Earthquakes and Structures, *Vol. 8, No. 4 (2015) 915-934*
DOI: http://dx.doi.org/10.12989/eas.2015.8.4.915

Metamodeling of nonlinear structural systems with parametric uncertainty subject to stochastic dynamic excitation

Minas D. Spiridonakos* and Eleni N. Chatzi[a]

Institute of Structural Engineering, D-BAUG, ETH Zurich Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland

(Received April 9, 2014, Revised October 28, 2014, Accepted November 2, 2014)

Abstract. Within the context of Structural Health Monitoring (SHM), it is often the case that structural systems are described by uncertainty, both with respect to their parameters and the characteristics of the input loads. For the purposes of system identification, efficient modeling procedures are of the essence for a fast and reliable computation of structural response while taking these uncertainties into account. In this work, a reduced order metamodeling framework is introduced for the challenging case of nonlinear structural systems subjected to earthquake excitation. The introduced metamodeling method is based on Nonlinear AutoRegressive models with eXogenous input (NARX), able to describe nonlinear dynamics, which are moreover characterized by random parameters utilized for the description of the uncertainty propagation. These random parameters, which include characteristics of the input excitation, are expanded onto a suitably defined finite-dimensional Polynomial Chaos (PC) basis and thus the resulting representation is fully described through a small number of deterministic coefficients of projection. The effectiveness of the proposed PC-NARX method is illustrated through its implementation on the metamodeling of a five-storey shear frame model paradigm for response in the region of plasticity, i.e., outside the commonly addressed linear elastic region. The added contribution of the introduced scheme is the ability of the proposed methodology to incorporate uncertainty into the simulation. The results demonstrate the efficiency of the proposed methodology for accurate prediction and simulation of the numerical model dynamics with a vast reduction of the required computational toll.

Keywords: nonlinear dynamics; earthquake excitation; metamodeling; nonlinear ARX models; polynomial chaos expansion; system identification; uncertainty quantification

1. Introduction

The assessment of the dynamic behavior of structural systems subjected to extreme loading conditions, such as earthquakes, is of particular importance for the safe operation of civil structures and forms a significant chapter within the broad topic of Structural Health Monitoring (SHM) (Farrar 2007). Nevertheless, for the vast majority of such structures, it is hardly ever possible to perform dynamic response testing on the actual structure or even realistic simulation tests on an appropriately scaled structural model.

*Corresponding author, Ph.D., E-mail: spyridonakos@ibk.baug.ethz.ch
[a]Assistant Professor, E-mail: chatzi@ibk.baug.ethz.ch

http://www.techno-press.org/?journal=eas&subpage=7 ISSN: 2092-7614 (Print), 2092-7622 (Online)

The advancements achieved in the field of Finite Element (FE) modeling methods have played a key role in overcoming this hurdle by enabling the realization of sophisticated simulation experiments emulating structural response even in the nonlinear range (Bathe 2009). However, despite the rapidly growing computational power and the continuous development of increasingly efficient algorithms, the accompanying growth in the complexity of FE models as well as the necessity for more detailed descriptions of both structural geometry and mechanical properties render the use of highly detailed FE models almost prohibitive for complex, large structures (Gholizadeh and Salajegheh 2009).

The problem is even more pronounced when taking into account that structural systems are commonly characterized by parameter uncertainty to what concerns for instance the mechanical properties of the structure (Hernandez and Bernal 2008, Christodoulou *et al*. 2008, Vanik *et al*. 2000, Beck and Katafygiotis 1998). Thus, the analyst is faced with the task of performing a number of simulations in order to obtain an accurate numerical model of an existing structure. Moreover, when the linearity assumption is relaxed in favor of increased modeling accuracy the numerical model must be tested for a number of different excitation scenarios. In the case of stochastic dynamic loading, which might often be the case (Katkhuda *et al*. 2005, Poulimenos and Fassois 2006, Lourens *et al*. 2012, Naets *et al*. 2013, Kalkan and Chopra 2010), the statistical characteristics of the input excitation need also be treated as uncertain random variables.

Thus, for processes requiring a mutliplicity of forward simulations, such as design optimization, or model updating procedures based on time history loading, a simpler representation of the FE model should be considered, able to accurately reproduce the behavior of the structure for a wide range of excitations. To this end, several researchers have dealt with the extraction and identification of models of reduced order both in the case of linear (Fraraccio *et al*. 2008, Caicedo *et al*. 2004, Faravelli *et al*. 2011) and nonlinear dynamics (Chatzi and Smyth 2009, Smyth *et al*. 2002, Yun and Shinozuka 1980, Corigliano and Mariani 2004, Katkhuda *et al*. 2005, Lin *et al*. 2001, Moaveni *et al*. 2010).

In this work, both structural properties and excitation uncertainties are treated by means of a time-series model with random parameters utilized for the description of uncertainty propagation through the nonlinear numerical model. The time series model is of a nonlinear autoregressive with exogenous input form. Methodologies of this type have previously been explored in the works of Samara *et al*. (2013), Cheng *et al*. (2007), Rutherfor *et al*. (2007), Kerschen *et al*. (2006), Adams and Farrar (2002), Worden and Tomlinson (2000), and have been proven efficient in simulating the dynamics of nonlinear systems or extracting nonlinear features. In the approach introduced herein, uncertainty of the structural and loading parameters is incorporated via implementation of a Polynomial Chaos (PC) expansion representation. More specifically, the random responses of a large-scale numerical model are approximated by a suitably defined Polynomial Chaos Nonlinear AutoRegressive with eXogenous input (PC-NARX) metamodel. The latter is able to describe both aforementioned types of uncertainties, i.e., regarding the structural properties and the input excitation, by expansion of its random model parameters onto a finite-dimensional PC basis (Spiridonakos and Chatzi 2012, Blatman and Sudret 2010, Sapsis and Lermusiaux 2009). Thus, the PC-NARX metamodel is described by a small number of deterministic coefficients of projection, and is therefore appropriate for simulating the structural system at hand under a significantly reduced computational effort.

The method's effectiveness is demonstrated through the identification of a metamodel for a five-storey building with nonlinear material properties. Toward this end, a limited number of simulation experiments is carried out, with the FE numerical model being subjected to different realizations of synthetic earthquakes. Synthetic earthquakes are produced by filtering a white noise process through a non-stationary impulse response filter and a non-stationary modulating function in order to simulate the time-varying characteristics for both the temporal and spectral content of a real earthquake as proposed in the work of Rezaeian and Der Kiureghian (2010). The non-stationary filter and the modulating functions are described by a small number of uncertain parameters, with sample observations of the latter being estimated through fitting of the model to real recorded earthquake ground motion signals. One of the important features of the introduced approach therefore lies in the parallel parameterization of the input.

The remainder of the paper is organized as follows: The description of the PC-NARX class of models is presented in Section 2 along with the proposed model parameter and model structure selection methods. The model for the parametric representation of real earthquakes, along with the estimation results obtained by modeling the PEER database earthquake accelerograms, are described in Section 3. The numerical application involving a five-storey frame, the conducted simulations and the metamodeling results are presented in Section 4, while the main conclusions of the study are drawn in Section 5.

2. Polynomial chaos NARX models

Metamodeling refers to the process of identifying a reduced order, computationally efficient representation of a large scale numerical model. In the present work, a metamodel is sought for the accurate representation of the nonlinear dynamics of a refined numerical model and the accurate simulation and/or prediction of its time history loading response.

Let us consider a structural system represented by a numerical model $\mathcal{M}$ that is characterized by a number of input parameters relating to the properties of the modeled structure (mechanical and/or geometric). It is assumed that M_s of these parameters are subject to uncertainty and that they may be described by independent random variables gathered in a random vector $\boldsymbol{\xi}_S = [\xi_1, \xi_2, \ldots, \xi_{M_S}]^{\mathrm{T}}$. Superscript T denotes transpose of a vector or a matrix while it is noted that for simplicity of notation no distinction is made between a random variable and its value(s).

Since the dynamic behavior of a nonlinear system is also a function of the input excitation characteristics, it is also considered that the set of excitation signals used for the dynamic loading of the numerical model may be parameterized by a small number of random variables $\boldsymbol{\xi}_X = [\xi_{M_S+1}, \xi_{M_S+2}, \ldots, \xi_{M_S+M_X}]^{\mathrm{T}}$.

