An efficient likelihood-free Bayesian computation for model selection and parameter estimation applied to structural dynamics

A. Ben Abdessalem[1], **N. Dervilis**[2], **D. Wagg**[2], **K. Worden**[2]

[1] LARIS, EA, 7315, Angers University, 62 Avenue Notre Dame du Lac, 49000 Angers, France

[2] Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield Mappin Street, Sheffield S1 3JD, United Kingdom

ABSTRACT

Model selection is a challenging problem that is of importance in many branches of the sciences and engineering, particularly in structural dynamics. By definition, it is intended to select the most plausible model among a set of competing models, that best matches the dynamic behaviour of a real structure and better predicts the measured data. The Bayesian approach is based essentially on the evaluation of a likelihood function and is arguably the most popular approach. However, in some circumstances, the likelihood function is intractable or not available even in a closed form. To overcome this issue, likelihood-free or approximate Bayesian computation (ABC) algorithms have been introduced in the literature, which relax the need of an explicit likelihood function to measure the degree of similarity between model prediction and measurements. One major issue with the ABC algorithms in general is the low acceptance rate which is actually a common problem with the traditional Bayesian methods. To overcome this shortcoming and alleviate the computational burden, a new variant of the ABC algorithm based on an ellipsoidal nested sampling technique is introduced in this paper. It has been called ABC-NS. This paper will demonstrate how the new algorithm promises drastic speedups and provides good estimates of the unknown parameters. To demonstrate its practical applicability, two illustrative examples are considered. Firstly, the efficiency of the novel algorithm to deal with parameter estimation is demonstrated using a moving average process based on synthetic measurements. Secondly, a real structure called the VTT benchmark, which consists of a wire rope isolators mounted between a load mass and a base mass, is used to further assess the performance of the algorithm in solving the model selection issue.

Keywords: Model selection, structural dynamics, likelihood-free Bayesian computation, moving average process, wire rope isolators

1 INTRODUCTION

Many branches of the sciences and engineering, including structural dynamics involve a choice between a set of models. This is quite challenging, mainly when a large set of competing models with different numbers of parameters are available. In practice, one may assume a set of plausible models and then estimate the evidence associated to each one using a sophisticated statistical tool. The Bayesian approach is one of the most popular techniques to deal with model selection and parameter estimation issues. Arguably, the most popular algorithm in this class of methods is the one proposed by Green and called Reversible-Jump Markov chain Monte Carlo (RJ-MCMC), readers can go for more details about the RJ-MCMC in [1] and for an extensive review of the different methods to [2]. The Bayesian method requires the definition of a likelihood function; however, in some circumstances, the likelihood function is computationally intractable or not available. To overcome this shortcoming, a class of methods called likelihood-free or approximate Bayesian computation (ABC) algorithms have been proposed in the literature. ABC has gained popularity in recent years owing to its easy implementation, flexibility and good performance. It has been applied in a wide range of applications and recently in structural dynamics [3][4][5][6]. The most popular variants are the ABC-MCMC [7] and ABC-SMC [8]. Those algorithms can be applied when it is possible to generate data from the model, offering the possibility to make Bayesian inference relying only on an appropriate choices of summary statistics. Despite their numerous advantages and popularity, the ABC algorithms suffer from a difficulty in maintaining a reasonable acceptance rate over the populations. In the framework of this paper, a novel ABC algorithm based on an efficient ellipsoidal sampling method [9][10] is proposed to enhance the capability of the ABC algorithm to deal with parameter estimation and model selection issues. To demonstrate the efficiency of the algorithm, two illustrative examples in structural dynamics have been proposed. The proposed examples will show that the parsimony principle (i.e., simpler models are automatically preferred unless a more complicated model provides a significantly better fit to the data) is naturally embedded in the ABC-NS algo-

rithm. Moreover, by employing a nested sampling technique, the ABC algorithm can be substantially accelerated, offering the possibility to consider a relatively large number of competing models.

The paper is organised as follows. Section 2 will shortly go through the basics of the Bayesian paradigm for parameter estimation and model selection and will introduce the novel algorithm. Sections 3 and 4 are devoted to numerical and experimental examples and forms the core of the paper. Section 5 concludes the paper.

