Computers and Geotechnics 89 (2017) 43–54

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Research Paper

Site response analysis using one-dimensional equivalent-linear method and Bayesian filtering

Rodrigo Astroza [a,*], César Pastén [b], Felipe Ochoa-Cornejo [b]

[a] Facultad de Ingeniería y Ciencias Aplicadas, Universidad de los Andes, Santiago, Chile
[b] Departamento de Ingeniería Civil, Universidad de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 14 December 2016
Received in revised form 1 March 2017
Accepted 10 April 2017

Keywords:
Site response
Equivalent-linear model
Unscented Kalman filter
Parameter estimation

ABSTRACT

Site response analysis is crucial to define the seismic hazard and distribution of damage during earthquakes. The equivalent-linear (EQL) is a numerical method widely investigated and used for site response analysis. Because several sources of uncertainty are involved in this type of analysis, parameters defining the numerical models need to be identified from in-situ measurements. In this paper, a Bayesian inference method to estimate the expected values and covariance matrix of the model parameters is presented. The methodology uses data from downhole arrays recorded during earthquakes. Two numerical applications show the good performance and prediction capabilities of the proposed approach.

1. Introduction

Near surface geological site conditions have a predominant role in defining the characteristics of the ground surface motion during earthquakes, and therefore, in the damage distribution. The changes in the amplitude and frequency content of the ground surface motion due to topographic effects and/or characteristics of the soil deposits, referred to as site effects, define the seismic hazard and have a direct influence on the behavior of structures during earthquakes. A site response analysis aims at estimating the ground motion at the surface given the seismic motion at the bedrock and the properties of the soil profile [1]. If topographic and basin effects are not significant, one-dimensional (1D) analysis (i.e., assuming horizontal soil layers, boundaries of infinite lateral extension, and vertically propagating shear waves) has proved adequate to model the propagation of the seismic waves through the soil profile (e.g., [1]).

Various numerical methods for 1D site response analysis, including the time-domain nonlinear (NL) method (e.g., [2]) and the frequency-domain equivalent-linear (EQL) method, have been proposed and investigated. The EQL method has been widely used in both research and engineering practice (e.g., [3,4]) due to its simplicity, flexibility, and robustness [2]. Research efforts have focused in pointing out limitations on the NL and the EQL methods and on comparing site response analyses using both methods, with numerically simulated data (e.g., [5,6]) and experimental data from downhole arrays [7,8].

A crucial aspect in site response analysis is the several sources of uncertainty involved in the phenomenon [4], such as model parameters, input motion, and modeling errors (e.g., 1D representation and soil constitutive models assumptions). Particularly, among model parameters, those related to the soil profile (e.g., shear wave velocity and thickness of soil layers) and constitutive models (dynamic properties) can be highly uncertain. Field data from downhole arrays have been used with system identification methods to estimate soil parameters, such as shear wave velocities and damping ratios. Glaser [9] summarized several efforts to characterize soil properties using system identification methods applied to data obtained from forced and ambient vibrations. Elgamal et al. [10], Zeghal et al. [11], and Glaser and Baise [12] used seismic downhole array data at the Lotung site to identify the dynamic properties of the soil profile and model the response during several seismic events. Kokusho et al. [13] employed data recorded at downhole seismic arrays during the 1995 Kobe Earthquake to estimate the shear wave velocities and damping ratios in shallow soil layers. Tsai and Hashash [14] proposed a nonparametric approach to identify the soil behavior at different layers using downhole array data, applying it to data from Lotung and La Cienaga sites [15]. Recently, Mercado et al. [16,17] proposed a system identification method to obtain non-parametric estimates of the shear stresses with downhole array data in the case of 1D and

* Corresponding author.
E-mail address: rastroza@miuandes.cl (R. Astroza).

http://dx.doi.org/10.1016/j.compgeo.2017.04.004
0266-352X/

bi-directional (2D) soil response analyses, respectively. These previous studies have proposed deterministic identification approaches, providing point estimates of the quantities or parameters of interest. In this context, to overcome the limitations of deterministic approaches, Bayesian inference methods allow assessing the uncertainty in the system identification, in addition to point estimates. In recent years, Bayesian methods (see [18] for introductory concepts) are attracting the attention in the field of geotechnical engineering for site characterization (e.g., [19,20]).

This paper proposes employing a Bayesian inference methodology to estimate the model parameters involved in site response analyses. The methodology uses the EQL method, implemented in the computational program SHAKE91 [21], and downhole array data. The proposed approach estimates the expected values and covariance matrix of the model parameters using the measured input (acceleration at the bedrock) and output (soil profile responses) data. Although this model calibration approach applies to any parametrized numerical model for site response analysis, including NL models, the EQL model is used here as an illustrative application due to its wide use in research and the engineering practice.

2. Nonlinear 1D site response analysis

2.1. Soil response to cyclic loading

Local ground conditions influence the propagation of seismic waves in the near surface, modifying the seismic site response and the ground motion. During seismic events, the behavior of soils is hysteretic and nonlinear and the seismic-induced cyclic shear strain decreases the secant shear modulus G and increases the damping ratio D (Fig. 1a). This nonlinear soil behavior is described as a function of the cyclic shear strain in the so-called degradation curves (Fig. 1b) [22], which are influenced by factors such as the cyclic shear strain amplitude [22], the soil plasticity [23], and the effective confining stress [24]. Also, at very small cyclic shear strain, of the order of 10^{-6}–10^{-3}%, the shear modulus can be assumed constant and calculated as a function of the mass density ρ and the shear wave velocity V_S of the soil. In this context, various studies consistently showed that the normalized degradation curve of the shear modulus with respect to the small strain shear modulus G/G_{max} (Fig. 1b) fall within a narrow band, depending on the soil type (i.e., sand, clay, or gravel).

2.2. Shear modulus degradation curve

The Hardin and Drnevich Hyperbolic Model (e.g., [22,25,26]) is one of the most used models to characterize the nonlinear behavior of soils. It considers the shear stress-shear strain curve of the soil as an asymptotic hyperbola. The combination of this model with the Masing Rule describes a hysteresis stress-strain loop, defined by the secant shear modulus G and the damping ratio D. Fig. 1a shows a hysteresis loop, the secant shear modulus, and the initial shear modulus G_{max}.

The Hyperbolic Model renders a degradation curve that cannot be fitted to experimental data and thus is better expressed as [26]

$$\frac{G}{G_{\max}} = \frac{1}{1+(\gamma/\gamma_r)^{\alpha}} \tag{1}$$

In Eq. (1), α is a curvature parameter and γ_r is the reference strain. The parameter γ_r is defined by the expression

$$\gamma_r = \gamma_{r1}(\sigma'_m/P_a)^k \tag{2}$$

In Eq. (2), $\sigma'_m = \sigma'_v(1+2K_0)/3$ is the mean horizontal effective stress, γ_{r1} is the reference strain at atmospheric pressure, used as a reference mean stress P_a, k is a sensitivity parameter, σ'_v is the vertical effective stress, and K_0 is the ratio of horizontal and vertical effective stresses. Zhang et al. [27] suggested that the reference strain γ_r is the shear strain at $G/G_{max} = 0.5$ in the degradation curves. To numerically simulate the site response data, this study calibrates the stiffness degradation curve for clays and sands with data from [28,29], respectively.

2.3. Hysteretic damping

The damping ratio D represents the energy dissipated from various mechanisms simultaneously occurring at the particle level within the soil, such as contact asperities, friction, strain rate effects, and nonlinearity of the material. Fig. 1a shows a hysteresis loop and the damping ratio D, which is calculated as

$$D = \frac{W_D}{4\pi W_S} \tag{3}$$

In Eq. (3), W_D is the area of the complete hysteresis loop and W_S is the equivalent elastic stored energy. The calculation of D is rather sensitive to experimental conditions and several efforts have aimed at identifying indirect ways for its estimation. One of

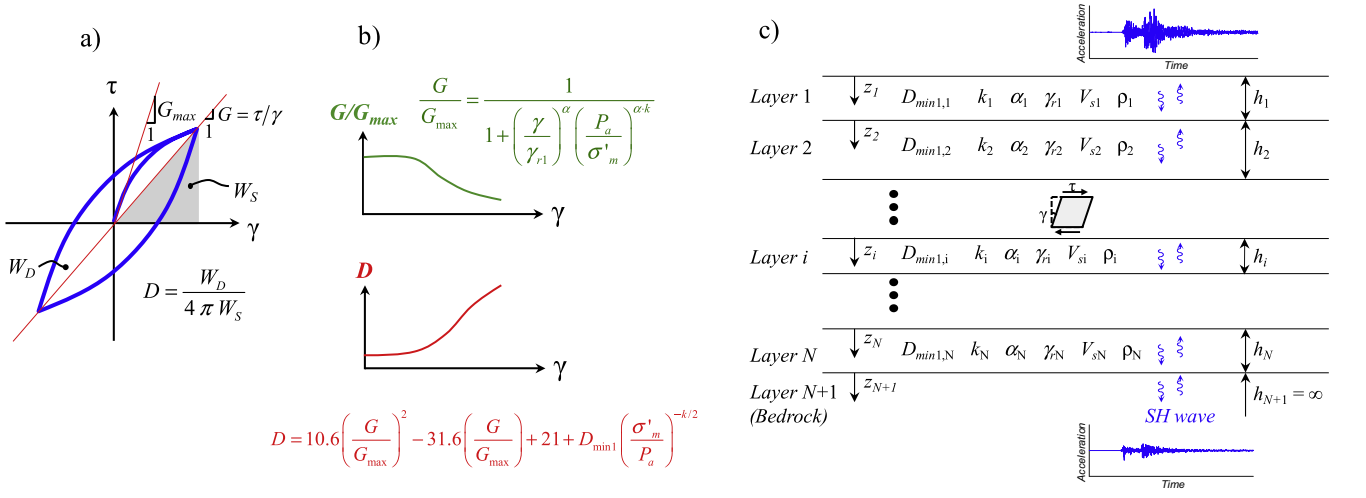

Fig. 1. Equivalent-linear technique for 1D site response analysis (a) hysteresis loop and evaluation of damping ratio, (b) shear modulus degradation and damping curves, and (c) sketch of 1D site response model.

the most widely accepted indirect methods to estimate D is through its relationship with the normalized shear modulus degradation G/G_{max} (e.g., [22]). In this context, D is related to G/G_{max} as suggested in [27] by

$$D = f(G/G_{max}) + D_{min} \tag{4}$$

In Eq. (4), D_{min} is the minimum damping ratio of the material and $f(G/G_{max})$ is a function of the normalized shear modulus. According to [27], this function can be expressed as

$$f(G/G_{max}) = 10.6(G/G_{max})^2 - 31.6(G/G_{max}) + 21 \tag{5}$$

The minimum damping ratio can be defined according to the expression

$$D_{\min} = D_{\min 1}(\sigma'_m/P_a)^{-k/2} \tag{6}$$

where D_{min1} is the damping ratio at the reference mean stress P_a.

