

A metamodeling framework for multi-degree-of-freedom nonlinear dynamic structural systems subject to stochastic excitation

Bowei Li¹, Wei-Chu Chuang, Ph.D.², and Seymour M.J. Spence, Ph.D.³

¹Graduate Student, Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA. Email: jacklbw@umich.edu

²Postdoctoral Scholar, Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA. Email: wechuang@umich.edu

³Associate Professor, Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA (corresponding author). Email: smjs@umich.edu

ABSTRACT

The ever-growing reliance on probabilistic performance-based frameworks in assessing and designing structural systems is creating a need for efficient tools for propagating uncertainty through general nonlinear and dynamic structural systems. This research is focused on the development of metamodeling strategies for rapid response evaluation of a class of non-linear multi-degree-of-freedom (MDOF) structural systems driven by stochastic excitations. In particular, the nonlinear auto-regressive with exogenous input (NARX) model has been demonstrated to be versatile and effective in this respect. However, significant difficulties in NARX model calibration and execution have been encountered when directly applying this approach to practical MDOF systems with large numbers of degree-of-freedom. To overcome this limitation, a new metamodeling approach is proposed in this work through combining projection-based model order reduction with multi-input multi-output NARX models. The effectiveness and accuracy of the proposed approach are illustrated on a 40-story nonlinear steel-frame subject to stochastic earthquake excitation.

Keywords: Metamodeling; Nonlinear dynamic systems; Stochastic excitation; Multi-degree-of-freedom systems; Model order reduction; Stochastic dynamics

INTRODUCTION

The rapid rise of available computational power has made the Monte Carlo method, or more in general the stochastic simulation method, a widely used strategy for uncertainty quantification. Indeed, these approaches enable the direct estimation of the uncertain response characteristics of a wide variety of engineering problems and are often at the core of the frameworks developed to estimate the performance metrics used in state-of-the-art probabilistic performance-based design frameworks (Yang et al. 2009; FEMA P-58-1 2012; Chuang and Spence 2017; Ouyang and Spence 2020; Ouyang and Spence 2021). Notwithstanding this increase in available computational power, these approaches require a large number of model realizations in order to provide reliable response statistics and can easily become computationally cumbersome in the case of nonlinear dynamic systems. To overcome this computational difficulty, approaches based on metamodeling techniques have recently been explored for estimating the stochastic responses of dynamic systems (Lucor and Karniadakis 2004; Lucor et al. 2004; Kundu and Adhikari 2014; Gidaris et al. 2015; Mai et al. 2016; Mai and Sudret 2017; Bhattacharyya et al. 2020). In particular, researchers have recently developed a promising metamodeling approach based on the use of nonlinear autoregressive with exogenous input (NARX) models (Spiridonakos and Chatzi 2015; Mai et al. 2016). This approach has been successfully applied to various nonlinear single-degree-of-freedom (SDOF) systems. While this approach has been further applied to multi-degree-of-freedom (MDOF) systems (Spiridonakos and Chatzi 2015; Mai 2016), difficulties in calibration and accuracy have been observed (Mai 2016). It should also be noted that, even in the case of MDOF systems, the approach is based on a single-input single-output (SISO) formulation. Therefore, a separate metamodel is required for each output of a MDOF system. These limitations create the need for alternative metamodeling approaches for MDOF nonlinear dynamic systems.

To effectively evaluate the response of MDOF nonlinear structures subject to stochastic excitations, methods based on model order reduction (MOR) have been investigated (Grigoriu 2009; Grigoriu 2012; Gidaris and Taflanidis 2013; Jensen et al. 2016; Bamer et al. 2017; Tehrani et al. 2018; Patsialis and Taflanidis 2020). The basic idea of these approaches is to represent the full

system in a nonlinear reduced order subspace that preserves, with sufficient accuracy, the main dynamic behavior of the system. The possibility of combining MOR with metamodeling for replicating the behavior of nonlinear MDOF systems was recently investigated in (Chuang and Spence 2019). Despite the capability of efficiently replicating the time evolution of the system, the approach outlined in (Chuang and Spence 2019) can only be applied to a special class of nonlinear system with limited and concentrated sources of nonlinearity. The limitations of the approach lie in the use of a normal mode MOR, difficulties associated with identifying appropriate NARX models that do not have spurious model terms, and in the inability to capture response coupling between the coordinates of the reduced space.

To address these limitations, this paper proposes an advanced metamodeling approach for a more general class of MDOF systems. The approach is based on combining a proper orthogonal decomposition (POD) based model order reduction and a multi-input multi-output (MIMO) NARX model. In particular, the POD model order reduction converts the original system into a low-dimensional space, while the MIMO NARX model captures the dynamics of the reduced order system, including any coupling between the reduced coordinates. To calibrate the metamodel, a non-intrusive least angle regression with pruning (LARP) scheme is developed for model structure identification and an ordinary least square (OLS) method is implemented for coefficient determination. A case study consisting in a nonlinear steel frame subject to non-stationary stochastic earthquake excitation is presented to illustrate the efficiency and practicability of the proposed approach.

PROBLEM DEFINITION

A general n -dimensional MDOF dynamic structural system driven by stochastic excitation can be modeled through a mapping, $\mathcal{M}(\cdot) : \mathbb{R}^n \times \mathcal{T} \mapsto \mathbb{R}^n \times \mathcal{T}$, between the spaces of the stochastic input and output as:

$$\mathcal{M}(\ddot{\mathbf{x}}(t), \dot{\mathbf{x}}(t), \mathbf{x}(t)) = \mathbf{f}(t), \quad t \in \mathcal{T} \quad (1)$$

where $\ddot{\mathbf{x}}(t), \dot{\mathbf{x}}(t), \mathbf{x}(t) \in \mathbb{R}^n \times \mathcal{T}$ are the stochastic acceleration, velocity, and displacement output

vectors while $\mathbf{f}(t) \in \mathbb{R}^n \times \mathcal{T}$ is the input stochastic excitation vector.

To model the stochastic input, it is generally convenient to consider a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ defined by a sample space Ω , the σ -algebra $\mathcal{B}$ on Ω , and the probability measure $\mathbb{P}$. The stochasticity of the excitation $\mathbf{f}(t)$ can then be described by a vector random process $\{\mathbf{w} : \mathbf{w}(t) \in \Omega, t \in \mathcal{T}\}$, under the influence of which the excitation becomes $\mathbf{f}(\mathbf{w}, t)$. The focus of this work is to define a computationally tractable approach based on advanced metamodeling techniques for propagating the uncertainty in $\mathbf{f}(\mathbf{w}, t)$ through the system $\mathcal{M}$ when $\mathcal{M}$ is nonlinear.

THE PROPOSED APPROACH

This section outlines the proposed metamodeling approach together with a non-intrusive training scheme. The metamodeling approach combines a proper orthogonal decomposition-based model order reduction (POD-MOR) and MIMO NARX, which respectively extracts the underlying low-dimensional reduced-order model from the general system (Eq. 1) and captures the dynamics of the reduced-order model. In particular, the key step in the metamodeling approach is MIMO NARX training, which entails MIMO NARX structure determination and coefficient estimation. A least angle regression with pruning (LARP) scheme is proposed in this work for structure determination, while an ordinary least square (OLS) method is implemented for estimating the coefficients.

Model order reduction

Most structural systems of practical interest have a large number of DOFs, which not only increases the computational effort required for estimating structural responses but causes difficulties in applying metamodeling techniques to represent the system (Spiridonakos and Chatzi 2015; Chuang and Spence 2019). To overcome this issue, a MOR is used in this work for reducing the order of the system. This approach is based on approximating the response of the system through the following transformation:

$$\mathbf{x}(t) \approx \mathbf{\Phi}_{n_r} \mathbf{q}(t) \quad (2)$$

where $\mathbf{\Phi}_{n_r}$ is an appropriate $n \times n_r$ coordinate transformation matrix with $n_r \ll n$, while $\mathbf{q}(t) \in \mathbb{R}^{n_r} \times \mathcal{T}$ is the response vector in the reduced space. From the above transformation, Eq. (1) can be

written in the following n_r -dimensional reduced-order form:

$$\mathbf{\Phi}_{n_r}^T \mathcal{M}(\mathbf{\Phi}_{n_r} \ddot{\mathbf{q}}(t), \mathbf{\Phi}_{n_r} \dot{\mathbf{q}}(t), \mathbf{\Phi}_{n_r} \mathbf{q}(t)) = \mathbf{p}(t; \mathbf{w}). \quad (3)$$

where $\mathbf{p}(t; \mathbf{w}) = \mathbf{\Phi}_{n_r}^T \mathbf{f}(t; \mathbf{w}) \in \mathbb{R}^{n_r} \times \mathcal{T}$ is the excitation in the reduced space. For simplicity, $\mathbf{p}(t; \mathbf{w})$ will be denoted in the following as $\mathbf{p}(t)$.

In this work, the coordinates transformation matrix, $\mathbf{\Phi}_{n_r} \in \mathbb{R}^{n \times n_r}$, is obtained through proper orthogonal decomposition (POD), an unsupervised learning approach that extracts principal components, or basis functions, from a set of known data (Holmes et al. 1996). To this end, various approaches have been proposed including, the method of Lagrangian multipliers (Volkwein 2013), eigen-decomposition (Kerschen and Golinval 2002; Volkwein 2013), and singular value decomposition (SVD) (Kerschen and Golinval 2002; Volkwein 2013). In general, the most widely used approach is SVD and is also adopted in this work. To apply this approach to the problems of interest to this work, it is first necessary to directly evaluate the full system of Eq. (1) for n_s samples of the stochastic excitation $\mathbf{f}(\mathbf{w}_i, t)$. From the output of Eq. (1), the following discrete time snapshot matrix can be defined:

$$\mathbf{X} = [\mathbf{x}_1(t_1), \dots, \mathbf{x}_1(t_{n_t}), \dots, \mathbf{x}_{n_s}(t_1), \dots, \mathbf{x}_{n_s}(t_{n_t})], \quad \mathbf{X} \in \mathbb{R}^{n \times n_t n_s} \quad (4)$$

where n_t is the total number of discrete time steps, i.e. snapshots, considered for each of the n_s samples. In general, n_t can be taken as a subset of the total number of time steps evaluated in solving Eq. (1). In generating $\mathbf{X}$, it is important to ensure that the snapshots, i.e. $\mathbf{x}_i(t_j)$ for $i = 1, \dots, n_s$ and $j = 1, \dots, n_t$, are capable of capturing not only the nonlinear behavior of the system, but also the stochasticity of the excitation. The snapshot matrix, $\mathbf{X}$, can then be decomposed through SVD as (Holmes et al. 1996):

$$\mathbf{X} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T \quad (5)$$

where $\mathbf{U}$ is a $n \times n$ orthonormal matrix containing the left singular vectors of $\mathbf{X}$, $\mathbf{V}$ is the $(n_s n_t) \times (n_s n_t)$

orthonormal matrix of the corresponding right singular vectors, while Λ is a $n \times (n_s n_t)$ pseudo-diagonal matrix containing the singular values with $\Lambda(j, j) = \lambda_j \in \mathbb{R}^+$ the j th largest singular value for $j = 1, 2, \dots, n$. In general, the size of V is extremely large as $n_s n_t \gg n$. Hence, a more computationally effective economy-size SVD is adopted in this work, in which only the first n columns of V and the first n singular values are estimated. As outlined in (Kerschen and Golinval 2002; Volkwein 2013), the left singular vectors, U , are the POD modes.