2 BAYESIAN INFERENCE

2.1 Bayesian method for parameter estimation

In this section, a brief introduction to the Bayesian and approximate Bayesian methods is given. In the Bayesian method, the posterior probability density, $p(\theta|u)$ given observed data u and a model $\mathcal{M}$, can be computed using Bayes' Theorem:

$$p(\theta|u) = \frac{p(\theta)\mathcal{L}(u|\theta)}{\int_\theta p(\theta)\mathcal{L}(u|\theta)\mathrm{d}\theta} \propto p(\theta)\mathcal{L}(u|\theta) \tag{1}$$

where $p(\theta)$ is the prior probability of θ and $\mathcal{L}(u|\theta)$ is the likelihood function. The denominator is a normalising constant.

However, as mentioned earlier, explicit forms for likelihood functions are rarely available. The ABC methods approximate the likelihood by evaluating the discrepancy between the observed data and the data generated by a simulation using a given model, yielding an approximate form of Bayes' Theorem:

$$p(\theta|\Delta(u,u^*) < \varepsilon) \propto p(\theta)p\left(\Delta(u,u^*) < \varepsilon|\theta\right) \tag{2}$$

where $u^* \sim f(\cdot|\theta)$ are the simulated data, $\Delta(\cdot)$ is a discrepancy metric, and $\varepsilon > 0$ is a tolerance threshold (when ε tends towards 0, the approximated posterior distribution is a good approximation of the true posterior distribution).

2.2 Bayesian method for model selection

In model selection, usually two or more competing models exist and may support the data u. Assume that k candidate models, $\mathcal{M}_j, j = 1, \ldots, k$, are under consideration, each with associated unknown parameters, θ_j. Here one would like to quantify the relative support from the data for each candidate model. Bayes' Theorem can be applied to compute the posterior probability associated to each candidate model:

$$p(\mathcal{M}_j|u) = \frac{p(u|\mathcal{M}_j)p(\mathcal{M}_j)}{\sum_r p(u|\mathcal{M}_r)p(\mathcal{M}_r)} \tag{3}$$

where $p(u|\mathcal{M}_j)$ is the posterior probability of $\mathcal{M}_j$ and $p(\mathcal{M}_j)$ is the prior probability of $\mathcal{M}_j$, in most cases, one assigns equal prior probabilities to each candidate model.

2.3 ABC-NS implementation

In this section, a detailed description of the novel ABC algorithm is given for parameter estimation. The ABC-NS algorithm broadly works following the same scheme as the ABC-SMC algorithm in [8]. The main novelties are in (i) the way of sampling, (ii) the weighting technique adopted from [11] and (iii) instead of dropping one particle per iteration, a proportion of particles is dropped based on the assigned weights, which speeds-up the algorithm without compromising the precision on the posterior estimates. The iterative process for parameter estimation is detailed in Algorithm 1. The algorithm starts by generating N particles from the prior satisfying the constraint $\Delta(u,u^*) < \varepsilon_1$ (here, u for observed data, u^* for simulated data and ε_1 is the first tolerance value defined by the user). The accepted particles are then weighted (see, Step 9) and the next tolerance threshold is defined based on the discrepancy values ranked in descending order (highest on top, see, step 11) as the $(\alpha_0 N)^{\text{th}}$ value where α_0 is a proportion of dropped particles defined by the user. Then, one assigns a weight of zero to the dropped particles. After that, the weights of the remaining particles are normalised. From the remaining particles, one selects $\beta_0 N$ particles based on the updated weight values, where β_0 is a proportion of particles the so-called "alive" particles. The alive particles are then enclosed in an ellipsoid in which the mass center μ_1 and covariance matrix $\mathcal{C}_1$ are estimated based on the values of those particles; one denotes this ellipsoid by $\mathcal{E}_1 = (\mu_1, \mathcal{C}_1)$. The generated ellipsoid could be enlarged by a factor f_0 to ensure that the particles on the borders are inside. It should be noted that ellipsoidal sampling was firstly proposed in [12] to improve the efficiency of the nested sampling algorithm which has been widely used for Bayesian inference, mainly in cosmology [13]. Finally, the population is replenished by resampling $(1-\beta_0)N$ particles inside the enlarged ellipsoid and a re-weighting step is carried out. The procedure is repeated until a stopping criterion defined by the user is met. It should be noted that for model selection, the same scheme shown in Algorithm 1 is followed by considering the candidate models as additional parameters. Thus, one treats the pair $(\mathcal{M}_j, \theta^{(j)})$ with $\mathcal{M}_j$ as a candidate model and $\theta^{(j)}$ its vector of unknown parameters. For a given $(\mathcal{M}_j, \theta^{(j)})$, the pair is accepted or rejected based on a discrepancy value. At the end of the algorithm, the model probability for $\mathcal{M}_j$ is approximated using Eq. (4).

$$p(\mathcal{M}_j|u^*) \approx \frac{\text{Accepted particles for } \mathcal{M}_j}{\text{Total number of particles } N} \tag{4}$$

In the considered examples, the tuning parameters used to run the ABC-NS algorithm are selected as follows: the number of samples is set to 1000, α_0, β_0 and f_0 are set to 0.3, 0.6 and 1.1, respectively.

