertical joints, whose opening-closing movement is related to thermal variations, time-dependent phenomena of mechanical parameters and the accuracy of measuring instruments.
Explanatory functions $\gamma\left(\mathbf{x}, \boldsymbol{\theta}_{\gamma}\right)$ (equation 4.2) are introduced to reduce the bias, by taking into account phenomena which cannot be directly considered in the model.
To capture the potential bias, which is independent of the parameters $\mathbf{x}$ and $\boldsymbol{\theta}_{\mathrm{m}}$, the function $\mathrm{h}_1 = 1$ is considered. In this case, θ_1 represents the previously discussed shift between the reference systems. When using more than one explanatory function (Gardoni et al. 2002b; Andreini et al. 2016), a stepwise deletion process can be performed in order to identify those functions which really contribute reducing the bias.

5.6 Prior distribution definition

The results of in-situ and laboratory tests on the dam's concrete and on the foundation rock are usually available, then the prior distributions of the model parameters $\boldsymbol{\theta}_{\mathrm{gPCE}}$, $\boldsymbol{\theta}_{\mathrm{gPCE},K}$ and $\boldsymbol{\theta}_{\mathrm{gPCE},G}$ can be deduced from these data. In this context, the other random variables collected in the vector $\boldsymbol{\theta}_{\mathrm{FA}}$ and $\boldsymbol{\theta}_{\gamma}$ and θ_{K} have no physical meaning, except for θ_1, and there is no prior information about them. The same consideration is valid for the global error standard deviation σ. For this reason, a set of non-informative prior distributions, introduced in Chapter 4.3.3, are selected.
As demonstrated in Chapter 4.3.3 a non-informative prior distribution is locally uniform near the likelihood function. In this way, the inference is not affected by information external to the observations. Therefore, in the present case the parameters $\boldsymbol{\theta}_{\mathrm{FA}}$, $\boldsymbol{\theta}_{\gamma}$, θ_{K} and σ are considered approximately independent, so

$$p\left(\boldsymbol{\theta}_{\mathrm{FA}}, \theta_{\mathrm{K}}, \boldsymbol{\theta}_{\gamma}, \sigma\right) \cong p\left(\boldsymbol{\theta}_{\mathrm{FA}}\right) p\left(\theta_{\mathrm{K}}\right) p\left(\boldsymbol{\theta}_{\gamma}\right) p\left(\sigma\right). \tag{5.6}$$

Moreover, by assuming a parametrization of $\boldsymbol{\theta}_{\mathrm{FA}}$, θ_{K} and $\boldsymbol{\theta}_{\gamma}$ such that is appropriate to take them locally uniform, the non informative prior distribution of the global error standard deviation becomes

$$p\left(\sigma\right) \propto \sigma^{-1}. \tag{5.7}$$

Finally, the other prior distributions become

$$p\left(\boldsymbol{\theta}_{\mathrm{FA}}, \theta_{\mathrm{K}}, \boldsymbol{\theta}_{\gamma}\right) \propto \frac{p\left(\boldsymbol{\theta}_{\mathrm{FA}}\right) p\left(\theta_{\mathrm{K}}\right) p\left(\boldsymbol{\theta}_{\gamma}\right)}{\sigma}. \tag{5.8}$$

Given the large amount of data recorded by the monitoring system, any reasonable choice of prior estimate has little influence on the posterior estimates of the parameters.

5.7 Likelihood function

Once the prior distributions of the unknown parameters are defined, the posterior ones must be calculated by using the Bayes's rule, as described in Chapter 4.3. In particular, equation 4.48 can be specialized for the uni-variate case of dam displacements, obtaining

$$L\left(\mathbf{x}, \boldsymbol{\Theta}\right) \propto \prod_{i=1}^{n} \left\{ \frac{1}{\sigma} \varphi \left[\frac{\mathbf{r}_i\left(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\theta}_{\gamma}\right)}{\sigma} \right] \right\}, \tag{5.9}$$

where $\mathbf{r}_i\left(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\theta}_{\gamma}\right)$, is the i-th residual which represents the discrepancy between the measurement and its prediction

$$\mathbf{r}_i\left(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\theta}_{\gamma}\right) = \ln\left(\delta_i\left(\mathbf{x}, \boldsymbol{\Theta}\right) + k\right) - \ln\left(\hat{\delta}_i\left(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}\right) + k\right) - \gamma\left(\mathbf{x}, \boldsymbol{\theta}_{\gamma}\right), \tag{5.10}$$

and $\varphi\left(\cdot\right)$ is the probability density function of a standard normal distribution. In this context, the vector of the unknown parameters is $\boldsymbol{\Theta} = \left[\boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\theta}_{\gamma}, \sigma\right]^T$.
The presented procedure uses the Metropolis-Hastings algorithm which requires, as previously described, the definition of the proposal distributions and the starting point. Normal distributions whit zero mean have been used as proposals. Their variances have been empirically defined in order to find a good agreement between calculation time and solution accuracy. The initial values of the random variables have been selected in order to maximize the likelihood function (Owen 2001).
The convergence of the probabilistic analysis has been checked as explained in paragraph 4.3.6, by performing two chains for each problem starting from different initial values. Every chain was stopped when an appropriate value of the acceptance ratio was reached, as indicated in literature (Gelman et al. 2004).

5.8 Final Remarks

The predictive model presented in this Chapter reproduces the displacements of a single point in only one direction. As mentioned before, more than one point or one direction can be considered by extending the proposed predictive model to the multi-variate case. In particular, every direction of every point would be a component of the multi-variate model, which would be a function of the position of the point itself. Moreover, the error terms would be correlated through a covariance matrix, thus defining a random field. With regard to the model proposed in this section, from the practical view point, the coefficient of the polynomial expansion $u^{(\alpha)}$, the combination coefficients of the Fourier series $\boldsymbol{\theta}_{\mathrm{FA}}$, the combination coefficient of the drift term θ_{K} and the error term $\sigma\epsilon$ would be functions of the point coordinates.
Another important assumption of the predictive model defined in this Chapter is that there is no statistical correlation in time among displacements. This assumption is based on the quasi-static nature of the dam displacements during normal operations. Therefore, the distribution of the error term does not vary during time and thus it is possible to sample from the same error distribution. From the physical point of view, the thermal component is the only one which varies in time (Chapter 5.3).

Chapter 6

Hierarchical Bayesian model for dynamic parameters updating via ambient vibrations

6.1 Overview

In this Chapter the probabilistic procedure for dynamic model parameters updating of concrete dams is described. This procedure, based on the Bayesian inference, allows reducing the epistemic uncertainties of a dynamic twin model of the dam. It can be used either to control the health state of the structure, if integrated in a SHM system, or to improve the estimation of the structural fragility, via the fragility curves calculation (Chapter 2.2). The concepts introduced in this Chapter are at the base of the dynamic SHM framework introduced in Chapter 7.

Figure 6.1 shows the proposed method. In this context, the reference dynamic measures are the modal characteristics of the system, i.e. frequencies and mode shapes, determined through the Operational Modal Analysis (OMA). As discussed in Chapter 2.3, ambient vibrations can be used as source of information in the updating process following two

different approaches. On the one hand, they can be directly used as reference measure, leading to a simplification from the measure point of view as no elaborations in this step are needed, but several complications from the computational side arise (Chapter 2.3). On the other hand, ambient vibrations can be elaborated through OMA techniques in order to determine experimental frequencies and mode shapes of the system. In the presented procedure the results of the OMA are used in order to reduce the computational burden and to avoid the previously discussed drawbacks connected to the direct use of ambient vibrations. However, as mentioned in Chapter 2.3 and discussed with more detail in this section, the use of modal characteristics as source of information within a probabilistic setting leads to complications related to the coherence between experimental and numerical modes. Experimental and numerical modes could be differently ordered, due to the variation of material characteristics, model error, measurement noises and so on. In dam engineering field this problem is even more evident because of the SSI in the numerical models. Indeed, by considering a particular frequency range, due to the SSI a large number of numerical modes related to the soil motion and with no experimental correlations are calculated. In this research work a modified version of the numerical algorithm MCMC is proposed in order to solve the coherence problem. More specifically, MCMC is modified by introducing a criterion which allows reordering the numerical modes coherently with the experimental ones. This also allows considering the SSI because the modes with no experimental correlation, can be recognized by the algorithm and discarded.
Since the reference measurements are modal characteristics, the main assumption is the linear elastic behaviour of the materials. Therefore, the aim of the procedure is the updating of the bulk and shear modulus, K and G, of the materials.
The gPCE is used to approximate the results of FEA in order to reduce the computational burden. More specifically, the eigenfrequencies collected in the vector $\boldsymbol{f}$ and the related mode shapes collected in the matrix $\boldsymbol{\Phi}$ are simulated through the gPCE 4.4. In this way, the set of unknown parameters $\boldsymbol{\theta}_{\mathrm{m}}$ introduced in Chapter 4.2.1 can be thought as composed by the gPCE random variables $\boldsymbol{\theta}_{\mathrm{gPCE}}$, that is $\boldsymbol{\theta}_{\mathrm{m}} = \boldsymbol{\theta}_{\mathrm{gPCE}}$.
The modal behaviour of a structure, both in terms of frequencies and mode shapes, is influenced by ambient conditions (Azzara et al. 2018). In the case of concrete gravity dams, water level, temperatures and humidity rates are the most influential factors. Therefore in this context, the measurable variable collected in the vector $\mathbf{x}$ of the probabilistic model (Chapter 4.2.1) are the environmental conditions.
In this case, a hierarchical Bayesian model is used to solve the inverse problem. This allows updating both the mean values and the standard deviations of the model parameters, thus

reducing the uncertainties at more levels. Therefore, with the aim to use the proposed procedure also to update a dynamic twin model of the dam for the calculation of fragility curves, this architecture of the probabilistic framework allows obtaining a more reliable estimation of the structural fragility.
The prior distributions of the mechanical parameters of the materials are defined starting from the results of material tests, while those related to model parameters with no-physical meaning are supposed non-informative, as reported in Chapter 4.3.3.

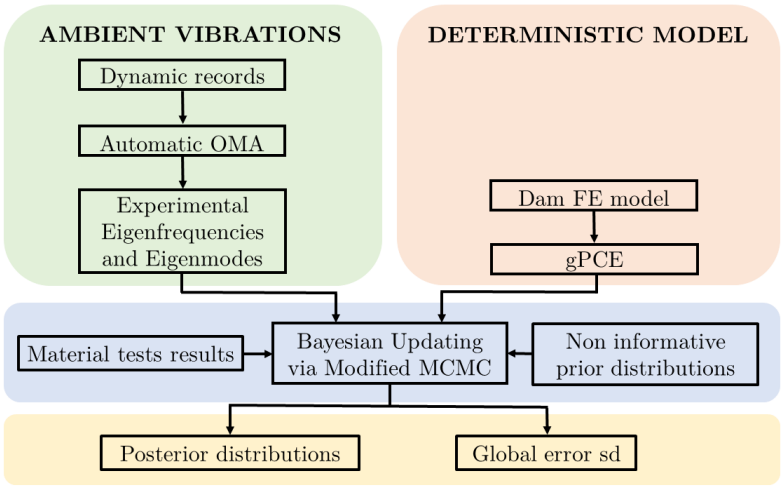


Fig. 6.1: Dynamic model parameters updating of concrete gravity dams.

The integration of the procedure within a SHM is needed to reduce the global error of the predictive model. Indeed the large amount of data continuously recorded could be used to update the proposed displacements model. Moreover, the use of a hierarchical Bayesian model allows introducing information in different levels, e.g. if new informations about the materials are available they can be inserted separately in terms of mean and variance.
Differently from the research works available in the literature (Chapter 2), the proposed approach is characterised by the following novelties:

- The proposed model is defined in a hierarchical way.
- The proposed modified version of MCMC allows directly solving the problem of coherence between experimental and numerical modes without using the concept of *system mode shapes* or objective functions.
- The proposed approach allows directly constructing a predictive model of the eigen-

frequencies and mode shapes. In this way, also the errors of the predictive models can be separately estimated.

- The use of gPCE-based meta models allows reducing the computational burden, thus making possible the application of the proposed procedure even when complex FE models are involved, or making possible its integration within a SHM framework.

6.2 The use of ambient vibration data as source of information in the model parameters calibration process

In Chapter 2.3.3.2 a literature review on the use of ambient vibration data as source of information for model parameters updating is introduced. The information embodied in ambient vibration are particularly important for seismic assessment purpose, since they can be used to calibrate dynamic parameters.
Ambient vibrations can be used directly for the model parameters updating or they can be used to determine the modal characteristics of the system, i.e. frequencies and mode shapes, which are used in the updating procedure. Both approaches have different advantages and disadvantages. On the one hand, the direct use of ambient vibrations within a probabilistic procedure leads to a grater computational burden, because transient analysis are needed. Moreover, modelling the unknown action which produces the vibration is a trivial task. Since ambient vibrations are just displacements, velocities or accelerations, then their errors can be easily characterised leading to a simplification from the measure viewpoint.
On the other hand, the use of OMA results allows defining the probabilistic problem in terms of modal characteristics. In this way, the modal analysis can be used thus reducing the computational time. However, in this case the reference QIs are derived quantities because experimental frequencies and mode shapes are a results of a post-process of the records. Therefore, the quantification of an error is difficult. The procedure proposed in this Chapter is defined with regard to the OMA results mainly for two reasons. First, the calibrated predictive model of the modal characteristics can be used for structural health monitoring purpose, because mode shapes are particularly sensitive to structural damage. The second reason is the idea to embody the proposed procedure within a SHM system, the continuous acquisition of information leads to a mitigation of the error related to the OMA.
In the context of OMA, ambient vibration data are measures of displacements, velocities or

accelerations recorded on a structure due to an unknown excitatory action. By assuming that the unknown action is a white noise, its power spectral density matrix is constant, the power spectral density matrix of the output and the frequency response function matrix are linked through a constant. Starting from this assumption the modal characteristics of the system can be easily determined (Brincker and Ventura 2015). Several different OMA techniques are available, some defined in the time domain and others in the frequency domain, in the recent years also OMA techniques set up in a Bayesian framework (Au et al. 2013) have been developed. The procedure introduced in this Chapter focuses the attention on the use of OMA results, rather than the development of new technique. However, since the guiding principle behind this work is the inclusion of the updating procedure in a SHM, some considerations about the nature of the reference information (OMA results) and the way in which they should be recorded and elaborated must be formulated. First, a dynamic monitoring system records continuously ambient vibrations, these records must be analysed in automatic way. Therefore, automated OMA techniques (Brincker and Ventura 2015) are preferable. Moreover, despite the influence of the measurement error on the result is mitigated by the large number of observations, OMA methods characterised by a higher level of accuracy are preferable.

In the last years several papers have been written on the use of OMA results as source of information for the model parameters calibration, most of them highlight the problem related to the coherence between numerical and experimental frequencies, regardless the deterministic or probabilistic nature of the updating procedure.

In the literature (**Tsogka2017**; Huang et al. 2018; Marwala 2010) two main solution strategies have been proposed: one is based on the concept of *system mode shapes* (Beck et al. 2001) and the other on the definition of objective functions (**Kodikara2016**; Jaishi and Ren 2005; Bassoli et al. 2018), containing a modal coherence criterion (Allemang 2003). Objective functions can be easily implemented, and the introduction of coherence criteria can help solve the coherence problem without increasing the dimension of the inverse problem. However, the error in this case is relative to the entire objective function, rather than to the QI. With the aim to define the predictive models directly in terms of modal characteristics, whose variation is sensitive to damage (Gopalakrishnan et al. 2011), the estimation of the error in term of QI is fundamental to avoid false warning. Conversely, the strategy based on the *system mode shapes* allows the predictive models to be defined directly in terms of the modal QI and for the coherence problem to be solved. However, since *system mode shapes* are treated as additional variables, the use of this strategy leads to an increase of the inverse problem dimension. In practical application, where the number of

the random variables is already high due to the parameters of the predictive models and the measurements are noised, further increase the dimension of the probabilistic problem by applying the *system mode shapes* approach could lead to numerical instability or low convergence rate.
The development of a robust procedure which accounts for the modes coherence is needed also with the prospect of controlling the health state of the structure, because a possible damage would affect the mode shapes, or their derivatives, more than the frequencies (Gopalakrishnan et al. 2011). For the same purpose the direct use of frequencies and mode shapes, instead objective functions, seems to be more appropriate, because it allows building directly a predictive model for the modal characteristics.
For the previous reasons the presented procedure is based on a multi-variate probabilistic model (Chapter 4.2.2) of the frequencies and mode shapes. They are defined and calibrated separately, even though in the same framework, in order to compare their different sensitivity with respect to the model parameters.

6.3 Definition of the proxy model for modal characteristics through general Polynomial Chaos

In the proposed procedure the gPCE is used to approximate the FEA results in terms of both frequencies $\boldsymbol{f}$ and mode shapes $\boldsymbol{\Phi}$. Due to the nature of the modal analysis, the aim of the procedure is the calibration of the elastic parameters of materials K and G. With regard to the notation introduced in Chapter 4 $\boldsymbol{\theta}_{\mathrm{gPCE}} = [\boldsymbol{\theta}_{\mathrm{gPCE},K}, \boldsymbol{\theta}_{\mathrm{gPCE},G}]^T$.
Conceptually the meta model is created in the same way as explained in Chapter 4.2.2 and 5.2, but in this case the gPCE is used to approximate the components of both the vector and the matrix. Therefore, the polynomial coefficients and the basis are function of the spatial coordinates $\mathbf{z}$ and the considered mode ϱ. That is, $u^{(\alpha)}(\mathbf{z}, \varrho)$ and $\boldsymbol{\Psi}_\alpha(\mathbf{z}, \varrho)$.

6.4 Definition of the probabilistic model for frequencies and mode shapes

In this section, the probabilistic model response $\hat{c}_k(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$ (equation 4.6) is specialised for the frequencies and mode shapes of concrete gravity dams.
The i-th value of the k-th numerical frequency approximated through the gPCE is

$$\hat{f}_{k,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = \hat{f}_{k,i}^{\mathrm{gPCE}}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}}), \tag{6.1}$$

where $\hat{f}_{k,i}^{\mathrm{gPCE}}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}})$ is the meta model created through the gPCE.
A transformation of the reference quantity is needed to satisfy the assumption at the base of the additive model, as reported in Chapter 4.2. Specifically, the normal logarithmic function is used to define $C(\mathbf{x}, \boldsymbol{\Theta})$ and $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})$,

$$\hat{c}_k(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = \ln\left(\hat{f}_k(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})\right). \tag{6.2}$$

Finally, the proposed probabilistic model for eigenfrequencies of a concrete gravity dams can be determined by combining equations 4.6, 6.1 and 6.2, thus obtaining

$$\begin{aligned} &\ln\left(f_{k,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\Sigma}_f)\right) = \ln\left(\hat{f}_{k,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}})\right) + \sigma_{f_k}\epsilon_{f_{k,i}} = \\ &= \ln\left(\hat{f}_{k,i}^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}})\right) + \sigma_{f_k}\epsilon_{f_{k,i}}, \end{aligned} \tag{6.3}$$

where $f_{k,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\Sigma}_f)$ is the i-th observation of the k-th experimental frequency and $\sigma_{f_k}\epsilon_{f_{k,i}}$ are the error terms collected in $\boldsymbol{\Sigma}_f$.
In the same way the probabilistic model for the mode shapes can be defined. Before starting with the development of the probabilistic model, the transformation of the mode shapes matrix into a vector is a convenient choice from the computational point of view. More specifically, by considering q modes of a system with m dynamic degrees of freedom (with $q \leq m$), the k-th mode shape vector $\boldsymbol{\phi}_k$ has dimension $m \times 1$, and the mode shapes matrix $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, ..., \boldsymbol{\phi}_k, ..., \boldsymbol{\phi}_q]$ has dimension $m \times q$. Therefore, by collecting every mode shape vector $\boldsymbol{\phi}_k$ in only one column, a total vector $\boldsymbol{\phi}^{\mathrm{total}}$ with dimension $m \cdot q \times 1$ can be obtained. By defining a global index h, with $1 \leq h \leq m \cdot q$, the meta model of the mode shapes becomes

$$\hat{\phi}_{h,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = \hat{\phi}_{h,i}^{\mathrm{gPCE}}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}}), \tag{6.4}$$

where $\hat{\phi}_{h,i}^{\mathrm{gPCE}}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}})$ is the i-th value of the h-th component of the numerical mode shapes vector approximated through the gPCE.
By combining equations 4.6 and 6.4, the predictive model for the mode shapes of the dam becomes

$$\begin{aligned} &\phi_{h,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}}) = \hat{\phi}_{h,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) + \sigma_{\phi_{h,i}}\epsilon_{\phi_h} = \\ &= \hat{\phi}_{h,i}^{\mathrm{gPCE}}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{gPCE}}) + \sigma_{\phi_{h,i}}\epsilon_{\phi_h}, \end{aligned} \tag{6.5}$$

where $\phi_{h,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}})$ is the i-th observations of the h-th component of the experimental mode shapes and $\sigma_{\phi_{h,i}}\epsilon_{\phi_h}$ are the error terms collected in the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\phi}}$.
The components of each the predictive model are correlated through the covariance matrix

(Chapter 4). With the assumption of correlation among the components of the predictive models, the dimension of the inverse problem increases considerably because also the error terms are objects of the updating process. By analysing the nature of the physical phenomenon and how the predictive models are used, some assumptions can be fixed in order to reduce the computational burden. Damage detection is based on the comparison between experimental and numerical modal characteristics, mode by mode, thus the components of a particular mode shape $\boldsymbol{\phi}_i$ must be assumed correlated, while the assumptions of correlation between different modes is not needed.
Therefore, the eigenfrequencies collected in $\boldsymbol{f}$ are uncorrelated and the related covariance matrix $\boldsymbol{\Sigma}_f$ has the form

$$\boldsymbol{\Sigma}_f = \begin{bmatrix} \sigma_{f_1}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{f_2}^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{f_q}^2 \end{bmatrix}, \tag{6.6}$$

where q is the number of considered modes.
Whereas, the covariance matrix $\boldsymbol{\Sigma}_\phi$ of the mode shapes has the form

$$\boldsymbol{\Sigma}_\phi = \begin{bmatrix} \blacksquare & & \\ & \blacksquare & \\ & & \blacksquare \end{bmatrix}, \tag{6.7}$$

where the blocks indicate the covariance matrices of each mode, otherwise the value is zero. Despite the number of terms within the correlation matrix $\boldsymbol{\Sigma}_\phi$ is strongly reduced, it could be still too high leading to an excessive computational burden. Therefore, the terms of each block within $\boldsymbol{\Sigma}_\phi$ are approximated through covariance functions with the form

$$\mathrm{COV}\left(\phi_i, \phi_j\right) = \frac{1}{\lambda^m} \exp\left[-\left(w_d^m d_{\phi_i,\phi_j}\right)\right], \tag{6.8}$$

where i and j are indices related to components within the same block, i.e. belonging to the same mode shape m. Moreover, d_{ϕ_i,ϕ_j} is the Euclidean distance between ϕ_i and ϕ_j. Finally, λ^m and w_d^m are combination coefficients, which are unknown and updated in the

proposed procedure. In this way, the number of random variables strongly decreases, since only two parameters for each mode are needed to describe the covariance matrix $\mathbf{\Sigma}_{\phi}$. The correlation function has been defined after a sensitivity analysis on some case studies, as shown in Chapter 8.3.7.
For the sake of simplicity, the unknown parameters λ^m are collected in the vector $\boldsymbol{\lambda}$ and w_d^m are collected in the vector $\mathbf{w}_d$. These vectors have dimension 1xq, where q is the number of considered modes.

6.5 Hierarchical Bayesian updating and likelihood function definition

One application of the procedure proposed in this Chapter is the calibration of the dynamic twin model parameters of the dam, in order to improve the prediction of the structural behaviour during seismic events. In particular, this last purpose can be reached by calculating the fragility curves, which, as explained in Chapter 2, account for both epistemic and aleatory uncertainties. With regard to this last aspect, both the mean values and the standard deviations of the model parameters distributions must be updated. Therefore, the presented procedure is defined in a hierarchical Bayesian framework. Conceptually a hierarchical Bayesian model has the same meaning of a normal one, and all properties described in Chapter 2 are still valid. In a hierarchical Bayesian model different levels are defined, each level is linked to the previous and the successive through statistical relationship. In other words, in a level the parameters governing the pdf of the successive level can be found. For instance, suppose that a random variable Y follows a normal distribution with mean θ and standard deviation equal to 1, that is $Y \sim \mathcal{N}(\theta, 1)$. Moreover, suppose that θ is normally distributed with mean μ and standard deviation 1, that is $\theta \sim \mathcal{N}(\mu, 1)$. Finally, assuming that μ is normally distributed with 0 mean and unit variance, $\mu \sim \mathcal{N}(0, 1)$, than the hierarchical model is defined because each level is linked to the next one with a statical relation.
Two important concepts in a hierarchical Bayesian model are the hyper-parameters and hyper-prior distributions. The hyper-parameters are the parameters of the prior distribution, as μ in the previous example. Whereas, the hyper-prior distributions are the pdf of the hyper-parameters (Gelman et al. 2004).
A hierarchical Bayesian model is needed when the entire prior distribution must be updated, i.e. when also the variance is influential, as in case of statistical structural analysis, i.e. when uncertainties are considered through their distributions. Moreover, hierarchical

model are important when the possibility of collecting information able to infer at different levels is available.
In the context of this procedure two levels are set up: the first one collecting the model parameters $\boldsymbol{\Theta}$, and the second one collecting the hyper-parameters $\boldsymbol{\Xi}$, as indicated in the following list.

- *Level 1*, Model Parameters $\boldsymbol{\Theta}$:
  - $\boldsymbol{\theta}_{\mathrm{m}}$: unknown model parameters,
  - $\boldsymbol{\Sigma}_f$: components of the covariance matrix of the frequencies probabilistic model,
  - $\boldsymbol{\lambda}$: coefficients λ^m of the covariance function of $\boldsymbol{\Sigma}_{\phi}$,
  - $\mathbf{w}_d$: coefficients w_d^m of the covariance function of $\boldsymbol{\Sigma}_{\phi}$,
- *Level 2*, Hyper-parameters $\boldsymbol{\Xi}$:
  - $\boldsymbol{\xi}_{\mathrm{m}}^{\mu}$: hyper-parameters of the mean values of the unknown model parameters $\boldsymbol{\theta}_{\mathrm{m}}$,
  - $\boldsymbol{\xi}_{\mathrm{m}}^{\sigma}$: hyper-parameters of the standard deviation of the unknown model parameters $\boldsymbol{\theta}_{\mathrm{m}}$.

With the aim to improve the estimation of the structural fragility, only the parameters $\boldsymbol{\theta}_{\mathrm{m}}$ play a fundamental role, thus only their distributions are fully updated. The previous parametrization reflects this idea, in fact the hyper-parameters are only related to $\boldsymbol{\theta}_{\mathrm{m}}$.
In the context of a hierarchical Bayesian model, the prior distribution $p(\boldsymbol{\Theta},\boldsymbol{\Xi}) = p(\boldsymbol{\Theta}|\boldsymbol{\Xi})\, p(\boldsymbol{\Xi})$ is updated in order to obtain the posterior distribution $p(\boldsymbol{\Theta},\boldsymbol{\Xi}|\mathbf{y})$. Equation 4.10 becomes,

$$p(\boldsymbol{\Theta},\boldsymbol{\Xi}|\mathbf{y}) = \kappa L(\mathbf{x},\boldsymbol{\Theta},\boldsymbol{\Xi}|\mathbf{y})\, p(\boldsymbol{\Theta}|\boldsymbol{\Xi})\, p(\boldsymbol{\Xi}). \tag{6.9}$$

Moreover, following equation 4.48, once n new observations collected in the vector $\mathbf{y}$ are recorded, the likelihood function for the eigenfrequencies of the dam is

$$L(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Sigma}_f,\boldsymbol{\Xi}) \propto \prod_{i=1}^{n} \frac{\exp\left[-\frac{1}{2}\mathbf{r}_i^{f\,T}(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi})\,\boldsymbol{\Sigma}_f^{-1}\mathbf{r}_i^f(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi})\right]}{\sqrt{|2\pi\boldsymbol{\Sigma}_f|}}. \tag{6.10}$$

Where $\mathbf{r}_i^f(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi}) = \left[r_{1,i}^f,...,r_{k,i}^f,...,r_{q,i}^f\right]^T$ represents the discrepancy between the i-th vector of experimental frequencies and the related numerical prediction. They assume the form

$$r_{k,i}^f = \ln\left(f_{k,i}(\mathbf{x},\boldsymbol{\Theta},\boldsymbol{\Xi})\right) - \ln\left(\hat{f}_{k,i}(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi})\right). \tag{6.11}$$

Following the same idea, the likelihood function for the mode shapes of a concrete gravity dam can be defined as

$$L\left(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Sigma}_{\phi},\boldsymbol{\Xi}\right)\propto\prod_{i=1}^{n}\frac{\exp\left[-\frac{1}{2}\mathbf{r}_{i}^{\phi^{T}}\left(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi}\right)\boldsymbol{\Sigma}_{\phi}^{-1}\mathbf{r}_{i}^{\phi}\left(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi}\right)\right]}{\sqrt{\left|2\pi\boldsymbol{\Sigma}_{\phi}\right|}}. \tag{6.12}$$

In this case $\mathbf{r}_{i}^{\phi}\left(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi}\right)=\left[r_{1,i}^{\phi},...,r_{h,i}^{\phi},...,r_{q\cdot m,i}^{\phi}\right]^{T}$ represents the discrepancy between the i-th observation of the experimental mode shape and related experimental one,

$$r_{h,i}^{\phi}=\phi_{h,i}\left(\mathbf{x},\boldsymbol{\Theta},\boldsymbol{\Xi}\right)-\hat{\phi}_{h,i}\left(\mathbf{x},\boldsymbol{\theta}_{\mathrm{m}},\boldsymbol{\Xi}\right). \tag{6.13}$$

In the next sections the definition of the hyper-prior distributions and the modified version of MCMC are treated.

6.6 Hyper-prior distribution definition

The results of in-situ and laboratory tests on the dam's concrete and on the foundation rock are usually available. Starting from these results, the distributions of the hyper-parameters $\boldsymbol{\Xi}$ of the mechanical parameters $\boldsymbol{\theta}_{\mathrm{m}}$ can be defined. Whereas there are no information about the parameters related to the error term, i.e. the elements of $\boldsymbol{\Sigma}_{f}$, $\boldsymbol{\lambda}$ and $\mathbf{w}_{d}$. For these random variables non informative prior distribution, as introduced in Chapter 4.3.3, are selected. Assuming that the parameters collected in $\boldsymbol{\Theta}$ are approximately independent, the prior distribution $p\left(\boldsymbol{\Theta},\boldsymbol{\Xi}\right)$ becomes

$$p\left(\boldsymbol{\Theta},\boldsymbol{\Xi}\right)=p\left(\boldsymbol{\Theta}|\boldsymbol{\Xi}\right)p\left(\boldsymbol{\Xi}\right)\cong p\left(\boldsymbol{\theta}_{\mathrm{m}}|\boldsymbol{\Xi}\right)p\left(\boldsymbol{\Xi}\right)p\left(\boldsymbol{\Sigma}_{f}\right). \tag{6.14}$$

Therefore, as discussed in Chapter 4, the non-informative prior distribution of the components of $\boldsymbol{\Sigma}$ can be written as

$$p\left(\boldsymbol{\Sigma}\right)\propto\left|\boldsymbol{\Sigma}\right|^{-(q+1)/2}. \tag{6.15}$$

Equations 6.15 can be directly used to define the prior distributions of the components of $\boldsymbol{\Sigma}_{f}$.

Since the components of $\boldsymbol{\Sigma}_{\phi}$ are approximated through covariance functions, the prior distributions of $\boldsymbol{\lambda}$ and $\mathbf{w}_{d}$ must be defined so that equation 6.15 is satisfied.

Nevertheless, when a large number of new data are collected, the choice of the prior distribution has little influence on the final result. Therefore, if the presented procedure

is integrated in a SHM system the procedure would lead to posterior distributions not influenced by the prior ones.

6.7 Modified Markov Chain Monte Carlo with Modal Assurance Criteria

The use of experimental modal characteristics, i.e. frequencies and mode shapes, as reference measure in the updating procedure involves the problem of modes matching. More specifically, the results of the deterministic model for a particular sample set $\bar{\boldsymbol{\Theta}}$ could be characterised by modes with a different sequence with respect to the experimental ones, leading to the comparison of frequencies related to different modes. Moreover, in the particular case of concrete gravity dams, considering the SSI in the numerical model, a large number of modes related only to the soil can be found in a selected frequency range. Usually, the largest part of these soil modes have no experimental correlation because the devices are installed only on the dam body or close to it. Therefore, alongside the need to guarantee the modes coherence, considering the SSI involves the need to discard the numerical modes which have no experimental correlation.
In the presented procedure the previous issues are solved by introducing a reordering criterion in the numerical algorithm MCMC. In particular, it changes the position of the numerical modes in order to ensure the coherence with the experimental ones, discarding numerical modes with no experimental correlation. Mode shapes, experimental and numerical ones, are used as reference information in the reordering criterion, through the calculation of the Modal Assurance Criterion (MAC) matrix. The MAC is a statistical indicator, a least squares based form of linear regression analysis which leads to an indicator that is most sensitive to the largest difference between experimental and numerical mode shapes (Allemang 2003). Therefore, by considering the i-th experimental mode shape $\boldsymbol{\phi}_i$ and the j-th numerical mode shape $\hat{\boldsymbol{\phi}}_j$, the component (i,j) of the MAC matrix is

$$\mathrm{MAC}\,(i,j) = \frac{|\boldsymbol{\phi}_i^T \hat{\boldsymbol{\phi}}_j^*|^2}{\left(\boldsymbol{\phi}_i^T \boldsymbol{\phi}_i^*\right)\left(\hat{\boldsymbol{\phi}}_j^T \hat{\boldsymbol{\phi}}_j^*\right)}. \tag{6.16}$$

Where the superscripts T and $*$ indicate respectively the transposed vector and the complex conjugated vector. Even when the eigenvectors are complex the MAC is real. The value of a MAC ranges from 0, in the case of no correlation, to 1, in the case of full correlation. The Metropolis-Hastings version of MCMC (algorithm 1) is modified by inserting the

calculation of the MAC matrix and the reordering criterion, as shown in algorithm 2. In particular, at every step of the algorithm the solution of the deterministic model is used to calculate the MAC matrix between numerical and experimental results. Therefore, numerical frequencies and mode shapes are reordered in order to move the higher values of the MAC coefficient toward the diagonal of the matrix. Whereas the numerical modes with no experimental correlation are discarded. In this way the coherence of the modes is ensured. This step needs the definition of a threshold beyond which the numerical mode is assumed equivalent to an experimental one. There are no rules to calibrate the threshold, but a preliminary sensitivity analysis could give helpful information. In the case of concrete gravity dams, as shown in the case study (Chapter 8), the modes are well defined, thus the MAC matrix is characterised by high values for coherent modes and small ones for the incoherent. Therefore, a threshold equal to 0.9 is set. A too high value of the threshold could lead to a stagnation of the algorithm. In fact, due to approximations both in the OMA, i.e. in the experimental results calculation, and in the numerical model, a perfect coherence, i.e. a MAC=1, is impossible to achieve.
The numerical algorithm, used in this procedure, needs a starting point in the space of the random variables, and the definition of proposal distributions (Chapter 4). The starting point of the chain is selected in order to maximize the likelihood function (Owen 2001), while normal distribution with zero mean and variance empirically calibrated are chosen. The calibration of the proposal distribution is a fundamental aspect, since it allows tuning the balance between calculation burden and accuracy of the results. The use of a meta model within the algorithm, instead of the real FE model, allows speeding up the procedure, thus several different calibration trials can be studied.

Algorithm 2 Metropolis-Hastings algorithm with MAC

procedure $(p(\boldsymbol{\Theta}|\boldsymbol{\Xi}), p(\boldsymbol{\Xi}), q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_i), N)$
draw hyper-parameter $\boldsymbol{\Xi}_0$ from the hyper-prior $p(\boldsymbol{\Xi})$
draw parameter $\boldsymbol{\Theta}_0$ from the prior $p(\boldsymbol{\Theta}|\boldsymbol{\Xi}_0)$
for $i = 1$ to N **do**
draw $\boldsymbol{\Xi}^*$ from the proposal distribution $q(\boldsymbol{\Xi}^*|\boldsymbol{\Xi}_{i-1})$
draw $\boldsymbol{\Theta}^*$ from the proposal distribution $q(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_{i-1})$
calculate the numerical frequencies and mode shapes by using $\boldsymbol{\Theta}^*$
for $j = 1$ to # experimental modes **do**
for $z = 1$ to # numerical modes **do**
calculate the MAC (j, z) (equation 6.16) by comparing the j-th experimental mode shape with the z-th numerical one
for $j = 1$ to # experimental modes **do**
for $z = 1$ to # numerical modes **do**
if MAC (j, z) >= treshold **then**
the z-th numerical mode is assumed correlated to the j experimental one
else
the z-th numerical mode is assumed uncorrelated to the j experimental one, so it is discarded
reorder the numerical modes coherently to the numerical ones, the higher values of the MAC coefficients are moved to the diagonal
compute the residual vector $\mathbf{r}_i$
evaluate the probability of acceptance
$\alpha = \min\left\{1, \frac{\pi(\boldsymbol{\Xi}^*,\boldsymbol{\Theta}^*)q(\boldsymbol{\Xi}_{i-1},\boldsymbol{\Theta}_{i-1}|\boldsymbol{\Xi}^*,\boldsymbol{\Theta}^*)}{\pi(\boldsymbol{\Xi}_{i-1},\boldsymbol{\Theta}_{i-1})q(\boldsymbol{\Xi}^*,\boldsymbol{\Theta}^*|\boldsymbol{\Xi}_{i-1},\boldsymbol{\Theta}_{i-1})}\right\}$
accept the next step with probability α
$\boldsymbol{\Xi}_i = \boldsymbol{\Xi}^*$ and $\boldsymbol{\Theta}_i = \boldsymbol{\Theta}^*$
or rejected with probability $1 - \alpha$
$\boldsymbol{\Xi}_i = \boldsymbol{\Xi}_{i-1}$ and $\boldsymbol{\Theta}_i = \boldsymbol{\Theta}_{i-1}$

Chapter 7

Probabilistic framework for dynamic SHM and devices optimization procedure

7.1 Overview

In this Chapter the proposed probabilistic framework for dynamic SHM of concrete gravity dams is shown. The first part of the Chapter is dedicated to the description of the methodology, while the last part focuses on the optimization procedure for the designing of the SHM layout.

