

Concrete gravity dams model parameters updating using static measurements.

Giacomo Sevieri^{a,*}, Marco Andreini^b, Anna De Falco^c, Hermann G. Matthies^d

^a*Department of Civil and Industrial Engineering, University of Pisa, Largo Lucio Lazzarino, 1 - 56126 Pisa, Italy. giacomo.sevieri@unifi.it*

^b*Department of Energy, Systems, Territory and Constructions Engineering, University of Pisa, Largo Lucio Lazzarino, 1 - 56126 Pisa, Italy. m.andreini@ing.unipi.it*

^c*Department of Energy, Systems, Territory and Constructions Engineering, University of Pisa, Largo Lucio Lazzarino, 1 - 56126 Pisa, Italy. a.defalco@ing.unipi.it*

^d*Institute for Scientific Computing, Technische Universität Braunschweig, Mühlenpfordtstrasse, 23 - 38106 Braunschweig, Germany. wire@tu-braunschweig.de*

Abstract

The structural control of concrete gravity dams is of primary importance. In this context, numerical models play a fundamental role both to assess the vulnerability of gravity dams and to control their behaviour during normal operativity and after extreme events. In this regard, data monitoring represents an important source of information for numerical model calibrations.

This study proposes a novel probabilistic procedure, defined in the Bayesian framework, to calibrate the parameters of finite elements models of dams. To this aim, monitoring data and the results of material tests are used as reference information. The computational burden is reduced by using a new hybrid-predictive model of the dam displacements. An application on an Italian dam shows the feasibility of the proposed procedure.

Keywords: Bayesian updating, probabilistic models, general polynomial chaos expansion, Fourier analysis, gravity dams, model parameters identification.

*

*Corresponding author

Email address: giacomo.sevieri@unifi.it (Giacomo Sevieri)

1. Introduction

Most of the concrete gravity dams in the world were built before the introduction of seismic regulations or were designed by means of out-of-date analysis methods which often underestimated seismic actions. In addition, the deterioration levels of these dams may have reached a critical value.

Due to economic and environmental reasons [1], only a few new dams are now being built, though dams represent an important resource in energy production. Therefore, the evaluation of the seismic vulnerability of the existing dams is needed, to extend their life expectancy. Those concerns have spurred the scientific community to pursue a better understanding of the seismic risk of existing dams. Furthermore, since no concrete gravity dam failures have occurred after an earthquake [2, 3], numerical modelling is the only means of predicting the seismic performance of this type of dams [4]. In this regard, several studies have addressed the problem in recent years by the methods used in performance-based earthquake engineering [5, 6, 7].

Uncertainties in the dam modelling process [8, 9, 10, 11] are a consequence of the lack of knowledge on the physical data such as material properties, dam geometry, rock mass profile, and failure modes. In order to achieve model reliability, all available information must be used to reduce the uncertainty of the analysis results [12].

Dams are usually monitored during normal operations in order to timely highlight any undesired structural behaviour. Displacements at different levels due to both water level variation and thermal variations are monitored in conjunction with the water level and the water and air temperature values [13], although dynamic measurements are very rare. Monitoring data also offers the opportunity to calibrate numerical models. Nevertheless, the common methodologies which employ quasi-static displacements under environmental actions are strongly affected by the subjectivity of the analyst and do not evaluate the model error [13] as further explained.

In this study, an original probabilistic procedure to build hybrid predictive mod-

els of concrete gravity dam displacements is proposed. Indeed, the displacements due to seasonal storage level are simulated by FE models while those due to thermal variations are simulated through Fourier series [14]. The proposed procedure involves a set of explanatory functions that allow further correction of the model bias, taking into account phenomena that cannot directly enter into the model [15]. The FE model unknown parameters, as well as the coefficients of the Fourier series, are treated as random variables. A Bayesian approach is adopted for the updating of the unknown parameters and for the model calibration based on pendulum observations. The computational burden for the model calibration is reduced using the spectral approach of the general Polynomial Chaos Expansion (gPCE) [16].

The adopted Bayesian approach [17] can cope with “ill-conditioning” or non-uniqueness problems, even assuring model flexibility with inference and confidence intervals not depending on sample estimations.

Therefore, the updated model parameters can be used either as starting point of a seismic vulnerability analysis or to control in real time the structural behaviour after seismic events. This latter is performed by comparing the recorded displacements with the ones calculated through the proposed probabilistic model. This research work aims to solve two practical open issues in dam engineering field: a) the calibration of mechanical parameters of static FE dam models, by using observations recorded by the monitoring system, and b) the definition of reliable dam predictive models, which are characterised by high computational speed, for control purposes. Both of them require FE analyses, thermal and static ones, which are computationally and time intensive. The common methodologies used in dam engineering field tackle these issues by splitting the components of the observations, and by elaborating each one separately, as explained below. The novelty of this research work lies in the definition of a probabilistic procedure, based on the common statistical tools, which allows updating the parameters of dam displacement predictive models through the Bayesian inference. The use of meta models allows strongly reducing the computational burden, so speeding up the procedure and making possible its use

in an on-line control system. The proposed procedure does not require the pre-process step of the observations, thus reducing the influence of the analyst and performing an automatic scheme. Thanks to the meta models and the Bayesian
65 approach, the final computational speed allows solving the previous open issues in dam engineering field.

This paper firstly presents a brief review of the most common methods to interpret the structural behaviour and to calibrate FE and analytical models of dams by means of displacement monitoring data. The proposed probabilistic
70 model is then described and the calibration of the unknown parameters is finally presented for an about 35-m-high concrete gravity dam built in the '50 in Italy.

2. Interpretation of the dam structural behaviour from monitoring data

The surveillance of a large concrete gravity dam during normal operations
75 is usually based on the monitoring of displacements caused by the storage level and seasonal thermal variations [18]. 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 rain-
80 fall, water and air temperatures and storage level are recored daily, while uplift pressures and water losses through seepages are generally recorded weekly.

Safety monitoring of a dam has the dual purpose of highlighting anomalous operations and understanding structural behaviour. In this context the main assumption is the linearity of the model and the materials. It allows us to apply
85 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 behaviors, creep and other phenomena. δ^H , which mainly depends on the mechanical characteristics of the materials, both concrete
90 and foundation rock, is usually represented in the literature by a polynomial se-

ries of material parameters [13]. 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 .

95 On the one hand, the observational approach assumes that the system response is stationary. It makes use of statistical procedures which correlate input variables such as temperature and water level values, with output ones, such as crest displacement during normal operations. It aims to find the contribution of external loads to structure deformation and to identify irreversible compo-
 100 nents in the structural response as well. 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
 105 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 re-
 110 sponse. 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
 115 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.
 120 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's parameters while avoiding burdensome thermal analyses, it is therefore necessary to separately analyse δ^H by subtracting δ^T from δ ,
125 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
130 degree of uncertainty and subjectivity mainly due to the selection of the analytical functions describing the output variables.

In the context of SHM and model parameters updating of concrete gravity dams, the definition of predictive models is needed. In order to correctly simulate the structural behaviour, these predictive models must be characterised by
135 high computational speed and relatively small error as well. For these reasons functional approaches are broadly used to define the predictive models of dam displacements as indicated in the following literature review. Moreover, despite more complex models which account for the interaction between mechanical and thermal phenomena could lead to a better approximation, they are not com-
140 monly used, since local phenomena are of little importance. It is finally worth noting that, with regard to the accuracy level required for the structural control of concrete gravity dams, predictive models based on the functional approximation are successfully applied.

As reported by Bukenya et al. [19], several different research works available in
145 the literature have proposed predictive models of dam displacements for structural control purposes. The largest part of them was based on the previously discussed functional approach whose coefficients were calibrated by regression onto the observations acquired by the monitoring system.

Kao and Loh [20] used Artificial Neural Networks (ANN) to build a predic-
150 tive model of the dam displacements, whose coefficients were trained with the recorded measurements. The authors studied the Fei-Tsui dam by comparing the use of three different kinds of ANN with an increasing level of non-linearity.

The authors found that the three versions of ANN were able to accurately approximate the observed behaviour.

155 Mata et al. [21] proposed multiple linear regression models to approximate the displacements of the Alto Lindoso dam, whose coefficients were calibrated by regression on the recorded observations. The authors also used different approaches to approximate the thermal part of the displacement. In particular, a first approach based on sinusoidal functions and a second approach based on the
160 real recorded temperatures were used. In the end, the comparison highlighted that the latter strategy allowed one to slightly improve the estimation of the dam behaviour but with a higher computational cost than the former one.

Kang et al. [22] proposed an Extreme Learning Machine-based predictive model for the displacements of the Fengman dam for structural health monitoring purposes. The authors compared the results of the proposed model with the clas-
165 sical ones, as those described at the beginning of this Section, showing that all of them led to a good agreement between predictions and observations.

Prakash et al. [23] used the idea of principal component analysis to build predictive models of the dam static behaviour, e.g. displacements and strains, whose
170 coefficients were calibrated by using the least squares method. The study of a concrete dam located in Bulgaria showed the applicability of the proposed procedure to real cases.

Dai et al. [24] studied a RCC gravity dam by applying the statistical model of the dam displacement and Random Forest Regression (RFR) to determine its
175 combination coefficients, thus reducing the bias of the prediction.