In defining the coordinates transformation matrix, Φ_{n_r} , for the reduced-order system, the first n_r POD modes, and so columns of U , are considered, i.e., $\Phi_{n_r} = [U_1, U_2, \dots, U_{n_r}]$. In terms of the snapshot matrix X , this corresponds to the following approximation:

$$X \approx \Phi_{n_r} \Lambda_{\text{Trunc}} V_{\text{Trunc}}^T + \epsilon_X \quad (6)$$

where Λ_{Trunc} is the diagonal matrix of the first n_r singular values, V_{Trunc} is the first n_r columns of V ; while ϵ_X is the error process given $\text{tr}(\epsilon_X \epsilon_X^T) = \sum_{j=n_r+1}^n \lambda_j^2$ (Volkwein 2013). As a trade-off between accuracy and computational efficiency, n_r can be chosen by ensuring the energy captured in the truncated representation of X , $\sum_{j=1}^{n_r} \lambda_j^2$, is not less than η of the total energy, $\text{tr}(X X^T) = \sum_{j=1}^n \lambda_j^2$, i.e., $\sum_{j=1}^{n_r} \lambda_j^2 \geq \eta \sum_{j=1}^n \lambda_j^2$, where η is typically assumed to be close to 1, e.g., 0.99 (Bamer and Markert 2017). A properly chosen η can bring significant dimensional reduction to the system, leading to a considerable reduction in dimensions from Eq. (1) to Eq. (3), i.e. $n_r \ll n$.

The MIMO NARX metamodel

Overview

Despite the computational savings gained through model order reduction, Eq. (3) is still a coupled nonlinear dynamic equation that must be solved through computationally intensive numerical integration schemes, e.g. Newmark or Runge-Kutta methods. It should also be observed that, in general, the nonlinear model, $\mathcal{M}$, still requires evaluation in the full space at each time step. In other words, the computational gains associated with directly integrating the reduced system of Eq. (3), are related mainly to the possibility of choosing a much larger time step as compared to

that used in integrating the full system. To overcome this computational barrier, the idea that will be explored in this work is to develop a non-intrusive metamodel, based on MIMO NARX, of the reduced space that, once calibrated, does not require the evaluation of the full model at each time step.

Under the assumption that the current output, $q(t_i)$, of a nonlinear single-degree-of-freedom dynamic system depends on its past output values, $[q(t_i - \Delta t), \dots, q(t_i - n_q \Delta t)]$ with n_q the maximum output lag and Δt the time step size, and current and past load inputs, $[p(t_i), p(t_i - \Delta t), \dots, p(t_i - n_f \Delta t)]$ with n_f the maximum load lag, the nonlinear dynamic behavior of the system can be captured through the following NARX model:

$$q(t_i) = G(p(t_i), p(t_i - \Delta t), \dots, p(t_i - n_f \Delta t), q(t_i - \Delta t), \dots, q(t_i - n_q \Delta t)) + \epsilon(t_i) \quad (7)$$

where $G(\cdot) : \mathbb{R}^{n_f + n_q + 1} \mapsto \mathbb{R}$ is the mapping from the recent inputs and outputs to the current output, and $\{\epsilon : \epsilon(t_i) \sim \mathcal{N}(0, \sigma_{\epsilon(t_i)}^2)\}$ is the error process which is generally assumed as a Gaussian process (Leontaritis and Billings 1987). Under the assumption that the dependence between the coordinates, $\mathbf{q}(t)$, of the reduced space is negligible, SISO NARX models can be applied to MDOF systems (Chuang and Spence 2019). This model, however, is incapable of capturing the inevitable response coupling between the reduced coordinates for general nonlinearity. To overcome this limitation, the possibility of applying a MIMO NARX (Billings et al. 1989) strategy is explored in this work as a means to capture nonlinear and coupled dynamic behavior of Eq. (3). The general form of the MIMO NARX model is:

$$\mathbf{q}(t_i) = \mathbf{G}(\mathbf{z}(t_i)) + \boldsymbol{\epsilon}(t_i) \quad (8)$$

where $\mathbf{G}(\cdot) : \mathbb{R}^{(n_f + n_q + 1)n_r} \mapsto \mathbb{R}^{n_r}$ is the MIMO NARX model to be identified, $\mathbf{z}(t_i) = [\mathbf{p}^T(t_i), \mathbf{p}^T(t_i - \Delta t), \dots, \mathbf{p}^T(t_i - n_f \Delta t), \mathbf{q}^T(t_i - \Delta t), \dots, \mathbf{q}^T(t_i - n_q \Delta t)]$ is the regression vector of current and past input and output values, and $\boldsymbol{\epsilon} : \boldsymbol{\epsilon}(t_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}(t_i)})$ is a vector-valued Gaussian error process.

A common structure for $\mathbf{G}(\cdot)$, and that will be considered in this work, is the following linear-

in-the-parameter form:

$$\mathbf{q}(t_i) = \mathbf{\Theta}^T \mathbf{g}(\mathbf{z}(t_i)) + \epsilon(t_i) \quad (9)$$

where $\mathbf{g}(\cdot) = [\mathbf{g}_1^T(\cdot), \mathbf{g}_2^T(\cdot), \dots, \mathbf{g}_{n_r}^T(\cdot)]^T$ is the vector collecting all n_r NARX model terms $\mathbf{g}_j(\cdot)$, in which $\mathbf{g}_j(\cdot) : \mathbb{R}^{(n_f+n_q+1)n_r} \mapsto \mathbb{R}^{l_j}$ is a l_j -dimensional function of the regression vector $\mathbf{z}(t)$ for the j th reduced coordinate; while $\mathbf{\Theta} = \text{diag}[\mathbf{\Theta}_1, \dots, \mathbf{\Theta}_{n_r}]$ is a block-diagonal matrix collecting the NARX coefficients of the n_r DOFs of the reduced system.

The LARP scheme

Model identification

In general, the identification of the MIMO NARX metamodel entails structure determination, i.e. selecting NARX terms, and coefficient calibration. An efficient approach based on implementing the least angle regression (LARs) algorithm for structure determination and the ordinary least square (OLS) method for coefficient calibration has been proposed for identifying the NARX model of SISO systems (Mai et al. 2016). In this work, the basic idea underpinning this approach is extended for the identification of the MIMO NARX model of Eq. (9). To this end, consider the following form for the j th reduced coordinate:

$$q_j(t_i) = \mathbf{\Theta}_j^T \mathbf{g}_j(\mathbf{z}(t_i)) + \epsilon_j(t_i) \quad (10)$$

The first step towards calibrating the MIMO NARX model is to obtain a set of potential NARX terms/features $\mathbf{g}_j^p(\mathbf{z}(t))$ for each reduced coordinate based on a pre-designated form of basis function (e.g. polynomial (Leontaritis and Billings 1987), rational (Billings and Chen 1996), wavelet (Wei and Billings 2004), neural network (Billings and Chen 1996)), and maximum time delays n_f and n_q . The potential NARX feature matrix $\mathbf{Z}_j^p$ of the j th reduced coordinate can then be written in the

following discrete form:

$$\mathbf{Z}_j^p = \begin{bmatrix} \mathbf{g}_j^p(\mathbf{z}(t_1))^T \\ \mathbf{g}_j^p(\mathbf{z}(t_2))^T \\ \dots \\ \mathbf{g}_j^p(\mathbf{z}(t_{n_t}))^T \end{bmatrix} \quad (11)$$

where $t_1, t_2, \dots, t_i, \dots, t_{n_t}$ is the discrete time sequence while $\mathbf{g}_j^p(\mathbf{z}(t_i))$ contains $l_s \geq l_j$ potential NARX features. It is important to note that the regression vector $\mathbf{z}(t_i)$ of the MIMO NARX model contains input and output values from all reduced coordinates, in contrast to the SISO NARX model, where only terms of the j th reduced coordinate are considered. This enables the coupling between the reduced coordinates to be captured.

The LARs algorithm (Efron et al. 2004) can then be employed to select the most relevant NARX features in $\mathbf{Z}_j^p$ by computing the correlation of each potential model term to the system output (Billings et al. 1989), leading to a candidate NARX model term that contains a subset of the potential NARX features. By simulating over n_s samples, a total of $n_{c,j}$, where $n_{c,j} \leq n_s$, unique candidate model terms will be identified for the j th reduced coordinate of the system. For the k th identified candidate model, the corresponding NARX coefficients can then be estimated by the following OLS method that minimizes the one-step-ahead prediction error:

$$\Theta_{j,k} = \arg \min_{\Theta_{j,k}} e_{\text{PE},j,k} = [\mathbf{Z}_{j,k}^T \mathbf{Z}_{j,k}]^{-1} \mathbf{Z}_{j,k}^T \mathbf{Q}_j^T \quad (12)$$

where $\mathbf{Z}_{j,k}$ for $k = 1, 2, \dots, n_{c,j}$ is the candidate feature matrix of the k th candidate model, while $\mathbf{Q}_j$ is the response of the j th reduced coordinate. The prediction error $e_{\text{PE},j,k}$ can be defined as (Chuang and Spence 2019; Mai et al. 2016):

$$e_{\text{PE},j,k} = \frac{\left\| \mathbf{Q}_j^T - \mathbf{Z}_{j,k} \Theta_{j,k} \right\|^2}{\left\| \mathbf{Q}_j^T - \boldsymbol{\iota} \mathbb{E}_t[\mathbf{Q}_j] \right\|^2} \quad (13)$$

where $\boldsymbol{\iota}$ is a all-ones vector, while $\mathbb{E}_t[\mathbf{Q}_j]$ is the expected value (in a time average sense) of the

response time series $\mathbf{Q}_j$. The prediction error $e_{\text{PE},j,k}$ measures one-step-ahead error, i.e. the error of the current output given that $\mathbf{z}(t)$, or the recent outputs and inputs, are perfectly accurate (one-step-ahead prediction).