The proposed probabilistic framework for dynamic SHM of concrete gravity dams is shown in figure 7.1. Through the elaboration of ambient vibrations and the use of predictive models of the dam behaviour, the presented SHM allows both detecting a possible damage and updating the parameters of the predictive models themselves. This last feature, as in the static case, allows reducing the epistemic uncertainties involved in the dam twin model, thus obtaining a more reliable SHM and a better estimation of the structural fragility. These two goals are achieved thanks to the architecture of the SHM, which is composed

by three phases based on structural analysis of the dam and Bayesian inference:

- *Training Phase*
- *Detection Phase*
- *Updating Phase*

The *Training Phase* is considered activated when the dam operates correctly, in this step the recorded ambient vibrations are used to update the parameters of the mode shapes predictive model described in Chapter 6. In this way a reliable predictive model of the undamaged dam is built. During the *Training Phase*, and coherently with Chapter 6, the measurable variables $\mathbf{x}$ represent the environmental conditions and the random variables $\boldsymbol{\theta}_{\mathrm{m}}$ are related to the elastic mechanical parameters, i.e. bulk and shear modulus. For the sake of simplicity, in this step the mode shapes predictive model is labelled with the superscript U, in order to indicate the undamaged behaviour of the dam, while the vector of unknown parameters $\boldsymbol{\theta}_{\mathrm{m}}$ is indicate with $\boldsymbol{\theta}_{\mathrm{el}}$, because K and G describe the elastic tensor.
The *Detection Phase* can be activated both during the regular used of the dam or when a seismic event occurs. In both cases a time window which balance recorded data and control must be defined. The aim of this step is to determine the presence and the position of a possible damage by comparing the experimental mode shapes with those obtained with the calibrated predictive model of the undamaged behaviour (previous step).
The *Updating Phase* is activated only when a seismic event occurs. The aim of this phase is the updating of the state of knowledge about the strength parameters of the materials, by exploiting the relationship between: damage, action and strength parameters. A second predictive model is used in this step to predict the possible damage of the structure, i.e. the mode shapes of the damaged dam. This predictive model is labelled with the superscript D. As further discussed, the damage propagation within a dam depends on the strength parameters of the materials, collected in $\boldsymbol{\theta}_{\mathrm{s}}$, and the action which produces the damage itself. In this context, the parameters which define the seismic action (IMs) are collected in the vector $\boldsymbol{\varpi}$. Therefore, the predictive model of the damaged behaviour is function of $\boldsymbol{\theta}_{\mathrm{s}}$ and $\boldsymbol{\varpi}$, while the elastic parameters are those calibrated in the training phase, i.e. $\boldsymbol{\theta}_{\mathrm{el}}$.
At the end of the Chapter a procedure to design the monitoring system is presented. The Optimal Bayesian Experimental Design (OBED), introduced in Chapter 4.5, is used to design the monitoring system of ambient vibrations. More specifically, the devices position is determined in order to maximize the acquisition of useful information with regard to

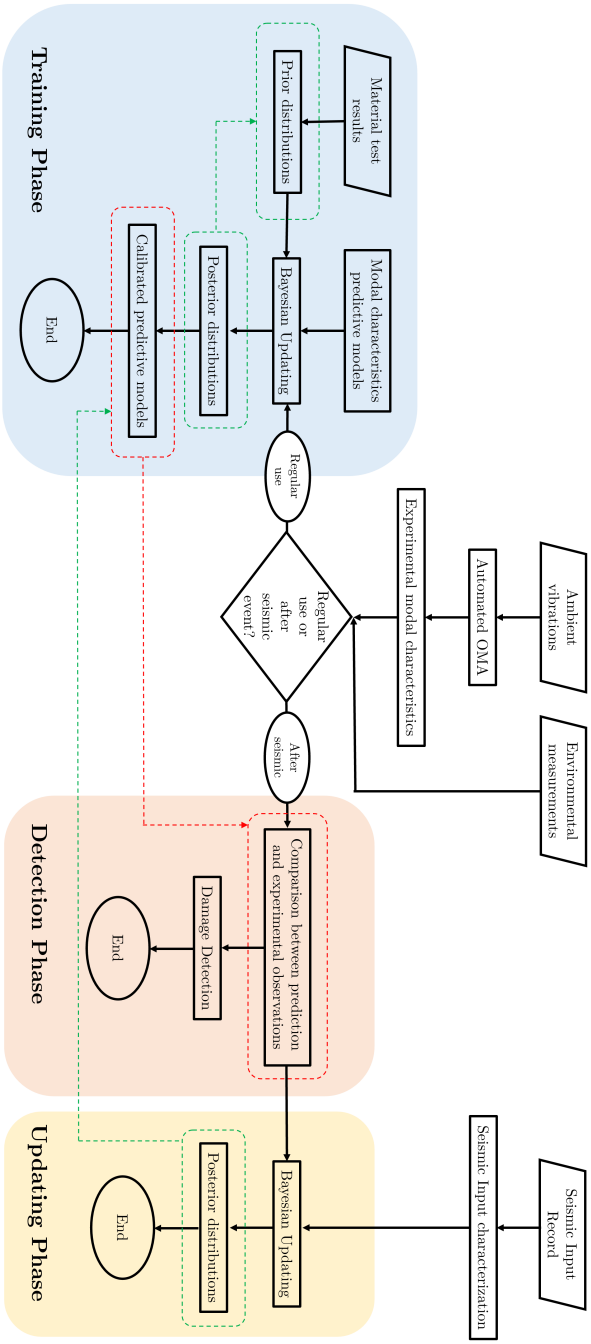


Figure 7.1: Dynamic SHM for concrete gravity dams.

the damage detection and to the updating of the material constitutive models strength parameters.
The gPCE-based meta models are used within the proposed SHM in order to reduce the computational burden, thus speeding up the whole procedure. This computational speed is an important feature of a SHM system, because the use of gPCE-based motel is fundamental.
The literature review presented in Chapter 2.3 shows that in dam engineering field there are only few examples of dynamic monitoring systems for structural control purpose. Most of the available research works face the problem of modes matching by using objective functions or by applying the *system mode shapes* concept, which are inappropriate in dam engineering field, because of the SSI in the numerical models. The probabilistic dynamic SHM framework proposed in this Chapter introduces the following novelties:

- the proposed SHM is directly based on the predictive model of the mode shapes of the dam.
- the predictive models are based on FE models of the system and calibrated by using recorded data.
- gPCE-based meta models are used instead of the FEA outputs, thus speeding up the procedure.
- the proposed framework allows both detecting the damage and updating the mechanical parameters of the materials of the predictive models, thus improving the SHM itself.
- the proposed optimization procedure allows designing the devices positions in order to maximize both the possibility of damage detection and the acquisition of information useful to infer the material strength parameters.

In the next sections, each step of the proposed SHM framework is described in detail, specialising the likelihood function and by specifying the parameters objects of the updating.

7.2 The training phase

The *Training Phase* is activated when the dam is regularly working. In this step the measurements recorded by the monitoring system are used to update the parameters of the predictive model of the undamaged mode shapes through the Bayesian Inference. In

particular, since the material behaviour is assumed linear during the regular use of the dam, the elastic parameters of the predictive model are updated, i.e. $\boldsymbol{\theta}_{\text{el}}$. For the sake of simplicity, the covariance matrix, whose components are updated in this step, is indicated as $\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\text{U}}}$, in the same way the combination coefficients of the covariance function (Equation 6.8) are collected in $\boldsymbol{\lambda}^{\text{U}}$ and $\mathbf{w}_d^{\text{U}}$.
The predictive model of the mode shapes of the undamaged dam are expressed by the equation

$$\phi_{h,i}^{\text{U}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\text{U}}}\right) = \hat{\phi}_{h,i}^{\text{U}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}\right) + \sigma_{\phi_h}^{\text{U}} \epsilon_{\phi_{h,i}^{\text{U}}}, \qquad h = 1, ..., m \cdot q \quad i = 1, ..., l. \tag{7.1}$$

In this context the likelihood function becomes

$$L\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\text{U}}}, \boldsymbol{\Xi}_{\text{el}}\right) \propto \prod_{i=1}^{l} \frac{\exp\left[-\frac{1}{2} \mathbf{r}_i^{\boldsymbol{\phi}^{\text{U}}\,T}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Xi}_{\text{el}}\right) \boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\text{U}}}^{-1} \mathbf{r}_i^{\boldsymbol{\phi}^{\text{U}}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Xi}_{\text{el}}\right)\right]}{\sqrt{|2\pi \boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\text{U}}}|}}, \tag{7.2}$$

where $\mathbf{r}_i^{\boldsymbol{\phi}^{\text{U}}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Xi}_{\text{el}}\right) = \left[r_{1,i}^{\boldsymbol{\phi}^{\text{U}}}, ..., r_{h,i}^{\boldsymbol{\phi}^{\text{U}}}, ..., r_{q \cdot m,i}^{\boldsymbol{\phi}^{\text{U}}}\right]^T$ are the residuals which express the difference between the i-th observation of the experimental mode shapes and related numerical ones

$$r_{h,i}^{\boldsymbol{\phi}^{\text{U}}} = \phi_{h,i}^{\text{U}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Sigma}_{\boldsymbol{\phi},\text{el}}\right) - \hat{\phi}_{h,i}^{\text{U}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}, \boldsymbol{\Xi}_{\text{el}}\right). \tag{7.3}$$

The computational burden related to the calculation of $\hat{\phi}_{h,i}^{\text{U}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{el}}\right)$, within the resolution numerical algorithm, is strongly reduced by the use of the gPCE. In this phase the gPCE based meta model is built on the results of the FEA analysis of the undamaged dam. The predictive model defined in this section is used for damage detection purposes. Whereas, the posterior distributions of the elastic parameters $\boldsymbol{\theta}_{\text{el}}$ are used within the second predictive model in the *Updating Phase*.

7.3 The detection phase

The *Detection Phase* is activated both during the regular use of the dam and when a seismic event occurs. In particular, it aims to detect a possible damage by comparing the experimental mode shapes, determined through the elaboration of ambient vibrations with OMA, with those simulated by the calibrated predictive model of the undamaged mode shapes. The processing speed, due to the use of a gPCE based meta model instead of the FE one, allows using the proposed predictive model for on-line structural control. A

suitable time window for the acquisition of the observations and the related parameters updating should be defined in order to optimise the performance of the SHM itself.
In this step two fundamental aspects must be discussed: the determination of experimental mode shapes from ambient vibrations and the definition of the threshold beyond which the dam behaviour is considered anomalous.
The experimental modal characteristics are determined through the elaboration of ambient vibrations by using OMA. Several different techniques are available, and in this context automatic ones are preferable because of practical reasons. However, the scope of this research work is the definition of the SHM framework rather than the selection of the best OMA techniques in dam engineering field.
The definition of the threshold of a particular damage level involves considerations on the physical phenomenon and the accuracies of both observation and prediction. From this latter point of view the proposed framework can facilitate the task. Indeed, it allows updating the covariance matrix $\mathbf{\Sigma}_{\phi^U}$ of the error terms, which are composed by different contributions. In the end, they give us an estimation of the degree of belief on the discrepancy between recorded and predicted QI. These indications are particularly important to define the threshold. Moreover, the continuous flow of information allows updating the covariance matrix, thus improving the capacity of the system itself in the damage detection.
The study of the physical phenomenon is needed to correlate a variation of the monitored QI with a particular damage level and the effects of the damage itself on the dam performance. This aspect is particularly important from the practical point of view where numerical studies and expert judgements assume great importance, due to the absence of case histories. Moreover, each dam has different features which can lead to different thresholds. Therefore, in every case a specific study of the correlation between damage, variation of the QI and dam performance should be set up. Also this aspect is out from the aims of this work.
Another important aspect related to damage detection is the relationship between minimal damage level and number of devices, which are strictly linked. Indeed fixing a minimal level of damage beyond which the SHM system must detect it, the optimization procedure presented in this Chapter could be used to optimize the minimum number of devices needed to detect the minimal damage. However, the definition of such threshold is very complicated, the procedure can be used to find the best devices layout once the minimum number of instrument is defined.

7.4 The updating phase

The *Updating Phase* is activated only when a seismic event occurs. It aims to updated the strength parameters of the dam concrete collected in $\boldsymbol{\theta}_{\mathrm{s}}$ by comparing the recorded mode shapes with those obtained with the predictive model of the damaged system. This is an important feature of the proposed framework because it allows improving the estimation of residual life expectancy of the dam after seismic event, for instance through the calculation of the fragility curves. Indeed, in this way the epistemic uncertainties related to the strength parameters are reduced, thus improving the estimation of the dam behaviour during seismic events. The damage propagation depends on the strength parameters of the concrete $\boldsymbol{\theta}_{\mathrm{s}}$ as well as the action which produces the damage itself, in this context the seismic one. With the aim to define a statistical relationship between QI, observations and error terms, a parametrization of the seismic input is needed. It is important to highlight the aleatory nature of the uncertainty related to the seismic action. Therefore, the state of knowledge about the seismic parameters can not be updated, they are needed only to consider the seismic action within the probabilistic model. For the sake of simplicity, the parameters of the seismic input are collected in the vector $\boldsymbol{\varpi}$. As for the other uncertainties involve in this work the parametrization is a fundamental aspect, so sensible IMs of the seismic action must be selected. Indications about the different IMs in the case of concrete gravity dams can be found in the literature (Chapter 2).
In this phase, the predictive model of the mode shapes of the damaged dam is expressed by the equation

$$\phi_{h,i}^{\mathrm{D}}\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}},\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\mathrm{D}}}\right)=\hat{\phi}_{h,i}^{\mathrm{D}}\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}}\right)+\sigma_{\phi_h}^{\mathrm{D}}\epsilon_{\phi_{h,i}^{\mathrm{D}}},\qquad h=1,...,m\cdot q\quad i=1,...,l. \tag{7.4}$$

As in the previous case the likelihood function is assumed normally distributed as the error term. Therefore, by indicating the covariance matrix as $\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\mathrm{D}}}$ the likelihood function becomes

$$L\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}},\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\mathrm{D}}},\boldsymbol{\Xi}_{\mathrm{s}}\right)\propto\prod_{i=1}^{l}\frac{\exp\left[-\frac{1}{2}\mathbf{r}_i^{\boldsymbol{\phi}^{\mathrm{D}}T}\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}},\boldsymbol{\Xi}_{\mathrm{s}}\right)\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\mathrm{D}}}^{-1}\mathbf{r}_i^{\boldsymbol{\phi}^{\mathrm{D}}}\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}},\boldsymbol{\Xi}_{\mathrm{s}}\right)\right]}{\sqrt{\left|2\pi\boldsymbol{\Sigma}_{\boldsymbol{\phi}^{\mathrm{D}}}\right|}}. \tag{7.5}$$

In the previous equation $\mathbf{r}_i^{\boldsymbol{\phi}^{\mathrm{D}}}\left(\mathbf{x},\boldsymbol{\varpi},\boldsymbol{\theta}_{\mathrm{s}},\boldsymbol{\Xi}_{\mathrm{s}}\right)=\left[r_{1,i}^{\boldsymbol{\phi}^{\mathrm{D}}},...,r_{h,i}^{\boldsymbol{\phi}^{\mathrm{D}}},...,r_{q\cdot m,i}^{\boldsymbol{\phi}^{\mathrm{D}}}\right]^T$ are the residuals which express the difference between the i-th value of the experimental mode shapes and

the related numerical one,

$$r_{h,i}^{\phi^{\mathrm{D}}} = \phi_{h,i}^{\mathrm{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Sigma}_{\phi^{\mathrm{D}}}, \boldsymbol{\Xi}_{\mathrm{s}}\right) - \hat{\phi}_{h,i}^{\mathrm{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}}\right). \tag{7.6}$$

The use of the gPCE for the construction of the meta model of $\hat{\phi}_{h,i}^{\mathrm{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}}\right)$ allows strongly reducing the computational burden. In this phase the gPCE based meta model is trained on the FEA results of the damaged dam. Therefore, a high-fidelity model must be used to simulate the damage of the system and the behaviour of the damaged dam. An important aspect, from the computational point of view is how to train the gPCE with respect to the seismic input. According to the literature the IMs, as those collected in $\boldsymbol{\varpi}$, are not sufficient to fully describe the variability of the seismic motion. Also the frequency-content parameters play a fundamental role. This problem can be solved by considering a set of seismic inputs as in the case of Multi-Record IDA. Therefore, for every analysis defined to train the meta model, a set of IMs is sampled from the prior distributions of $\boldsymbol{\varpi}$ and a seismic event is sampled from the set of recorded motions. Afterwards, the seismic event is scaled until its IMs are equal to the sampled ones, this step is illustrated in figure 7.2. Due to the aleatory nature of the earthquake, the prior distribution of $\boldsymbol{\varpi}$ and the set of reference seismic events must defined following practical concepts. In particular, the reference earthquakes must be chosen with regard to the site characteristics as in the case of MR-IDA. Whereas, the prior distributions of $\boldsymbol{\varpi}$ should be selected in order to reduce as much as possible the computational burden. This can be done by avoiding the analysis which do not lead to the damage of the dam. Preliminary analyses can be helpful to select the ranges of variation of the IMs.

In this phase, as in the previous one, only two parameters for each mode are needed to define the correlation matrix. For the sake of simplicity these are collected in the vectors $\boldsymbol{\lambda}^{\mathrm{D}}$ and $\mathbf{w}_d^{\mathrm{D}}$.

Finally, it is worth noting that even no damage, after a seismic event, is an important information. Indeed, if the prior mean values of the material strengths is too small, and for that the prediction would lead to a damaged structure for a particular seismic event, but no damage are determined on the real structure, the posterior distribution will be moved toward highest value. The possibility to determine the posterior distributions of the parameters collected in $\boldsymbol{\theta}_{\mathrm{s}}$ allows improving the estimation of the residual life expectancy of the structure, this step falls in the prognosis phase.

In this context, the seismic input, which is parametrised by $\boldsymbol{\varpi}$, influences the damage as well as the strength parameters. Therefore, once an earthquake occurs must be characterised, determining the parameters of the vector $\boldsymbol{\varpi}$, which are usually defined with regard to the

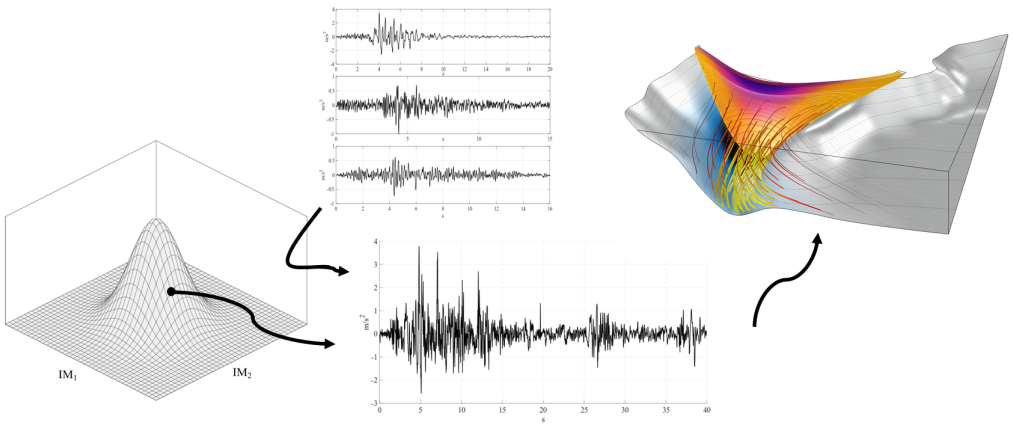


Figure 7.2: Seismic action modelling in the *Updating Phase*.

bed rock. Usually, dam are built on rock soil, then it could be sufficient to install devices on the ground around the dam in order to characterise the seismic event. Once the seismic input is characterised, the vector $\boldsymbol{\varpi}$ is defined and it can be fixed in the predictive model of equation 7.4. In this way, only the strength parameters $\boldsymbol{\theta}_s$ remain unknown.
Since the quantities of interest are the mode shapes, the problems related to the SSI and the experimental numerical coherence arise. Therefore, the modified version of MCMC, proposed in Chapter 2, must be used to solve these problems.
Finally, figure 7.3 shows the procedure presented in this Chapter.

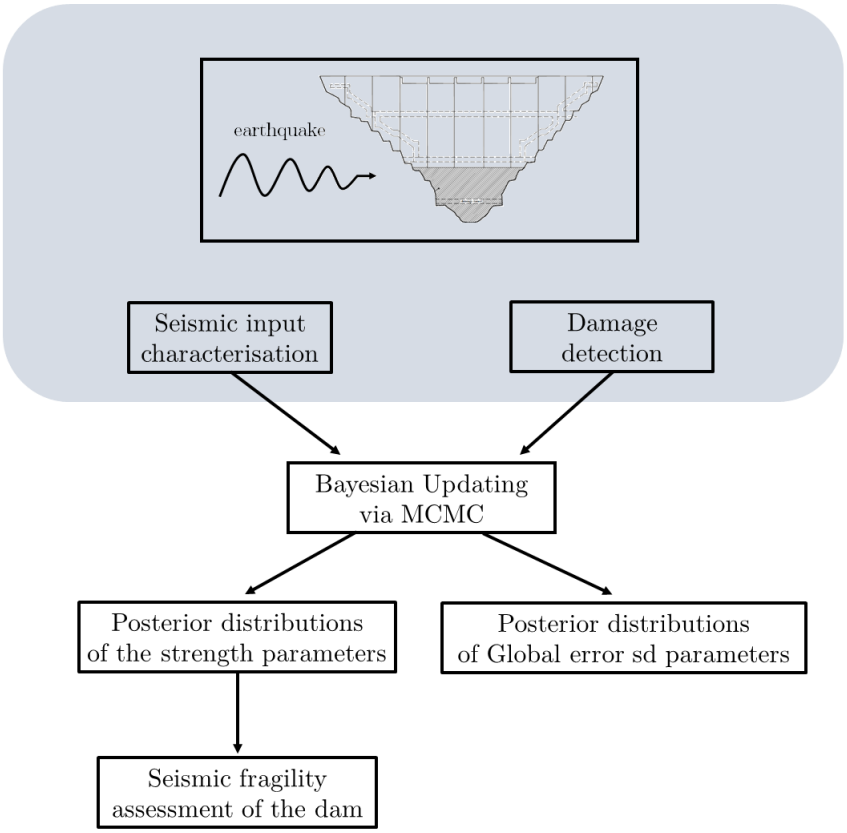


Fig. 7.3: Flowchart of the material constitutive models strength parameters updating.

7.5 Hyper-prior distributions definition

The probabilistic dynamic SHM proposed in this Chapter requires the definition of the hyper-prior distributions. In particular those of the mechanical parameters of the materials, namely $p\left(\boldsymbol{\theta}_{\mathrm{el}}, \boldsymbol{\Xi}_{\mathrm{el}}\right) = p\left(\boldsymbol{\theta}_{\mathrm{el}} | \boldsymbol{\Xi}_{\mathrm{el}}\right) p\left(\boldsymbol{\Xi}_{\mathrm{el}}\right)$ and $p\left(\boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}}\right) = p\left(\boldsymbol{\theta}_{\mathrm{s}} | \boldsymbol{\Xi}_{\mathrm{s}}\right) p\left(\boldsymbol{\Xi}_{\mathrm{s}}\right)$, can be defined starting from the results of material tests. Whereas there are usually no prior information about the parameters $\boldsymbol{\lambda}^{\mathrm{U}}$, $\mathbf{w}_d^{\mathrm{U}}$, $\boldsymbol{\lambda}^{\mathrm{D}}$ and $\mathbf{w}_d^{\mathrm{D}}$, thus for which non-informative prior distributions must be used.

7.6 Devices layout optimization through Optimal Bayesian Design of Experiment

The last part of the presented research works aim to define a robust procedure for designing the SHM layout. In particular, the one proposed in this Chapter is based on the Optimal Bayesian Experimental Design (Chapter 4.5) and it allows defining the devices position in order to maximize both the damage detection, and the acquisitions of information useful to infer the strength parameters of the materials. This double aim can be achieved by exploiting the relationship between damage development and strength parameters.
The predictive model of the damaged dam (Equation 7.4) and the relative likelihood function (Equation 7.5) are used within the optimization procedure. Also in this context, particular attention must be placed to guarantee the coherence of modes when Equation 4.113 is solved. Indeed, Equation 4.113 can be solved with a double Monte Carlo procedure for which the sampling from the likelihood function is needed both in the external and internal calculation of Monte Carlo estimator. Before the likelihood function construction, and in particular before the residual calculation, the MAC matrix is determined and the numerical modes reordered coherently with the experimental ones. More specifically, in this context the experimental modes are unknown a priori, so they are simulated within the algorithm and indicated as $\mathbf{y}$. The proposed algorithm is shown in algorithm 3.
An important aspect of the Optimal Bayesian Experimental Design is the parametrization of the design variable $\mathbf{d}$, as discussed in Chapter 4.5. Assuming that the devices can be installed only on the downstream face of the dam, the use of polar coordinates r and α is a convenient choice to parametrize the devices position, as shown in figure 7.4. However, the FEA solution is calculated only in few points of the design domain, which can be thought as the mesh nodes. Therefore, the gPCE is trained only in these solution points. An interpolation procedure must be set up within the numerical algorithm in order to ensure the continuity of the design domain. In this way once r and α are sampled, a candidate P of the devices position is selected. The quantity of interest can be determined by interpolating the values of the element nodes in which the candidate point lies. The definition of a maximum and minimum value of r and α is needed in order to limit the domain. Despite this parametrization, some areas can be outside the design domain. A control on the candidate point position is added in the numerical algorithm 3. If the candidate point P lies outside the design domain, it is projected onto the closest edge. It is noticeable that this projection procedure is possible thanks to the use of the SPSA algorithm, as discussed in Chapter 4.5.3.

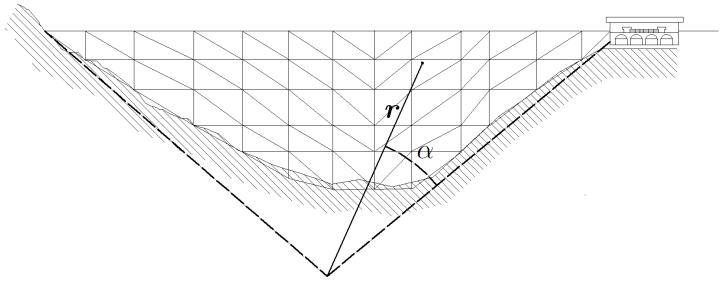


Fig. 7.4: Design variable **d** for the SHM system optimization.

The SPSA algorithm, described in Chapter 4.5.3, is used to maximize the expected utility, so to find the best position of a device. Since more than one device can be used, a sequential Bayesian optimal experimental design is set up. Once the device position in the i-th step is determined, the reference predictive model is changed in order to consider this last device. Therefore, in the next step, $U(\mathbf{d}_{i+1})$ is calculated considering a layout composed by the devices determined in the previous steps, plus the one sampled by the procedure in the current step $i+1$. In other words, once a device position is determined through SPSA the prior knowledge is updated to embodied the information related to the new device. This updated state of knowledge becomes the prior knowledge of the next step. The optimization of the SHM system is shown in figure 7.5.
The whole algorithm, implemented in MATLAB® R2017A (Mathworks 2016), is shown in algorithm 3. Despite the use of the gPCE to speed up the procedure, it still requires long calculation time because of the presence of a double loop within the optimization algorithm. Moreover, as discussed in Chapter 4.5.3, the number of steps of the outer and inner loops can not be reduced too much, because they control the bias and the variance of the expected utility. Therefore, the algorithm is parallelized exploiting the features of MATLAB® R2017A.

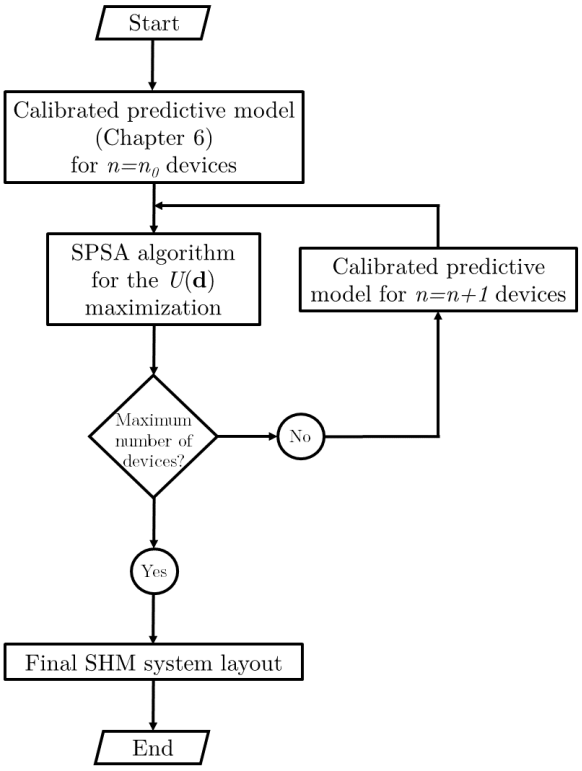


Fig. 7.5: Flowchart of the SHM system optimization procedure.

Algorithm 3 Double Monte Carlo estimator algorithm for OED

procedure OED $M, N, p(\boldsymbol{\theta}_{\mathrm{M}})$
  for $i = 1$ to M (max number of devices) **do**
    select the initial candidate point $\mathbf{d}_{i,0} = (r_{i,0}, \alpha_{i,0})$
    if $\mathbf{d}_{i,0}$ is inside allowed areas **then**
      $\mathbf{d}_{i,0} = \mathbf{d}_{i,1}$
    else
      $\mathbf{d}_{i,0}$ is projected onto the closest edge and that point becomes $\mathbf{d}_{i,1}$
    end if
    for $j = 1$ to N (max number of steps in the SPSA algorithm) **do**
      calculation of the SPSA parameters $a_j, c_j, \boldsymbol{\Delta}_j$
      $\mathbf{d}_{\mathrm{plus},j} = \mathbf{d}_{i,j} + c_j\boldsymbol{\Delta}_j$ and $\mathbf{d}_{\mathrm{minus},j} = \mathbf{d}_{i,j} - c_j\boldsymbol{\Delta}_j$
      if $\mathbf{d}_{\mathrm{plus},j}$ is inside allowed areas **then**
        $\mathbf{d}_{\mathrm{plus},j} = \mathbf{d}_{\mathrm{plus},j}$
      else
        $\mathbf{d}_{\mathrm{plus},j}$ is projected onto the closest edge and that point becomes $\mathbf{d}_{\mathrm{plus},j}$
      end if
      if $\mathbf{d}_{\mathrm{minus},j}$ is inside allowed areas **then**
        $\mathbf{d}_{\mathrm{minus},j} = \mathbf{d}_{\mathrm{minus},j}$
      else
        $\mathbf{d}_{\mathrm{minus},j}$ is projected onto the closest edge and that point becomes $\mathbf{d}_{\mathrm{minus},j}$
      end if
      for $k = 1$ to n_{out} (sample number in the outer Monte Carlo estimate) **do**
        draw $\boldsymbol{\theta}_{\mathrm{m}}^{k}$ from the prior distribution $p(\boldsymbol{\theta}_{\mathrm{m}})$
        calculation of the reference measure $\mathbf{y}^{(k)}$
        calculation of $\hat{\phi}_{h,i}^{\mathrm{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}^{k}, \mathbf{d}_{\mathrm{plus},j}\right)$
        MAC matrix calculation and mode shapes reordering
        calculation of the residual
        calculation $p\left(\mathbf{y}^{(k)} | \boldsymbol{\theta}_{\mathrm{m}}^{k}, \mathbf{d}_{\mathrm{plus},j}\right)$
        for $h = 1$ to n_{in} (sample number in the inner Monte Carlo estimate) **do**
          draw $\boldsymbol{\theta}_{\mathrm{m}}^{(k,h)}$ from the prior distribution $p(\boldsymbol{\theta}_{\mathrm{m}})$
          calculation of $\hat{\phi}_{h,i}^{\mathrm{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}^{(k,h)}, \mathbf{d}_{\mathrm{plus},j}\right)$
          MAC matrix calculation and mode shapes reordering
          calculation of the residual
          calculation $p\left(\mathbf{y}^{(k)} | \boldsymbol{\theta}_{\mathrm{m}}^{(k,h)}, \mathbf{d}_{\mathrm{plus},j}\right)$
        end for
        calculation of $p\left(\mathbf{y}^{(k)} | \mathbf{d}_{\mathrm{plus},j}\right)$
      end for
      $U\left(\mathbf{d}_{\mathrm{plus},j}\right)$ calculation

34: **for** $k = 1$ to n_{out} (sample number in the outer Monte Carlo estimate) **do**
35: draw $\boldsymbol{\theta}_{\text{m}}^{k}$ from the prior distribution $p\left(\boldsymbol{\theta}_{\text{m}}\right)$
36: calculation of the reference measure $\mathbf{y}^{(k)}$
37: calculation of $\hat{\phi}_{h,i}^{\text{D}}\left(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\text{s}}^{k}, \mathbf{d}_{\text{minus},j}\right)$
38: MAC matrix calculation and mode shapes reordering
39: calculation of the residual
40: calculation of $p\left(\mathbf{y}^{(k)} | \boldsymbol{\theta}_{\text{m}}^{k}, \mathbf{d}_{\text{minus},j}\right)$
41: **for** $h = 1$ to n_{in} (sample number in the inner Monte Carlo estimate) **do**
42: draw $\boldsymbol{\theta}_{\text{m}}^{(k,h)}$ from the prior distribution $p\left(\boldsymbol{\theta}_{\text{m}}\right)$
43: $\frac{\partial \hat{\boldsymbol{\phi}}\left(\mathbf{x}, \boldsymbol{\theta}_{\text{m}}^{(k,h)}, \mathbf{d}_{\text{minus},j}\right)}{\partial \boldsymbol{\theta}_{\text{m}}^{(k,h)}}$
44: MAC matrix calculation and mode shapes derivative selection
45: calculation of the residual
46: calculation of $p\left(\mathbf{y}^{(k)} | \boldsymbol{\theta}_{\text{m}}^{(k,h)}, \mathbf{d}_{\text{minus},j}\right)$
end for
47: calculation of $p\left(\mathbf{y}^{(k)} | \mathbf{d}_{\text{minus},j}\right)$
end for
48: $U\left(\mathbf{d}_{\text{minus},j}\right)$ calculation
49: $g_{i,j}\left(\mathbf{d}_{i,j}\right) = \frac{U^{-1}\left(\mathbf{d}_{\text{plus},j}\right) - U^{-1}\left(\mathbf{d}_{\text{minus},j}\right)}{2c_j \boldsymbol{\Delta}_j}$
50: $\mathbf{d}_{i,j} = \mathbf{d}_{i,j} - a_j g_{i,j}\left(\mathbf{d}_{i,j}\right)$
end for
51: devices set $\mathbf{d}_{\text{set}} = \left[\mathbf{d}_{\text{set}}, \mathbf{d}_{i,j}\right]$
end for
52: **end procedure**

Chapter 8

The case studies

8.1 Introduction

In this chapter two real Italian concrete gravity dams are described and analysed, applying the procedures developed in the present research work. The choice of studying two different cases stems from the need to use real measurements. In addition to material test results, static measurements are available for the Gramolazzo dam, while dynamic records are available for the Scandarello dam.
The Gramolazzo dam is used as benchmark to test the static SHM framework. In particular, both the *Updating Phase* and the *Detection Phase* are applied.
The results of an experimental campaign with the aim of determining the modal characteristics of the system through ambient vibrations are available for the Scandarello dam. Therefore, it has been chosen as case study for the application of the dynamic SHM framework described in Chapters 6 and 7 and the optimization procedure proposed at the end of Chapter 7. The application of the procedure proposed in Chapter 6 allows updating the dynamic parameters of the FE model and performing a dam twin model. The updated model is used both to predict the structural behaviour during seismic events through the fragility curves calculation and to control the health state of the structure.

Finally, the Scandarello dam is also used as case study for the application of the procedure proposed in chapter 7. The monitoring system based on ambient vibrations able to detect damage and to acquire useful information for the strength parameters updating is designed.
It is worth noting that since there are no information about the static or dynamic behaviour of the damaged dams, they have been simulated through calibrated high-fidelity models.

8.2 Static SHM for concrete gravity dams: the case of Gramolazzo dam

8.2.1 Introduction

In this section the numerical results related to the study of the Gramolazzo dam are shown. The static SHM framework presented in Chapter 5 is applied, using the results of material tests, environmental measurements and static displacements as source of information.
Dam displacements recorded during the regular use of the dam are small and their variation is very slow, then materials can be assumed as liner elastic, as previously discussed in Chapter 5. Therefore, the elastic parameters of both dam concrete and foundation soil, namely the bulk modulus K and the shear one G, must be update. In this application concrete parameters are indicated with the subscript C, while those of the foundation soil with the subscript G.
The first step of the procedure is the training of the meta model needed to reproduce the static behaviour of the dam. Displacements due to the basin level variation are approximated through the gPCE (Chapter 5.2), while those related to the thermal variation are approximated by using a Fourier analysis (Chapter 5.3). According to the notation introduced in Chapter 5, the parameters can be indicated as $\theta_{\mathrm{gPCE},K_C} = K_C$, $\theta_{\mathrm{gPCE},K_G} = K_G$, $\theta_{\mathrm{gPCE},G_C} = G_C$, $\theta_{\mathrm{gPCE},G_G} = G_G$. Therefore, the vector $\boldsymbol{\theta}_{\mathrm{gPCE},K} = [\theta_{\mathrm{gPCE},K_C}, \theta_{\mathrm{gPCE},K_G}]^T$ and $\boldsymbol{\theta}_{\mathrm{gPCE},G} = [\theta_{\mathrm{gPCE},G_C}, \theta_{\mathrm{gPCE},G_G}]^T$. Whereas, the dimension of the vectors $\boldsymbol{\theta}_{\mathrm{FA},a}$ and $\boldsymbol{\theta}_{\mathrm{FA},b}$ (Chapter 5.3) of the Fourier coefficients is related to the number of harmonics assumed in the expansion. In this application, the number of harmonics is varied in order to asses its influence in the procedure result.
A high-fidelity model is performed in order to simulate the dam behaviour after damage.

8.2.2 Dam description

The Gramolazzo dam is a large dam located in north-centre Italy. Daily environmental measurements, such as temperatures and basin levels, recorded displacements at different heights and the results of material tests, are available. With a total length of 96 m, the dam is made up of 11 monoliths with a maximum height of 37 m from the base and a downstream slope of 0.75. The structure, whose shape in plan is perfectly straight, is inserted in a canyon with monoliths separated by trapezoidal key joints. The shape of the vertical joints suggests a complete and monolithic 3D model.

8.2.3 Static and environmental measurements

The displacements are recorded by two inverted pendulums, one is installed in the central spillway monolith and the other is installed on the non-overflow monolith as shown in figure 8.1. This latter pendulum, fixed at the rock foundation at 534 m a.s.l., acquires displacements in both up-stream/down-stream and cross-valley directions at three different heights: the dam crest at 604 m a.s.l. (P3), the upper tunnel at 593 m a.s.l. (P4) and the lower tunnel at 574 m a.s.l. (P5). Only P4 displacements, shown in figure 8.2, are considered in the present application. Vertical joints opening-closing displacements are monitored by removable mechanical strain gauges located inside the two horizontal tunnels and shown in figure 8.1.

The highest mean annual variations of the upstream-downstream displacements hover around 3.5 mm at the dam crest, whereas the lowest mean annual variations are of about 0.5 mm in the lower tunnel. The opening - closing displacements of contraction joints in both tunnels show average annual oscillations that are well correlated and hence with water and air thermal variations.

