Approximate Bayesian Computation for structural identification of ancient tie-rods using noisy modal data

Silvia Monchetti [a], Chiara Pepi [c,*], Cecilia Viscardi [b], Massimiliano Gioffrè [c]

[a] *Department of Civil and Environmental Engineering, University of Florence, Florence, Italy*
[b] *Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy*
[c] *Department of Civil and Environmental Engineering, University of Perugia, Perugia, Italy*

ARTICLE INFO

ABSTRACT

Masonry arches and vaults are common historic structural elements that frequently experience asymmetric loading due to seismic action or abutment settlements. Over the past few decades, numerous studies have sought to enhance our understanding of the structural behavior of these elements for the purpose of preventive conservation. The assessment of the structural performance of existing constructions typically relies on effective numerical models guided by a set of unknown input parameters, including geometry, mechanical characteristics, physical properties, and boundary conditions. These parameters can be estimated through deterministic optimization functions aimed at minimizing the discrepancy between the output of a numerical model and the measured dynamic and/or static structural response. However, deterministic approaches overlook uncertainties associated with both input parameters and measurements. In this context, the Bayesian approach proves valuable for estimating unknown numerical model parameters and their associated uncertainties (posterior distributions). This involves updating prior knowledge of model parameters (prior distributions) based on current measurements and explicitly considering all sources of uncertainties affecting observed quantities through likelihood functions. However, two significant challenges arise: the likelihood function may be unknown or too complex to evaluate, and the computational costs for approximating the posterior distribution can be prohibitive. This study addresses these challenges by employing Approximate Bayesian Computation (ABC) to handle the intractable likelihood function. Additionally, the computational burden is mitigated through the use of accurate surrogate models such as Polynomial Chaos Expansions (PCE) and Artificial Neural Networks (ANN). The research focuses on setting up numerical models for simple structural systems (tie-rods) and inferring unknown input parameters, such as mechanical properties and boundary conditions, through Bayesian updating based on observed structural responses (modal data, strains, displacements). The main novelties of this research regard, on the one hand, the proposal of a methodology for obtaining a reliable estimate of the axial force in ancient tie-rods accounting for different sources of uncertainty and, on the other hand, the application of ABC to obtain the structural identification inverse problem solution.

1. Introduction

The structural identification of ancient tie-rods is still an open issue in the scientific community due to the inherent uncertainties associated with the geometrical configuration and the mechanical properties of the material, as well as the difficulties in the acquisition of experimental data. In this class of problem, the definition of a proper Bayesian framework could be crucial to account for all the sources of uncertainties and clarify their role both in the data generative process and in the structural identification analysis. Moreover, the integration of prior knowledge is of paramount importance when the lack of experimental

data, or their non-informative features, threatens parameter identifiability and, consequently, the reliability of the results. However, the Bayesian approach is not yet widely used in this field, or it is not always well formalized.

In order to predict the structural behavior of existing tie-rods, numerical models are commonly applied. The main parameters characterizing these numerical models are the transversal cross-section shape and dimension, the effective length, the bending stiffness, and the boundary conditions used to reproduce the tie-rod insertion length into the lateral walls, as well as the tensile axial force. However,

---

as direct measurements of these quantities are not feasible, indirect methods have been developed on the base of in-situ non-destructive tests, such as dynamic monitoring systems. In particular, the dynamic monitoring systems of the structural modal parameter under environmental vibration loading are widely recognized as an effective method for investigating the evolution of structural health over time [1–3]. This approach applied to the tie-rods could provide useful information about the entire structure; indeed, changes in the structural behavior of tie-rods could reveal some changes in the structural configuration of the historical construction where they are placed. However, the experimental data, as well as the output of the numerical model, cannot provide the exact representation of the system output. Despite this consideration, usually, indirect methods rely on calibration strategies that disregard either measurement errors or any other source of uncertainty, since they are based on the mere minimization of the difference between numerical response and experimental data acquired from dynamic monitoring systems. In [4], the authors include an evaluation of the measurement error as the discrepancy, given in percentage, between the experimental data and the analytical results provided by the numerical modeling. Other research, see [5,6], proposes the well-established method of minimization of the error in frequency in order to update the numerical model. The issue of the development of a structural health monitoring system of tie-rods under uncertain environmental conditions is studied in [7], where the authors proposed an approach to separate the effect of environmental conditions on the variation of modal parameters from structural damage. Eventually, in other studies, experimental methods are exploited for estimating in situ tensile force in tie-rods through static, mixed static/dynamic, and fully dynamic methods in a deterministic framework [4–6,8–13]. However, to the best of our knowledge, all the existing works disregard uncertainties both in the model parameters and in the measurement errors, that are considered fixed quantities to be computed. Indeed, the discrepancy between the experimental data and the results of the numerical model is quantified without defining any likelihood function and the model parameters are not considered random variables. In other words, any effect of the inherent uncertainties in the physical and mechanical properties of tie-rods is overlooked. Therefore, it becomes crucial to define a procedure that formally quantifies the uncertainty and updates knowledge on unobserved quantities as experimental data are gathered. Even though proper uncertainty quantification is essential, the comprehensive formalization of this aspect is still an open issue and the implementation of uncertainty quantification strategies is promising to improve the reliability of the structural identification results and of the parameters themselves [14]. In this scenario, Bayesian inference emerges as a feasible solution; widely recognized in civil engineering, its application to existing buildings for model updating is driven by the need to solve inverse problems, and some attempts are provided in scientific literature, see [15–17]. Indeed, an inverse problem involves inferring the unobserved parameters governing a specific observed physics, especially when various sources of uncertainty contribute to collecting noisy and incomplete data. In such cases, due to measurement errors and other sources of uncertainties, the same input parameter can lead to observing different quantities. Bayes' theorem becomes of paramount importance in this scenario: it allows the use of pieces of evidence to update the probability of each unknown quantity. As a result, the inferential procedure provides a probabilistic solution that accounts both the uncertainty in the data generating process and the uncertainty around model parameters [18,19]. This paper proposes and formalizes a full Bayesian approach to estimate posterior distributions that quantify the uncertainty about the parameters accounting both for the prior uncertainty and the variability of the observable quantities. In particular, the variability around observed data is included by considering both measurement errors and the uncertainty driven by the lack of information about some quantities that interact with structural identification parameters and affect the value of experimental data. This is allowed by the definition of a statistical model,

rather than a deterministic one, that expresses observable quantities as a function of the unknown parameters and measurement errors but includes also latent variables. It is worth noting that, while measurement errors are often considered in the scientific literature, the introduction of latent variables is a novelty in this context. Indeed, commonly, the numerical model that links the uncertain parameters to the observable quantities is defined as a deterministic model where all quantities that are not parameters to be inferred are fixed to values arbitrarily chosen. However, as an effect of the introduction of latent variables, the likelihood function associated with the proposed statistical model is complex, thus making the computation of the posterior distribution infeasible. This paper shows two computational solutions to approximate posterior distributions. The first solution is a well-known Markov Chain Monte Carlo (MCMC) algorithm that introduces a data argumentation step to address the uncertainty about latent variables. A second solution is a likelihood-free approach known as Approximate Bayesian Computation (ABC). The two methods are compared at work on the same experiment, in order to assess the reliability of ABC methods, which are a less consolidated practice in this framework, and to highlight their flexibility that would allow relaxing some of the assumptions needed to implement MCMC algorithms. Furthermore, building upon the recent attempt to employ General Polynomial Chaos Expansion (PCE) method in [20] in the estimation of axial load in ancient tie-rods, we investigate the potential use of high-fidelity surrogate models within the framework of ABC. In particular, we evaluate the efficacy and feasibility of using such surrogate models to enhance the computational efficiency and accuracy of ABC-based solutions for structural problems by means of a direct comparison of the results obtained from the exact and the surrogate model. Furthermore, we propose a strategy based on preliminary sensitivity analyses to select inferring parameters and the latent variables among all the unknown quantities. The case at hand in this paper is a real-world case study. Finally, the choice of the optimal dimension of the observed data is discussed through a simulation study.

2. Bayesian inference and uncertainty quantification

The problem of structural identification can be seen as an *inferential* problem, meaning that we want to *learn* some characteristics of a structure relying on observed dynamics. In particular, one should rely on statistical inference when the relationship that links unknown quantities – i.e. the structure – with the observed dynamics includes some sources of randomness. In other words, even if the characteristics of the structure were known, we do not expect to always observe the same behavior of the system and collect the same data. This sampling variability can arise from the measurement process and/or from the way other variables, which are not the main target of the inferential process, interact with the considered system.

