Bayesian optimization for inverse identification of cyclic constitutive law of structural steels from cyclic structural tests

Bach Do[a,*], Makoto Ohsaki[a]

[a]*Department of Architecture and Architectural Engineering, Graduate School of Engineering, Kyoto University, Kyoto-Daigaku Katsura, Nishikyo, Kyoto 615-8540, Japan*

ARTICLE INFO

Keywords:
Elastoplastic consititutive law
Parameter identification
Structural steels
Bayesian optimization
Noise-free and noisy observations
Cyclic loading

Abstract

Properly modeling cyclic elastoplastic behavior of structural steels is essential for establishing accurate analyses of structures subjected to earthquake excitation. However, identifying the underlying parameters to simulate such behavior is commonly hindered by the computational burden of carrying out many nonlinear analyses. This work proposes using Bayesian optimization (BO) for solving an inverse problem by which certain parameters for the nonlinear combined isotropic/kinematic hardening model are inferred from cyclic responses of a specimen or a structural component. BO minimizes an error function that represents the difference between the simulated responses and those measured experimentally while providing a global optimization framework for parameter identification, reducing the number of simulations, and addressing observational noise. It is found that BO has higher robustness as compared with some population-based optimization algorithms when expending the same number of simulations. Identification results for a specimen and a cantilever show a good ability of identified parameters to capture the behavior of structural steels under different cyclic loadings. They also suggest a possibility of identifying the parameters for multiple materials from cyclic tests of a structural component that is remarkable because cyclic material tests are difficult and usually not carried out before structural tests. Experimental measures from various loading histories should be simultaneously used for identification as they can mitigate the bias toward a specific loading history, which may lead the parameters to inaccurate prediction of material behavior under other loading histories.

---

1. Introduction

Cyclic elastoplastic behavior of steel structures under earthquake excitation has been extensively studied using detailed finite element (FE) models [1, 2]. These models require construction of an accurate cyclic constitutive law for structural steels to establish reliable analyses. Solid mechanics has provided linear isotropic and kinematic hardening models to quantify elastoplastic deformations of structural steels as well as various nonlinear models to simulate the cyclic behavior of metallic materials [3, 4]. Each of these hardening models is characterized by certain parameters that should be identified based on available experimental measures before being used as input to future analyses. The problem formulated for finding these material parameters is known as an inverse problem. This problem is computationally too expensive to solve as it requires carrying out many nonlinear analyses. Therefore, it is desirable to find a new parameter identification method that can reduce the number of simulation calls considerably.

Methods for inverse identification of material parameters can be categorized into deterministic [5–9] and probabilistic approaches [10–12]. By penalizing the errors in numerical predictions, the deterministic approach minimizes an error function that represents the difference between the simulated and actual structural responses, and produces a deterministic estimate for each parameter. This approach is often applied to identifying parameters from uniaxial cyclic tests

with few loading cycles because the calculation of numerical responses for these tests does not arise a major computational issue that, in turn, facilitates the use of conventional optimization algorithms for minimizing the associated error functions. Meanwhile, the probabilistic approach adopts Bayesian inference to handle uncertainty in the material parameters as well as noise involved in the experimental measures [12, 13]. It specifies a prior probability density function (PDF) of parameters and constructs a likelihood PDF that describes the information on observations for a given set of parameter values. Uncertainty in the parameters and observational noise are flexibly incorporated into the prior and likelihood PDFs, respectively. Bayes' rule then uses the prior and likelihood PDFs for finding a posterior PDF and requires support from a sampling technique, e.g., the Metropolis algorithm [14] or Monte-Carlo simulation [15], for generating posterior samples of each parameter. With these samples, statistical estimates of each parameter can be found for use of future analyses. A shortcoming of the probabilistic approach is that it may lead parameter identification to a wrong direction because it does not provide a rigorous method for specifying a proper prior PDF. Moreover, structural engineers are still hesitant to adopt this approach as it requires sophisticated probabilistic computations. By taking the advantages of the two aforementioned approaches, this work focuses on the deterministic approach, while considering observational noise in parameter identification, which is handled by the probabilistic approach.

Research on identification of cyclic elastoplastic parameters for structural steels using the deterministic approach has evolved in two directions. The first direction proposes

*Corresponding author

✉ se.do@archi.kyoto-u.ac.jp (B. Do); ohsaki@archi.kyoto-u.ac.jp (M. Ohsaki)

new hardening models by modifying the current models to describe the yield plateau [6, 16, 17] or a decrease of the yield stress for structural steels [9]. Material parameters underpinning the new hardening models are identified by solving the associated inverse problems using conventional optimization algorithms, e.g., population-based [5, 7, 18], gradient-based [9], or hybrid algorithms [19, 20]. Although its mathematical foundation is solid, this direction does not account for the computational complexity arising from calculation of the error function as well as its derivatives. Thus, it may be inefficient to simply apply the methods in this direction to complicated problems, for example, where multiaxial cyclic behavior of the material or a large number of loading cycles is of interest. The second direction incorporates advanced machine learning in solving the inverse problem [21, 22]. It conducts validated computer experiments [22] that generate synthetic training datasets for different sets of parameters. The generated datasets are used for training a neural network, where the synthetic responses are considered as input. The trained neural network inputs the available experimental measures and produces the parameters as its output. By doing so, this direction faces the following two issues: (1) obtaining a useful neural network needs a large number of training points, or equivalently, a large number of costly simulations, and (2) validating a computer experiment by tuning its parameters is intrinsically equivalent to solving an inverse problem.