In this work the displacements of the control point P04 are used as QI in the procedure proposed in Chapter 5. Figure 8.2 shows the displacements of the point P04 in the period from 2004 to 2011. Environmental measurements, such as basin level variation, water and air temperatures, recorded in the same period, are also available, as shown in figure 8.3 and 8.4. Finally, the observation recorded in the period 2004-2008 are used to calibrate the model parameters (Chapter 5), while those registered from 2008 to 2011 are used to validate the calibrated probabilistic model of dam displacements.

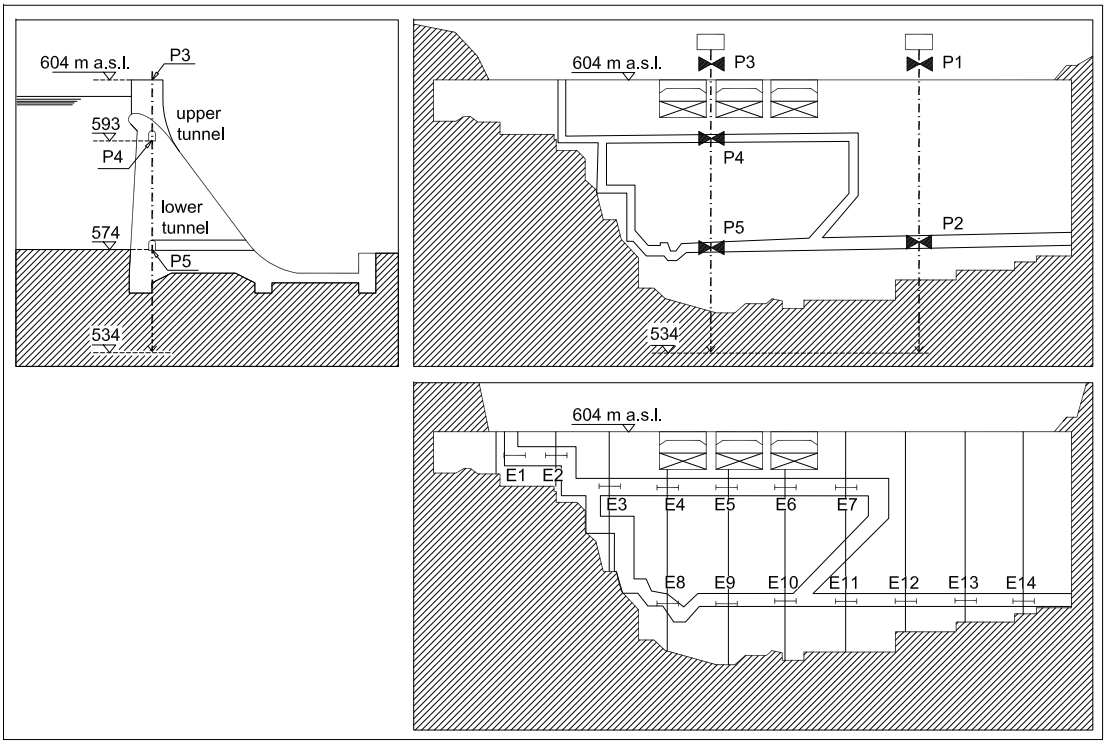


Fig. 8.1: Static monitoring system layout of the Gramolazzo dam.

8.2.4 Material tests results

Several experimental campaigns have been conducted on the Gramolazzo dam. Usually, during experimental campaigns both in-situ and laboratory tests are conducted, with the aim to determine the material characteristics. In the case of Gramolazzo dam both dam concrete and foundation soil have been investigated, through laboratory tests as compression ones and in situ tests as cross-hole. The results, in terms of Young's modulus E and Poisson's coefficient ν, are synthesized in table 8.1, where the parameters with subscript C are related to the dam concrete, while those with the subscript G are related to the soil ground.

Table 8.1: Gramolazzo dam: results of the material characterisation campaign.

	E_C [MPa]	ν_C	E_G [MPa]	ν_G
Mean	17543	0.176	29653	0.313
s.d.	4870	0.072	18873	0.075

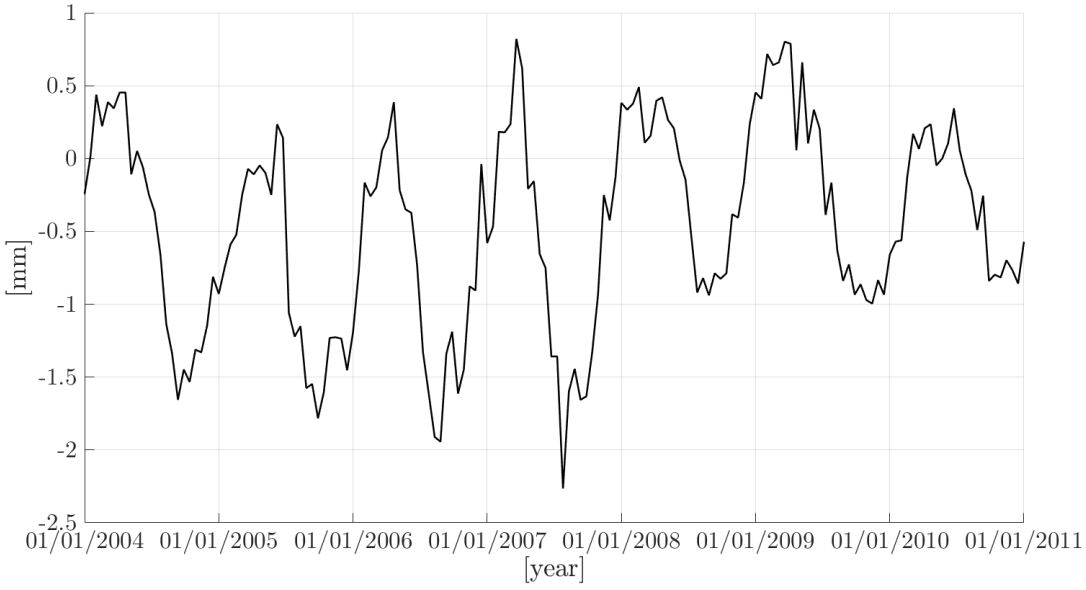


Fig. 8.2: Displacements of the point P4, recorded on the Gramolazzo dam.

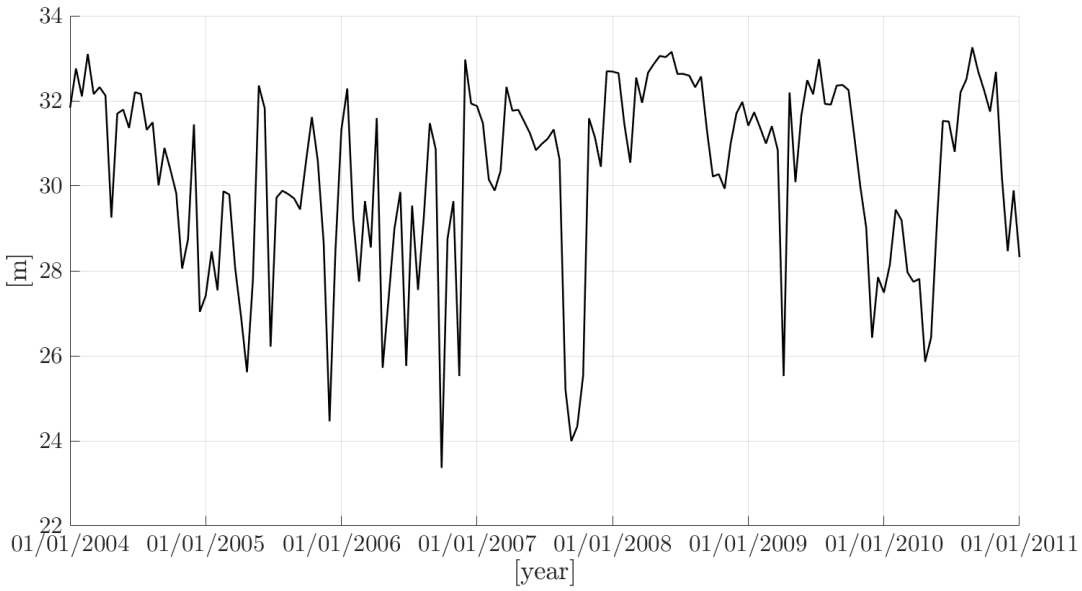


Fig. 8.3: Basin level variation recorded by the monitoring system of the Gramolazzo dam.

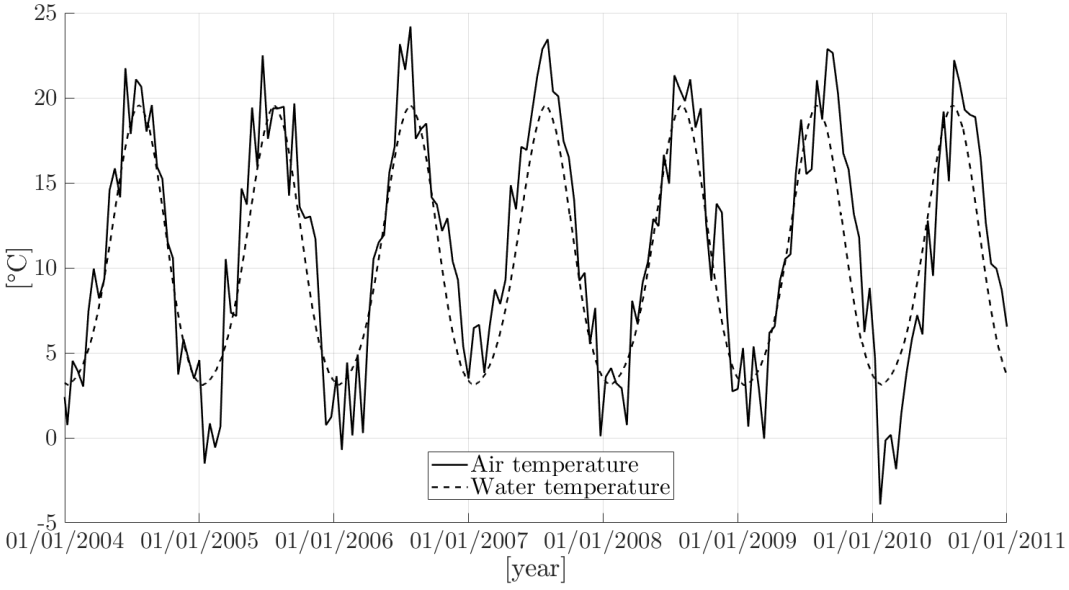
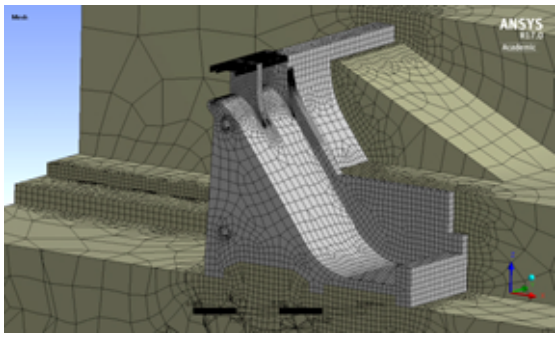


Fig. 8.4: Air and water temperatures recorded by the monitoring system of the Gramolazzo dam.

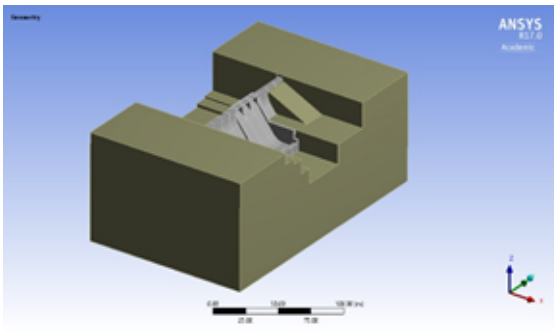
The parameters indicated in table 8.1 are used to define the prior distribution of the model parameters with physical meaning, as indicated in paragraph 8.2.6.

8.2.5 FE Models description

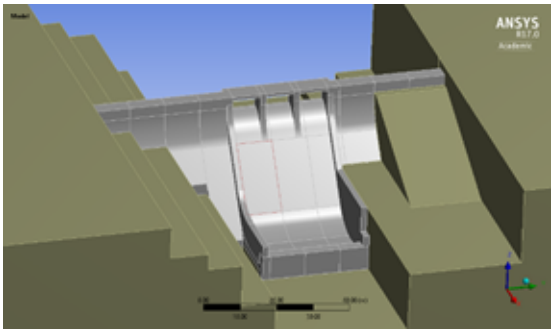
Two FE models of the Gramolazzo dam are performed: a first model, to train the gPCE for hydrostatic displacements, and a high-fidelity model used both in static and dynamic analysis, to simulate the dam behaviour after damage due to seismic events.
The geometry of the structure and the shape of the ground are built in a CAD program, based on the original drawings of the dam and the orographic map of the region.
The first model, used only in static analyses, is performed by importing the geometry in ANSYS r.17 (ANSYS 2013). The mesh is composed by 231254 quadratic hexahedral elements, for a total of 870787 nodes, as shown in figure 8.5. Hertzian bonded contacts are introduced between the monoliths and the materials, concrete and rock, are assumed to be homogeneous linear and elastic.
Since there are no information about the damaged behaviour of the dam, it is simulated through a high-fidelity model. Starting from the geometry, previously imported in a CAD



(a) Model mesh.



(b) Global view of the model.

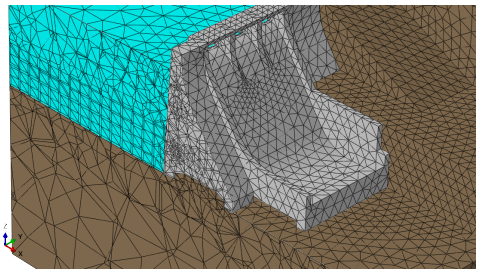


(c) Detail of the dam.

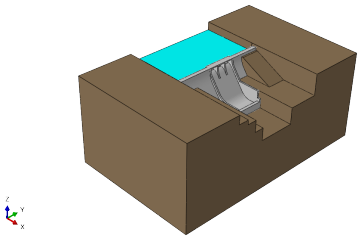
Fig. 8.5: Static FE model of the Gramolazzo dam.

program, the high-fidelity model is built in ABAQUS 6.14 (ABAQUS 2014). In particular, the SSI is modelled by considering a foundation with mass and by using Infinite Elements as boundary conditions in order to account for the unboundedness of the soil, described in Chapter 3.3. Whereas, the FSI is considered by modelling the basin with fluid elements and by using low reflecting boundary conditions to account for the unboundedness of the basin, as explained in Chapter 3.4. Figure 8.6 shows the high-fidelity FE model of the Gramolazzo dam.

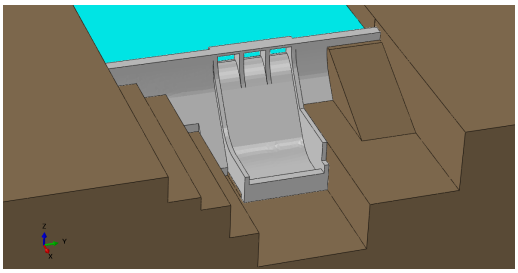
Since this model is used in dynamic analysis with seismic input, the minimum mesh size is defined in order to be one tenth of the smallest elastic wavelength, considering a significant frequency range of the seismic input. Therefore, the mesh is composed by 40500 quadratic tetrahedral mechanical elements (type C3D10), used for the dam concrete and the foundation soil, 9506 quadratic tetrahedral fluid elements (type AC3D10), used for the basin, and 309 linear hexahedral elements (type CIN3D8) as infinite elements. Since the dam has vertical contraction key joints, the connections between monoliths can be modelled as bonded.



(a) Model mesh.



(b) Global view of the model.



(c) Detail of the dam.

Fig. 8.6: High-fidelity FE model of the Gramolazzo dam.

The foundation soil is assumed a homogeneous isotropic linear elastic material, while the dam concrete is modelled through a homogeneous isotropic plastic damage constitutive model which has been proposed by Lee and Fenves 1998b and described in Chapter 3.2.3. This constitutive material model requires the definition of the strength parameters, those related to the damage behaviour and those related to the yield surface. The results of the material tests allows determining the compressive and tensile strengths of the dam concrete, indicated in table 8.2.

Table 8.2: Compressive and tensile strengths of the Gramolazzo dam.

$f_{c,k}$ [MPa]	$f_{t,k}$ [MPa]
32.5	1.45

The strength parameters are used to define the post elastic behaviour both in tension and compression. In particular, the compressive branch is composed of three parts. By defining $f_{c,0} = 0.85 f_{c,k}$ as the value of the compressive stress beyond which the material starts to show a non linear behaviour, $f_{c,u} = 0.3 f_{c,k}$ as the value of the compressive stress beyond which the strength material decreases exponentially, and respectively characterised by the deformation values $\epsilon_{c,0} = f_{c,0}/E$ and $\epsilon_{c,u} = 0.005$, and $f_{c,r} = 0.1 f_{c,k}$ as the residual value

of strength, the constitutive law can be defined as:

- *Hardening branch*: the branch where $\epsilon_{c,0} < \epsilon \leq \epsilon_{c,k}$,

$$\sigma_c(\epsilon_c) = f_{c,0} + (f_{c,k} - f_{c,0}) \sqrt{\frac{2(\epsilon_c - \epsilon_{c,0})}{\epsilon_{c,k} - \epsilon_{c,0}} - \frac{(\epsilon_c - \epsilon_{c,0})^2}{(\epsilon_{c,k} - \epsilon_{c,0})^2}}. \tag{8.1}$$

- *Softening branch 1*: the branch where $\epsilon_{c,k} < \epsilon \leq \epsilon_{c,u}$,

$$\sigma_c(\epsilon_c) = f_{c,k} - (f_{c,k} - f_{c,u}) \left(\frac{\epsilon_c - \epsilon_{c,k}}{\epsilon_{c,u} - \epsilon_{c,k}} \right)^2. \tag{8.2}$$

- *Softening branch 2*: the branch where $\epsilon > \epsilon_{c,u}$,

$$\sigma_c(\epsilon_c) = f_{c,r} + (f_{c,u} - f_{c,r}) \exp\left(2 \frac{f_{c,u} - f_{c,k}}{\epsilon_{c,u} - \epsilon_{c,k}} \frac{\epsilon_c - \epsilon_{c,u}}{f_{c,u} - f_{c,r}} \right). \tag{8.3}$$

Where $\epsilon_{c,k} = 0.002$ is the deformation value for which σ_c is equal to $f_{c,k}$. The resulting behaviour in compression is shown in figure 8.7.

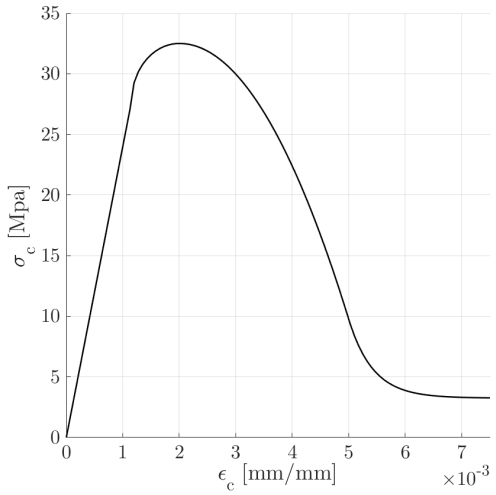


Fig. 8.7: Compressive constitutive behaviour of the Gramolazzo dam concrete.

The tensile behaviour of the dam concrete follows an exponential law after the elastic branch, by defining the constitutive law in terms of tensile stress σ_t and displacements u_t,

in order to avoid the mesh dependency, it can be written as

$$\sigma_t\left(u_t\right) = \alpha_t \exp\left(-\beta_t u_t\right). \tag{8.4}$$

In equation 8.4 α_t and β_t are the parameters which define the shape of the exponential branch, they are calibrated in order to obtain a fracture energy (Hillerborg et al. 1976) equal to 150 N/m, as usually indicated in literature (Lee and Fenves 1998b).

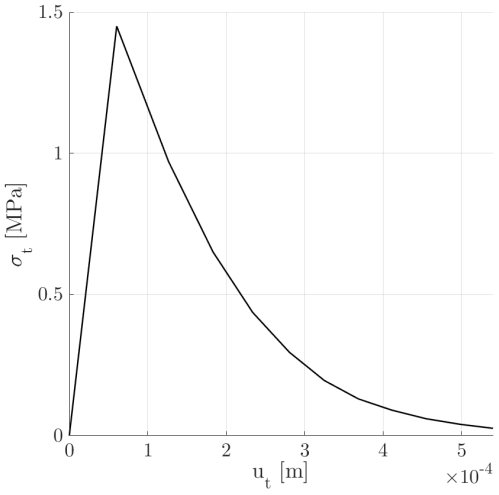


Fig. 8.8: Tensile constitutive behaviour of the Gramolazzo dam concrete.

The Concrete Damage Plasticity constitutive model considers the effect of the damage as a degradation of the undamaged elastic tensor $\mathbb{C}$, as reported in Chapter 3.2.3. Therefore, the definition of the law of the tension scalar degradation parameters d is needed. In this application the law of d_t is defined through a quadratic function to approximate the d_t behaviour. The relationship plotted in figure 8.9, between d_t and the cracking displacements u_t^{ck}, is adopted.
Finally, the parameters of the yield criteria, introduced in Chapter 3.2.3, are defined according to the values which can be found in the literature (Omidi et al. 2013), since no information about them was available. They are reported in table 8.3.

Table 8.3: Yield criteria parameters of the Gramolazzo dam concrete.

ψ [°]	ε	σ_{b0}/σ_{c0}	K_c
36.31	0.1	1.16	0.66

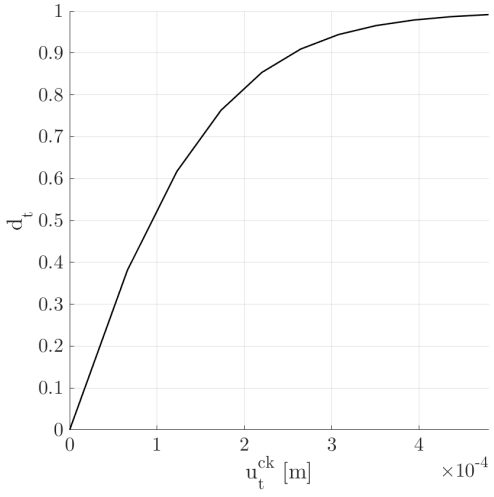


Fig. 8.9: Tensile damage law of the Gramolazzo dam concrete.

The Rayleigh's damping model has been used in the dynamic analyses and it has been calibrated in order to have a damping factor equal to 5% with regard to the first and second mode of the dam.

8.2.6 Prior distributions of the model parameters

The mechanical parameters to be identified are the bulk modulus $\boldsymbol{\theta}_{\mathrm{gPCE},K}$ and the shear modulus $\boldsymbol{\theta}_{\mathrm{gPCE},G}$ of both the concrete and foundation rock. The results of the in-situ and laboratory tests, reported in table 8.1, are given in terms of Young modulus E and Poisson's coefficient ν. Therefore, the well known equations

$$\begin{aligned} K &= \frac{E}{3\,(1-2\nu)} \\ G &= \frac{E}{2\,(1+\nu)}, \end{aligned} \tag{8.5}$$

are used to determine the prior distributions of the mechanical parameters of both materials in terms of K and G. Log-normal distributions are used for K and G and their parameters are shown in table 8.4. The prior distributions of the other parameters with no physical meaning, such as the combination coefficients of the Fourier series $\boldsymbol{\theta}_{\mathrm{FA}}$, the combination coefficient of the drift term θ_{K}, and the error standard deviation σ, are defined as non-informative (Chapter 5.6).

Table 8.4: Gramolazzo dam: Prior distributions of the mechanical parameters.

	K_C [MPa]	G_C [MPa]	K_G [MPa]	G_G [MPa]
Distr.	Log-Normal	Log-Normal	Log-Normal	Log-Normal
Mean	9024	7458	26429	11292
s.d.	5028	2073	16352	7608

8.2.7 Proxy model for hydrostatic displacements

In order to build the proxy model of the structure with the gPCE four random variables: θ_{gPCE,K_C}, θ_{gPCE,K_G}, θ_{gPCE,G_C}, θ_{gPCE,G_G}, and the deterministic variable of the water level are selected. Three trials are performed by changing the polynomial expansion degree in order to check the error and to find the best compromise between solution accuracy and computational burden.

The gPCE of the dam displacement is trained by using the results of the simplest static FE model, previously described. Figure 8.10 shows the maximum relative error in terms of different displacements between FEM and the proxy model, by varying the expansion degree. A polynomial degree of three allows obtaining the best agreement between the number of analyses and the error value. The maximum error is less than 1% of the displacement which is considered acceptable. The number of analyses to build the proxy model, depending on the selected random variables and on the degree of the polynomial expansion, is 1024.

One of the results of the gPCE is the response surfaces. Figure 8.11a shows displacement due to the hydrostatic load by varying the bulk moduli K of both materials. It can be observed that the variation of the concrete elastic modulus provides a strong variation of the displacement, i.e. for a selected value of K_G the result ranges between 0.5 and 0.7 mm. Finally, the response surface involving K_C and G_C is shown in figure 8.11b.

The sensitivity analysis is a fundamental task, both to recognize the most influential model parameters and to asses the model validity (Chapter 4.4.4). Since gPCE is used to build the meta model, Sobol's coefficients are calculated, figure 8.12. Since the basin level is considered as deterministic variable in the gPCE, then the Sobol's coefficients account also for this parameter, which is the most influential one. This is the reason why the values of the coefficients shown in figure 8.12 are so low, so they must be analysed in a comparative way.

The most influential parameters on the model output are the bulk modulus of both materials θ_{gPCE,K_C} and θ_{gPCE,K_G}.

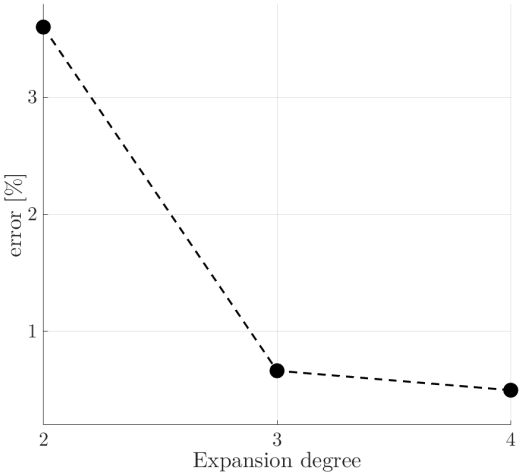


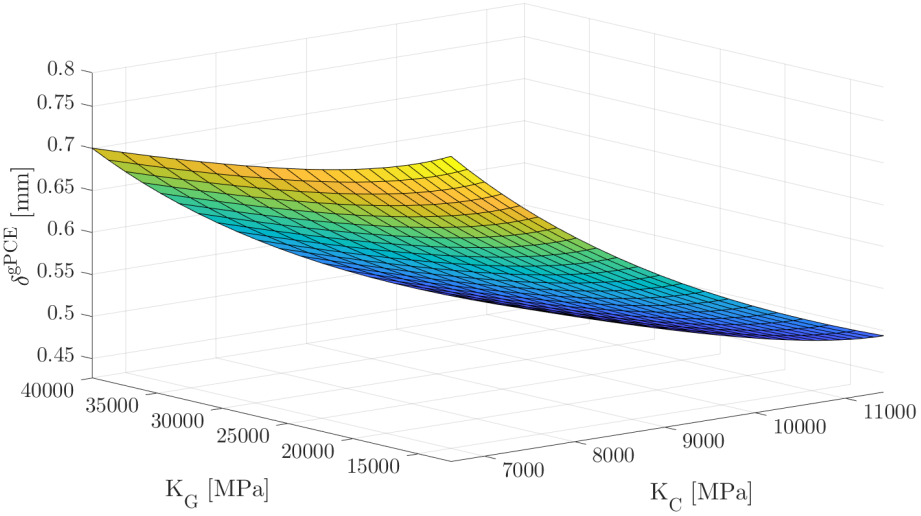
Fig. 8.10: Maximum relative gPCE error versus expansion degree of the hydrostatic displacements meta model.

8.2.8 Proxy model for thermal displacements

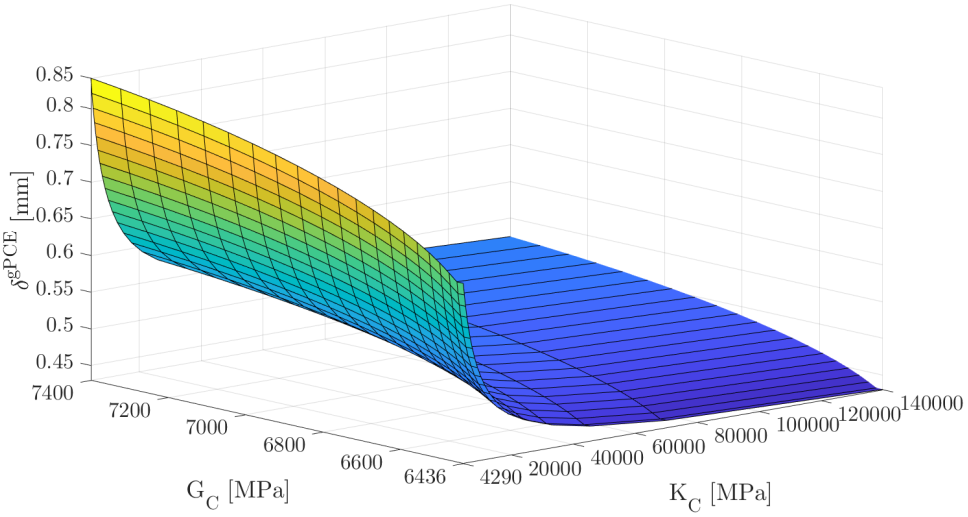
Three analyses are performed with different numbers of harmonics in order to evaluate the error related to the degree of expansion of the Fourier series. The number of harmonics considered in the first analysis are two, three in the second and four in the last analysis. The combination coefficients of the Fourier series are treated as random variables and collected in the vector $\boldsymbol{\theta}_{\mathrm{FA}}$.

8.2.9 Bayesian updating and posterior statistics

Bayesian updating of the probabilistic model is performed using the displacement data of the point placed in the upper part of the central spillway monolith (figure 8.2), together with the data of water and air temperatures and basin levels (figures 8.3 and 8.4), recorded from 2004 to 2008. The results of the updating procedure for the three analyses are reported in tables 8.5, 8.6 and 8.7. The posterior mean values of the bulk modules of concrete and ground soil, K_C and K_G, have a very small variance. The reason for this can be found in the response surface in figure 8.11 and in the Sobol's coefficients in figure 8.12, which shows the stronger influence of K_C and K_G on the dam's response in comparison to the shear modules of both materials. Moreover, the posterior mean values of θ_1 and θ_{K} have a small variation in every case, proving that they are little influenced by the variation in the harmonics number.



(a) Model output response surface versus K_C and K_G for a basin level equal to 30 m.



(b) Model output response surface versus K_C and G_C for a basin level equal to 30 m.

Fig. 8.11: Response surfaces of the hydrostatic displacements meta model.

Regarding the combination coefficients of the Fourier series, the mean values of the first two coefficients of the first harmonic are higher with respect to the others and a small variation is observed in the three cases. In figure 8.13, prior and posterior distributions

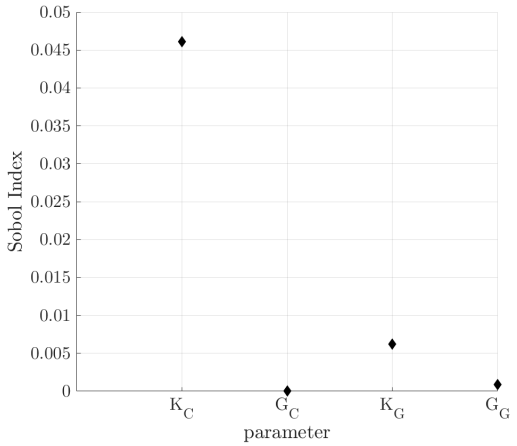


Fig. 8.12: Hydrostatic displacements Sobol's coefficients.

of the model parameters are represented in the case of 2, 3 and 4 harmonics. One can observe that the mean values of the error standard deviations σ are around 0.039 in all cases. The substantial invariance of σ suggests that there are many possible solutions in terms of posterior distributions for different choices related to the Fourier series order. The mean values of the output parameters are different for each case, but they do not exceed 10% of their average value. Since no further references are available to calibrate the result, a second order Fourier series seems to be the best choice to simulate the thermal displacements of the structure, while adding only two random variables to the probabilistic problem. Moreover, for practical reasons, engineers usually work with E and ν, so the comparison between prior and posterior distributions is reported also with regard to these parameters. Figure 8.14 shows the results in terms of E and ν, in accordance to the results indicated in figure 8.13 the mean values of the parameters don't show notable changes, while the standard deviation values are strongly reduced.
The values of $\hat{R}^p$ varying between 1 and 1.4 endorse the convergence of the analyses, as reported in paragraph 4.3.6. Figure 8.15 shows both recorded and calculated displacements. One can note the agreement between them for every value of the Fourier series order.
Figures 8.16 shows the relationship between recorded and calculated displacements at every step of the procedure. The crosses indicate the displacements calculated by assuming material characteristics from experimental tests, while empty circles represent the updated values of the material characteristics. Finally, filled circles indicate displacements when thermal effects and introducing explanatory functions are considered. Dashed lines represent the standard deviation. Figure 8.16 reports the results obtained for two and four

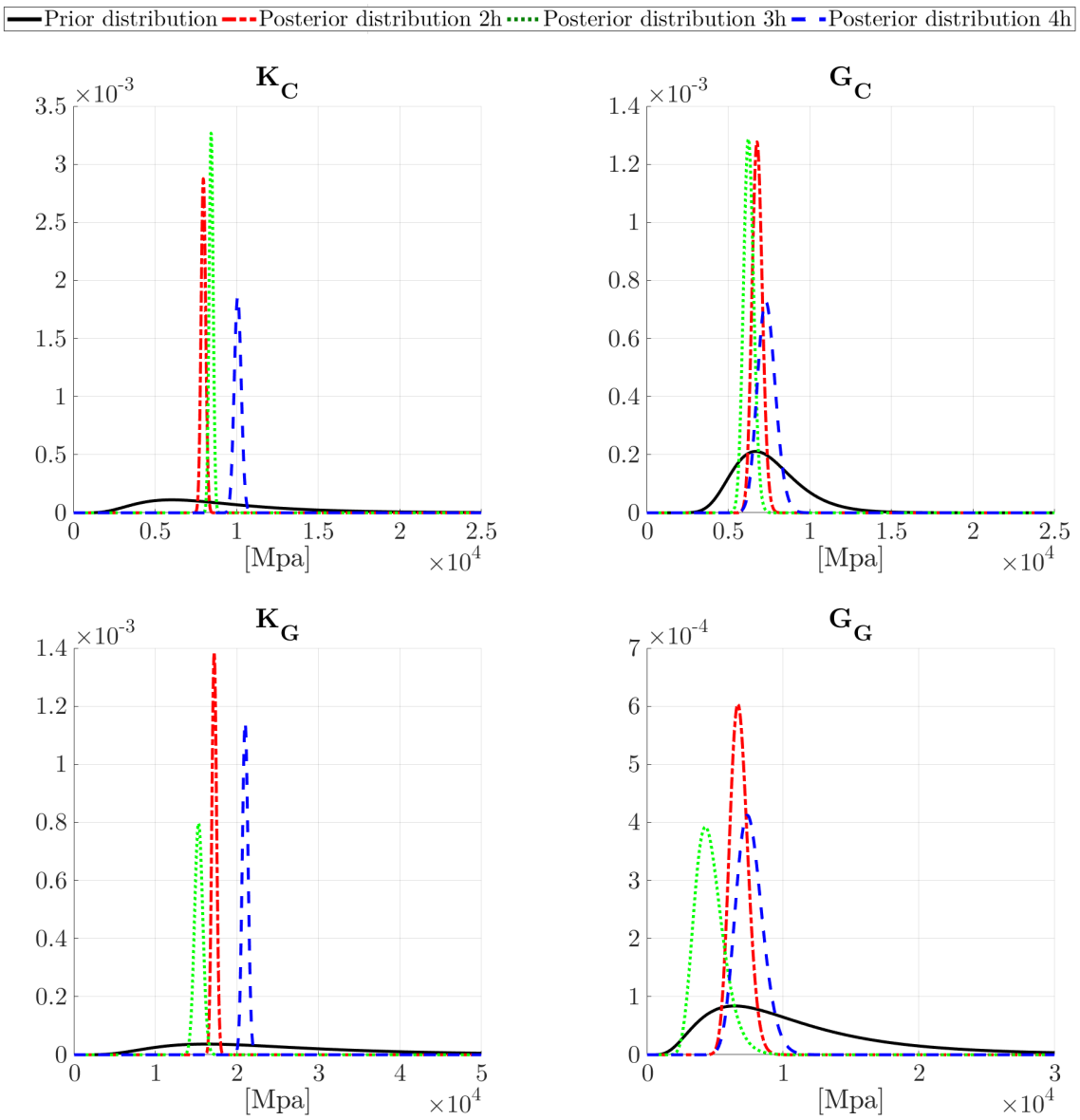


Fig. 8.13: Updating procedure in static case, comparison between prior and posterior distributions of K and G.

harmonics of the Fourier series for simulating thermal displacements. At the end of the procedure, the agreement between recorded and calculated displacements is notable in both cases.