Lin et al. [25] proposed a method to split dam displacements into two parts: the one related to the dam behaviour and the second one related to the foundation. The idea of partitioned FEM was used to define hybrid equations which allowed one to separate these two contributions. The observations recorded by
180 the monitoring system were used to obtain the mechanical parameters of the materials. The study of a concrete dam was used to assess the validity of the separation method.

From the literature review two basic concepts can be found in each application:

the predictive models of the dam displacements are defined through functional approximations and they are calibrated in a deterministic setting. These two
185 widespread assumptions are mainly related to computational aspects. Indeed, the final aim of a calibrated dam predictive model is to be used within a real-time SHM, so it must be characterised by a high computational speed. This important feature is also needed in the calibration process, which requires a
190 large number of solutions in order to achieve a good quality of the prediction. In this context, FE models become prohibitive, as well as probabilistic procedures. One of the biggest drawbacks in the use of functional approximations is the impossibility to directly link the prediction with the mechanical parameters of the materials.

Thanks to the gPCE based meta model to reproduce the outputs of the FEA, the
195 predictive model proposed in this work allows one to calibrate the FE model parameters, thus determining the properties of the materials. Moreover, the high computational speed of the gPCE allows the use of the proposed predictive model both in real-time SHM and within probabilistic procedure for model
200 updating.

The use of Bayesian inference in SHM [26] and, from a broader point of view, in civil engineering fields has recently gained interest as shown by the high number of papers [27, 28, 29, 30].

Focusing the attention in dam engineering field, the most interesting examples
205 aim to determine the model parameters for the risk assessment of earth fill dams [31, 32, 33, 34].

3. Development of the proposed probabilistic model

In this section, the general formulation for the probabilistic model adopted in the following for concrete gravity dam displacements is presented. The model
210 takes into account both the epistemic and aleatory uncertainties. Moreover, the correction of the model bias is performed by including a set of explanatory functions, providing insight on the underlying behavioural phenomena. The

Bayesian updating technique used to perform the calibration of the unknown parameters, e.g. the discussed material parameter identification, and the Markov Chain Monte Carlo (MCMC) adopted for the inverse problem resolution are finally described.

3.1. General formulation of the probabilistic model

Referring to [15, 35], an additive corrected form is adopted for the probabilistic model, which is characterised by the presence of a correction term, described below, to further reduce the model bias. In this section the general formulation of the probabilistic model is introduced in its univariate version. By defining: \mathbf{x} as the vector of measurable model variables, $\boldsymbol{\theta}_m$ as the vector of the unknown model parameters and $\boldsymbol{\theta}_\gamma = (\theta_0, \dots, \theta_j, \dots, \theta_p)^T$ as the vector of combination coefficients of the explanatory functions which constitute the correction term, the probabilistic model can be written as:

$$C(\mathbf{x}, \boldsymbol{\theta}_m, \boldsymbol{\theta}_\gamma, \sigma) = \hat{c}(\mathbf{x}, \boldsymbol{\theta}_m) + \gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma) + \sigma\epsilon, \quad (1)$$

where

$$\gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma) = \theta_0 + \sum_{j=1}^p \theta_j h_j(\mathbf{x}). \quad (2)$$

In this context $C(\mathbf{x}, \boldsymbol{\theta}_m, \boldsymbol{\theta}_\gamma, \sigma)$ is the quantity of interest or a transformation of that, $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_m)$ is the deterministic model output or a transformation of that and $\gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma)$ is the correction term reducing the bias through the combination of the p explanatory functions h_j . Finally, ϵ is a normal random variable with zero mean and unit variance, while σ represents the standard deviation of the model error.

It is well known that the additive model is based on three assumptions:

- Normality: the global error ϵ is normally distributed.
- Homoscedasticity: the model error standard deviation σ is independent from \mathbf{x} .
- Additivity: the model error can be added to the probabilistic model.

Since these three assumptions are generally not satisfied in real problems, a variance stabilizing transformation is used to approximately satisfy these assumptions within the selected range of data [36].

In the context of the present work, the reference measurement is a displacement, which could be positive or negative. In order to use a natural logarithm form, a translation of the reference system which makes it possible to obtain only positive values is needed:

$$C_i = \ln(y_i + l) \quad (3)$$

where C_i is the transformation of the i -th measurement, y_i is the i -th measurement and l is the translation.

230

3.2. Bayesian Updating via Markov Chain Monte Carlo

The unknown model parameters are estimated through a Bayesian approach following the updating rule [37]. By defining the vector of the parameters which will be updated, $\Theta = [\theta_m, \theta_\gamma, \sigma]^T$, the updating rule becomes:

$$p(\Theta|\mathbf{y}) = \kappa L(\Theta|\mathbf{y}) p(\Theta). \quad (4)$$

Where $p(\Theta|\mathbf{y})$ is the posterior distribution which represents the updated state of knowledge on the random variables Θ , $L(\Theta|\mathbf{y})$ is the likelihood function which transforms the prior distribution to a posterior distribution by updating the model parameters once the new n data, collected in \mathbf{y} , are gathered. Finally, $p(\Theta)$ is the prior distribution which represents the state of the knowledge before the introduction of new data \mathbf{y} , and $\kappa = [\int L(\Theta|\mathbf{y}) p(\Theta) d\Theta]^{-1}$ is the normalising factor. Since the model error is normally distributed, the likelihood function can be written as [37, 15]:

$$L(\mathbf{x}, \Theta) \propto \prod_{i=1}^n \left\{ \frac{1}{\sigma} \varphi \left[\frac{r_i(\mathbf{x}, \theta_m, \theta_\gamma)}{\sigma} \right] \right\}. \quad (5)$$

Where $r_i(\mathbf{x}, \theta_m, \theta_\gamma)$, is the i -th residual which represents the discrepancy between the measurement and its prediction

$$r_i(\mathbf{x}, \theta_m, \theta_\gamma) = C_i - \hat{c}_i(\mathbf{x}, \theta_m) - \gamma(\mathbf{x}, \theta_\gamma) \quad (6)$$

and $\varphi(\cdot)$ is the probability density function of a standard normal distribution. The determination of the likelihood and the posterior distribution of random variables is a challenging task in real cases. Generally, this problem can not be solved in a closed form, so numerical techniques are needed. The Markov Chain Monte Carlo (MCMC) method [38] is one of the most common techniques when stochastic FE are involved. In this paper, the Metropolis-Hastings algorithm is used. In the Metropolis-Hastings algorithm, a sequence of samples Θ^s is generated for each step s . The value of Θ^s depends only on the previous sample in the chain, while the value of the new state Θ^{s+1} is generated according to the proposal distribution $g_k(\Theta^{s+1}|\Theta^s)$. For several practical reasons, a symmetric distribution is often chosen as proposal [39]

$$q(\Theta^{s+1}|\Theta^s) = q(\Theta^s|\Theta^{s+1}). \quad (7)$$

Once the new sample is generated, it can be accepted with probability r or rejected with probability $1-r$. As explained in [40], when the regularity conditions of aperiodicity, irreducibility, and positive recurrence of the Markov Chain [41] are satisfied, the distribution of sampling sequence Θ^s converges to the target posterior distribution, regardless of the starting point. Nevertheless, checking the compliance with the previous conditions can be very difficult in real problems, so convergence diagnostics are applied.

To speed up the procedure, the FEM response can be simulated by a proxy model. So, the error term of Equation 1 accounts also for the difference between the FEM response and the proxy model [42].

3.3. Convergence Diagnostics

Convergence diagnostics allow us to check whether the samples generated by MCMC are representative of the underlying equilibrium distribution. In this paper, the widely used diagnostics metric proposed by Brooks and Gelman [43] is employed. The convergence of the Markov Chain simulation has been reached when inferences for quantities of interest do not depend on the starting point. Therefore, monitoring convergence is performed 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 it represents the upper bound of the maximum of the univariate Potential Scale Reduction Factor (PSRF) statistics \hat{R} among Θ 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 [44], but when the dimension of the problem increases, a convergence criterion $\hat{R}^p = 1.5$ is allowed.

$$\hat{R}^p = \max_{\mathbf{a}} \frac{\mathbf{a}^T \hat{\mathbf{V}} \mathbf{a}}{\mathbf{a}^T \mathbf{W} \mathbf{a}}, \quad (8)$$

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 the vector used to
245 achieve the maximum value of the ratio in equation 8.

4. Formulation of probabilistic model for existing concrete gravity dams

In this section, a probabilistic model for mechanical parameters identification in concrete gravity dam analysis is described. Due to the large amount of mon-
250 itoring data available, model identification procedures based on the Bayesian statistic are successfully applied. Generally, environmental data are recorded together with the displacements of the structure due to seasonal temperature variations and changes in the level of the reservoir. Since displacements vary slowly, the problem can be approached as static. The presence of both thermal
255 and mechanical actions requires performing coupled analyses, which are computationally demanding, in particular when complex FE models are involved. For this reason, a spectral approach based on gPCE is used to approximate the FE model response for the water level variation δ_i^H . Following the same idea, thermal displacements are approximated through a Fourier series, in order to avoid
260 expensive thermal analyses [45]. In the next sections, the proposed procedure is

described. This way, the set of unknown parameters $\boldsymbol{\theta}_m$ can be thought as composed by the gPCE random variables, collected in $\boldsymbol{\theta}_{\text{gPCE}}$, and the combination coefficients of the Fourier analysis $\boldsymbol{\theta}_{\text{FA}}$, that is $\boldsymbol{\theta}_m = [\boldsymbol{\theta}_{\text{gPCE}}, \boldsymbol{\theta}_{\text{FA}}]^T$.