Let us denote by θ the vector of variables object of our inference, and by y_0 some observed data. In the framework of structural identification, θ are the characteristics of the structure that we want to identify (e.g. mechanical properties of the structure), while y_0 are data regarding the observed dynamic of the system (e.g., natural frequencies). To learn θ starting from y_0, there exist at least two different approaches. The frequentist approach considers θ as a vector of parameters – i.e. fixed but unknown quantities – while y_0 is just one of the possible realizations of a random variable Y defined over $\mathcal{Y}$ and whose probability distribution is governed by θ. In the Bayesian framework, the vector θ is in turn considered a random vector because it includes unknown quantities that are unobserved and unobservable. Thus, Bayesian methods aim at deriving a probability distribution for θ, in order to formalize and quantify the uncertainty about it. This output comes up as the result of a learning process. Usually, we are able to elicit a prior probability distribution $\pi(\cdot)$ over the space on which θ is defined, say Θ. This distribution formalizes the experts' prior belief about possible values for θ and their associated probabilities, while y_0 drives information that can be exploited to update our knowledge

about θ. Given a certain value of θ, different realizations of Y in $\mathcal{Y}$ are plausible. Thus, y_0 is just one of them and the probability function $p(y_0 \mid \theta)$, when seen as a function of θ, is called likelihood function, as it expresses how likely is a configuration of the system described by θ, in the light of observed data y_0.[1] Bayesian inference aims to update the prior knowledge after having observed data, and compute the posterior distribution relying on the Bayes' formula

$$\pi(\theta \mid y_0) = \frac{\pi(\theta)p(y_0 \mid \theta)}{\int_\Theta \pi(\theta)p(y_0 \mid \theta)d\theta}. \tag{1}$$

The denominator in Eq. (1) is equal to $p(y_0)$, known as the marginal likelihood, and acts as a normalizing constant. Except for the case of conjugacy [21], its computation requires numerical solutions usually based on Monte Carlo methods (see next section for further details).

In this work, the choice of relying on Bayesian approaches is to exploit the expert knowledge in the inferential process. Furthermore, we want to take into account all the sources of uncertainty and provide a posterior distribution to avoid the false sense of confidence given by point estimates. Finally, a formal quantification of the uncertainty is particularly relevant when one is interested in making predictions about the future behavior of the system: the uncertainty described through posterior distributions can be straightforwardly propagated to future dynamics of the system. However, the likelihood function is a key ingredient both in frequentist and Bayesian inference. It follows that a formal definition of all the unobserved quantities that interact with θ in the determination of the observed data, as well as their probability distributions, is needed in both the inferential frameworks. To this end, we adopt the formalization given in [22] to define all the relevant quantities.

Let X be a vector of latent variables, and $\xi = (\theta, X)$ the vector of all the unknown quantities. We can suppose that there exists a function $y^R(\theta)$ that maps the mechanical properties of interest to observable dynamics and describes the *real process* (indicated by the apex R). However, data are always collected with some random measurement errors, thus $y_0 = y^R(\theta) + e$. Usually, the function that describes the real process is unknown and the best we can do is to define a mathematical model that tries to reproduce it. This is often based on systems of partial differential equations whose solutions are approximated using a simulator based on a Finite Element Model (FEM). This latter can be thought of as a computer program that takes all the unknown quantities, i.e. the vector ξ, as input and outputs a simulated dynamic $y^M(\xi)$. Unfortunately, no model is perfect and there is always a discrepancy between the model and the reality, here called *bias* and denoted by $b(\xi) = y^R(\theta) - y^M(\xi)$. It may come from incorrect or missing physical characteristics, as well as the simplification of the problem needed to put it in a digital framework (e.g. space discretization in FEM). To sum up we can state the following equation

$$y_0 = y^R(\theta) + e = y^M(\xi) + b(\xi) + e.$$

A common practice in this field is to consider observed data just as the output of the mathematical model. This means that all the sources of uncertainty – i.e. measurement errors, latent variables, and the bias of the model – are ignored or considered equal to fixed quantities. Under this assumption, the probability of the observed data is a Dirac function, $p(y_0 \mid b, e, x, \theta) = \delta_{y_0}\{y^M(\theta, x) + b + e\}$, such that the probability of observing y_0 is 1 when the output of the mathematical model leads to $y^M(\theta, x) + b + e = y_0$, and 0 otherwise.

Here, we consider the case in which, x, e and b are unknown and specific probability distributions are assumed for each of them. In such

a case, the evaluation of the likelihood function requires the following marginalization:

$$\begin{aligned}
p(y_0 \mid \theta) &= \iiint p(y_0, e, b, x, \mid \theta)de\,db\,dx \\
&= \iiint p(y_0 \mid e, b, x, \theta)\,p(e, b, x \mid \theta)de\,db\,dx \\
&= \iiint \delta_{y_0}\{y^M(\theta, x) + b + e\}\,p(e, b, x \mid \theta)de\,db\,dx. \tag{2}
\end{aligned}$$

Looking at Eq. (2) it is apparent that point-wise evaluations of the likelihood function require numerical approximations of multiple integrals. Depending on the analytical form of $p(e, b, x \mid \theta)$ and the structure of dependence assumed among the random quantities involved in the model, the solution of those integrals may be more or less complex, but in any case, requires several calls to the simulator to evaluate the Dirac function. Note that, both frequentist and Bayesian methods use several computations of the likelihood function that lead to a high computational cost. A possible escape is to introduce an *emulator*, which is a surrogate model that mimics the mathematical model but is less expansive in computational terms.

The emulator. To understand what an emulator represents, we should start by considering the mathematical model as a function that maps each pair $\xi = (\theta, x)$ to a dynamic: $y^M : (\Theta, \mathcal{X}) \to \mathcal{Y}$. Sometimes, the computational cost of computing $y^M(\xi)$ is high, thus one can take advantage of the use of a function that approximates the model but has a lower computational cost. This function, here denoted by y^M_{em}, is called emulator and still is a map from $(\Theta, \mathcal{X})$ to $\mathcal{Y}$.

In practice, the definition of an emulator is based on three steps: (1) choose a strategy to select D evaluation points onto $\Theta \times \mathcal{X}$: $\xi_1 = (\theta_1, x_1), \ldots, \xi_D = (\theta_D, x_D)$; (2) compute $y^M(\xi_d)$ for each $d \in \{1, \ldots, D\}$; (3) fit a mathematical/statistical model (e.g. neural networks, regression models) by estimating its coefficients ϕ (e.g. weights and bases of the neural network or the regression coefficients) relying on the dataset $\{y^M_d, \xi_d\}_{d=1}^D$ where y^M_d and ξ_d are the response and explanatory variables, respectively. The estimated coefficients $\hat{\phi}$ will be used to compute predictions $y^M_{\text{em}}(\xi) = f(\xi; \hat{\phi})$.

3. Bayesian simulated inference: MCMC vs ABC

To evaluate the posterior distribution, Bayesian inference often relies on methods based on simulations, such as Monte Carlo methods. The main reason is that the marginal likelihood – i.e. the normalizing constant in Eq. (1) – cannot be computed analytically and requires numerical approximations. Moreover, models used to describe complex phenomena, often include several latent variables and nuisance parameters. In such cases, Bayesian methods involve further challenging integrals or summations over high dimensional spaces, such as those in Eq. (2). To solve this problem, at least two alternative solutions are available:

- MCMC for getting samples from a posterior distribution defined on an augmented space: $\pi(\theta, x \mid y_0)$;
- ABC algorithms for drawing samples from an approximate posterior distribution $\pi_\epsilon(\theta \mid y_0)$.

While MCMC methods are a well-established strategy, also in the structural identification framework, ABC methods are still less investigated. In this work, we test the use of ABC for structural identification and provide a comparison with a MCMC approach to assess the reliability of ABC and show its potential and flexibility.

Markov chain Monte Carlo. To overcome the problem of evaluating the intractable normalizing constant, one can resort to MCMC methods. These algorithms usually form a Markov Chain $\theta^{(0)}, \ldots, \theta^{(S)}$ whose limiting distribution is the target $\pi(\theta \mid y_0)$. Thus, after the check of the convergence of the chain, samples drawn using MCMC methods can be considered as samples from the true posterior distribution and can be

---

[1] For the sake of simplicity, our notation does not discriminate between probability density functions and mass functions, both denoted by $p(\cdot \mid \theta)$ that can be distinguished from the context.