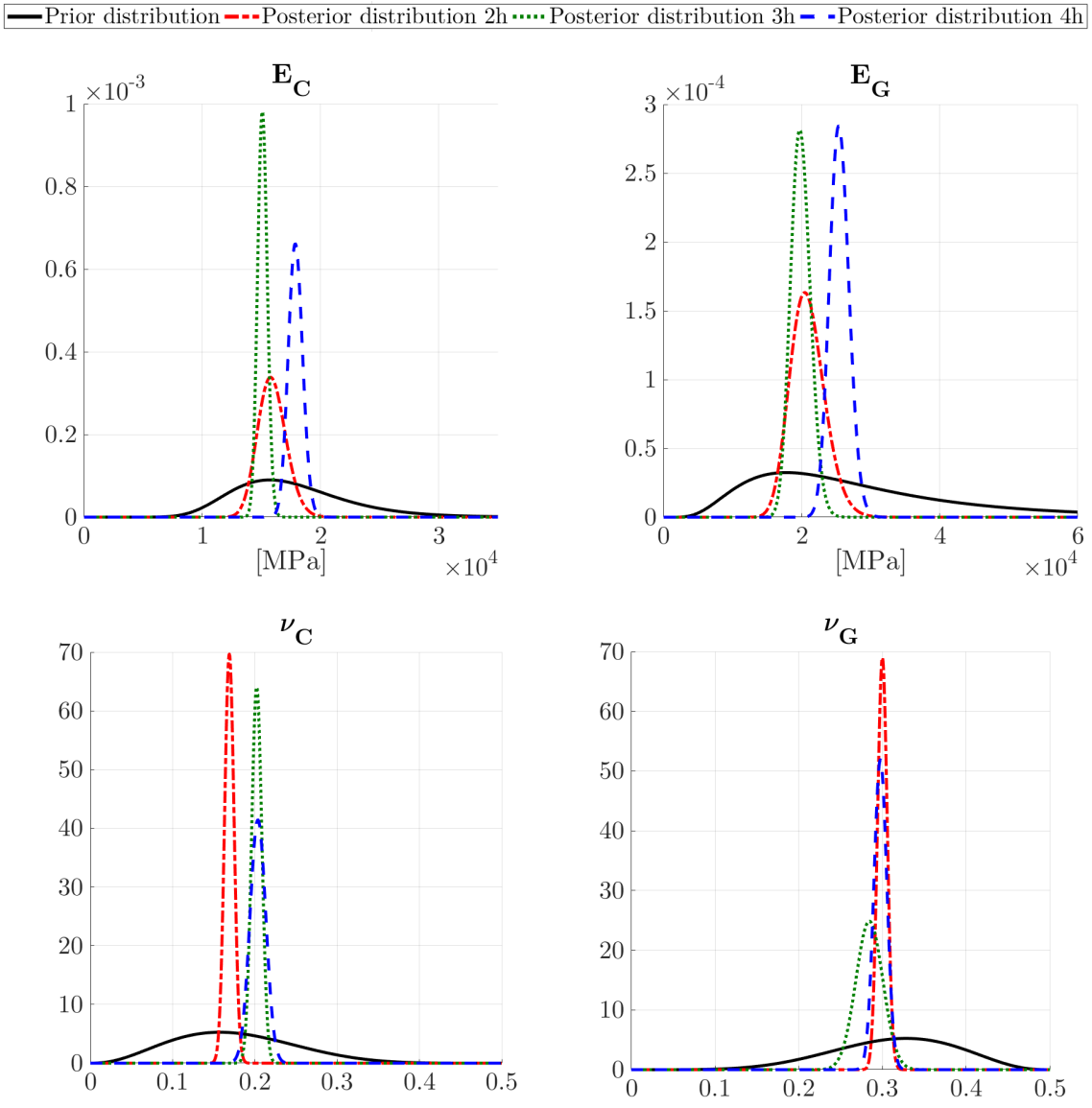


Fig. 8.14: Updating procedure in static case, comparison between prior and posterior distributions in terms of E and ν.

8.2.10 Procedure validation

In this section, the predictions of the dam displacements predictive model and those recorded on the dam body are compared in order to asses the effectiveness of the calibrated model.

The calibrated model with only two harmonics is used to predict the crest displacement

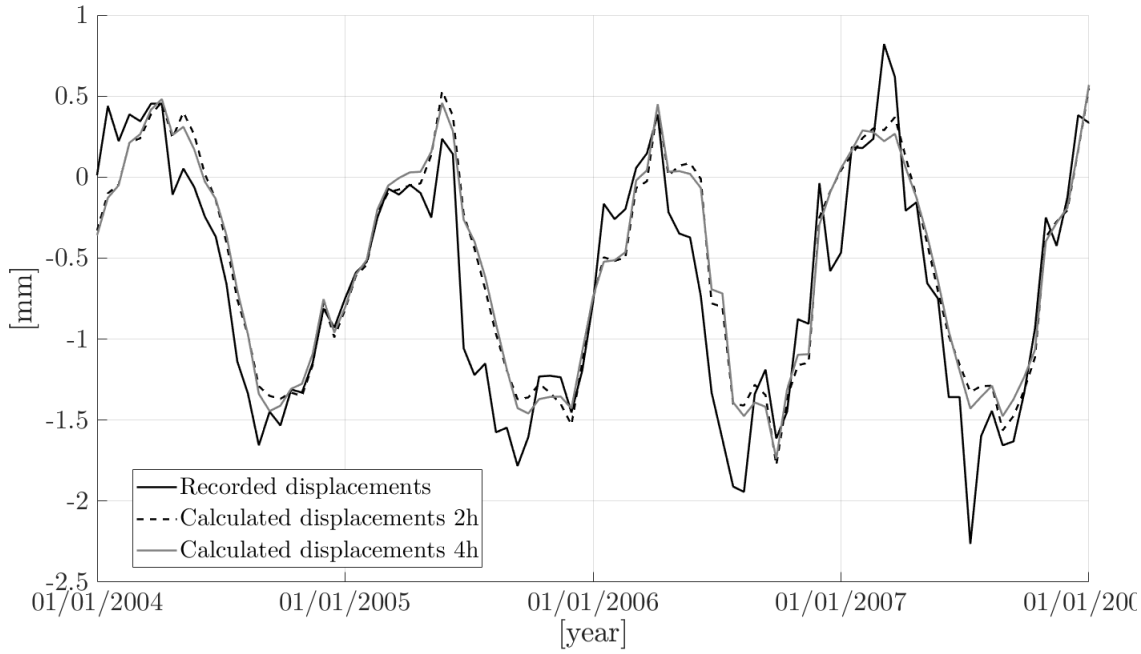


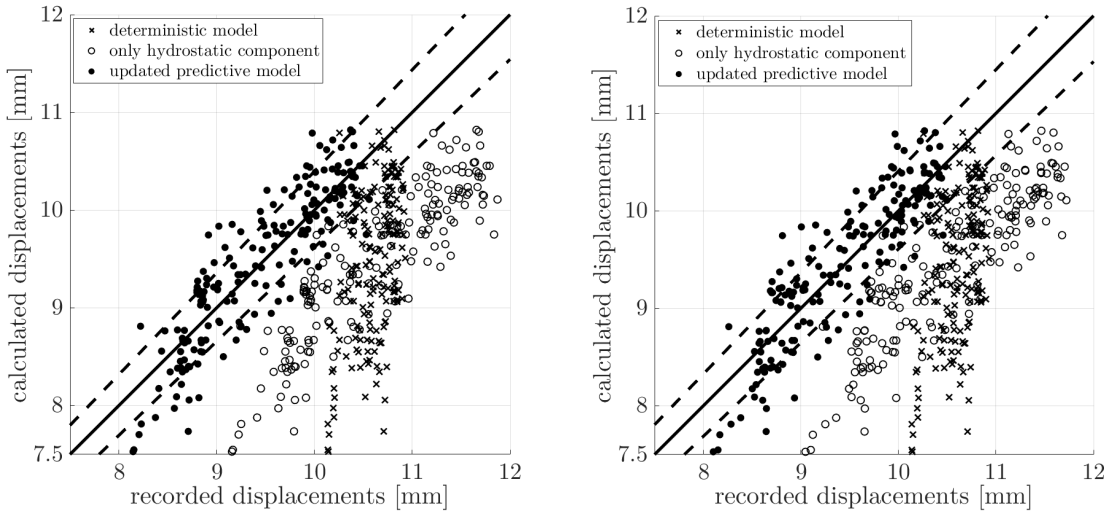
Fig. 8.15: Updating procedure in static case, comparison between recorded and calculated displacements.

Table 8.5: Updating procedure in static case, posterior statistics 2 harmonics.

	$\theta_{\mathrm{gPCE},K_C}$ [MPa]	$\theta_{\mathrm{gPCE},G_C}$ [MPa]	$\theta_{\mathrm{gPCE},K_G}$ [MPa]	$\theta_{\mathrm{gPCE},G_G}$ [MPa]	θ_1	θ_{FA,a_1}
Mean	7960	6788	17206	7987	-0.117	0.837
s.d.	20.14	469.23	33.89	5057	0.006	0.045
	θ_{FA,b_1}	θ_{FA,a_2}	θ_{FA,b_2}	θ_{K}	σ	
Mean	0.417	-0.102	-0.117	$3.368 * 10^{-10}$	0.039	
s.d.	0.041	0.045	0.038	$1.342 * 10^{-12}$	0.002	

in the period from 2008 to 2011. Such predictions are compared with the displacements recorded by the monitoring system. Results are illustrated in figure 8.17, they show a good agreement between recorded and calculated displacements. Recorded displacement curve almost totally falls in the interval between the prediction plus and minus standard deviation of the probabilistic model.

The calibrated probabilistic model for the dam displacement could be used for a real time structural control, thanks to its calculation speed. The proposed model allows evaluat-



(a) Thermal displacements calculated with 2 harmonics.

(b) Thermal displacements calculated with 4 harmonics.

Fig. 8.16: Updating procedure in static case, effects of the updating procedure in terms of residual.

Table 8.6: Updating procedure in static case, posterior statistics 3 harmonics.

	θ_{gPCE,K_C} [MPa]	θ_{gPCE,G_C} [MPa]	θ_{gPCE,K_G} [MPa]	θ_{gPCE,G_G} [MPa]	θ_1	θ_{FA,a_1}	θ_{FA,b_1}
Mean	8437	6275	15302	7722	-0.121	0.841	0.443
s.d.	17.28	629.13	61.64	10687	0.003	0.052	0.046
	θ_{FA,a_2}	θ_{FA,b_2}	θ_{FA,a_3}	θ_{FA,b_3}	θ_{K}	σ	
Mean	-0.112	-0.104	0.055	-0.037	$2.974 * 10^{-10}$	0.039	
s.d.	0.047	0.039	0.050	0.0391	$9.875 * 10^{-13}$	0.002	

Table 8.7: Updating procedure in static case, posterior statistics 4 harmonics.

	θ_{gPCE,K_C} [MPa]	θ_{gPCE,G_C} [MPa]	θ_{gPCE,K_G} [MPa]	θ_{gPCE,G_G} [MPa]	θ_1	θ_{FA,a_1}	θ_{FA,b_1}	θ_{FA,a_2}
Mean	10067	7425	21008	9808	-0.110	0.831	0.436	-0.108
s.d.	30.44	1154	41.18	8247	0.004	0.038	0.045	0.039
	θ_{FA,b_2}	θ_{FA,a_3}	θ_{FA,b_3}	θ_{FA,a_4}	θ_{FA,b_4}	θ_{K}	σ	
Mean	-0.104	0.054	-0.039	-0.018	0.017	$3.422 * 10^{-10}$	0.040	
s.d.	0.042	0.040	0.043	0.039	0.034	$1.233 * 10^{-12}$	0.002	

ing the global error standard deviation, which can be used as starting point to define a threshold beyond which the dam behaviour is considered anomalous. In the next section, the behaviour of the damaged dam is simulated in order to show how use the calibrated predictive model for structural control purpose.

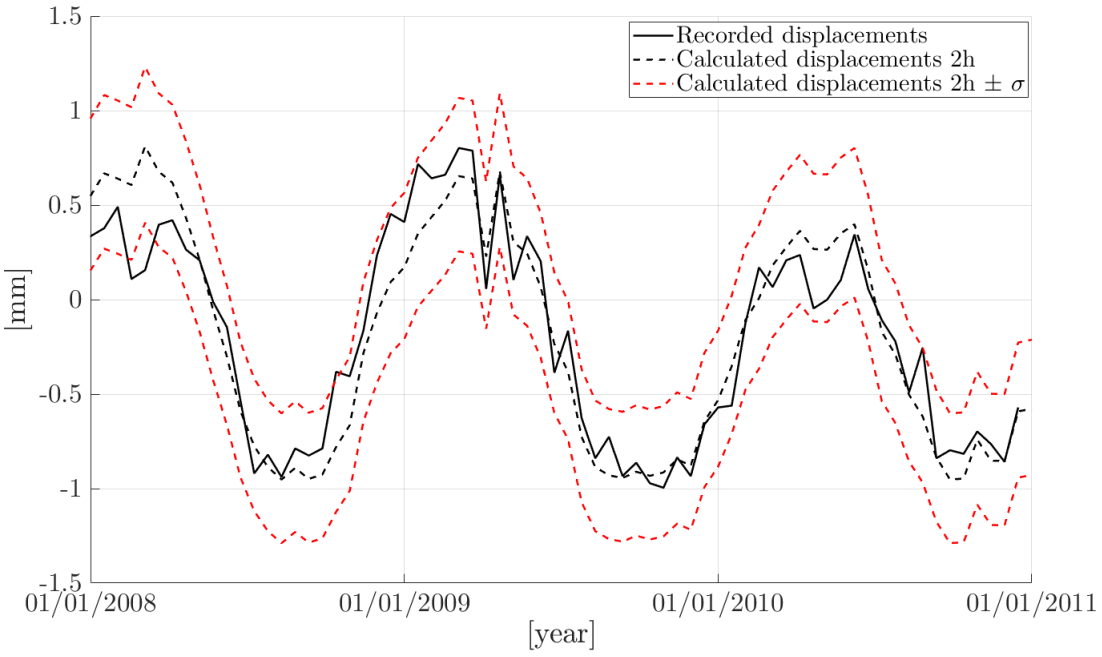
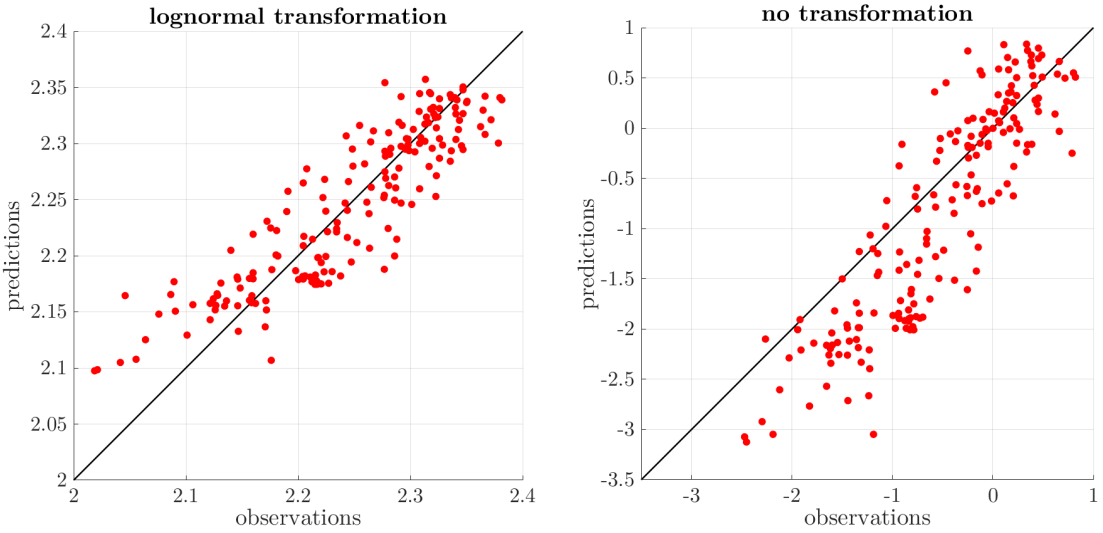


Fig. 8.17: Updating procedure in static case, comparison between recorded and calculated displacements of P04.

Finally, the homoscedasticity assumption has been verified by comparing the residuals obtained with or without the use of the logarithmic transformation function. Figure 8.18 shows that only with the use of the logarithmic transformation function the residuals have a constant error. In the case without logarithmic transformation function the residuals show an error which increases moving away from zero.

8.2.11 The use of the displacements probabilistic model for structural control purpose

The last part of the present study has the aim to show how to use the calibrated predictive model (Chapter 5) for structural control purpose. The high-fidelity model described in section 8.2.2 is used to simulate the structural response of the damaged dam after a



(a) Residuals calculated with the use of the transformation function.

(b) Residuals calculated withouth the use of the transformation function.

Fig. 8.18: Homoscedasticity assumption verification for the Gramolazzo dam.

seismic event. In particular, after the seismic analysis a basin level variation history and thermal loads, both for water and air, are used to simulate static displacements of the damaged dam during its regular use. Finally, the resulting displacements are compared with those predicted by the calibrated probabilistic model by assuming the same basin level variation and thermal loads. An excessive difference between them is associate to a damage even though the problem related to the threshold definition remains. In order to build a predictive model of the high-fidelity response the procedure proposed in Chapter 5 is repeated to calibrate the parameters of the probabilistic model. Finally, two damage scenarios are assumed in order to validate the application of the predictive model for different conditions.

Firstly the high-fidelity model is used to simulate the structural behaviour of the undamaged dam for two years, by assuming a basin level variation history sampled from a uniform distribution, bounded in the interval 20 - 30 m, and considering temperatures of water and air as in figure 8.4 during the period 2004-2006. A transient thermal analysis is performed and the resulting strain field variation applied in the mechanical model with the basin level variation during a structural transient analysis. In this way the dam behaviour is simulated by considering both thermal and mechanical actions. The model parameters used in this work are indicated in table 8.8.

Table 8.8: Reference parameters of the high-fidelity model of Gramolazzo dam.

K_C [MPa]	G_C [MPa]	K_G [MPa]	G_G [MPa]	ρ_C [kg/m^3]	ρ_G [kg/m^3]
7960	6788	17206	7987	2444	1800
α_C [C$^{\circ -1}$]	α_G [C$^{\circ -1}$]	λ_C [W/(mC$^\circ$)]	λ_G [W/(mC$^\circ$)]	c_C [J/(kgC$^\circ$)]	c_G [J/(kgC$^\circ$)]
1 10^{-5}	1 10^{-5}	2.3	0.04	880	840

The procedure presented in Chapter 5 is repeated in order to calibrate the parameters of the static twin model with respect to the outputs of the undamaged one. Only the predictive model with only two harmonics is used. Since the prior distributions of all parameters collected in $\boldsymbol{\Theta}$ are the same as in the previous case (section 8.2.6), the gPCE built in the case of real measurements is still valid.
The results of the updating procedure in terms of comparison between prior and posterior distributions are shown in figure 8.19. It highlights that the bulk modulus of both materials are more sensitive to the information embedded in the observations than the other parameters, as in the previous case.
Figure 8.20 shows the comparison between displacements simulated through the high-fidelity model and those obtained with the predictive model. The comparison shows a good agreement between them. By comparing the results shown in figure 8.20 with those of figure 8.17 it is easy to see that the use of simulated measurements leads to a better agreement between reference records and predictions than in the case of real observations. This fact is due to the absence of the measure error in the case of the simulated measurements.
Once the predictive model is calibrated it is used to control the structural behaviour. The Fivizzano seismic event ((LU) 12/07/2013) is used as reference ground motion in this application. This earthquake, characterised by a magnitude of 3.3, is one of the strongest seismic events occurred in this area and the epicentre was only 12 km far from the Gramolazzo dam. The registration recorded in the municipality of Minucciano (LU), only 2.78 km far from the dam, is used (Luzi et al. 2016). The registration is scaled in order to achieve two levels of damage scenarios. The reference input, i.e. the registration, is reported in figure 8.21, and it is characterised by a horizontal $S_a\left(\mathrm{T}_1\right)$ equal to 0.04 g. The energy contained in this seismic event is not sufficient to damage the structure, which in fact did not show damage after the real seismic event, so it is scaled until the damage occurs. As introduced before, two damage scenarios are considered in order to evaluate

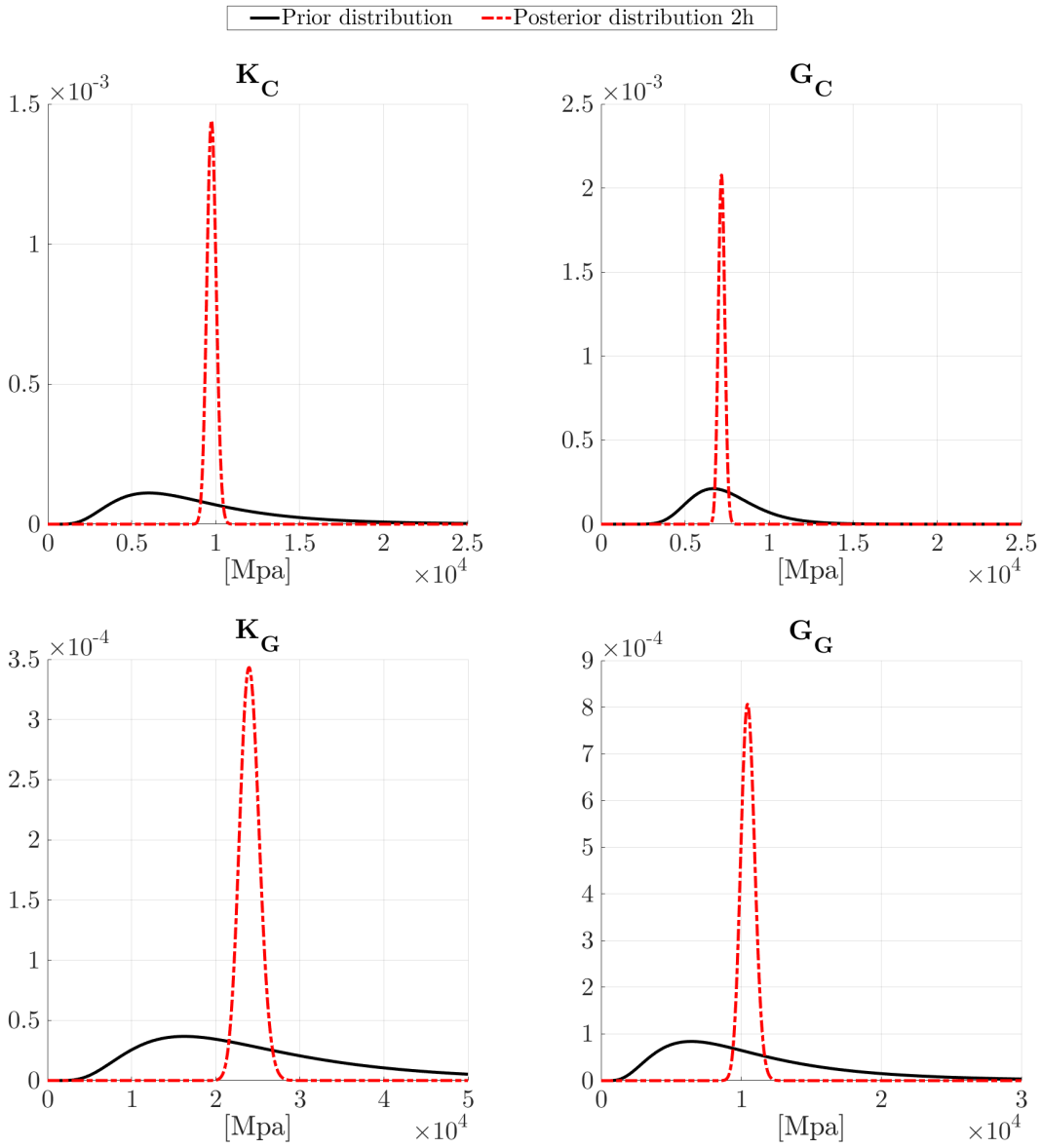


Fig. 8.19: Updating procedure in static case, comparison between prior and posterior distributions of K and G in the case of simulated observations.

the effectiveness of the calibrated predictive model of dam displacements for structural control purpose. In particular, they are based on the crack propagation achieved at the end of the seismic event. In this context, a crack path is assumed to be significant if the tensile damage variable d_t of the constitutive law is equal or greater than 0.3, i.e. the damage leads to a reduction of the material elastic modulus equal or greater than 30%. A

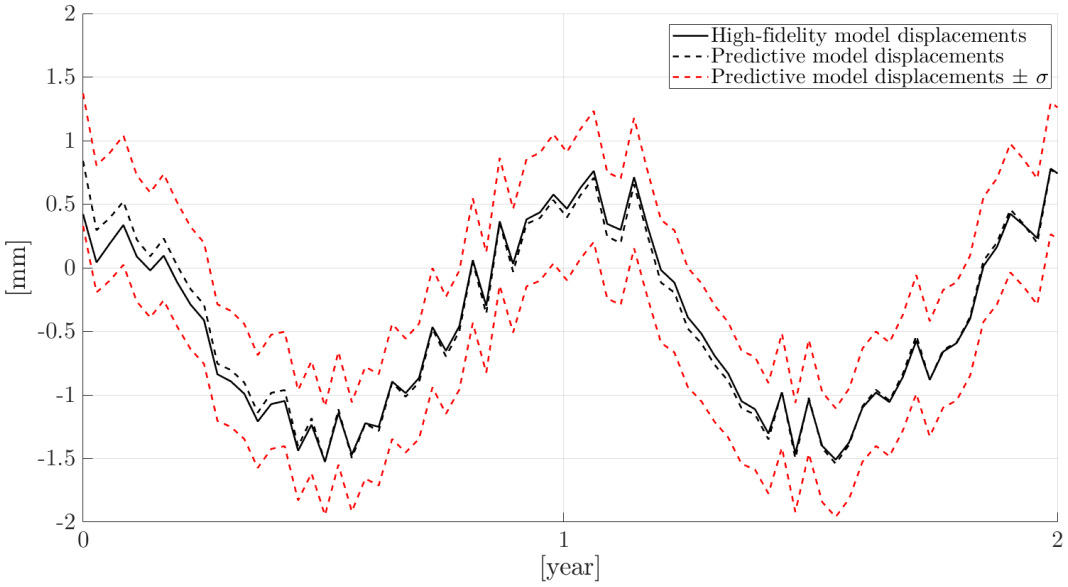
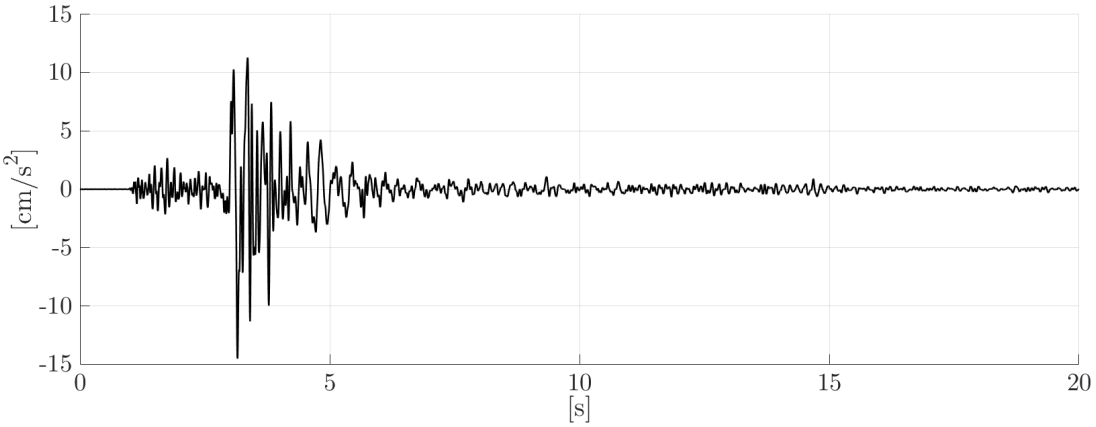


Fig. 8.20: Updating procedure in static case, comparison between displacements calculated with high-fidelity and predictive models.

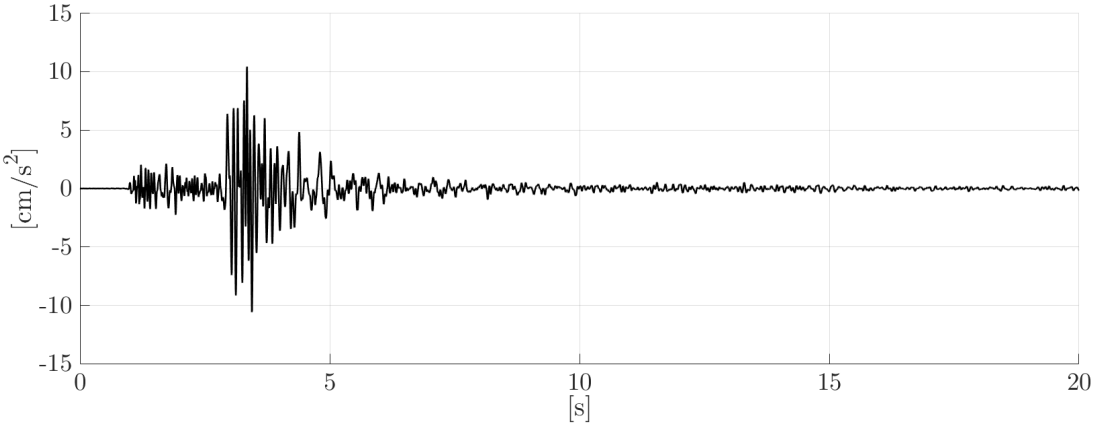
localised soft damage is assumed to be achieved when the most damaged section shows no thorough crack paths, and they do not reach the tunnel inside the dam body. A strong damage is assumed to be reached when the most damaged section shows thorough crack paths. Therefore, $S_a\,(\mathrm{T}_1)$ is scaled in order to obtain the previous damage scenario. The scaled input is deconvolved, according to Sooch and Bagchi (2014), and introduced in the model as displacements boundary condition.

The soft damage condition is reached by scaling the seismic input until a value of $S_a\,(\mathrm{T}_1) = 0.1$ g. The most damaged section is near to the right abutment, this area is particularly stressed because of the different the heights of the monoliths, as shown in figure 8.22. Figure 8.23 shows the comparison between the predictive model plus and minus the global error standard deviation and the results of the high-fidelity model simulation in the case of soft damage. The results shows that in this case the damage detection is particularly difficult, because the static behaviour is not changes so strongly. Indeed, the history of displacements of the damaged dam remains within the confidence interval of the predictive model.

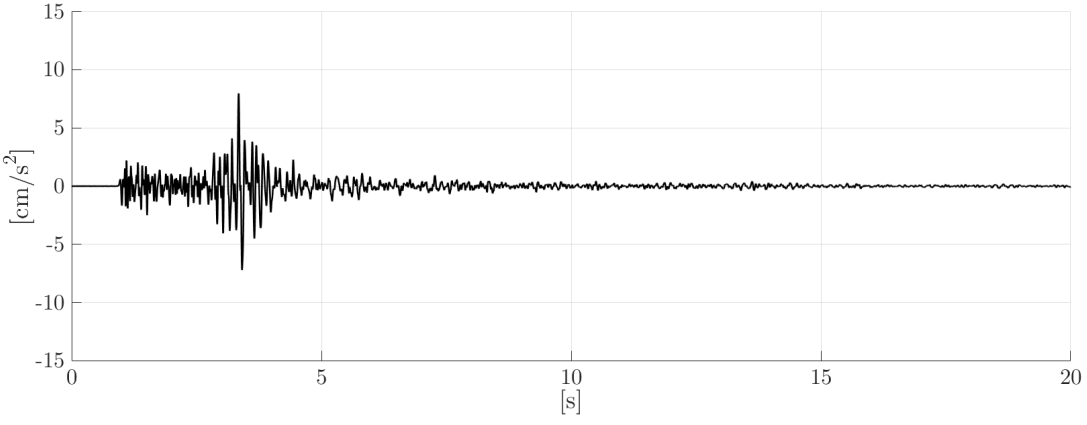
The strong damage is reached by multiplying the reference seismic action until $S_a\,(\mathrm{T}_1)$ is equal to 0.21 g. Through the analysis of the crack propagation, figure 8.24, it is pos-



(a) North-South direction record.

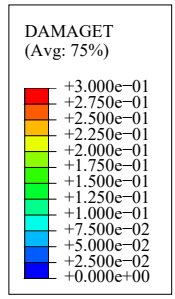


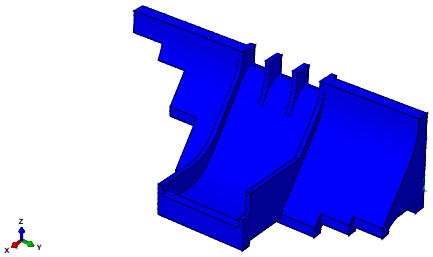
(b) East-West direction record.



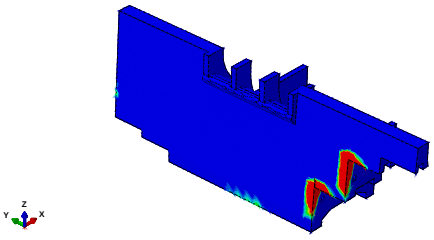
(c) Vertical direction record.

Fig. 8.21: Seismic event of Fivizzano recorded in Minucciano on the 12/07/2013.

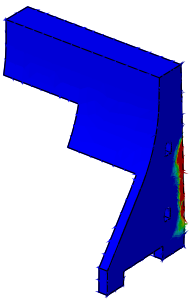




(a) Downstream view.



(b) Upstream view.



(c) Most damaged section.

Fig. 8.22: Gramolazzo dam: Damage Level 1.

sible to observe that the most damaged section is close to the right abutment, as in the previous case. The comparison between damaged and undamaged behaviour in terms of displacements is shown in figure 8.25. Clearly this time some differences can be found in the comparison, even tough these two behaviours are very similar. This time the displacements of the damaged dam exceed the confidence interval, then an abnormal phenomenon can be determined.

The application of the proposed static SHM system to the case of the Gramolazzo dam shows that the displacements of the control point, placed in the central monolith, are little affected by the crack propagation in a section close to the abutments. This phenomenon can be considered as the failure of a lateral vertical joint.

In this case a devices network which record the structural displacements in several different points could lead to a better detection of a possible damage. However, displacements are generally little affected by crack propagation because it leads to a reduction of the global stiffness which is evident only when damage is widespread.

From the general point of view a good SHM system should detect also soft localised dam-

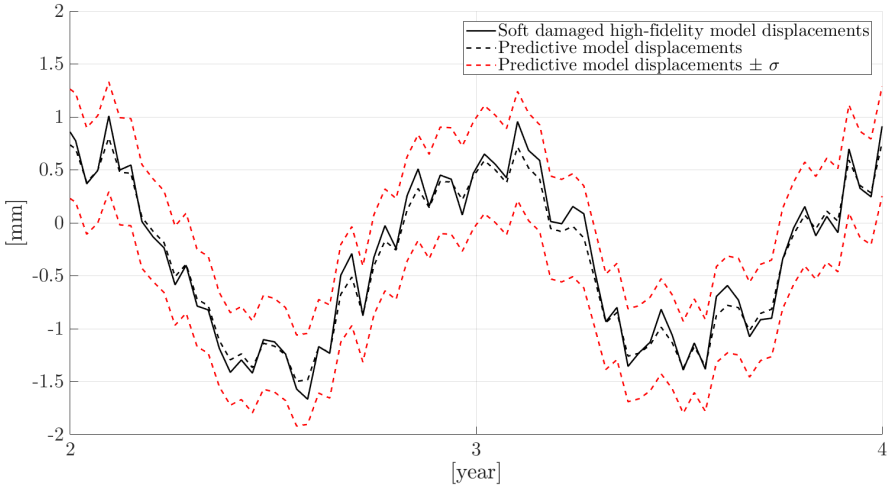
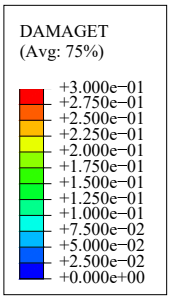
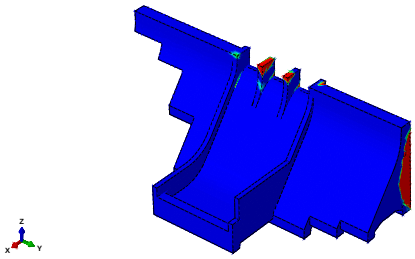
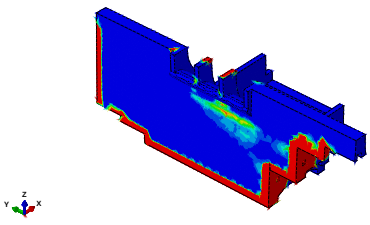


Fig. 8.23: Comparison between soft damaged high-fidelity model displacements and predictive model of the Gramolazzo dam.

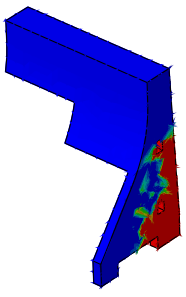




(a) Downstream view.



(b) Upstream view.



(c) Most damaged section.

Fig. 8.24: Gramolazzo dam: strong damage.

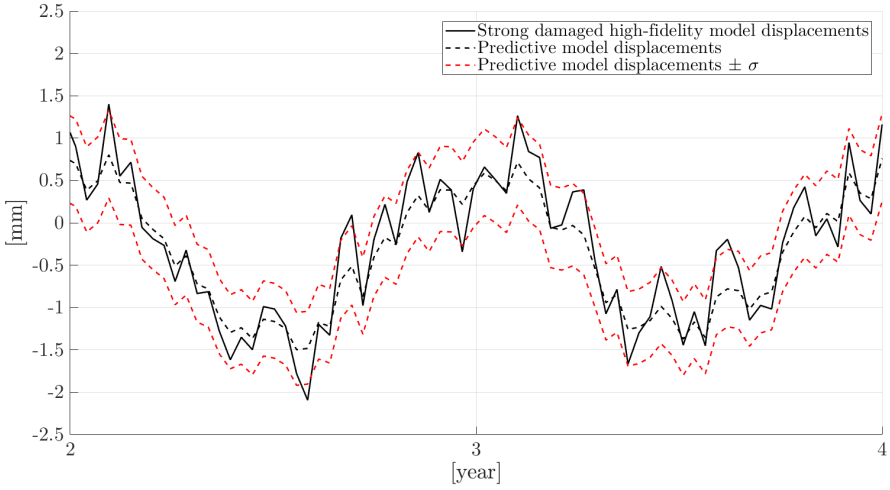


Fig. 8.25: Comparison between strong damaged high-fidelity model displacements and predictive model of the Scandarello dam.

ages. Therefore, dynamic measurements recorded in different points of the dam body can be more efficient for damage detection purpose, as shown in the next application.
Finally, in this example the threshold is assumed to be equal to the standard deviation of the predictive model. In general this value could be inappropriate, a study of the relationship between displacement variation and damage should be done in order to set a right threshold.

8.3 Dynamic SHM for concrete gravity dams: the case of Scandarello dam

8.3.1 Introduction

In this section the numerical results related to the study of the Scandarello dam are shown. The dynamic SHM presented in Chapters 6 and 7 is applied, by using the results of material tests and the results of the OMA as source of information.
In particular experimental frequencies and mode shapes are used, the numerical model is analysed by using a modal analysis, as previously discussed in Chapter 6. Therefore, the elastic parameters of both dam concrete and foundation soil, namely the bulk modulus K and the shear one G, are the subjects of the updating procedure.

The first step of the SHM is the definition of the proxy model which reproduce the modal behaviour of the undamaged dam. As introduced in Chapter 6.3, the proxy model is built through the gPCE.
Furthermore, the model parameters are updated by using either the predictive model of the frequencies or that one of the mode shapes, in order to compare them (*Training Phase*).
The effect of the updating procedure on the fragility curves calculation is assessed as Prognosis step. In particular, the fragility curve calculated with the prior distributions of the model parameters is compared with the one calculated by using the posterior distributions.

8.3.2 Dam description

The Scandarello dam is a concrete arch-gravity dam located in the center of Italy. This particular construction has a total length of 199.20 m and a maximum high of 55.5 m. The structure is characterised by a curvature radius in plant of 150.0 m, while the slope of the upstream face is 4% and that one of the downstream face is 70%. Since the dam has no key joints, the use of a 3D model is the best choice.
Several information about the dam are available due to the numerous experimental campaigns. Particularly important is the geological composition of the foundation soil, which is composed by two different kinds of rocks: an arenaceous mass and a marl mass. These two soil typologies have completely different mechanical characteristics, then they are modelled by using different elastic parameters.