As a result of the uncertainty propagation, the dynamic response of the numerical model to a given input excitation will also be a random variable which in addition depends on time, that is

$$y[t, \boldsymbol{\xi}] = \mathcal{M}\left(x[1, \boldsymbol{\xi}_X], x[2, \boldsymbol{\xi}_X], \ldots, x[t, \boldsymbol{\xi}_X], \boldsymbol{\xi}_S\right) \tag{1}$$

with $t = 1,2,\ldots,T$ designating normalized by the sampling period discrete time, $x[t, \xi_X]$ the excitation input signal, $y[t,\xi]$ the corresponding numerical model response signal, and $\boldsymbol{\xi} = [\boldsymbol{\xi}_S^{\mathrm{T}}, \boldsymbol{\xi}_X^{\mathrm{T}}]^{\mathrm{T}}$ the $M-$dimensional ($M=M_S+M_X$) complete vector of input random variables with known joint probability density function (pdf) $f(\xi)$.

In order to approximate the numerical model dynamic response $y[t, \boldsymbol{\xi}]$ for every realization of $\boldsymbol{\xi}$ in an efficient way, a novel metamodeling method based on Polynomial Chaos Nonlinear AutoRegressive with eXogenous input (PC-NARX) models is introduced in the present study. The nonlinear form being key to the accurate representation of a structural system subjected to significant amplitudes of input ground motion. The general PC-NARX model, in the linear-in-the-parameters form, is given by the following relationship (for the purposes of notational simplicity the dependency of the input and output signals on $\boldsymbol{\xi}$ is not indicated explicitly in the following relationships)

$$y[t] = \sum_{i=1}^{n_\theta} \theta_i\left(\boldsymbol{\xi}\right) \cdot g_i\left(\boldsymbol{z}[t]\right) + e[t] \tag{2}$$

where $g_i(\boldsymbol{z}[t])$ are the nonlinear model terms generated from the regression vector $\boldsymbol{z}[t]=[y[t-1],\ldots,y[t-n_a],x[t],\ldots,x[t-n_b]]^{\mathrm{T}}$ with n_a,n_b designating the maximum output and input time lags, respectively, and $e[t] \sim \mathrm{NID}(0,\sigma_e^2)$ the model's residual sequence with NID $(\cdot,\cdot)$ denoting a Normally Independently Distributed process with the indicated mean and variance. It should be mentioned that the model terms $g_i(\boldsymbol{z}[t])$ may be constructed from a variety of local or global basis functions including polynomials, splines, neural networks, wavelets and others (Wei *et al.* 2004).

The important feature of PC-NARX models, in comparison with the conventional NARX models (Chen and Billings 1989), is that they are characterized by parameters $\theta_i(\boldsymbol{\xi})$ which are random variables themselves. These parameters are actually represented by a deterministic mapping which describes their relation to the input random variables. More specifically, assuming that the PC-NARX model parameters $\theta_i(\boldsymbol{\xi})$ have finite variance, they admit the following polynomial chaos representation (Soize and Ghanem2004)

$$\theta_i(\boldsymbol{\xi}) = \sum_{j=1}^{\infty} \theta_{i,j} \cdot \varphi_{\boldsymbol{d}(j)}(\boldsymbol{\xi}) \tag{3}$$

where $\theta_{i,j}$ are unknown deterministic coefficients of projection, $\boldsymbol{d}(j)$ is the multi-index of the multivariate polynomial basis, and $\varphi_{\boldsymbol{d}(j)}$ are multivariate basis functions that are orthonormal with respect to the joint pdf of $\boldsymbol{\xi}$, that is

$$\mathrm{E}\left[\varphi_{\boldsymbol{\alpha}}\left(\boldsymbol{\xi}\right), \varphi_{\boldsymbol{\beta}}\left(\boldsymbol{\xi}\right)\right] = \delta_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{cases} 1 & \text{for } \boldsymbol{\alpha} = \boldsymbol{\beta} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Each probability density function may be associated with a well-known family of orthogonal polynomials. For instance, the normal distribution is associated with Hermite polynomials while the uniform distribution with Legendre. A list of the most common probability density functions along with the corresponding orthogonal polynomials and the relations for their construction may be found in (Soize and Ghanem 2004).

For purposes of practicality, the infinite series of expansion of Eq. (3) must be truncated by selecting an appropriate functional subspace consisting of a finite number of terms p. In this way, the resulting PC-NARX model is fully parametrized in terms of a finite number ($n_\theta \times p$) of deterministic coefficients of projection $\theta_{i,j}$, while the complete PC-NARX identification problem consists of the subproblems of model parameter estimation and model structure selection, which are discussed in the following sections.

At this point it should be noted that PC-ARIMA models with a-priori known deterministic coefficients of projection have been considered before for the characterization of terrain topology by Wagner and Ferris (Wagner and Ferris 2007). Moreover, linear PC-ARX models have been used in the metamodeling context in previous of work of the authors (Spiridonakos and Chatzi

2012). Similarly linear ARX models with functionally dependent parameters have also been used in a number of studies for the purposes of structural identification and damage detection (for instance see Kopsaftopoulos and Fassois (2013) and the references therein), and have been extended to the nonlinear case for aircraft virtual sensor design in Samara *et al*. (2013). However, in these studies the input parameters are not related with a known probability density function and thus the basis is constructed from an arbitrarily selected family of orthogonal basis functions.

2.1 Parameter estimation

As already mentioned, the estimation of a PC-NARX metamodel refers to the determination of the coefficients of projection parameter vector $\boldsymbol{\theta}$

$$\boldsymbol{\theta} = \left[\theta_{1,1}, \theta_{1,2}, \ldots, \theta_{n_\theta, p}\right]^T \tag{5}$$

The calculation has to be based on the availability of time history data for the input excitation and output response of the numerical model. These may be acquired for a small number of simulations, conducted for different realizations of the input random vector using the full scale numerical model.

Let us consider a series of K simulations conducted for a corresponding number of input random vector realizations $\boldsymbol{\xi}_k = \left[\xi_{k,1}, \xi_{k,2}, \ldots, \xi_{k,M}\right]^T$ (for k =1,2,…,K), and the resulting set of excitation signals $x_k^T = \{x_k[1], x_k[2], \ldots, x_k[T]\}$.

The corresponding dynamic response of the full scale numerical model is indicated as $y_k^T = \{y_k[1], y_k[2], \ldots, y_k[T]\}$ and as already mentioned it is assumed to follow the general PC-NARX model of Eq. (2)

$$y_k[t] = \sum_{i=1}^{n_\theta} \theta_i(\xi_k) \cdot g_i(z_k[t]) + e_k[t], \quad t = 1, \ldots, T, \qquad k = 1, \ldots K \tag{6a}$$

$$e_k[t] \sim \mathrm{NID}(0, \sigma_{e_k}), \tag{6b}$$

$$\mathrm{E}\{e_i[t], e_j[t-\tau]\} = \begin{cases} \sigma_{e_i}^2 \cdot \delta_{t,\tau} & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \tag{6c}$$

In the relationships above, uncorrelatedness of the residual series between different simulation experiments is also assumed. For these simulations, the input vector ξ_k is generated from the input parameter space either randomly or by using a structured sampling technique, such as the Latin Hypercube Sampling (LHS) (Helton and Davis 2003).

According to the intended use of the metamodel, the estimation of the parameter vector $\boldsymbol{\theta}$ may be based upon the minimization of either the Simulation Error (SE) criterion or Prediction Error (PE) criterion.

2.1.1 Simulation error method

The SE estimation method must be employed when a PC-NARX metamodel that may replace the numerical model for additional simulations is sought.

Let us consider again the same set of K input excitation signals $x_k^T (k = 1, 2, \ldots, K)$. For each x_k^T

the simulated response $\bar{y}_k[t]$ of a PC-NARX metamodel may be obtained by using the input excitation signals and the following relationship applied recursively with respect to time t

$$\bar{y}_k[t] = \sum_{i=1}^{n_\theta} \theta_i(\boldsymbol{\xi}_k) \cdot g_i\left(\bar{z}[t]\right) \tag{7}$$

with given initial conditions $\{\bar{y}_k[0], \ldots, \bar{y}_k[-n_a]\}$. The responses $\bar{y}_k[t]$ of the metamodel should be as close as possible to that of the numerical model $y_k[t]$, thus the estimation of a PC-NARX metamodel may be based on the minimization of the following SE criterion

$$\hat{\boldsymbol{\theta}}_s = \arg\min_{\boldsymbol{\theta}_s}\left\{\sum_{k=1}^{K}\sum_{t=1}^{T}\left(y_k[t] - \bar{y}_k[t]\right)^2\right\} = \arg\min_{\boldsymbol{\theta}_s}\left\{\sum_{k=1}^{K}\sum_{t=1}^{T}\varepsilon_k[t]^2\right\} \tag{8}$$

where $\varepsilon_k[t]$ designates the simulation error and $\arg\min$ stands for the minimizing argument. This optimization problem may be solved by iterative nonlinear optimization methods. However, such methods are normally amenable to local minima convergence problems when arbitrary initial conditions are used, resulting to models with poor simulation performance or even models with unstable simulated responses. For this reason the optimization procedure should be initialized either by derivative-free search methods or by the parameter vector $\boldsymbol{\theta}$ obtained through the PE method of the following section.