Algorithm 1. Metropolis-Hastings

1: **Inizialize** $\xi^{(0)} = (\theta^{(0)}, x^{(0)})$
2: **for** s in $1 : S$ **do**
3:     **Draw** $\xi^* \sim q(\cdot \mid \xi^{(s-1)})$
4:     **Solve** the deterministic FE model to compute $y^M(\xi^*)$
5:     **Compute** $\alpha = \min \left\{ 1, \dfrac{\pi(\xi^*) p(y_0 - y^M(\xi^*) \mid \xi^*) q(\xi^{(s-1)} \mid \xi^*)}{\pi(\xi^{(s-1)}) p(y_0 - y^M(\xi^{(s-1)}) \mid \xi^{(s-1)}) q(\xi^* \mid \xi^{(s-1)})} \right\}$
6:     **Sample** $u \sim U(0,1)$
7:     **if** $u < \alpha$ **then**
8:         **Set** $\xi^{(s)} = \xi^*$
9:     **else**
10:         **set** $\xi^{(s)} = \xi^{(s-1)}$
11:     **end if**
12: **end for**

---

used to approximate posterior quantities, such as mean, quantiles, and densities. However, MCMC algorithms require a point-wise evaluation of the likelihood function $p(y_0 \mid \theta)$ at each step. Looking at Eq. (2), it is apparent that these computations, in the structural identification framework, have a high computational cost due to the presence of multiple integrals and the need to call the mathematical model at each evaluation of the Dirac function.

A possible strategy is to define a posterior distribution on the augmented space $\Theta \times \mathcal{X}$:

$$\pi(\theta, x \mid y_0) \propto \pi(\theta, x) p(y_0 \mid x, \theta), \tag{3}$$

where the likelihood function is $p(y_0 \mid x, \theta) = p(y_0 \mid \xi)$. Similar to what has been derived in (2) we have that

$$p(y_0 \mid \xi) = \iint p(y_0 \mid e, b, \xi) \, p(e, b \mid \xi) de \, db$$
$$= \iint \delta_{y_0}(y^M + b + e) \, p(e, b \mid \xi) de \, db.$$

This function still involves multiple integrals but, under the assumption that (a) there is no bias in the model (or it is known); (b) measurement errors have a simple distribution, e.g. Standard Normal distribution; it can be evaluated as follows

$$p(y_0 \mid \xi) = \int \delta_{y_0}(y^M + e) \, p(e \mid \xi) de = p(y_0 - y^M \mid \xi)$$

where $p(\cdot \mid \xi)$ is the Standard Normal density. In such a case, the widespread Metropolis–Hastings (MH) algorithm [23] can be implemented, as displayed in Alg 1.

The algorithm produces a Markov chain whose limiting distribution is $p(\xi \mid y_0)$. The speed of convergence to the target distribution depends also on the choice of the proposal distribution $q(\cdot \mid \cdot)$ and the induced autocorrelation in the chain. However, after the check of the convergence, we can disregard samples $x^{(1)}, \dots, x^{(S)}$ and take $\theta^{(1)}, \dots, \theta^{(S)}$ as an Independent and Identically Distributed (i.i.d.) sample from the posterior distribution $p(\theta \mid y_0)$. For further details about MCMC methods, we refer the reader to [24].

Note that to reduce the computational cost of evaluating the acceptance probability α, one can replace the FE model with the emulator to compute the prediction $y^M_{\text{em}}(\xi^*)$ at step 4 in Alg 1. In such a case, also the acceptance probability at step 5 would be based on $y^M_{\text{em}}(\xi^*)$ and $y^M_{\text{em}}(\xi^{(s-1)})$ rather than $y^M(\xi^*)$ and $y^M(\xi^{(s-1)})$.

Approximate Bayesian computation. ABC is a class of likelihood-free methods. They allow Bayesian simulated inference without exact likelihood computation and only require the availability of a generative model – i.e. a computer program that takes the parameters object of the inference as inputs, performs stochastic calculations that reproduce the data generative process, and outputs pseudo data y. The key idea can be traced back to [25–27] and consists of replacing the evaluation of the likelihood function with a Monte Carlo (MC) estimate. The basic ABC

algorithm follows a rejection scheme made of the three steps displayed in Alg 2.

Algorithm 2 Rejection ABC

1: **Draw** $\theta^{(1)}, \dots, \theta^{(S)}$ as an i.i.d. sample from the prior distribution $\pi(\cdot)$
2: **Run** the *generative model* to draw $y^{(s)} \sim p(\cdot \mid \theta^{(s)})$ for each $s \in \{1, \dots, S\}$
3: **Retain** $\theta^{(s)}$ such that $d(y_0; y^{(s)}) \leq \epsilon$

Algorithm 3 Generative Model

1: **Take** θ as an input
2: **Sample** x from its prior distribution
3: **Compute** $y^M(\theta, x)$ using the simulator (FEM) – or the emulator
4: **Sample** e and b from their distributions
5: **Return** $y = y^M(\xi) + b + e$

The algorithm is based on a positive tolerance threshold ϵ and a distance function $d(\cdot; \cdot)$. The output is an i.i.d. sample from an *approximate* posterior distribution:

$$\pi_\epsilon(\theta \mid y_0) \propto \pi(\theta) \int \mathbb{1}\{d(y_0; y) \leq \epsilon\} p(y \mid \theta) dy$$
$$= \pi(\theta) \Pr(d(y_0; y) \leq \epsilon \mid \theta)$$

where the likelihood function is replaced by the probability that $\Pr(d(y_0; y) \leq \epsilon \mid \theta)$ – i.e., the expected value of the indicator function which can be easily approximated using an MC estimate. Let us consider the case of discrete data and impose an equality constraint with $\epsilon = 0$ – i.e. accept only if $y_0 = y$. The probability $\Pr(d(y_0; y) \leq \epsilon \mid \theta)$ would be $\Pr(Y = y_0 \mid \theta)$ and the output of the algorithm would be a sample from the exact posterior distribution. Generally speaking, the smaller is ϵ, the smaller is the approximation and as $\epsilon \to 0$ the approximate posterior converges to the true posterior.

This algorithm has several drawbacks. First of all, sampling parameter proposals from the prior distribution can be highly inefficient when the posterior is far from the prior. Furthermore, the choice of the tolerance threshold establishes a trade-off: the smaller is ϵ, the lower the approximation but also the acceptance rate. In the statistical literature, many methodological advances have been made to overcome these problems. Most of them rely on sequential sampling schemes that adapt, at each iteration, both the proposal distribution and the tolerance threshold (see [28] for further details on the method). However, all of them still require several samples from the simulator that is often computationally expansive. Alg 3 summarizes how the generative model produces pseudo-data. Looking at the algorithm it is apparent that most of the computational effort is attributable to the mathematical model. Thus, also in ABC methods, one can take advantage by replacing the mathematical model with the emulator at step 3 in Alg 3.

Note that ABC does not require the assumptions formulated for the MCMC method described in the previous section. In fact, b is
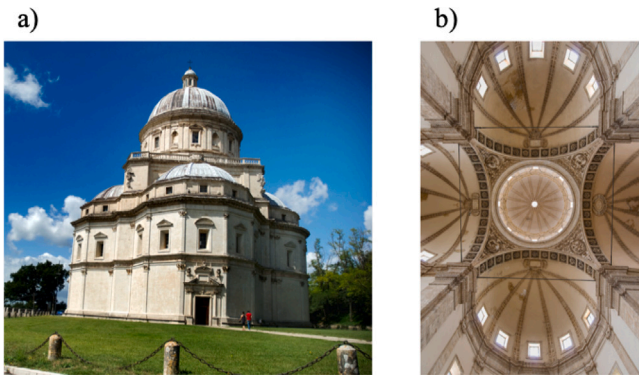
a) b)

Fig. 1. Santa Maria della Consolazione Temple (a) and tie-rods (b).

straightforwardly included via simulation in Alg 3. Furthermore, more complex distributions can be assumed for the measurement errors since the method only requires the ability to simulate them. In summary, ABC in contrast with MCMC, produces an approximated version of the posterior distribution but offers a more flexible method, suitable under milder distributional assumptions.

As a general comment, MCMC methods should be preferred when dealing with models whose likelihood function is available, since they do not introduce sources of approximation and avoid calls to the generative model to evaluate the likelihood. However, in this framework, also MCMC algorithms require calls to the generative model. This suggests that ABC, in this scenario, may outperform likelihood-based methods. In addition, it is worth comparing the framework of ABC method with the calibration method commonly used in model optimization. The presented Rejection ABC is indeed based on a tolerance threshold and a distance function in accordance with the deterministic optimization functions for the minimization of the discrepancy between the observed data and the numerical output. The novelty herein introduced with the ABC framework is the possibility of taking into account the different sources of uncertainties involved in the identification process.