3 MOVING AVERAGE PROCESS

In this section, one demonstrates the efficiency of the ABC-NS algorithm to deal with parameter estimation (or model calibration). One considers a moving average (MA) process defined

Algorithm 1 ABC-NS SAMPLER

Require: u: observed data, $\mathcal{M}(\cdot)$: model, ε_1, N, α_0, β_0, f_0

set $t = 1$
for $i = 1, \ldots, N$ **do**
  repeat
    Sample θ^* from the prior distributions $p(\cdot)$
    Simulate u^* using the model $\mathcal{M}(\cdot)$
  until $\Delta(u, u^*) < \varepsilon_1$
  set $\Theta_i = \theta^*$, $e_i = \Delta(u, u^*)$
end for
Associate a weight to each particle: $\omega_i \propto \frac{1}{\varepsilon_1}\left(1 - (\frac{e_i}{\varepsilon_1})^2\right)$
Sort e_i in descending order and store them in e^t.
Define the next tolerance threshold $\varepsilon_2 = e^t(\alpha_0 N)$
Drop particles with $\Delta(u, u^*) \geq \varepsilon_2$, $\omega_{j=1:\alpha_0 N} = 0$
Normalise the weights such that $\sum_{i=1}^{(1-\alpha_0)N} \omega_i = 1$
Select $\mathcal{A}^t = \beta_0 N$ particles from the remaining based on the weights
Define the ellipsoid by its centre of the mass and covariance matrix $\mathcal{E}_t = \{\mu_t, \mathcal{C}_t\}$
Enlarge the ellispoid by f_0 ▷ For simplicity the same notation for the updated ellipsoid is kept
for $t = 2, \ldots, T$ **do**
  for $j = 1, \ldots, (1 - \beta_0)N$ **do**
  repeat
    Sample one particle θ^* inside $\mathcal{E}_{t-1}$
    Simulate u^* using the model $\mathcal{M}(\cdot)$
  until $\Delta(u, u^*) < \varepsilon_t$
  set $\Theta_j = \theta^*$, $e_j = \Delta(u, u^*)$
  end for
    Store the new particles in $\mathcal{S}_t$
    Obtain the new particle set, $\mathcal{N}_{new} = [\mathcal{A}_{t-1}; \mathcal{S}_t]$ with their correspondent distance values e^t
    Sort e^t and define $\varepsilon_{t+1} = e^t(\alpha_0 N)$
    Associate a weight to each particle as in step (9)
    Define the new set of selected particles $\mathcal{A}^t$ as in step (14)
    Update the ellipsoid hyperparameters using $\mathcal{A}^t$, $\mathcal{E}_t = \{\mu_t, \mathcal{C}_t\}$ ▷ The enlargement factor is kept constant
end for

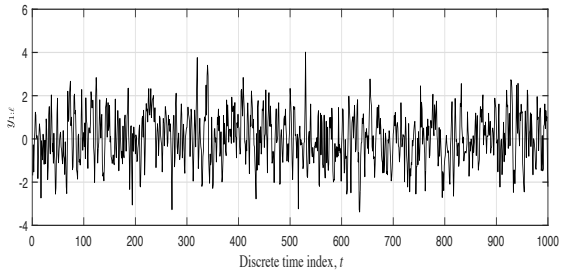


Figure 1: Simulated data from moving average model MA(2) with $(\theta_1, \theta_2) = (0.6, 0.2)$.

by:

$$y_\ell = e_\ell + \sum_{i=1}^{d} \theta_i e_{l-i} \tag{5}$$

where e_ℓ is an independently and identically distributed (iid) sequence of a standard Gaussian $\mathcal{N}(0, 1)$; d and ℓ are set to 2 and 1000, respectively.