8.3.3 Material tests results

Several different experimental campaigns have been conducted on the Scandarello dam with the aim of material characterisations. In particular, the results of the experimental campaigns performed in the 2012 are considered in this work. Both dam concrete and foundation soil have been investigated, through laboratory tests as compression ones, and in-situ tests as tomography. Regarding the foundation soil two different rock typologies have been found, an arenaceous mass and a marl mass. Figure 8.26 shows the arenaceous mass (orange) and the marl one (green). Therefore in this application, the parameters of the arenaceous mass have the subscript A, those of the marl mass have the subscript M and those of the concrete have the subscript C.
Finally, the results of the material tests related to the elastic properties are synthesised in table 8.9.
The results of the strength tests of the concrete are indicated in table 8.10.

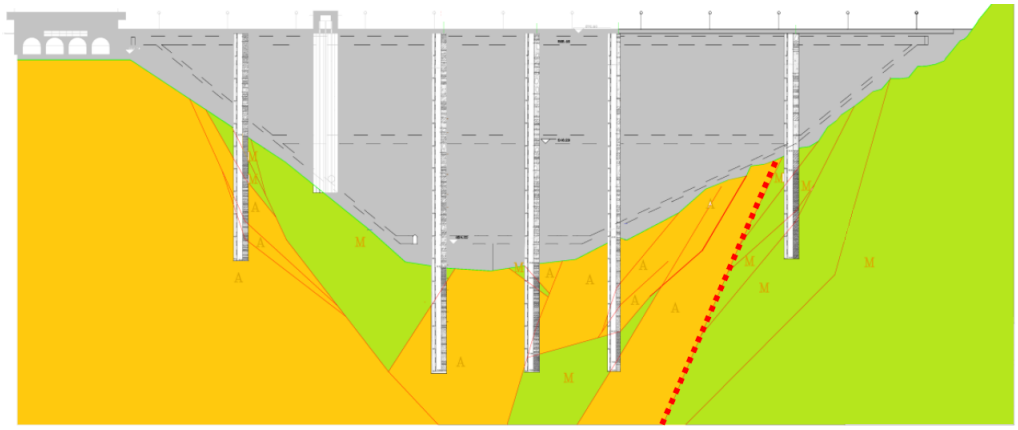


Fig. 8.26: Geological characterisation of the Scandarello dam.

Table 8.9: Scandarello dam: results of the material characterisation campaign.

	E_C [MPa]	ν_C	ρ_C [kg/m^3]	E_A [MPa]	ν_A	ρ_A [kg/m^3]	E_M [MPa]	ν_M	ρ_M [kg/m^3]
Mean	34620	0.16	2270	21000	0.16	1800	7000	0.22	1800
s.d.	12498	0.06	189.17	9545	0.07	-	3181.8	0.1	-

Table 8.10: Scandarello dam: results of the concrete strength tests.

	$f_{t,C}$ [MPa]	$f_{c,C}$ [MPa]
Mean	2	16.7
s.d.	1.1	5.16

8.3.4 Ambient vibration records and OMA results

In the 2016 an experimental campaign was performed by the Italian Civil Protection with the aim to record ambient vibrations in order to determine the modal characteristics of the system. In particular, 15 LE-3Dlite velocimeters, placed in the upper part of the structure, were used as shown in figure 8.27. Every registrations was characterised by a sampling frequency f_s equal to 200 Hz and a duration of 2 hours. The basin level variation was equal to 856.2 meters above the sea level, i.e. 40.8 m from the bottom of the dam. The ambient vibration were elaborated by the Italian Civil Protection by using the

Frequency Domain Decomposition (FDD) technique (Brincker and Ventura 2015). The ambient records and their OMA elaboration can be found in the experimental campaign report (*MISURE E ANALISI DELLE VIBRAZIONI DIGA DI SCANDARELLO*) available on-line (*ISS: Indagini conoscitive e monitoraggi nell'ambito dell'Osservatorio Sismico delle Strutture*).

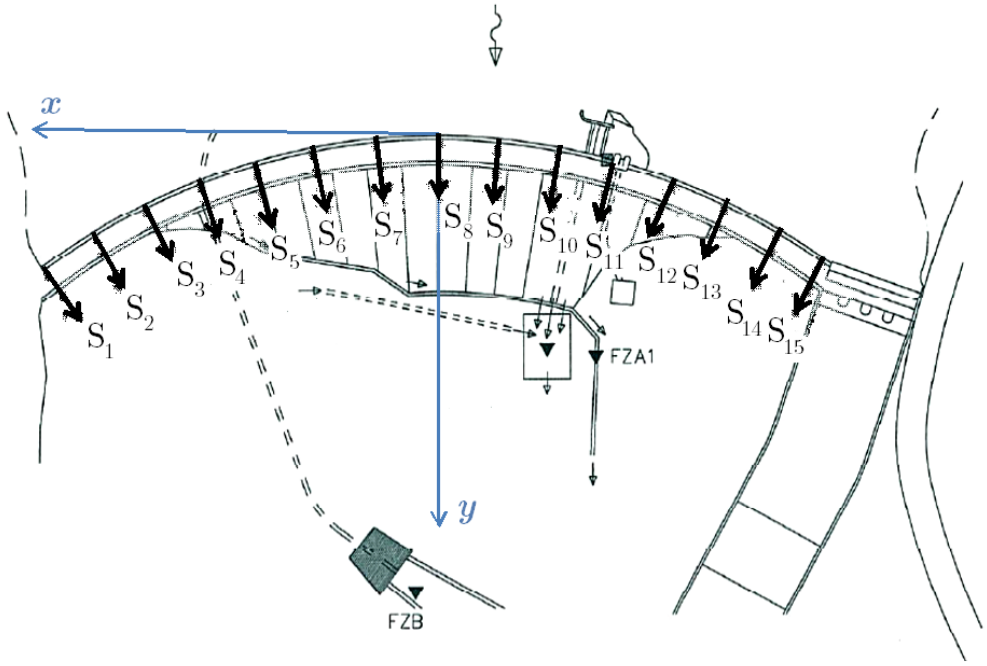
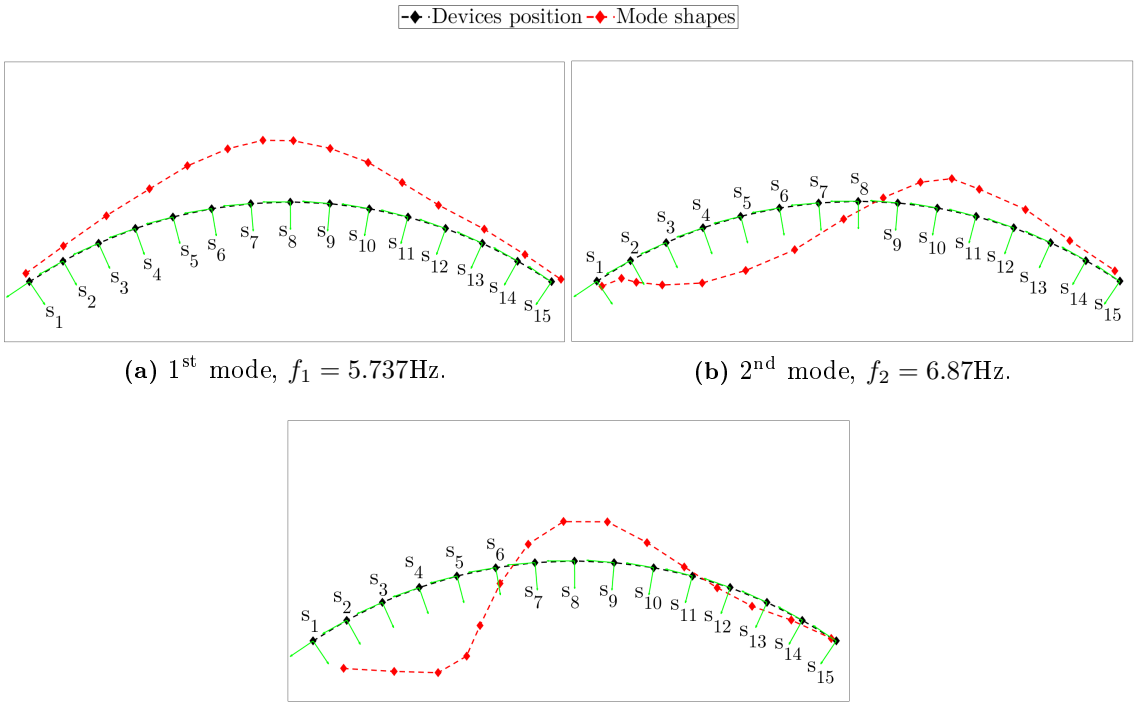


Fig. 8.27: Ambient vibration test layout of the Scandarello dam.

The first three experimental modes were identified, they were characterised by the frequencies reported in table 8.11. Whereas, the relative mode shapes are indicated in figure 8.28. Finally, figure 8.29 shows the auto-MAC matrix of the three experimental modes.

Table 8.11: Scandarello dam: experimental frequencies.

Mode number	Frequency [Hz]
1	5.737
2	6.87
3	8.896



(a) 1st mode, $f_1 = 5.737$Hz.

(b) 2nd mode, $f_2 = 6.87$Hz.

(c) 3rd mode, $f_3 = 8.896$Hz.

Fig. 8.28: Scandarello dam: experimental mode shapes.

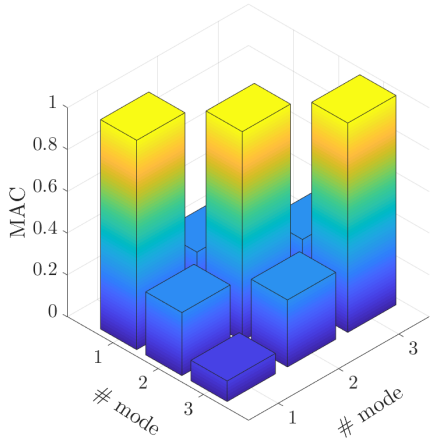
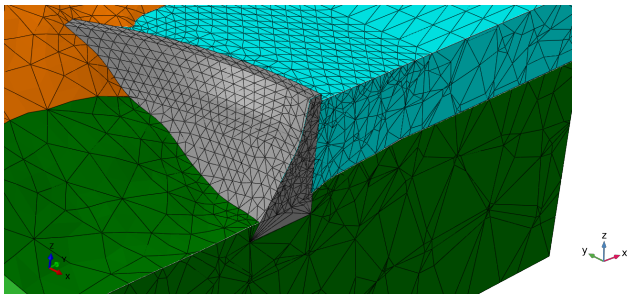


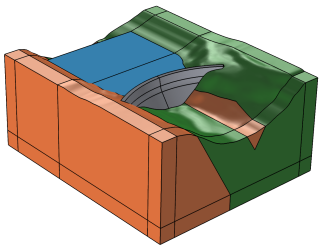
Fig. 8.29: Auto-MAC matrix of the Scandarello experimental modes.

8.3.5 Description of the FE models

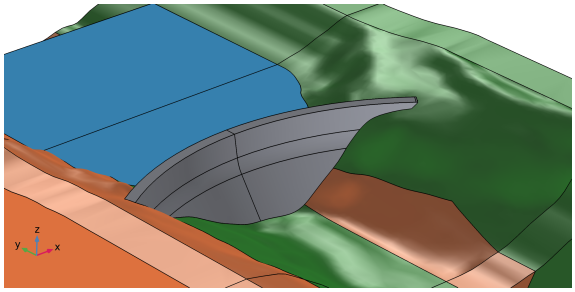
The dynamic SHM system proposed in Chapters 6 and 7 requires the definition of four FE models. A first FE model is needed within the *Training Phase* in order to build the gPCE of the predictive models of the modal characteristics of the undamaged dam. Therefore this is an elastic model. The geometry of the structure and the shape of the ground are built in a CAD program, based on the original drawings of the dam and the orographic map of the region. The model, performed in ABAQUS 6.14 (ABAQUS 2014), accounts for the FSI by considering the basin composed by fluid elements. Moreover, according to the indication reported in Chapter 3.4, low reflecting boundary conditions are imposed at the end of the fluid domain. Whereas, the SSI is considered by using infinite elements (Chapter 3.3) as boundary condition for the soil domain. Figure 8.30 shows the FE model of the Scandarello dam. In this case, a basin level equal to the one recorded during the experimental campaign, so 40.8m, is considered. This model is composed by 13280 quadratic tetrahedral C3D10 elements for the mechanical components, i.e. dam and foundation soil, 6137 linear tetrahedral AC3D4 elements for the basin and 237 linear hexahedral CIN3D8 elements for the IEs used as boundary condition. Since the reference analysis is the modal one, the materials are considered linear elastic.
A non-linear FE model is used to calculate the seismic fragility of the structure. As introduced in Chapter 2.2, the structural behaviour during seismic events is predicted by calculating the fragility curves. In this application fragility curves are calculated both considering prior and posterior distributions in order to evaluate the effect of the procedure in reducing epistemic uncertainties. The FE model performed to predict the seismic behaviour of the dam through the fragility curves calculation has the same features of the previous one, in terms of FSI and SSI assumptions. However, the basin level is fixed at the maximum height, equal to 52.9 m. Therefore, it is composed by 13280 quadratic tetrahedral C3D10 elements for the mechanical components, i.e. dam and foundation soil, 7671 linear tetrahedral AC3D4 elements for the basin and 237 linear hexahedral CIN3D8 elements for the IEs used as boundary condition. This model is used for dynamic analyses with seismic input, so the minimum mesh size is set as one tenth of the smallest wavelength, considering a significant frequency range of the seismic input.
The dam behaviour is modelled by using the concrete damage plasticity, introduced in Chapter 3.2.3, in order to simulate the cyclic behaviour of the material. By assuming that the structural failure is only related to the dam body the foundation soil is modelled by using a linear elastic material. The concrete damage plasticity model requires the definition of the strength parameters, those related to the damage behaviour and those related



(a) Model mesh.



(b) Global view of the model.



(c) Detail of the dam.

Fig. 8.30: FE model of the Scandarello dam.

to the yield surface. The results of the material tests allows determining the compressive and tensile strengths of the dam concrete, as indicated in table 8.10.
Likewise the static case, the strength parameters are used to define the post elastic behaviour both in tension and compression. The compressive post-elastic behaviour is assumed composed of three parts, whose shapes are respectively defined by equations 8.1, 8.2 and 8.3. The shape of the compressive behaviour is the same of the previous application, figure 8.7, but the compressive strength in the fragility curves calculation is considered a random variable, so there is not a single diagrams.
The tensile behaviour of the dam concrete follows an exponential law after the elastic branch (Equation 8.4), as in the case of the Gramolazzo dam. The tensile strength and the compressive one are treated as random variables, also in this case the plot of the tensile behaviour of the material has no meaning because it varies, even tough it has the same shape of figure 8.9. Finally, in the parametrization of the material behaviour during the fragility curves calculation the fracture energy is kept constant equal to 150 N/m, as discussed in detail in section 8.3.9.
The tensile damage law has the same shape of the previous application, and it is shown in figure 8.9. Since the tensile strength is considered a random variable with a constant

fracture energy, the damage law must be defined for every analysis. Indeed, equation 3.15 shows the relationship between plastic strain, tensile stresses and cracking strain, and so the need to define the damage law for every value of tensile strength.
Finally, the parameters of the yield criteria, described in Chapter 3.2.3, are defined with regard to the values which can be found in the literature, because no information about them is available for the Gramolazzo dam. The parameters are reported in table 8.12. Since dynamic analyses are performed Rayleigh's damping of the materials, both dam concrete and foundation soil, is defined in order to have a structural damping equal to 5% for the first two frequencies of the system.

Table 8.12: Yield criteria parameters of the Scandarello dam concrete.

ψ [°]	ε	σ_{b0}/σ_{c0}	K_c
36.31	0.1	1.16	0.66

The third FE model is used to reproduce the modal behaviour of the dam when a damage occurs (*Updating Phase*). The FE model used to train the gPCE is the same described for the fragility curves calculation. Also in this case the tensile strength varies, while the fracture energy remains constant. Since the described model is used to simulate the dam modal behaviour after seismic events, the FE analysis is composed by different steps. The analysis steps are summarised below:

- *gravity 1*: in the first step the gravity load is applied only to the soil.
- *gravity 2*: in the second step the gravity load is applied to the dam body.
- *hydrostatic force*: in this step hydrostatic forces are introduced.
- *seismic*: in this step a transient analysis is performed. The seismic action is applied at the boundary of the model as displacement field after a deconvolution procedure.
- *modal analysis*: in this step a modal analysis on the damaged configuration is performed.

The last FE model is a high-fidelity one with refined mesh, which allows simulating the damaged behaviour of the dam in order to test the *Detection Phase* and the possibility to update the strength parameters of the materials. This last FE model is defined by refining the mesh of the previous non-linear model, in order to obtain a more accurate result. The refined FE model is composed by 15936 quadratic tetrahedral C3D10 elements for

the mechanical components, i.e. dam and foundation soil, 6873 linear tetrahedral AC3D4 elements for the basin and 308 linear hexahedral CIN3D8 elements for the IEs used as boundary condition. This model is a deterministic one and is used to obtain the simulated reference measures of the damaged behaviour. A fixed value of the tensile strength equal to 1.45 MPa is considered, while the parameters needed to define the CDP are the same of the previous FE model, summarised in table 8.12. In order to assess the validity of the procedure with regard to the damage detection, three damage scenarios are considered. In particular, by considering only crack paths with tensile damage variable d_t higher than 0.3, DL1 is achieved when cracks appear in one of the critical areas (localised damage); DL2 is achieved when crack paths grow toward the core of the dam or others critical areas (diffuse damage); DL3 widespread damage or crossing crack path.

8.3.6 Prior distributions of *Training Phase*

Starting from the results of the material tests, described in section 8.3.3, the prior distributions of the hyper-parameters related to the mechanical characteristics of the material in the hierarchical predictive model (Chapter 6.5) can be defined. Indeed, the procedure introduced in Chapter 6 is defined with regard to a hierarchical predictive model, in order to update also the prior knowledge about the standard deviations of the mechanical parameters.
According to the notation introduced in Chapter 6, the hyper-parameters related to the mean values of the model random variables $\boldsymbol{\theta}_{\text{el}} = \boldsymbol{\theta}_{\text{gPCE}}$ can be collected in the vector $\boldsymbol{\xi}^{\mu}_{\text{el}}$, where:

- mean value of the concrete bulk modulus $\xi^{\mu}_{\text{gPCE},K_C}$;
- mean value of the concrete shear modulus $\xi^{\mu}_{\text{gPCE},G_C}$;
- mean value of the arenaceous mass bulk modulus $\xi^{\mu}_{\text{gPCE},K_A}$;
- mean value of the arenaceous mass shear modulus $\xi^{\mu}_{\text{gPCE},G_A}$;
- mean value of the marl mass bulk modulus $\xi^{\mu}_{\text{gPCE},K_M}$;
- mean value of the marl mass shear modulus $\xi^{\mu}_{\text{gPCE},G_M}$.

Moreover, the hyper-parameters related to the standard deviations of the model random variables $\boldsymbol{\theta}_{\text{gPCE}}$ can be collected in the vector $\boldsymbol{\xi}^{\sigma}_{\text{el}}$, where:

- standard deviation of the concrete bulk modulus $\xi^{\sigma}_{\text{gPCE},K_C}$;

- standard deviation of the concrete shear modulus $\xi^{\sigma}_{\mathrm{gPCE},G_C}$;
- standard deviation of the arenaceous mass bulk modulus $\xi^{\sigma}_{\mathrm{gPCE},K_A}$;
- standard deviation of the arenaceous mass shear modulus $\xi^{\sigma}_{\mathrm{gPCE},G_A}$;
- standard deviation of the marl mass bulk modulus $\xi^{\sigma}_{\mathrm{gPCE},K_M}$;
- standard deviation of the marl mass shear modulus $\xi^{\sigma}_{\mathrm{gPCE},G_M}$.

Since the materials tests results are in terms of Young Modulus and Poisson's coefficient, the prior distributions of the hyper-parameters can be defined by using equations 8.5. In the case of the standard deviation, since no information are available about their variance a Coefficient of Variation equal to 10% is considered. The final hyper-prior distributions are reported in table 8.13.

Table 8.13: Scandarello dam: Prior distributions of the mechanical parameters.

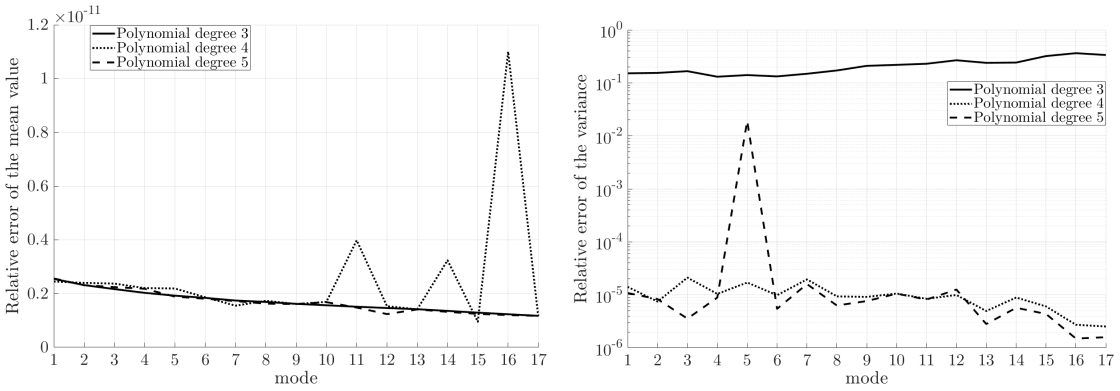
	$\xi^{\mu}_{\mathrm{gPCE},K_C}$ [MPa]	$\xi^{\mu}_{\mathrm{gPCE},G_C}$ [MPa]	$\xi^{\mu}_{\mathrm{gPCE},K_A}$ [MPa]	$\xi^{\mu}_{\mathrm{gPCE},G_A}$ [MPa]	$\xi^{\mu}_{\mathrm{gPCE},K_M}$ [MPa]	$\xi^{\mu}_{\mathrm{gPCE},G_M}$ [MPa]
Distr.	Log-Normal	Log-Normal	Log-Normal	Log-Normal	Log-Normal	Log-Normal
Mean	16971	14922	10294	9051.7	4166.7	2868.9
s.d.	6109.6	5372.1	4411.8	3879.3	1785.7	1229.5
	$\xi^{\sigma}_{\mathrm{gPCE},K_C}$ [MPa]	$\xi^{\sigma}_{\mathrm{gPCE},G_C}$ [MPa]	$\xi^{\sigma}_{\mathrm{gPCE},K_A}$ [MPa]	$\xi^{\sigma}_{\mathrm{gPCE},G_A}$ [MPa]	$\xi^{\sigma}_{\mathrm{gPCE},K_M}$ [MPa]	$\xi^{\sigma}_{\mathrm{gPCE},G_M}$ [MPa]
Distr.	Log-Normal	Log-Normal	Log-Normal	Log-Normal	Log-Normal	Log-Normal
Mean	6109.6	5372.1	4411.8	3879.3	1785.7	1229.5
s.d.	611.0	537.2	441.2	387.9	178.6	123.0

The prior distributions of the parameters with no physical meaning, such as the components of the covariance matrix $\mathbf{\Sigma}_f$, and the parameters of the covariance functions collected in the vectors $\boldsymbol{\lambda}^{\mathrm{U}}$, $\mathbf{w}_d^{\mathrm{U}}$, are defined as non-informative (Chapter 5.6).

8.3.7 Proxy model for modal characteristics of *Training Phase*

According to Chapter 6 the first step of the proposed procedure is the composition of the meta model based on the gPCE approach. This methodology, within the numerical algorithm MCMC, allows significantly reducing the computational burden, because it replaces the FEA outputs. Moreover, both frequencies and mode shapes must be approximated, so two different meta models are set up.
The procedure described in Chapter 4.4.3 is used to determine the gPCE coefficients, of

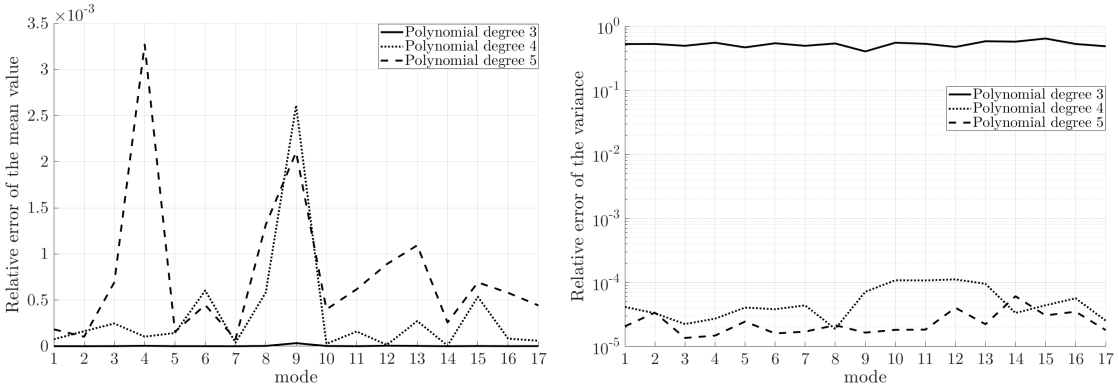
the basis functions assumed as hermitian. Whereas the analysis number and the polynomial expansion degree are changed in order to determine the best agreement between accuracy and computational burden. In particular, the final analysis number is 150 and the polynomial expansion degree varies from 3 to 5. Figure 8.31 shows the relative errors in terms of mean and variance for the frequencies approximation, while figure 8.32 shows the maximum relative errors of the mode shapes meta models. The errors, in both cases, are calculated for each numerical frequency. In particular, the number of numerical frequencies considered in this work is equal to 17 in order to reduce the computational burden. A sensitivity analysis, varying the model parameters, is performed in order to find the number of numerical modes to consider, in this case a maximum number of 17 modes in the frequency range from 4 to 20 Hz allows catching the dam behaviour.



(a) Error of the mean value. **(b)** Error of the standard deviation.

Fig. 8.31: Relative errors of the frequencies meta model of the Scandarello dam.

The analysis of the two figures highlights the different relationship between errors, modes and polynomial degree, even tough the error values are very low. Figure 8.31 shows that the mean value error is minimum with a polynomial degree equal to 3 or 5, while the variance approximation is better with a polynomial expansion equal to 4. Therefore, a polynomial expansion equal to 4 is chosen for the frequency proxy model, in order to reduce as much as possible the variance error. The same strategy is used for the mode shapes proxy model, thus analysing figure 8.32 is evident that the best choice is the polynomial expansion equal to 5 since it reduce the variance.
Finally, the Sobol's coefficients (Chapter 4.4.4) are calculated in order to understand the influence of each parameter on the model output, as shown in figure 8.33. Each parameter variation has a strong influence on the model output, so no one of them can be neglected



(a) Error of the mean value. **(b)** Error of the standard deviation.

Fig. 8.32: Relative errors of the mode shape meta model of the Scandarello dam.

in the updating analysis.

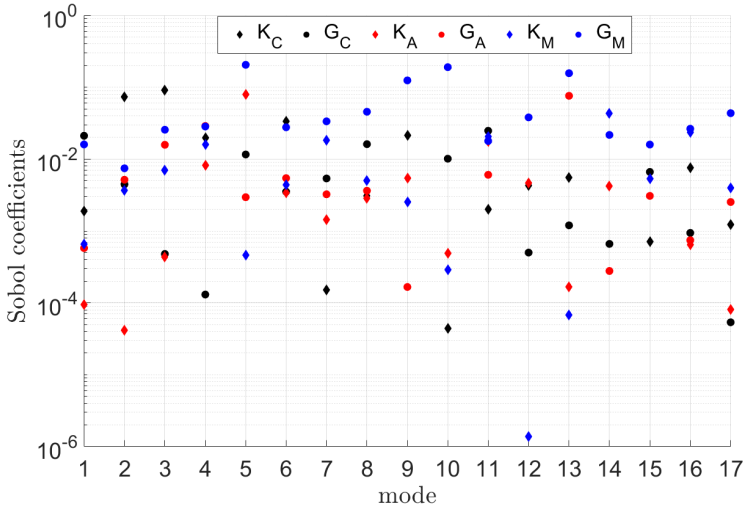


Fig. 8.33: Scandarello dam: Sobol's coefficients.

8.3.8 Bayesian updating and posterior statistics

Once a reliable proxy model is built, it can be used within the proposed procedure to reduce the computational burden. The experimental modal characteristics, i.e. frequencies and mode shapes, are used to updated the parameters of the predictive models, namely $\boldsymbol{\xi}^{\theta}_{\text{el}}$, so $\boldsymbol{\theta}_{\text{gPCE}}$, $\boldsymbol{\lambda}^{\text{U}}$, $\mathbf{w}^{\text{U}}_{d}$ and $\boldsymbol{\Sigma}_{f}$. The model parameters updating was done separately by

using frequencies and mode shapes for comparison reasons.
Regarding the frequency case, the posterior distributions are reported in table 8.14. Whereas figure 8.37 shows the comparison between prior and posterior distributions. The concrete elastic parameters are more influenced by the procedure than those of the soil. Indeed, the posterior distributions of K_C and G_C show a shift of the mean values and a reduction of the variances. Whereas, the mean values of the soil elastic parameters posterior distributions are very close to the prior ones. Regarding the variances of the soil parameters, they show a reduction if compared to the prior ones, but less than those of the concrete. The mean values of the error terms are very close to each other but they increase toward high frequencies. Table 8.16 shows a good agreement between experimental and numerical frequencies, while figure 8.34 shows the MAC matrix calculated by using the three experimental modes and the 17 numerical ones. In this frequency range only three numerical modes shows a high MAC coefficient (higher than 0.85) and so they can be recognised as correlated to the experimental ones. This shows the importance of the use of the reordering criterion within the procedure, as the proposed one. In particular, it allows ensuring the coherence between experimental and numerical modes, discarding the modes without experimental coherence. The three selected numerical mode shapes are shown in figure 8.35. The convergence is checked by performing three chains starting from different points composed by 100000 steps (1000 burn-in). Therefore, according to Chapter 4.3.6, the convergence of the numerical algorithm is verified by calculating $\hat{R}^p$ which is equal to 1.3 in this case.

Table 8.14: Scandarello dam: Posterior distributions of the mechanical parameters calculated by means of experimental frequencies.

	$\xi^{\mu}_{\text{gPCE},K_C}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_C}$ [MPa]	$\xi^{\mu}_{\text{gPCE},K_A}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_A}$ [MPa]	$\xi^{\mu}_{\text{gPCE},K_M}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_M}$ [MPa]
Mean	10965	5265.5	9438.3	7212.7	4802.1	2631.7
s.d.	471.49	136.91	462.48	230.80	292.92	94.53
	$\xi^{\sigma}_{\text{gPCE},K_C}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_C}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},K_A}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_A}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},K_M}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_M}$ [MPa]
Mean	1823.2	502.21	3870.4	939.19	1473.2	511.3
s.d.	145.86	25.11	313.50	54.47	114.91	34.76
	σ_{f_1}	σ_{f_2}	σ_{f_3}			
Mean	0.0042	0.0053	0.0089			
s.d.	0.0021	0.0034	0.0051			

In the updating procedure of the mode shape predictive model, the prior distributions are the same as the ones of the frequency case (Chapter 8.3.6). In Chapter 6.4 the use of

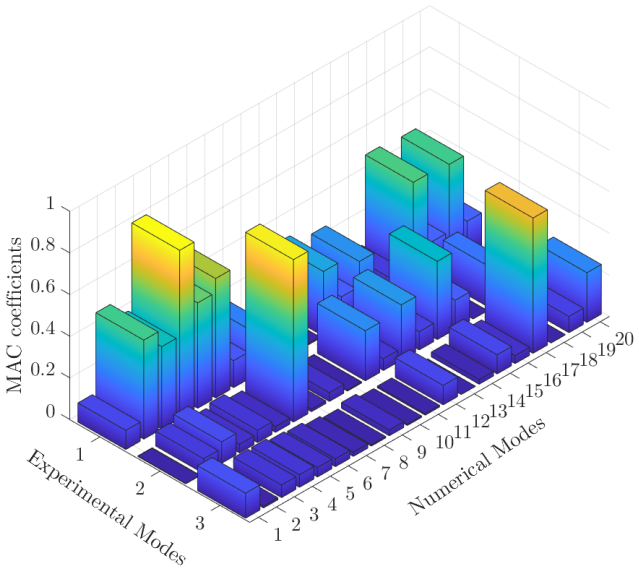
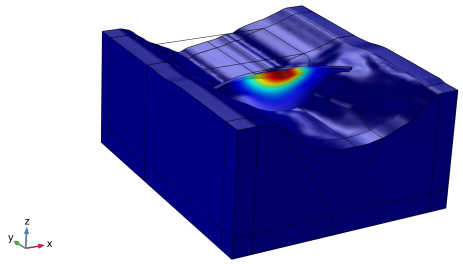
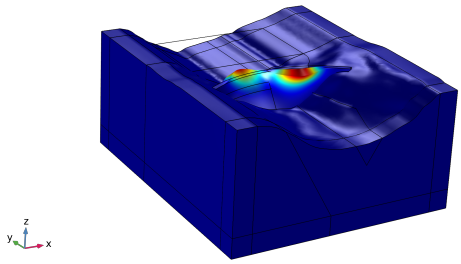


Fig. 8.34: MAC matrix between experimental frequencies and numerical ones calculated by using the frequency predictive model.

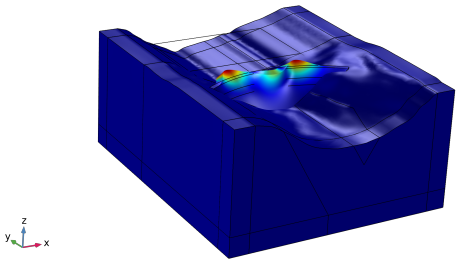
the covariance function is introduced. More specifically, an exponential function is used (equation 6.8) to approximate the covariance matrix terms. Since this assumption is particularly important for the predictive model, it must be verified. The analyses used to train the gPCE are also used to calculate the semivariogram and the estimation of the correlation between two points (Kottegoda and Rosso 2008). Figure 8.36a shows the plot of the correlation coefficients versus the Euclidean distance of the first three modes. An exponential function is fitted through these points, showing the goodness of the exponential covariance function. The modified MCMC (Chapter 2) is used to determined the marginal posterior distributions. In this case, as in the frequencies one, the selection of the mode shapes through a reordering criterion is a fundamental aspect both to ensure the coherence between experimental and numerical modes, and to discard the modes related to the soil. The results of the procedure are reported in table 8.15, while figure 8.37 shows the comparison between prior and posterior distributions. The results show how the procedure is more influential to infer the concrete parameters than those of the soil. A strong shift of the mean values of the concrete parameters can be observed with a reduction of the variances. Whereas the mean values of the soil elastic parameters remain close the prior ones. The variances of these last parameters are reduced with respect to the prior ones, but not as in the case of the concrete parameters, while the posterior parameters of the covariance functions shows higher values toward the highest frequencies. Furthermore,



(a) 1st mode, $f_1 = 5.737$Hz.



(b) 2nd mode, $f_2 = 6.87$Hz.



(c) 3rd mode, $f_3 = 8.896$Hz.

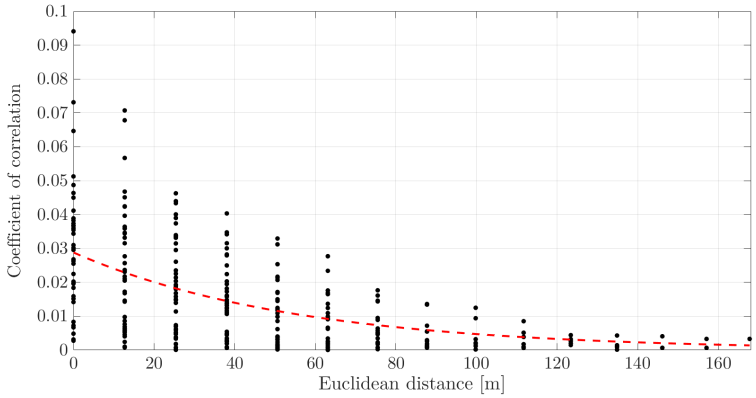
Fig. 8.35: Scandarello dam: experimental mode shapes.

figure 8.38 reports the comparison between experimental and numerical characteristics in terms of MAC matrix, it highlights the need of a coherence criterion within the numerical algorithm which discards the modes related to the soil, selecting only right candidates with respect to the experimental ones. Finally, table 8.16 shows a good agreement between experimental and calculated frequencies.

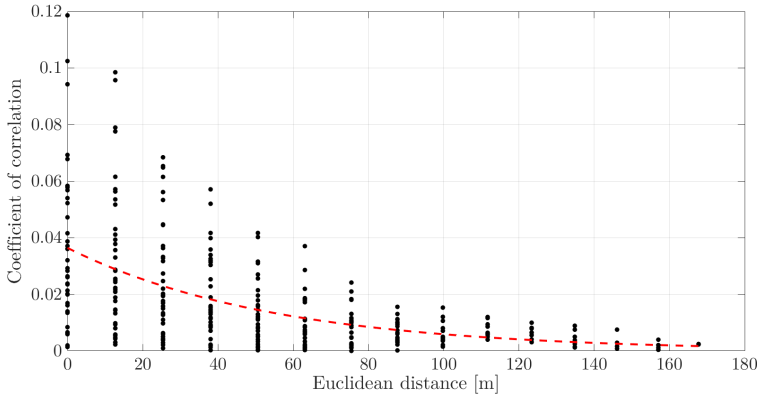
By comparing the results related to the two approaches (figure 8.37 and table 8.16) different comments can be derived.

First, both procedures lead to a stronger inference of the concrete parameters than those of the soil. The reasons of this fact can be found by reasoning about the nature of the information. Indeed, the experimental modes are calculated by using the registrations acquired on the dam body, so the effect of the soil is only a stiffening of the boundary conditions. Although dams strongly interact with the soil, differently from other kind of structures, and so the experimental frequencies are those of a complex system, if the ambient vibrations are measured on the dam body the modes of the dam are more evident than those of the soil.

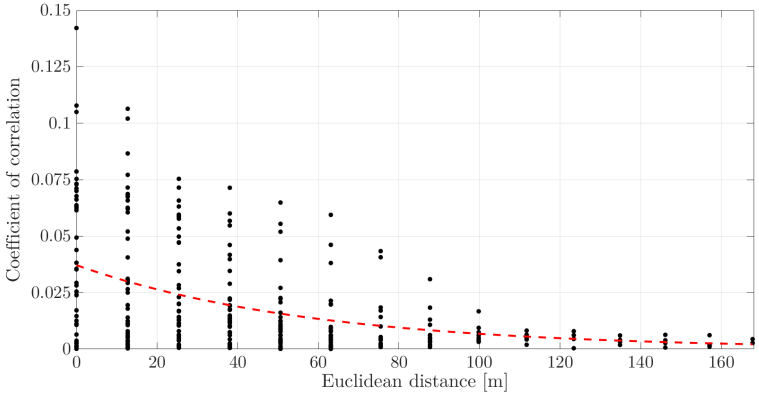
The two procedures lead to very similar posterior distributions and posterior modal char-



(a) Numerical mode 1.