4. Case study

4.1. Experimental activities

One of the four tie-rods of the Santa Maria della Consolazione Temple in Umbria Region, Central Italy, was used as a case study to assess the capability of the ABC approach to identify the tensile axial force and other unknowns that govern the problem and allow the structural identification (Fig. 1 - a). The contactless measurements on these inaccessible tie-rods are proposed in [29]. The church follows a Greek cross plan having dimensions of 42 m by 48 m. It includes a prominent central dome reaching a 50 m height from the temple floor. Additionally, there are four semi-shells on the lateral sides of the structure. The central dome is constructed as a single shell with an ogival shape supported by a circular drum, further enhanced with ribs for reinforcement. At its pinnacle, the dome is crowned with a lantern. Further details about the architectural and structural characteristics of this Temple can be found in [30,31].

Four tie-rods attached to the spring line of the vaults by bolts and plates were installed at the end of 1800s as reinforcement (Fig. 1 - b). Currently, the four tie-rods geometrical and mechanical properties are uncertain. Consequently, numerical analyses need to account for uncertainties related to various variables including cross-sectional dimensions, mass density, elastic properties, and support boundary conditions.

Experimental activities were carried out in order to determine the tie-rods dynamic properties (i.e. natural frequencies). Two different contactless sensors were used for vibration measurements: laser vibrometer and radar interferometer. This choice was due to the fact that these tie-rods are located at 16 m height from the ground making challenging to install contact sensors such as accelerometers. For the same reason, free vibrations were induced giving a force pulse with a thin rope suitably connected to a drone using a special equipment. The drone was used to overpass the thin rope over the four tie-rods. Displacement time histories were recorded using the laser vibrometer and the radar interferometer. Natural frequencies were obtained from the power spectral density spectrum using peak picking technique. Further details about this topic can be found in a previous work [32].

In this work, the structural identification problem of the ancient tie-rods is solved for just one of the four tie-rods. Displacement time histories and associated power spectral density spectrum are shown in Fig. 2-a,b, respectively. Natural frequencies were obtained using the Pick Picking technique.

The first three natural frequencies, used as observed data, were equal to $y_0 = \{4.492, 9.288, 13.252\}$. They are considered as a realization of a random vector whose probability density function is defined according to the statistical model described in Section 2.

4.2. Statistical model

To identify the structure, we exploit information driven by y_0 to infer the following quantities: (1) N, the axial load; (2) E, the Young's modulus; (3) ρ, the mass density; (4) a, the square cross-section transversal dimension. Thus, $\theta = (N, E, \rho, a)$, is a four-dimensional vector. Its prior distribution has been defined by assuming four independent Zero left-truncated Gaussian distributions (see Table 1 for further details). As regards measurement errors we assume the following Multivariate Normal distribution: $e = (e_1, e_2, e_3) \sim \text{MVN}(\mu = (0, 0, 0); \Sigma = 10^{-3} \cdot \mathbf{1}_3)$, where $\mathbf{1}_3$ denotes the identity matrix of size 3. Note that both the assumption of normality and independence simplify the evaluation of the likelihood function and allow a straightforward implementation of the MCMC strategy. For the same reason, we ignore the bias of the model $b(\xi)$. Despite it can be easily taken into account when implementing an ABC algorithm, for the sake of comparison, we consider the same statistical model in the MCMC and ABC implementation.

To describe the mathematical model that relates θ to y_0, we relied on the computational model described in the following section.

4.3. Computational model

The structural identification problem for tie-rods is addressed using the computational model $y^M(\xi)$, as depicted in Fig. 3. The model consists of an Euler–Bernoulli beam with a uniform square cross-section, hinged at both ends and subject to two rotational springs. A constant axial load N is applied to the beam. The rotational springs account for the influence of the anchorage length of tie-rods on the lateral walls. In addition to the axial load, the input parameters governing the free transversal vibration problem in Fig. 3 include Young's modulus E, mass density ρ, square cross-section transversal dimension a of the tie-rod, the stiffness of the left and right rotational springs k_l and k_r, and the length l (distance between lateral walls). Three other variables, namely k_l, k_r, and l, contribute to determining the frequency values alongside the parameters that are the target of inference. The tie-rod length l is assumed to be a fixed and known quantity, equal to the distance between the two lateral walls. This choice is due to the fact that the effect of the embedding length of the rod within the lateral walls is reproduced by the computational model through the use of the two rotational springs at the ends of the axially loaded beam like structure in Fig. 3. Indeed, the uncertainties about the stiffnesses of the two rotational springs, i.e. k_l and k_r, were addressed by introducing into the model the vector of latent variables $x = (k_l, k_r)$, characterized by two independent Zero left-truncated Normal distributions (refer to
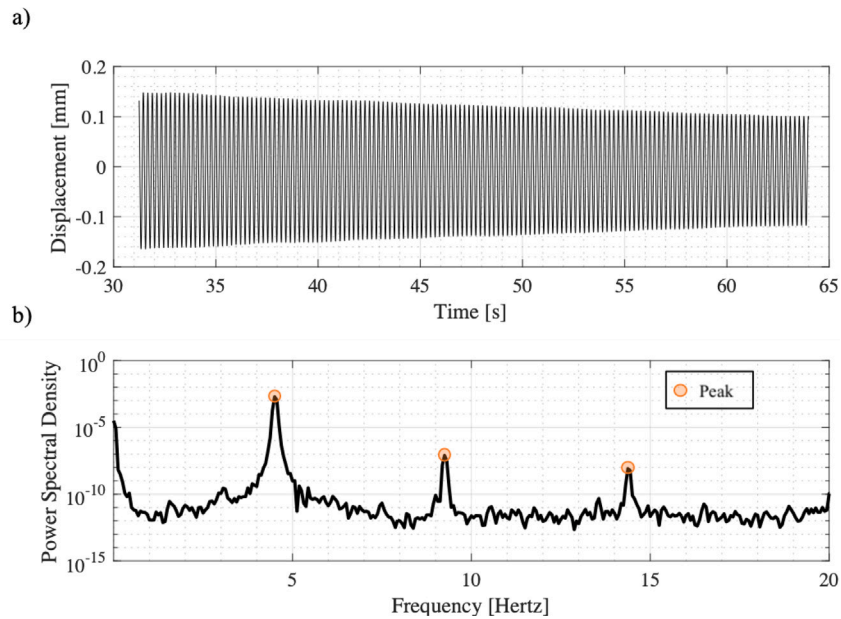
a)

b)

Fig. 2. Displacement time history recorded by laser vibrometer on the tie-rod (a) and associated power spectral density spectrum (b).

Table 1). The choice of setting the two rotational spring stiffnesses k_l and k_r as latent random variables was based on the outcomes of a preliminary sensitivity analysis aimed at identifying the computational model input parameters influencing the natural frequencies of the tie-rod. This analysis was carried out using the Sobol index decomposition method [33]. The computed Sobol indices showed that the computational model responses (i.e. first three natural frequencies) were mainly influenced by the elastic modulus, the mass density, the cross section transversal dimension and the axial force in the tie-rod. It further turned out that the computational model output has a low sensitivity to the stiffnesses of the two rotational springs. As an example, Fig. 4 depicts the variation of the first three natural frequencies obtained from the computational model using as input parameters for the elastic modulus, the mass density, the cross section transversal dimension and the axial force equal to the mean values of the distributions in Table 1 for a 14 m length beam. Meanwhile, the stiffnesses of the rotational springs assume variable values between the two conditions corresponding to hinged and fixed boundary conditions. Further details about this topic can be found in [34]. This behavior suggests that the values of the two spring stiffnesses cannot be learned as a result of the inference; however, k_l and k_r do have an impact, albeit small, on the variability of the frequencies. It is noteworthy that in the existing literature on this subject, in order to have an easy-to-handle mathematical (non-stochastic) model, these kinds of quantities are usually fixed to arbitrary values based on strong assumptions, such as symmetry in the stiffness of lateral rotational springs describing unknown boundary conditions. The framework proposed in the present work allows us to overcome all these limitations by considering non-symmetric boundary conditions and introducing a greater number of random variables than typically used in the literature for addressing the same problem.

As it is typically done, the mathematical model has been approximated through a finite element (FE) numerical model. In this study, we employ two-dimensional beam elements with two Degrees of Freedom (DoFs) at each node, representing deflection and rotation. Modal analyses are performed with a total of 50 beam elements. The eigenvalue problem is solved modifying the local stiffness matrix of the Euler Bernoulli beam elements introducing an additional matrix able to consider the additional stiffness due to the axial load. The boundary conditions, including rotational springs, are incorporated by adjusting the global stiffness matrix to the corresponding degrees of freedom accordingly.
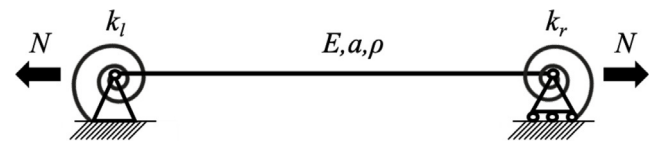

Fig. 3. Computational model reproducing the dynamic behavior of the tie-rod and uncertain model parameters θ.