Once the candidate NARX model terms and the associated coefficients are determined, the next step is to select the most appropriate MIMO NARX model from the candidates for representing the system of interest. An appropriate error measure must be defined to this end. Since the goal of metamodeling is to reproduce the whole time history with only inputs and a few initial conditions, the MIMO NARX model must run recursively to generate the entire time history. The prediction error criterion, however, is incapable of taking into account error accumulation during the recursive process, making it unsuitable for model selection to be used defining metamodels. To avoid such issues, the simulation error criterion can be employed to estimate the error produced by recursively running the model. This approach, nevertheless, requires NARX models of all reduced coordinates to run simultaneously. Given that each reduced coordinate has $n_{c,j}$ candidate NARX models, the total number of candidate MIMO NARX models, $\prod_{j=1}^{n_r} n_{c,j}$, can become extremely large. In this work, it is proposed to overcome this issue by decoupling the identification of the NARX models for each reduced coordinate. This is achieved by considering the following form for the NARX model of the j th reduced coordinate during identification:

$$\tilde{q}_{j,k}(t_i) = \mathbf{\Theta}_{j,k}^T \mathbf{g}_{j,k}(\tilde{\mathbf{z}}(t_i)) \quad (14)$$

where $\tilde{\mathbf{z}}(t_i)$ is the following modified regression vector:

$$\begin{aligned} \tilde{\mathbf{z}}(t_i) = & [\mathbf{p}^T(t_i), \mathbf{p}^T(t_i - \Delta t), \dots, \mathbf{p}^T(t_i - n_f \Delta t), \\ & q_1(t_i - \Delta t), \dots, q_{j-1}(t_i - \Delta t), \tilde{q}_{j,k}(t_i - \Delta t), q_{j+1}(t_i - \Delta t), \dots, q_{n_r}(t_i - \Delta t), \\ & \dots, \\ & q_1(t_i - n_q \Delta t), \dots, q_{j-1}(t_i - n_q \Delta t), \tilde{q}_{j,k}(t_i - n_q \Delta t), q_{j+1}(t_i - n_q \Delta t), \dots, q_{n_r}(t_i - n_q \Delta t)] \end{aligned}$$

in which only the responses of the j th reduced coordinate, $\tilde{q}_{j,k}(t_i - \Delta t), \dots, \tilde{q}_{j,k}(t_i - n_q \Delta t)$, are estimated from the MIMO NARX model. The responses of all other reduced coordinates are directly obtained from the high-fidelity data therefore decoupling the identification of the j th NARX model from the identification of the other NARX models without losing the effects of reduced coordinate interdependence. The corresponding error measure $\tilde{e}_{\text{SE},j,k}$ is defined as:

$$\tilde{e}_{\text{SE},j,k} = \frac{\|\mathbf{Q}_j - \tilde{\mathbf{Q}}_{j,k}\|^2}{\|\mathbf{Q}_j^T - \mathbf{I} \mathbb{E}_t[\mathbf{Q}_j]\|^2} \quad (15)$$

where $\tilde{\mathbf{Q}}_{j,k} = [\tilde{q}_{j,k}(t_1), \tilde{q}_{j,k}(t_2), \dots, \tilde{q}_{j,k}(t_{n_t})]$. By simulating over n_s samples, the accuracy of the k th candidate model can be measured by the expected error measure, $\tilde{\tilde{e}}_{\text{SE},j,k}$, from which the optimal NARX model can be determined.

In selecting the optimal NARX model, a simpler model with less terms are generally preferred as spurious NARX terms and features have been found to not only cause deleterious effects on the accuracy of the model, e.g. over-fitting, but also induce spurious dynamics (Billings 2013; Piroddi and Spinelli 2003; Mai et al. 2016). The optimal model, i.e. model terms $\mathbf{g}_j(\cdot)$ and associated coefficients $\mathbf{\Theta}_j$, is therefore chosen as the NARX model with the least number of NARX terms that achieves a sufficiently small overall error, i.e. $\tilde{\tilde{e}}_{\text{SE},j,k} \leq \tilde{E}$, where $\tilde{E}$ is a predefined threshold value. The aforementioned process is then carried out for each reduced coordinate over all samples. The final MIMO NARX metamodel is given by:

$$\hat{\mathbf{q}}(t) = \tilde{\mathbf{\Theta}}^T \mathbf{g}(\hat{\mathbf{z}}(t)) \quad (16)$$

where $\hat{\mathbf{z}}(t) = [\mathbf{p}^T(t), \mathbf{p}^T(t - \Delta t), \dots, \mathbf{p}^T(t - n_f \Delta t), \hat{\mathbf{q}}(t - \Delta t)^T, \dots, \hat{\mathbf{q}}(t - n_q \Delta t)^T]$ in which output feedback from all reduced coordinates is considered, and $\tilde{\mathbf{\Theta}}$ is the expected value of $\mathbf{\Theta}$ over all n_s samples.

Model refinement

The MIMO NARX model of Eq. (16), however, can still include spurious NARX terms even though the simplest model is selected. This problem originates from the LARs approach, which selects candidate terms based on correlation analysis that is not necessarily a reflection of the contribution of a term to the model (Piroddi and Spinelli 2003). To address this issue, an iterative simulation error based pruning procedure (Piroddi and Spinelli 2003) is introduced in this work to identify and remove the spurious NARX terms. In particular, starting from the model of Eq. (16), the procedure progressively identifies and deletes the most deleterious NARX term at each iteration until an predefined error tolerance is met. Within each iteration, a set of trial models is first generated, with each of them obtained by removing one unique term from the current MIMO NARX model, and then compared with the current model. The coefficients associated with each trial model are estimated by the OLS approach. To assess the performance of trial models, a MIMO NARX simulation error measure is defined for a user-defined DOF of interest, as follows:

$$\hat{e}_{SE,m} = \frac{\|X_m - \Phi_{n_r}^m \hat{Q}\|^2}{\|X_m - \mathbb{E}[X_m]\|^2} \quad (17)$$

where m is the DOF of interest, X_m is the m th row of the snapshot matrix X of Eq. (4) (i.e. the response of the m th DOF), $\Phi_{n_r}^m$ is the m th row of Φ_{n_r} , $\hat{Q} = [\hat{q}(t_1), \hat{q}(t_2), \dots, \hat{q}(t_{n_t})]$. The error describes the goodness of the MIMO NARX model in reproducing the response of the m th DOF in the physical/full space. The performance of the current MIMO NARX model in each iteration can be evaluated by taking the expectation over all training samples, $\bar{e}_{SE,m}$. Similarly, the error measure for each trial model can be calculated from the expectation over all samples, denoted as $\bar{e}'_{SE,m}$. The deleterious effect of removing each term is then evaluated by calculating the error increase of each trial model against the current MIMO NARX model i.e. $\bar{e}'_{SE,m} - \bar{e}_{SE,m}$. Within each iteration, the current MIMO NARX model is then replaced by the optimal trial model with $\min\{\bar{e}'_{SE,m} - \bar{e}_{SE,m}\}$. The pruning process then proceeds to the next iteration with the new MIMO NARX model serving as the current model and reevaluates until a user-defined error change threshold $\hat{E}$ is satisfied, i.e.,

$$\tilde{e}'_{SE,m} - \tilde{e}_{SE,m} \geq \hat{E}.$$

The algorithm and overall procedure

The LARP algorithm of this section is outlined in Algorithm 1. The data flow associated with the algorithm is schematically illustrated in Fig. 1(a), while in Fig. 1(b) a flowchart illustrates the three main phases of the algorithm, which can be summarized as follows:

- Phase 1: Data collection and model order reduction
 - Generate n_s samples of the stochastic excitation $f(\mathbf{w}_i, t)$, solve Eq. (1) for the high-fidelity response samples, and define the snapshot matrix $\mathbf{X}$.
 - Estimate the coordinate transformation matrix Φ_{n_r} through SVD on $\mathbf{X}$. Solve the reduced-order model of Eq. (3) therefore defining n_s reduced-order input $\mathbf{P}_i$ and output $\mathbf{Q}_i$ samples.
- Phase 2: LARs based MIMO NARX training
 - For each reduced coordinate, propose a set of potential NARX terms (e.g. polynomial, rational, wavelet, neural network). Loop over all training samples to identify the most relevant NARX terms via the LARs algorithm, and estimate the NARX coefficients by OLS. Keep the $n_{c,j}$ unique candidate NARX models over all samples.
 - Estimate the error measure $\bar{e}_{SE,j,k}$ for all candidate NARX models of Eq. (14). Keep the most appropriate NARX models, in the sense of both accuracy and simplicity, and define the MIMO NARX model.
- Phase 3: MIMO NARX pruning
 - Apply the simulation error based pruning procedure to identify and remove the unnecessary terms from the MIMO NARX iteratively, therefore defining the final MIMO NARX metamodel.

Algorithm 1: Least angle regression with pruning algorithm

Data: $\mathbf{Q}_i, \mathbf{P}_i$ for $i = 1, 2, \dots, n_s, \mathbf{\Phi}_{n_r}$
Result: $\hat{\mathbf{r}}, \bar{\mathbf{\Theta}}$
 // I.1 LARs to identify relevant features
 Set errors threshold $\bar{E}$ and user defined DOF m ;
for $j \leftarrow 1$ **to** n_r **do**
 for $i \leftarrow 1$ **to** n_s **do**
 Construct $\mathbf{Z}_{i,j}^p$ by Eq. (11);
 LARs (Efron et al. 2004) select relevant features (column indices $\mathbf{r}_{i,j}$);
 if $\mathbf{r}_{i,j}$ is unique **then**
 $\mathbf{r}_{j,k} \leftarrow \mathbf{r}_{i,j}$;
 $\mathbf{Z}_{i,j,k} \leftarrow \mathbf{Z}_{i,j}^p(:, \mathbf{r}_{i,k})$;
 end
end
 // I.2 Select the NARX model
for $j \leftarrow 1$ **to** n_r **do**
 for $k \leftarrow 1$ **to** $\tilde{n}_c(j)$ **do**
 for $i \leftarrow 1$ **to** n_s **do**
 $\mathbf{\Theta}_{i,j,k} \leftarrow [\mathbf{Z}_{i,j,k}^T \mathbf{Z}_{i,j,k}]^{-1} \mathbf{Z}_{i,j,k}^T \mathbf{Q}_{i,j}^T$; // Eq. 12
 for $t \leftarrow t_0$ **to** t_{final} **do**
 $\tilde{\mathbf{g}}_j \leftarrow \mathbf{g}_j([\mathbf{p}^T(t), \dots, \mathbf{p}^T(t - n_f \Delta t), q_1(t - \Delta t), \dots, \tilde{\mathbf{Q}}_{i,j,k}(t - \Delta t), \dots, q_{n_r}(t - \Delta t),$
 $\dots, q_1(t - n_q \Delta t), \dots, \tilde{\mathbf{Q}}_{i,j,k}(t - n_q \Delta t), \dots, q_{n_r}(t - n_q \Delta t)])$;
 $\tilde{\mathbf{Q}}_{i,j,k}(t) \leftarrow \mathbf{\Theta}_{i,j,k}^T \tilde{\mathbf{g}}_j(\mathbf{r}_{j,k})$; // Eq. 14
 end
 $\tilde{e}_{SE}(i, j, k) \leftarrow \|\tilde{\mathbf{Q}}_{i,j} - \tilde{\mathbf{Q}}_{i,j,k}\| / \|\tilde{\mathbf{Q}}_{i,j} - \iota \mathbb{E}[\tilde{\mathbf{Q}}_{i,j}]\|$; // Eq. 15
 end
 $\tilde{\bar{e}}_{SE}(j, k) \leftarrow \mathbb{E}_i[\tilde{e}_{SE}(i, j, k)]$;
 end
 $k_{opt,j} \leftarrow \arg \min_{k \in \{k: \tilde{\bar{e}}_{SE}(j, k) \geq \bar{E}\}} \text{cardinality}(\mathbf{r}_{j,k})$; // Accurate & simplest
end
 $\bar{\mathbf{\Theta}}_{opt} \leftarrow \text{diag}\{\mathbb{E}_i[\mathbf{\Theta}_{i,1,k_{opt,1}}] \dots \mathbb{E}_i[\mathbf{\Theta}_{i,n_r,k_{opt,n_r}}]\}$;
 Collect $\mathbf{r}_{j,k_{opt,j}} + \sum_{j_j=1}^{j-1} l_{j_j}$ for all j into $\hat{\mathbf{r}}$;
 // I.3 Run the NARX model over the training data set
for $i \leftarrow 1$ **to** n_s **do**
 for $t \leftarrow t_0$ **to** t_{final} **do**
 $\hat{\mathbf{g}} = \mathbf{g}([\mathbf{p}^T(t), \mathbf{p}^T(t - \Delta t), \dots, \mathbf{p}^T(t - n_f \Delta t), \hat{\mathbf{Q}}_i(t - \Delta t)^T, \dots, \hat{\mathbf{Q}}_i(t - n_q \Delta t)^T])$;
 $\hat{\mathbf{Q}}_i(t) \leftarrow \bar{\mathbf{\Theta}}_{opt}^T \hat{\mathbf{g}}(\hat{\mathbf{r}})$; // Eq. 16
 end
 $\hat{e}_{SE,m}(i) \leftarrow \|\mathbf{X}_{i,m} - \mathbf{\Phi}_{n_r}^m \hat{\mathbf{Q}}_i\|^2 / \|\mathbf{X}_{i,m} - \iota \mathbb{E}[\mathbf{X}_{i,m}]\|^2$; // Eq. 17
end
 $\bar{\hat{e}}_{SE,m} = \mathbb{E}_i[\hat{e}_{SE,m}]$; // The error of the NARX model

```

// II. Pruning phase
Set error threshold  $\hat{E} \in \mathbb{R}$ ;
while  $\tilde{e}'_{SE,j} - \tilde{e}_{SE,j} \leq \hat{E}$  do
    for  $i_{term} \leftarrow 1$  to  $cardinality(\hat{r})$  do
         $\hat{r}' \leftarrow \hat{r} \setminus \hat{r}(i_{term})$ ; // Remove  $i_{term}$ th term
        for  $j \leftarrow 1$  to  $n_r$  do
            for  $i \leftarrow 1$  to  $n_s$  do
                 $\Theta'_{i,j} \leftarrow [Z^T(:, \hat{r}')Z(:, \hat{r}')]^{-1} Z^T(:, \hat{r}')Q_{i,j}^T$ ;
            end
        end
         $\bar{\Theta}_{i_{term}} \leftarrow \text{diag}\{\mathbb{E}_i[\Theta'_{i,1}] \dots \mathbb{E}_i[\Theta'_{i,n_r}]\}$ ;
        for  $i \leftarrow 1$  to  $n_s$  do
            for  $t \leftarrow t_0$  to  $t_{final}$  do
305          $\hat{g} \leftarrow g([p^T(t), \dots, p^T(t - n_f \Delta t), \hat{q}(t - \Delta t)^T, \dots, \hat{q}(t - n_q \Delta t)^T])$ ;
                 $\hat{Q}'_i(t) \leftarrow \bar{\Theta}_{i_{term}}^T \hat{g}(\hat{r}')$ ;
            end
             $\hat{e}'_{SE,m}(i) \leftarrow \|X_{i,m} - \Phi_{n_r}^m \hat{Q}'_i\|^2 / \|X_{i,m} - \iota \mathbb{E}[X_{i,m}]\|^2$ ;
        end
         $\tilde{e}'_{SE,m}(i_{term}) = \mathbb{E}_i[\hat{e}'_{SE,m}]$ ;
    end
    if  $\min(\tilde{e}'_{SE,m}) - \tilde{e}_{SE,m} \leq \hat{E}$  then
         $i_{delete} \leftarrow \text{the index of } \min(\tilde{e}'_{SE,m}) \text{ in } \tilde{e}'_{SE,m}$ ;
         $\hat{r} \leftarrow \hat{r} \setminus \hat{r}(i_{delete})$ ;
         $\bar{\Theta} = \bar{\Theta}_{i_{delete}}$ ;
         $\tilde{e}_{SE,m} \leftarrow \tilde{e}'_{SE,m}(i_{delete})$ 
    end
end