Discussions regarding the existence of a minimum damping ratio in the elastic regime of deformation are beyond the scope of this study. In addition, the reference damping ratio D_{min1} is considered as a constant in this study, although it could be expressed as a function of the plasticity index (PI). To numerically simulate site response data, this study calibrates the damping ratio curves for clays and sands with the data reported in [30].

2.4. Nonlinear 1D site response analysis

A methodology for addressing strong motion and site response analysis during earthquakes is the 1D nonlinear wave propagation model, which assumes that the subsoil is a horizontally layered media, through which horizontally polarized shear waves (*SH* waves) travel from-and-to the bedrock (Fig. 1c). In this study, each layer is characterized by an elastic shear modulus and a viscous damping ratio. The EQL method is an iterative procedure based on vertically propagating *SH* waves in a layered media that accounts for the degradation of the dynamic soil properties with the cyclic shear strain. The method updates the shear modulus and the damping ratio as a function of the induced shear strain, using the normalized degradation curves, until compatible deformations are achieved [21,31]. Based on Sections 2.1 to 2.3, the response of the soil deposit can be written as a function of the shear wave velocity V_{S_i}, the mass density ρ_i, the thickness h_i (with $i = 1, \ldots, N+1$ and $h_{N+1} = \infty$), and the parameters defining the shear modulus degradation and damping curves (i.e., G/G_{max} *vs.* γ and D *vs.* γ) of each soil layer.

This study uses the EQL procedure implemented in SHAKE91 [21] to compute the soil deposit response, and it is summarized as indicated in Fig. 2.

3. Bayesian filtering for parameter estimation in site response problems

The soil deposit response predicted with the EQL model at time step $(n+1)$, $\hat{\mathbf{y}}_{n+1} \in \Re^{m\times 1}$ (with m = number or response measurements), can be expressed mathematically as:

$$\hat{\mathbf{y}}_{n+1} = \mathbf{h}_{n+1}(D_{\min 1_j}, k_j, \alpha_j, \gamma_{r1_j}, V_{S_i}, h_i, \rho_i, \ddot{\mathbf{u}}_{b_{1:n+1}}) = \mathbf{h}_{n+1}(\boldsymbol{\theta}, \ddot{\mathbf{u}}_{b_{1:n+1}}) \tag{7}$$

In Eq. (7), $\mathbf{h}(\cdot)$ is the nonlinear response function of the EQL model, j is the soil type identifier ($j = 1, \ldots, M$, M = number of types of soils in the profile, e.g., clay, sand, and gravel), i is the soil layer identifier ($i = 1, \ldots, N+1$, N = number of soil layers), $\ddot{\mathbf{u}}_{b_{1:n+1}} \in \Re^{(n+1)\times 1}$ is the input acceleration time history at the bedrock from time step 1 to $(n+1)$, $\boldsymbol{\theta} = [D_{\min 1_j}, k_j, \alpha_j, \gamma_{r1_j}, V_{S_i}, h_i, \rho_i]^T \in \Re^{n_\theta \times 1}$ is the vector of time-invariant parameters of the EQL model (model parameter vector) with $n_\theta = 3N + 2 + 4M$, and the subscript $n = 0, \ldots, N_s - 1$, being N_s the number of time samples, represents the discrete-time step. $D_{\min 1}$, k, α, and γ_{r1} are referred to as soil type parameters, while V_{S_i}, h_i, and ρ_i are referred to as soil layer parameters. On the other hand, the field response of the soil deposit can be recorded with an array of sensors, such as accelerometers, seismometers, and transducers to measure strains (e.g., [32]) and stresses (e.g., [33]). Assuming that $\ddot{\mathbf{u}}_b$ is known, the measured response of the soil ($\mathbf{y}_{n+1} \in \Re^{m\times 1}$) can be related to the EQL model predicted response at the time step $(n+1)$ as

$$\mathbf{y}_{n+1} = \hat{\mathbf{y}}_{n+1} + \mathbf{v}_{n+1} \tag{8}$$

In Eq. (8), $\mathbf{v}_{n+1} \in \Re^{m\times 1}$ is the simulation error vector accounting for the misfit between the measured and model predicted responses and arises from modeling errors (including uncertainties

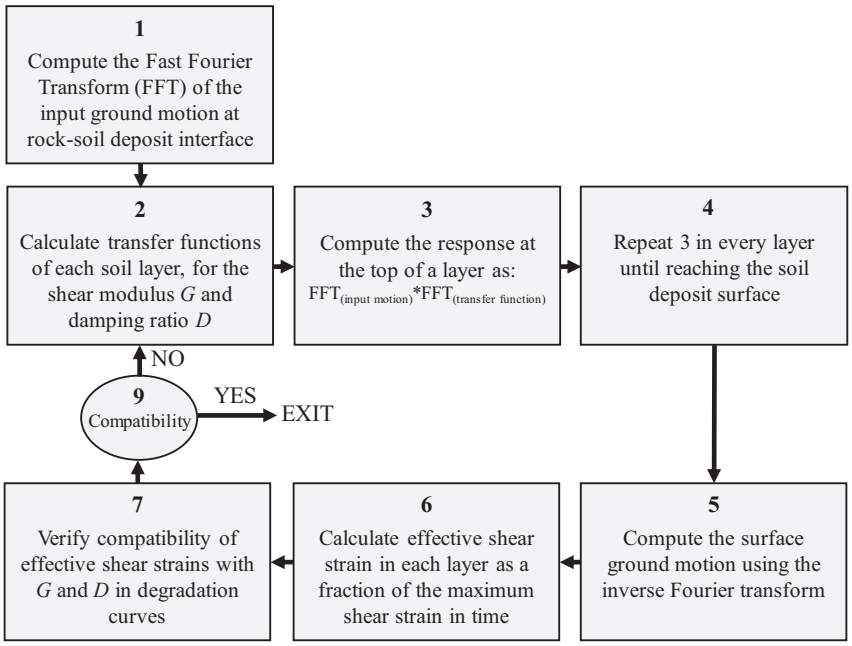

Fig. 2. Iterative procedure of the EQL method to estimate the dynamic soil response.

in the model parameters and in the model itself) and measurement noise. The vector of time-invariant model parameters, $\boldsymbol{\theta}$, is modeled as a random walk process (i.e., the value of $\boldsymbol{\theta}$ at time step ($n+1$) is equal to the value of $\boldsymbol{\theta}$ at time step n plus an error term defined as a white noise) and assuming that model uncertainty is neglected and that the measurement noise is a stationary, zero-mean and independent white Gaussian noise (i.e., statistically independent between measurement channels and across time), the following parameter estimation problem can be formulated

$$\begin{aligned}\boldsymbol{\theta}_{n+1} &= \boldsymbol{\theta}_n + \mathbf{w}_n \\ \mathbf{y}_{n+1} &= \mathbf{h}_{n+1}(\boldsymbol{\theta}_{n+1}, \ddot{\mathbf{u}}_{b_{1:n+1}}) + \mathbf{v}_{n+1}\end{aligned} \tag{9}$$

Using the notation $\mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ = normal probability density function with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$, then $\mathbf{w}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_n)$ and $\mathbf{v}_{n+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{n+1})$ are the process and measurement noise vectors, both with zero mean and covariance matrices $\mathbf{Q}_n$ and $\mathbf{R}_{n+1}$, respectively.

The nonlinear state-space model in Eq. (9) allows using Bayesian filtering techniques to estimate the model parameter vector $\boldsymbol{\theta}$ by a prediction-correction scheme. In this work, for the initial model parameter it is assumed that $\boldsymbol{\theta}_0 \sim \mathcal{N}\left(\hat{\boldsymbol{\theta}}_{0|0}, \hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0|0}\right)$ and the Unscented Kalman filter (UKF) [34,35], a so-called Sigma-Point (SP) Kalman Filter, is employed as estimation tool. The UKF is a deterministic sampling approach that does not require analytical linearization of the nonlinear state-space model in Eq. (9); therefore, model response sensitivities with respect to the parameters to be estimated do not need to be computed. The scaled unscented transformation [35] is used to generate the deterministic samples (called sigma points or SPs) of the UKF. More details about the UKF can be found in Appendix A and Fig. 3 shows the proposed algorithm to calibrate the soil profile model using the measured input at the bedrock and the soil response. Likewise, similar formulations are employed in fields such as aerospace [36], structural [37], and environmental [38] engineering. It is noted that other estimation approaches (e.g., *maximum likelihood estimation* method) could also be used to tackle this identification problem.

The proposed approach requires the process and measurement noise covariance matrices ($\mathbf{Q}_n$ and $\mathbf{R}_{n+1}$) and initial estimates of the mean and covariance matrix of the model parameters to be identified ($\hat{\boldsymbol{\theta}}_{0|0}$ and $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0|0}$). The process and measurement noise covariance matrices are taken as diagonal and time-invariant, i.e., $\mathbf{Q}_n = \mathbf{Q} = E[\mathbf{w}\mathbf{w}^T] \in \Re^{n_\theta \times n_\theta}$ and $\mathbf{R}_{n+1} = \mathbf{R} = E[\mathbf{v}\mathbf{v}^T] \in \Re^{m \times m}$. $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0|0}$ represents the uncertainty in the initial parameter estimate $\hat{\boldsymbol{\theta}}_{0|0}$. A larger $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0|0}$ implies less confidence (more uncertainty) in $\hat{\boldsymbol{\theta}}_{0|0}$, then the filter relies more on the measured responses ($\mathbf{y}$) than on the prior information of $\hat{\boldsymbol{\theta}}_{0|0}$. This is expected to speed up the convergence rate of the estimation but may adversely influence its stability (e.g., [39]). Larger diagonal entries of $\mathbf{Q}$ increase the estimation uncertainty and larger relative importance is given to the response measurement $\mathbf{y}_{n+1}$ than to the latest prior estimate of the model parameters. If the model uncertainty is neglected, diagonal entries of the matrix $\mathbf{R}$ are estimates of the actual variances of the noise in the different measurements. $\hat{\boldsymbol{\theta}}_{0|0}$, $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0|0}$, $\mathbf{Q}$, and $\mathbf{R}$ are chosen based on experience and engineering judgment.

4. Validation examples

Two numerical examples are presented to validate the proposed approach. In each of them, the response of a soil profile is first simulated using the EQL method with a set of realistic model parameter values, referred to as true model parameter values $\boldsymbol{\theta}^{true}$. Then, different response time histories, referred to as true response $\mathbf{y}^{true}$, are polluted with additive white Gaussian noise and used as measured site responses $\mathbf{y}$ in the parameter estimation stage. The estimation phase assumed unknown model parameters that are estimated with the proposed method.