Table 1
Summary of the probability density functions assumed for all the variables in the model.

Model Parameters θ			
Variable	**Distribution**	**Mean**	**CV**
E	Left-truncated Normal	2.0×10^{11} [Pa]	0.30
a	Left-truncated Normal	0.07 [m]	0.10
N	Left-truncated Normal	800000 N	0.20
ρ	Left-truncated Normal	7500 [Kg/m^3]	0.10
Latent variables X			
Variable	**Distribution**	**Mean**	**CV**
k_l	Left-truncated Normal	400000 [N × m]	0.30
k_r	Left-truncated Normal	400000 [N × m]	0.30
Fixed quantities			
Variable	**Value**		
l	14 [m]		

4.4. PCE emulator

In this work we also want to test whether the use of an emulator allows us to obtain an accurate representation of all the uncertainties characterizing the simulator, even simplifying the problem and making Bayesian updating more feasible and practical in structural identification applications. Indeed, Bayesian methods may resort to physics-based or emulator-based models to solve associated inverse problems. These latter can replicate the actual behavior of the structural system with a significant reduction of computational costs making Bayesian updating feasible in structural identification applications. Since emulator-based models are introduced when physics-based models become impractical due to the significant computational costs, there is a noticeable gap in the literature comparing the results of the Bayesian updating procedures obtained with emulator-based models
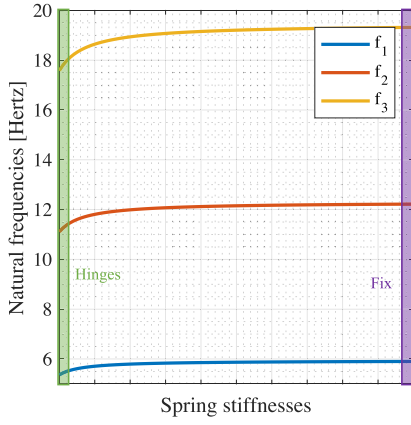
Fig. 4. Natural frequencies variation due to variation in the rotational spring stiffnesses k_l and k_r.

versus physics-based models, particularly with reference to ABC methods. To address this gap, we have selected a case study characterized by reduced computational costs. This choice allows us to compare the two approaches in the same updating procedure. To this end, we investigate the use of a PCE emulator in comparison with the physic-based method described in the previous section, both in the MCMC and ABC approach. While the use of PCE emulators in MCMC procedures is well established in the literature, no contributions can be found regarding the comparison between its use in MCMC and ABC procedures [35–38].

In this section, the PCE emulator mathematical setting is defined and the PCE emulator calibration procedure is discussed. It involves decomposing the response of the simulator $y^M(\xi)$ into orthogonal polynomial terms in order to replace the simulator dynamic response [39–43]. Let us consider the multivariate input ξ characterized by probability distribution $\pi(\xi)$ and the univariate output $y^M(\xi)$. Assuming that the model response has a finite variance, the emulator is represented in terms of the input random variables as:

$$y_{\text{em}}^M(\xi) = \sum_\alpha c_\alpha \cdot \Phi_\alpha(\xi) \tag{4}$$

where $\alpha = \{\alpha_1, \cdot, \alpha_N\}$ is a multi-index denoting the degree of each polynomial term, c_α are the expansion coefficients to be determined and Φ_α are multivariate orthogonal polynomials. The basis functions Φ_α can be expressed as a product of univariate polynomials corresponding to each of the J components of the multivariate input ξ:

$$\Phi_\alpha(\xi) = \phi_{\alpha_1}(\xi_1) \cdot \phi_{\alpha_2}(\xi_2) \cdot \ldots \cdot \phi_{\alpha_J}(\xi_J) \tag{5}$$

The coefficients c_α are determined using the regression technique (least square estimation), ensuring an accurate approximation of the model response. The data training set is built using the physics model. Further details about this topic can be found in [36,37].

In this case study, since no direct information (e.g. measures of the mechanical characteristics) are available for the input model random variables, a set of four statistically independent truncated Gaussian random variables are set as uncertain computational model input parameters (ξ_j, $j = 1, \ldots, 4$). The mean values and the CV of these PDFs are reported in Table 1. Therefore, Hermite polynomials were selected ensuring the orthogonality condition with respect to the probability distribution of each input variable.

The first three natural frequencies y_0 identified from the experimental tests (see Section 5.2) are set as the output of the model. Therefore, the corresponding first three numerical frequencies estimated from the simulator y_M are set as Quantity of Interest (QoIs). The PC expansion in Eq. (5) is used to build a surrogate model for each of the selected QoIs. Each of the three surrogate models is built using a procedure based on (i) selection of the polynomial order (ii) selection of the regression

points using the Gaussian quadrature rule method (iii) evaluation of deterministic coefficients using regression (iv) validation of the surrogate models. This validation is carried out using the twofold procedure discussed in [36,37] in order to ensure the accuracy of the emulator. Polynomial order 5 is found to be the best selection.

In particular, the difference between the response vectors $y_{\text{em}}^M(\xi)$ and $y^M(\xi)$, i.e. the error vector, e, is evaluated on the 46656 sample combinations, used for calibrating the emulator. The mean and the variance of the error vector is estimated for polynomial order 5, obtaining a perfect matching between numerical and surrogate models with the mean and variance of the error vector e lower than 1.8127×10^{-5} and 1.364×10^{-5}, respectively. Furthermore, the accuracy of the emulator is performed also estimating the error vector outside the grid used to calibrate the proxy model, driving the simulation of the parameters in the tail values of the joint PDF of the random variables (ξ_j, $j = 1, \ldots, 4$) obtaining maximum errors equal to 5×10^{-3}, 5×10^{-2} and 4×10^{-2}, for each of the first three natural frequencies.

5. Results

5.1. Real noisy modal data

Using $y_0 = \{4.492, 9.288, 13.252\}$ as observed data, the posterior marginal probability density functions of the four-dimensional random vector ξ are estimated using both MCMC and ABC following the procedures discussed in Section 3, in order to compare the accuracy of the obtained results. The results of the two Bayesian updating procedures are reported in Fig. 5. Convergence of MCMC-MH was assessed using visual inspections on the trace plots (i.e. plots showing the sampled values of parameters over iterations), autocorrelation plots and estimating the Gelman–Rubin diagnostic [44]. A total of $S = 450.000$ runs of the algorithm are required to ensure convergence. ABC is carried out using $S = 10.000$ runs of the algorithm. The proposal distribution is the prior as displayed in Alg 2 and observed and simulated data have been compared using the Euclidean distance. The threshold ϵ has been selected using the method of the αth quantile ($\alpha = 0.01$) [45]. It is worth noting that while the MCMC-MH algorithm necessitates 450.000 runs and 2 h of computing time, the ABC algorithm accomplishes the task with only 10.000 runs, taking just 5 min. This comparison underscores the efficiency and computational advantage of the ABC algorithm in handling large-scale simulations.

The updating framework carried out using the two different algorithms is able to modify the probability distributions of the updating parameters E, ρ, a and N. The obtained posterior mean values and coefficients of variations (CV) are reported in Table 2. Fig. 6 shows the comparison between the posterior mean value and CV obtained using MCMC-MH and ABC algorithm for the updated parameters.

Fig. 6 - (a) shows the Kullback–Leibler divergence [46] between posterior distributions approximated using MCMC and ABC samples (i.e. Case C) for all the updated distributions. It is worth recalling that Kullback–Leibler divergence is used to measure the difference between two probability distributions, with values approaching zero meaning a high level of fidelity. The divergences derived from the probability distributions of all the updated parameters are remarkably close to zero. Notably, only the posterior distributions of N exhibit a divergence exceeding 0.2. Furthermore, Fig. 6 - (b) and (c) illustrate the ratio between the posterior mean and the posterior coefficient of variation (CV) obtained using the two different algorithms. Values equal to 1 indicate that the posterior mean values and the posterior CV are identical. Posterior mean ratio values close to 1 are obtained for all the updated distributions. In contrast, posterior CV ratio values between 0.8 and 1.2 are observed. This suggests that while the posterior mean values are consistent between the two algorithms, there may be slight differences in the estimation of the posterior uncertainty. Overall, these observations suggest that ABC algorithms are capable of

Table 2
Posterior distributions characteristics of the updating parameters.