2.1.2 Prediction error method

Beyond providing initial estimates for the nonlinear optimization procedure of SE method, the PE-based estimation method may also provide PC-NARX metamodels for the online prediction of the structural dynamic response, when the structural system corresponding to the numerical model is appropriately instrumented. This metamodel may in turn be used in real-time implementations within the context of vibration-based SHM and control. In such an effort issues such as the coupling of the model with the real structure (model updating) and possibly the modelling of responses of unobserved degrees-of-freedom should be addressed, however a more detailed investigation of these points lies outside the scope of the present study.

The PE criterion consists in the sum of squares of the model's one-step-ahead prediction errors for the complete set of simulation experiments. As it may be easily demonstrated, the one-step-ahead prediction errors coincide with the model's residual sequence, and thus $\boldsymbol{\theta}$ has to be estimated by the following minimization problem

$$\hat{\boldsymbol{\theta}} = \underbrace{\operatorname{argmin}}_{\boldsymbol{\theta}}\left\{\sum_{k=1}^{K}\sum_{t=1}^{T}\left(y_k[t] - \hat{y}_k[t|t-1]\right)^2\right\} = \underbrace{\operatorname{argmin}}_{\boldsymbol{\theta}}\left\{\sum_{k=1}^{K}\sum_{t=1}^{T} e_k[t]^2\right\} \tag{9}$$

with $\hat{y}_k[t\,|\,t-1]$ designating the model's one-step-ahead prediction, and $e_k[t]$ the prediction error. Toward this end, and by using Eq. (3), Eq. (6a) may be rewritten as

$$\begin{aligned} y_k[t] &= \sum_{i=1}^{n_\theta}\sum_{j=1}^{p}\theta_{i,j}\cdot\varphi_{\boldsymbol{d}(j)}\left(\xi_k\right)\cdot g_i\left(z_k[t]\right) + e_k[t] \Rightarrow \\ y_k[t] &= \underbrace{\left[\varphi_{d(1)}\left(\xi_k\right)g_1\left(z_k[t]\right)\cdots\varphi_{\boldsymbol{d}(p)}\left(\xi_k\right)g_{n_\theta}\left(z_k[t]\right)\right]}_{\boldsymbol{\pi}[\xi_k,t]^T}\cdot\begin{bmatrix}\theta_{1,1}\\ \vdots \\ \theta_{n_\theta,p}\end{bmatrix} + e_k[t] \end{aligned} \tag{10}$$

or by stacking all time-instants in a single vector

$$\underbrace{\begin{bmatrix} y_k[1] \\ y_k[2] \\ \vdots \\ y_k[T] \end{bmatrix}}_{y_{k[T\times 1]}} = \underbrace{\begin{bmatrix} \varphi[\xi_k,1]^T \\ \varphi[\xi_k,2]^T \\ \vdots \\ \varphi[\xi_k,T]^T \end{bmatrix}}_{\Phi(\xi_k)_{[T\times n_\theta p]}} \cdot \underbrace{\begin{bmatrix} \theta_{1,1} \\ \theta_{1,2} \\ \vdots \\ \theta_{n_\theta,p} \end{bmatrix}}_{\theta_{[n_\theta p\times 1]}} + \underbrace{\begin{bmatrix} e_k[1] \\ e_k[2] \\ \vdots \\ e_k[T] \end{bmatrix}}_{e_{k[T\times 1]}}$$

where subscripts in brackets indicate the respective matrix/vector dimensions. Finally, by pooling all the available simulation experiments, the following linear regression model is obtained

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_K \end{bmatrix}}_{\mathbf{Y}_{[KT\times 1]}} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}(\xi_1) \\ \boldsymbol{\Phi}(\xi_2) \\ \vdots \\ \boldsymbol{\Phi}(\xi_K) \end{bmatrix}}_{\boldsymbol{\Phi}(\xi)_{[KT\times n_\theta p]}} \cdot \underbrace{\begin{bmatrix} \theta_{1,1} \\ \theta_{1,2} \\ \vdots \\ \theta_{n_\theta,p} \end{bmatrix}}_{\boldsymbol{\theta}_{[n_\theta p\times 1]}} + \underbrace{\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_K \end{bmatrix}}_{\boldsymbol{E}_{[KT\times 1]}} \tag{11}$$

where $\boldsymbol{\Phi}(\xi)$ is the regression matrix. Thus, the minimization problem of Eq. (9) may be rewritten as follows

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \left\{ \mathbf{E}^T \cdot \mathbf{E} \right\} \tag{12}$$

which due to the linear dependence of the residual sequence $\mathbf{E}$ on the parameter vector $\boldsymbol{\theta}$ leads to a linear ordinary least squares estimator for the latter, that is

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{\Phi}^{\mathsf{T}}(\xi) \cdot \boldsymbol{\Phi}(\xi) \right)^{-1} \cdot \left(\boldsymbol{\Phi}(\xi)^{\mathsf{T}} \cdot \mathbf{Y} \right) \tag{13}$$

2.2 Model structure selection

Model structure selection may be considered as the optimization procedure during which metamodels corresponding to various candidate “structures” are estimated (through the methods of the previous section), and the one providing the “*best fitness*” to the simulation data is selected.

Particularly, the PC-NARX model structure selection procedure concerns the determination of: *i*) the nonlinear model terms $g_i(z[t])$ and *ii*) the PC basis functional subspace.

2.2.1 NARX terms

Given the specific family of functions (e.g., polynomials, splines, neural networks, or wavelets), and the maximum output and input lags n_a and n_b, a finite number of appropriate nonlinear terms must be selected. This selection cannot be based on a trial-and-error approach, since the dimension of the initial search space is usually excessive, while also the nonlinear optimization procedure required for the SE estimation method is computationally intensive. For these reasons a more automated and computationally inexpensive approach is proposed herein.

The nonlinear terms selection procedure is initialized by selecting the simulation dataset exhibiting the highest nonlinearity, assuming that this set will require the maximum number of nonlinear terms for its accurate representation. This dataset may be chosen based on either physical insight (e.g., simulated response corresponding to the highest excitation forces) or nonlinearity quantification criteria such as coherence or cross-bicoherence functions (Choudhury *et al*. 2008).

Assuming that experiment k_o is selected for this purpose, the excitation signal $x_{k_o}^T$ and the simulated response $y_{k_o}^T$ are utilized along with a Genetic Algorithm (GA) (Coley 1999) in order to determine promising subregions of the complete search space. Each GA individual is a bit-string, with each bit representing the existence (1) or not (0) of the corresponding nonlinear term from the initial search space. For example considering the initial search space, given as $\{y[t\text{-}1], y[t\text{-}1]^2, y[t\text{-}1]^3, x[t], x[t]^2, x[t]^3\}$, each individual will be represented by a 6-bit vector, with a hypothetical optimal solution [101100] corresponding to $\boldsymbol{g}(\boldsymbol{z}[t])=\{y[t\text{-}1], y[t\text{-}1]^3, x[t]\}$.

Each candidate NARX model created by the GA, is estimated by means of the PE method, although judged in terms of its simulation capabilities. This simplification is based on an implicit assumption on the proximity of the PE-based and SE-based estimated models and it is employed in order to circumvent the nonlinear optimization procedure of the SE method.

In order to refine, and potentially reduce model dimensionality, the concept of backward elimination may be used as a second step. Starting from the GA solution, the nonlinear term whose removal has the minimum effect on the SE criterion is dropped, with the procedure being repeated till no more regressors are available. In this way, only the nonlinear terms which play a significant role in SE criterion minimization are retained.

This phase is concluded by estimating "local" NARX models for each simulation experiment. The local NARX model parameters $\theta_i^{\text{local}}(\boldsymbol{\xi}_k)(i=1,\ldots,n_\theta \text{ and } k=1,\ldots,K)$ are estimated by means of the SE method.