(b) Numerical mode 2.



(c) Numerical mode 3.

Fig. 8.36: Scandarello dam: Correlation coefficients vs Euclidean distance of the first three modes.

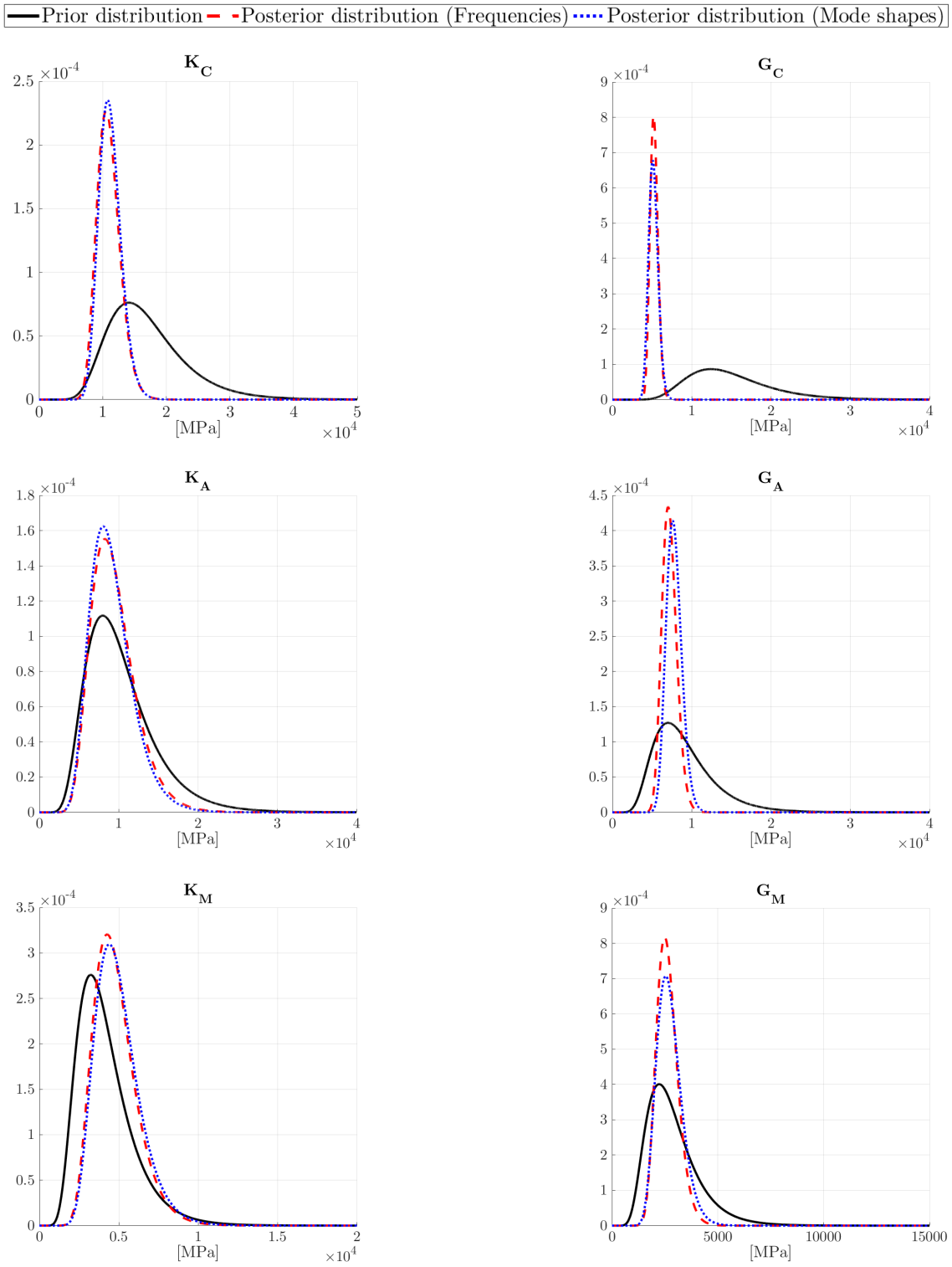


Fig. 8.37: Updating procedure with dynamic measurements, comparison between prior and posterior distributions of K and G.

Table 8.15: Scandarello dam: Posterior distributions of the mechanical parameters, calculated with the experimental mode shapes.

	$\xi^{\mu}_{\text{gPCE},K_C}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_C}$ [MPa]	$\xi^{\mu}_{\text{gPCE},K_A}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_A}$ [MPa]	$\xi^{\mu}_{\text{gPCE},K_M}$ [MPa]	$\xi^{\mu}_{\text{gPCE},G_M}$ [MPa]
Mean	11184.2	5186.5	9156.3	7789.7	4958.1	2724.6
s.d.	559.22	378.61	631.78	545.28	242.94	168.93
	$\xi^{\sigma}_{\text{gPCE},K_C}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_C}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},K_A}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_A}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},K_M}$ [MPa]	$\xi^{\sigma}_{\text{gPCE},G_M}$ [MPa]
Mean	1749.1	598.31	2732.74	982.26	1422.50	598.3
s.d.	155.66	45.77	240.42	68.03	137.29	45.01
	$\lambda^{1,\text{U}}$	$\lambda^{2,\text{U}}$	$\lambda^{3,\text{U}}$	$w_d^{1,\text{U}}$	$w_d^{2,\text{U}}$	$w_d^{3,\text{U}}$
Mean	25.16	39.09	34.28	0.0074	0.0261	0.0053
s.d.	5.24	6.25	2.92	0.0096	0.1615	0.0132

Table 8.16: Scandarello dam: Comparison between experimental and numerical frequencies.

	Experimental frequencies	Numerical frequencies frequencies predictive model	Numerical Frequencies mode shapes predictive model
f_1 [Hz]	5.737	5.693	5.592
f_2 [Hz]	6.870	6.722	6.691
f_3 [Hz]	8.896	9.029	9.183

acteristics. This is a proof of the goodness of the procedures.
The calibrated FE model is used in the next section to assess the structural fragility, showing how the proposed procedures can be used to improve the prediction of the structural behaviour during seismic events.

8.3.9 The use of the dynamic twin model for the fragility curves calculation

The procedure introduced in Chapter 6 allows updating the parameters of a dynamic twin model of the dam, thus reducing the epistemic uncertainties related to the elasticity of the materials. Therefore, the calibrated twin model can be used for different purposes, e.g. to detect damage (Chapter 7) or to predict the seismic behaviour of the dam. In this latter case the seismic behaviour can be assessed through the fragility curves calculation, as explained in Chapter 2.2.
More specifically, the ASOSM procedure proposed by Liel et al. (2009) and explained in detail in Chapter 2.2.2 is used to calculate the fragility curves reducing the computational

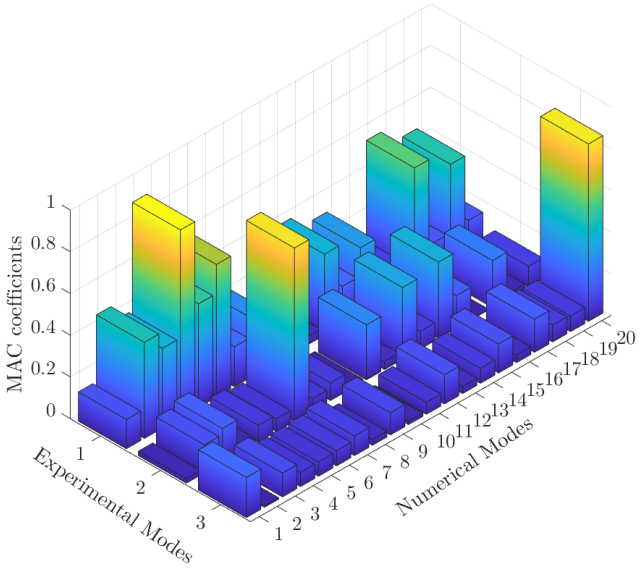


Fig. 8.38: MAC matrix between experimental frequencies and numerical ones calculated using the mode shape predictive model.

burden. The Multi-Record Incremental Dynamic Analysis (MR-IDA) is used to calculate the structural failure for each sample of aleatory and epistemic uncertainty.
The record-to-record variation is the only aleatory uncertainty, the MR-IDA analysis allows considering it but a set of earthquakes must be selected. In this application 15 events are chosen and reported in table 8.17. They are selected with regard to their magnitude, which is higher than 6 and their epicentral distance which less than 50 km. Moreover, they are recorded on site class A (rock soil) as that of the dam.
In MR-IDA every seismic event must be scaled until the structure reaches the collapse, i.e. a fixed LS. This process can be particularly expensive from the computational point of view, because the selected seismic event could have characteristics which does not affect the integrity of the dam. The *hunt & fill* algorithm (Vamvatsikos and Allin Cornell 2002) is used in this work. This procedure ensures that the record scaling levels are appropriately selected in order to minimize the number of required runs. More specifically, once the maximum analysis number and the desired accuracy for demand and capacity is selected, the algorithm starts increasing rapidly the IM until numerical non-convergence is encountered, indicating global dynamic instability. Furthermore, additional analyses are run at intermediate IM levels to sufficiently bracket the global collapse and increase the accuracy at lower IMs. This algorithm is implemented in MATLAB® R2017A (Mathworks 2016). Finally, every scaled seismic input composed of three components is applied at the

Table 8.17: Scandarello dam: Selected Ground Motion Records.

#	Event	Year	Station ID	Mw	Epicentral Distance [km]	EC8 site class
1	Campano Lucano	1980	ST96	6.9	32	A
2	Friuli	1976	ST20	6.5	23	A
3	Campano Lucano	1980	ST98	6.9	25	A
4	Bingol	2003	ST539	6.3	14	A
5	South Iceland (aftershock)	2000	ST2558	6.4	5	A
6	Duzce 1	1999	ST1252	7.2	34	A
7	Tabas	1978	ST54	7.3	12	A
8	Umbria Marche	1997	ST238	6	21	A
9	Montenegro	1979	ST64	6.9	21	A
10	Basso Tirreno	1978	ST49	6	34	A
11	Golbasi	1986	ST161	6	29	A
12	Duzce 1	1999	ST3136	7.2	23	A
13	South Iceland	2000	ST2556	6.5	35	A
14	Izmit	1999	ST575	7.6	9	A
15	Friuli (aftershock)	1976	ST36	6	28	A

bottom of the soil domain once properly deconvolved (Sooch and Bagchi 2014). The use of MR-IDA as reference analysis requires the definition of a IM. As deeply discussed in Chapter 2.2.2, several works are available in the literature which analyse the influence if the IM choice in the fragility curves calculation. The largest part of them uses $S_a(T_1)$ the spectral acceleration of the seismic input for a period equal to the fundamental one of the structure and considering a structural damping of 5%. Therefore, according to the indications available in literature $S_a(T_1)$ is considered as IM in this work.

For the sake of simplicity only one limit state is considered, following Tekie and Ellingwood (2003), the limit state function is defined in terms of defection of the top of the dam relatively to the heel considering a limit value equal to 0.03% of the dam height, which corresponds to 16.65 mm (equation 8.6). This LS can be related to large deformations of the dam during seismic shaking, which might impair the internal drainage or function of appurtenant structures.

$$\Delta\delta^{\text{top}-\text{heel}} \geq 16.65\text{mm} \tag{8.6}$$

In this application the epistemic uncertainties are those related to the elastic parameters of the materials, namely K and G, and the values of the concrete strength parameters. The elastic parameters of the materials are the objectives of the updating procedures, both

their prior distributions and the posterior ones are used to calculate the fragility curves in order to assess the effects of the proposed procedure. The distributions of the concrete strength parameters can be defined starting from the results of the material tests, shown in table 8.10. Both the compressive and tensile strengths are considered log-normally distributed. The post-elastic tensile constitutive behaviour and the tensile damage law are calculated every time that a new value of the tensile strength $f_{t,C}$ is sampled. The post-elastic tensile branch can be approximated by assuming a functional expression, e.g. exponential in this case, whose parameters are determined in order to kept constant the fracture energy. Therefore, the tensile damage law must be defined by using equation 3.15. A matlab procedure is implemented to automatize this task.
The correlation between epistemic uncertainties could lead to a variation of the results in terms of fragility curves. The ASOSM procedure, introduced in Chapter 2.2, allows easily considering correlations between random variables. Therefore, the fragility curves are calculated assuming the RVs in a case uncorrelated, and in the other case correlated, considering a correlation matrix based on expert judgement, as shown in table 8.18.

Table 8.18: Scandarello dam: Correlation Matrix of the concrete constitutive model parameters.

	K_C	G_C	$f_{c,C}$	$f_{t,C}$
K_C	1			
G_C	0	1		
$f_{c,C}$	0.83	0.83	1	
$f_{t,C}$	0.83	0.83	0.9	1

The first step of the ASOSM method is the calculation of the fragility curve only related to the record-to-record variation. Whereas, the effects of the epistemic uncertainties can be considered in a simplified way, evaluating the gradients of the results corresponding a particular RV (Chapter 2.2). Figure 8.39 shows the mean MR-IDA curve, and its confidence interval, calculated considering only the aleatory uncertainty. The curve interval shown in figure 8.39 that is small around $\Delta\delta^{\text{top}-\text{heel}} = 16.65$ mm, expresses the influence of the record-to-record uncertainty on the determination of the LS.

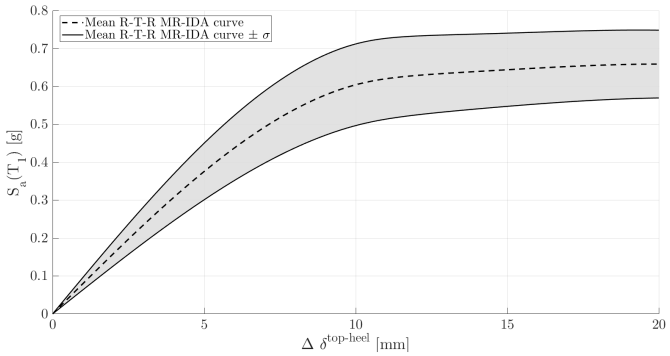


Fig. 8.39: MR-IDA curve calculated considering only the Record-to-Record uncertainty.

The fragility curves related to the prior distributions of the mechanical parameters are calculated, and shown in figure 8.40. The figure highlights the effects of the epistemic uncertainty, which leads to an increase of the variance. By comparing the curve calculated considering the epistemic uncertainties uncorrelated and the one considering them correlated, it is evident that the latter is characterised by the highest value of variance. The mean value of the fragility curve decreases considering also the epistemic uncertainties, regardless their correlation.

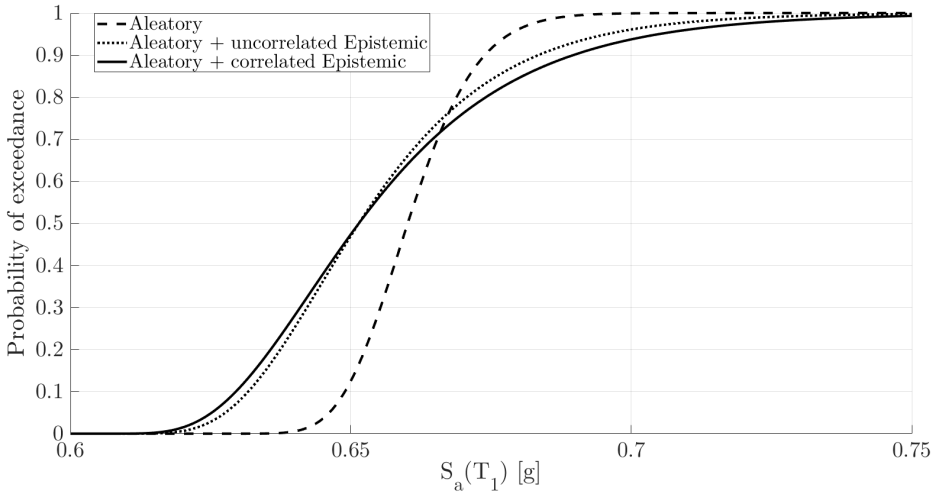


Fig. 8.40: Scandarello dam: fragility curve calculated using the prior distributions of the mechanical parameters.

The procedure is performed again by considering the posterior distributions of the mechan-

ical parameters, for comparison purposes. Figure 8.41 shows the fragility curves related to the posterior distributions, as in the previous case the effect of the epistemic uncertainties can be seen in terms of mean values shift toward lower values and an increment of the variance. Anyway, the effects of the epistemic uncertainties are significantly reduced with respect to the case of the prior distributions.

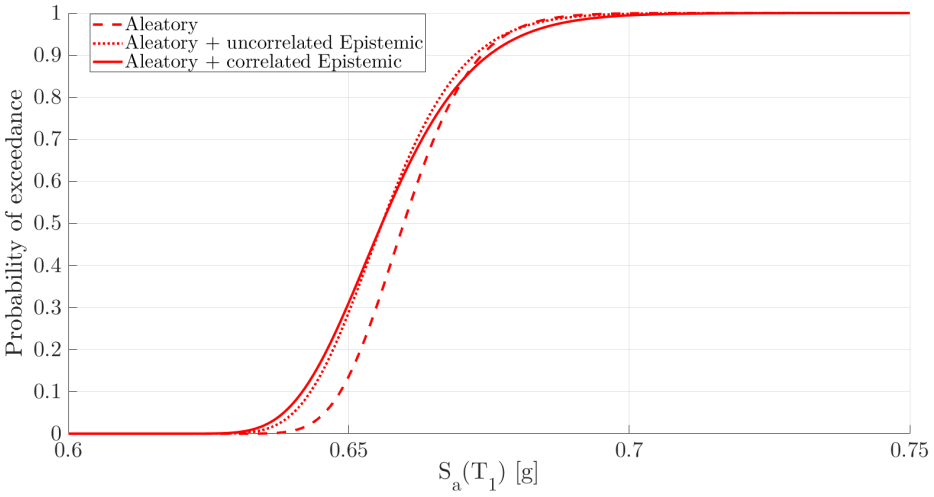


Fig. 8.41: Scandarello dam: fragility curve calculated using the posterior distributions of the mechanical parameters.

Finally, the comparison of prior and posterior fragility curves in the case of uncorrelated epistemic uncertainties (figure 8.42) and in that one considering them correlated (figure 8.43) shows a significant change in the resulting fragility curves. In both cases the values of the standard deviations significantly decrease while the mean values are very close. The parameters of the fragility curves, calculated by using the Method of Moments (MOM) explained in Chapter 2.2.3, are reported in table 8.19.

Table 8.19: Scandarello dam: Fragility curves parameters.

		Mean value	Standard deviation
Prior	Aleatory	0.661	0.0097
	Aleatory + Uncorrelated Epistemic	0.652	0.0213
	Aleatory + Correlated Epistemic	0.652	0.0257
Posterior	Aleatory	0.6605	0.0099
	Aleatory + Uncorrelated Epistemic	0.657	0.0114
	Aleatory + Correlated Epistemic	0.657	0.0133

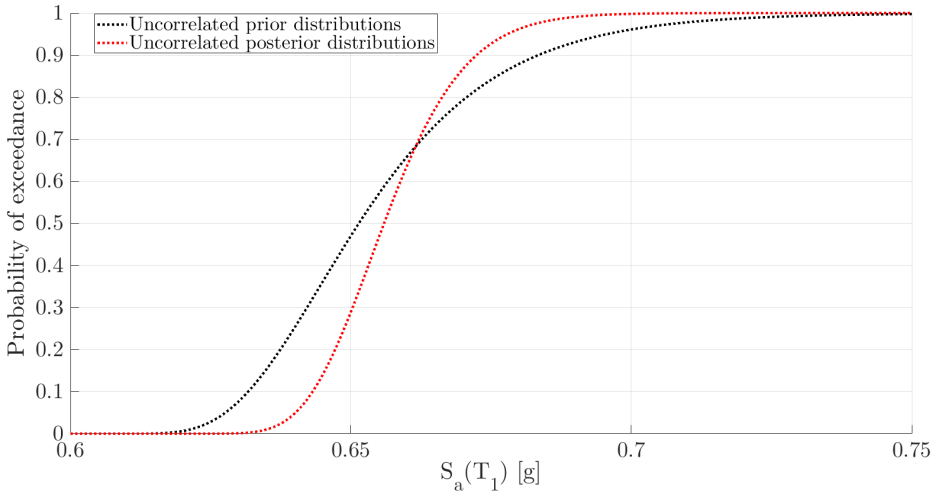


Fig. 8.42: Scandarello dam: comparison between fragility curve calculated using prior and posterior distributions of the uncorrelated mechanical parameters.

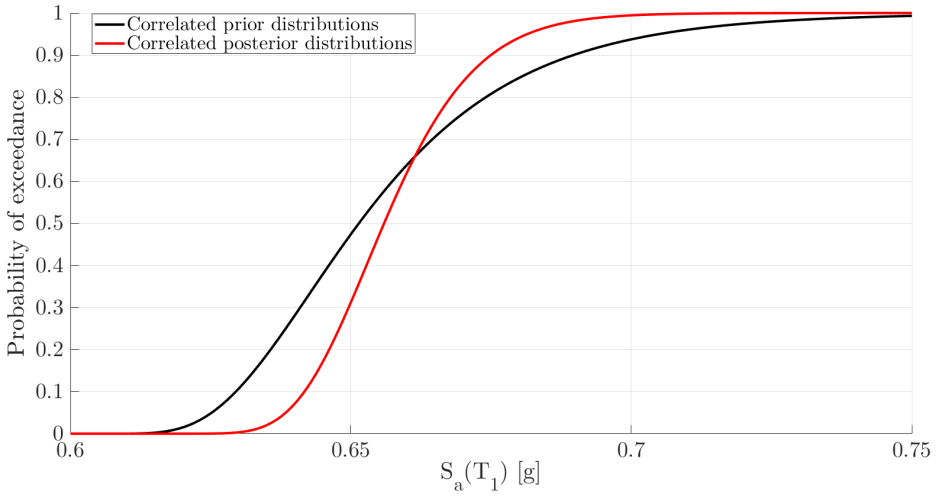


Fig. 8.43: Scandarello dam: comparison between fragility curve calculated using prior and posterior distributions of the correlated mechanical parameters.

Finally, figure 8.44 A shows the initial crack path, which corresponds to the first damage state of the Scandarello dam. This figure shows the importance of the 3D model, because the crack path in the downstream surface starts from the left abutment. Figures 8.44 B and 8.44 C show the final damage state of the structure, which is related to the overcoming

of the selected LS. Anyway, the analysis of the final damage state is interesting because it shows that the crack path starts from the left abutment and it propagates toward the right one and the crest of the dam. The application of the procedure proposed in Chapter

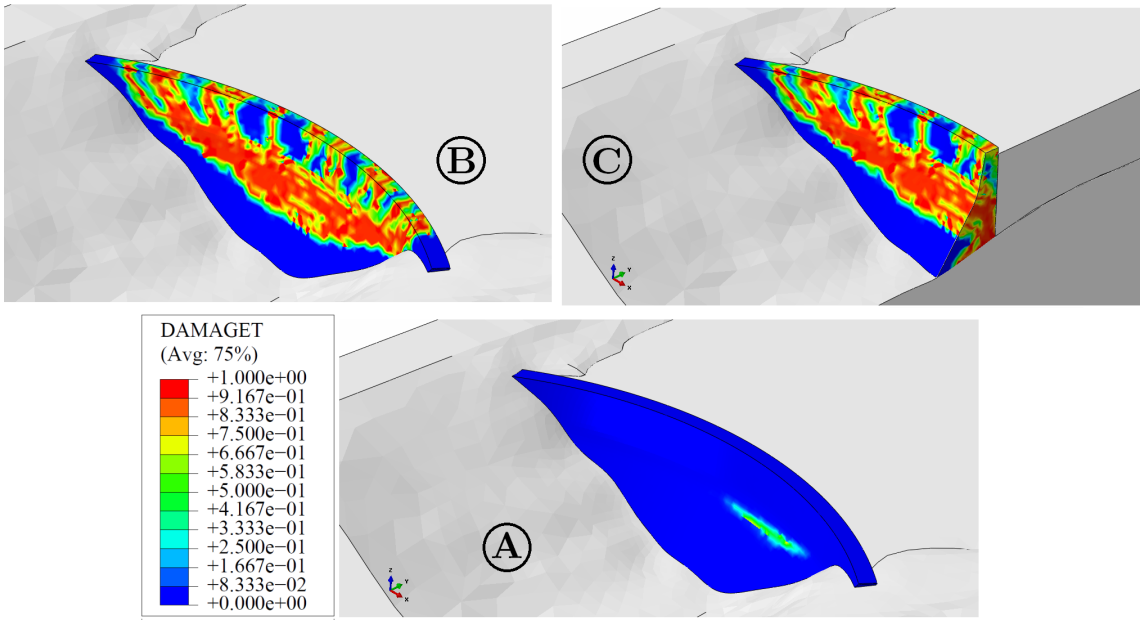


Fig. 8.44: Damage development in the Scandarello dam: a) first damage, b) final damage, c) final damage cross section.

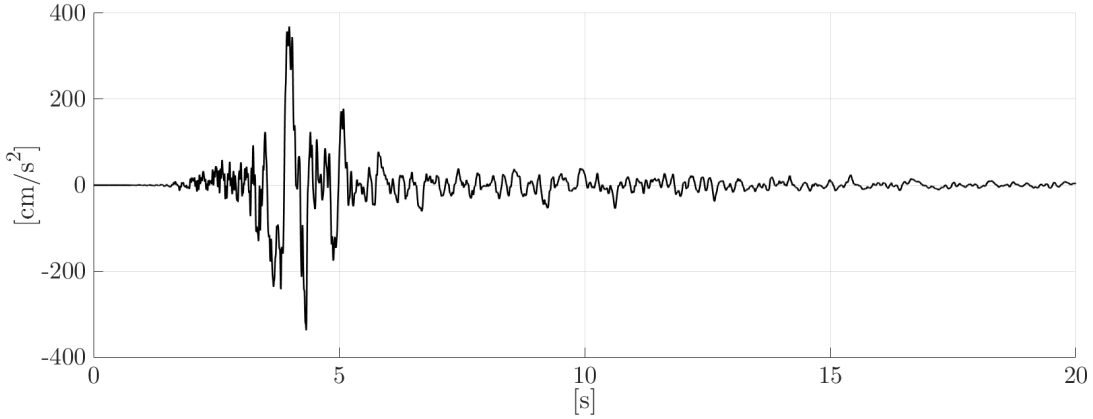
6 shows the possibility to reduce the epistemic uncertainties by using ambient vibrations. The elaboration of the fragility curves highlights the need to account both for aleatory and epistemic uncertainties, as also reported in literature (Chapter 2.2.3).
Finally, the analysis of the collapse mechanism of the dam shows how only an accurate model can be able to catch this particular failure mode. In particular, the crack path starts from the left abutment in the downstream face of the dam, so only a 3D model is able to detect this particular behaviour. Obviously, this result is not generalizable because only the Scandarello dam has been analysed so deeply. This fact and the reasons reported in Chapter 3 show that significant differences arise when a 3D high-fidelity model is used to assess the structural behaviour instead a 2D one, also in terms of localization of the failure.

8.3.10 *Detection Phase*

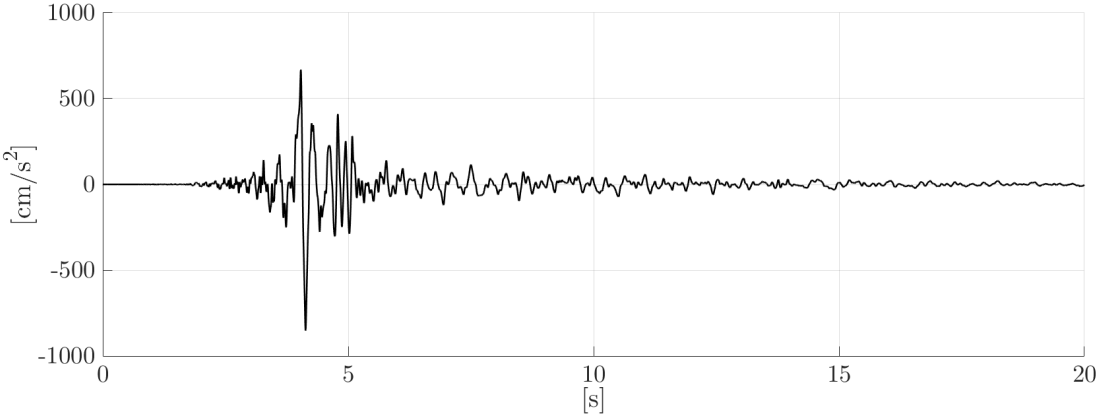
Once the predictive model of the undamaged dam modal characteristics has been trained, it can be used for structural control purpose by comparing the prediction of the undamaged mode shapes with the experimental ones determined by the elaboration of the ambient vibrations. The damaged behaviour of the dam must be simulated through a high-fidelity model, as previously described.

Three Damage Levels (DLs) are defined with regard to the critical areas defined in Figure 2.21. In particular, by considering only crack paths with tensile damage variable d_t higher than 0.3, DL1 is achieved when cracks appear in one of the critical areas (localised damage); DL2 is achieved when crack paths grow toward the core of the dam or others critical areas (diffuse damage); DL3 widespread damage or crossing crack path. As in the static case the spectral acceleration at the first period of the structure $S_a\left(T_1\right)$ is used as IM of the seismic input. The Accumoli seismic event of the 24/08/2016 is selected (Figure 8.45) as reference action, because the epicentre was 6 km far from the dam and the seismic station of Amatrice, which recorded the event, is just 6 km faraway. The magnitude of the event is equal to 6, the maximum horizontal PGA of the selected registrations is equal to 850.804 cm/s^2, while the vertical one is equal to 43.459 cm/s^2 (Luzi et al. 2016) and the horizontal spectral acceleration is 1.5622 g. The Accumoli seismic event is scaled with respect to $S_a\left(\mathrm{T}_1\right)$ until the damage scenario is achieved. Before the application of the scaled seismic event as boundary conditions at the soil bottom, it is properly deconvolved (Sooch and Bagchi 2014). DL1 is achieved for $S_a\left(T_1\right) = 0.2107$g, DL2 is achieved for $S_a\left(T_1\right) = 0.4328$g and DL3 is achieved for $S_a\left(T_1\right) = 0.472$g. These damage scenarios can be achieved only by reducing the initial seismic input. This means that the dam did not experience the recorded entity of the seismic action, which was recorded less than 3 km far from the epicentre. Moreover, the tensile strength is assumed equal to 1.45MPa, which is a small value if compared to the results of the material test. The damage levels are reported in figure 8.46. Following the proposed procedure, damage is detected by comparing the mode shapes of the undamaged system, calculated through the calibrated predictive model, with those recorded after the seismic event. Two different layouts of the SHM system are considered, in the first case, as shown in Figure 8.47. The first layout is composed by 15 devices placed along the dam crest, as during the experimental campaign. In the second layout the same 15 devices are placed in a different way, as shown in figure 8.48. More specifically, 7 devices are placed along the dam crest, 4 devices are installed 10 m under the crest and the last 4 devices 20 m under the crest level.

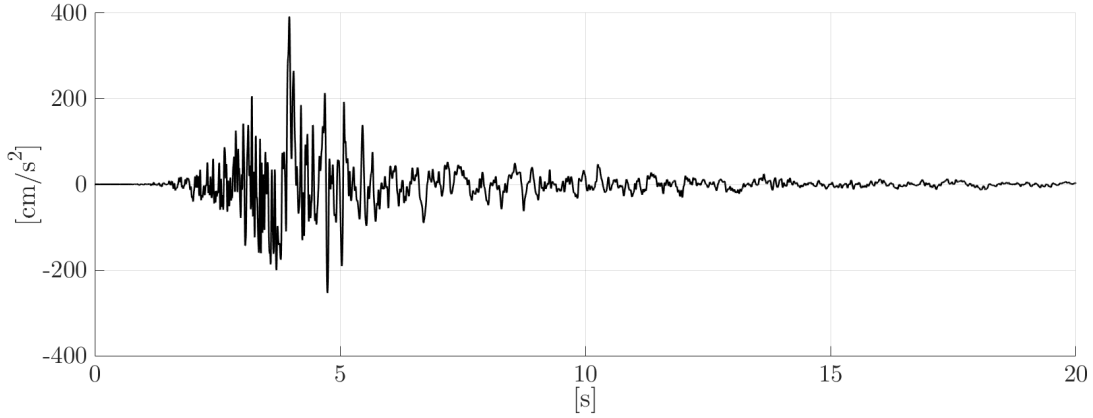
In both cases the first three modes determined with the calibrated predictive model are



(a) North-South direction record.

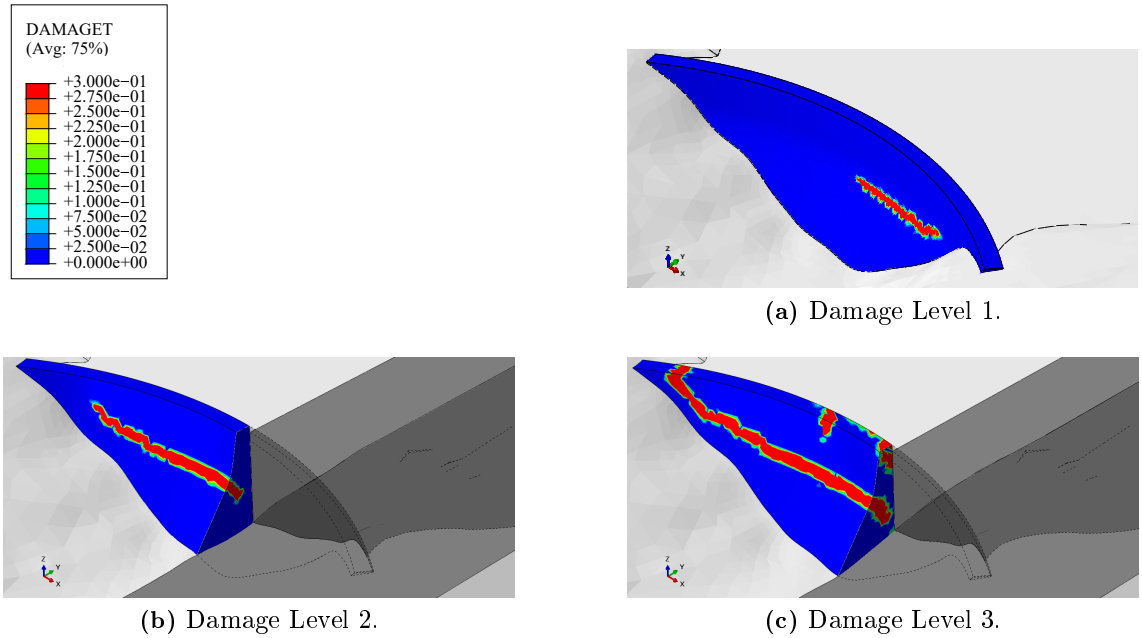


(b) East-West direction record.



(c) Vertical direction record.

Fig. 8.45: Seismic event of Accumoli recorded in Amatrice on the 24/08/2016.



(a) Damage Level 1.

(b) Damage Level 2.

(c) Damage Level 3.

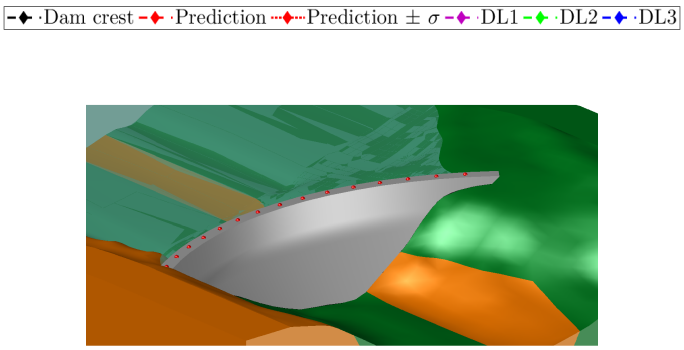
Fig. 8.46: Scandarello dam: Damage Levels, *Detection Phase*.

compared with the corresponding ones related to the damaged dam. The predictions and the values of the DL1 are always very close, even though this last is out from the stripe defined by the prediction plus and minus the standard deviation of the error. This is due to the little extension of the crack path in the first scenario, as shown in figure 8.46. In the other two cases the predictions and mode shapes of the damaged system are well separated and distinguishable.

Although the final results are qualitatively very close, the second layout in which the devices are not aligned allows drawing the three-dimensional behaviour of the dam. This leads to a wider control.

8.3.11 Prior distributions and seismic action parametrization in the *Updating Phase*

The last phase of the procedure aims to use the information gathered during a seismic event to update the state of knowledge about the strength parameters of the materials. Also in this case the computational burden related to the Bayesian inference is reduced by the use of gPCE based meta model. However, differently from the previous step, in this one the gPCE is trained with regard to the damaged dam. In particular, observing that the compressive stresses in the concrete mass are low if compared to the relative



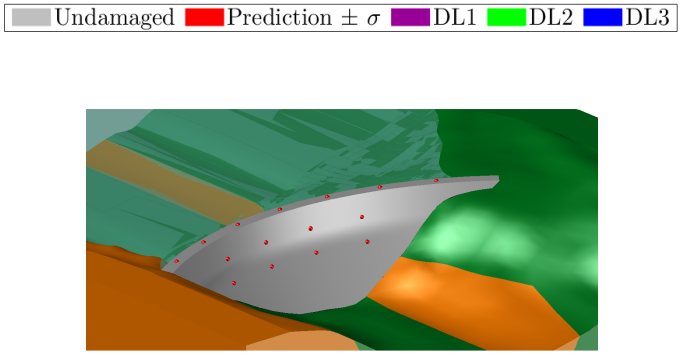
(a) SHM layout.

(b) First mode.

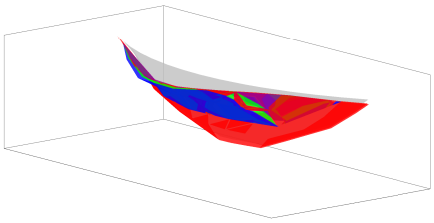
(c) Second mode.

(d) Third mode.

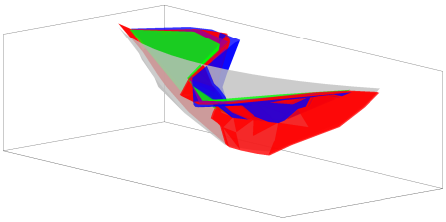
Fig. 8.47: Scandarello dam: comparison between experimental mode shapes and predictions, *Detection Phase*, Layout 1.



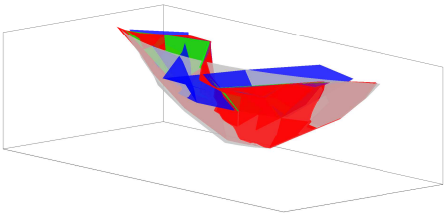
(a) SHM layout.



(b) First mode.



(c) Second mode.