Parameter	MCMC-MH		ABC		MCMC-MH PCE		ABC PCE	
	Mean	CV	Mean	CV	Mean	CV	Mean	CV
E [Pa]	1.58×10^{11}	0.318	1.59×10^{11}	0.350	1.53×10^{11}	0.349	1.53×10^{11}	0.347
ρ [Kg/m^3]	7520	0.096	7580	0.082	7299	0.086	7540	0.084
a[m]	0.076	0.074	0.075	0.082	0.074	0.087	0.075	0.081
N [N]	605 000	0.168	558 000	0.206	573 000	0.206	564 000	0.194

producing reliable results that closely resemble those obtained via the well-established likelihood-based MCMC algorithms.

The effectiveness of the PCE based emulator within ABC context is assessed by a direct comparison between the results (i.e. Kullback–Leibler divergence, posterior mean value ratio and posterior CV ratio) obtained from ABC algorithm using the true and the emulator solution (i.e. Case B in Fig. 6). For all the updated parameters, Kullback–Leibler divergence values lower than 0.016 are obtained. Additionally, posterior mean ratio values between 0.98 and 1.04 and posterior CV ratio values between 0.97 and 1.06 are observed for all the updated parameters. This assessment highlights the effectiveness of the PCE-based emulator integrated within the ABC framework, as it accurately captures the key characteristics and statistical properties of the true solution, showcasing its reliability and utility within the ABC context.

In order to compare the effect of employing the PCE-based emulator in the ABC and MCMC-MH algorithms, a direct comparison is carried out between the results obtained from the MCMC-MH algorithm using the true solution and the emulator solution (i.e., Case A in Fig. 6), focusing on the same metrics of previous case. For all the updated parameters, Kullback–Leibler divergence values lower than 0.15 are obtained. Additionally, posterior mean ratio values between 1.02 and 1.06 and posterior CV ratio values between 0.81 and 1.11 are observed for all the updated parameters.

It is worth to observe that the Kullback–Leibler divergence values are generally lower in Case B compared to Case A; the posterior mean ratio values in Case B are closer to 1, compared to Case A, indicating a slightly better agreement between the emulator and true solution in Case B. However, the posterior CV ratio values show a wider range in Case B compared to Case A, suggesting that the variability captured by the posterior CVs may vary more in Case B. Overall, these comparisons provide insights into the performance and accuracy of the PCE-based emulator in both ABC and MCMC-MH algorithms. These comparisons suggest two significant conclusions: firstly, the PCE emulator is capable of providing accurate solutions even in the MCMC-MH algorithm case, and secondly, the ABC algorithm appears to be less sensitive to the use of the surrogate model while still maintaining a good approximation ability compared to the true solution.

Finally, the same metrics are utilized to compare the posterior distributions obtained for the updating parameters using the emulator solutions in both the ABC and MCMC-MH algorithms (i.e. Case D in Fig. 6). Kullback–Leibler divergence values lower than 0.12 are observed, indicating a close match between the distributions. Additionally, the posterior mean value ratio values are in between 0.97 and 1.02, suggesting consistency in the mean values of the distributions. The posterior CV ratio values indicate variability in the coefficient of variation of the distributions, with values ranging between 1 and 1.07.

5.2. Simulation study

In this section, the ABC algorithm is used to tackle the structural identification problem associated with tie-rods, aiming to assess the efficacy of the updating procedure by varying the number of natural frequencies employed as observed data y_0. This evaluation is of utmost importance due to the considerable challenges involved in identifying numerous natural frequencies through in-situ experimental activities on tie-rods. In historical masonry structures, tie-rods are often located at considerable heights above the ground, making the use of

contact dynamic measurements impractical. Moreover, the restricted access points inherent in historical masonry structures present obstacles in the installation of dynamic measurement devices, thereby adding complexity to the structural identification process. This indicates that identifying a large number of frequencies is frequently unfeasible using non contact measurement devices. Hence, the reliability of the results obtained using varying numbers of natural frequencies becomes crucial, and it is of utmost importance to assess the reliability of the structural identification of the ancient tie-rod, as discussed in the previous section. For this reason, five simulated natural frequencies y_0 are obtained from the computational model driven by values of the uncertain parameters ξ denoted as true values in Table 3. Zero left truncated Gaussian distributions are selected for all the uncertain parameters whose mean values are reported in Table 3. A different number of natural frequencies nf ranging from 1 to 5 are set as reference dataset for the Bayesian updating procedures. Having assessed the accuracy of the ABC algorithm to solve the structural identification problem, only ABC is implemented for subsequent analyses.

The outcomes of the ABC framework (i.e. posterior distributions of the updating parameters ξ) are reported in Fig. 7. As it was expected, the posterior distributions mean values of the updated parameters ξ move towards the true values (i.e. actual values used as to simulate the reference data set y_0) augmenting the dimension of the reference dataset vector. This further highlights the accuracy of the Bayesian updating framework in providing useful information on the single point solution of the structural identification problem.

The posterior mean value obtained for the five considered cases are reported in Table 3 and compared in terms of ratio (i.e. posterior mean/true value) in Fig. 8-(a). As it was expected, using only the first natural frequency give unreliable single point solutions for the updated parameters, with maximum differences occurring for the E and N parameters. Conversely, reliable results are obtained when number of natural frequencies greater than 2 are used as a reference.

Finally, we computed the Bayesian Mean Absolute Percentage Error (BMAPE) as follows:

$$BMAPE = E_\theta \left[\frac{|\theta - \theta^{\text{true}}|}{\theta^{\text{true}}} \mid y_0 \right] \approx \frac{1}{S} \sum_{s=1}^{S} \frac{|\theta^{(s)} - \theta^{\text{true}}|}{\theta^{\text{true}}}$$

where θ^{true} are true values used to generate data. This quantity evaluates parameter posterior estimates accounting also for posterior uncertainty and preferring more concentrated distributions. Since posterior distributions become less dispersed as the number of frequencies increases, BMPAE slightly decreases moving from 3 to 4 or 5 frequencies, despite posterior means being closer to true values in the case of three observed frequencies. Overall, from these results, it seems that three frequencies are enough to get good point estimates even taking into account the posterior uncertainty.

6. Discussion and conclusion

In this paper, we have developed a robust framework aimed at tackling the identification problem from a statistical perspective, marking a notable advancement by integrating various sources of uncertainty into our model. Through the incorporation of latent variables, in addition to conventional measurement errors, we have achieved a more exhaustive representation of uncertainty, despite the ensuing statistical complexity with an intractable likelihood.
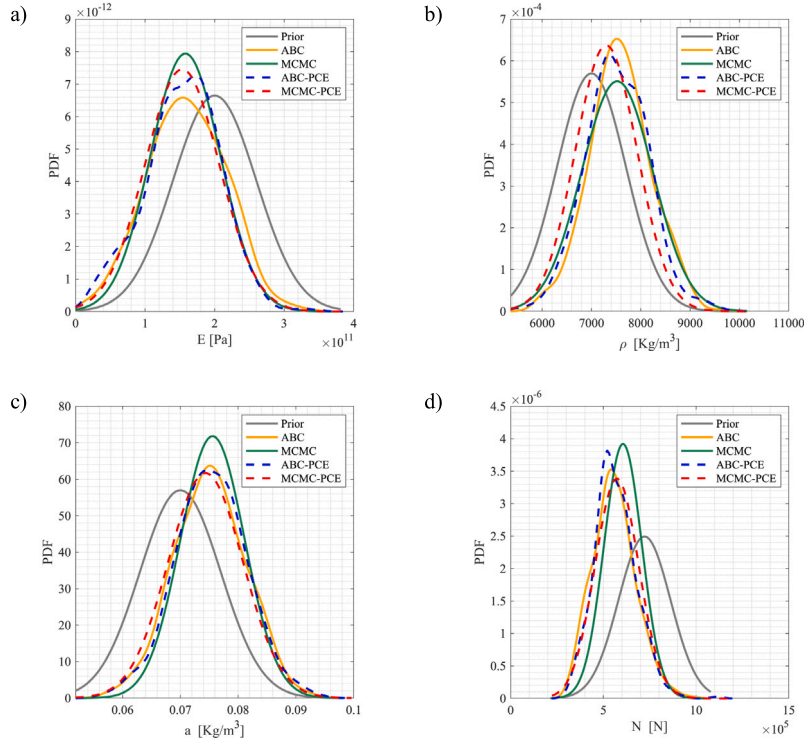
Fig. 5. Posterior PDF of the updated parameters using ABC, MCMC and PCE based ABC and MCMC: (a) Young modulus E, (b) mass density ρ, (c) cross section transversal dimension a and (d) axial load N.
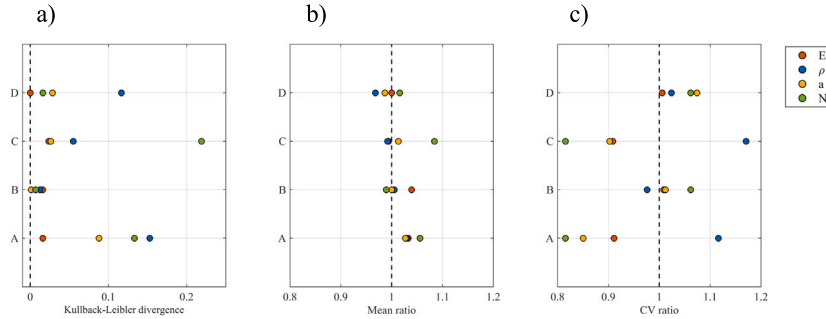

Fig. 6. Kullback–Leibler distances (a) posterior mean value ratio (b) and posterior CV ratio (c). Case A: comparison between the results obtained from MCMC-MH and MCMC-MH PCE; Case B: comparison between the results obtained from ABC and ABC-PCE; Case C: Comparison between the results obtained from MCMC-MH vs ABC; Case D: Comparison between the results obtained from MCMC-MH PCE and ABC PCE.