4.1. Soil profile and earthquake input motions

The analysis considered a 46 m depth soil profile that consists of six sandy and clayey layers (the same profile is examined in the

Initialization: $\hat{\boldsymbol{\theta}}_{0\|0}$ and $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{0\|0}$	Initial estimates of model parameter vector and parameter covariance matrix
for $n = 0,1,2,\ldots$	Loop over time steps
Prediction:	
(i) $\hat{\boldsymbol{\theta}}_{n+1\|n} = \hat{\boldsymbol{\theta}}_{n\|n}$	A priori estimate of model parameter vector
(ii) $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{n+1\|n} = \hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{n\|n} + \mathbf{Q}_n$	A priori estimate of model parameter covariance matrix
(iii) Generate SP $\boldsymbol{\vartheta}^{(i)}_{n+1\|n}$ (i=1,...,$2n_\theta+1$) based on $\hat{\boldsymbol{\theta}}_{n+1\|n}$ and $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{n+1\|n}$	Generate sigma points (SPs)
(iv) $\boldsymbol{\psi}^{(i)}_{n+1} = \mathbf{h}_{n+1}\left(\boldsymbol{\vartheta}^{(i)}_{n+1\|k}, \ddot{\mathbf{u}}_{b_{1:n+1}}\right)$	Output vector for each SP
(v) $\hat{\mathbf{y}}_{n+1\|n} = \sum_{i=1}^{2n_\theta+1} W_m^{(i)} \boldsymbol{\psi}^{(i)}_{n+1}$	Predicted output vector
(vi) $\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1\|n} = \sum_{i=1}^{2n_\theta+1} W_c^{(i)} \left[\boldsymbol{\psi}^{(i)}_{n+1} - \hat{\mathbf{y}}_{n+1\|n}\right]\left[\boldsymbol{\psi}^{(i)}_{n+1} - \hat{\mathbf{y}}_{n+1\|n}\right]^T + \mathbf{R}_{n+1}$	Estimated response covariance matrix
(vii) $\hat{\mathbf{P}}^{\boldsymbol{\theta}\mathbf{y}}_{n+1\|n} = \sum_{i=1}^{2n_\theta+1} W_c^{(i)} \left[\boldsymbol{\vartheta}^{(i)}_{n+1\|n} - \hat{\boldsymbol{\theta}}_{n+1\|n}\right]\left[\boldsymbol{\psi}^{(i)}_{n+1} - \hat{\mathbf{y}}_{n+1\|n}\right]^T$	Estimated cross-covariance matrix
Correction: output measurement $\mathbf{y}_{n+1}$ is recorded	
(viii) $\mathbf{K}_{n+1} = \hat{\mathbf{P}}^{\boldsymbol{\theta}\mathbf{y}}_{n+1\|n}\left(\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1\|n}\right)^{-1}$	Kalman gain
(ix) $\hat{\boldsymbol{\theta}}_{n+1\|n+1} = \hat{\boldsymbol{\theta}}_{n+1\|n} + \mathbf{K}_{n+1}\left(\mathbf{y}_{n+1} - \hat{\mathbf{y}}_{n+1\|n}\right)$	A posteriori estimate of model parameter vector
(x) $\hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{n+1\|n+1} = \hat{\mathbf{P}}^{\boldsymbol{\theta}\boldsymbol{\theta}}_{n+1\|n} - \mathbf{K}_{n+1}\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1\|n}\mathbf{K}^T_{n+1}$	A posteriori estimate of model parameter covariance matrix
end for	

Fig. 3. Proposed method to calibrate one-dimensional EQL models.

Shake91 user's manual [21]). The details of the soil profile and the shear wave velocities V_S of each layer are shown in Fig. 4a. Despite a slight V_S inversion near the surface, the profile gradually increases the stiffness with depth. In Fig. 4b, the G/G_{max} and D versus γ curves for sand with $\sigma'_m = 300$ kPa are plotted. Here, the "true" soil parameters for sand $D_{\min 1}^{sand} = 0.82$, $k^{sand} = 0.316$, $\alpha^{sand} = 0.88$, and $\gamma_{r1}^{sand} = 0.0004$ were calibrated using experimental data reported in [29,30] and the "true" soil type parameters for clay $D_{\min 1}^{clay} = 1.06$, $k^{clay} = 0.207$, $\alpha^{clay} = 0.9$, and $\gamma_{r1}^{clay} = 0.00108$ were calibrated with data reported in [28,30]. To analyze the sensitivity of the shear modulus degradation and damping curves with respect to the soil type parameters (i.e., $D_{\min 1}$, k, α, and γ_{r1}), Fig. 4b shows the curves for sand when each of these parameters takes values of 0.5, 1.0, and 2.0 times its true value with all other parameters kept equal to their true values (i.e., each parameter is varied individually). The continuous lines correspond to the curves with the true parameter value, while dotted and dashed ones are the curves with 0.5 and 2.0 times the true parameter value, respectively. G/G_{max} and D versus γ curves are very sensitive to α and γ_{r1}, but considerably less sensitive to $D_{\min 1}$ and k. Similar results are obtained for shear modulus degradation and damping curves of clays. The input acceleration record corresponds to the 90° component acceleration measured at the Diamond Heights station during the 1989 Loma Prieta Earthquake (Fig. 5). The original 40 s duration record has a peak acceleration of 0.11 g.

4.2. Case study 1: Estimation of shear modulus and damping curves and shear wave velocities using limited acceleration response

4.2.1. Acceleration-only and heterogeneous response measurements

The first case study (CS) analyzes the estimation of the model parameters associated with the modulus reduction and damping curves (i.e., soil type parameters) as well as the layers' initial shear wave velocities. The model parameter vector is defined by $\boldsymbol{\theta} = \left[D_{\min 1}^{clay}, k^{clay}, \alpha^{clay}, \gamma_{r1}^{clay}, D_{\min 1}^{sand}, k^{sand}, \alpha^{sand}, \gamma_{r1}^{sand}, V_{S_1}, V_{S_2}, V_{S_3}, V_{S_4}, V_{S_5}, V_{S_6}, V_{S_7}\right]^T \in \Re^{15\times 1}$. To simulate the response of the soil deposit, the true model parameters are taken as $\boldsymbol{\theta}^{true} = [1.06, 0.207, 0.9, 0.00108, 0.82, 0.316, 0.88, 0.0004, 305, 275, 335, 365, 455, 520, 1220]^T \in \Re^{15\times 1}$, where shear wave velocities are in units of m/s. The EQL method is used to simulate the response of the soil deposit of Fig. 4a when the 90° component acceleration time history recorded at the Diamond Heights during the 1989 Loma Prieta Earthquake (Fig. 5) amplitude-scaled to a peak acceleration (PA) of 0.5 g is used as input at the bedrock. Different response quantities of this true response $\mathbf{y}^{true}$ are then polluted with artificial measurement noise and used to define the response measurement $\mathbf{y}$. Various cases of response measurements are analyzed, when the acceleration response at the surface (CS1_01 and CS1_02 in Table 1), acceleration responses at different depths (CS1_03 in Table 1), and heterogeneous responses (CS1_04 in Table 1) of the soil deposit are measured. The first scenario is the typical situation when only a single seismic station is available at the surface of a soil deposit ($acc1$ in Fig. 4a). The second one corresponds to when a seismic array is deployed in a borehole (acc1 to acc6 in Fig. 4a). Additional scenarios were investigated considering that stresses and strains are measured in addition to acceleration time histories. Each scenario is summarized in Table 1 and the measured responses are depicted in Fig. 4a. Uncorrelated white Gaussian measurement noises with zero-mean and standard deviations (root-mean-square or RMS) of 0.5% g (0.005 g), 1 mm/m, and 50 Pa are used to pollute the acceleration, strain, and stress response time histories, respectively. Note that the amplitude of these measurement

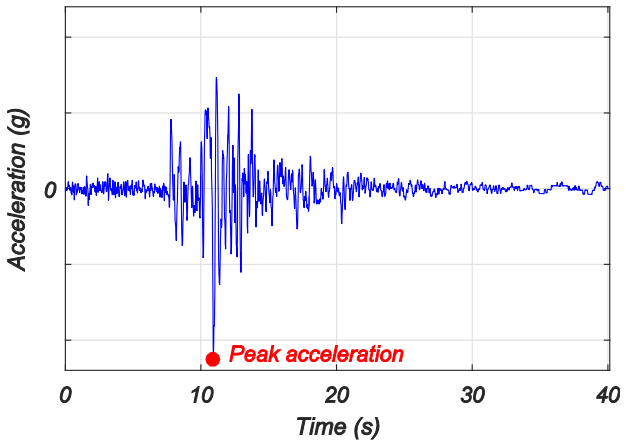

Fig. 5. Acceleration time history recorded at the Diamond Heights station (90° component) during the 1989 Loma Prieta Earthquake (M_W 6.9).

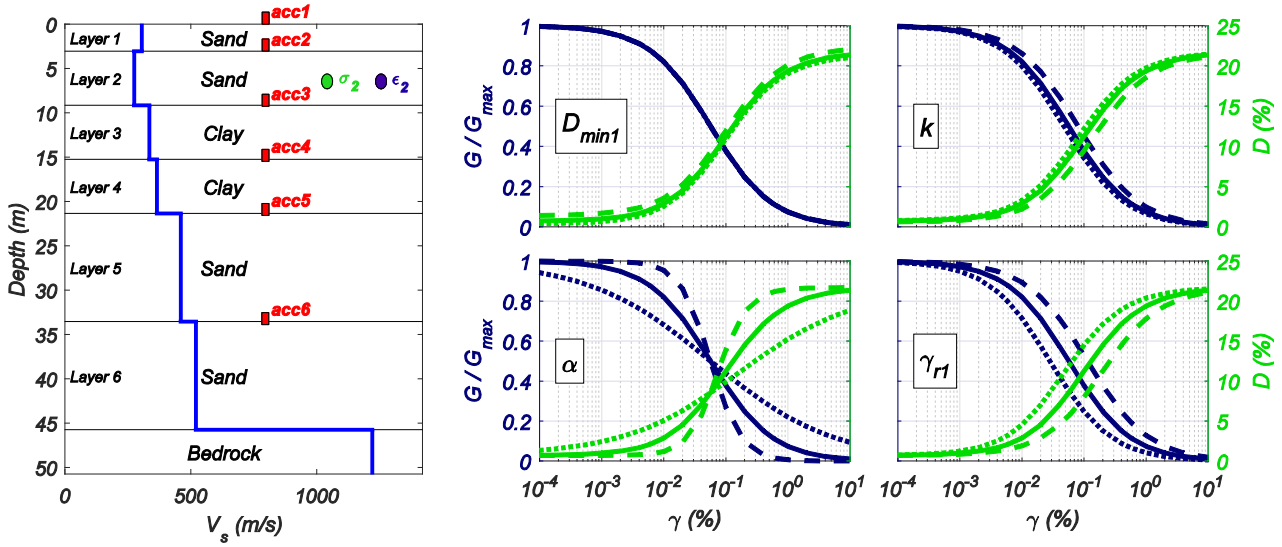

Fig. 4. (a) Soil profile and recorded responses used in the example, (b) Sensitivity of G/G_{max} and D versus γ curves for sand with $\sigma'_m = 300$ kPa (continuous: true value, dotted: 0.5 × true value, dashed: 2.0 × true value).