4.1. Definition of the proxy model for hydrostatic displacements through general Polynomial Chaos Expansion (gPCE)

265

Orthogonal polynomials are often used to approximate a continuous function [46]. The general Polynomial Chaos Expansion (gPCE) is a spectral method based on the representation of the random variable $\boldsymbol{\theta}_{\text{gPCE}}$ through a series of orthogonal polynomials. It allows us to define the response surface $f_K(\boldsymbol{\theta}_{\text{gPCE}})$ of the model output $f(\boldsymbol{\theta}_{\text{gPCE}})$. This way, statistics can be calculated from integrals over the probability space which is defined by the distributions of the variables collected in $\boldsymbol{\theta}_{\text{gPCE}}$.

In the D -variate case, $\boldsymbol{\theta}_{\text{gPCE}} = (\theta_{\text{gPCE},1}, \dots, \theta_{\text{gPCE},d}, \dots, \theta_{\text{gPCE},D})^T$ is composed of mutual independent components. Moreover, by defining $\{\phi_k(\theta_{\text{gPCE},d})\}_{k=0}^K \in \mathbb{P}_K(\theta_{\text{gPCE},d})$ as the univariate gPC basis functions of $\theta_{\text{gPCE},d}$, in the space of polynomials whose degree is up to K , with $1 < d < D$, and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_D) \in \mathbb{N}_0^D$ as multi-index, with $|\boldsymbol{\alpha}| = \alpha_1 + \dots + \alpha_D$, the D -variate K^{th} degree gPC basis functions $\Phi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\text{gPCE}})$ are obtained as a product among the univariate gPC polynomials $\Phi_{\alpha_d}(\theta_{\text{gPCE},d})$. Therefore, the orthogonality condition can be written as:

$$\mathbb{E}[\Phi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\text{gPCE}}) \Phi_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\text{gPCE}})] = \int \Phi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\text{gPCE}}) \Phi_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\text{gPCE}}) dF_{\boldsymbol{\theta}_{\text{gPCE}}} = \eta_{\boldsymbol{\alpha}} \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}, \quad (9)$$

where $\eta_{\boldsymbol{\alpha}} = \mathbb{E}[\Phi_{\boldsymbol{\alpha}}^2]$ are the normalization factors and $\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ is the d-variate Kronecker delta function.

Once the proxy model is defined, all the statistics can be retrieved from gPCE in a straightforward manner. In the general case of K^{th} degree gPC approximation, $f(\boldsymbol{\theta}_{\text{gPCE}})$ can be written as:

$$f_K(\boldsymbol{\theta}_{\text{gPCE}}) = \sum_{|\boldsymbol{\alpha}| \leq K} a_{\boldsymbol{\alpha}} \Phi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\text{gPCE}}) \in \mathbb{P}_K^D \quad (10)$$

where a_{α} is the matrix of the combination coefficients of the basis of functions and may be calculated using different techniques, e.g. the collocation method, regression, the orthogonal projection and the Galerkin projection [16]. In this paper, a regression method is used, then the response of the model is calculated in more points, of the random variables space, than the number of the basis functions.

The mean value of f can be approximated as:

$$\mu_f \triangleq \mathbb{E}[f(\boldsymbol{\theta}_{\text{gPCE}})] \approx \mathbb{E}[f_K(\boldsymbol{\theta}_{\text{gPCE}})] = \int \left(\sum_{|\alpha| \leq K} a_{\alpha} \Phi_{\alpha}(\boldsymbol{\theta}_{\text{gPCE}}) \right) dF_{\boldsymbol{\theta}_{\text{gPCE}}} = a_0 \quad (11)$$

Other statistics can also be readily approximated by applying their definitions directly to the gPC approximation f_K . The family of polynomials may be selected on the basis of the probability distribution functions of the random variables [16], while the maximum degree of polynomials must be chosen in such a way as to reduce the approximation of the gPCE.

In this work the gPCE has been used to approximate the displacements of the dam due to the reservoir 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 described by the elastic constitutive matrix \mathbb{C} [47], whose components, in the most general cases of anisotropic material, may be treated as random variables in a Bayesian updating procedure. Modelling materials as anisotropic is convenient only when it is strictly necessary and there are measurements which are able to infer all the components of the constitutive matrix \mathbb{C} . 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 the measurements in the upstream-downstream direction and in the cross-valley direction were available.

Conversely, when only upstream-downstream measurements are available, an isotropic material is the best choice. For an elastic isotropic material, matrix \mathbb{C}

is as follows

$$\mathbb{C} = \begin{bmatrix} K + 4G/3 & K - 2G/3 & K - 2G/3 & 0 & 0 & 0 \\ K - 2G/3 & K + 4G/3 & K - 2G/3 & 0 & 0 & 0 \\ K - 2G/3 & K - 2G/3 & K + 4G/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix}. \quad (12)$$

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 physically and statistically independent. Therefore, in this context the bulk and the shear modulus of the materials are respectively collected in $\boldsymbol{\theta}_{\text{gPCE},K}$ and $\boldsymbol{\theta}_{\text{gPCE},G}$, then $\boldsymbol{\theta}_{\text{gPCE}} = [\boldsymbol{\theta}_{\text{gPCE},K}, \boldsymbol{\theta}_{\text{gPCE},G}]^T$.

4.2. Definition of the proxy model for thermal displacements through Fourier analysis (FA)

The aim of this work is to identify the model mechanical parameters in order to perform a static twin model of the structure for control purpose. The response of the structure due to the reservoir level variation is related to the values of materials mechanical parameters, while the response of the structure due to thermal effects is related to thermal materials parameters and it is relatively unaffected by mechanical parameters variation. Thermal displacement is the largest part of the total displacement, so it cannot be neglected. Thermal displacements are approximated through a Fourier analysis (FA) because their variation is similar to that of measured temperatures having a periodic behaviour over one year [48]. Using the Fourier series, and by defining the vector of the combination coefficients $\boldsymbol{\theta}_{\text{FA}} = [\boldsymbol{\theta}_{\text{FA},a}, \boldsymbol{\theta}_{\text{FA},b}]^T$, the target function $s(t, \boldsymbol{\theta}_{\text{FA}})$ is approximated by $s_M(t, \boldsymbol{\theta}_{\text{FA}})$, which is the sum of simple waves [14]. Usually, it is expressed as

$$s_M(t, \boldsymbol{\theta}_{\text{FA}}) = \sum_{m=1}^M [\theta_{\text{FA},a_m} \cos(2\pi m t \omega) + \theta_{\text{FA},b_m} \sin(2\pi m t \omega)], \quad (13)$$

where the combination coefficients θ_{FA,a_m} and θ_{FA,b_m} are respectively collected in the vectors $\boldsymbol{\theta}_{\text{FA},a}$ and $\boldsymbol{\theta}_{\text{FA},b}$. Moreover, t is the time and ω is the frequency of the wave. In this paper, 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.

4.3. Definition of the probabilistic model for total displacements

The way to handle the various components of the displacement is described in the previous paragraphs. In this section, the probabilistic model response $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_m)$ (equation 1), written in the case of concrete gravity dams, is defined. In this context, the i -th simulated displacement of a point on the dam can be written as

$$\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_m) = \hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\text{gPCE}}) + \hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\text{FA}}), \quad (14)$$

where $\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_m)$ is the i -th simulated total displacement, $\hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\text{gPCE}})$ is the part related to the basin level variation approximated through the gPCE, and $\hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\text{FA}})$ is the thermal term approximated by the Fourier analysis. As reported in paragraph 3.1, 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 $\hat{c}(\mathbf{x}, \boldsymbol{\theta}_m)$, and a translation of the reference system is needed to obtain positive values of displacements,

$$\hat{c}(\mathbf{x}, \boldsymbol{\theta}_m) = \ln \left(\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_m) + l \right). \quad (15)$$

Finally, the proposed probabilistic model for the static displacements of a concrete gravity dams can be determined by combining equations 1, 14 and 15, obtaining:

$$\begin{aligned} \ln(\delta_i(\mathbf{x}, \boldsymbol{\Theta}) + l) &= \ln \left(\hat{\delta}_i(\mathbf{x}, \boldsymbol{\theta}_m) + l \right) + \gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma) + \sigma\epsilon = \\ &= \ln \left(\hat{\delta}_i^{gPCE}(\mathbf{x}, \boldsymbol{\theta}_{\text{gPCE}}) + \hat{\delta}_i^{FA}(\boldsymbol{\theta}_{\text{FA}}) + l \right) + \gamma(\mathbf{x}, \boldsymbol{\theta}_\gamma) + \sigma\epsilon. \end{aligned} \quad (16)$$

The predictive model presented in this section 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. oreover, 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 a_{α} , the combination coefficients of the Fourier series $\theta_{FA,a}$ and $\theta_{FA,b}$, and the error term $\sigma\epsilon$ would be functions of the point coordinates.

Another important assumption of the predictive model defined in this section 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 during time (section 4.2).