2.2.2 PC basis

This second phase, is based on the estimated local NARX model parameters in order to define appropriate PC basis subspaces for their expansion. Toward this end, the estimated parameter vector $\boldsymbol{\theta}^{\text{local}}(\boldsymbol{\xi})$ is initially projected onto a PC basis consisting of a small number of functions by using the following linear regression model and ordinary least squares optimization

$$\boldsymbol{\theta}_i^{\text{local}}(\boldsymbol{\xi}) = \sum_{j=1}^{p_1} \theta_{i,j} \cdot \varphi_{\boldsymbol{d}(j)}(\boldsymbol{\xi}) + \boldsymbol{w}_i \tag{14}$$

where $\boldsymbol{w}_i$ designates the vector of the expansion residuals. Then, it is examined whether the addition of more PC basis functions significantly contributes to the reduction of the residual sum of squares criterion

$$\text{RSS} = \sum_{i=1}^{n_\theta} \sum_{k=1}^{K} w_{i,k}^2 \tag{15}$$

The significance of the added term is checked by the partial F-test, that is the F-statistic

$$F = \frac{\dfrac{\text{RSS}_1 - \text{RSS}_2}{q_2 - q_1}}{\dfrac{\text{RSS}_2}{Kn_\theta - q_2}} \tag{16}$$

Table 1 PC-NARX model structure selection procedure

Step 1.	Select k_o dataset (either by taking into consideration characteristics of the input random vector realization or based on a nonlinearity measure).
Step 2.	Define initial search space for nonlinear terms.
Step 3.	GA for the selection of nonlinear terms.
Step 4.	Refine GA results of the previous step by dropping excessive terms.
Step 5.	Estimate local NARX models for each simulation dataset (SE method).
Step 6.	Use partial F-test in order to determine an appropriate sparse PC basis (maximum total degree has to be preselected).
Step 7.	Refine PC-NARX model by using coefficients of projection of the previous step and SE method.

is calculated, with RSS_1 designating the residual sum of squares of the initial model (model 1) with q_1 coefficients of projection and RSS_2 that of model 2 with q_2 coefficients of projection, while Kn_θ is the total dimension of the $\boldsymbol{\theta}^{\text{local}}$ vector. Under the null hypothesis, model 2 does not provide a significantly better fit than model 1. The F-statistic will follow an F distribution, with (q_2-q_1, Kn_θ-q_2) degrees of freedom, while the null hypothesis will be rejected if the F-statistic is greater than the critical value of the F-distribution for a selected false-rejection probability. Thus, PC basis functions may be added till the F-statistic becomes lower than the critical value of the F-distribution.

Normally, the process above may be started with the expansion of the local NARX parameters onto the constant basis $\varphi_{\boldsymbol{d}(1)}(\xi)$ with $\boldsymbol{d}_{(1)}=[0\ 0\ 0\ldots0]^{\mathrm{T}}$ and the addition each time of a single PC basis function - the one which leads to the highest reduction of the RSS criterion. The final coefficients of projection $\theta_{i,j}$ are used as initial estimates for the final PC-NARX model which is refined through SE-based nonlinear optimization. The basic steps of the PC-NARX model structure selection procedure are summarized in Table 1.

3. Parametric representation of synthetic earthquakes

In the present study, the stochastic ground motion model proposed in (Rezaeian and Der Kiureghian 2010) is employed for the parametric representation of synthetic earthquake ground motion acceleration signals. According to this model the ground motion is produced by time-modulating a normalized filtered white noise process. In order to simulate the time-varying characteristics of both the temporal and spectral content of a real earthquake, the modulating function and the Impulse Response Filter (IRF) of the aforementioned model have non-stationary properties.

More specifically, the non-stationary modulating function is defined as a gamma function of the following form

$$q(t,\boldsymbol{\alpha}) = \alpha_1 t^{\alpha_2 - 1} \exp(-\alpha_3 t) \tag{17}$$

where $\boldsymbol{\alpha}=[\alpha_1,\ \alpha_2,\ \alpha_3]^{\mathrm{T}}$, with $\alpha_1>0$ the parameter of the process intensity, $\alpha_2>1$ the parameter controlling the shape of the modulating function, and $\alpha_3>0$ the duration of the motion. These parameters may be directly estimated from three time-domain characteristics of a ground motion accelerogram: *i)* the expected Arias intensity I_a, which forms a measure of the total energy contained in the motion, *ii)* the effective duration of the motion $D_{5\text{-}95}$ (the time interval between the instants at which the 5 and 95% of the expected Arias intensities are reached), and finally, *iii)* the time t_{mid} at which a 45% level of the expected Arias intensity is reached.

On the other hand, the non-stationary IRF is given by the following expression

$$h[t-\tau]=\frac{\omega_f(\tau)e^{-\zeta_f\omega_f(\tau)(t-\tau)}\cdot\sin\left[\omega_f(\tau)(t-\tau)\sqrt{1-\zeta_f^2}\right]}{\sqrt{1-\zeta_f^2}},\ \tau\le t \tag{18}$$

with ζ_f designating the damping ratio of the filter and $\omega_f(\tau)$ the filter's frequency. The latter is defined as $\omega_f(\tau)=\omega_{\mathrm{mid}}+\omega'(\tau-t_{\mathrm{mid}})$ with ω_{mid} representing the filter frequency at t_{mid}, and ω' the rate of change of the filter frequency with time.

Therefore, given a target accelerogram, the parameters ω_{mid},ω', ζ_f, I_a, $D_{5\text{-}95}$, and t_{mid} may be identified by matching the properties of the recorded motion with the corresponding statistical measures of the stochastic ground motion model (Rezaeian and Der Kiureghian 2010). The identified model may be subsequently used for the generation of synthetic ground motion accelerograms with similar time-frequency characteristics with that of the real recorded motion.

In a similar way, a pdf may be identified for each of the stochastic ground motion model parameters in order to represent not a single but a set of real earthquake accelerograms. This procedure is presently applied for the modeling of the 3297 ground motion signals of the PEER database (horizontal components only with less than 10000 samples; (Peer 2012)). From the identified sets of parameters, only the first 2000 with the best model fitting are kept for the subsequent analysis. The resulting histograms for the identified parameters and the fitted, by the means of the maximum likelihood estimation method, pdfs are shown in Fig. 1. $D_{5\text{-}95}$ and t_{mid} are shown in normalized discrete time t/T, while the results of fitting ζ_f by a parametric pdf were not satisfactory. It is also noted that the square root of the Arias intensity estimated values are used for this parametric fitting.

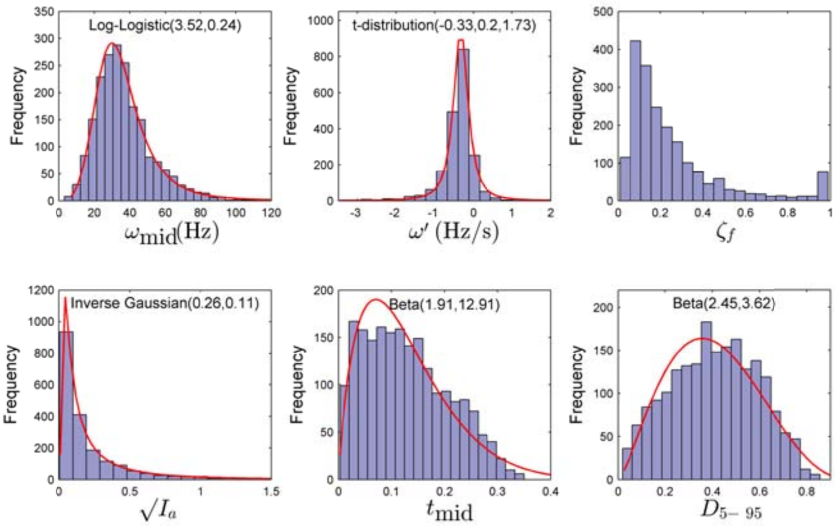

Fig. 1 Histograms of the stochastic ground motion model parameters along with the fitted probability density function

4. Numerical example

A FE model of a building frame structure is presently considered for the validation of the introduced method (Fig. 2). The identification of the frame's metamodel is based on recordings of the horizontal velocity of the top floor of the building measured at node 12 (Fig. 2) obtained by the time history loading of the FE model with various synthetic earthquake ground motion signals.