Table 1
Final posterior estimates and *CV*s (in parenthesis) of the model parameters for CS1.

Case	Output	$\hat{\boldsymbol{\theta}}_{0\|0}$	Final estimates of model parameters $\hat{\theta}_i^*/\theta_i^{true}$ and coefficient of variations (in parenthesis)														
			$D_{\min 1}^{clay}$	k^{clay}	α^{clay}	γ_{r1}^{clay}	$D_{\min 1}^{sand}$	k^{sand}	α^{sand}	γ_{r1}^{sand}	V_{S_1}	V_{S_2}	V_{S_3}	V_{S_4}	V_{S_5}	V_{S_6}	V_{S_7}
			$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$
CS1_01	acc1	$\hat{\boldsymbol{\theta}}_{0\|0_1}$	1.18 (16.8)	1.40 (11.9)	1.45 (10.6)	1.53 (13.3)	1.29 (17.3)	1.23 (16.8)	0.80 (12.3)	1.46 (8.3)	1.18 (16.6)	1.55 (8.3)	0.81 (3.2)	**0.94 (5.1)**	0.82 (3.7)	**0.90 (3.2)**	**0.99 (8.3)**
CS1_02	acc1	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	1.18 (18.4)	**1.02 (16.0)**	**1.05 (14.1)**	**0.99 (13.6)**	1.41 (18.3)	1.18 (18.7)	0.71 (10.1)	0.62 (11.9)	1.28 (18.5)	**1.02 (5.1)**	0.87 (4.8)	1.22 (16.7)	**0.93 (4.6)**	**1.08 (7.7)**	**1.04 (7.9)**
CS1_03	acc1 acc3 acc5	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	1.75 (5.3)	0.31 (3.9)	**1.06 (1.0)**	0.78 (1.2)	1.68 (2.0)	0.84 (5.8)	1.47 (0.27)	1.60 (0.7)	2.14 (8.4)	**0.91 (0.1)**	**0.94 (0.1)**	1.11 (0.2)	**0.92 (0.1)**	**0.95 (0.1)**	**0.91 (0.2)**
CS1_04	acc1 acc3 acc5 str sts	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	1.77 (2.7)	**0.98 (1.7)**	0.84 (0.9)	1.32 (0.7)	1.48 (1.4)	0.33 (3.6)	0.68 (0.18)	0.63 (0.3)	**1.05 (1.0)**	**0.98 (0.1)**	**0.97 (0.1)**	**0.91 (0.1)**	**0.96 (0.1)**	**0.99 (0.1)**	**1.00 (0.2)**

Estimates with an error less than or equal to 10% are shown in bold.

noises are larger than those expected when using current sensing technologies in real-world applications. They are assumed high in order to analyze the robustness of the proposed approach to the output measurement noise.

In the estimation phase, CS1_01 and CS1_02 consider initial estimates of the model parameters $\hat{\boldsymbol{\theta}}_{0|0_1} = 1.3\boldsymbol{\theta}^{true}$ and $\hat{\boldsymbol{\theta}}_{0|0_2} = [1.1, 0.8, 1.2, 0.9, 1.3, 1.2, 0.8, 0.8, 1.3, 1.4, 0.8, 1.2, 0.7, 1.3, 1.2]^T \times \boldsymbol{\theta}^{true}$ (with $\mathbf{a} \times \mathbf{b}$ denoting element-wise multiplication of vectors $\mathbf{a}$ and $\mathbf{b}$), respectively. In both cases, the initial covariance matrix is assumed diagonal with entries computed as $\left(p \times \hat{\theta}_{0|0}^i\right)^2$ with $i = 1, 2, \ldots, 15$ and $p = 0.20$. The diagonal entries of the process noise covariance matrix are taken equal $\left(q \times \hat{\theta}_{0|0}^i\right)^2$ with $i = 1, 2, \ldots, 15$ and $q = 1 \times 10^{-5}$. The diagonal entries of the measurement noise covariance matrix $\mathbf{R}$ are selected as $(0.3\%\,\mathrm{g})^2$, $(0.5\ \mathrm{mm/m})^2$, and $(25\ \mathrm{Pa})^2$. The elements of matrix $\mathbf{R}$ are purposely selected different from the variances of the measurement noises used to contaminate the true responses, because only an estimate of the amplitude of measurement noise is possible in real-life applications.

Table 1 summarizes final posterior model parameter estimates $\left(\hat{\boldsymbol{\theta}}^* = \hat{\boldsymbol{\theta}}_{N|N}\right)$ normalized with respect to the true parameter values and the final posterior coefficient of variation (*CV*) estimates for the four case studies. If only the acceleration at the surface is recorded (CS1_01 and CS1_02), the shear wave velocities of the deeper layers (V_{S_4} to V_{S_7}) are reasonably well estimated, while soil type parameters for clay and sand, as well as shear wave velocities of shallow layers are not well estimated and their *CV*s remain high. From cases CS1_01 and CS1_02 is observed that estimation results are consistent when different initial model parameter estimates $\hat{\boldsymbol{\theta}}_{0|0}$ are considered. If acceleration response at layers 3 and 5 (*acc*3 and *acc*5 in Fig. 4a) are considered in addition to the surface acceleration response (CS1_03), estimation of shear wave velocities improves, approaching their true values and reducing the estimation uncertainty (*CV*), except for V_{S_1}. When strain and stress in layer 2 (σ_2 and ε_2 in Fig. 4a) are considered as measured response in addition to acceleration responses at layers 1, 3 and 5 (CS1_04), shear wave velocities of all layers approach their true values with relative errors less than or equal to 5%. However, estimation of soil type parameters is still not accurate in case CS1_04. Parameters with good final estimate but with a high *CV* mean that the estimation is not reliable because a high estimation uncertainty is associated with the model parameter estimates.

To analyze the sensitivity of the different output response measurements considered (accelerations, strain, and stress) with respect to the fifteen model parameters to be estimated, tornado diagrams are employed [40]. In the tornado diagram analysis, each model parameter is perturbed from its true value to lower and upper bounds – selected here as 0.9 and 1.1 times the true model parameter value, respectively – keeping the other parameter values fixed (equal to the true values), i.e., each model parameter is varied individually. Then, the response for each perturbed model parameter is computed by subjecting the soil deposit model to the Diamond Heights record (Fig. 5) amplitude-scaled to a *PA* of 0.5 g. The range of response obtained for different perturbations, called *swing*, is considered as a measure of sensitivity of the response with respect to the perturbed model parameter. Here, the swing in the response is evaluated using the relative root-mean-square error (*RRMSE*) between the response with the perturbed parameter and the response with the true parameter. The *RRMSE* between vectors $\mathbf{a} \in \Re^{L\times 1}$ and $\mathbf{b} \in \Re^{L\times 1}$ is defined by

$$RRMSE(\mathbf{a}, \mathbf{b}) = \frac{\sqrt{1/L\sum_{k=1}^{L}(a_k - b_k)^2}}{\sqrt{1/L\sum_{k=1}^{L}(a_k)^2}} \times 100\ [\%] \tag{10}$$

Fig. 6a shows the tornado diagram for the surface acceleration and stress in layer 2 responses. The surface acceleration response is mostly sensitive to model parameters V_{S_6}, V_{S_5}, and V_{S_2}, while it is practically insensitive to parameters k^{clay}, k^{sand}, α^{clay}, V_{S_1}, $D_{\min 1}^{clay}$, and $D_{\min 1}^{sand}$. The stress in layer 2 is more sensitive (Fig. 6b) to all parameters compared to surface acceleration, but a similar order of sensitivities for both responses is observed. Those parameters with higher sensitives are expected to be better estimated because more information about them is contained in the measured responses. On the other hand, model parameters with low sensitivities are non-identifiable because the measured responses contain very little information about them. This is confirmed by the estimation results presented in Table 1, e.g., for CS1_04, all the shear wave velocities, except V_{S_1}, are well estimated with low *CV*s, while clay parameters are not well identified and/or have high *CV* estimates.

To check the proper calibration of the soil profile model, Table 2 reports the *RRMSE*s between the true responses and the estimated responses based on the initial estimate of the model parameters ($\hat{\boldsymbol{\theta}}_{0|0}$) and based on the final posterior estimate of the model parameters ($\hat{\boldsymbol{\theta}}^*$). The large misfit between the true and estimated

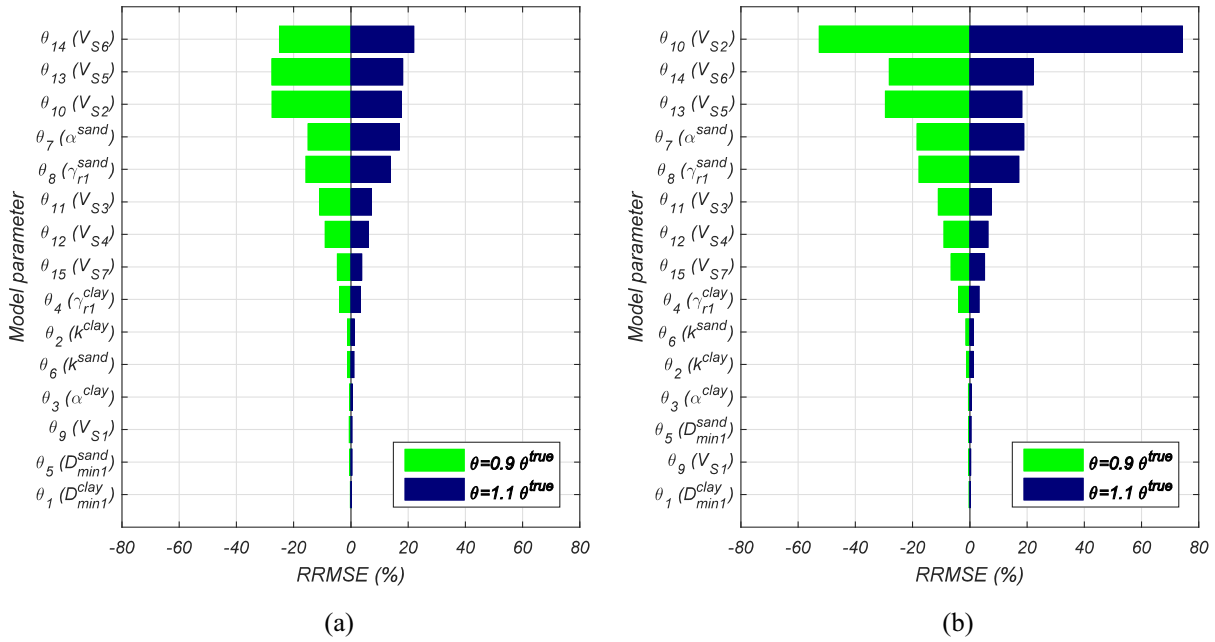

Fig. 6. Tornado diagram of *RRMSE* of response measurements with $0.90\boldsymbol{\theta}^{true}$ and $1.10\boldsymbol{\theta}^{true}$ with respect to $\boldsymbol{\theta}^{true}$ (a) surface acceleration (acc1 in Fig. 4a) and (b) stress in layer 2 (σ^2 in Fig. 4a).