(d) Third mode.

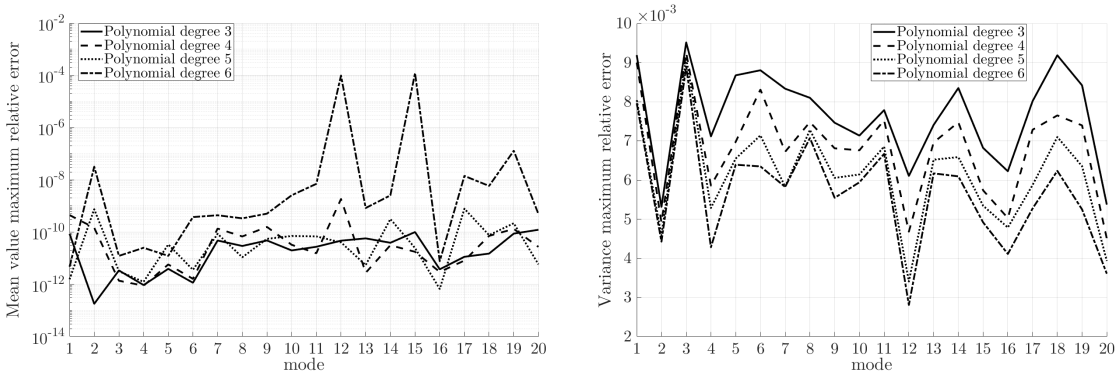
Fig. 8.48: Scandarello dam: comparison between experimental mode shapes and predictions, *Detection Phase*, Layout 2.

characteristic strength and the rock soil is characterised by high resistance, then the only collapse mechanism is related to tensile crack propagation inside the dam body. Since the concrete behaviour is modelled by using the CDP constitutive model (Chapter 3.2.3), and the fracture energy is kept constant, the only epistemic uncertainty is the tensile strength $f_{t,C}$. With regard to the notation introduced in Chapter 7, $\boldsymbol{\theta}_{\mathrm{s}}$ is equal to $f_{t,C}$, that is $\boldsymbol{\theta}_{\mathrm{s}} = f_{t,C}$. The distributions of the hyper-prior parameters of $f_{t,C}$ are defined as those of the elastic ones. Therefore, the mean value $\xi^{\mu}_{f_{t,C}}$ is assumed log-normally distributed with mean value and standard deviation equal to those obtained by the material tests, shown in table 8.10. In the same way $\xi^{\sigma}_{f_{t,C}}$ is assumed log-normally distributed with mean value equal to the standard deviation indicated in table 8.10 and CoV equal to 10%.
In the construction of the meta model of the damaged behaviour $\hat{\phi}^{\mathrm{D}}_{h,i}(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}})$ particularly attention must be placed on the parametrization of the seismic input, as explained in Chapter 7.4. In particular, $S_a(T_1)$ is used as IM of the seismic action, that is $\boldsymbol{\varpi} = S_a(\mathrm{T}_1)$. Whereas, a set of recorded motion is considered in order to account for the variability of the frequency-content. The record-to-record variation is considered by sampling the seismic input from the set defined in table 8.17. The recorded motions used in this application to train the gPCE are the same used for the MR-IDA, reported in table 8.17, while $S_a(T_1)$ is assumed uniformly distributed between 0.1 g and 0.8 g. This range has been defined starting from the fragility curves in order to reduce the number of analyses. To sum up, the seismic input is modelled by sampling a $S_a(\mathrm{T}_1)$ value from the interval 0.1-0.8 g, and an event is sampled from the set reported in table 8.17. Therefore, the sampled seismic event is scaled in order to have $S_a(\mathrm{T}_1)$ equal to the sampled value. Once a seismic action is defined, as indicated in figure 7.2, it is applied as boundary condition at the bottom of the near field after a proper deconvolution Sooch and Bagchi 2014.

8.3.12 Proxy model for modal characteristics of the damaged dam, *Updating Phase*

The predictive model of the modal behaviour of the damaged dam $\hat{\phi}^{\mathrm{D}}_{h,i}(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}})$ is built by using Hermitian polynomials, while 150 analyses are considered for the combination coefficient calculation. The parameters of each analysis are sampled from the prior distributions. A sensitivity analysis is performed with respect to the polynomial expansion degree, in order to select that one which minimizes the errors in terms of mean values and variances. Figure 8.49 shows the maximum relative errors of the mean values and variances among all mesh joints for every considered mode. On the one hand, the mean values

maximum errors increase with the polynomial expansion order. Anyway, polynomial order equal to 3/4/5 still give an acceptable result. On the other hand, the variance maximum relative error decreases with the polynomial expansion order, even though these values are very close. Therefore, a 4th order polynomial expansion is chosen. This polynomial expansion degree leads to small relative errors both for mean values and variances.



(a) Error of the mean value.

(b) Error of the standard deviation.

Fig. 8.49: Maximum relative error of the mode shapes gPCE of the damaged Scandarello dam.

The Sobol's coefficients provide useful information about how the tensile strength influences the mode shapes derivative. For instance, figure 8.50 shows the Sobol's coefficients of the first mode of the dam. Beyond the particular values of the coefficients, it is important to note the differences between them. Indeed, focusing the attention on the central points of the first line, namely d-7, d-8 and d-9, they are little influenced by the variation of the tensile strength because in every case the first mode of the dam has a sinusoidal shape in the upper part of the structure, characterized by only one harmonic. Whereas the other points are influenced by the parameter variation.

8.3.13 Tensile strength parameter updating

Once the gPCE is built, it can be used to update the strength parameters of the material. This is done by solving equation 6.9. In particular, the modified version of MCMC (Chapter 2) is used to solve this task. In particular, after 40000 steps a $\hat{R}^p$ equal to 1.223 is achieved and thus the convergence is ensured. An additional condition is added to the three damage levels previously defined. Indeed, also the case of no damage is simulated in this application. By considering the Amatrice's seismic event this is achieved for $S_a(T_1) = 0.0953$g. The

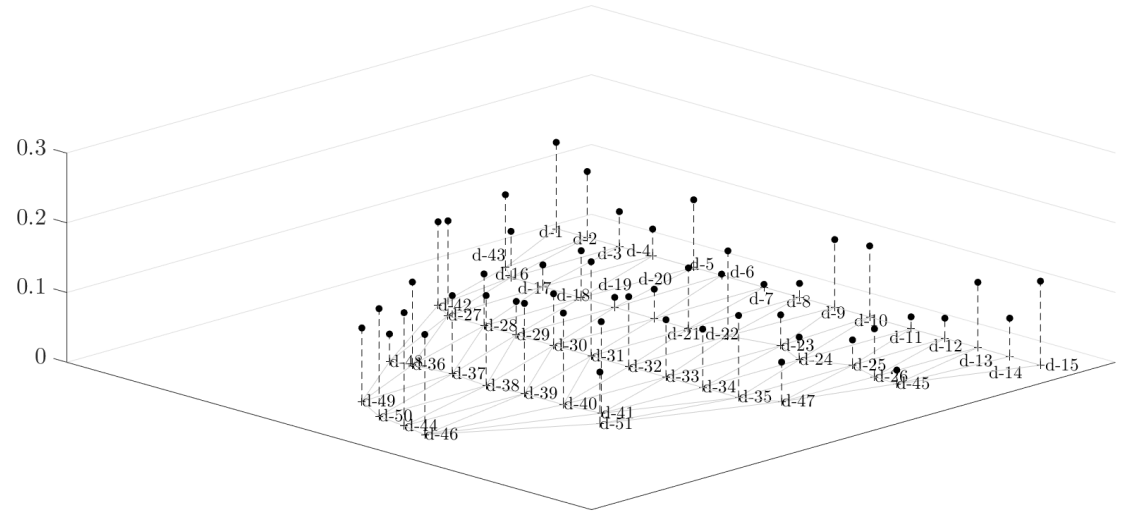


Fig. 8.50: Predictive model Sobol's coefficients of the damaged mode shapes of the Scandarello dam.

updating process is performed only for the Layout 2.
The comparison between prior and posterior distributions is shown in figure 8.51, while table 8.20 summarise their parameters. The results shows that every scenario allows

Table 8.20: Posterior distributions of $f_{t,C}$.

	DL0		DL1		DL2		DL3	
	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]
Mean	1.92	1.02	1.85	0.98	1.72	0.94	1.58	0.87
s.d.	0.15	0.09	0.15	0.07	0.12	0.06	0.12	0.04

updating the strength parameters, but in different way. Increasing the seismic action, and so the DL, the updating is more evident. In every case the mean value of the distribution shows a shift toward the right value, 1.45 MPa. This means that the procedure allows updating the damage parameters but the value of the information is not sufficient to rightly update the distribution with only one measure. Moreover, this results show that the value of the information varies with the intensity of the seismic action.

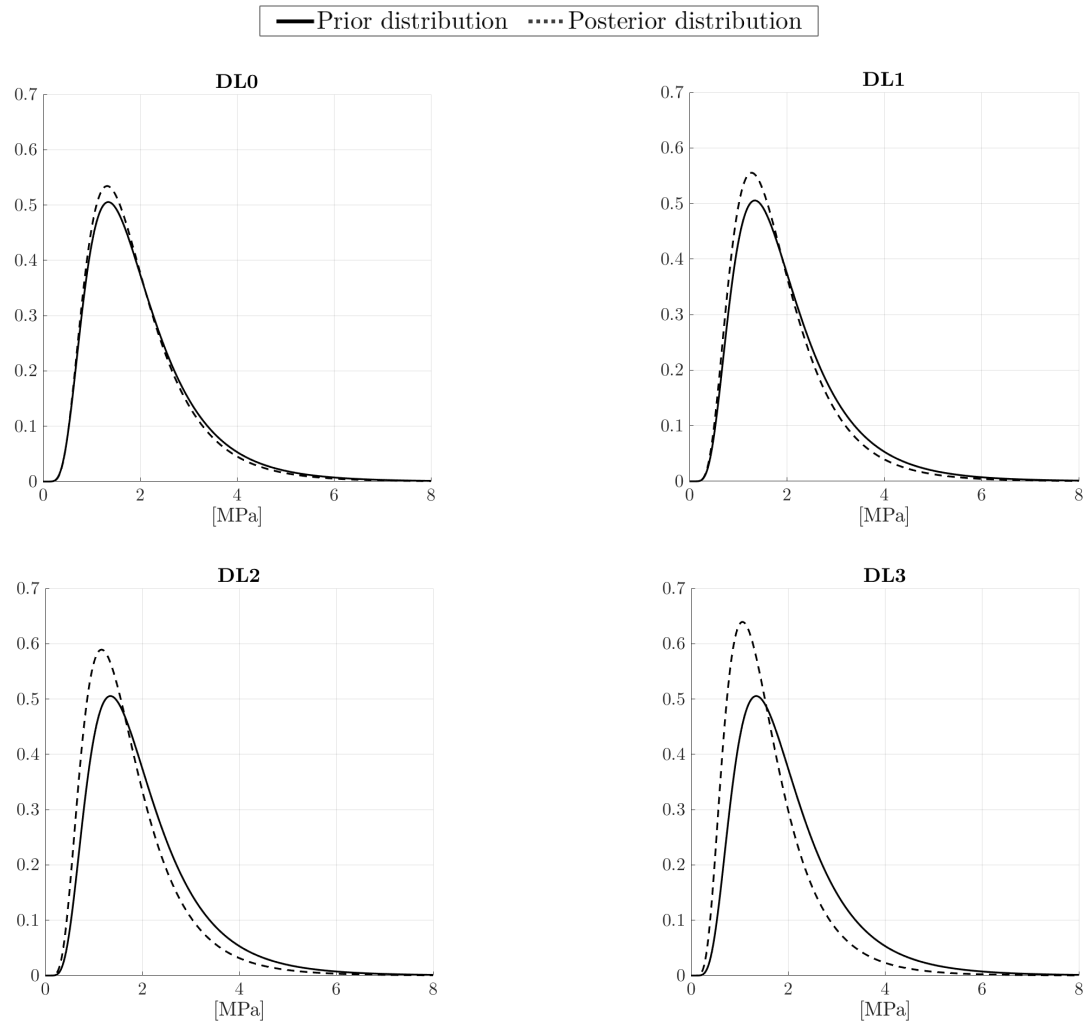


Fig. 8.51: Scandarello dam: Comparison between prior and posterior distributions of $f_{t,C}$, Layout 2.

8.3.14 Parametrization of the experimental domain

In this last application, the procedure proposed in Chapter 7 is used to design the monitoring system of the Scandarello dam. The optimal Bayesian experimental design introduced in Chapter 4.5, is applied in order to maximize the expected utility $U(\mathbf{d})$ with regard to the mode shapes variation, which is the QI for the damage detection. Since the crack initiation and propagation is strictly linked to the strength material parameters, a SHM system able to detect the damage can be also used to acquire useful information for the updating of material strength parameters (Chapter 7).

In the context of the Optimal Bayesian Experimental Design also the experimental variable $\mathbf{d}$ must be parametrized. As introduced in Chapter 7, the polar coordinates are used to parametrize $\mathbf{d} = [r, \alpha]^T$, as shown in figure 8.52. This requires the definition of a sampling range for r and α, which depends on the geometry of the dam. In the present case the ranges of the polar parameters are indicated in table 8.21. Despite the parametrisation of $\mathbf{d}$, some candidate points could lie out from the design domain, as the grey areas shown in figure 8.52. The algorithm 3, introduced in Chapter 7 projects the candidate points which lies externally to the design domain onto the closest edge. The design domain is discretized by considering nodes coming from the FE model. Therefore, the analysis solutions, used to calibrate the gPCE, are determined at nodes of these elements. The value of the quantity of interest of a candidate point is calculated by interpolating the values of the nodes of the region in which the point itself lies. Theoretically, these nodes can coincide with those of the mesh used in the FEA, but in this case only few of them are selected in order to reduce the computational burden. The final mesh of the OED is shown in figure 8.52.

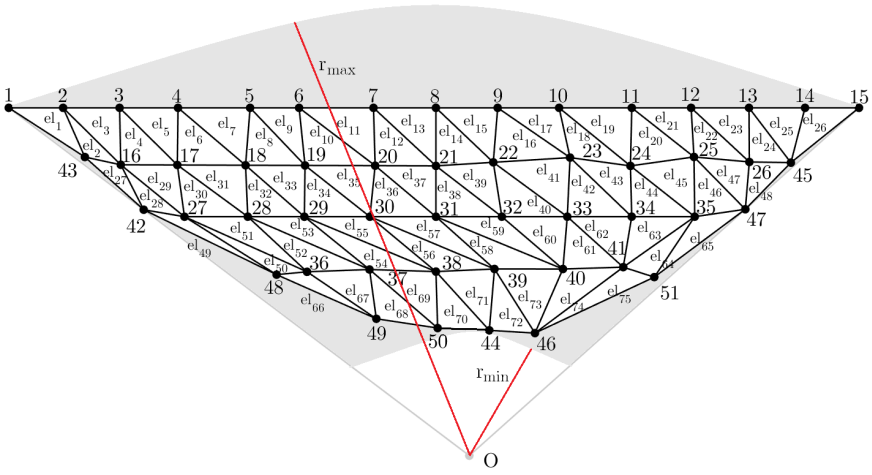


Fig. 8.52: OED mesh of the Scandarello dam.

Table 8.21: Bounds of the design variable $\mathbf{d}$.

	r [m]	α [rad]
max	110.93	1.30
min	25.84	0.00

8.3.15 SHM system optimization through OED

Following the procedure introduced in Chapter 7 the predictive model of the behaviour of the damaged dam $\hat{\phi}^{\mathrm{D}}_{h,i}(\mathbf{x}, \boldsymbol{\varpi}, \boldsymbol{\theta}_{\mathrm{s}}, \boldsymbol{\Xi}_{\mathrm{s}})$ (*Updating Phase*) can be used within the optimization algorithm instead of the FEA solution, thus speeding up the procedure.
The SPSA algorithm requires the definition of some parameters (Chapter 4.5.3), which are indicated in table 8.22. These parameters are derived from indications available in the literature, and empirically checked for the analysed case. A maximum number of 10 devices

Table 8.22: Parameters of the SPSA algorithm for the devices optimization.

a	A	α	c	γ	number of steps
20	100	0.602	10	0.101	3000

is chosen, the optimized layout is shown in figure 8.53, where the devices are plotted in red. The results show that the algorithm prefers to place the devices in the upper part

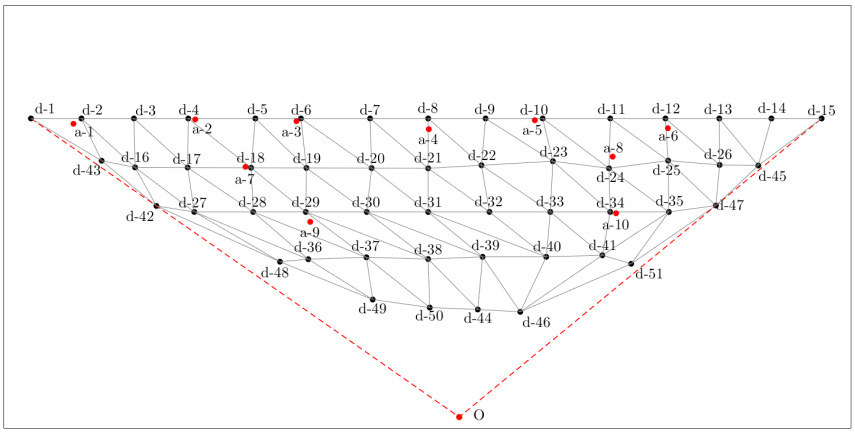


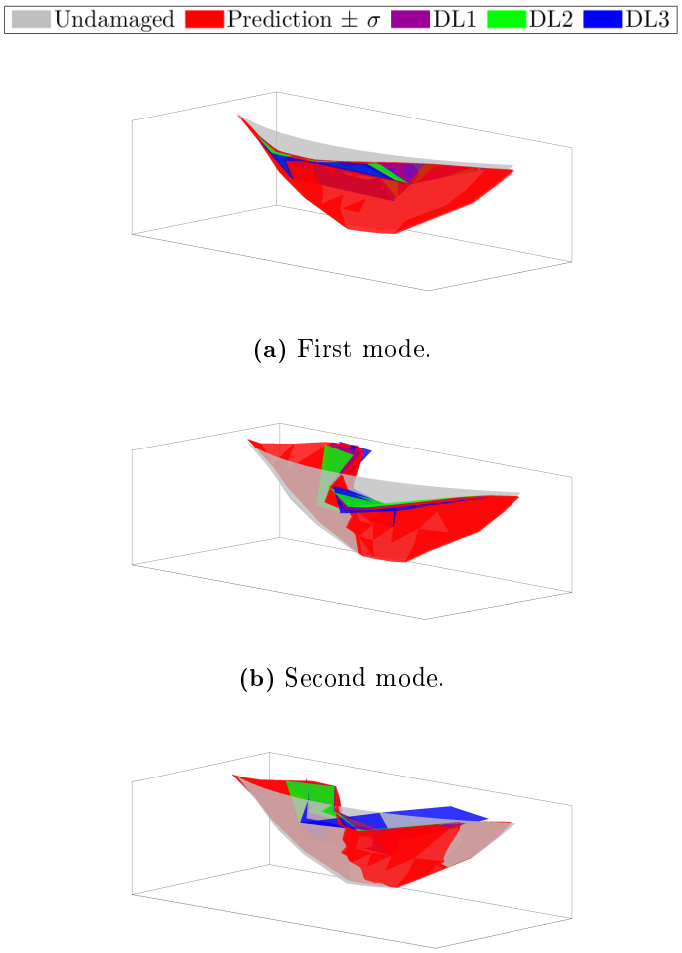
Fig. 8.53: Final SHM system of the Scandarello dam.

of the dam and around the section where the downstream face changes the slope. This means that in these positions the expected utility, defined with regard to damage detection, is maximized. Moreover, these results are coherent with the crack pattern found in the fragility curves calculation (figure 8.44). Indeed, the proposed procedure puts the devices around the section of the downstream where the slope changes. The crack paths obtained in the seismic analyses performed for the fragility curves calculation are concentrated around

the section where the downstream face changes its slope, highlighting how this area can be considered as a geometrical weakness of the dam.

8.3.16 The use of the optimised SHM system for damage detection and material strength parameters updating

In this last section the optimised dynamic SHM system is used for damage detection and strength parameters updating. Therefore, the *Detection Phase* and the *Updating Phase* described in Chapter 7 are repeated for the new optimised layout. As in the case of the initial monitoring system, also in this application the results of the high-fidelity model for the three DLs are used as reference measure (section 8.3.10). Figure 8.54 shows the comparison between the prediction of the undamaged dam behaviour obtained with the predictive model $\hat{\phi}^{\mathrm{U}}_{h,i}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{el}})$ calibrated in the *Training Phase* and the experimental mode shapes calculated with the optimised layout. As observed also in the previous case (Figure 8.48) a non aligned layout allows gathering more information about the dam behaviour, and thus the probability of damage detection increases. Also in this case the proposed SHM system is able to detect the damage for all of the three DLs. Indeed, the experimental mode shapes of the damaged dam are out of the bounds, even tough it is defined considering the standard deviation of the prediction as threshold. As in the previous case the third mode is more sensitive to damage and thus the damaged mode shape is very different to the undamaged one. Finally, from the Layout 2 (Figure 8.48) to the optimised one (Figure 8.54) the accuracy of the damage detection decreases, but the SHM is able to the detect damage with a small number of devices.
These results are used to update the state of knowledge about the tensile strength of the concrete in the *Updating Phase*. As in the previous case, also the DL0, which corresponds to the no-damage scenario, has been considered. The modified version of MCMC, introduced in Chapter 2, is used to solve the inverse problem by considering 40000 steps. In this way, a $\hat{R}^p$ equal to 1.223 is achieved. The posterior distributions of the tensile strength of the dam concrete are shown in figure 8.55, while the statistics are collected in table 8.23.
The posterior distributions of the concrete tensile strength obtained with the optimised layout are very close to the ones obtained in the case of Layout 2. However, this result is achieved by using a smaller number of devices. Also in this case, for every damage scenario the proposed SHM system allows updating the state of knowledge, even tough stronger seismic motions lead to posterior distributions closer to the real value of the tensile strength (1.45 MPa in this application). The quality of the information embedded in the mode shapes after damage is not sufficient to obtain the real value with only one



(a) First mode.

(b) Second mode.

(c) Third mode.

Fig. 8.54: Scandarello dam: comparison between experimental mode shapes and predictions, *Detection Phase*, Optimised Layout.

Table 8.23: Posterior distributions of $f_{t,C}$ determined with the Optimised Layout.

	DL0		DL1		DL2		DL3	
	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]	$\xi^{\mu}_{f_{t,C}}$ [MPa]	$\xi^{\sigma}_{f_{t,C}}$ [MPa]
Mean	1.94	1.07	1.89	0.92	1.75	0.94	1.61	0.92
s.d.	0.17	0.11	0.12	0.07	0.15	0.08	0.11	0.04

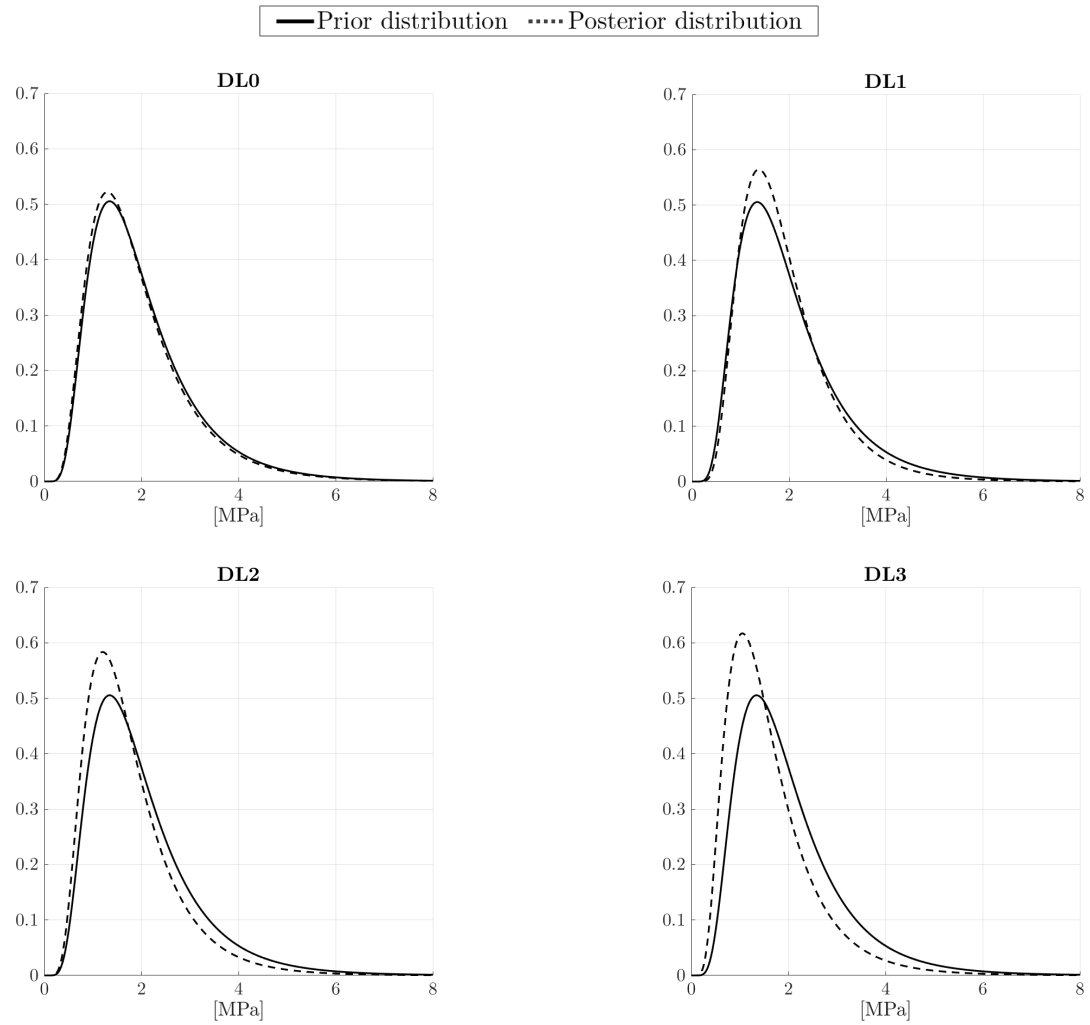


Fig. 8.55: Scandarello dam: Comparison between prior and posterior distributions of $f_{t,C}$, Optimised Layout.

record. The reason for that, as in the previous case, can be found in several factors as the seismic motion parametrization or other bias in the predictive model.

8.4 Concluding remarks

In this last Chapter two dams have been studied: the Gramolazzo dam, in order to test the proposed static SHM framework (Chapter 5), and the Scandarello in order to apply the proposed dynamic SHM framework (Chapter 7).
In the static case the proposed procedure has been successfully used to update the pa-

rameters of a static dam twin model, which can be used to improve the estimation of the structural fragility. The calibrated predictive model has been used to control the structural behaviour during the regular use of the dam. In the case of the Gramolazzo dam, the presence of a structural weakness close to the right abutment, in which the damage is concentrated, leads to a difficult damage detection if only one point is used. Indeed, a network of devices could lead to an improvement in damage detection. Although the use of only one point in the presence of strong damage the system has been able to find it. It is worth noting that the proposed gPCE-based predictive model is very fast and reliable, so it can be successfully used in an on-line SHM.

The application of the probabilistic dynamic SHM to the case of the Scandarello dam showed that it can be successfully used for damage detection purpose. The particular architecture of the proposed SHM system, based on three different steps, allows both detecting a possible damage and updating the mechanical parameters of the materials, thus reducing the uncertainties. In this way, the accuracy of the predictive model (Diagnosis step) and the prediction of the remaining life expectancy of the dam are improved. In particular, the *Training Phase* allows reducing the uncertainties related to the elastic mechanical parameters, thus calibrating the predictive model of the undamaged dam. In the *Detection Phase* the calibrated predictive model is used to control the health state of the structure by comparing the prediction of mode shapes of the undamaged dam with the experimental mode shapes. In the *Updating Phase* the observations recorded during the seismic events are used to update the strength parameters of the materials.

The effects of the uncertainty reduction are quantified in terms of fragility curves. Indeed, the fragility curves of the Scandarello dam are calculated by using first the prior distributions and the posterior ones. The results show that the fragility estimation calculated with the posterior distributions is characterised by less dispersion if compared to the one calculated with the prior distributions.

The last part of the Chapter is dedicated to the optimization of the devices layout in order to maximise both the probability of damage detection and the updating of the strength parameters of the materials. By exploiting the relationship between damage development, concrete tensile strength and mode shape variation, the expected utility, based on the predictive model defined in the *Updating Phase*, can be defined and optimized, thus leading to the final layout of the SHM system.

It is worth noting that the posterior distributions of the tensile strength show the effectiveness of the procedure in strength parameters updating. However, the posterior distributions have mean values higher than the tensile strength fixed in the high-fidelity model. This

is mainly due to the biases contained in the predictive model used to update the state of knowledge on the strength parameters. Moreover, the *Updating Phase* requires the characterisation of the seismic input acquired on the soil around the dam. Therefore, specific devices installed on the soil close to the dam are needed.

Chapter 9

Conclusions

9.1 Introduction

This Chapter summarizes the most important conclusions of the thesis contributing appropriately to the research issues stated in Chapter 1, and which can be listed in the following two points:
(a) *Reduce the epistemic uncertainties involved in the seismic assessment of concrete gravity dams by using all available information about the structures.*

(b) *Improve the dam control by developing suitable SHM system for concrete gravity dams.*

In order to deal with these two points, two probabilistic SHM framework have been developed in this work. More specifically, the two probabilistic SHM framework aim both to control the structural health state and to update the parameters of the predictive models, thus reducing the epistemic uncertainties. The first SHM system is based on static measurements like dam displacements recorded during the regular use, while the second SHM system is based on ambient vibrations records. This two procedures, developed in a Bayesian setting, are applied to real Italian dams, thus showing the feasibility of reduc-

ing uncertainties and of integrating them within a SHM system to improve the structural control. The predictive models can be used in an on-line structural control thanks to its calculation speed achieved with the definition of gPCE-based meta models.
In the last part of the present research work a procedure to design the dynamic SHM system has been introduced. It aims to optimize the expected utility with respect to the damage detection and the strength parameters of the materials. The objective function, which drives the devices optimization, aims to maximize the acquisition of information useful for the calibration of material strength parameters. The procedure, applied also in this case to an Italian dam, shows how to successfully detect damage and how to update the strength parameters of material constitutive models.

9.2 Summary of the work

This research project is composed by 9 Chapters. Chapter 1 presents the topic with the main research questions and objectives.
In the Chapter 2, a literature overview on the seismic assessment of concrete gravity dams and on the use of SHM for model calibration purpose is illustrated.
Chapter 3 shows the most influential issues and their effects in the seismic analysis of concrete gravity dams. At the end of the Chapter the main epistemic uncertainty sources involved in the seismic analysis of concrete gravity dams are introduced.
Chapter 4 introduces the theoretical background and the statistical tools used to develop the procedures introduced in this research work.
Chapter 5 shows the static SHM framework based on the dam displacements.
Chapter 6 describes the procedure to calibrate the parameters of a dynamic twin model of the dam by using the modal characteristics experimentally determined through the operational modal analysis.
Chapter 7 shows the dynamic SHM framework and the procedure developed to design the SHM system itself.
Chapter 8 presents the case studies on which the proposed procedures are applied.
Chapter 9 summarizes the main achievements of the research work.

9.3 Main outcomes

The efficiency of the three proposed procedures are demonstrated by making use of two case studies represented by two Italian concrete gravity dams for which material test results and

real measurements are available. Generally, probabilistic procedures lead to huge numerical efforts. Meta models, based on the gPCE, are used to reduce the computational burden thus making the procedure applicable to real cases.
In the first case the static SHM system is tested on the Gramolazzo dam, because material test results, static measurements and environmental records are available. In particular, half of these information is used to train the gPCE-based meta model, while the other part is used to demonstrate the ability of the probabilistic model to well predict the dam behaviour during its regular use. The procedure to integrate the predictive model of the dam displacements within the SHM system is shown. Since no information about the behaviour of the damaged dam are available, a high-fidelity model of the structure is used to simulate it by assuming two different damage scenarios. The proposed procedure is deeply analysed in Chapter 5. The main results of this application can be summarized as follow:

- The use of a gPCE-based meta model instead of the FEA solution allows speeding up the procedure.
- Material test results, environmental measurements and static displacements can be successfully used to calibrate a probabilistic predictive model of the dam displacements. The calibration procedure defined in a Bayesian setting allows determining the posterior distribution of the model parameters.
- The calibrated predictive model can be used within a SHM framework to control the structural behaviour during the regular use of the dam.
- The procedure allows determining the global error standard deviation, which can be also used to define a threshold beyond which the dam shows abnormal phenomena.
- The procedure can be repeated every time that new measurements are recorded. If the procedure is integrated in SHM systems it allows continuously improving the model parameters estimation, and the probabilistic model prediction.

A second case study is investigated in order to test the dynamic SHM framework introduced in this research work. In particular, the Scandarello dam is selected because real ambient vibration records are available in addition to the material test results. First of all, the parameters of a dynamic twin model and those of the predictive models of the modal characteristics are calibrated by using the results of the OMA, through the procedure introduced in Chapter 6. The computational burden is strongly reduced by using

gPCE-based meta models to reproduce the FEA results in terms of frequencies and mode shapes. For comparison reasons, the predictive models of frequencies and mode shapes are separately calibrated.
The introduction of covariance functions for the error terms, in the case of the mode shape predictive model, allows reducing the computational burden.
The use of modal characteristics as reference measure in the updating process requires ensuring the coherence between experimental and numerical modes. More specifically, for a particular sample numerical modes could be ordered differently from the experimental ones. Before the likelihood calculation numerical results must be reordered coherently with the experimental modes. Considering the SSI in the numerical model a large number of numerical modes related to the soil mass with no experimental correlation are calculated. A selection criterion is introduced in the numerical algorithm MCMC in order to solve this problems, thus reordering the numerical modes coherently with the experimental ones, and discarding the modes with no experimental correlation.
The calibrated dynamic twin model is used to estimate the structural fragility through the MR-IDA. The comparison between fragility curves, calculated considering prior and posterior distributions of the model parameters, shows an improvement in the estimation of the seismic behaviour of the dam. The procedure is introduced and deeply described in Chapter 6. The main results of this application can be summarized as follows:

- The use of gPCE-based meta models instead of the FEA outputs allows speeding up the procedure.
- Material test results and the OMA results calculated by the elaboration of ambient vibrations allows calibrating the parameters of dynamic twin models of the dam. A dynamic twin model can be used both to predict the structural behaviour during seismic events, and to build a predictive model of the modal characteristics which can be integrated in a SHM framework, thus improving the damage detection system.
- The proposed modified version of MCMC, defined by the introduction of a selection criterion in order to ensure the coherence between experimental and numerical modes, allows considering the SSI within dam numerical models.
- The use of covariance functions for the definition of the error terms of the mode shapes predictive model allows reducing the computational burden. The case of study demonstrates that exponential functions can be successfully used as covariance ones.
- The calibrated dynamic twin model of the dam leads to a better estimation of the

structural fragility, because of the reduction of epistemic uncertainties obtained with the use of the proposed procedure.

Starting from the results obtained with the application of the procedure introduced in Chapter 6, the dynamic SHM framework proposed in Chapter 7 is applied to the Scandarello dam. The particular architecture of the proposed dynamic SHM system, based on three different phases, allows both controlling the health state of the structure and updating the material mechanical parameters, elastic and strength ones. The use of two different predictive models, the one defined in Chapter 6, and one which reproduces the behaviour of the damaged dam, allows reducing the computational burden, thus speeding up the procedure and making possible its use in real time control. Three damage scenarios are simulated through the use of a high-fidelity model in order to test the effectiveness of the proposed SHM both in damage detecting and model parameters updating. The results show that the procedure has been successfully used to achieve these two goals, even tough the posterior distributions of the tensile strength parameters are not fully updated. This fact can be due to the bias of the predictive model of the damaged dam.
The optimization procedure described at the end of Chapter 7 is used to design the dynamic SHM system. The resolution of the experimental design problem requires the parametrization of the design domain. Therefore, a projection/interpolation procedure is implemented and integrated in the algorithm for the resolution of the OED, in order to ensure the continuity of the design domain and to limit it by the definition of geometrical criteria.
Once the maximum devices number is decided, the proposed procedure allows determining the best device position in order to maximise the probability of damage detection and the acquisition of information for strength parameters updating.
The results obtained in the case of the optimised layout are compared with those obtained in the case of a non optimised layout. The comparison shows that the optimised layout allows reaching the same performance of the non optimised one but with a less number of devices. The main results of this application can be summarized as follows:

- The use of gPCE-based meta models instead of the FEA outputs allows speeding up the procedure, thus making possible the use of the proposed SHM in real time structural control.

- The proposed dynamic SHM system allows both controlling the health state of the structure and updating the mechanical parameters of the materials, thus reducing the epistemic uncertainties. For this purpose, some devices of the dynamic SHM system must be installed on the soil around the dam in order to make possible the

seismic input characterisation.

- The proposed optimisation algorithm allows solving the problem for a continuous design space, and if necessary limiting it by inserting geometrical criteria, and selecting only the significant modes.

- The designed SHM system shows better performance than a non-optimised SHM system, once defined the minimum level of damage to be recorded or the maximum number of devices.

The two proposed SHM systems can be used in a real time structural control both to control the health state of the dam and to update the mechanical parameters of the materials, thus improving the prediction itself. The use of gPCE-based meta models and the parallel algorithms developed in MATLAB allows speeding up the procedure even without High-Performance Computing (HPC). Finally, the proposed updating procedures can be employed every time new information are available. If the procedures are integrated in a SHM system, the calibration of the parameters and the prediction of the probabilistic models are improved.

9.4 Future developments

On the basis of the present work, with the challenge of making even more possible the integration of calibrated predictive models within the SHM system of concrete gravity dams, some efforts are still necessary:

- *Probabilistic framework for static SHM system of concrete gravity dams*:
  - More studies on the relationship between environmental conditions, e.g. humidity and thermal radiation, and dam behaviour are necessary to improve the predictive model. This improvement should reduce the global error standard deviation, thus improving the estimation of the QI.
  - The use of devices network, in conjunction to two or three dimensional displacement records, could lead to a better resolution and then to a better control during the regular use of the structure.
  - A sensitivity analysis to investigate the relationship between damage and static displacements after seismic events could lead to a good setting of the displacement threshold.