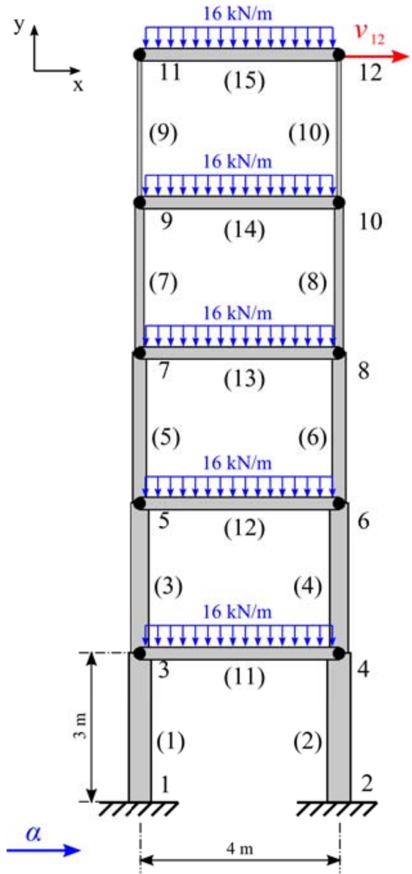

Fig. 2 The five-storey shear frame model

Table 2 Geometric and mechanical properties of the five-storey shear frame model

Geometric		Mechanical	
Cross-sectional area	cm^2		
1st storey columns	484	Poisson ratio	0.29
2nd storey columns	400	Density (kg/m^3)	7850
3rd storey columns	324	Yield stress (MPa)	50
4th storey columns	256	Tangent modulus (GPa)	10
5th storey columns	196		
Horizontal beams	324		

Table 3 Random input variables

Variable	Distribution	pdf parameters
E (GPa)	Uniform	min=180, max=220
ω_{mid} (Hz)	Log-Logistic	μ = 3.52, σ = 0.24
$\sqrt{I_a}$	Inverse Gaussian	μ = 0.26, λ = 0.11

The elements of the shear frame are considered to have square cross sections and being made of steel with isotropic behavior described by a bilinear stress-strain curve. The constant, characterized by negligible uncertainty, mechanical and geometric properties of the structure are summarized in Table 2. On the other hand, the steel's Young modulus E is considered to be uncertain, modeled by a random variable following a uniform distribution E~U(180,220) (GPa).

The uncertain input parameter vector of this numerical example includes also the stochastic ground motion model parameters ω_{mid} and $\sqrt{I_a}$ which follow the distributions indicated in Fig. 1. For this case study, the rate of change of the filter frequency ω', the IRF filter damping ratio ζ_f, and the time variables t_{mid} and $D_{5\text{-}95}$ are considered constant and equal to the PEER database mean estimated values for all synthetic earthquake accelerograms. That is, $\omega'=-0.4160$, ζ_f=0.2514, t_{mid} =0.1296, and $D_{5\text{-}95}$ =0.4080.

Summarizing, in total three independent input random variables with known pdfs are considered for this metamodeling problem (see Table 3).

4.1 Simulation experiments

A total number of 100 simulation is conducted (K=100) for a corresponding number of input random vector realizations ξ_k(k =1,2,…,100) sampled by the LHS method. The values of the sampled variables along with their histograms are shown in Fig. 3.

For each simulation experiment, the shear frame model is excited by a synthetic ground motion accelerogram applied in the x-axis direction (sampling frequency f_s=40 Hz; 1000 samples for each accelerogram). The Peak Ground Acceleration (PGA) values of the synthetic accelerograms

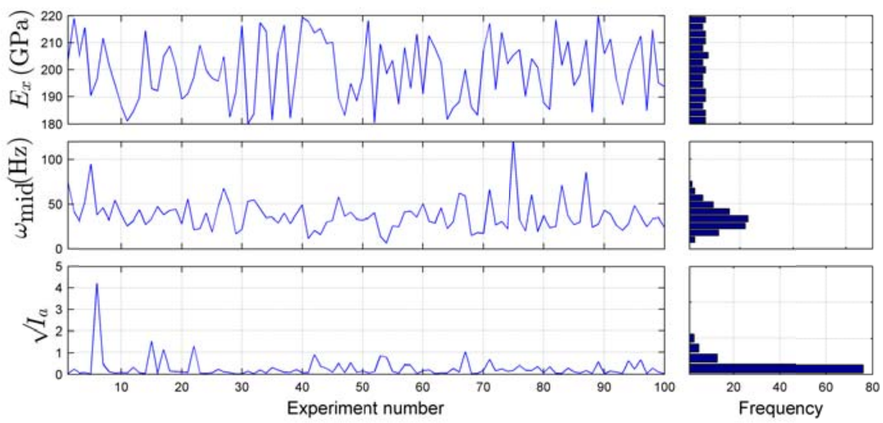

Fig. 3 The input random vector realizations for the 100 simulations conducted

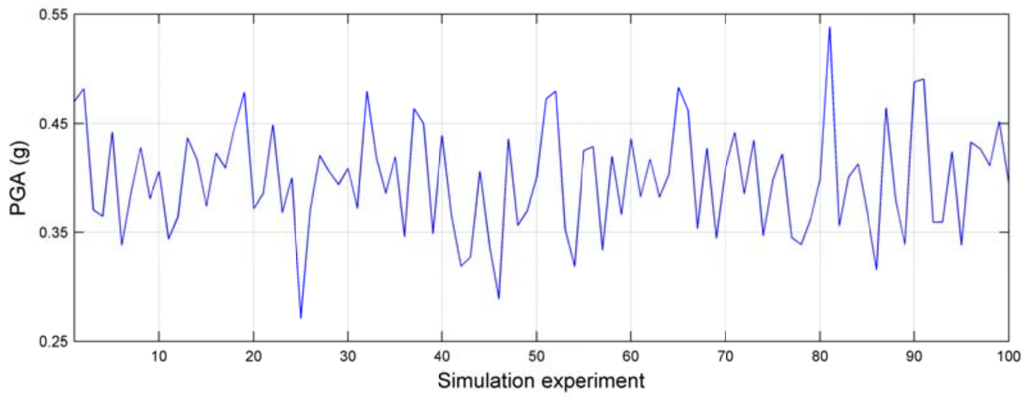

Fig. 4 Peak ground acceleration level for the simulation experiments synthesized earthquakes

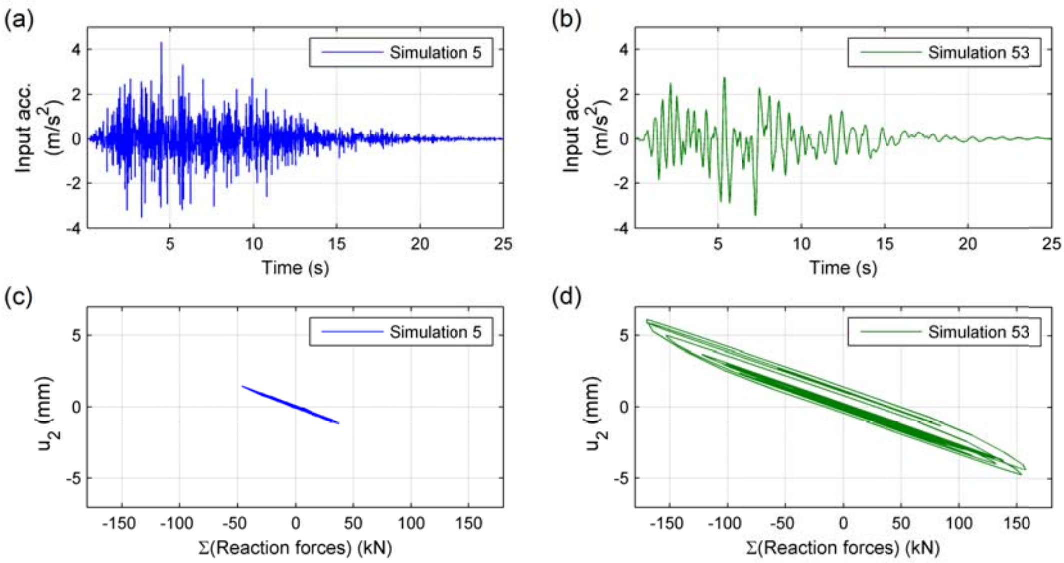

Fig. 5 (a) Input acceleration signal x_5^T and (b) the corresponding plot of the total reaction-force calculated at the base of the shear frame versus 1st floor displacement calculated at node 2. (c) Input acceleration signal x_{53}^T and (d) the corresponding plot of the total reaction-force calculated at the base of the shear frame versus 1st floor displacement calculated at node 2

vary within an interval of 0.25 to 0.55 g (Fig. 4). The selected range is indeed capable of exciting the nonlinear dynamics of the FE model, as indicated in Fig. 5(d) in which the displacement of the FE model calculated for node 2 (1st floor), versus the corresponding sum of reaction forces calculated at the base of the frame for simulation experiment 53 are depicted. In Fig. 5(b) the corresponding plot for simulation experiment 5 is shown. These simulation experiments are those with the lowest (experiment 5) and highest (experiment 53) degree of nonlinearity, according to the mean value of the coherence function criterion (Choudhury *et al*. 2008, Spiridonakos and Chatzi 2012).