4.4. Model correction

Other kinds of uncertainties are involved in the static analysis of concrete gravity dams. 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 is related to thermal variations, time-dependent phenomena of mechanical parameters and the accuracy of measuring instruments.

Explanatory functions $\gamma(\mathbf{x}, \theta_{\gamma})$ (equation 1) are introduced to reduce the bias, by taking into account phenomena which cannot be directly accounted for in the model.

To capture the potential bias which is independent of the parameters \mathbf{x} and θ_m , the function $h_1 = 1$ is considered. In this case, θ_1 represents the previously discussed shift between the reference systems. When using more than one ex-

planatory functions [35, 31], a stepwise deletion process can be performed in order to identify those functions which really contribute reducing the bias.

4.5. Prior distribution definition

Prior distributions represent the state of knowledge before the acquisition of new information. When using a probabilistic model based on a Bayesian framework, the choice of the prior distributions is a fundamental step because it influences the rate of convergence and the accuracy of the results. This fact is particularly evident when the amount of new information \mathbf{y} is small compared to the state of prior knowledge [37]. Usually, the results of in-situ and laboratory tests on the dam's concrete and on the ground rock are available. The prior distributions of the model parameters $\boldsymbol{\theta}_{\text{gPCE}}$, $\boldsymbol{\theta}_{\text{gPCE},K}$ and $\boldsymbol{\theta}_{\text{gPCE},G}$ can be therefore deduced from these data. The other random variables, collected in the vector $\boldsymbol{\theta}_{\text{FA}}$ and $\boldsymbol{\theta}_{\gamma}$, have no physical meaning in this case, 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 are selected. Referring to Box 1992 [37], it is possible to prove that 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. In the univariate case, the parameters collected in $\boldsymbol{\theta}_{\text{FA}}$, $\boldsymbol{\theta}_{\gamma}$ and σ are considered approximately independent [44], then:

$$p(\boldsymbol{\theta}_{\text{FA}}, \boldsymbol{\theta}_{\gamma}, \sigma) \cong p(\boldsymbol{\theta}_{\text{FA}}) p(\boldsymbol{\theta}_{\gamma}) p(\sigma). \quad (17)$$

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

$$p(\sigma) \propto \sigma^{-1}. \quad (18)$$

Therefore, the other prior distributions become:

$$p(\boldsymbol{\theta}_{\text{FA}}, \boldsymbol{\theta}_{\gamma}) \propto \frac{p(\boldsymbol{\theta}_{\text{FA}}) p(\boldsymbol{\theta}_{\gamma})}{\sigma}. \quad (19)$$

315 Thus, given the large amount of data observed in this case, any reasonable
 choice of prior estimate has little influence on the posterior estimates of the
 parameters.

4.6. Calibration Procedure

320 The presented procedure uses the Metropolis-Hastings algorithm which re-
 quires, as previously described, the definition of the proposal distributions and
 the starting point. In this paper, normal distributions with 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
 325 the likelihood function [49].

The convergence of the probabilistic analysis has been checked as explained in
 section 3.3, by performing two chains for each problem starting from different
 initial values. Every chain was stopped when the acceptance ratio reached an
 appropriate value, as indicated in literature [44]. The flow chart of the procedure
 is shown in figure 1.

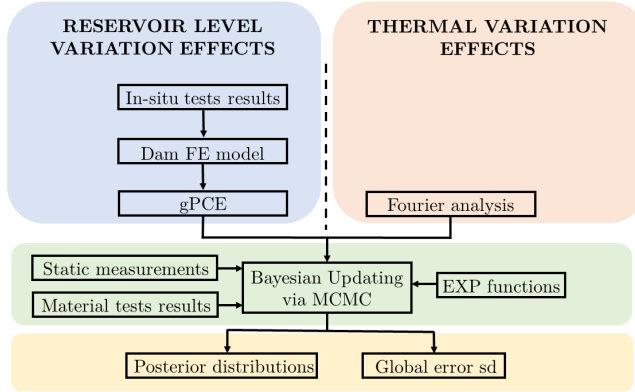


Figure 1: Procedure flow chart.

5. Case Study

5.1. Dam description and recorded measurements

In this paragraph, an application of the probabilistic model for a large dam in north-centre Italy is presented. Daily environmental measurements such as temperatures and reservoir levels are available as well as recorded displacements of the dam at different heights and the results of in-situ tests. With a crest 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. Finally, a sensitivity analysis on the influence of the Fourier series degree on the results is performed.

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 2, in the period from 2004 to 2011. This latter pendulum is fixed at the rock foundation at 534 m a.s.l. and measures displacements in both up-stream/down-stream and cross-valley direction 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 3, have been considered in the present work. Vertical joints opening-closing displacements are monitored by removable mechanical strain gauges located inside the two horizontal tunnels and shown in figure 2.

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.

Moreover, basin level, water and air temperatures are recorded in the same period (figures 4 and 5). In this work the observations recorded from 2004 to 2008 are used to update the model parameters, while the records from 2008 to 2011

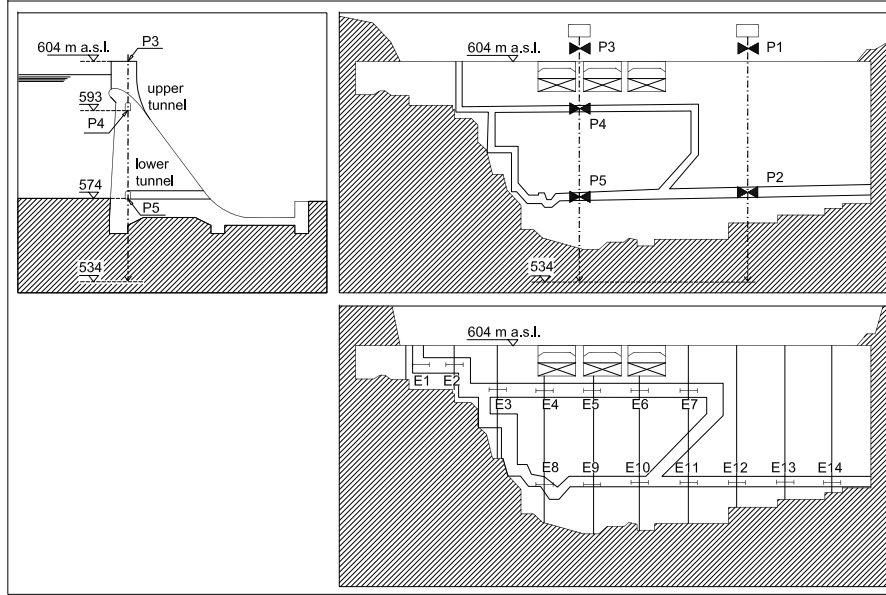


Figure 2: Static monitoring system layout.

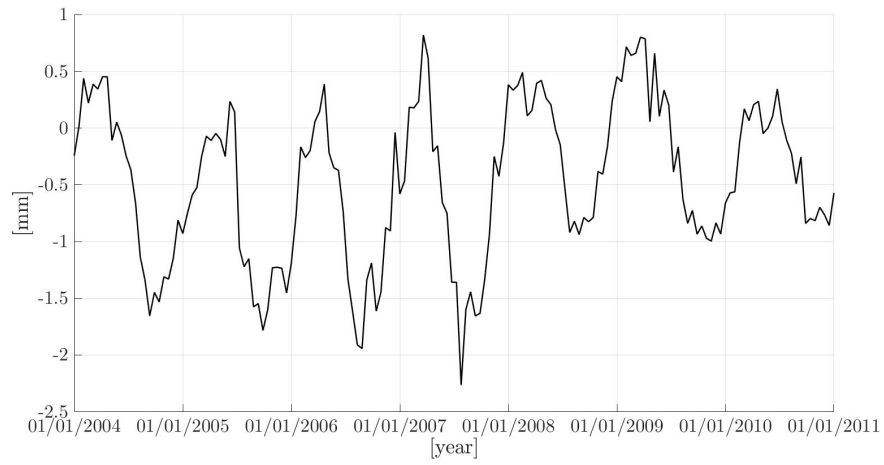


Figure 3: Displacements of the point P04.

are used for validation purpose.

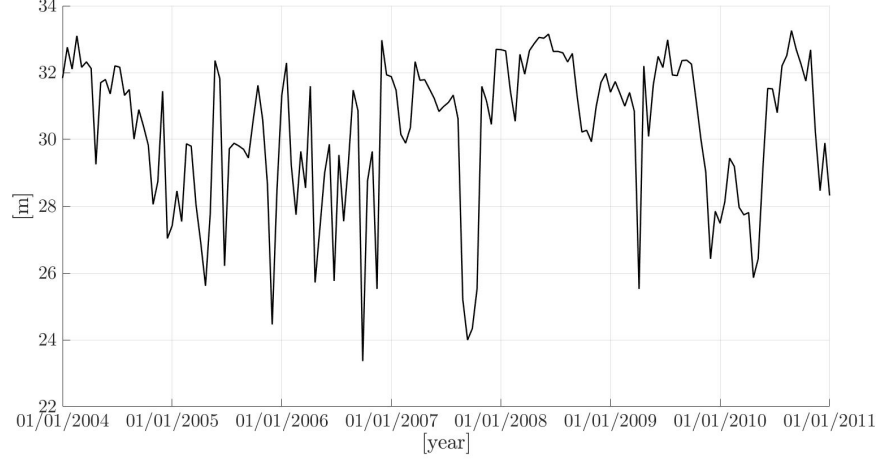


Figure 4: Recorded reservoir levels.