Table 2
RRMSE (in %) between the true responses and the corresponding model predictions based on the final estimates of the model parameters for case study 1.

Case	$\hat{\boldsymbol{\theta}}$	Output response measurement ($\mathbf{y}_i$)				
		acc1	acc3	acc5	ε_2	σ_2
CS1_01	$\hat{\boldsymbol{\theta}}_{0\|0_1}$	145.7	–	–	–	–
	$\hat{\boldsymbol{\theta}}^*$	15.3	–	–	–	–
CS1_02	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	62.1	–	–	–	–
	$\hat{\boldsymbol{\theta}}^*$	6.2	–	–	–	–
CS1_03	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	62.1	68.7	64.7	–	–
	$\hat{\boldsymbol{\theta}}^*$	6.9	5.9	7.3	–	–
CS1_04	$\hat{\boldsymbol{\theta}}_{0\|0_2}$	62.1	68.7	64.7	86.1	57.8
	$\hat{\boldsymbol{\theta}}^*$	2.0	2.0	2.7	2.3	2.2

response with $\hat{\boldsymbol{\theta}}_{0|0}$ ($57.8\% \leq RRMSE \leq 145.7\%$), is significantly reduced after the model is updated. The *RRMSE*s between the true and the estimated response with $\hat{\boldsymbol{\theta}}^*$ are less than or equal to 15.3% and the effect of adding additional response measurements in improving the calibration of the model is clearly observed, e.g., in CS1_04 the *RRMSE*s are less than or equal to 2.7%, which implies an almost exact match between the true response and the estimated response with $\hat{\boldsymbol{\theta}}^*$. Fig. 7 compares the surface acceleration response based on the true, initial, and final estimated model parameters for CS1_02. The excellent agreement between the response computed from the final estimated parameters and the corresponding true response proves the successful updating of the model.

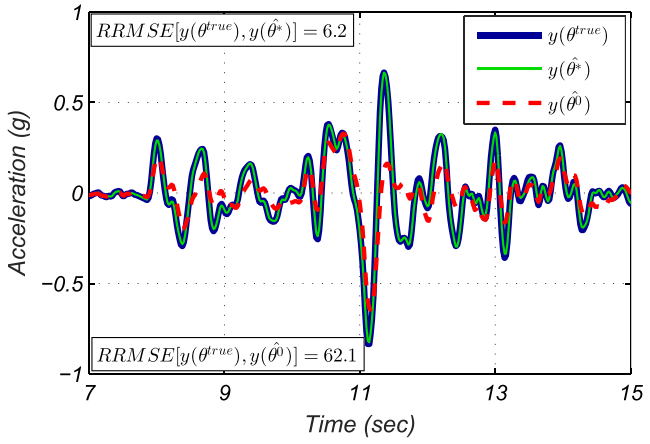

Fig. 7. Comparison of true and estimated responses with $\hat{\boldsymbol{\theta}}_{0|0}$ and $\hat{\boldsymbol{\theta}}^*$ for CS1_02.

4.2.2. Verification analysis

In this section, the prediction capabilities of the models calibrated in previous section are investigated by considering a different input motion at the bedrock. Component 180° recorded at 252 m depth at La Cienaga station during the 2001 Hollywood Earthquake (Fig. 8a) amplitude-scaled to *PA* = 0.4 g is used as excitation at the bedrock. The intensity of this excitation is similar to that used for model calibration in Section 4.2.1 (90° component at Diamond Heights station during the 1989 Loma Prieta Earthquake with *PA* = 0.5 g). The true response of the soil deposit (i.e., using $\boldsymbol{\theta}^{true}$) and the predicted response using the final estimates of the model parameters (i.e., $\hat{\boldsymbol{\theta}}^*$) obtained in CS1_01 are simulated using La Cienaga record amplitude-scaled to a *PA* of 0.4 g as input

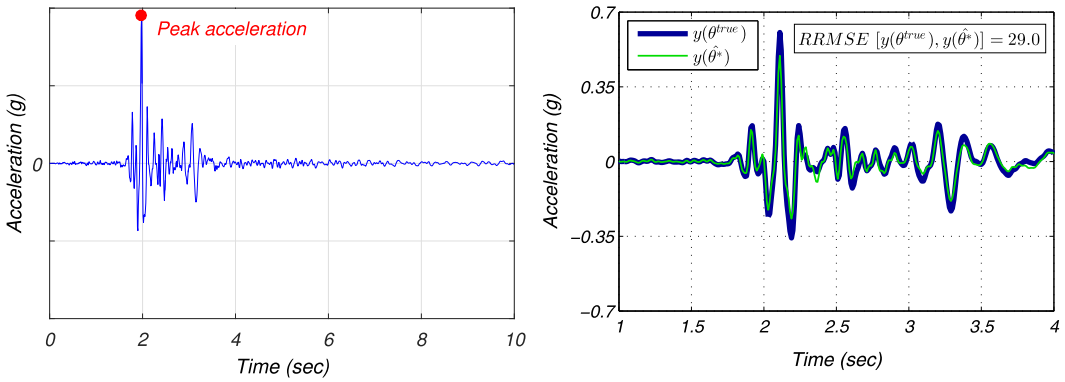

Fig. 8. (a) Acceleration time history recorded at La Cienaga Geotechnical Array (180° component at 252 m depth) during the 2001 Hollywood Earthquake and (b) Comparison of the true and estimated (with $\hat{\boldsymbol{\theta}}^{*}$ obtained in CS1_01) surface acceleration response for La Cienaga input with *PA* = 0.4 g (time window 1.0–4.0 s).

Table 3
Final estimate and *CV* (in parenthesis) of the model parameters for case study 2.

Case	Output	Final estimates of model parameters $\hat{\theta}_i^{*}/\theta_i^{true}$ and coefficient of variations (in parenthesis)																				
		$D_{\min 1}^{clay}$ $i=1$	k^{clay} $i=2$	α^{clay} $i=3$	γ_{r1}^{clay} $i=4$	$D_{\min 1}^{sand}$ $i=5$	k^{sand} $i=6$	α^{sand} $i=7$	γ_{r1}^{sand} $i=8$	V_{S_1} $i=9$	V_{S_2} $i=10$	V_{S_3} $i=11$	V_{S_4} $i=12$	V_{S_5} $i=13$	V_{S_6} $i=14$	V_{S_7} $i=15$	h_1 $i=16$	h_2 $i=17$	h_3 $i=18$	h_4 $i=19$	h_5 $i=20$	h_6 $i=21$
CS2_01	acc1	1.36 (17.2)	0.66 (16.3)	1.21 (16.7)	0.80 (13.6)	1.22 (17.4)	1.26 (18.5)	0.60 (14.4)	0.69 (14.1)	1.36 (17.7)	1.13 (6.4)	**1.03 (5.9)**	**1.09 (8.8)**	**1.01 (6.9)**	1.24 (7.5)	1.11 (6.0)	0.75 (16.8)	1.26 (15.3)	1.33 (16.4)	**1.10 (18.2)**	0.79 (14.6)	1.31 (13.7)
CS2_02	acc1 acc4 acc6	1.51 (1.4)	1.25 (1.9)	1.11 (1.5)	1.14 (1.1)	1.42 (2.3)	0.84 (2.6)	0.46 (0.8)	**0.92 (1.1)**	1.30 (1.7)	**1.00 (0.1)**	**0.95 (0.1)**	**1.09 (0.2)**	**1.03 (0.1)**	**1.04 (0.1)**	**1.07 (0.3)**	**1.04 (2.6)**	0.74 (1.0)	1.30 (1.1)	**1.06 (6.4)**	1.15 (3.4)	**1.03 (0.4)**

Estimates with an error less than or equal to 10% are shown in bold.

excitation. Fig. 8b depicts the comparison between the true and estimated (with $\hat{\boldsymbol{\theta}}^{*}$ obtained in CS1_01) surface acceleration responses. A good agreement between both responses is observed with a $RRMSE[\mathbf{y}(\boldsymbol{\theta}^{true}), \mathbf{y}(\hat{\boldsymbol{\theta}}_{CS1\_01}^{*})] = 29.0\%$, proving that the model updated for a given earthquake excitation can be successfully used to predict the response of the deposit for a different earthquake. It is noted that the error between the true response and the response obtained with the initial model parameters in CS1_01 is much larger, with a $RRMSE\left[\mathbf{y}(\boldsymbol{\theta}^{true}), \mathbf{y}(\hat{\boldsymbol{\theta}}_{0|0_1})\right] = 172.7\%$.

4.3. Case study 2: Estimation of shear modulus reduction and damping curves, shear wave velocities, and layers thicknesses using limited acceleration response

4.3.1. Acceleration-only response measurements

This case study analyzes the identification of the modulus reduction and damping curves, the layers' initial shear wave velocities, and the layers' thickness when the acceleration time history is known at the surface of the soil deposit (CS2_01 in Table 3) and at different depths (CS2_02 in Table 3). The model parameter vector $\boldsymbol{\theta} = [D_{\min 1}^{clay}, k^{clay}, \alpha^{clay}, \gamma_{r1}^{clay}, D_{\min 1}^{sand}, k^{sand}, \alpha^{sand}, \gamma_{r1}^{sand}, V_{S_1}, V_{S_2}, V_{S_3}, V_{S_4}, V_{S_5}, V_{S_6}, V_{S_7}, h_1, h_2, h_3, h_4, h_5, h_6]^T \in \Re^{21\times 1}$ is then identified using $\mathbf{y}$ and the input excitation. The true model parameters used to simulate the true response of the soil deposit are $\boldsymbol{\theta}^{true} = [1.06, 0.207, 0.9, 0.00108, 0.82, 0.316, 0.88, 0.0004, 305, 275, 335, 365, 455, 520, 1220, 10, 20, 20, 20, 40, 40]^T \in \Re^{21\times 1}$ where shear wave velocities (entries 9 to 15) are in m/s and layers' thickness (entries 16 to 21) in m. The 90° component at the Diamond Heights station recorded in the 1989 Loma Prieta Earthquake (Fig. 5) amplitude-scaled to a peak acceleration (*PA*) of 0.3 g is used as excitation at the bedrock. Uncorrelated white Gaussian measurement noises with zero-mean and 0.5% g *RMS* are used to contaminate the true acceleration responses and define the measured response $\mathbf{y}$.