4.2 PC-NARX metamodel identification

For the metamodel structure selection problem PC-NARX models with n_a=n_b=10 are considered, since the FE model may be approximated by a 5-dof system corresponding to a NARX model with maximum lag equal to 10, while the appropriate nonlinear terms $g_i(\boldsymbol{z}[t])$ are searched among functions of the following form

$$g_i(\boldsymbol{z}[t]) = z_i[\mathrm{t}] \quad \text{or} \quad g_i(\boldsymbol{z}[t]) = z_i[\mathrm{t}] \cdot |y[t-j]| \tag{19}$$

with $\boldsymbol{z}[t]=[y[t-1],\ldots,y[t-10],x[t],x[t-1],\ldots,x[t-10]]^{\mathrm{T}}$ and j=1,…,10. This type of nonlinear functions have been shown to be appropriate for modelling bilinear material properties and in general systems represented through a hysteretical mechanism (Spiridonakos and Chatzi 2014). The application of the GA algorithm for the selection of the most significant terms is employed along with the dataset of the input excitation and FE model output response of the simulation experiment number 53 which, as already mentioned, corresponds to the simulation experiment depicting the higher degree of nonlinearity. The GA leads to the selection of 124 linear and nonlinear terms, while the backward elimination procedure applied on the GA indicates, as illustrated in Fig. 6, that a significant number of terms may be dropped since their contribution in the reduction of the SE criterion is insignificant. However, as shown in Fig. 6 there is no clear indication of the number of functions that should be rejected since the simulation error criterion depicts a linearly increasing trend with no clear minimum. In some cases, an educated guess is required for the final selection of the salient nonlinear terms, or alternatively a trial-and-error procedure may be followed. In this case study, and in order to compromise the dimensionality of the approximation (included number of terms) with the accuracy attained the first 60 terms are dropped since their rejection leads to an minor increase of the minimization criterion, lying below 0.2%. Thus, the number of finally selected nonlinear terms is equal to 64 (Table 4).

It must be noted that the selection of the appropriate functional form of the nonlinear regressors and the subsequent selection of a subset of these regressors from a usually infinite set is a demanding task (Piroddi and Spinelli 2003, Wei *et al.* 2004).

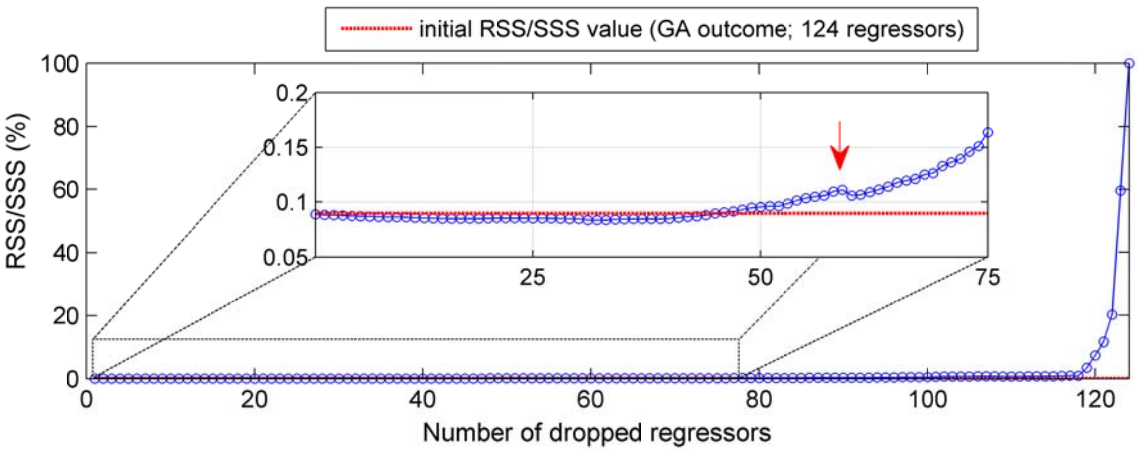

Fig. 6 Backward elimination procedure results for the selection of the nonlinear terms

Table 4 Selected nonlinear terms

$x[t]$	$x[t-2]\|y[t-9]\|$	$x[t-10]\|y[t-1]\|$	$y[t-2]\|y[t-9]\|$	$y[t-7]\|y[t-2]\|$
$x[t-1]$	$x[t-2]\|y[t-10]\|$	$x[t-10]\|y[t-2]\|$	$y[t-3]\|y[t-3]\|$	$y[t-7]\|y[t-8]\|$
$x[t-6]$	$x[t-3]\|y[t-7]\|$	$x[t-10]\|y[t-4]\|$	$y[t-3]\|y[t-10]\|$	$y[t-7]\|y[t-10]\|$
$x[t-10]$	$x[t-4]\|y[t-1]\|$	$x[t-10]\|y[t-5]\|$	$y[t-4]\|y[t-1]\|$	$y[t-8]\|y[t-1]\|$
$y[t-1]$	$x[t-4]\|y[t-2]\|$	$y[t-1]\|y[t-2]\|$	$y[t-4]\|y[t-2]\|$	$y[t-8]\|y[t-2]\|$
$y[t-2]$	$x[t-4]\|y[t-6]\|$	$y[t-1]\|y[t-3]\|$	$y[t-4]\|y[t-9]\|$	$y[t-8]\|y[t-9]\|$
$y[t-6]$	$x[t-6]\|y[t-6]\|$	$y[t-1]\|y[t-4]\|$	$y[t-4]\|y[t-10]\|$	$y[t-9]\|y[t-2]\|$
$y[t-9]$	$x[t-6]\|y[t-10]\|$	$y[t-1]\|y[t-7]\|$	$y[t-5]\|y[t-2]\|$	$y[t-9]\|y[t-3]\|$
$x[t]\|y[t-1]\|$	$x[t-7]\|y[t-5]\|$	$y[t-1]\|y[t-9]\|$	$y[t-5]\|y[t-3]\|$	$y[t-9]\|y[t-7]\|$
$x[t]\|y[t-3]\|$	$x[t-7]\|y[t-7]\|$	$y[t-1]\|y[t-10]\|$	$y[t-5]\|y[t-8]\|$	$y[t-9]\|y[t-9]\|$
$x[t]\|y[t-4]\|$	$x[t-7]\|y[t-9]\|$	$y[t-2]\|y[t-1]\|$	$y[t-6]\|y[t-1]\|$	$y[t-9]\|y[t-10]\|$
$x[t-1]\|y[t-8]\|$	$x[t-8]\|y[t-1]\|$	$y[t-2]\|y[t-2]\|$	$y[t-6]\|y[t-8]\|$	$y[t-10]\|y[t-4]\|$
$x[t-2]\|y[t-1]\|$	$x[t-9]\|y[t-8]\|$	$y[t-2]\|y[t-3]\|$	$y[t-7]\|y[t-1]\|$	

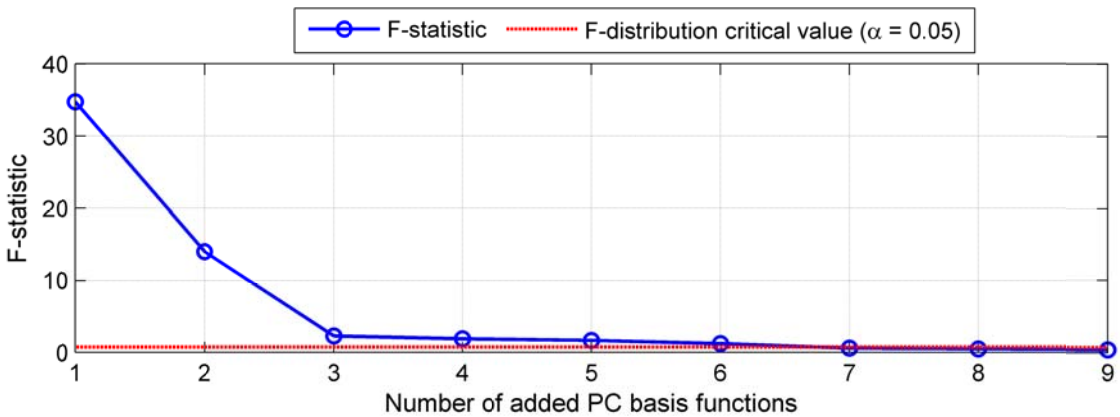

Fig. 7 Partial F-test procedure for the selection of the sparse PC basis

The F-test procedure outlined in section 2.2.2 is then utilized for the selection of the appropriate PC basis functions. The initial search space is defined by the truncated set of multivariate Legendre polynomials basis functions of maximum total degree equal to two. It is noted that the input random variables listed in Table 3 are first transformed into standard uniform variables by using the inverse cumulative density functions of the corresponding pdfs. The calculation of the F-statistic and the corresponding critical values while increasing the dimension of the PC basis is shown in Fig 7. The multi-indices vectors of the finally selected functions are given in Table 5.