Denote by $y = [y_1, \ldots, y_l, \ldots, y_\ell]$ the observed data. The objective is to illustrate the ability of the algorithm to sample from the ABC posterior relying on available data and an appropriate choice of summary statistics. Fig. 1 shows a series of data from a moving average model of order 2 denoted by MA(2). In this example, the quadratic distance between the $d = 2$ first autocovariances is taken as a metric to measure the degree of similarity between simulated and observed data, as in [14]:

$$\Delta(\eta(x), \eta(y)) = \sum_{q=1}^{d} (\tau_{y,q} - \tau_{x,q})^2 \tag{6}$$

where the terms $\tau_{y,q}$ and $\tau_{x,q}$ are the autocovariances of y and x (x is the simulated data given a candidate (θ_1, θ_2)), respectively, which are used as summary statistics. They are obtained as $\tau_{y,q} = \sum_{k=q+1}^{\ell} y_k y_{k-q}$ and $\tau_{x,q} = \sum_{k=q+1}^{\ell} x_k x_{k-q}$, respectively.

To avoid unnecessary difficulties, a standard identifiability condition is imposed on this model (see, [14] for further details). In this case of $d = 2$, this condition is fulfilled when the explored space is delimited by imposing the following constraints on the input parameters:

$$-2 < \theta_1 < 2; \quad \theta_1 + \theta_2 > -1; \quad \theta_1 - \theta_2 < -1 \tag{7}$$

One uses synthetic data for y by generating it from Eq. (5) considering $\theta_{true} = (0.6, 0.2)$. The numerical simulation can now be performed using the ABC-NS algorithm with the hyperparameters given in Section 2.3. From Fig. 2, one can see the evolution of the sample distribution over some selected populations for different tolerance threshold values (from left to right, the tolerance values are equal to $\varepsilon_{up-l} = 6.043 \times 10^5$, $\varepsilon_{up-r} = 6.544 \times 10^4$, $\varepsilon_{down-l} = 1.002 \times 10^3$, $\varepsilon_{down-r} = 0.178$) enclosed in an ellipsoid. The histograms of the particles at

the last population are shown in Fig. 3 from where one can see that those histograms are well peaked around the true values. Table 1 shows the statistics of the posterior estimates, one can see that the posterior estimates are well estimated from the last population with reduced uncertainty.

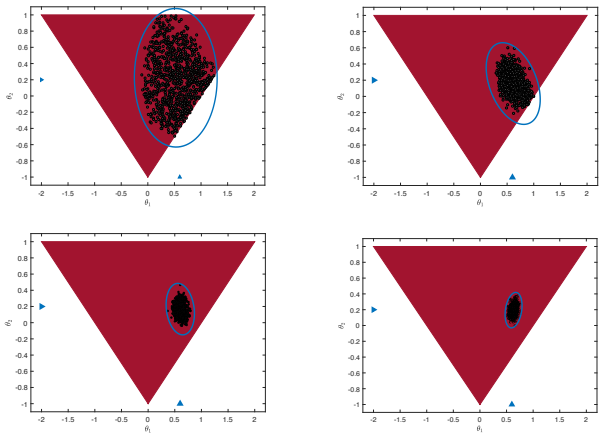


Figure 2: Evolution of the particles distribution over some selected populations, the triangle in red delimits the input space, the blue triangles are the true values.

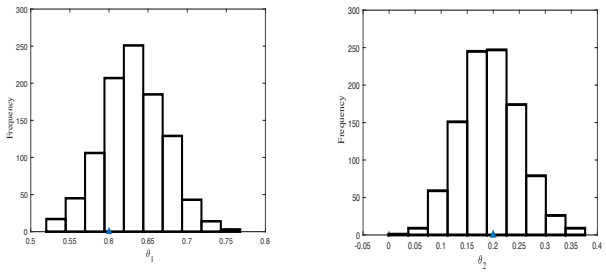


Figure 3: Histograms of the MA(2) model parameters.