CS2_01 considers only the acceleration response at the surface of the soil profile, while CS2_02 assumes that acceleration responses are recorded at the surface, layer 4, and layer 6 (see Table 3). The estimation of the thickness of the layers implies that the total depth of the soil deposit is unknown, which can be the case in some real-world applications. If the total depth of the deposit is known, a constraint on the total height of the deposit (sum of the heights of the soil layers) can be added by using a constrained nonlinear filter (e.g., [41,42]).

In the estimation, $\hat{\boldsymbol{\theta}}_{0|0} = [1.1, 0.8, 1.2, 0.9, 1.3, 1.2, 0.8, 0.8, 1.3, 1.4, 0.8, 1.2, 0.7, 1.3, 1.2, 0.8, 1.2, 1.3, 1.1, 0.8, 1.3]^T \times \boldsymbol{\theta}^{true}$ is assumed as the initial estimate of the model parameters and the diagonal entries of $\hat{\mathbf{P}}_{0|0}^{\boldsymbol{\theta\theta}}$ are taken as $(p \times \hat{\theta}_{0|0}^{i})^2$ with $i = 1, 2, \ldots, 21$ and $p = 0.20$. The diagonal entries of $\mathbf{Q}$ are taken equal $(q \times \hat{\theta}_{0|0}^{i})^2$ with $i = 1, 2, \ldots, 21$ and $q = 1 \times 10^{-5}$ and the diagonal entries of $\mathbf{R}$ are selected as $(0.3\%\ \mathrm{g})^2$. Table 3 presents the normalized final posterior model parameter estimates and the final posterior *CV* estimates for CS2_01 and CS2_02. Estimation results are consistent with those presented in Section 4.2.1. Shear wave velocities are properly estimated and soil-type parameters are not accurately estimated because the measured responses are not very sensitive to these parameters. Estimation of layers' thickness improves in CS2_02 compared to CS2_01 but their final *CV* estimates are still large in some cases.

Figs. 9 and 10 show the time history of the posterior estimate and *CV* of the model parameters for CS2_01 and CS2_02. It can be observed for any given model parameter that the estimation

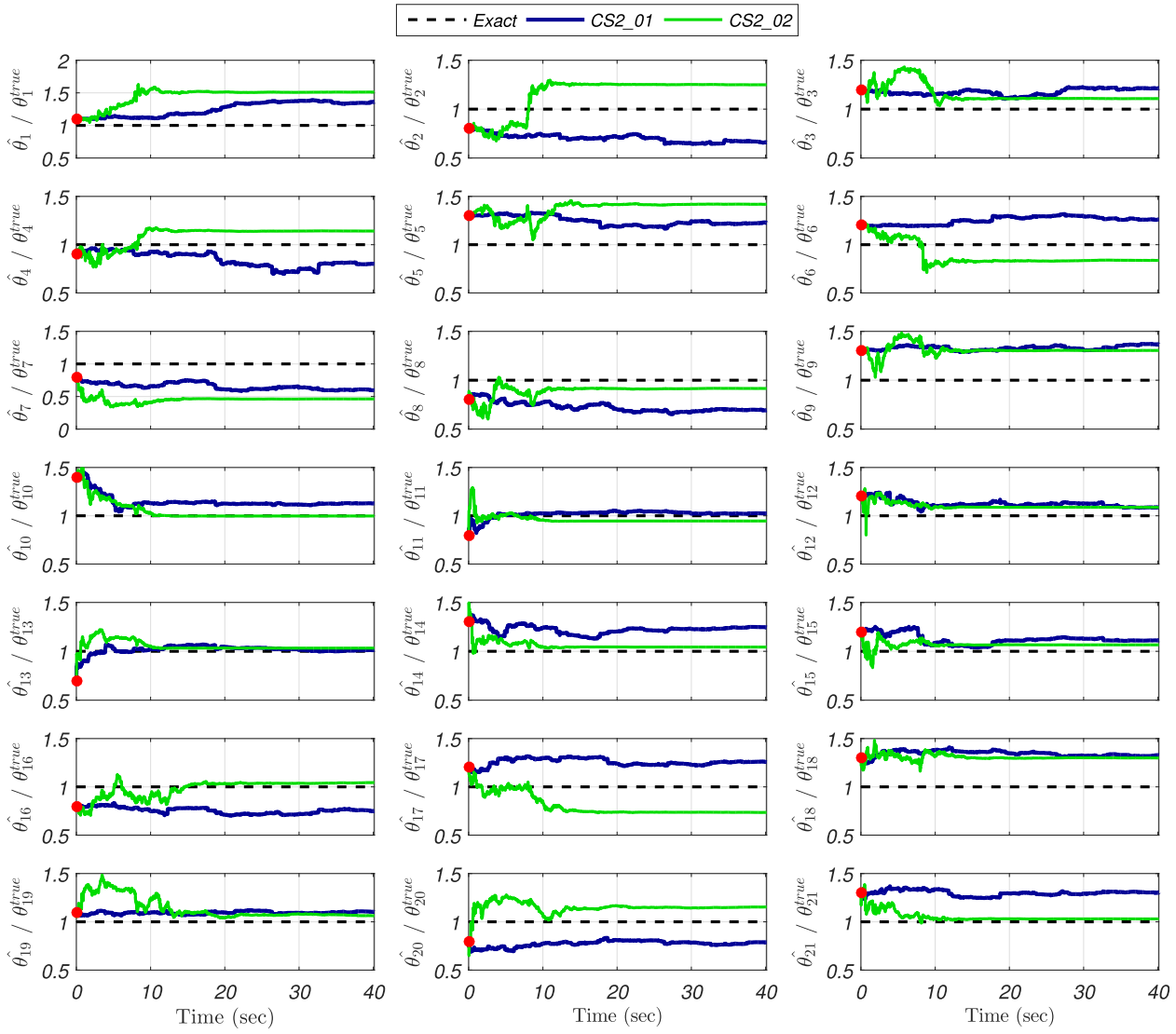

Fig. 9. Time history of the posterior estimates of the model parameters for CS2_01 and CS2_02.

uncertainty is always lower in CS2_02 than in CS2_01 because more information is contained in the response measurements considered in CS2_02. Although all parameters start updating from the beginning of the time history, most of them are not sensitive enough to the measured response and therefore their final estimates do not converge to the true values and the final *CV*s estimates are large. This observation does not necessarily imply poor model predictability.

Table 4 summarizes the *RRMSE*s between the true responses and those estimated based on $\hat{\boldsymbol{\theta}}^{*}$. An excellent agreement between these responses is observed and the significant improvement compared to the estimated response based on $\hat{\boldsymbol{\theta}}_{0|0}$ is evident, confirming the proper calibration of the soil profile model. Fig. 11 depicts the comparison between the true and estimated surface acceleration responses with $\hat{\boldsymbol{\theta}}_{0|0}$ and $\hat{\boldsymbol{\theta}}^{*}$ for CS2_01, where a reduction in the *RRMSE* from 85.8 to 4.8% confirms the accurate updating of the model.

4.3.2. Verification analysis

As in Section 4.2.2, a verification analysis is conducted for CS2. Here, the acceleration time history recorded at La Cienaga (Fig. 8a) with a *PA* = 0.8 g is used as excitation at the bedrock. This excitation is considerably larger than that used for model calibration in Section 4.3.1 and therefore allows investigating the prediction capabilities when the model is calibrated with an earthquake excitation of smaller amplitude and different characteristics, which is usually the case in engineering practice. Fig. 12 compares the true surface acceleration response of the soil deposit with the estimated response based on $\hat{\boldsymbol{\theta}}^{*}$ obtained in CS2_01. An excellent agreement is observed, with $RRMSE\left[\mathbf{y}(\boldsymbol{\theta}^{true}), \mathbf{y}(\hat{\boldsymbol{\theta}}^{*}_{\text{CS2\_01}})\right] = 17.9\%$.

5. Conclusions

This paper presented a novel approach that uses the Bayesian inference methodology to calibrate numerical models to perform site response analysis, fundamental in assessing seismic hazard and earthquake damage. The approach estimates the expected values and covariance matrix of model parameters, such as shear wave velocity, damping ratio, and thickness of the constituent layers, that define the numerical model. The equivalent-linear (EQL) method is used as the numerical model because of its popularity and wide usage. Synthetic data recorded on downhole arrays

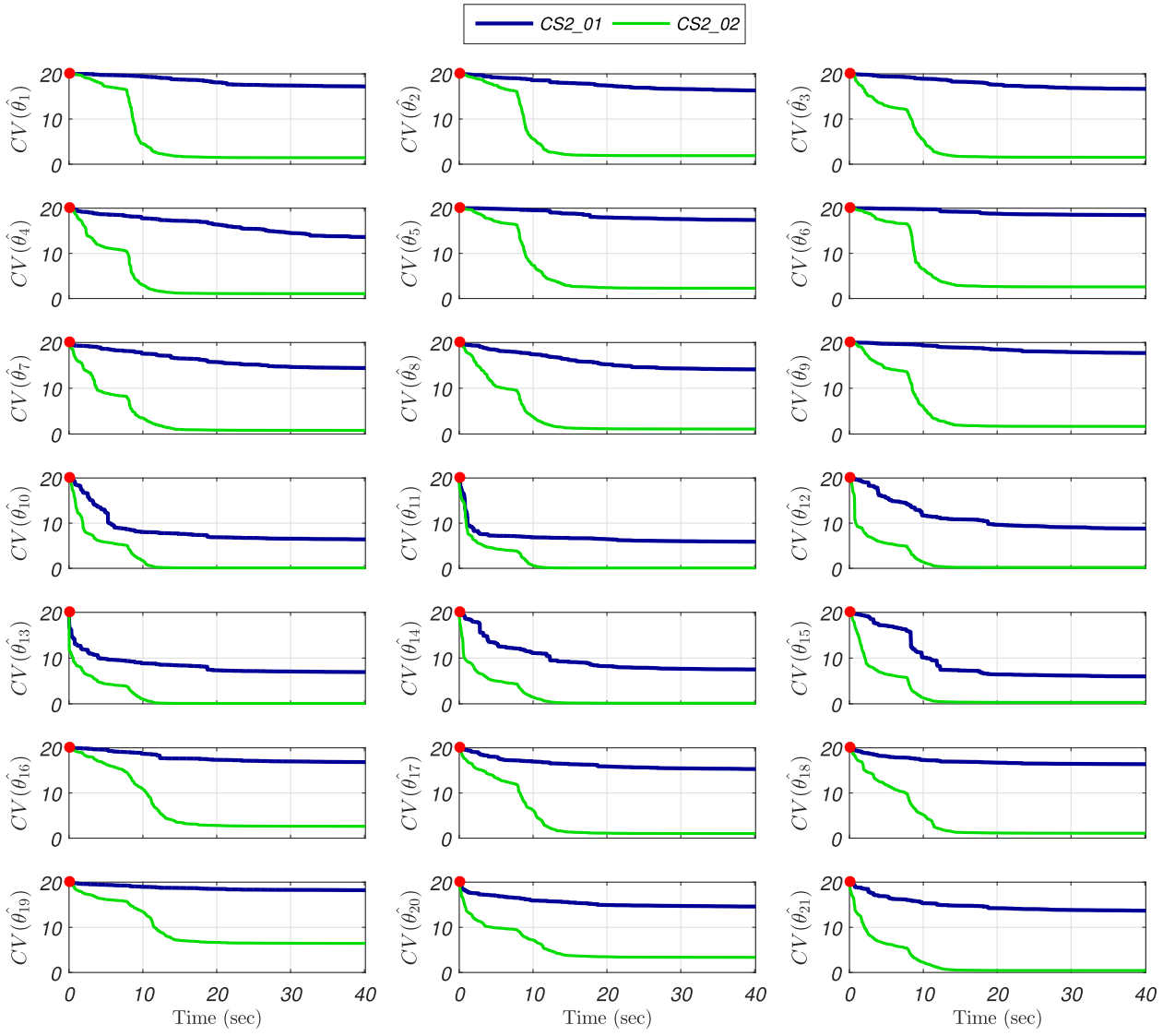

Fig. 10. Time history of the posterior coefficient of variation of the model parameters for CS2_01 and CS2_02.