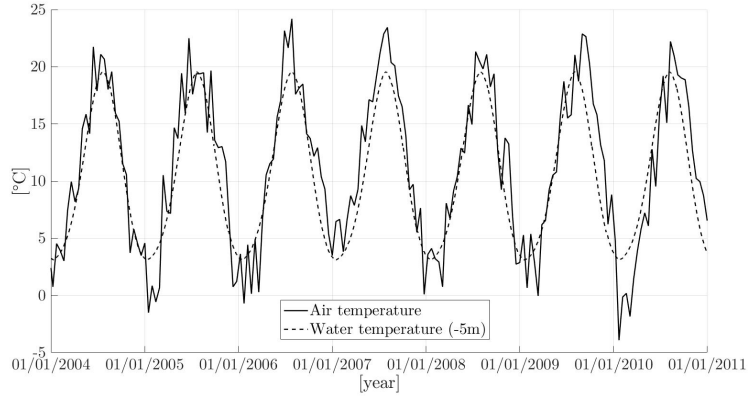


Figure 5: Recorded temperatures.

5.2. Prior distributions of the mechanical parameters

The mechanical parameters to be identified are the bulk modulus $\theta_{\text{gPCE},K}$ and the shear modulus $\theta_{\text{gPCE},G}$ of both the concrete and foundation rock. The results of the in-situ and laboratory tests, reported in table 1, are given in terms of Young modulus E and Poisson's coefficient ν , where subscript C indicates the concrete parameters and the subscript G indicates the ground parameters. Therefore, the well known equations

Table 1: Results of materials tests.

	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

$$K = \frac{E}{3(1-2\nu)}$$

$$G = \frac{E}{2(1+\nu)}, \quad (20)$$

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 2 [50]. The prior distributions

Table 2: 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

365

of the other parameters with no physical meaning, such as the combination coefficients of the Fourier series and the error standard deviation, are defined as non-informative (section 4.5).

5.3. Forward problem solution and response surfaces

370

In order to build the proxy model of the structure with the gPCE technique, four random variables, $\theta_{\text{gPCE},K_C} = K_C$, $\theta_{\text{gPCE},K_G} = K_G$, $\theta_{\text{gPCE},G_C} = G_C$, $\theta_{\text{gPCE},G_G} = G_G$, and the deterministic variable of the water level were 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.

375

The FE model represents the whole structure and a sufficiently wide portion of the ground around it, as indicated in the literature [51]. 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
380 geometry is imported in ANSYS r.17 [52]. The mesh is composed by 231254
quadratic hexahedral elements, for a total of 870787 nodes, as shown in figure 6.
Hertzian bonded contacts are introduced between the monoliths and the con-
stitutive law of the materials concrete and rock are linear elastic.

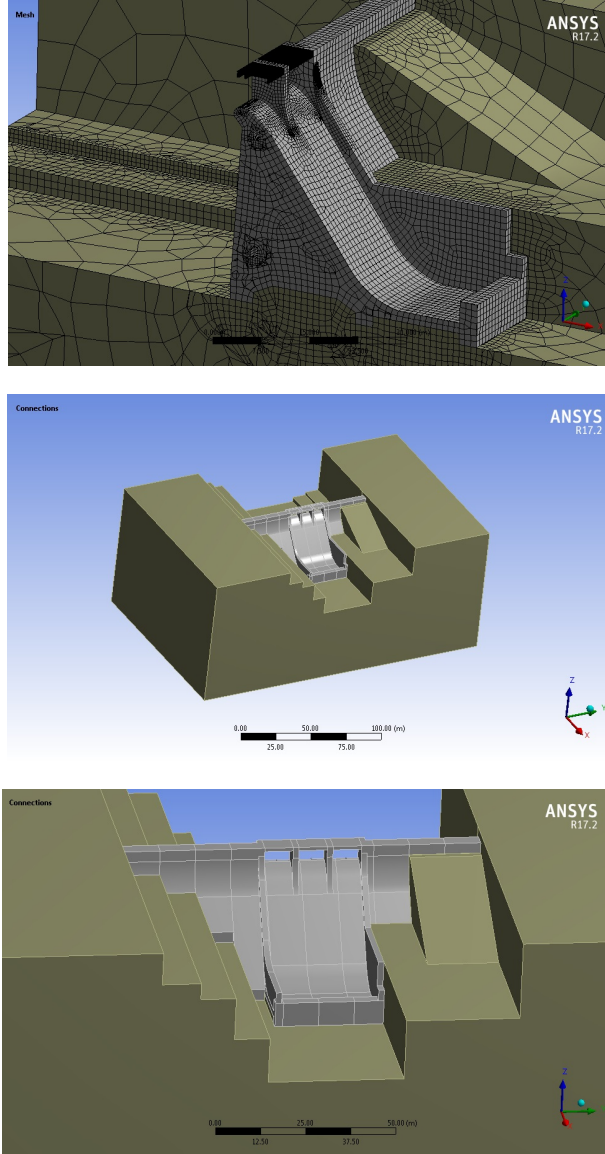


Figure 6: Finite elements model.

385 Figure 7 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 us to obtain 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
 390 build the proxy model, depending on the selected random variables and on the
 degree of the polynomial expansion, is 1024 [16]. The tools for the probabilis-
 tic analysis and Bayesian updating are developed in MATLAB R2017 [53]. The
 gPCE technique provides the response surfaces. Figure 8(a) shows displacement
 due to the hydrostatic load by varying the bulk moduli K of both materials. It
 395 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(b).

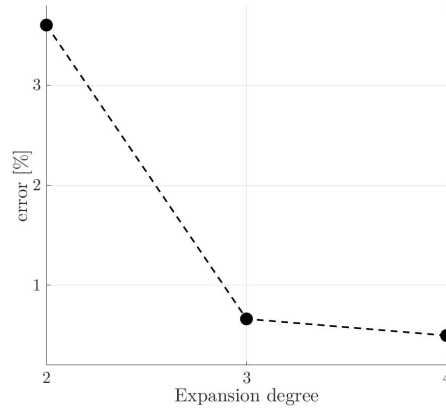
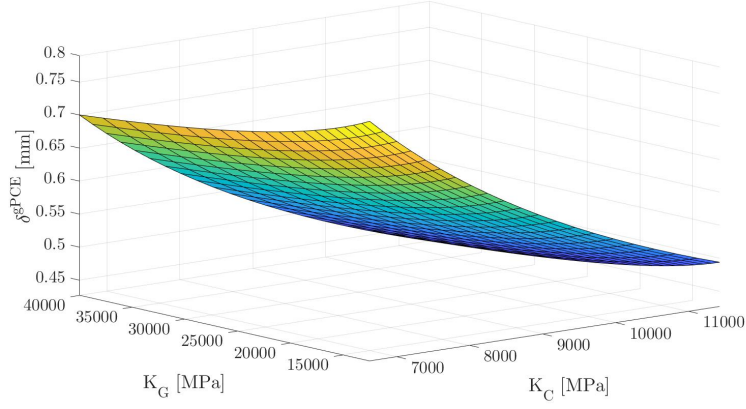
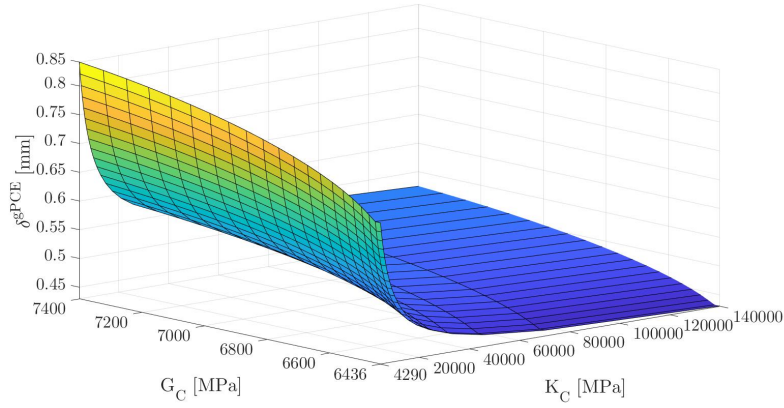


Figure 7: gPCE error.



(a) Model output response surface versus K_C and K_G for a reservoir level equal to 30 m.



(b) Model output response surface versus K_C and G_C for a reservoir level equal to 30 m.

Figure 8: Response surfaces.

5.4. The Fourier analysis

400 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.

405 *5.5. The Bayesian updating and the posterior statistics of the unknown parameters*

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 3), together with the data of water and air temperatures and reservoir levels (figures 4 and 5), recorded from 2004 to 2008. The results of the updating procedure for the three analyses are reported in tables 3, 4 and 5. The posterior mean values of the bulk moduli 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, 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 have a small variation in every case, proving that they are little influenced by the variation in the harmonics number.

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 9, prior and posterior distributions 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 nevertheless 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 , then the comparison between prior and posterior distributions is reported also with regard to these parameters. Figure 10 shows the results in terms of E and ν , in accordance to the results indicated

in figure 9 the mean values of the parameters do not show notable changes, while
the standard deviation values are strongly reduced. In this case of study only
the mean values of the distributions are updated, thus in Figure 9 and Figure
10 the prior distributions are compared with the posterior distributions of the
mean values.