Table 3
True values, prior distribution mean values, and posterior distribution mean values augmenting the number of frequencies.

	E (N/mm^2)	ρ (kg/m^3)	a (m)	N (N)
True Value	1.8×10^{11}	7500	0.07	700 000
Prior Mean	2.0×10^{11}	7000	0.07	800 000
Posterior Mean (nf = 1)	1.29×10^{11}	7550	0.0736	582 910
Posterior Mean (nf = 2)	1.86×10^{11}	7270	0.0724	713 170
Posterior Mean (nf = 3)	1.75×10^{11}	7363	0.0719	716 660
Posterior Mean (nf = 4)	1.72×10^{11}	7349	0.0716	712 490
Posterior Mean (nf = 5)	1.70×10^{11}	7347	0.0718	714 490

Our study presents a meticulous comparative assessment of two prominent inference methodologies: data augmentation MCMC-MH and ABC, both enhanced with emulator-based approximations, with reference to a simple structural system consisting in an axial load beam like structure.

In particular, an ancient tie-rod was object of an extensive experimental campaign aimed at identifying its modal characteristics in terms of natural frequencies. Three experimental natural frequencies were identified and used as target in the updating framework using both algorithms. A computational and a PCE based emulator model were also set up and used to reproduce the dynamic response.

The efficacy of the ABC algorithm, even in its simplest form, was evaluated through a direct comparison of the posterior distributions of the updating parameters obtained from ABC and the well-established standard MCMC-MH. This comparison was conducted using various metrics capable of assessing probability distribution shape and moments, such as mean and coefficients of variation. Through this comprehensive analysis, we sought to gauge the performance and reliability of ABC in capturing the underlying parameter distributions and to assess its effectiveness relative to the widely used MCMC-MH approach.

The effectiveness of high fidelity surrogate models based on PCE method within the framework of ABC is demonstrated and analyzed through a direct comparison of the posterior distributions of the updating parameters obtained from the ABC algorithm using both the true model solutions and the emulator model solutions. Furthermore, the comparison between the posterior distributions of the updating
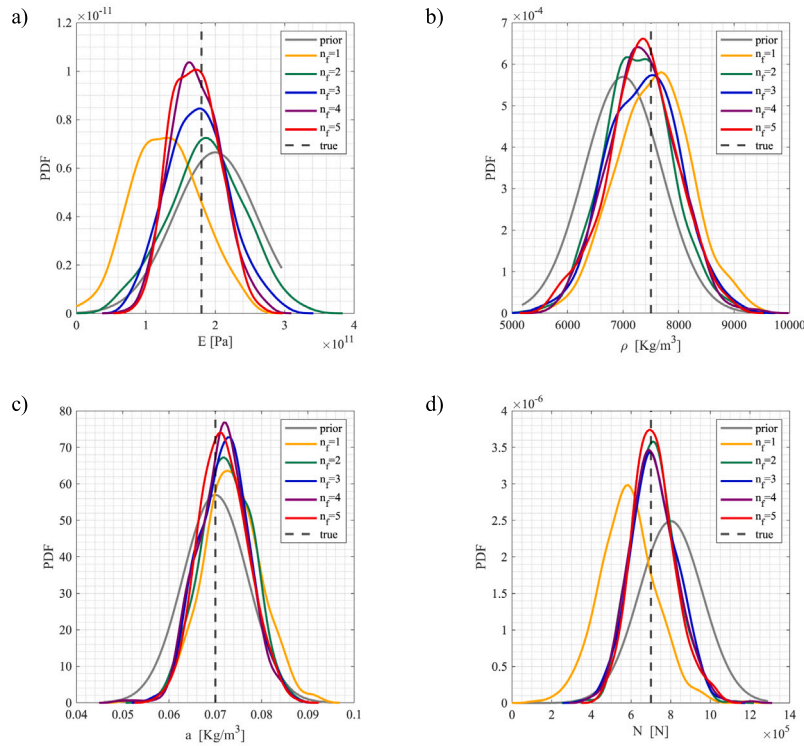
Fig. 7. Posterior PDF of the updated parameters using ABC and different number of natural frequencies y_0, $i = 1,..,5$: (a) Young modulus E, (b) mass density ρ, (c) cross section transversal dimension a and (d) axial load N.
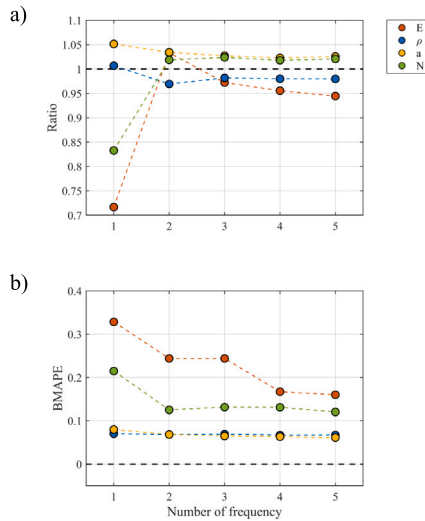

Fig. 8. Ratio between the ABC posterior mean of the updated parameters using different numbers of natural frequencies, from 1 to 5, and the true value (a); BMAPE values (b).

parameters obtained from the MCMC-MH algorithm using the true model solutions and the emulator model solutions revealed that the ABC algorithm exhibited less sensitivity to the utilization of emulators compared to MCMC-MH. This observation underscores the robustness of ABC in accommodating model approximations, thereby enhancing its applicability in scenarios where computational efficiency or model complexity necessitates the use of surrogate models.

Finally, our investigation into frequency estimation sheds light on the practicality of our proposed approach, revealing that merely three frequencies are sufficient for yielding reliable posterior inference.

In summary, our contribution lies in advancing statistical methodologies tailored for addressing intricate identification problems, emphasizing the importance of comprehensive uncertainty modeling and the exploration of alternative inference strategies. Note that the proposed statistical framework is flexible enough to tackle the problem regardless of the defined physics and emulator models. However, note that defining a complex physical model will necessarily require the use of the emulator-based approximation to make the procedure computationally feasible. Furthermore, the implementation of the MCMC procedure would become infeasible in cases where the number of components to be identified is too high. In such cases, since the algorithm must explore a high dimensional parameter space, the resulting Markov chain would suffer from auto-correlation problems. Moreover, the procedure can be applied to other structural dynamic characteristics, meaning that the type of observed data y_0 changes. This may require a different definition of the probabilistic assumptions about the error and bias components of the model. However, once the appropriate probability distributions for these quantities are specified, proposed procedures are still applicable. Moreover, ABC algorithms are also well-suited for addressing cases where multiple types of data are used simultaneously. In such a case, it would be necessary to summarize observed data through proper summary statistics and adapt the ABC distance function to compare all simulated and observed quantities simultaneously.

CRediT authorship contribution statement

Silvia Monchetti: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chiara Pepi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cecilia Viscardi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Massimiliano Gioffrè:** Writing – review & editing, Software, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

[1] Filippo Ubertini, Nicola Cavalagli, Alban Kita, Gabriele Comanducci, Assessment of a monumental masonry bell-tower after 2016 Central Italy seismic sequence by long-term SHM, Bull. Earthq. Eng. 16 (2018) 775–801.

[2] Antonella Saisi, Carmelo Gentile, Antonello Ruccolo, Continuous monitoring of a challenging heritage tower in Monza, Italy, J. Civ. Struct. Health Monit. 8 (2017) 77–90.