Table 5 Multi-indices vectors of the selected PC basis functions

	$\xi_1(E)$	$\xi_2\,(\omega_{\text{mid}})$	$\xi_3\,(I_a)$
$\boldsymbol{d}(1)$	0	0	0
$\boldsymbol{d}(2)$	1	0	0
$\boldsymbol{d}(3)$	1	1	0
$\boldsymbol{d}(4)$	0	1	0
$\boldsymbol{d}(5)$	0	2	0
$\boldsymbol{d}(6)$	0	1	1
$\boldsymbol{d}(7)$	0	0	1

It should be observed that according to these results the PC-NARX metamodel parameters θ_i and thus the dynamic properties of the corresponding numerical model seem to be more sensitive to the input variable ω_{mid} which defines the dominant frequency of the synthetic ground motion signal used for the excitation of the FE model. On the other hand, the results of Table 5 indicate lower sensitivity to input variables E and I_a, i.e., the Young modulus and the intensity of the ground motion parameter. However, this fact may be attributed to the relatively narrow range of their variability.

The performance of the estimated metamodel is assessed through its application on both prediction and simulation of the dynamic response of the FE model excited by the N-S component of the El Centro earthquake accelerogram. It is noted here that for the SE method the Levenberg-Marquardt algorithm, i.e., the *lsqnonlin* MATLAB function is utilized with TolFun=1×10^{-3}, TolX=1×10^{-6}. The Young modulus of the shear frame for this test case is selected as E=200 GPa, while the rest of the input random variables are estimated from the actual accelerogram as ω_{mid}=24.02 (Hz), I_a=0.95 (sampling frequency set equal to 40 Hz; 2688 samples).

The predictions of the PE-based and the simulations of the SE-based estimated PC-NARX metamodels are contrasted to the dynamic response of the numerical model in Fig. 8. As it may be observed, the estimated metamodel is capable of reproducing the dynamic response of the numerical model with excellent accuracy. More specifically, the prediction error sum of squares normalized by the sum of squares of the simulated dynamic response for the PE is equal to 0.1357% and the normalized simulation error sum of squares amounts to 1.3077% for the SE. It is worth noting that these results are obtained for an excitation signal for which the silent assumptions of $\omega'=-0.4160$, $\zeta_f=0.2514$, $t_{\text{mid}}=0.1296$, and $D_{5-95}=0.4080$ are no longer valid.

Finally, it should be added that the PC-NARX based simulated response was calculated more than 100 times faster than that of the FE model. A single simulation run of the PC-NARX metamodel for an excitation of 1000 samples requires approximately 2 seconds (mean value of 100 simulations in MATLAB on a PC with quad-core Xeon 3.5 GHz CPU, 8 GB RAM), while a simulation of the corresponding shear frame FEM on the same PC requires approximately 4 minutes. The reduction in computational time furnished by the proposed method is larger as the dimension of the reference FE model increases; naturally it is not uncommon to deal with structural FE models of thousands or even millions of DOF. However, it should be added that the estimation of the PC-NARX model based on the SE method is followed by additional effort required by the user for the PC-NARX structure selection and computational cost for the model estimation, which usually comprises a trial-and-error procedure and may not be directly quantified. This however takes place only once in a training stage of the overall analysis with the selected configuration kept unaltered for subsequent runs.

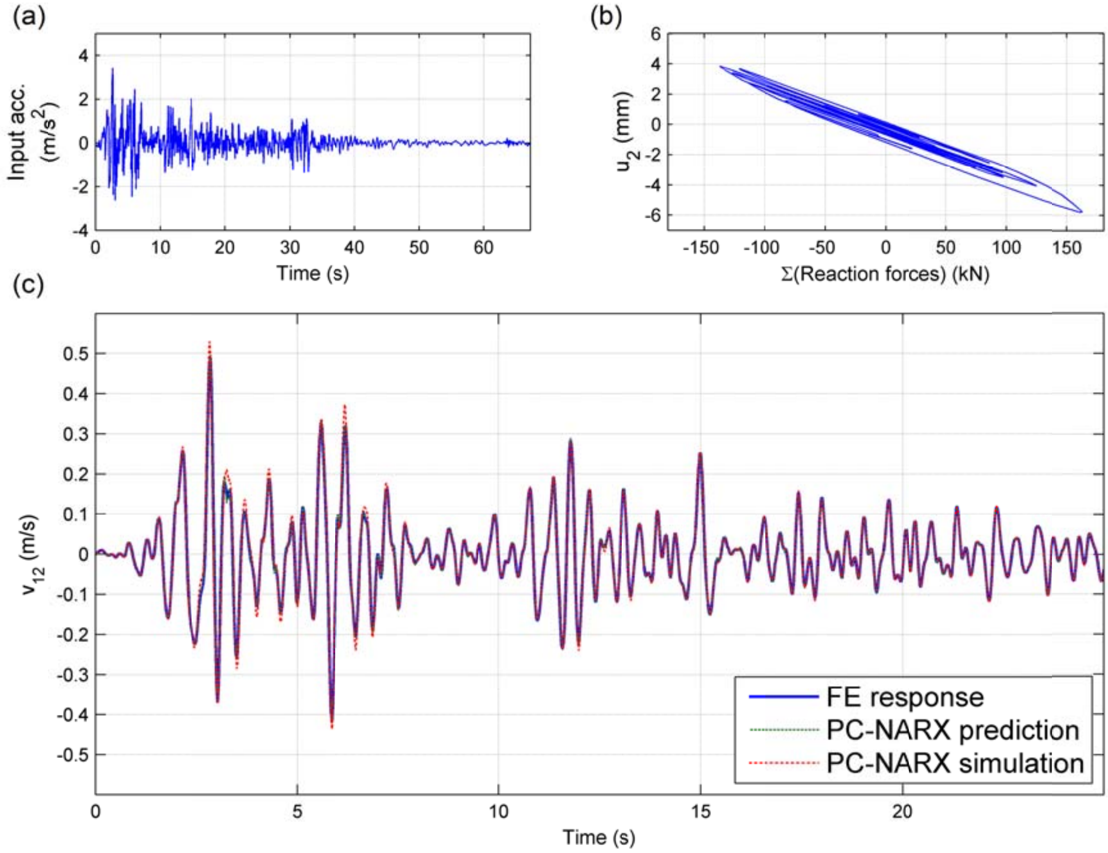

Fig. 8 (a) El Centro earthquake time history, (b) FE shear frame model displacement of the 1st floor (node 2) versus the total reaction force calculated at the base of the frame, and (c) top floor velocity response (v_{12}) as calculated by the FE shear frame model and the corresponding PC-NARX one-step-ahead prediction and simulation response signals

5. Conclusions

This work introduces a methodology for the metamodeling of large scale structural models with uncertain parameters and under stochastic excitation.

•The proposed metamodel, termed PC-NARX, is based on the fusion of the Polynomial Chaos expansion method with Nonlinear ARX models.

•The NARX model is able to account for the nonlinearity of the response for increased levels of input (earthquake) excitation.

•The proposed framework is suitable for incorporating uncertainties into the simulation, since the NARX model coefficients are treated as stochastic parameters, which are dependent upon the input random variables. These variables include structural properties as well as characteristics of the earthquake excitation. The aforementioned dependency is described via expansion on a properly constructed polynomial chaos basis.

•The framework introduced herein provides a reliable metamodel of complex engineering systems, with significantly reduced computational toll, applicable for purposes of system identification, SHM, control and uncertainty quantification.

When discussing applications in real-time control, this framework delivers an online model, which is able to reliably estimate structural response of complex systems even in the nonlinear range. Within a SHM framework, the proposed metamodel can substitute the full order FE models that are often utilized in time consuming inverse problem formulations. Within an uncertainty quantification framework, the proposed PC-NARX approach delivers a functional relationship between structural response and the uncertain input variables, which can be straightforwardly used for determining for instance the robustness of certain types of structures in earthquake prone regions. The method has been applied for the metamodeling problem of a five-storey shear frame FE model subject to synthetic earthquake ground motion, resulting in response of that lies outside the linear elastic region. Random variables are used to describe the uncertainties of both the material properties of the simulated structure and the input excitation signals. The PC-NARX model succeeds in reproducing the complex response of the system while integrating uncertainty into the simulation.