4 EXAMPLE 2: CHARACTERISATION OF THE DYNAMICS OF A WIRE ROPE ISOLATOR USING ABCNS

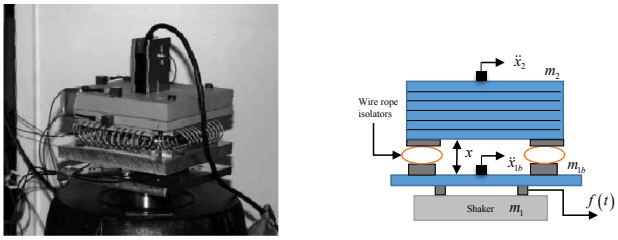
4.1 Experimental set-up

The second example consists of characterising the dynamics of a wire rope isolator (WRI) used for vibration isolation. WRIs have found a vast number of application in medical equipment and military hardware to mention just a few, due to their superior performance for the isolation of impact and vibration. However, the dynamical properties of mechanical isolators are typically nonlinear and these characteristics are seldom well defined, which may cause problems for the design calculations

Parameter	Mean value	Std. dev.	Quantiles	
			2.5%	97.5%
θ_1	0.6318	0.04044	0.5522	0.7115
θ_2	0.1942	0.0569	0.0871	0.3108

TABLE 1: Statistics of the posterior estimates for the MA(2) model.

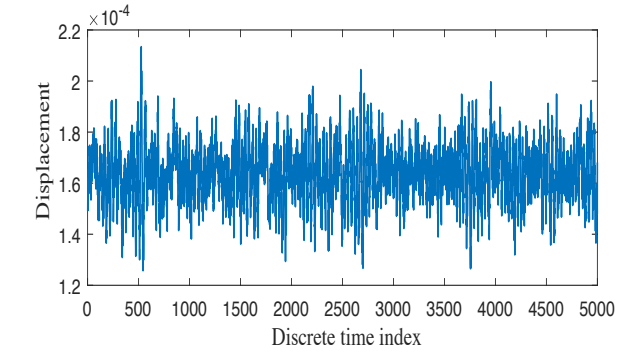
and computer simulations. The system considered in this paper has been proposed within the framework of the European COST Action F3 working group in "Identification of non-linear systems" [15]. The aim of this benchmark was to identify the dynamic properties of resilient mounts used for vibration isolation in industrial applications using different methods. Fig. 4a shows the experimental set-up of the WRI mounted between a load mass m_2 and a base mass m_{1b} while Fig. 4b is a schematic illustration. The applied excitation is produced by an electro-dynamic shaker and corresponds to a white noise sequence, low-pass filtered at 400 Hz. The motion and forces experienced by the isolators are measured; in particular, the acceleration responses $\ddot{x}_2$ and $\ddot{x}_{1b}$ of the load mass and bottom plate, the applied f and the relative displacement x_{12} between the top and bottom plates. For more details concerning the experimental set-up and the methods presented for the identification of the system, the reader is referred to the following references [16] [17] [18] [19].



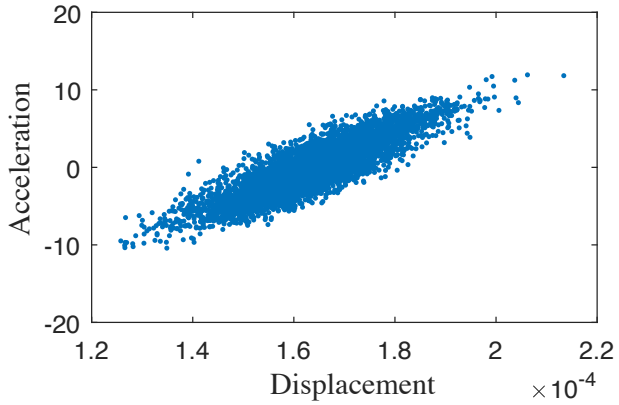
(a) (b)

Figure 4: (a) Experimental set-up: configuration of the experiment, (b) schematic illustration of the dynamical system under consideration.

In this Benchmark, 5 excitation amplitude levels have been considered, varying from 0.5 to 8 Volts. In the present work, one is interested in the experimental data obtained for an excitation amplitude equal to 4 Volts. Fig. 5a shows the relative displacement while Fig. 5b shows the inertial force versus the relative displacement. From Fig. 5b, one may observe a hysteretic behaviour of the WRI. This leads one to select the popular Bouc-Wen model as a competing model. The model has been used previously in similar problems in [20] [21].



(a)



(b)

Figure 5: (a) Displacement, (b) inertial force versus relative displacement under an excitation amplitude of 4 Volts.