- *Probabilistic framework for dynamic SHM system of concrete gravity dams*:
  - Since the method is based on the results of the OMA, a sensitivity analysis of the OMA technique on the final results could lead to determine the best OMA technique.
  - Different selection criteria within the proposed modified version of MCMC should be assumed in order to determine the most efficient in the mode selection.
  - Different covariance functions should be tested in order to assess their influence on the final results.
  - Verify the effect of the epistemic uncertainty reduction with respect to fragility curves calculated by using other LSs function.
  - Concrete gravity dams with vertical contraction joints should be analysed with the aim to evaluate the possibility to update the shear strength parameters.
  - Investigate the effect of the strength parameters uncertainty reduction in terms of fragility curves.
  - Detect the structural damage by combining the mode shapes with other quantity of interest, e.g. frequencies, deformations, vibrations, etc.
  - Verify the influence of environment conditions, e.g. humidity, water and air temperatures, on the modal characteristic predictive models.
  - Investigate the effect of relationship between damage development and structural behaviour in order to define a threshold beyond which the damage must be detected.
  - Use the proposed procedure to determine the minimum number of devices needed to detect a minimal level of damage, and optimize their positions.

Bibliography

ABAQUS (2014). *ABAQUS documentation.*

Akköse, M. and E. Şimşek (2010). "Non-linear seismic response of concrete gravity dams to near-fault ground motions including dam-water-sediment-foundation interaction". In: *Appl. Math. Model.* 34.11, pp. 3685–3700.

Allemang, R. J. (2003). "The modal assurance criterion - Twenty years of use and abuse". In: *Sound Vib.* 37.8, pp. 14–21.

American Society of Civil Engineers (2013). *Report Card for America's Infrastructure.*

Anderson, T. W. and D. A. Darling (1954). "A Test of Goodness of Fit". In: *J. Am. Stat. Assoc.*

Andreini, M., P. Gardoni, S. Pagliara, and M. Sassu (2016). "Probabilistic Models for Erosion Parameters and Reliability Analysis of Earth Dams and Levees". In: *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.*

Ansari, M. I. and P. Agarwal (2016). "Categorization of Damage Index of Concrete Gravity Dam for the Health Monitoring after Earthquake". In: *J. Earthq. Eng.*

ANSYS (2013). "ANSYS Mechanical APDL Theory Reference". In: *ANSYS Inc* Release15, p. 909.

ASDSO (2011). "State and federal oversight of dam safety must be improved". In: *Mag. Assoc. State Dam Saf. Off.*

ATC-13 (1985). *Earthquake Damage Evaluation Data for California.* Tech. rep.

Au, S. K., F. L. Zhang, and Y. C. Ni (2013). "Bayesian operational modal analysis: Theory, computation, practice". In: *Comput. Struct.* 126.1, pp. 3–14.

Azzara, R. M., G. De Roeck, M. Girardi, C. Padovani, D. Pellegrini, and E. Reynders (2018). "The influence of environmental parameters on the dynamic behaviour of the San Frediano bell tower in Lucca". In: *Eng. Struct.*

Baker, J. W. (2015). "Efficient analytical fragility function fitting using dynamic structural analysis". In: *Earthq. Spectra.*

Banon, H. and D. Veneziano. "Seismic safety of reinforced concrete members and structures". In: *Earthq. Eng. Struct. Dyn.* 10.2 (), pp. 179–193.

Bassoli, E., L. Vincenzi, A. M. D'Altri, S. de Miranda, M. Forghieri, and G. Castellazzi (2018). "Ambient vibration-based finite element model updating of an earthquake-damaged masonry tower". In: *Struct. Control Heal. Monit.*, e2150.

Bažant, Z. P. and L. Cedolin (1979). "Blunt Crack Band Propagation in Finite Element Analysis". In: *J. Eng. Mech. Div.* 105.2, pp. 297–315.

Bažant, Z. P. and B. H. Oh (1983). "Crack band theory for fracture of concrete". In: *Matériaux Constr.* 16.3, pp. 155–177.

Beck, J. L. (2010). "Bayesian system identification based on probability logic". In: *Struct. Control Heal. Monit.* 17, pp. 825–847.

Beck, J. L., A. Siu-Kui, and M. W. Vanik (2001). "Monitoring structural health using a probabilistic measure". In: *Comput. Civ. Infrastruct. Eng.* 16.1, pp. 1–11.

Behrouz, A. N. (2002). "Multivariate statistical analysis of monitoring data for concrete dams". PhD thesis. McGill University, Montreal.

Berenger, J.-P. (1994). "A perfectly matched layer for the absorption of electromagnetic waves". In: *J. Comput. Phys.*

Bernier, C., J. E. Padgett, J. Proulx, and P. Paultre (2014). "Seismic Fragility of Concrete Gravity Dams With Modeling Parameter Uncertainty and Spatial Variation". In: *Am. Soc. Civ. Eng.*

— (2016a). "Seismic Fragility of Concrete Gravity Dams with Spatial Variation of Angle of Friction: Case Study". In: *J. Struct. Eng.*

Bernier, C., R. Monteiro, and P. Paultre (2016b). "Using the conditional spectrum method for improved fragility assessment of concrete gravity dams in Eastern Canada". In: *Earthq. Spectra.*

Bianchi, M. and R. Bremen (2001). "Health monitoring of arch dams: recent developments TT - Bauinstandsetzen und Baudenkmalpflege: eine internationale Zeitschrift". In: *Restor. Build. Monum. an Int. J. = Bauinstandsetz. und Baudenkmalpfl. eine Int. Zeitschrift.*

Bobrowski, A. (2005). *Functional Analysis for Probability and Stochastic Processes* .

Box, G. E. P. and D. R. Cox (1964). "An Analysis of Transformations". In: *J. R. Stat. Soc. Ser. B* 26.2, pp. 211–252.

Box, G. E. P. and G. C. Tiao (1992). *Bayesian Inference in Statistical Analysis*, p. 588.

Brincker, R. and C. E. Ventura (2015). *Introduction to Operational Modal Analysis.*

Brooks, S. P. B. and A. G. Gelman (1998). "General methods for monitoring convergence of iterative simulations". In: *J. Comput. Graph. Stat.* 7.4, pp. 434–455.

Brownjohn, J. M. W. (1990). *Dynamic investigation of Hermitage Dam, Jamaica.* Tech. rep. Bristol: UBCE-EE-90-13 University of Bristol Department of Civil Engineering.

Brownjohn, J. M. W., R. T. Severn, and C. A. Taylor (1986). *Ambient vibration survey of Contra dam.* Tech. rep. Bristol: University of Bristol Department of Civil Engineering.

Brühwiler, E. and F. H. Wittmann (1990). "Failure of dam concrete subjected to seismic loading conditions". In: *Eng. Fract. Mech.* 35.1-3, pp. 565–571.

Bukenya, P., P. Moyo, and C. Oosthuizen (2012a). *Comparative study of operational modal analysis techniques using ambient vibration measurements of a concrete dam.*

— (2012b). "Modal parameter estimation from ambient vibration measurements of a dam using stochastic subspace identifcation methods." In: *Proc. 3rd Int. Conf. Concr. Repair, Rehabil. Retrofit.* Cape Town.

Cantieni, R., U. Wiberg, and S. Pietrzko (2004). "Modal investigation of a dam." In: *Proc. 16th Int. Modal Anal. Conf.* Santa Barbara.

Celarec, D. and M. Dolšek (2013). "The impact of modelling uncertainties on the seismic performance assessment of reinforced concrete frame buildings". In: *Eng. Struct.*

Celik, O. C. and B. R. Ellingwood (2010). "Seismic fragilities for non-ductile reinforced concrete frames - Role of aleatoric and epistemic uncertainties". In: *Struct. Saf.* 32.1, pp. 1–12.

Chatzi, E. N. (2016). *Identification Methods for Structural Health Monitoring.* Ed. by Eleni N. Chatzi and Costas Papadimitriou. 1st ed. Vol. 567. Springer International Publishing.

Chen, W. F. (2007). *Limit Analysis and Soil Plasticity.* Elsevier Scientific Pub. Co Amsterdam ; New York, p. 638.

Cheng, L., J. Yang, D. Zheng, B. Li, and J. Ren (2015). "The Health Monitoring Method of Concrete Dams Based on Ambient Vibration Testing and Kernel Principle Analysis". In: *Schock Vib.*

Chopra, A. K. (1967). "Hydrodynamic Pressures on Dams During Earthquakes". In: *J. Eng. Mech. Div.* 93.6, pp. 205–224.

Chouinard, L., R. Larivière, P. Côté, and W. Zhao (2006). "Analysis of irreversible displacements in multiple arch concrete dam using principal component analysis". In: *Jt. Int. Conf. Comput. Decis. Mak. Civ. Build. Eng.* Montréal.

Clough, R. W., K. W. Chang, and R. W. Stephen (1986). *Vibration behaviour of Xiang Hong Dian dam.* Tech. rep. UC, Berkeley.

Clough, R.W. (1980). "Non-linear mechanisms in the seismic response of arch dams". In: *Proc. Int. Conf. Earthq. Eng.* Skopje, pp. 669–684.

COMSOL Multiphysics (2009). "Introduction to COMSOL Multiphysics". In: *Manual*.

Cornell, C. A., F. Jalayer, R. O. Hamburger, and D. A. Foutch (2002). "Probabilistic Basis for 2000 SAC Federal Emergency Management Agency Steel Moment Frame Guidelines". In: *J. Struct. Eng.*

Cosenza, E. and G. Manfredi (2000). "Damage indices and damage measures". In: *Prog. Struct. Eng. Mater.*

Council, National Research (1990). *Earthquake Engineering for Concrete Dams: Design, Performance, and Research Needs*. Washington, DC: The National Academies Press.

Cover, T. M. and J. A. Thomas (2005). *Elements of Information Theory*.

Dai, B., C. Gu, E. Zhao, and X. Qin (2018). "Statistical model optimized random forest regression model for concrete dam deformation monitoring". In: *Struct. Control Heal. Monit.* 25, e2170.

Daniell, W. E. and C. A. Taylor (1999). "Effective ambient vibration testing for validating numerical models of concrete dams". In: *Earthq. Eng. Struct. Dyn.*

Darbre, G. R., C. A. M. De Smet, and C. Kraemer (2000). "Natural frequencies measured from ambient vibration response of the arch dam of Mauvoisin". In: *Earthq. Eng. Struct. Dyn.*

Davis, J. (2016). *Methods of Applied Mathematics with a Software Overview*. 2nd ed. Birkhäuser Basel, p. 781.

De Falco, A., M. Mori, G. Sevieri, and N. Zani (2017). "Simulation of concrete crack development in seismic assessment of existing gravity dams". In: *XVII CONVEGNO ANIDIS "L'Ingegneria Sismica Ital.* Pistoia.

De Sortis, A. and P. Paoliani (2007). "Statistical analysis and structural identification in concrete dam monitoring". In: *Eng. Struct.*

Demirkaya, S. and B. Balcilar (2012). "The Contribution of Soft Computing Techniques for the Interpretation of Dam Deformation". In: *Proc. FIG Work. week 2012*. Rome.

Der Kiureghian, A. and O. Ditlevsen (2009). "Aleatory or epistemic? Does it matter?" In: *Struct. Saf.* 31.2, pp. 105–112.

Ditlevsen, O. and H. O. Madsen (2007). "Structural Reliability Methods". In: *Mech. Eng.*

Dolšek, M. (2012). "Simplified method for seismic risk assessment of buildings with consideration of aleatory and epistemic uncertainty". In: *Struct. Infrastruct. Eng.* 8.10, pp. 939–953.

Donoho, D. L. (2006). "Compressed sensing". In: *IEEE Trans. Inf. Theory*.

Efron, B. and C. Stein (1981). "The Jackknife Estimate of Variance THE JACKKNIFE ESTIMATE OF VARLANCE". In: *Source Ann. Stat. Ann. Stat.*

Ellis, E., Z. Duron, N. Von Gersdorff, and N. Knarr (2010). "Dynamic characterization of a large multiple arch dam." In: *Proc. 30th Annu. USSD Conf.* Sacramento.

Faber, M. H. (2005). "On the Treatment of Uncertainties and Probabilities in Engineering Decision Analysis". In: *J. Offshore Mech. Arct. Eng.*

Feltrin, G. (1997). *Absorbing boundaries for the time-domain analysis of dam-reservoir-foundation systems.* Tech. rep. Report Swiss Federal Institute of Technology Zurich.

Fisher, R. A. (1922). "On the Mathematical Foundations of Theoretical Statistics". In: *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 222.594-604, pp. 309–368.

— (1925). "Theory of Statistical Estimation". In: *Math. Proc. Cambridge Philos. Soc.*

Gamerman, D. and H. F. Lopes (2006a). "Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference". In: *CRC Press.*

— (2006b). *Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference.* Chapman and Hall/CRC, p. 342.

Gardoni, P., A. Der Kiureghian, and K. M. Mosalam (2002a). "Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations". In: *J. Eng. Mech.* 128.10, pp. 1024–1038.

— (2002b). *Probabilistic models and fragility estimates for bridge components and systems.* Tech. rep. Pacific Earthquake Engineering Research Center, College of Engineering, Univ. of California, Berkeley, CA.

Gehl, P., D. M. Seyedi, and J. Douglas (2013). "Vector-valued fragility functions for seismic risk evaluation". In: *Bull. Earthq. Eng.*

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis.*

Ghaemian, M. and A. Ghobarah (1999). "Nonlinear seismic response of concrete gravity dams with dam-reservoir interaction". In: *Eng. Struct.* 21.4, pp. 306–315.

Ghanaat, Y. (2004). *Failure modes approach to safety evaluation of dams.*

Ghanaat, Y., S. Hashimoto Philip, O. Zuchuat, and R. P. Kennedy (2011). "Seismic fragility of Muhlberg dam using nonlinear analysis with Latin Hypercube Simulation". In: *31st Annu. USSD Conf. San Diego, California, April 11-15.*

Ghanaat, Y., R. C. Patev, and A. K. Chudgar (2012). "Seismic fragility analysis of concrete gravity dams". In: *15th World Conf. Earthq. Eng. Lisbon Port.*

— (2015). "Seismic fragility for risk assessment of concrete gravity dams". In: *Proceeding 2015 USSD Annu. Conf.* Pp. 645–60.

Ghanem, R. G. and P.D. Spanos (1991). *Stochastic Finite Elements: A Spectral Approach*, p. 222.

Ghrib, F. and R. Tinawi (1995). "An application of damage mechanics for seismic analysis of concrete gravity dams". In: *Earthq. Eng. Struct. Dyn.* 24.2, pp. 157–173.

Gioia, G., Z. P. Bažant, and B. P. Pohl (1992). "Is no-tension dam design always safe? – A numerical study". In: *De* 3, pp. 23–34.

Givoli, D. (2004). "High-order local non-reflecting boundary conditions: A review". In: *Wave Motion*.

Gopalakrishnan, S., M. Ruzzene, and S. Hanagud (2011). *Computational Techniques for Structural Health Monitoring*. 1st. Springer Publishing Company, Incorporated.

Guanglun, W., O. A. Pekau, Z. Chuhan, and W. Shaomin (2000). "Seismic fracture analysis of concrete gravity dams based on nonlinear fracture mechanics". In: *Eng. Fract. Mech.* 65.1, pp. 67–87.

Hall, J. F. (1988). "The dynamic and earthquake behaviour of concrete dams: review of experimental behaviour and observational evidence". In: *Soil Dyn. Earthq. Eng.* 7.2, pp. 58–121.

Hariri-Ardebili, M. A. and V. Saouma (2015). "Quantitative failure metric for gravity dams". In: *Earthq. Eng. Struct. Dyn.*

Hariri-Ardebili, M. A. and V. E. Saouma (2016a). "Collapse Fragility Curves for Concrete Dams: Comprehensive Study". In: *J. Struct. Eng.*

— (2016b). "Probabilistic seismic demand model and optimal intensity measure for concrete dams". In: *Struct. Saf.*

— (2016c). "Seismic fragility analysis of concrete dams: A state-of-the-art review". In: *Eng. Struct.* 128.October, pp. 374–399.

— (2016d). "Sensitivity and uncertainty quantification of the cohesive crack model". In: *Eng. Fract. Mech.*

Hariri-Ardebili, M. A., S. Sattar, and H. E. Estekanchi (2014). "Performance-based seismic assessment of steel frames using endurance time analysis". In: *Eng. Struct.*

Hariri-Ardebili, M. A., S. M. Seyed-Kolbadi, and M. R. Kianoush (2016a). "FEM-based parametric analysis of a typical gravity dam considering input excitation mechanism". In: *Soil Dyn. Earthq. Eng.* 84, pp. 22–43.

Hariri-Ardebili, M. A., V. E. Saouma, and K. A. Porter (2016b). "Quantification of seismic potential failure modes in concrete dams". In: *Earthq. Eng. Struct. Dyn.*

Hariri-Ardebili, M. A., S. M. Seyed-Kolbadi, V. E. Saouma, J. Salamon, and B. Rajagopalan (2018). "Random finite element method for the seismic analysis of gravity dams". In: *Eng. Struct.* 171, pp. 405–420.

Hariri-Ardebili, M. A., S. Mahdi Seyed-Kolbadi, V. E. Saouma, J. W. Salamon, and L. K. Nuss (2019). "Anatomy of the vibration characteristics in old arch dams by random field theory". In: *Eng. Struct.* 179, pp. 460–475.

Hebbouche, A., M. Bensaibi, and H. Mroueh (2013). "Seismic Fragility and uncertainty Analysis of Concrete Gravity Dams under Near-Fault Ground Motions". In: *Civ. Environ. Res.* 5, pp. 123–9.

Henriques, M. J., J. N. Lima, and S. Oliviera (2012). "Measuring Inclinations in Cabril Dam with an Optoelectronic Sensor". In: *FIG Work. Week 2012*. Rome.

Hillerborg, A., M. Modéer, and P. E. Petersson (1976). "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements". In: *Cem. Concr. Res.* 6.6, pp. 773–781.

Hoeffding, W. (1948). "A Class of Statistics with Asymptotically Normal Distribution". In: *Ann. Math. Stat.*

Holzapfel, G. (2000). *Nonlinear solid mechanics: A continuum approach for engineering.*

Homma, T. and A. Saltelli (1996). "Importance measures in global sensitivity analysis of nonlinear models". In: *Reliab. Eng. Syst. Saf.*

Hu, J., F. Ma, and S. Wu (2018a). "Anomaly identification of foundation uplift pressures of gravity dams based on DTW and LOF". In: *Struct. Control Heal. Monit.* 25, e2153.

— (2018b). "Comprehensive investigation of leakage problems for concrete gravity dams with penetrating cracks based on detection and monitoring data: A case study". In: *Struct. Control Heal. Monit.* 25, e2127.

Huan, X. and Y. M. Marzouk (2013). "Simulation-based optimal Bayesian experimental design for nonlinear systems". In: *J. Comput. Phys.* 232.1, pp. 288–317.

Huang, Y., C. Shao, B. Wu, J. L. Beck, and H. Li (2018). "State-of-the-art review on Bayesian inference in structural system identification and damage assessment". In: *Adv. Struct. Eng.*

Ibrahimbegovic, A. (2009). "Nonlinear solid mechanics". In: *Solid Mech. its Appl.* 160, pp. 1–594.

ICOLD (2000). *Bullettin 118: Automated Dam Monitoring Systems Guidelines and Case Histories.* Tech. rep. Paris: ICOLD.

Ingraffea, A. R. (2007). "Computational Fracture Mechanics". In: *Encycl. Comput. Mech.*

Iooss, B. and P. Lemaître (2014). "A review on global sensitivity analysis methods". In: *Uncertain. Manag. Simulation-Optimization Complex Syst. Algorithms Appl.*, p. 23.

ISS: Indagini conoscitive e monitoraggi nell'ambito dell'Osservatorio Sismico delle Strutture.

Jaishi, B. and W.-X. Ren (2005). "Structural Finite Element Model Updating Using Ambient Vibration Test Results". In: *J. Struct. Eng. ASCE* 131.4, pp. 617–628.

Johnson, R. A. (1967). "An Asymptotic Expansion for Posterior Distributions". In: *Ann. Math. Stat.*

Johnson, S. G. (2008). "Notes on perfectly matched layers (PMLs)". In: *Lect. notes, Massachusetts Inst. ...*

Ju, B. S. and W. Jung (2015). "Evaluation of seismic fragility of weir structures in South Korea". In: *Math. Probl. Eng.*

Kadkhodayan, V., S. Meisam Aghajanzadeh, and H. Mirzabozorg (2015). "Seismic Assessment of Arch Dams Using Fragility Curves". In: *Civ. Eng. J.*

Kang, F., J. Liu, J. Li, and S. Li (2017). "Concrete dam deformation prediction model for health based on extreme learning machine". In: *Struct. Control Heal. Monit.* 24.e1997.

Kao, C.-Y. and C.-H. Loh (2013). "Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches". In: *Struct. Control Heal. Monit.* 20, pp. 282–303.

Kemp, B. G. (1996). "Ambient vibration assessment of Ruskin dam: dynamic properties." PhD thesis. University of British Columbia (UBC).

Kennedy, R. P., C. A. Cornell, R. D. Campbell, S. Kaplan, and H. F. Perla (1980). "Probabilistic seismic safety study of an existing nuclear power plant". In: *Nucl. Eng. Des.*

Kim, D. K. and C. B. Yun (2000). "Time-domain soil - Structure interaction analysis in two-dimensional medium based on analytical frequency-dependent infinite elements". In: *Int. J. Numer. Methods Eng.*

Kottegoda, N. T. and R. Rosso (2008). *Applied Statistics for Civil and Environmental Engineers*, p. 718.

Kuhlemeyer, R. L. and J. Lysmer (1973). "Finite Element Method Accuracy for Wave Propagation Problems". In: *J. Soil Mech. Found. Div.* 99.5, pp. 421–427.

Lahmer, T. (2011). "Optimal experimental design for nonlinear ill-posed problems applied to gravity dams". In: *Inverse Probl.* 27.12, p. 125005.

Lee, J. and G. L. Fenves (1998a). "A plastic-damage concrete model for earthquake analysis of dams". In: *Earthq. Eng. Struct. Dyn.* 27.9, pp. 937–956.

Lee, J. H. and G. L. Fenves (1998b). "Plastic-damage model for cyclic loading of concrete structures". In: *J. Eng. Mech.* 124.8, pp. 892–900.

Léger, P. (2007). "Reducing the earthquake induced damage and risk in monumental structures: experience at Ecole Polytechnique de Montreal for large concrete dams supported by hydro-Quebec and Alcan ALCAN". In: *Extrem. Man-Made Nat. Hazards Dyn. Struct.* Ed. by Adnan Ibrahimbegovic and Ivica Kozar. Dordrecht: Springer Netherlands, pp. 285–309.

Léger, P. and S. Seydou (2009). "Seasonal Thermal Displacements of Gravity Dams Located in Northern Regions". In: pp. 166–174.

Liel, A. B., C. B. Haselton, G. G. Deierlein, and J. W. Baker (2009). "Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings". In: *Struct. Saf.* 31.2, pp. 197–211.

Lilliefors, H. W. (1967). "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown". In: *J. Am. Stat. Assoc.*

Lin, C., T. Li, X. Liu, L. Zhao, S. Chen, and H. Qi (2019). "A deformation separation method for gravity dam body and foundation based on the observed displacements". In: *Struct. Control Heal. Monit.* 26, e2304.

Lin, L. and J. Adams (2007). "Lessons for the fragility of Canadian hydropower components under seismic loading". In: *Proc. Ninth Can. Conf. Earthq. Eng.* 1771, pp. 1762–1771.

Lindley, D. V. (1956). "On a Measure of the Information Provided by an Experiment". In: *Ann. Math. Stat.* 27.4, pp. 986–1005.

— (1972). "Bayesian statistics : a review." In: *Book.*

— (2000). "The philosophy of statistics". In: *J. R. Stat. Soc. Ser. D Stat.*

Loh, C.-H. and T.-S. Wu (1996). "Identification of Fei-Tsui arch dam from both ambient and seismic response data". In: *Soil Dyn. Earthq. Eng.*

Loh, C.-H., C.-H. Chen, and T.-Y. Hsu (2011). "Application of advanced statistical methods for extracting long-term trends in static monitoring data from an arch dam". In: *Struct. Heal. Monit.*

Lorenz, D. A., M. E. Pfetsch, and A. M. Tillmann (2015). "Solving Basis Pursuit: Heuristic Optimality Check and Solver Comparison". In: *ACM Trans. Math. Softw.*

Lotfi, V., J. M. Roesset, and Tassoulas J. L. "A technique for the analysis of the response of dams to earthquakes". In: *Earthq. Eng. Struct. Dyn.* 15.4 (), pp. 463–489.

Lubliner, J., J. Oliver, S. Oller, and E. Oñate (1989). "A plastic-damage model for concrete". In: *Int. J. Solids Struct.* 25.3, pp. 299–326.

Lucchesi, M., C. Padovani, G. Pasquinelli, and N. Zani (2008). *Masonry constructions: mechanical models and numerical applications.* Springer - Verlag Berlin Heidelberg.

Lucchesi, M., B. Pintucchi, and N. Zani (2017a). "A 3D masonry-like model with bounded shear stress". In: *Key Eng. Mater.*

— (2017b). *A 3D masonry-like model with bounded shear stress.* Vol. 747 KEM.

— (2017c). "Influence of the modelling approach in the dynamic analysis of masonry towers". In: *XVII CONVEGNO ANIDIS "L'Ingegneria Sismica Ital.* Pistoia.

Lupoi, A. and C. Callari (2012). "A probabilistic method for the seismic assessment of existing concrete gravity dams". In: *Struct. Infrastruct. Eng. Maintenance, Manag. Life-Cycle Des. Perform.* 8, pp. 37–41.

Luzi, L., R. Puglia, E. Russo, and ORFEUS WG5 (2016). *Engineering Strong Motion Database. Istituto Nazionale di Geofisica e Vulcanologia, Observatories & Research Facilities for European Seismology.*

Lysmer, J. and R. L. Kuhlemeyer (1969). "Finite Dynamic Model For Infinite Media". In: *J. Eng. Mech. Div.*

Marwala, T. (2010). *Finite Element Model Updating Using Computational Intelligence Techniques: Applications to Structural Dynamics.* 1st. Springer Publishing Company, Incorporated.

Mata, J. (2011). "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models". In: *Eng. Struct.*

Mata, J., A. Tavares de Castro, and J. Sá da Costa (2013). "Time-frequency analysis for concrete dam safety control: Correlation between the daily variation of structural response and air temperature". In: *Eng. Struct.*

— (2014). "Constructing statistical models for arch dam deformation". In: *Struct. Control Heal. Monit.* 21, pp. 423–437.

Mathworks (2016). *MATLAB - Mathworks - MATLAB & Simulink.*

Matthies, H. G., E. Zander, B. V. Rosić, and A. Litvinenko (2016). "Parameter estimation via conditional expectation: a Bayesian inversion". In: *Adv. Model. Simul. Eng. Sci.*

Mendes, P. and S. Oliveira (2009). "Influence of the Intake Tower Dynamic Behaviour on Modal Identification of Cabril Dam". In: *Proc. 3rd Int. Oper. Modal Anal. Conf.* Portonovo.

Mendes, P., S. Oliveira, S. Guerreiro, M. A. Baptista, and A. Campos Costa (2004). "Dynamic behaviour of concrete dams monitoring and modelling." In: *Proc. 13th World Conf. Earthq. Eng.* Vancouver.

Mendes, P., C. Oliveira Costa, J. Almeida Garrett, and S. Oliveira (2007). "Development of a monitoring system to Cabril Dam with operational modal analysis." In: *Proc. 2nd Exp. Vib. Anal. Civ. Eng. Struct.* Porto.

Milton De Araiijo, J. and A. M. Awruchb (1998). "Probabilistic finite element analysis of concrete gravity dams". In: *Adv. Eng. SoNare*.

Mirzahosseinkashani, S. and M. Ghaemian (2009). "Seismic fragility assessment of concrete gravity dams". In: *29th Annu. USSD Conf. Nashville, Tennessee, April 20-24, 2009*.

MISURE E ANALISI DELLE VIBRAZIONI DIGA DI SCANDARELLO. Tech. rep. Dipartimento di Protezione Civile: Ufficio Rischio Sismico E Vulcanico Servizio Monitoraggio Sismico Del Territorio Osservatorio Sismico delle Strutture, Ministero delle Infrastrutture e dei trasporti: Dipartimento per le Infrastrutture, i Sistemi Informa.

Mivehchi, M. R., M. T. Ahmadi, and A. Hajmomeni (2003). "Effective Techniques for Arch Dam Ambient Vibration Test: Application on Two Iranian Dams". In: *J. Seismol. Earthq. Eng.* 5.2.

Moyo, P. and C. Oosthuizen (2010). *Ambient vibration survey trials of two arch dams in South Africa*.

Muntasir Billah, A. H. M. and M. Shahria Alam (2015). "Seismic fragility assessment of highway bridges: a state-of-the-art review". In: *Struct. Infrastruct. Eng.*

Noorzaei, J., K. H. Bayagoob, W. A. Thanoon, and M. S. Jaafar (2006). "Thermal and stress analysis of Kinta RCC dam". In: *Eng. Struct.* 28.13, pp. 1795–1802.

Novak, M. "Dynamic soil-structure interaction, John P. Wolf, Prentice-Hall, Englewood Cliffs, N.J., 1985. No. of pages: £466. Price: £61.45". In: *Earthq. Eng. Struct. Dyn.* 14.6 (), p. 951.

Okuma, N., Y. Etou, K. Kanzawa, and K. Hirata (2008). *Evaluation of dynamic properties of an aged large arch dam.* Tech. rep. Fukuoka: Civil Engineering Group. Research laboratory report, Kyushu Eletric Power Co.

Oliveira, S., J. Rodrigues, P. Mendes, and A. C. Costa (2004). "Damage characterization in concrete dams using output only modal analysis". In: *Conf. 2004 IMAC-XXII Conf. Expo. Struct. Dyn.*

Omidi, O., S. Valliappan, and V. Lotfi (2013). "Seismic cracking of concrete gravity dams by plastic-damage model using different damping mechanisms". In: *Finite Elem. Anal. Des.* 63, pp. 80–97.

Owen, A. B. (2001). *Empirical Likelihood*. Chapman and Hall/CRC, p. 304.

Pan, J., Y. Feng, F. Jin, C. Zhang, and D. Roger Jones Owen (2014). "Comparison of different fracture modelling approaches to gravity dam failure". In: *Eng. Comput.* 31.1, pp. 18–32.

Perner, F. and P. Obernhuber (2010). "Analysis of arch dam deformations". In: *Front. Archit. Civ. Eng. China.*

Peter, S. J., A. Siviglia, J. Nagel, S. Marelli, R. M. Boes, D. Vetsch, and B. Sudret (2018). "Development of Probabilistic Dam Breach Model Using Bayesian Inference". In: *Water Resour. Res.* 54.7, pp. 4376–4400.

Popescu, T. D. (2011). "A new approach for dam monitoring and surveillance using blind source separation". In: *Int. J. Innov. Comput. Inf. Control* 7.7 A.

Porter, K. A. (2003). "An Overview of PEER's Performance-Based Earthquake Engineering Methodology". In: *9th Int. Conf. Appl. Stat. Probab. Civ. Eng.*

— (2017). "A Beginner's guide to fragility, vulnerability, and risk". In: *Univ. Color. Boulder.*

Prakash, G., A. Sadhu, S. Narasimhan, and J. M. Brehe (2018). "Initial service life data towards structural health monitoring of a concrete arch dam". In: *Struct. Control Heal. Monit.* 25.e2036.

Pytharouli, S. I. and S. C. Stiros (2005). "Ladon dam (Greece) deformation and reservoir level fluctuations: Evidence for a causative relationship from the spectral analysis of a geodetic monitoring record". In: *Eng. Struct.*

Rashid, Y. R. (1968). "Ultimate strength analysis of prestressed concrete pressure vessels". In: *Nucl. Eng. Des.* 7.4, pp. 334–344.

Rosić, B. and H. G. Matthies (2017). "Sparse bayesian polynomial chaos approximations of elasto-plastic material models". In: *XIV Int. Conf. Comput. Plast. Fundam. Appl.* Barcelona, pp. 256–267.

Roth, S. N., P. Léger, and A. Soulaïmani (2015). "A combined XFEM-damage mechanics approach for concrete crack propagation". In: *Comput. Methods Appl. Mech. Eng.* 283, pp. 923–955.

Saini, S. S. and J. Krishna (1973). "Overturning of top profile of the Koyna Dam during severe ground motion". In: *Earthq. Eng. Struct. Dyn.* 2.3, pp. 207–217.

Salazar, F., M. Á. Toledo, J. M. González, and E. Onate (2017). "Early detection of anomalies in dam performance: A methodology based on boosted regression trees". In: *Struct. Control Heal. Monit.* 24.e2012.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola (2008). *Global Sensitivity Analysis. The Primer.*

Severn, R. T., A. P. Jeary, and B. R. Ellis (1980). "Forced vibration tests and theoretical studies on dams". In: *Proc. Inst. Civ. Eng.* ICE, London.

Sevim, B., A. Bayraktar, and A. C. Altunisik (2010). "Finite element model calibration of berke arch dam using operational modal testing". In: *J. Vib. Control.*

Shao, C., C. Gu, M. Yang, Y. Xu, and H. Su (2018). "A novel model of dam displacement based on panel data". In: *Struct. Control Heal. Monit.* 25.e2037.

Sobol, I. M. (1993). "Sensitivity analysis for non-linear mathematical models". In: *Math. Model. Comput.*

Song, C. and J. P. Wolf (2000). "Scaled boundary finite-element method - a primer: Solution procedures". In: *Comput. Struct.*

Sooch, G. S. and A. Bagchi (2014). "A New Iterative Procedure for Deconvolution of Seismic Ground Motion in Dam-Reservoir-Foundation Systems". In: *J. Appl. Math.*

Spall, J. C. (1998a). "An Overview of the Simultaneous Perturbation Method for Efficient Optimization". In: *Johns Hopkins Apl Tech. Dig. Vol. 19, Issue 4.*

— (1998b). "Implementation of the simultaneous perturbation algorithm for stochastic optimization". In: *IEEE Trans. Aerosp. Electron. Syst.*

Su, H., Z. Chen, and Z. Wen (2016). "Performance improvement method of support vector machine-based model monitoring dam safety". In: *Struct. Control Heal. Monit.* 23, pp. 252–266.

Su, H., Z. Wen, X. Sun, and X. Yan (2018). "Multisource information fusion-based approach diagnosing structural behavior of dam engineering". In: *Struct. Control Heal. Monit.* 25, e2073.

Sudret, B. (2008). *Global sensitivity analysis using polynomial chaos expansions.*

Tan, H. and A. K. Chopra (1995). "Earthquake analysis of arch dams including dam-water-foundation rock interaction". In: *Earthq. Eng. Struct. Dyn.*

Tekie, P. B. and B. R. Ellingwood (2003). "Seismic fragility assessment of concrete gravity dams". In: *Earthq. Eng. Struct. Dyn.* 32.14, pp. 2221–2240.

Tierney, L. (1994). "Markov Chains for Exploring Posterior Distributions". In: *Ann. Stat.* 22.4, pp. 1701–1762.

Tillmann, A. M. and M. E. Pfetsch (2014). "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing". In: *IEEE Trans. Inf. Theory.*

Timoshenko, S. and J. N. Goodier (1986). *Theory of Elasticity.*

Tipping, M. (2001). "Sparse Bayesian Learning and the Relevance Vector Mach". In: *J. Mach. Learn. Res.*

Tweedie, R. L. (1975). "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space". In: *Stoch. Process. their Appl.* 3.4, pp. 385–403.

USACE - U.S Army Corps of Engineers (2015). "National Inventory of Dams". In: *CorpsMap*.

Uğurlu, A. (2007). "Elasticity Theory of Concrete and Prediction of Static E-Modulus for Dam Concrete Using Composite". In: *J. Eng. Stud. Res.* 18.1, pp. 4055–4067.

Vamvatsikos, D. and C. Allin Cornell (2002). "Incremental dynamic analysis". In: *Earthq. Eng. Struct. Dyn.* 31.3, pp. 491–514.

Vamvatsikos, D. and M. Fragiadakis (2010). "Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty". In: *Earthq. Eng. Struct. Dyn.* 39.2, pp. 141–163.

Vogel, C. R. (1987). *Inverse and Ill-Posed Problems*.

Wang, M.-L. and S. P. Shah. "Reinforced concrete hysteresis model based on the damage concept". In: *Earthq. Eng. Struct. Dyn.* 15.8 (), pp. 993–1003.

Wei, B., M. Gu, H. Li, W. Xiong, and Z. Xu (2018a). "Modeling method for predicting seepage of RCC dams considering time-varying and lag effect". In: *Struct. Control Heal. Monit.* 25, e2081.

Wei, B., D. Yuan, Z. Xu, and L. Li (2018b). "Modified hybrid forecast model considering chaotic residual errors for dam deformation". In: *Struct. Control Heal. Monit.* 25, e2188.

Westergaard, H. M. (1933). "Water pressures on dams during earthquakes". In: *Trans. ASCE* 58.6, pp. 605–42.

Wiener, N. (1938). "The Homogeneous Chaos". In: *Am. J. Math.* 60.4, pp. 897–936.

Worden, K. and G. R. Tomlinson (2000). *Nonlinearity in Structural Dynamics: Detection, Identification and Modelling*. CRC Press, p. 659.

Wu, Y., S. Xu, Q. Li, G. Ruiz, and R. C. Yu (2016). "Estimation of real fracture parameters of a dam concrete with large size aggregates through wedge splitting tests of drilled cylindrical specimens". In: *Eng. Fract. Mech.* 163, pp. 23–36.

Xiu, D. (2010). *Numerical Methods for Stochastic Computations*, p. 142.

Xiu, D. and G. E. Karniadakis (2002). "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations". In: *SIAM J. Sci. Comput.*

Yang, Z. (2006). "A modified family of power transformations". In: *Econ. Lett.* 92.1, pp. 14–19.

Yao, X. W. and J. Jiang (2011). "Analytical Seismic Fragility Analysis of Concrete Arch Dams". In: *15th World Conf. Earthq. Eng.*

Yazdchi, M., N. Khalili, and S. Valliappan (1998). "Non-linear seismic behaviour of concrete gravity dams using coupled finite element–boundary element technique". In: *Int. J. Numer. Methods Eng.* 44.1, pp. 101–130.

Yun, C. B., D. K. Kim, and J. M. Kim (2000). "Analytical frequency-dependent infinite elements for soil-structure interaction analysis in two-dimensional medium". In: *Eng. Struct.*

Zangar, C. N. (1953). "Hydrodynamic Pressures on Dams due to Horizontal Earthquakes," in: *Proc. Soc. Exp. Stress Anal.* 10, pp. 93–102.

Zhang, L., M. Peng, D. Chang, and Y. Xu (2016). *Dam Failure Mechanisms and Risk Assessment*. John Wiley & Sons Singapore Pte. Ltd., p. 476.

Zhang, S., G. Wang, and X. Yu (2013). "Seismic cracking analysis of concrete gravity dams with initial cracks using the extended finite element method". In: *Eng. Struct.* 56, pp. 528–543.

Zhu, X. and O. A. Pekau (2007). "Seismic behavior of concrete gravity dams with penetrated cracks and equivalent impact damping". In: *Eng. Struct.* 29.3, pp. 336–345.

Zienkiewicz, O. C., C. Emson, and P. Bettess (1983). "A novel boundary infinite element". In: *Int. J. Numer. Methods Eng.*