Table 4
RRMSE (in %) between the true responses and the corresponding model predictions based on the final estimates of the model parameters for case study 2.

Case	$\hat{\boldsymbol{\theta}}$	Output response measurement ($\mathbf{y}_i$)				
		$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$	$\mathbf{a}_5$
CS2_01	$\hat{\boldsymbol{\theta}}_{0\|0}$	85.8	–	–	–	–
	$\hat{\boldsymbol{\theta}}^*$	4.8	–	–	–	–
CS2_02	$\hat{\boldsymbol{\theta}}_{0\|0}$	85.8	89.5	69.9	–	–
	$\hat{\boldsymbol{\theta}}^*$	1.2	1.1	2.2	–	–

during an earthquake are employed to estimate the parameters and calibrate the numerical model of the soil deposit. The unscented Kalman filter is employed as estimation tool, facilitating the identification of any model parameter, because the response sensitivities with respect to the parameters to be estimated are not required. Two numerical applications show the performance and prediction capabilities of the proposed approach. It is concluded that the numerical model is accurately updated, even if a very large number of model parameters need to be estimated using a very low number of response measurements. In addition, the

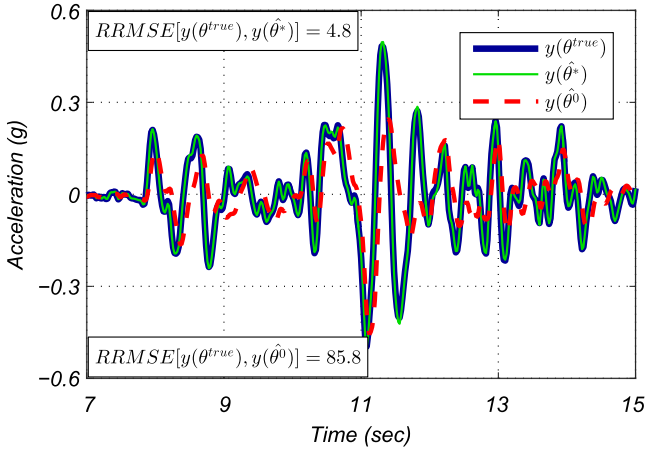

Fig. 11. Comparison of true and estimated responses with $\hat{\boldsymbol{\theta}}_0$ and $\hat{\boldsymbol{\theta}}^*$ for CS2_01.

effects of heterogeneous response measurement on the estimation results are presented. Although the proposed framework is used with the EQL method, it has the capability to be employed with

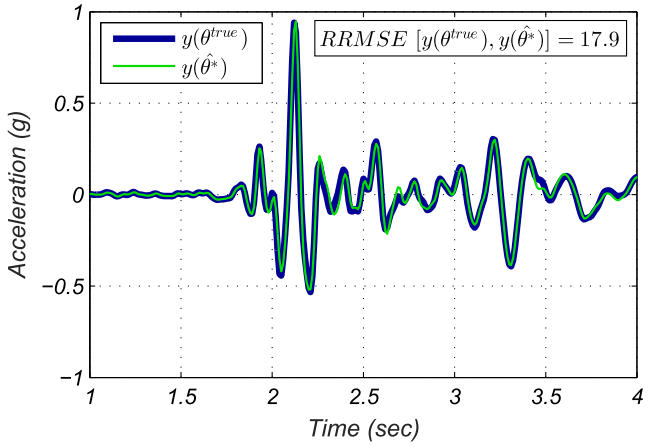

Fig. 12. Comparison of true and estimated (with $\hat{\boldsymbol{\theta}}^{*}$ obtained in CS2_01) surface acceleration response for La Cienaga input with PA = 0.8 g (time window 1.0–4.0 s).

other site response analysis techniques, such as fully nonlinear models. In future work, the presented approach will be used with experimental data recorded from downhole arrays.

Acknowledgements

R. Astroza acknowledges the financial support from the Universidad de los Andes – Chile through the research grant *Fondo de Ayuda a la Investigación* (FAI) and from the Chilean National Commission for Scientific and Technological Research (CONICYT), FONDECYT-Iniciación research project No. 11160009. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect those of the sponsors.

Appendix A

Consider a nonlinear discrete-time state-space model:

$$\mathbf{x}_{n+1} = \mathbf{f}_n(\mathbf{x}_n, \mathbf{u}_n) + \mathbf{w}_n \tag{A1}$$

$$\mathbf{y}_{n+1} = \mathbf{h}_{n+1}(\mathbf{x}_{n+1}, \mathbf{u}_{n+1}) + \mathbf{v}_{n+1} \tag{A2}$$

where $\mathbf{x}_n \in \Re^{n_x}$, $\mathbf{u}_n \in \Re^{n_u}$, and $\mathbf{y}_n \in \Re^{n_y}$ are the unknown system state, input, and measurement vectors at time t_n, respectively, $\mathbf{w}_n$ and $\mathbf{v}_n$ are mutually independent, zero-mean white Gaussian random vectors with covariance matrices $\mathbf{Q}_n$ and $\mathbf{R}_n$, respectively, and $\mathbf{f}_n$ and $\mathbf{h}_n$ are deterministic and known nonlinear vector-valued functions. It is assumed that the probability density functions (*PDFs*) $p(\mathbf{x}_{n+1}|\mathbf{y}_{1:n}) = \mathcal{N}\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n}\right)$ and $p(\mathbf{y}_{n+1}|\mathbf{y}_{1:n})$ are Gaussian (with $\mathbf{y}_{1:n} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_n^T\right]^T$), where $\mathcal{N}\left(\mathbf{z}; \hat{\mathbf{z}}, \hat{\mathbf{P}}^{\mathbf{zz}}\right)$ denotes a multivariate Gaussian distribution for the random vector $\mathbf{z}$ and $\hat{\mathbf{z}}$ and $\hat{\mathbf{P}}^{\mathbf{zz}}$ denote the mean and covariance matrix of $\mathbf{z}$. The mean $\hat{\mathbf{x}}_{n+1|n}$ and covariance matrix $\hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n}$ can be derived as

$$\hat{\mathbf{x}}_{n+1|n} = \int \mathbf{f}_n(\mathbf{x}_n, \mathbf{u}_n)\mathcal{N}\left(\mathbf{x}_n; \hat{\mathbf{x}}_{n|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n|n}\right)d\mathbf{x}_n \tag{A3}$$

$$\hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n} = \int \mathbf{f}_n(\mathbf{x}_n, \mathbf{u}_n)\mathbf{f}_n^T(\mathbf{x}_n, \mathbf{u}_n)\mathcal{N}\left(\mathbf{x}_n; \hat{\mathbf{x}}_{n|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n|n}\right)d\mathbf{x}_n - \hat{\mathbf{x}}_{n+1|n}\hat{\mathbf{x}}^T_{n+1|n} + \mathbf{Q}_n \tag{A4}$$

The predicted mean and covariance matrix of $\mathbf{y}_{n+1}$ given $\mathbf{y}_{1:n}$ can be expressed as

$$\hat{\mathbf{y}}_{n+1|n} = \int \mathbf{h}_{n+1}(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\mathcal{N}\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n}\right)d\mathbf{x}_{n+1} \tag{A5}$$

$$\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1|n} = \int \mathbf{h}_{n+1}(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\mathbf{h}^T_{n+1}(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\mathcal{N}\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n}\right)d\mathbf{x}_{n+1} - \hat{\mathbf{y}}_{n+1|n}\hat{\mathbf{y}}^T_{n+1|n} + \mathbf{R}_{n+1} \tag{A6}$$

Similarly, the cross-covariance matrix $\hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n}$ can be obtained as

$$\hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n} = \int \mathbf{x}_{n+1}\mathbf{h}^T_{n+1}(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\mathcal{N}\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n}\right)d\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\hat{\mathbf{y}}^T_{n+1|n} \tag{A7}$$

Based on the previous assumptions, it follows that $p(\mathbf{x}_{n+1}|\mathbf{y}_{1:n+1})$ is also Gaussian. Then, it follows that

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{1:n+1}) = \mathcal{N}\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n+1}, \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n+1}\right) \tag{A8}$$

with

$$\hat{\mathbf{x}}_{n+1|n+1} = \hat{\mathbf{x}}_{n+1|n} + \hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n}\left(\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1|n}\right)^{1}\left(\mathbf{y}_{n+1} - \hat{\mathbf{y}}_{n+1|n}\right) \tag{A9}$$

$$\hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n+1} = \hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n} - \hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n}\left(\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1|n}\right)^{1}\left(\hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n}\right)^{T} \tag{A10}$$