4.2 Selection of the competing models

The general single-degree-of-freedom (SDOF) hysteretic system described in the terms of Wen [22], is represented below:

$$m\ddot{y} + g(y, \dot{y}) + z(y, \dot{y}) = f(t) \tag{8}$$

where $g(y, \dot{y})$ is the polynomial part of the restoring force, $z(y, \dot{y})$ is the hysteretic part and $f(t)$ is the excitation force, m is the mass, and the polynomial part of the restoring force is assumed to be linear given by the following equation:

$$g(y, \dot{y}) = c\dot{y} + ky \tag{9}$$

The hysteretic component is defined by Wen [22] via the additional equation of motion:

$$\dot{z} = \begin{cases} -\alpha|\dot{y}|z^n - \beta\dot{y}|z^n| + A\dot{y}, & \text{for } n \text{ odd} \\ -\alpha|\dot{y}|z^{n-1}|z| - \beta\dot{y}|z^n| + A\dot{y}, & \text{for } n \text{ even} \end{cases} \tag{10}$$

The parameters α, β and n govern the shape and the smoothness of the hysteresis loop. It should be noted that the equations offer a simplification from the point of view of parameter estimation, in that the stiffness term in Eq. (9) can be combined with the $A\dot{y}$ term in the state equation for z. The reader can refer to [23] for full details.

A set of competing models have been proposed to capture the dynamics of wire rope isolators. First, although the problem is nonlinear, a linear model is considered at the aim to analyse the behaviour of the algorithm. Then, four models based on the Bouc-Wen model have been defined by varying n in the equations of motion from 1 to 4. In total, five competing models are considered and denoted by:

$$\mathcal{M}_1 \quad : \quad m\ddot{y} + c\dot{y} + Ay = f(t) \tag{11}$$

$$\mathcal{M}_{2:5} \quad : \quad \text{Eqs. (8) - (10), } n = 1:4 \tag{12}$$

After selecting a set of competing models, one aims to determine the most likely model among the competing ones by estimating the evidence associated to each one based on the available data. The data set contains 1000 samples representing a short recording period of the acceleration of the top plate. The data set is split into training data (from 501 to 1000) and testing data (from 1001 to 1500) as shown in Fig. 6. It should be noted that the transient part (from 0 to 500) has been ignored to reduce the effect of initial conditions.

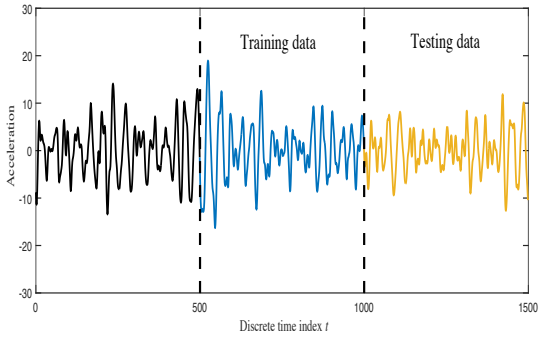


Figure 6: Training and testing data sets using an excitation amplitude of 4 Volts.

For the implementation of the ABC-NS algorithm, the same set of hyperparameters used previously is kept. Here, one considers the case where there is no prior information preferring some models over the other, so the prior is $p(\mathcal{M}_{i=1:5}) = \frac{1}{5}$. In this example, the tolerance threshold sequence is adaptively defined as mentioned in Algorithm 1. The normalised mean square error (MSE) given by Eq. (13) is selected as a metric to measure the discrepancy between the observed and simulated data. Finally, the stopping criterion chosen here is when the difference between two consecutive tolerance thresholds is less than 10^{-5}.

$$\Delta(u^*, u) = \frac{100}{n\sigma_{u^*}^2} \sum_{i=1}^{n} (u_i^* - u_i)^2 \tag{13}$$

where n is the size of the training data, $\sigma_{u^*}^2$ is the variance

of the observed acceleration; u^* and u are the observed and simulated accelerations given by the model, respectively.

4.3 Results and discussion

Fig. 7 shows the model posterior probabilities over some selected populations. One can see that the algorithm converges to $\mathcal{M}_2$. From the same figure, one can see that at higher tolerance values, the algorithm tends to select the linear model (see, the posterior probabilities at populations 7, 13 and 22 for example). This tendency shows that the parsimony principle is naturally embedded in the ABC-NS algorithm. Although, the system is nonlinear, the algorithm tries first to favour simpler models in an automatic way, while in the classical methods for model selection based on the estimation of an information criterion, a penalty term is defined to enforce parsimony. The ABC-NS algorithm switches to a more complex model only when those models are more accurate. Fig. 8 shows the histograms of the selected model parameters obtained at the last population as the mean values shown with red triangles.