```

CASE STUDY

In this section, the proposed metamodeling approach is illustrated on the 2D steel frame of Fig. 2 that was extracted from the archetype structure outlined in Hutt et al. (2016). The inter-story heights are 6.1 m for the 1st floor and 3.9 m for the others, leading to a total height of 154.7 m. The influence width of the frame is taken as 12.2 m. The steel frame consists of AISC (American Institute of Steel Construction) wide flange beams with 6.1 m spans and square box columns. Table 1 reports a summary of the section sizes used in defining the frame, as suggested in Hutt et al. (2016). In addition to the self-weight of the structure, a carried weight of $13.53h_j$ [kN/m²], where h_j is the story height, was considered for each floor.

Structure and earthquake modeling

The structure was considered as a shear building model, where the mass at each floor was assumed to be lumped at its center and was directly calculated from the corresponding self and carried weight. The inherent damping was modeled as Rayleigh damping, with damping ratios of the 1st and 2nd modes equal to 1.5%. The nonlinearity was lumped at each floor and modeled by the following Bouc-Wen model:

$$\begin{cases} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{B}^T \mathbf{K} \mathbf{y}(t) = \mathbf{M} \boldsymbol{\iota} \ddot{w}(w, t) \\ \dot{\mathbf{y}}(t) = \mathbf{B} \dot{\mathbf{x}}(t) - \alpha |\mathbf{B} \dot{\mathbf{x}}(t)| \circ \mathbf{y}(t) \end{cases} \quad (18)$$

where $\mathbf{M}$ and $\mathbf{C}$ are the $n \times n$ mass and damping matrices of the system; $\mathbf{K}$ is a diagonal $n \times n$ matrix collecting the lateral stiffness at each floor; $\boldsymbol{\iota}$ is a $n \times 1$ vector of ones; $\ddot{w}(w, t)$ is the base acceleration with stochasticity defined by the white noise process $\{w : w(t) \sim \text{i.i.d. } \mathcal{N}(0, 1), t \in [t_0, t_{\text{final}}]\}$; $\mathbf{y}$ and $\dot{\mathbf{y}}(t)$ are the non-observable hysteretic parameter and its first derivative; $\mathbf{B}$ is the global displacements to inter-story drifts transformation matrix; $|\cdot|$, $\circ$ are element-wise absolute value and multiplication; and α is a Bouc-Wen nonlinearity parameter taken as 10. In particular, the choice of $\alpha = 10$ was made so as to produce a nonlinear response similar to that reported in Hutt et al. (2016) from which the frame was extracted.

The structure was assumed to be located in downtown San Francisco with subsurface ground conditions consistent with Site Class D (ASCE/SEI 7-16 2017) and subjected to a 10% exceedance probability in 50 years ground motion hazard (Hutt et al. 2016). Synthetic ground motions were generated from the model proposed by Rezaeian and Der Kiureghian (2010) with target spectrum constructed from the USGS unified hazard tool. This model assumes that the ground motion $\tilde{w}(t)$ is a time-modulated and filtered white noise process, as follows:

$$\tilde{w}(t) = A(t, \alpha) \frac{1}{\sqrt{\sum_{\tau=t_0}^{t_{\text{final}}} h^2(t - \tau, \kappa(\tau))}} \sum_{\tau=t_0}^{t_{\text{final}}} [h(t - \tau, \kappa(\tau)) w(\tau)] \quad (19)$$

where $A(t, \alpha)$ is a time modulating function defining the temporal characteristics; $h(t, \kappa(\tau))$ is the impulse-response function of the time-varying filter. In particular, the time modulating function, $A(t, \alpha)$, is assumed to be of the following gamma type:

$$A(t, \alpha) = \alpha_1 t^{\alpha_2 - 1} \exp(-\alpha_3 t) \quad (20)$$

where $\alpha = [\alpha_1, \alpha_2, \alpha_3]$; $\alpha_1 \in \mathbb{R}^+$ controls the intensity; $\alpha_2 \in (1, +\infty)$ controls shape; $\alpha_3 \in \mathbb{R}^+$ controls duration. These parameters are related to Arias intensity $\bar{I}_a$, effective duration D_{5-95} , defined as the time interval between the instants in which 5% $\bar{I}_a$ and 95% $\bar{I}_a$ are reached, and the time instant t_{mid} when 45% of $\bar{I}_a$ is reached. The impulse-response function $h(t, \kappa(\tau))$ of the time-varying filter is given by:

$$h(t - \tau, \kappa(\tau)) = \begin{cases} \frac{\omega_f(\tau)}{\sqrt{1 - \zeta_f^2(\tau)}} \exp[-\zeta_f(\tau) \omega_f(\tau)(t - \tau)] \sin[\omega_f(\tau) \sqrt{1 - \zeta_f^2(\tau)}(t - \tau)], & \tau \leq t \\ 0, & \tau > t \end{cases} \quad (21)$$

where $\kappa(\tau) = [\omega_f(\tau), \zeta_f(\tau)]$ contains the undamped circular frequency $\omega_f(\tau)$ and damping ratio $\zeta_f(\tau)$ of the filter. In this case, $\omega_f(\tau)$ and $\zeta_f(\tau)$ were assumed to be:

$$\omega_f(\tau) = \omega_{\text{mid}} + \omega'(\tau - t_{\text{mid}}) \quad (22)$$

$$\zeta_f(\tau) = \zeta_f \quad (23)$$

where the ω_{mid} is the undamped circular frequency at the time instant t_{mid} ; ω' is a constant slope of the varying $\omega_f(\tau)$.

The initial ground motion process $\tilde{w}(t)$ is then high-pass filtered to remove unrealistic velocity and displacement residuals (Rezaeian and Der Kiureghian 2010). The final ground motion input $\ddot{w}(t)$ can be obtained by

$$\ddot{w}(t) + 2\omega_c \dot{w}(t) + 2\omega_c^2 w(t) = \tilde{w}(t) \quad (24)$$

where ω_c is the high-pass filter frequency, which is suggested to be $\omega_c/2\pi = 0.1$ Hz (Rezaeian and Der Kiureghian 2010).

In summary, the parameter vector $\xi = [\bar{I}_a, D_{5-95}, t_{\text{mid}}, \omega_{\text{mid}}, \omega', \zeta_f]$ uniquely defines the ground motion model. Among the parameters, ω_{mid} and ζ_f were calibrated by the Nelder-Mead simplex algorithm to fit the target spectrum, while all other parameters were obtained from the Loma Prieta records (Moment magnitude = 6.93, Rupture distance = 18.3 km). The process leads to a model setting of $\xi = [0.045, 12.62, 4.73, 2\pi \times 3.27, -2\pi \times 0.08, 0.48]$. The total time duration $t_{\text{final}} - t_0$ was taken as 30 s with a time step of $\Delta t = 0.005$ s. Fig. 3 shows the comparison between the target spectrum and the spectra of 300 synthetic ground motions.

Under these excitations, the system experiences a significant nonlinear hysteretic behavior, as illustrated in Fig. 4 for one of the ground motion samples of Fig. 3. Fig. 5 shows the peak and residual inter-story drift ratios (IDR) for all 300 synthetic ground motions. It is seen that the magnitude and distribution of the peak and residual IDR is similar to that reported in Hutt et al. (2016) for the full 3D archetype building, therefore ensuring the case study of this section is representative of practical engineering problems.

Model Training

A three-dimensional reduced space was considered in defining the reduced space, i.e., $n_r = 3$, as the corresponding sum of squares of the first 3 singular values is greater than 99.9% of the total, i.e. $\eta = 0.999$. Based on the identified reduced basis Φ_{n_r} , Eq. (18) can be rewritten in the reduced space as:

$$\begin{aligned} & \Phi_{n_r}^T M \Phi_{n_r} \ddot{q} + \Phi_{n_r}^T C \Phi_{n_r} \dot{q} + \Phi_{n_r}^T B^T K B \Phi_{n_r} q = \\ & \Phi_{n_r}^T M \ddot{w}(w, t) + \alpha \Phi_{n_r}^T B^T K |B \Phi_{n_r} \dot{q}(t)| \circ [B^T K]^{-1} [M \ddot{w}(w, t) - M \Phi_{n_r} \ddot{q} - C \Phi_{n_r} \dot{q}] \end{aligned} \quad (25)$$

The high-fidelity references solutions were determined by directly solving Eq. (18) and (25) through the 4th order Runge-Kutta (RK) algorithm, which were then used for calibrating the metamodel.