To evaluate the integrals in (A3)–(A7), the Unscented Kalman Filter (UKF) employs the unscented transformation (UT), which expresses a random vector $\mathbf{s}$ by a set of deterministically chosen sample points (referred to as sigma points or SPs) such that the weighted sample mean and weighted sample covariance matrix of the SPs match exactly the true mean and covariance matrix of the random vector $\mathbf{s}$. If the SPs are propagated through a nonlinear function, they capture the true mean and covariance matrix up to the second order of the Taylor series expansion of the nonlinear function [35]. The UKF evaluates (A3)–(A7) as follows:

$$\hat{\mathbf{x}}_{n+1|n} = \sum_{i=1}^{2n_x} W_m^{(i)}\boldsymbol{\chi}^{(i)}_{n+1|n} \tag{A11}$$

$$\hat{\mathbf{P}}^{\mathbf{xx}}_{n+1|n} = \sum_{i=1}^{2n_x} W_c^{(i)}\left[\boldsymbol{\chi}^{(i)}_{n+1|n} - \hat{\mathbf{x}}_{n+1|n}\right]\left[\boldsymbol{\chi}^{(i)}_{n+1|n} - \hat{\mathbf{x}}_{n+1|n}\right]^T + \mathbf{Q}_n \tag{A12}$$

$$\hat{\mathbf{y}}_{n+1|n} = \sum_{i=1}^{2n_x} W_m^{(i)}\mathbf{h}_{n+1}\left(\boldsymbol{\chi}^{(i)}_{n+1|n}\right) \tag{A13}$$

$$\hat{\mathbf{P}}^{\mathbf{yy}}_{n+1|n} = \sum_{i=1}^{2n_x} W_c^{(i)}\left[\mathbf{h}_{n+1}(\boldsymbol{\chi}^{(i)}_{n+1|n}) - \hat{\mathbf{y}}_{n+1|n}\right]\left[\mathbf{h}_{n+1}(\boldsymbol{\chi}^{(i)}_{n+1|n}) - \hat{\mathbf{y}}_{n+1|n}\right]^T + \mathbf{R}_{n+1} \tag{A14}$$

$$\hat{\mathbf{P}}^{\mathbf{xy}}_{n+1|n} = \sum_{i=1}^{2n_x} W_c^{(i)}\left[\boldsymbol{\chi}^{(i)}_{n+1|n} - \hat{\mathbf{x}}_{n+1|n}\right]\left[\mathbf{h}_{n+1}(\boldsymbol{\chi}^{(i)}_{n+1|n}) - \hat{\mathbf{y}}_{n+1|n}\right]^T \tag{A15}$$

where the weighting coefficients of the SPs to estimate the means and covariance matrices, $W_m^{(i)}$ and $W_c^{(i)}$ respectively, are

$$W_m^{(0)} = \frac{\lambda}{n_x + \lambda}; \quad W_c^{(0)} = \frac{\lambda}{n_x + \lambda} + (1 - \tilde{\alpha}^2 + \tilde{\beta});$$
$$W_m^{(i)} = W_c^{(i)} = \frac{1}{2(n_x + \lambda)}, \quad i = 1, \ldots, 2n_x \tag{A16}$$

and the SPs are calculated as

$$\boldsymbol{\chi}^{(i)}_{n+1|n} = \begin{cases} \hat{\mathbf{x}}_{n|n} & \text{if } i = 0 \\ \hat{\mathbf{x}}_{n|n} + \left[\left(\gamma\sqrt{\hat{\mathbf{P}}^{\mathbf{nxx}}_{n|}}\right)_i\right]^T & \text{if } i = 1, \ldots, n_x \\ \hat{\mathbf{x}}_{n|n} - \left[\left(\gamma\sqrt{\hat{\mathbf{P}}^{\mathbf{nxx}}_{n|}}\right)_i\right]^T & \text{if } i = n_x + 1, \ldots, 2n_x \end{cases} \tag{A17}$$

with $\sqrt{\mathbf{P}}$ = square-root of the covariance matrix $\mathbf{P}$, $(\cdots)_i$ represents the ith row of the matrix inside the parentheses, $\gamma = \sqrt{n_x + \lambda}$, $\lambda = \tilde{\alpha}^2(n_x + \kappa) - n_x$, $\tilde{\alpha} \in \left[10^{-4}, 1\right]$ is a constant related to the spread of the SPs around the mean, and κ is a secondary scaling parameter.

References

[1] Kramer SL. Geotechnical earthquake engineering. New Jersey: Prentice Hall; 1996.
[2] Hashash Y, Groholski DR, Phillips C. Recent advances in non-linear site response analysis. In: 5th international conference on recent advances in geotechnical earthquake engineering and soil dynamics, San Diego, CA; 2010.
[3] Kramer SL, Paulsen SB. Practical use of geotechnical site response models. In: International workshop on uncertainties in nonlinear soil properties and their impact on modeling dynamic soil response, PEER Center Headquarters, Richmond, CA; 2004.
[4] Kaklamanos J, Bradley BA, Thompson EM, Baise LG. Critical parameters affecting bias and variability in site-response analyses using Kik-net. Bull Seismol Soc Am 2013;103(3):1733–49.
[5] Kim B, Hashash YMA, Stewart JP, Rathje EM, Harmon JA, Musgrove MI, et al. Relative differences between nonlinear and equivalent-linear 1-D site response analyses. Earthq Spectra 2016;32(3):1845–65.
[6] Régnier J, Bonilla L-F, Bard P-Y, et al. International benchmark on numerical simulations for 1D, nonlinear site response (PRENOLIN): Verification phase based on canonical cases. Bull Seismol Soc Am 2016;106(5):2112–35.
[7] Kaklamanos J, Baise LG, Thompson EM, Dorfmann L. Comparison of 1D linear, equivalent-linear, and nonlinear site response models at six KiK-net validation sites. Soil Dyn Earthq Eng 2015;69:207–19.
[8] Zalachoris G, Rathje EM. Evaluation of one-dimensional site response techniques using borehole arrays. J Geotech Geoenviron Eng 2015;141(12):04015053.
[9] Glaser S. System identification and its application to estimating soil properties. J Geotech Eng 1995;121(7):553–60.
[10] Elgamal A, Zeghal M, Parra E, Gunturi R, Tang HT, Stepp JC. Identification and modeling of earthquake ground response I. Site amplification. Soil Dyn Earthq Eng 1996;15(8):499–522.
[11] Zeghal M, Elgamal A, Tang H, Stepp J. Lotung downhole array II: evaluation of soil nonlinear properties. J Geotech Eng 1995;121(4):363–78.
[12] Glaser SD, Baise LG. System identification estimation of soil properties at the Lotung. Soil Dyn Earthq Eng 2000;19(7):521–31.
[13] Kokusho T, Aoyagi T, Wakunami A. In situ soil-specific nonlinear properties back-calculated from vertical array records during 1995 Kobe earthquake. J Geotech Geoenviron Eng 2005;131(12):1509–21.
[14] Tsai C-C, Hashash Y. A novel framework integrating downhole array data and site response analysis to extract dynamic soil behavior. Soil Dyn Earthq Eng 2008;28(3):181–97.
[15] Tsai C-C, Hashash Y. Learning of dynamic soil behavior from downhole arrays. J Geotech Geoenviron Eng 2009;135(6):745–57.
[16] Mercado V, El-Sekelly W, Zeghal M, Abdoun T. Identification of soil dynamic properties through an optimization analysis. Comput Geotech 2015;65:75–186.
[17] Mercado V, El-Sekelly W, Zeghal M, Abdoun T. Identification of soil dynamic properties of sites subjected to bi-directional excitation. Soil Dyn Earthq Eng 2017;92:215–28.
[18] Ang AH-S, Tang WH. Probability concepts in engineering: emphasis on applications to civil and environmental engineering. 2nd ed. Hoboken (New Jersey): John Wiley & Sons Inc; 2007.
[19] Ching J, Wang J-S. Application of the transitional Markov chain Monte Carlo algorithm to probabilistic site characterization. Eng Geol 2016;203:151–67.
[20] Li J, Cassidy MJ, Huang J, Zhang L, Kelly R. Probabilistic identification of soil stratification. Géotechnique 2016;66(1):16–26.
[21] Idriss IM, Sun JI. SHAKE91: user's manual, center for geotechnical modeling. Davis (California): Department of Civil and Environmental Engineering, University of California; 1992.
[22] Hardin BO, Drnevich VP. Shear modulus and damping in soils. J Soil Mech Found Div ASCE 1972;98:667–92.
[23] Vucetic M, Dobry R. Effect of soil plasticity on cyclic response. J Geotech Eng 1991;117(1):89–107.
[24] Hashash Y, Park D. Non-linear one-dimensional seismic ground motion propagation in the Mississippi embayment. Eng Geol 2001;62(1):185–206.
[25] Pyke RM. Modeling of dynamic soil properties. Guidelines for determining design bases ground motions, Electric Power Research Institute; 1993. Appendix 7A, p. 7.A-1–7.A-90.
[26] Stokoe KH, Darendeli MB, Andrus RD, Brown LT. Dynamic soil properties: laboratory, field and correlation studies. In: 2nd international conference on earthquake geotechnical engineering, vol. 3, Lisbon, Portugal; 1999. p. 811–45.
[27] Zhang J, Andrus RD, Juang CH. Normalized shear modulus and material damping ratio relationships. J Geotech Geoenviron Eng 2005;131(4):453–64.
[28] Vardanega P, Bolton MD. Stiffness of clays and silts: normalizing shear modulus and shear strain. J Geotech Geoenviron Eng 2013;139(9):1575–89.
[29] Oztoprak S, Bolton MD. Stiffness of sands through a laboratory test database. Géotechnique 2013;63(1):54–70.
[30] Phillips C, Hashash Y. Damping formulation for nonlinear 1D site response analyses. Soil Dyn Earthq Eng 2009;29(7):1143–58.
[31] Schnabel PB, Lysmer JL, Seed HB. SHAKE: a computer program for earthquake response analysis of horizontally layered sites. Report no. EERC 72–12. Berkeley: University of California; 1972.
[32] Biancon L, Nancey A, Caquel F, Villard P. New technology for strain measurements in soil and the survey of reinforced earth constructions. In: 3rd European geosynthetics conf; 2004.
[33] Paikowsky S, Palmer C, Rolwes L. The use of tactile sensor technology for measuring soil stress distribution. Atlanta (GA): In GeoCongress; 2006.
[34] Julier SJ, Uhlmann JK. A new extension of the Kalman filter to nonlinear systems. In: 11th international symposium on aerospace/defense sensing, simulation and controls, Orlando, FL; 1997.
[35] Wan EA, van der Merwe R. The unscented Kalman filter for nonlinear estimation. In: IEEE 2000 adaptive systems for signal processing, communications, and control symposium, Lake Louise, AB, Canada; 2000.
[36] Chowdhar G, Jategaonkar R. Aerodynamic parameter estimation from flight data applying extended and unscented Kalman filter. Aerosp Sci Technol 2010;14:106–17.
[37] Astroza R, Ebrahimian H, Conte JP. Material parameter identification in distributed plasticity FE models of frame-type structures using nonlinear stochastic filtering. J Eng Mech ASCE 2015;141(5):04014149.
[38] Mansouri M, Dumont B, Destain M-F. Modeling and prediction of nonlinear environmental system using Bayesian methods. Comput Electron Agric 2013;92:16–31.
[39] Simon D. Optimal state estimation: Kalman, H∞, and nonlinear approaches. Hoboken (New Jersey): John Wiley & Sons Inc.; 2006.
[40] Porter KA, Beck JL, Shaikhutdinov RV. Sensitivity of building loss estimates to major uncertain variables. Earthq Spectra 2002;18(4):719–43.
[41] Kolås S, Fossa BA, Scheic TS. Constrained nonlinear state estimation based on the UKF approach. Comput Chem Eng 2009;33:1386–401.
[42] Wu B, Wang T. Model updating with constrained unscented Kalman filter for hybrid testing. Smart Struct Syst 2014;14(6):1105–12.