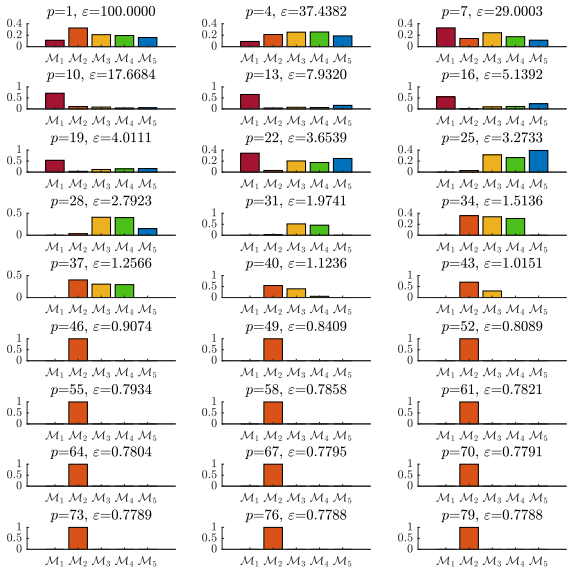


Figure 7: Model posterior probabilities over some selected populations using an excitation amplitude of 4 Volts.

Using the selected model with the posterior estimates from the last population, one can make predictions. Fig. 9 shows the model predictions on the training and testing data sets. As one can see, the predicted results match satisfactorily with the experimental results, which means that the selected model captures perfectly well the dynamics of the WRI. The normalised MSE values on the training and testing data sets are equal to

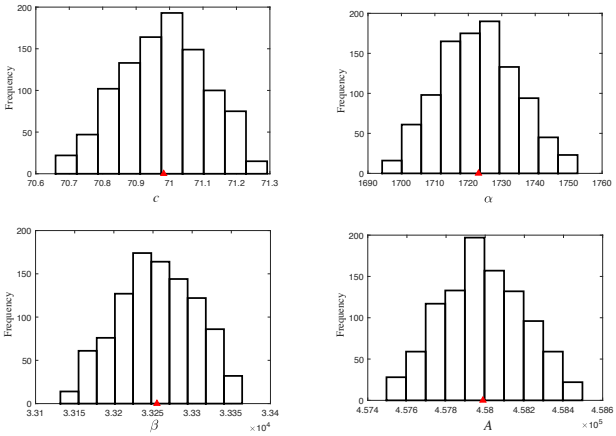


Figure 8: Histograms of the model parameters.

0.77 and 0.63, respectively, indicating an excellent fit.

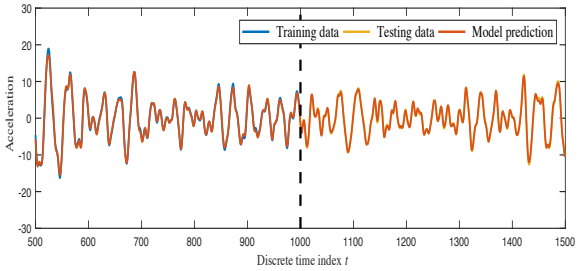


Figure 9: Model predictions on training and testing data sets.

To check if the choice of the best model is dependent on the training data, a new simulation is performed using a longer time series. The same tendency using a short time series has been observed and the algorithm finishes by converging to $\mathcal{M}_2$ (the results are not shown to simplify presentation). Therefore, it becomes obvious that model $\mathcal{M}_2$ is the model with the best predictive ability among the competing models.

5 CONCLUSIONS

A novel algorithm named ABC-NS, for parameter estimation and model selection based on an ellipsoidal nested sampling has been developed. The efficiency and robustness of the novel algorithm to deal with parameter estimation and model selection issues in structural dynamics has been demonstrated through two examples: a moving average process and a wire rope isolator. It has been shown that the algorithm is a promising alternative, offering the possibility to make Bayesian inference for complex systems based on the selection of an appropriate summary statistics. The algorithm is able to provide a good estimation of the unknown parameters within a reasonable computational requirement. In addition, it makes

a comparison between a set of competing models with different dimensionality in an automatic way, as the parsimony principle is naturally embedded in the ABC-NS algorithm. The efficiency of the novel ABC algorithm to deal with more challenging scenarios using different kinds of features and metrics from either the time domain or the frequency domain will be investigated in an upcoming work.