The MIMO NARX model was trained for representing the velocity of the reduced coordinates, $\dot{q}(t)$, with a maximum time delay of 3 for both inputs and outputs, i.e. $n_f = n_q = 3$. The

displacement responses, $\mathbf{q}(t)$, are then obtained by integrating $\dot{\mathbf{q}}(t)$. For each reduced coordinate, the potential NARX terms include: $1, \dot{q}_j(t-l\Delta t), \ddot{w}(t-l\Delta t), |\Phi_{n_r}^m \dot{\mathbf{q}}(t-\Delta t)|, |\Phi_{n_r}^m \dot{\mathbf{q}}(t-\Delta t)|\dot{q}_j(t-l\Delta t), |\Phi_{n_r}^m \dot{\mathbf{q}}(t-\Delta t)|\ddot{w}(t-l\Delta t)$ for $l, j = 1, 2, 3$, and $m = 1, 2, \dots, 40$, which leads to a total of 575 terms.

To investigate the convergence properties of the proposed approach, training set sizes of n_s equal to 10, 50, 100, 200, 300, and 400 were considered. In calibrating the metamodel to each training set, an error tolerance of $\hat{E} = -10^{-4}$ was considered. The expected simulation errors over the training sets for roof displacement, $\bar{e}_{SE,40}$, are summarized in Table 2. The corresponding convergence curves associated with $\bar{e}_{SE,40}$ are shown in Fig. 6 for both the three-dimensional reduced space and the full space. As can be seen, for $n_s \geq 200$, the expected simulation error becomes, for all intents and purposes, constant. It is interesting to observe how, even for small training sets ($n_s < 200$), the proposed approach still succeeds in achieving relatively low simulation errors, e.g. less than 0.1 in the full space. In the following, results will refer to the case of a training sample size of $n_s = 300$, for which the most appropriate MIMO NARX models identified by the LARP scheme contain 209, 232, and 226 terms respectively for the three coordinates of the reduced space. The velocity responses of the reduced coordinates estimated from the metamodel were compared with the high-fidelity reference solutions, as shown in Fig. 7 for a representative sample with median level of error. As can be seen, the metamodel reproduced the responses for all three reduced coordinates with remarkable accuracy. Fig. 8(a) presents the comparison for the representative sample between the reference and reproduced top floor velocity responses, $\dot{x}_{40}(t)$, of the system, obtained by transforming the responses of the reduced space back to the full space. Similar to the results seen for the reduced space, the responses reconstructed by the MIMO NARX model were in excellent agreement with the reference solutions. In addition, a strong correspondence can be observed between the reference and reconstructed peak velocities over all samples, as shown in Fig. 8(b), indicating that a high level of accuracy was maintained over the entire training set. The corresponding displacement responses were then obtained by integrating the velocity responses. Fig. 9 reports the comparison between the reference and reconstructed displacements at the top floor. It is noteworthy that the top floor displacement was reproduced with the same level of

accuracy as the velocity. Both the exceedance probability and peak displacements were reproduced with excellent accuracy over the entire training set.

Simulation Results

To investigate the performance of the trained MIMO NARX model, a new set of 300 samples, denoted “test set” in the following, were generated to evaluate the performance of the calibrated metamodel in simulation/prediction mode. It is important to observe that the aforementioned test set is generated independently from the set that was considered in training the metamodel (i.e. none of the test set samples were used in training the metamodel). This will allow the prediction capability of the metamodeling approach to be investigated in this section, therefore providing an initial verification of the proposed approach. The reference solutions were once again estimated using the 4th order Runge-Kutta (RK) algorithm. In particular, reference solutions in the reduced and full spaces were obtained. Fig. 10 shows the comparison between the reference and simulated velocity responses in the reduced space for a representative test sample. It can be seen that the accuracy of the metamodel for predicting responses maintained a similar level of accuracy as seen in the training set (Fig. 7). By transforming the responses in the reduced space back to the full space, the velocity responses shown in Fig. 11 were obtained. As can be seen, the trained MIMO NARX model accurately predicted both the time history response and the peak responses over all test samples.

To illustrate the validity of the reduced model for representing the displacements responses of the full system, Fig. 12 reports the comparison of the displacements at the top floor, x_{40} , with the references solutions determined from the transformation of the reduced space response to the full space. As can be seen, remarkable accuracy is seen in both the individual responses, as illustrated in Fig. 12(c) for a typical sample, as well as over all samples, as illustrated in Fig. 12(a) that shows the exceedance probability associated with the top floor response as well as Fig. 12(b) that shows the peak displacement responses over all test samples. To illustrate the effectiveness of the reduce model of the “Model order reduction” section, Fig. 13 reports the comparison of the displacement response at the top floor obtained from the proposed metamodel with those obtained by directly

integrating the full system. As can be seen, while there is an increase in error as compared to Fig. 12, the metamodel retains remarkable accuracy over all test samples.

Overall, and without a particular optimization of the codes, the metamodel was well over an order of magnitude more efficient than the full model while preserving a remarkable level of accuracy. It should also be observed that the proposed metamodeling approach provides the output for all the DOFs of the MDOF system, i.e. the response vectors $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$ and $\ddot{\mathbf{x}}(t)$, with $\ddot{\mathbf{x}}(t)$ derived from the knowledge of $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$. This property makes the proposed approach particularly well suited for integration into probabilistic performance-based frameworks that generally require the entire response of the system for evaluating the performance metrics. While this work illustrated the applicability of the proposed approach to a shear type building, it is believed that the framework can have a wide range of practical applications involving various types of nonlinearity. This belief stems from how the POD-based model order reduction has been shown to be effective in reducing complex nonlinear structural systems, e.g. (Bamer et al. 2017), while NARX-based metamodeling of low-dimensional nonlinear dynamic systems has shown promise for various types of nonlinearity, e.g. (Mai et al. 2016). These properties, together with the efficiency and accuracy shown in this section, illustrates the strong potential of the proposed metamodeling approach.

CONCLUSION

This paper proposes a metamodeling approach which combines reduced-order modeling with multi-input multi-output nonlinear auto-regressive models with exogenous input for representing a class of nonlinear hysteretic MDOF systems subject to general stochastic excitation. An important property of the approach is that it enables the representation of the response of the entire system through a single metamodel. For calibration, a framework base on the combination of a least angle regression algorithm with pruning scheme and an ordinary least squares approach was developed. To demonstrate the applicability of the approach, a case study consisting in a high-dimensional structural system with distributed hysteretic nonlinearity and subject to general stochastic earthquake excitations was presented. The proposed approach was seen not only to be capable of reproducing the dynamic response of the system with remarkable accuracy, but also to be over an

order of magnitude faster than classic integration approaches. The general versatility of POD-based reduction of nonlinear systems, coupled with the general capability of NARX-based metamodels in capturing nonlinearity in low-dimensional spaces, points towards the applicability of the proposed approach to various problems of practical interest with a variety of nonlinear behaviors. Future research will focus on better understanding the true versatility of the proposed approach.

DATA AVAILABILITY STATEMENT

All data, models, and code generated or used during the study appear in the submitted article.

ACKNOWLEDGMENTS

This research effort was supported in part by the National Science Foundation (NSF) under Grant No. CMMI-1750339. This support is gratefully acknowledged.

REFERENCES

- American Society of Civil Engineers (2017). *Minimum design loads and associated criteria for buildings and other structures*. Reston, Virginia.
- Bamer, F., Amiri, A. K., and Bucher, C. (2017). “A new model order reduction strategy adapted to nonlinear problems in earthquake engineering.” *Earthq. Eng. Struct. Dyn.*, 46(4), 537–559.
- Bamer, F. and Markert, B. (2017). “An efficient response identification strategy for nonlinear structures subject to nonstationary generated seismic excitations.” *Mech. Based Des. Struc.*, 45(3), 313–330.
- Bhattacharyya, B., Jacquelin, E., and Brizard, D. (2020). “A Kriging–NARX model for uncertainty quantification of nonlinear stochastic dynamical systems in time domain.” *J. Eng. Mech.*, 146(7), 04020070.
- Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.
- Billings, S. A. and Chen, S. (1996). “The determination of multivariable nonlinear models for dynamic systems using neural networks.” *Neural Network Systems Techniques and Application*, C. T. Leondes, ed., Academic Press, 231–278.

- Billings, S. A., Chen, S., and Korenberg, M. J. (1989). "Identification of mimo non-linear systems using a forward-regression orthogonal estimator." *Int. J. Control*, 49(6), 2157–2189.
- Chuang, W. C. and Spence, S. M. J. (2017). "A performance-based design framework for the integrated collapse and non-collapse assessment of wind excited buildings." *Eng. Struct.*, 150, 746–758.
- Chuang, W. C. and Spence, S. M. J. (2019). "Rapid uncertainty quantification for non-linear and stochastic wind excited structures: a metamodeling approach." *Meccanica*, 54(9), 1327–1338.
- Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). "Least angle regression." *Ann. Stat.*, 32(2), 407–499.
- Federal Emergency Management Agency (FEMA) (2012). *Seismic Performance Assessment of Buildings Volume I-Methodology*. Washington, DC.
- Gidas, I. and Taflanidis, A. A. (2013). "Parsimonious modeling of hysteretic structural response in earthquake engineering: Calibration/validation and implementation in probabilistic risk assessment." *Eng. Struct.*, 49, 1017–1033.
- Gidas, I., Taflanidis, A. A., and Mavroeidis, G. P. (2015). "Kriging metamodeling in seismic risk assessment based on stochastic ground motion models." *Earthq. Eng. Struct. Dyn.*, 44(14), 2377–2399.
- Grigoriu, M. (2009). "Reduced order models for random functions. application to stochastic problems." *Appl. Math. Model.*, 33(1), 161–175.
- Grigoriu, M. (2012). "A method for solving stochastic equations by reduced order models and local approximations." *J. Comput. Phys.*, 231, 6495–6513.
- Holmes, P., Lumley, J. L., and Berkooz, G. (1996). *Turbulence, Coherent Structures*. Cambridge University Press, Cambridge, UK.
- Hutt, C. M., Almufti, I., Willford, M., and Deierlein, G. (2016). "Seismic loss and downtime assessment of existing tall steel-framed buildings and strategies for increased resilience." *J. Struct. Eng.*, 142(8), C4015005.
- Jensen, H., Muñoz, A., Papadimitriou, C., and Millas, E. (2016). "Model reduction techniques for