Acknowledgments

The research described in this paper was financially supported by the Federal Roads Office - FEDRO (Bundesamt für Strassen - ASTRA), Project #AGB 2012/015.

References

Adams, D. and Farrar, C. (2002), "Classifying linear and nonlinear structural damage using frequency domain arx models", *Struct. Hlth. Monit.*, **1**(2), 185-201.

Bathe, K.J. (2009), *Finite element method*, *Wiley Encyclopedia of Computer Science and Engineering*, Ed. Wah B, Wiley & Sons, Inc.

Beck, J.L. and Katafygiotis, L.S. (1998), "Updating models and their uncertainties. I: Bayesian statistical framework", *J. Eng. Mech.*, **124**(4), 455-461.

Blatman, G. and Sudret, B. (2010), "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis", *Prob. Eng. Mech.*, **25**(2), 183-197.

Caicedo, J., Dyke, S. and Johnson, E. (2004), "Natural excitation technique and eigensystem realization algorithm for phase I of the IASC-ASCE benchmark problem, Simulated data", *J. Eng. Mech.*, **130**(1), 49-60.

Chatzi, E. and Smyth, A.W. (2009), "The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing", *Struct. Control Hlth. Monit.*, **16**(1), 99-123.

Chen, S. and Billings, S.A. (1989), "Modelling and analysis of non-linear time series", *Int. J. Control*, **50**(6), 2151-2171.

Chen, Q., Worden, K., Peng, P. and Leung, A. (2007), "Genetic algorithm with an improved fitness function for (n)arx modeling", *Mech. Syst. Signal Pr.*, **21**(2), 994-1007.

Choudhury, S.M., Shah, S.L. and Thornhill, N.F. (2008), "Linear or nonlinear? A bicoherence-based measure of nonlinearity", Chap. 6, *Adv. Indust. Control*, Springer-Verlag, 77-91.

Christodoulou, K., Ntotsios, E., Papadimitriou C. and Panetsos, P. (2008), "Structural model updating and prediction variability using pareto optimal models", *Comput. Meth. Appl. Mech. Eng.*, **198**(1), 138-149.

Coley, D.A. (1999), An introduction to genetic algorithms for scientists and engineers, World Scientific,

Singapore.
Corigliano, A. and Mariani, S. (2004), "Parameter identification in explicit structural dynamics, performance of the extended Kalman filter", *Comput. Meth. Appl. Mech. Eng.*, **193**(36-38), 3807-3835.
Faravelli, L., Ubertini, F. and Fuggini, C. (2011), "System identification of a super high-rise building via a stochastic subspace approach", *Smart Struct. Syst.*, **7**(2), 133-152.
Farrar, C. and Worden, K. (2007), "Structural health monitoring - preface", *Philosoph. Transact. Roy. Soc. A*, **365**(1851), 299-301.
Fraraccio, G., Brügger, A. and Betti, R. (2008), "Identification and damage detection in structures subjected to base excitation", *Exper. Mech.*, **48**(4), 521-528.
Gholizadeh, S. and Salajegheh, E. (2009), "Optimal design of structures subjected to time history loading by swarm intelligence and an advanced metamodel", *Comput. Meth. Appl. Mech. Eng.*, **198**(37-40), 2936-2949.
Helton, J.C. and Davis, F.J. (2003), "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems", *Reliab. Eng. Syst. Saf.*, **81**(1), 23-69.
Hernandez, E.M. and Bernal, D. (2008), "State estimation in structural systems with model uncertainties", *J. Eng. Mech.*, **134**(3), 252-257.
Kalkan, E. and Chopra, A.K. (2010), "Practical guidelines to select and scale earthquake records for nonlinear response history analysis of structures", *Usgs open file report 2010-1068*, *126 pgs.*, U.S. Geological Survey, Menlo Park, CA.
Katkhuda, H., Martinez, R. and Haldar, A. (2005), "Health assessment at local level with unknown input excitation", *J. Struct. Eng.*, **131**(6), 956-965.
Kerschen, G., Worden, K., Vakakis, A. and Golinval, J. (2006), "Past, present and future of nonlinear system identification in structural dynamics", *Mech. Syst. Signal Pr.*, **20**(3), 505-592.
Kopsaftopoulos, F.P. and Fassois, S.D. (2013), "A functional model based statistical time series method for vibration based damage detection, localization, and magnitude estimation", *Mech. Syst. Signal Pr.*, **39**(1-2), 143-161.
Lin, J.W., Betti, R., Smyth, A.W. and Longman, R.W. (2001), "On-line identification of non-linear hysteretic structural systems using a variable trace approach", *Earthq. Eng. Struct. Dyn.*, **30**(9), 1279-1303.
Lourens, E., Papadimitriou, C., Gillijns, S., Reynders, E., De Roeck, G. and Lombaert, G. (2012), "Joint input-response estimation for structural systems based on reduced-order models and vibration data from a limited number of sensors", *Mech. Syst. Signal Pr.*, **29**, 310-327.
Moaveni, B., He, X., Conte, J.P. and Restrepo, J.I. (2010), "Damage identification study of a seven-story full-scale building slice tested on the UCSD-NEES shake table", *Struct. Safety*, **32**(5), 347-356.
Naets, F., Pastorino, R., Cuadrado, J., and Desmet W. (2013), "Online state and input force estimation for multibody models employing extended Kalman filtering", *Multibody Syst. Dyn.*, **July**, 1-20.
Piroddi, L. and Spinelli, W. (2003), "An identification algorithm for polynomial NARX models based on simulation error minimization", *Int. J. Control*, **76**(17), 1767-1781.
Poulimenos, A.G. and Fassois, S.D. (2006), "Parametric time-domain methods for non-stationary random vibration modelling and analysis - a critical survey and comparison", *Mech. Syst. Signal Pr.*, **20**(4), 763-816.
PEER (2012), Peer ground motion database, http://peer.berkeley.edu/peer ground motion database.
Rezaeian, S. and Der Kiureghian, A. (2010), "Simulation of synthetic ground motions for specified earthquake and site characteristics", *Earthq. Eng. Struct. Dyn.*, **39**(10), 1155-1180.
Rutherford, A., Park, G. and Farrar, C. (2007), "Nonlinear feature identification based on self-sensing impedance measurement for structural health assessment", *Mech. Syst. Signal Pr.*, **21**(1), 322-333.
Samara, P., Sakellariou, J., Fouskitakis, G., Hios, J. and Fassois, S. (2013), "Aircraft virtual sensor design via a time-dependent functional pooling narx methodology", *Aerospace Sci. Technol.*, **29**(1), 114-124.
Sapsis, T.P. and Lermusiaux, P.F.J. (2009), "Dynamically orthogonal field equations for continuous stochastic dynamical systems", *Physica D: Nonlin. Phenom.*, **238**(23), 2347-2360.
Smyth, A.W., Masri, S.F., Kosmatopoulos, E.B., Chassiakos, A.G. and Caughey, T.K. (2002),

"Development of adaptive modeling techniques for non-linear hysteretic systems", *Int. J. Nonlin. Mech.*, **37**(8), 1435-1451.
Soize, C. and Ghanem, R. (2004), "Physical systems with random uncertainties, chaos representations with arbitrary probability measure", *J. Scientific Comput.* SIAM, **26**(2), 395-410.
Spiridonakos, M. and Chatzi, E. (2012), "Metamodeling of structural systems through polynomial chaos arx models", *International Conference on Uncertainty in Structural Dynamics (USD2012)*, Leuven, Belgium.
Wagner, S.M. and Ferris, J.B. (2007), "A polynomial chaos approach to ARMA modeling and terrain characterization", *SPIE 6564, Modeling and Simulation for Military Operations II, 65640M*.
Wei, H.L, Billings, S.A. and Liu, J. (2004), "Term and variable selection for non-linear system identification", *Int. J. Control*, **77**(1), 86-110.
Worden, K. and Tomlinson, G. (2000), *Nonlinearity in Structural Dynamics, Detection, Identification and Modelling*, Taylor & Francis.
Vanik, M.W., Beck, J.L. and Au, S.K. (2000), "Bayesian probabilistic approach to structural health monitoring", *J. Eng. Mech.*, **126**(7), 738-745.
Yun, C.B. and Shinozuka, M. (1980), "Identification of nonlinear structural dynamics systems", *J. Struct. Mech.*, **8**(2), 187-203.

CN