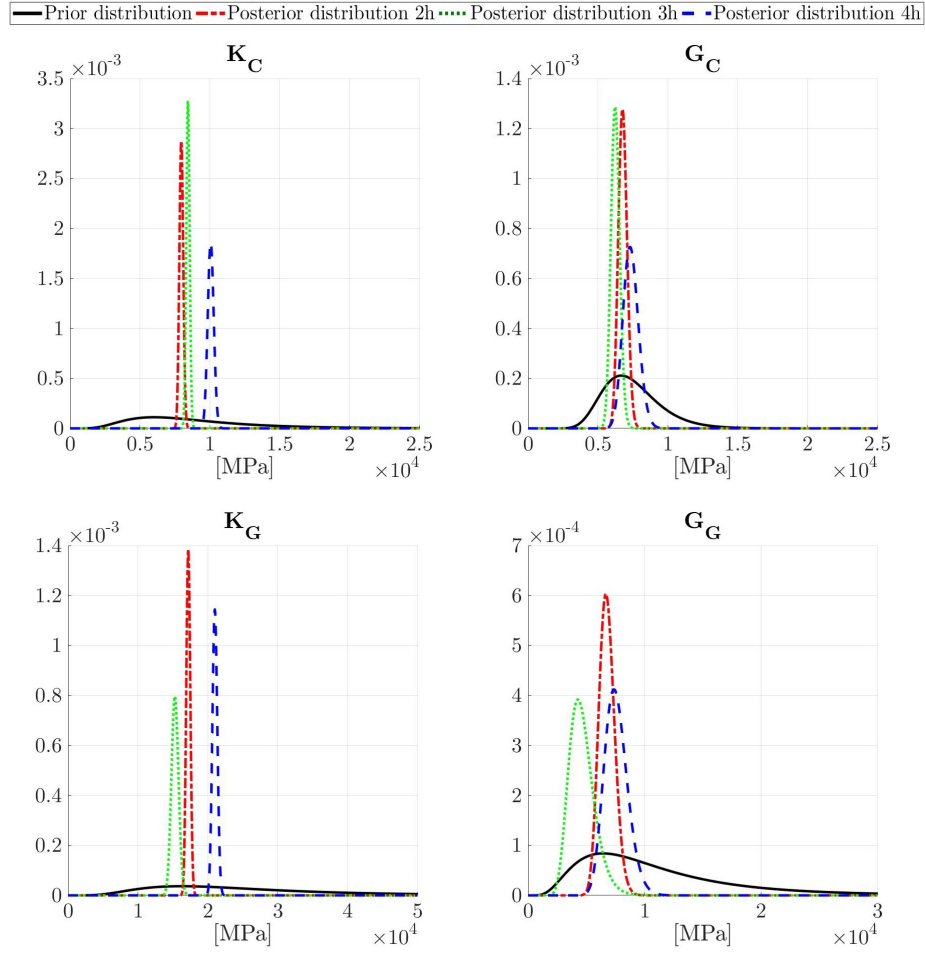


Figure 9: Comparison between prior and posterior distributions in terms of K and G .

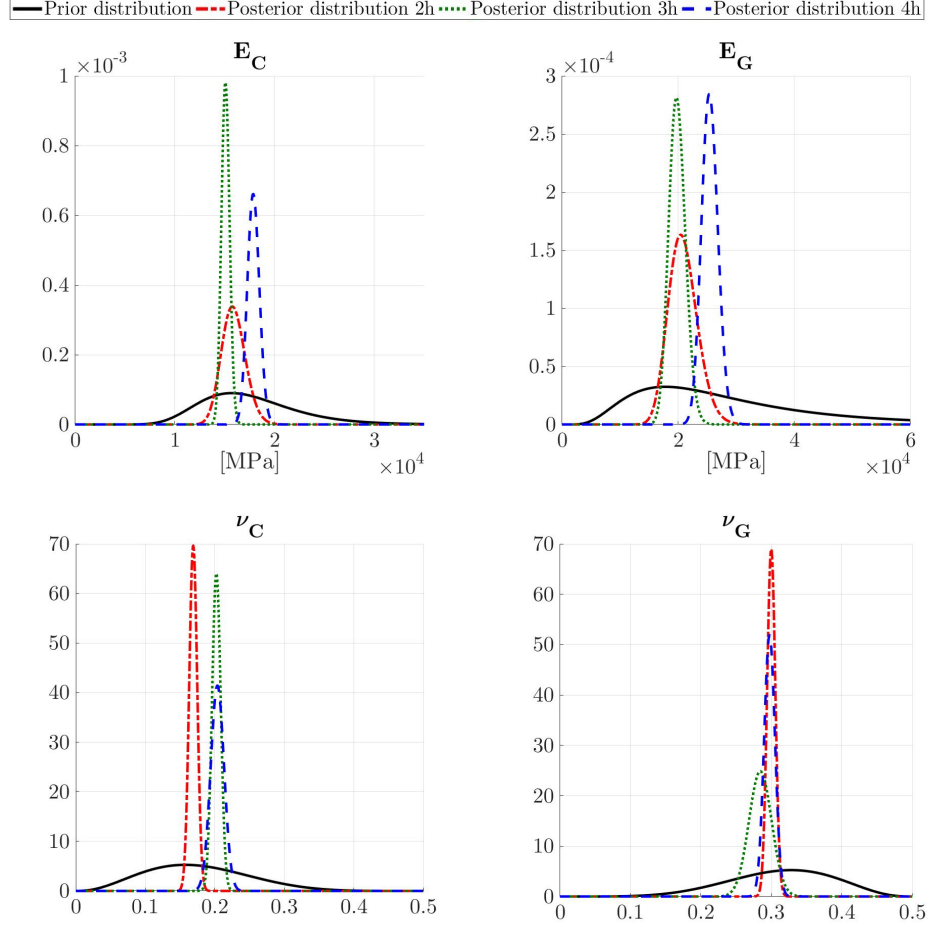


Figure 10: Comparison between prior and posterior distributions in terms of E and ν .

The numerical algorithm has been set up in order to interrupt the analysis
 440 when \hat{R}^p attains a value within the range 1 - 1.4. According to section 3.3 this
 value endorses the convergence of the analyses. Figure 11 shows both recorded
 and calculated displacements. One can note the agreement between them for
 every value of the Fourier series order.

Figures 12 shows the relationship between recorded and calculated displace-
 445 ments at every phase of the procedure for both two and four harmonics of the
 Fourier series. At the end of the procedure, the agreement between recorded

and calculated displacements is notable in both cases.

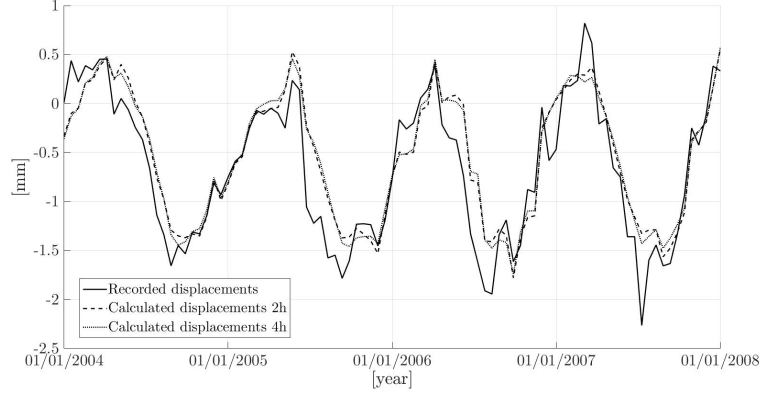


Figure 11: Comparison between recorded and calculated displacements.

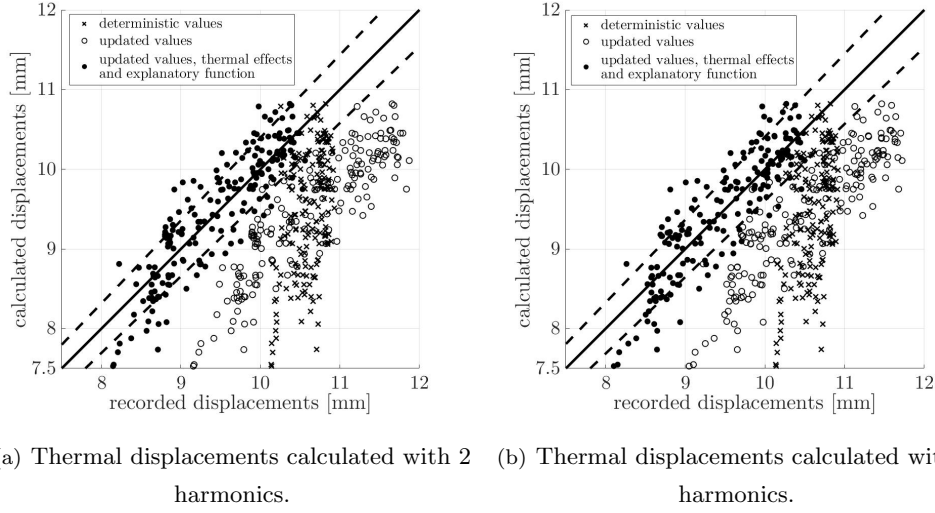


Figure 12: Effects of the updating procedure in terms of residual.