- reliability-based design problems of complex structural systems.” *Reliab. Eng. Syst. Saf.*, 149, 204–217.
- Kerschen, G. and Golinval, J. C. (2002). “Physical interpretation of the proper orthogonal modes using the singular value decomposition.” *J. Sound Vib.*, 249(5), 849–865.
- Kundu, A. and Adhikari, S. (2014). “Transient response of structural dynamic systems with parametric uncertainty.” *J. Eng. Mech.*, 140(2), 315–331.
- Leontaritis, I. J. and Billings, S. A. (1987). “Experimental design and identifiability for non-linear systems.” *Int. J. Syst. Sci.*, 18(1), 189–202.
- Lucor, D. and Karniadakis, G. E. (2004). “Adaptive generalized polynomial chaos for nonlinear random oscillators.” *SIAM J. Sci. Comput.*, 26(2), 720–735.
- Lucor, D., Su, C. H., and Karniadakis, G. E. (2004). “Generalized polynomial chaos and random oscillators.” *Int. J. Numer. Meth. Eng.*, 60(3), 571–596.
- Mai, C. V. (2016). “Polynomial chaos expansions for uncertain dynamical systems-applications in earthquake engineering.” Ph.D. thesis, ETH Zürich, Zürich, Switzerland.
- Mai, C. V., Spiridonakos, M. D., Chatzi, E. N., and Sudret, B. (2016). “Surrogate modeling for stochastic dynamical systems by combining nonlinear autoregressive with exogenous input models and polynomial chaos expansions.” *Int. J. Uncertainty Quantification*, 6(4).
- Mai, C. V. and Sudret, B. (2017). “Surrogate models for oscillatory systems using sparse polynomial chaos expansions and stochastic time warping.” *SIAM/ASA J. Uncertain. Quantif.*, 5(1), 540–571.
- Ouyang, Z. and Spence, S. M. J. (2020). “A performance-based wind engineering framework for the envelope system of engineered buildings subject to directional wind and rain hazards.” *J. Struct. Eng.*, 146, 5.
- Ouyang, Z. and Spence, S. M. J. (2021). “Performance-based wind-induced structural and envelope damage assessment of engineered buildings through nonlinear dynamic analysis.” *J. Wind. Eng. Ind. Aerodyn.*, 208, 104452.
- Patsialis, D. and Taflanidis, A. A. (2020). “Reduced order modeling of hysteretic structural response and applications to seismic risk assessment.” *Eng. Struct.*, 209, 110135.

- 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.
- 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.
- Spiridonakos, M. D. and Chatzi, E. N. (2015). "Metamodeling of dynamic nonlinear structural systems through polynomial chaos narx models." *Comput. Struct.*, 157, 99–113.
- Tehrani, M. H., Harvey Jr., P. S., Gavin, H. P., and Mirza, A. M. (2018). "Inelastic condensed dynamic models for estimating seismic demands for buildings." *Eng. Struct.*, 177, 616–629.
- Volkwein, S. (2013). "Proper orthogonal decomposition: Theory and reduced-order modelling." *Lecture Notes, University of Konstanz*, 4(4), 1–29.
- Wei, H. L. and Billings, S. A. (2004). "A unified wavelet-based modelling framework for non-linear system identification: the wanarx model structure." *Int. J. Control*, 77(4), 351–366.
- Yang, T. Y., Moehle, J., Stojadinovic, B., and Der Kiureghian, A. (2009). "Seismic performance evaluation of facilities: methodology and implementation." *J. Struct. Eng.*, 135(10), 1146–1154.

553	List of Tables	
554	1	Element sections used in the steel frame. 27
555	2	Simulation errors associated with roof displacement and various training set sizes. . 28

TABLE 1. Element sections used in the steel frame.

Floors	Beams	Interior columns	Exterior columns
1	W36×282	66×7.6	66×6.4
2-10	W36×282	56×7.6	51×6.4
11-20	W36×194	51×5.0	51×5.0
21-30	W33×169	46×2.5	46×2.5
31-40	W27×84	46×1.9	46×1.9
Columns sections: (outer side size)×(wall thickness)			

TABLE 2. Simulation errors associated with roof displacement and various training set sizes.

Training set size	10	50	100	200	300	400
$\hat{e}_{SE,40}$ in the reduced space	0.0627	0.0436	0.0411	0.0289	0.0354	0.0369
$\hat{e}_{SE,40}$ in the full space	0.0867	0.0793	0.0679	0.0624	0.0652	0.0631

List of Figures

1	Schematic illustration of the proposed approach: (a) data flow; (b) flow chart. . . .	31
2	Case study steel frame: (a) schematic of the frame layout; (b) shear building idealization.	32
3	Comparison between the target and simulated spectra.	33
4	Typical nonlinear restoring force at floor 1 for a representative ground motion. . . .	34
5	Inter-story drift ratio response over all 300 synthetic ground motions.	35
6	Variation of the expected simulation error for the roof displacement.	36
7	Comparison between the reconstructed and reference velocity solutions in the reduced space for a typical sample: (a) first reduced coordinate; (b) second reduced coordinate; (c) third reduced coordinate.	37
8	Comparison between the reference and reconstructed solutions in the full space: (a) velocity responses at the top floor for a representative sample of the training set; (b) peak velocities at the top floor over all training samples.	38
9	Comparison between the reference and reconstructed solutions in the full space: (a) exceedance probabilities of peak displacements at the top floor; (b) peak displacements at the top floor over all samples; (c) displacement responses at the top floor for a representative sample.	39
10	Comparison for a representative test set sample between the simulated and reference reduced space velocity response: (a) first reduced coordinate; (b) second reduced coordinate; (c) third reduced coordinate.	40
11	Comparison for the test set samples between the reference and simulated velocity after transformation to the full space: (a) velocity responses at the top floor for a representative sample of the test set; (b) peak velocity at the top floor over all test samples.	41

581	12	Comparison for the test set samples between the displacements at the 40th floor	
582		estimated from the metamodel and the reference reduced model after transformation	
583		to the full space: (a) exceedance probability; (b) peak values; (c) representative	
584		time history and overall error evolution.	42
585	13	Comparison for the test set samples between the displacements at the 40th floor	
586		estimated from the metamodel and the reference full model: (a) exceedance prob-	
587		ability; (b) peak values; (c) representative time history and overall error evolution. .	43

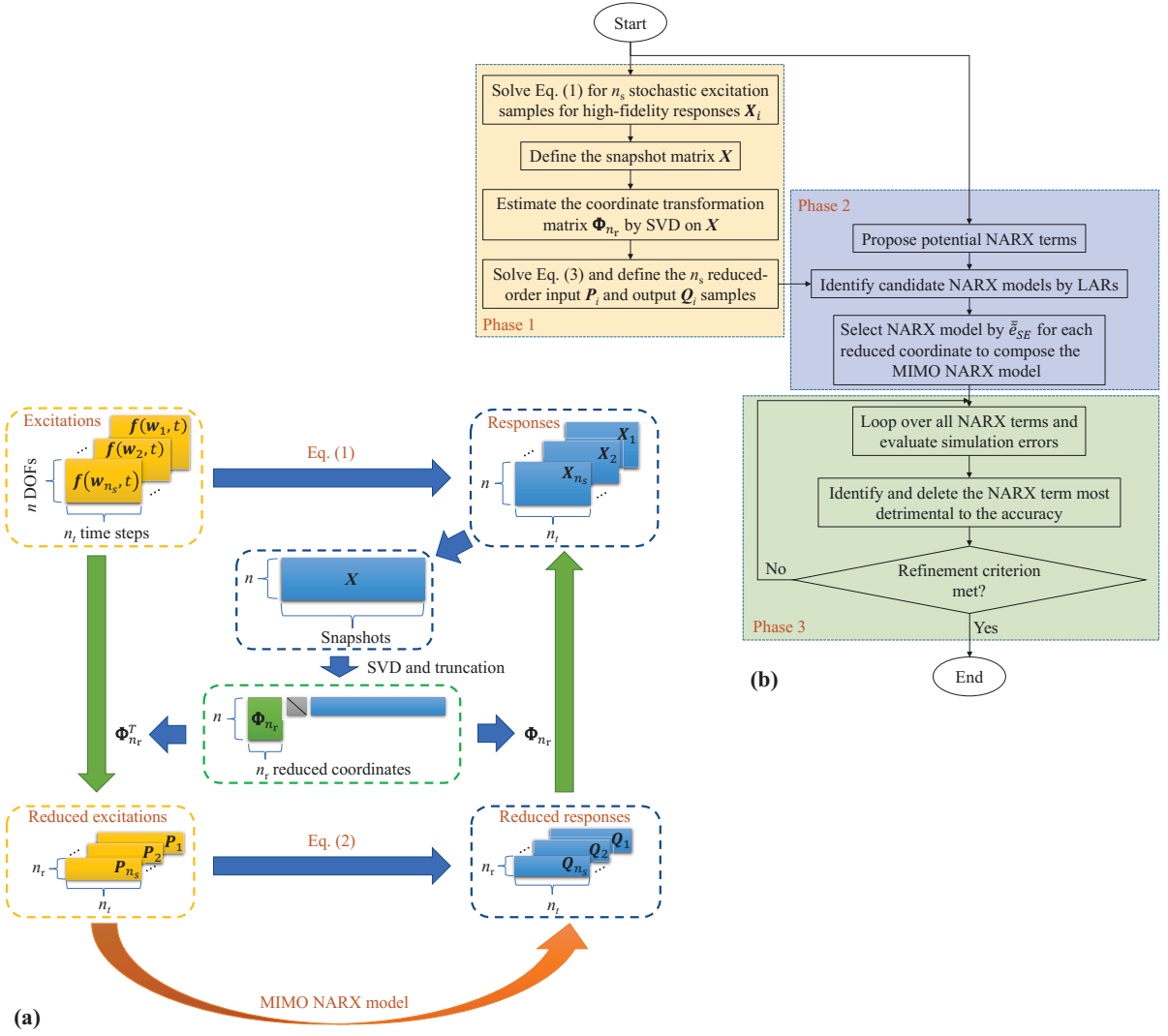


Fig. 1. Schematic illustration of the proposed approach: (a) data flow; (b) flow chart.

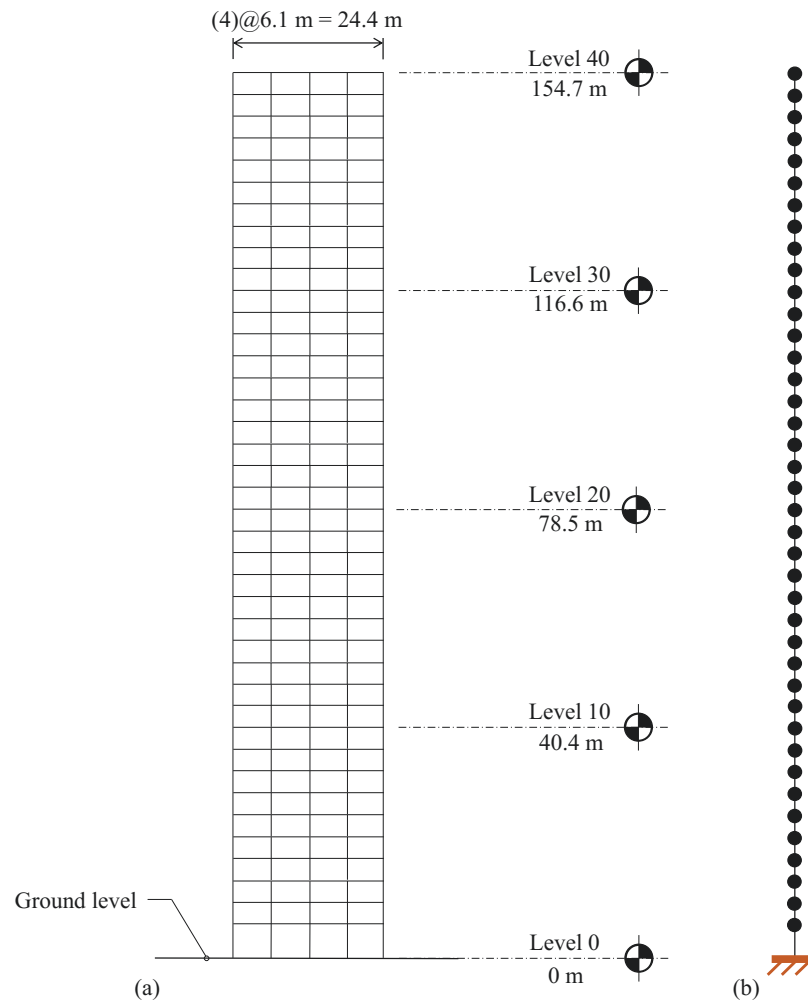


Fig. 2. Case study steel frame: (a) schematic of the frame layout; (b) shear building idealization.