ACKNOWLEDGMENT

The support of the UK Engineering and Physical Sciences Research Council (EPSRC) through grant reference no. EP/K003836/1 is greatly acknowledged.

REFERENCES

[1] **Green**, P.J., Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika. 82(4) (1995) 711–732.

[2] **Vehtari**, A., Ojanen, J., A survey of Bayesian predictive methods for model assessment, selection and comparison. Statistics Surveys, 6:142–228, 2012.

[3] **Ben Abdessalem**, A., Dervilis, N., Wagg, D. and Worden, K., Identification of nonlinear dynamical systems using approximate Bayesian computation based on a sequential Monte Carlo sample,. International Conference on Noise and Vibration Engineering, September 19-21, 2016, Leuven (Belgium).

[4] **Ben Abdessalem**, A., Worden, K., Dervilis, N. and Wagg, D., Recent advances in approximate Bayesian computation methodology: application in structural dynamics, ENL Workshop, January 2017, Bristol, United Kingdom.

[5] **Ben Abdessalem**, A., Dervilis, N., Wagg, D. and Worden, K., ABC-NS: a new computational inference method applied to parameter estimation and model selection in structural dynamics. Conference: 23 Congrès Français de Mécanique, May 2017, Lille, France.

[6] **Ben Abdessalem**, A., Dervilis, N., Wagg, D. and Worden, K., Model selection and parameter estimation in structural dynamics using approximate Bayesian computation, Mech. Syst. Signal Process. 99 (2018) 306–325.

[7] **Marjoram**, P., Molitor, J., Plagnol, V. and Tavare, S., Markov chain Monte Carlo without likelihoods, Proc. Natl. Acad. Sci. USA, 100, 15324–15328.

[8] **Toni**, T., Welch, D., Strelkowa, N., Ipsen, A. and Stumpf, M.P.H., Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, J. of the Royal Society Interface, 6 (2009) 187–202.

[9] **Skilling**, J., Nested sampling, in: R. Fischer, R. Preuss, U.V. Toussaint (Eds), American, Institute of Physics Conference Series, 2004, 395-405.

[10] **Skilling**, J., Nested sampling for general Bayesian computation, Bayesian Analysis, (2006) 1(4), 833-860, doi:10.1214/06-BA127.

[11] **Beaumont**, M.A., Zhang, W. and Balding, D.J., Approximate Bayesian computation in population genetics, Genetics, 162 (4) (2002) 2025–2035.

[12] **Mukherjee**, P., Parkinson, D. and Liddle, A.R., 2006, ApJ, 638, L51.

[13] **Feroz**, F., Hobson, M.P. and Bridges, M., MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics, Mon Not R Astron Soc 398 (4):(2009) 1601–1614.

[14] **Marin**, J.M., Pudlo, P., Robert, C.P. and Ryder, R.J., Approximate Bayesian computational methods, Stat. Comput. 22 (2012), 1167–1180.

[15] Technical Report, COST Action F3, VTT Technical Research Centre of Finland, 1999.

[16] **Juntunen**, M., Linjama, J., Presentation of the VTT benchmark. Mech. Syst. Signal Process. (2003) 17(1), 179–182

[17] **Kerschen**, G., On the Model Validation in Non-linear Structural Dynamics. Thèse de doctorat, Université de Liège, 2002.

[18] **Kerschen**, G., Worden, K., Vakakisc, A.F. and Golinval, J.C., Past, present and future of nonlinear system identification in structural dynamics, Mech. Syst. Signal Process. 20 (2006) 505–592.

[19] **Peifer**, M., Timmer, J. and Voss, H.U., Non-parametric identification of non-linear oscillating systems. Journal of Sound and Vibration 267 (2003) 1157–1167.

[20] **Demetriades**, G.F., Constantinou, M.C. and Reinhorn, A.M., Study of wire rope systems for seismic protection of equipment in buildings. Volume 15, Issue 5, September 1993, Pages 321–334.

[21] **Constantinou**, M.C., Tadjbakhsh, I.G., Hysteretic Dampers in Base Isolation: Random Approach. Journal of Structural Engineering 111(4), April 1985

[22] **Wen**, Y., Method for random vibration of hysteretic systems, ASCE Journal of the Engineering Mechanics Division. 102(2) (1976) 249–263.

[23] **Worden**, K., Hensman, J.J., Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference, Mech. Syst. Signal. Process. 32 (2012) 153–169.