In order to assess the validity of the proposed procedure, the calibrated proposed predictive model is compared to a regressive predictive model commonly used in the literature [13, 18]. By following the indications of the literature, and

Table 3: Posterior statistics 2 harmonics.

	θ_{gPCE,K_C}	θ_{gPCE,G_C}	θ_{gPCE,K_G}	θ_{gPCE,G_G}	θ_1
	[MPa]	[MPa]	[MPa]	[MPa]	
Mean	7960	6788	17206	7987	-0.117
s.d.	20.14	469.23	33.89	5057	0.006
	θ_{FA,a_1}	θ_{FA,b_1}	θ_{FA,a_2}	θ_{FA,b_2}	σ
Mean	0.837	0.417	-0.102	-0.117	0.039
s.d.	0.045	0.041	0.045	0.038	0.002

Table 4: Posterior statistics 3 harmonics.

	θ_{gPCE,K_C}	θ_{gPCE,G_C}	θ_{gPCE,K_G}	θ_{gPCE,G_G}	θ_1	θ_{FA,a_1}
	[MPa]	[MPa]	[MPa]	[MPa]		
Mean	8437	6275	15302	7722	-0.121	0.841
s.d.	17.28	629.13	61.64	10687	0.003	0.052
	θ_{FA,b_1}	θ_{FA,a_2}	θ_{FA,b_2}	θ_{FA,a_3}	θ_{FA,b_3}	σ
Mean	0.443	-0.112	-0.104	0.055	-0.037	0.039
s.d.	0.046	0.047	0.039	0.050	0.0391	0.002

Table 5: Posterior statistics 4 harmonics.

	θ_{gPCE,K_C}	θ_{gPCE,G_C}	θ_{gPCE,K_G}	θ_{gPCE,G_G}	θ_1	θ_{FA,a_1}	θ_{FA,b_1}
	[MPa]	[MPa]	[MPa]	[MPa]			
Mean	10067	7425	21008	9808	-0.110	0.831	0.436
s.d.	30.44	1154	41.18	8247	0.004	0.038	0.045
	θ_{FA,a_2}	θ_{FA,b_2}	θ_{FA,a_3}	θ_{FA,b_3}	θ_{FA,a_4}	θ_{FA,b_4}	σ
Mean	-0.108	-0.104	0.054	-0.039	-0.018	0.017	0.040
s.d.	0.039	0.042	0.040	0.043	0.039	0.034	0.002

summarised in Section 2, the i -th value of the displacement obtained through the regressive predictive model is,

$$\begin{aligned}
\delta_i &= \delta_i^H + \delta_i^T + \delta_i^K = \\
&= a + bQ_i + cQ_i^2 + dQ_i^3 + e \sin(\omega t_i) + f \cos(\omega t_i) + gt_i.
\end{aligned} \tag{21}$$

Where, a, b, c, d, e, f, g are the combination coefficients of the regressive predic-

tive model, Q_i is the i -th value of the basin level, t_i is the i -th time step and ω is
450 the value of the assumed period. The resulting combination coefficients and the
coefficient of determination R^2 are reported in Table 6, while Figure 13 shows
the comparison between the predictions obtained with the proposed model and
those obtained with the regressive one (indicated as REG). Figure 13 shows

Table 6: Combination coefficients and R^2 of the regressive predictive model.

a	b	c	d	e	f	g	R^2
-85.4755	9.0582	-0.3244	0.0039	0.7416	0.3710	0.0001	0.8390

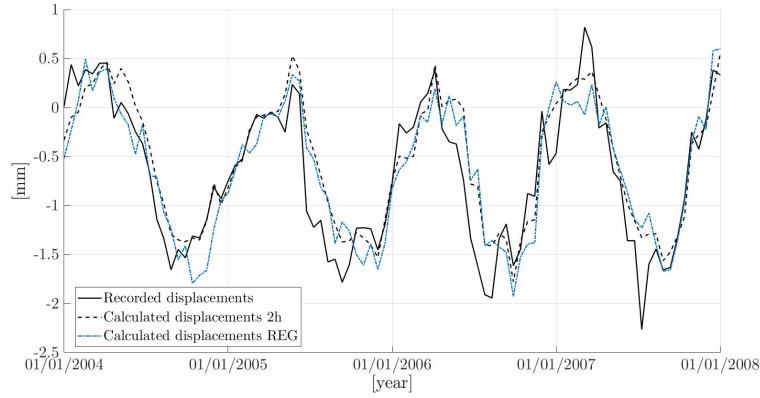


Figure 13: Comparison between the proposed predictive model and the regressive one.

that the displacements obtained through the two different predictive models in
455 the calibration period provide comparable results. Indeed, both of them allow
one to obtain a good agreement between records and predictions.

5.6. Procedure validation

In this section, a possible use of the procedure for the displacement moni-
toring of the dam structure during its regular use is presented. The proposed
460 calibrated model with two harmonics is used to predict the crest displacement
in the period within 2008 to 2011. Such predictions are then compared with

the displacements recorded by the monitoring system. Results are illustrated in figure 14, showing 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.

465 The calibrated probabilistic model for the dam displacement could be used for a real time structural control, due to its calculation speed. The proposed model allows evaluating the global error standard deviation, which can be used to define a threshold.

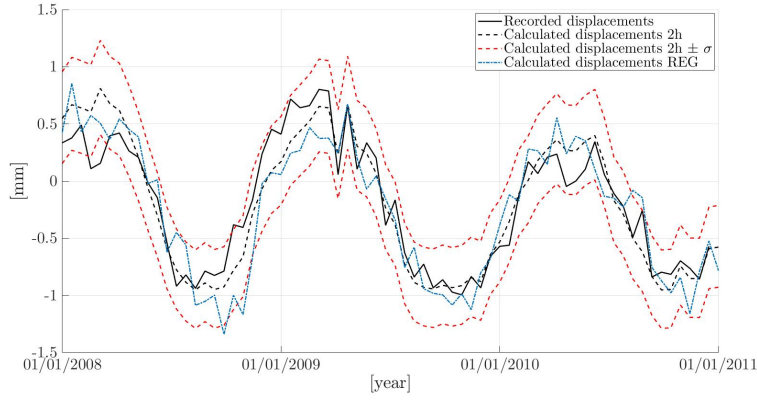


Figure 14: Comparison between recorded and calculated displacements of P04.

470 In Figure 14 also the prediction obtained with the regressive model is shown. By comparing the proposed predictive model with the regressive one it may be observe that the two predictions provide comparable results. Nevertheless, the proposed predictive model shows better performance since it is closer to the recorded displacements than the regressive one.

475 It is worth noting that, unlike the common approaches, the proposed hybrid predictive model allows one to calibrate the mechanical parameters of the materials thanks to the gPCE based meta model. Moreover, the definition of the procedure in the Bayesian setting allows estimating the error of the prediction itself and improving the reliability of the structural control every time new measurements are available.

480

6. Concluding remarks

The evaluation of the seismic safety of existing dams, and their control after earthquakes, are important and challenging tasks for the scientific community. The main issues are the uncertainties and the complexity of the system. In order
485 to obtain reliable models, all the information around the dam can be used to improve our knowledge. In this regard, monitoring data recorded during normal operations offer the opportunity to calibrate the numerical model.

The present study has adequately employed advanced techniques based on Bayesian updating to reduce model uncertainties related to material proper-
490 ties and the geometry of concrete gravity dams. The procedure proposed in this paper makes use of proxy models simulating the response of the dam in order to significantly reduce the computational burden due to the large number of analyses required by the MCMC technique. More specifically, the FE model response related to water level variation is approximated through the general polynomial
495 chaos expansion technique which creates a response surface depending on the selected parameters, thus making the uncertainties propagation computationally simple. The structural response to thermal variations is instead simulated by a Fourier analysis. The application of the procedure to the case study of a large Italian dam shows the advantages of the method. The procedure can be used
500 not only to update model parameters but also to identify the most significant ones and to select the best model, also evaluating the posterior error standard deviation. Finally, the procedure validation proves the possibility to directly use the updated proposed probabilistic model to control the structural behaviour after seismic events. In fact, by comparing the recorded displacements with
505 those predicted with the proposed model, in real time, abnormal behaviours can be detected.

Acknowledgements

This work has been developed under the research agreement between the Department of Energy, Systems, Territory and Constructions Engineering of the,

510 University of Pisa and the Directorate of Dams of the Italian Ministry of Infrastructure and Transport, whose financial support is gratefully acknowledged.