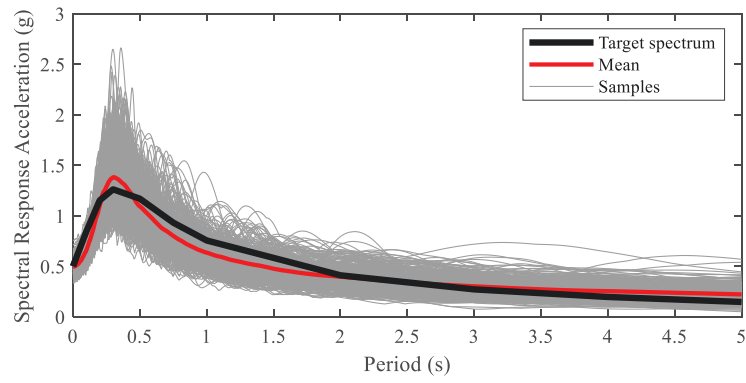


Fig. 3. Comparison between the target and simulated spectra.

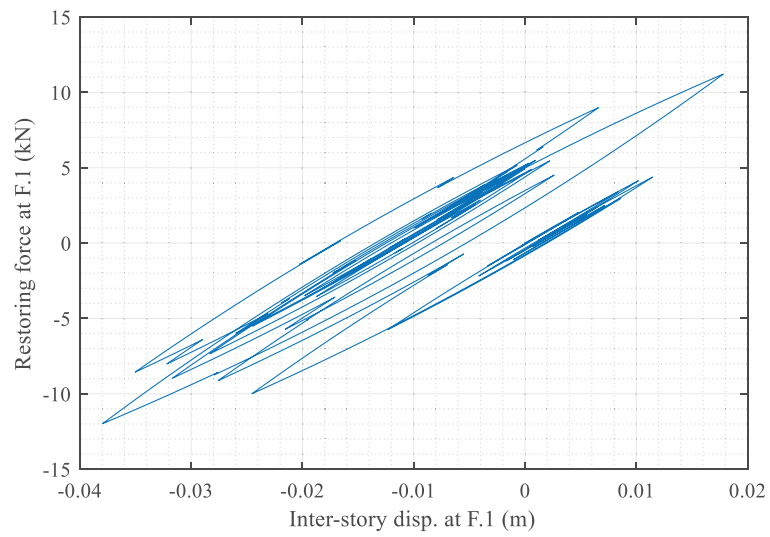


Fig. 4. Typical nonlinear restoring force at floor 1 for a representative ground motion.

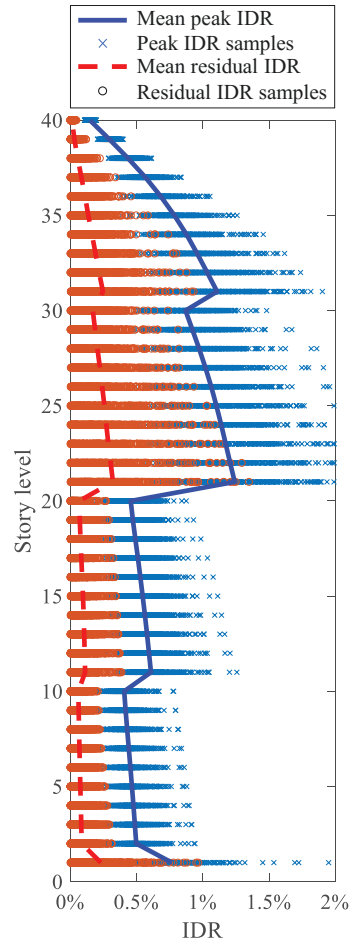


Fig. 5. Inter-story drift ratio response over all 300 synthetic ground motions.

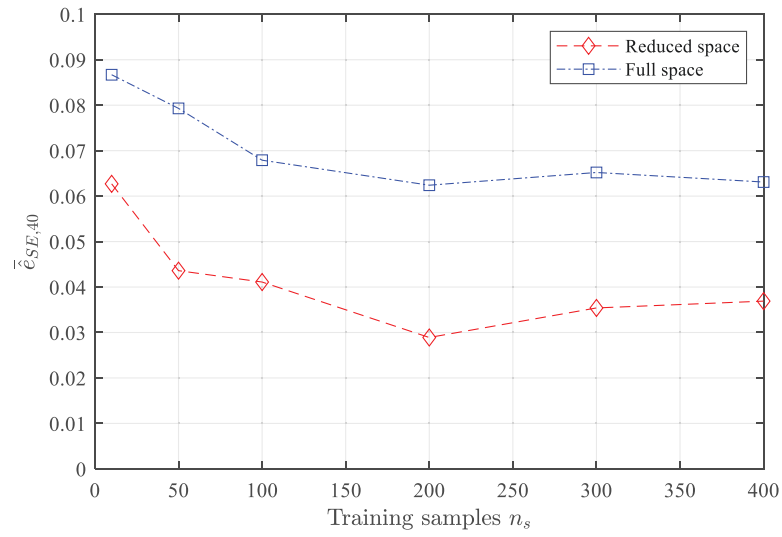


Fig. 6. Variation of the expected simulation error for the roof displacement.

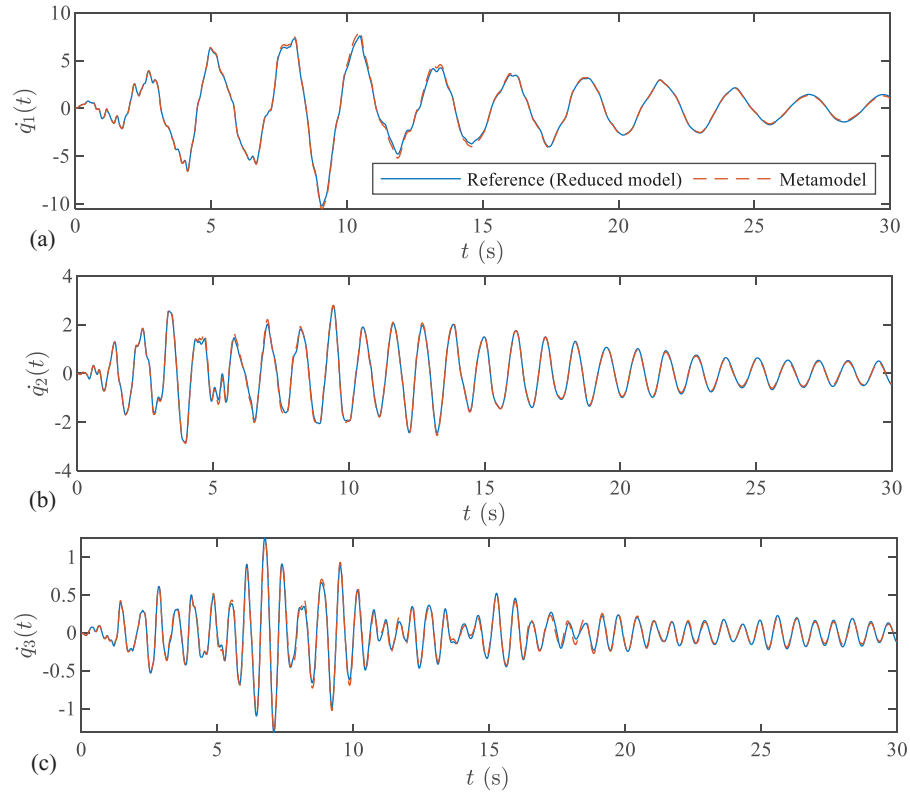


Fig. 7. Comparison between the reconstructed and reference velocity solutions in the reduced space for a typical sample: (a) first reduced coordinate; (b) second reduced coordinate; (c) third reduced coordinate.

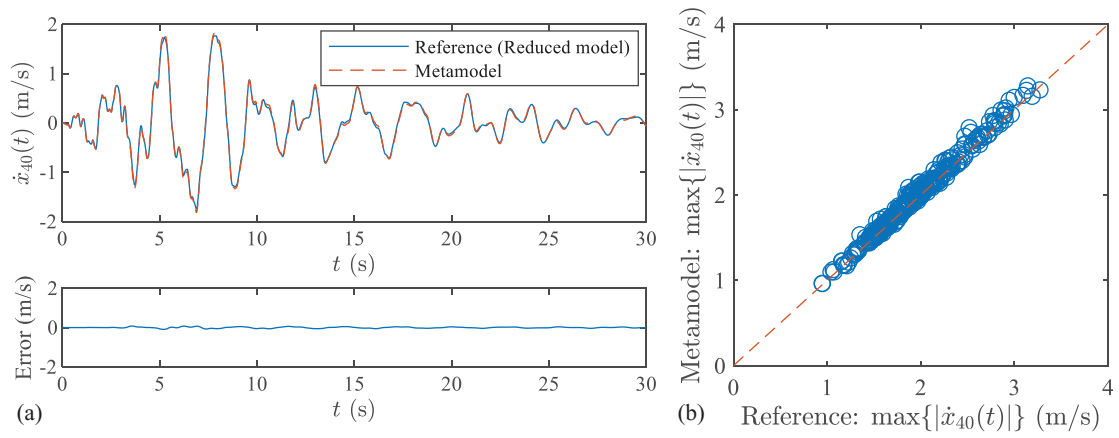


Fig. 8. Comparison between the reference and reconstructed solutions in the full space: (a) velocity responses at the top floor for a representative sample of the training set; (b) peak velocities at the top floor over all training samples.