[3] Giacomo Zini, Michele Betti, Gianni Bartoli, A pilot project for the long-term structural health monitoring of historic city gates, J. Civ. Struct. Health Monit. (2023).

[4] Silvia Briccoli Bati, Ugo Tonietti, Experimental methods for estimating in situ tensile force in tie-rods, J. Eng. Mech. 127 (12) (2001) 1275–1283.

[5] E. Cescatti, F. Da Porto, C. Modena, Axial force estimation in historical metal tie-rods: methods, influencing parameters, and laboratory tests, Int. J. Archit. Herit. 13 (3) (2019) 317–328.

[6] Nirvan Makoond, Luca Pelà, Climent Molins, Robust estimation of axial loads sustained by tie-rods in historical structures using artificial neural networks, Struct. Health Monit. (2022) 14759217221123326.

[7] Francescantonio Luca, Stefano Manzoni, Alfredo Cigada, L Frate, A vibration-based approach for health monitoring of tie-rods under uncertain environmental conditions, Mech. Syst. Signal Process. 167 (2022) 108547.

[8] Domenico Camassa, Anna Castellano, Aguinaldo Fraddosio, Giuseppe Miglionico, Mario Daniele Piccioni, Dynamic identification of tensile force in tie-rods by interferometric radar measurements, Appl. Sci. 11 (8) (2021) 3687.

[9] S Campagnari, F Di Matteo, Stefano Manzoni, Matteo Scaccabarozzi, M Vanali, Estimation of axial load in tie-rods using experimental and operational modal analysis, J. Vib. Acoust. 139 (4) (2017).

[10] Ivan Duvnjak, Suzana Ereiz, Domagoj Damjanović, Marko Bartolac, Determination of axial force in tie-rods of historical buildings using the model-updating technique, Appl. Sci. 10 (17) (2020) 6036.

[11] Carmelo Gentile, Carlo Poggi, Antonello Ruccolo, Mira Vasic, Vibration-based assessment of the tensile force in the tie-rods of the Milan Cathedral, Int. J. Archit. Herit. 13 (3) (2019) 411–424.

[12] Giovanni Rebecchi, Nerio Tullini, Ferdinando Laudiero, Estimate of the axial force in slender beams with unknown boundary conditions using one flexural mode shape, J. Sound Vib. 332 (18) (2013) 4122–4135.

[13] Stefano Sorace, Parameter models for estimating in-situ tensile force in tie-rods, J. Eng. Mech. 122 (9) (1996) 818–825.

[14] James L. Beck, Siu-Kui Au, Michael K. Vanik, Monitoring structural health using a probabilistic measure, Comput.-Aided Civ. Infrastruct. Eng. 16 (2001) 1–11.

[15] Roberto Rocchetta, Matteo Broggi, Quentin Huchet, Edoardo Patelli, On-line Bayesian model updating for structural health monitoring, Mech. Syst. Signal Process. (2018) 174–195.

[16] Chiara Pepi, Massimiliano Gioffrè, Mircea Grigoriu, Bayesian inference for parameters estimation using experimental data, Probab. Eng. Mech. 60 (2020) 103025.

[17] Yong Huang, Changsong Shao, Biao Wu, James L. Beck, Hui Li, State-of-the-art review on Bayesian inference in structural system identification and damage assessment, Adv. Struct. Eng. 22 (6) (2018) 1329–1351.

[18] Silvia Monchetti, Cecilia Viscardi, Michele Betti, Francesco Clementi, Comparison between Bayesian updating and approximate Bayesian computation for model identification of masonry towers through dynamic data, Bull. Earthq. Eng. (2023).

[19] Silvia Monchetti, Cecilia Viscardi, Michele Betti, Gianni Bartoli, Bayesian-based model updating using natural frequency data for historic masonry towers, Probab. Eng. Mech. (2022).

[20] Anna De Falco, Carlo Resta, Giacomo Sevieri, Sensitivity analysis of frequency-based tie-rod axial load evaluation methods, Eng. Struct. 229 (2021) 111568.

[21] Andrew Gelman, John B Carlin, Hal S Stern, Donald B Rubin, Bayesian Data Analysis, Chapman and Hall/CRC, 1995.

[22] James O. Berger, Leonard A. Smith, On the statistical formalism of uncertainty quantification, Annu. Rev. Statist. Appl. 6 (2019) 433–460.

[23] W. Keith Hastings, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, Oxford University Press, 1970.

[24] Christian P. Robert, George Casella, George Casella, Monte Carlo Statistical Methods, vol. 2, Springer, 1999.

[25] Donald B. Rubin, Bayesianly justifiable and relevant frequency calculations for the applied statistician, Ann. Statist. (1984) 1151–1172.

[26] Simon Tavaré, David J. Balding, Robert C. Griffiths, Peter Donnelly, Inferring coalescence times from DNA sequence data, Genetics 145 (2) (1997) 505–518.

[27] Jonathan K Pritchard, Mark T Seielstad, Anna Perez-Lezaun, Marcus W Feldman, Population growth of human Y chromosomes: a study of Y chromosome microsatellites, Mol. Biol. Evol. 16 (12) (1999) 1791–1798.

[28] Scott A. Sisson, Yanan Fan, Mark Beaumont, Handbook of Approximate Bayesian Computation, CRC Press, 2018.

[29] M. Gioffré, N. Cavalagli, C. Pepi, M. Trequattrini, Laser doppler and radar interferometer for contactless measurements on unaccessible tie-rods on monumental buildings: Santa Maria della Consolazione Temple in Todi, J. Phys. Conf. Ser. 778 (1) (2017) 012008.

[30] Umberto Nofrini, Arnaldo Bruschi, Il tempio del Bramante a Todi, 1970, Res Tudertinae 11, Comune di Todi.

[31] L. Battistoni, Il tempio di Santa Maria della Consolazione a Todi 1508-1670, 2012, ETAB-Associazione Pro Todi.

[32] M. Gioffré, N. Cavalagli, C. Pepi, M. Trequattrini, Laser doppler and radar interferometer for contactless measurements on unaccessible tie-rods on monumental buildings: Santa Maria della Consolazione Temple in Todi, J. Phys. Conf. Ser. 778 (1) (2017) 012008.

[33] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, Stefano Tarantola, Global Sensitivity Analysis: the Primer, John Wiley & Sons, 2008.

[34] Chiara Pepi, Massimiliano Gioffrè, Bayesian updating for random tensile force identification of ancient tie rods using modal data, Appl. Sci. 14 (9) (2024) 3698.

[35] C. Pepi, M. Gioffre, M.D. Grigoriu, H.G. Matthies, Bayesian updating of cable stayed footbridge model parameters using dynamic measurements, in: Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2019, 2019, pp. 330–342, Cited by: 8.

[36] Chiara Pepi, Massimiliano Gioffre', Mircea D. Grigoriu, Parameters identification of cable stayed footbridges using Bayesian inference, Meccanica 54 (2019) 1403–1419.

[37] Chiara Pepi, Massimiliano Gioffre, Mircea Grigoriu, Bayesian inference for parameters estimation using experimental data, Probab. Eng. Mech. 60 (2020) 103025.

[38] Bojana V. Rosić, Anna Kučerová, Jan Sýkora, Oliver Pajonk, Alexander Litvinenko, Hermann G. Matthies, Parameter identification in a probabilistic setting, Eng. Struct. 50 (2013) 179–196, Engineering Structures: Modelling and Computations (special issue IASS-IACM 2012).

[39] Roger G. Ghanem, Pol D. Spanos, Stochastic Finite Elements: A Spectral Approach, Courier Corporation, 2003.

[40] Olivier Le Maître, Omar M. Knio, Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics, Springer Science & Business Media, 2010.

[41] John D Jakeman, Fabian Franzelin, Akil Narayan, Michael Eldred, Dirk Plfüger, Polynomial chaos expansions for dependent random variables, Comput. Methods Appl. Mech. Engrg. 351 (2019) 643–666.

[42] Nora Lüthen, Stefano Marelli, Bruno Sudret, A spectral surrogate model for stochastic simulators computed from trajectory samples, Comput. Methods Appl. Mech. Engrg. 406 (2023) 115851.

[43] Xujia Zhu, Bruno Sudret, Stochastic polynomial chaos expansions to emulate stochastic simulators, Int. J. Uncertain. Quantif. 13 (2) (2023).

[44] Andrew Gelman, Donald B. Rubin, Inference from iterative simulation using multiple sequences, Stat. Sci. 7 (4) (1992) 457–472.

[45] Mark A. Beaumont, Wenyang Zhang, David J. Balding, Approximate Bayesian computation in population genetics, Genetics 162 (4) (2002) 2025–2035.

[46] Solomon Kullback, Richard A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1) (1951) 79–86.