References

- [1] ASDSO, State and federal oversight of dam safety must be improved, Mag. Assoc. State Dam Saf. Off. (2011).
- 515 [2] J. F. Hall, The dynamic and earthquake behaviour of concrete dams: review of experimental behaviour and observational evidence, *Soil Dyn. Earthq. Eng.* 7 (2) (1988) 58–121.
- [3] L. Zhang, M. Peng, D. Chang, Y. Xu, *Dam Failure Mechanisms and Risk Assessment*, John Wiley & Sons Singapore Pte. Ltd., 2016.
- 520 [4] M. A. Hariri-Ardebili, V. E. Saouma, Seismic fragility analysis of concrete dams: A state-of-the-art review, *Eng. Struct.* 128 (October) (2016) 374–399.
- [5] FERC-PFMA, FERC guidance document: potential failure mode analysis, Tech. rep., Federal Emergency Regulatory Committee (2005).
- 525 [6] FEMA-PFM, Selecting analytic tools for concrete dams to address key events along potential failure mode paths, Tech. rep., Colorado: Federal Emergency Management Agency Denver (2011).
- [7] USBR-manual, Dam safety risk analysis, best practices training manual, version 2.2, Tech. rep., Denver, Colorado: U.S. Department of the Interior Bureau of Reclamation in corporation with the U.S. Army Corps of Engineers (2011).
- 530 [8] M. Andreini, A. D. Falco, G. Marmo, M. Mori, G. Sevieri, Modelling issues in the structural analysis of existing concrete gravity dams, in: *Proc. 85th ICOLD Annu. Meet.*, Prague, 2017, pp. 363–383.

- 535 [9] A. De Falco, M. Mori, G. Sevieri, FE models for the evaluation of hydrodynamic pressure on concrete gravity dams during earthquakes, in: Proc. 6th Eur. Conf. Comput. Mech. (ECCM 6) 7th Eur. Conf. Comput. Fluid Dyn. (ECFD 7), Glasgow, 2018, pp. 1731–1742.
- [10] A. De Falco, M. Mori, G. Sevieri, Simplified Soil-Structure Interaction models for concrete gravity dams, in: Proc. 6th Eur. Conf. Comput. Mech. (ECCM 6) 7th Eur. Conf. Comput. Fluid Dyn. (ECFD 7), Glasgow, 2018, pp. 2269–2280.
540
- [11] A. De Falco, M. Mori, G. Sevieri, N. Zani, Simulation of concrete crack development in seismic assessment of existing gravity dams, in: Proc. XVII Convegno ANIDIS “L’Ingegneria Sismica Italiana”, Pistoia, 2017.
545
- [12] M. A. Hariri-Ardebili, V. E. Saouma, Collapse Fragility Curves for Concrete Dams: Comprehensive Study, J. Struct. Eng. 142 (10).
- [13] A. De Sortis, P. Paoliani, Statistical analysis and structural identification in concrete dam monitoring, Eng. Struct. 29 (1) (2007) 110–120.
- 550 [14] J. Davis, Methods of Applied Mathematics with a Software Overview, 2nd Edition, Birkhäuser Basel, 2016.
- [15] P. Gardoni, A. Der Kiureghian, K. M. Mosalam, Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations, J. Eng. Mech. 128 (10) (2002) 1024–1038.
- 555 [16] D. Xiu, Numerical Methods for Stochastic Computations, Princeton University Press, 2010.
- [17] T. Marwala, Finite Element Model Updating Using Computational Intelligence Techniques: Applications to Structural Dynamics, 1st Edition, Springer Publishing Company, Incorporated, 2010.
- 560 [18] ICOLD, Bulletin 118: Automated Dam Monitoring Systems Guidelines and Case Histories, Tech. rep., ICOLD, Paris (2000).

- [19] P. Bukenya, P. Moyo, H. Beushausen, C. Oosthuizen, Health monitoring of concrete dams: A literature review, *J. Civ. Struct. Heal. Monit.* (2014).
- [20] C.-Y. Kao, C.-H. Loh, Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches, *Struct. Control Heal. Monit.* 20 (2013) 282–303.
- [21] J. Mata, A. Tavares de Castro, J. Sá da Costa, Constructing statistical models for arch dam deformation, *Struct. Control Heal. Monit.* 21 (2014) 423–437.
- [22] F. Kang, J. Liu, J. Li, S. Li, Concrete dam deformation prediction model for health based on extreme learning machine, *Struct. Control Heal. Monit.* (2017) 24 (e1997).
- [23] G. Prakash, A. Sadhu, S. Narasimhan, J. M. Brehe, Initial service life data towards structural health monitoring of a concrete arch dam, *Struct. Control Heal. Monit.* (2018) 25 (e2036).
- [24] B. Dai, C. Gu, E. Zhao, X. Qin, Statistical model optimized random forest regression model for concrete dam deformation monitoring, *Struct. Control Heal. Monit.* 25 (2018) e2170.
- [25] C. Lin, T. Li, X. Liu, L. Zhao, S. Chen, H. Qi, A deformation separation method for gravity dam body and foundation based on the observed displacements, *Struct. Control Heal. Monit.* 26 (2019) e2304.
- [26] E. N. Chatzi, *Identification Methods for Structural Health Monitoring*, 1st Edition, Vol. 567, Springer International Publishing, 2016.
- [27] B. Conde, P. Eguía, G. E. Stavroulakis, E. Granada, Parameter identification for damaged condition investigation on masonry arch bridges using a Bayesian approach, *Eng. Struct.* (2018).
- [28] A. de Falco, M. Girardi, D. Pellegrini, L. Robol, G. Sevieri, Model parameter estimation using Bayesian and deterministic approaches: the case study of the Maddalena Bridge, *Procedia Struct. Integr.* 11 (2018) 210–217.

- 590 [29] T. Janda, M. Šejnoha, J. Šejnoha, Applying Bayesian approach to predict deformations during tunnel construction, *Int. J. Numer. Anal. Methods Geomech.* (2018).
- [30] Y. Huang, C. Shao, B. Wu, J. L. Beck, H. Li, State-of-the-art review on Bayesian inference in structural system identification and damage assess-
595 ment, *Adv. Struct. Eng.* (2018).
- [31] M. Andreini, P. Gardoni, S. Pagliara, M. Sassu, Probabilistic Models for Erosion Parameters and Reliability Analysis of Earth Dams and Levees, *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* 2 (4).
- [32] A. Kalinina, M. Spada, P. Burgherr, S. Marelli, B. Sudret, A bayesian
600 hierarchical modelling for hydropower risk assessment, in: *Risk, Reliab. Saf. Innov. Theory Pract. - Proc. 26th Eur. Saf. Reliab. Conf. ESREL 2016*, 2017.
- [33] S. J. Peter, A. Siviglia, J. Nagel, S. Marelli, R. M. Boes, D. Vetsch, B. Sudret, Development of Probabilistic Dam Breach Model Using Bayesian In-
605 ference, *Water Resour. Res.* (2018).
- [34] M. Andreini, P. Gardoni, S. Pagliara, M. Sassu, Probabilistic models for the erosion rate in embankments and reliability analysis of earth dams, *Reliab. Eng. Syst. Safe.* 181 (2019) 142–155.
- [35] P. Gardoni, A. Der Kiureghian, K. M. Mosalam, Probabilistic models and
610 fragility estimates for bridge components and systems, Tech. rep., Pacific Earthquake Engineering Research Center, College of Engineering, Univ. of California, Berkeley, CA (2002).
- [36] Z. Yang, A modified family of power transformations, *Econ. Lett.* 92 (1) (2006) 14–19.
- 615 [37] G. E. P. Box, G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Wiley-Interscience, 1992.

- [38] B. V. Rosić, J. Sýkora, O. Pajonk, A. Kučerová, H. G. Matthies, Comparison of Numerical Approaches To Bayesian Updating, in: Computational Methods for Solids and Fluids, Springer International Publishing, 2016, Ch. 16.
- [39] D. Gamerman, H. F. Lopes, Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference, Chapman and Hall/CRC, 2006.
- [40] L. Tierney, Markov Chains for Exploring Posterior Distributions, Ann. Stat. 22 (4) (1994) 1701–1762.
- [41] R. L. Tweedie, Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space, Stoch. Process. their Appl. 3 (4) (1975) 385–403.
- [42] A. Kučerová, J. Skora, B. Rosić, H. G. Matthies, Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials, J. Comput. Appl. Math. 236 (18) (2012) 4862–4872.
- [43] S. P. Brooks, A. Gelman, General methods for monitoring convergence of iterative simulations, J. Comput. Graph. Stat. 7 (4) (1998) 434–455.
- [44] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, Bayesian Data Analysis, Chapman and Hall/CRC, 2004.
- [45] J. Noorzaei, K. H. Bayagoob, W. A. Thanoon, M. S. Jaafar, Thermal and stress analysis of Kinta RCC dam, Eng. Struct. 28 (13) (2006) 1795–1802.
- [46] K.-K. Kim, R. D. Braatz, Generalised polynomial chaos expansion approaches to approximate stochastic model predictive control, Int. J. Control 86 (8) (2013) 1324–1337.
- [47] S. Timoshenko, J. N. Goodier, Theory of Elasticity, 3rd Edition, McGraw-Hill College, 1970.
- [48] P. Léger, S. Seydou, Seasonal Thermal Displacements of Gravity Dams Located in Northern Regions, J. Perform. Constr. Fac. 23 (2009) 166–174.

- [49] A. B. Owen, Empirical Likelihood, Chapman and Hall/CRC, 2001.
- 645 [50] JCSS, Probabilistic Model Code, Tech. rep. (2001).
- [51] H. Wang, M. Feng, H. Yang, Seismic nonlinear analyses of a concrete gravity dam with 3D full dam model, Bull. Earthq. Eng. 10 (6) (2012) 1959–1977.
- [52] ANSYS, ANSYS Mechanical APDL Theory Reference, ANSYS Inc Release15 (2013) 909.
- 650 [53] The Mathworks Inc., MATLAB - MathWorks (2016).