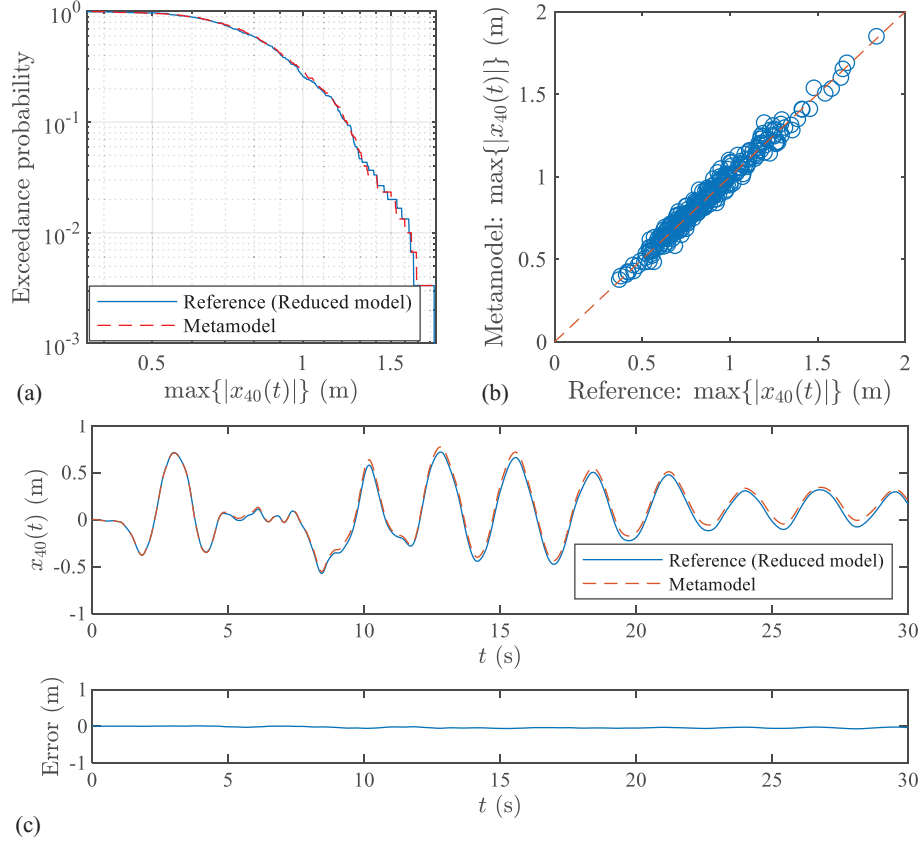


Fig. 9. Comparison between the reference and reconstructed solutions in the full space: (a) exceedance probabilities of peak displacements at the top floor; (b) peak displacements at the top floor over all samples; (c) displacement responses at the top floor for a representative sample.

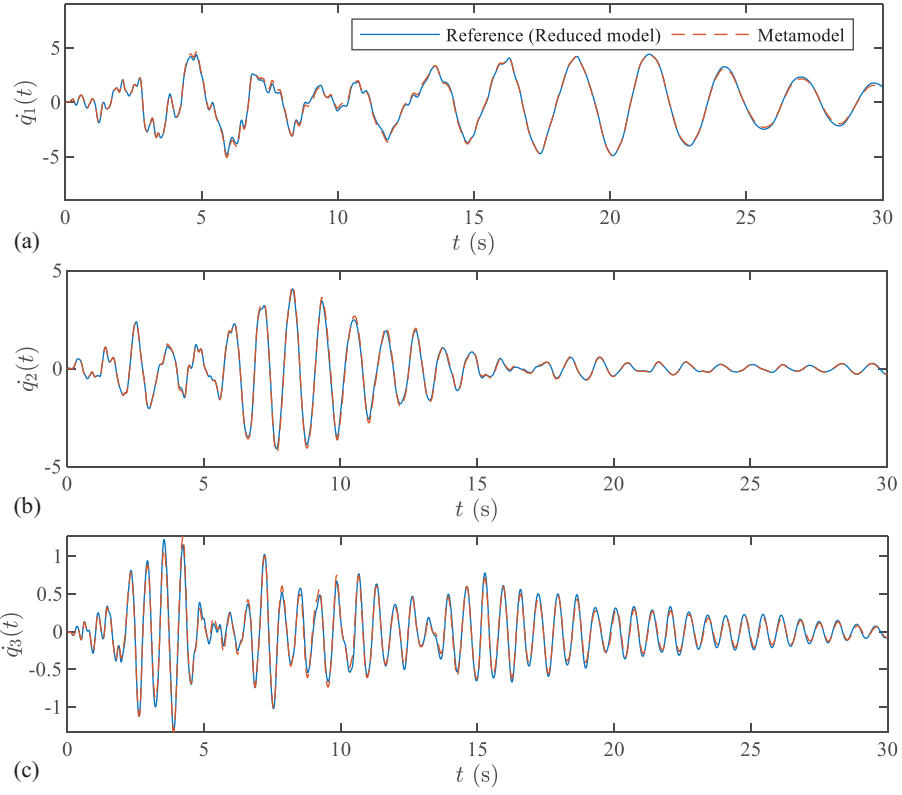


Fig. 10. Comparison for a representative test set sample between the simulated and reference reduced space velocity response: (a) first reduced coordinate; (b) second reduced coordinate; (c) third reduced coordinate.

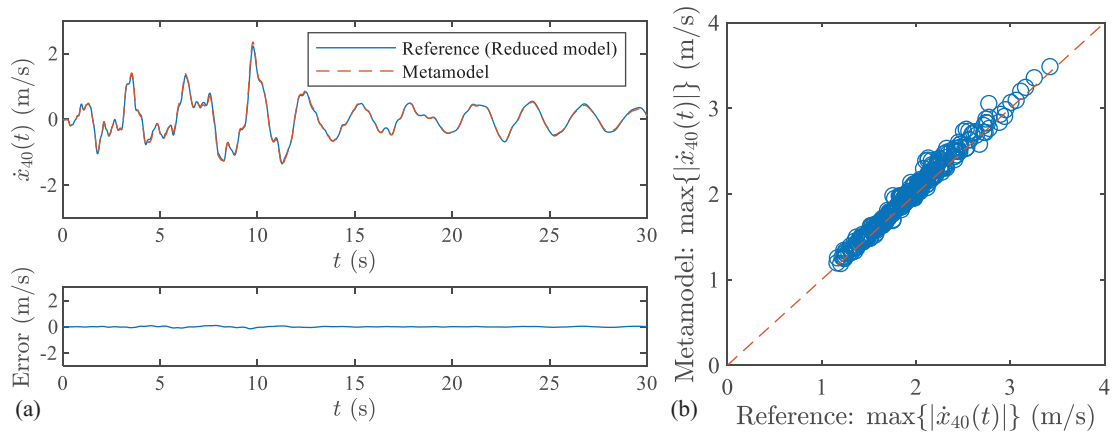


Fig. 11. Comparison for the test set samples between the reference and simulated velocity after transformation to the full space: (a) velocity responses at the top floor for a representative sample of the test set; (b) peak velocity at the top floor over all test samples.

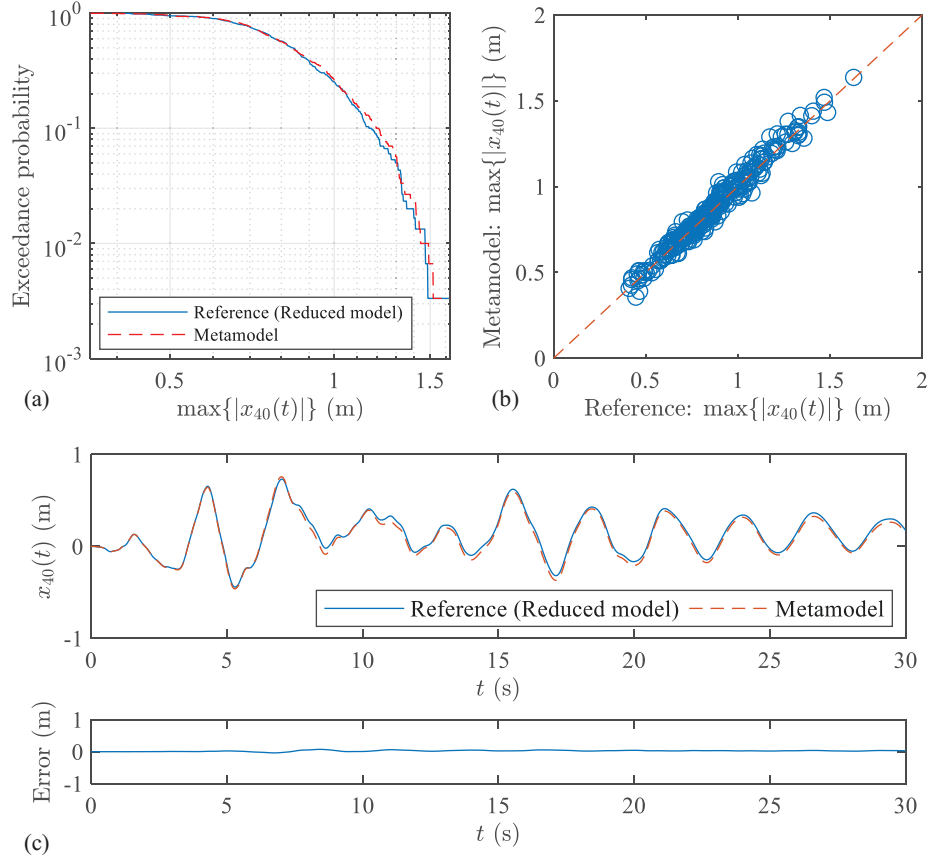


Fig. 12. Comparison for the test set samples between the displacements at the 40th floor estimated from the metamodel and the reference reduced model after transformation to the full space: (a) exceedance probability; (b) peak values; (c) representative time history and overall error evolution.

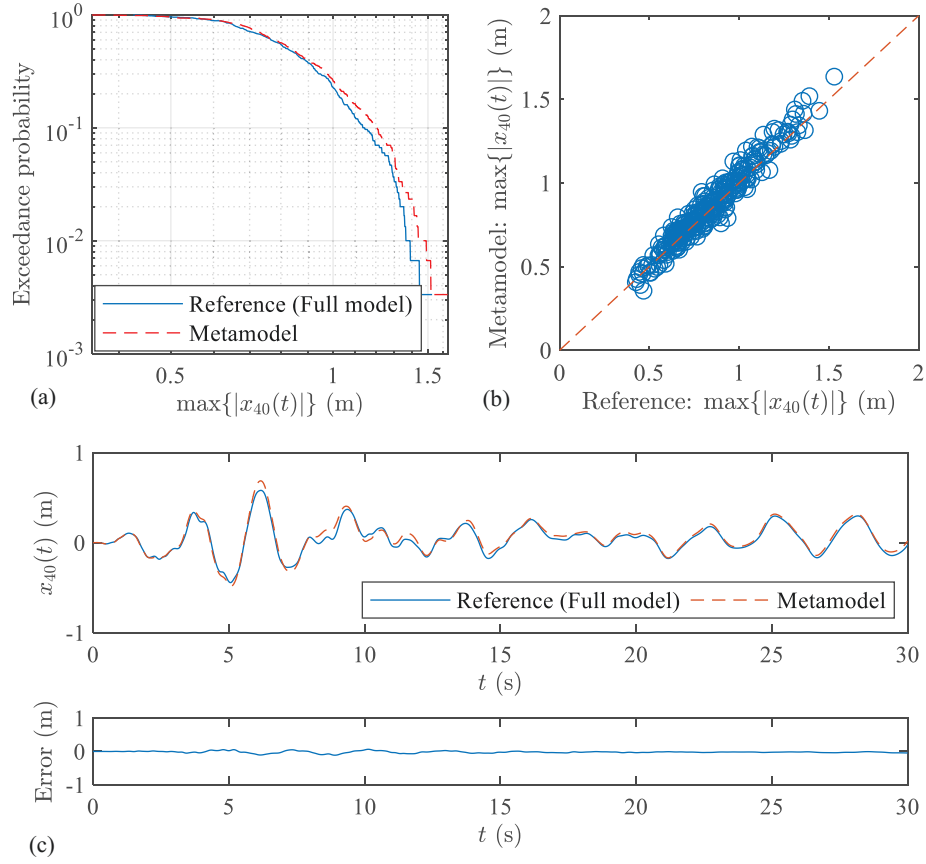


Fig. 13. Comparison for the test set samples between the displacements at the 40th floor estimated from the metamodel and the reference full model: (a) exceedance probability; (b) peak values; (c) representative time history and overall error evolution.