In addition to the computational complexity, observational noise [23] and dataset-specific bias [24] have not been fully addressed by the current methods of the deterministic approach. The former issue leads to uncertain calculation of the error function for a particular set of parameters. This is inevitable because noise contaminating the experimental measures cannot be ignored. The latter issue may bias the parameters toward a set of experimental results from a specific loading history used for parameter identification. Consequently, the parameters may inaccurately predict the material behavior under other loading histories as the cyclic hardening behavior depends on loading conditions [25].

Bayesian optimization (BO) [26–28] is a sequential design strategy for solving optimization problems that have an expensive-to-evaluate objective function. BO evaluates the objective function values for a small number of design points and uses them as input to construction of a Gaussian process (GP) model [29] that probabilistically describes the dependence of the objective function on the design variables or parameters. The GP model without considering noise gives exact values or Gaussian distributions of the objective function at the observed or unobserved design points, respectively. Nevertheless, it can incorporate noise involved in the objective function into every prediction [29]. BO uses the GP model to formulate an acquisition function that assists the algorithm in selection of a new, good design point in the next iteration without calling any simulation, thereby considerably reducing the number of simulations needed to find a good solution. Specifically, the new design point from maximizing the acquisition function is likely to be ideal in

the design space as it balances exploitation, i.e., improving the best-observed objective value in its neighborhood, and exploration, i.e., searching in regions where uncertainty in the GP prediction is large. This represents a key advantage of BO as compared with other surrogate-based optimization algorithms using deterministic surrogate models, e.g., radial basis functions, polynomial basis functions, or support vector regression, which are commonly limited to exploitation. Recently, BO has been successfully applied to designing materials [30], solving structural optimization problems [31, 32], and identifying material parameters for multiscale crystal plasticity models [33].

The primary objective of this study is to present the applications of BO to inverse identification of cyclic elastoplastic parameters for structural steels while addressing the computational issue arising from using conventional optimization algorithms of the deterministic approach. The secondary objective is to demonstrate the possibility of identifying the elastoplastic parameters for multiple materials from cyclic structural tests that is remarkable because cyclic material tests are difficult and usually not carried out before structural tests. The parameters are within the scope of the nonlinear combined isotropic/kinematic hardening model and the modeling error is not considered within the scope of this study. An inverse problem is formulated for either noise-free or noisy experimental measures. BO sequentially solves the inverse problem by maximizing the acquisition function formulated in each iteration with and without observational noise. To investigate the effect of dataset-specific bias on the identification results, different experimental datasets are used for parameter identification. The parameters obtained from BO as well as their prediction ability are compared with those from some well-known population-based optimization algorithms.

The remainder of this paper progresses as follows. Section 2 describes the nonlinear combined isotropic/kinematic hardening model. Section 3 details the BO framework for parameter identification with and without considering observational noise. Section 4 provides two identification examples of a steel specimen and a bi-material cantilever. Section 5 summarizes and concludes this paper. Mathematical foundation of the GP model is provided in Appendix.

2. Nonlinear combined isotropic/kinematic hardening material model

For completeness of the paper, this section briefly describes the nonlinear combined isotropic/kinematic hardening model used throughout this study for simulating the cyclic elastoplastic behavior of structural steels.

Let σ and ϵ denote the stress and strain tensors at a point of a steel body. With the infinitesimal strain theory, ϵ and its rate can be decomposed into elastic part ϵ^{e} and plastic part ϵ^{p} as

$$\epsilon = \epsilon^{\mathrm{e}} + \epsilon^{\mathrm{p}}, \quad \dot{\epsilon} = \dot{\epsilon}^{\mathrm{e}} + \dot{\epsilon}^{\mathrm{p}} \tag{1}$$

where the overdot denotes the time derivative of the quantity.

The elastic stress-strain relation and its rate form are formulated using the plastic strain tensor as

$$\sigma = \mathbf{D} : (\epsilon - \epsilon^{\mathrm{p}}), \quad \dot{\sigma} = \mathbf{D} : (\dot{\epsilon} - \dot{\epsilon}^{\mathrm{p}}) \tag{2}$$

where $\mathbf{D}$ is the isotropic tensor of elastic moduli. The stress rate $\dot{\sigma}$ corresponding to a strain rate $\dot{\epsilon}$ is determined once $\dot{\epsilon}^{\mathrm{p}}$ has been specified. This can be done by using the following von Mises yield criterion, in combination with an associative flow rule:

$$F = \|\xi\| - \sqrt{\frac{2}{3}}\sigma_{\mathrm{y}} \leq 0 \tag{3}$$

where $\xi = \mathrm{dev}[\sigma] - \mathrm{dev}[\alpha]$ the shifted-stress tensor; $\mathrm{dev}[\cdot]$ the deviatoric part of $[\cdot]$; $\|\cdot\|$ the 2-norm of the tensor; α the back-stress tensor; and σ_{y} the yield stress. The associative flow rule governing the evolution of ϵ^{p} gives

$$\dot{\epsilon}^{\mathrm{p}} = \lambda \frac{\xi}{\|\xi\|} = \lambda\mathbf{n} \tag{4}$$

where λ is the non-negative plastic consistency parameter and $\mathbf{n} = \xi/\|\xi\|$ is a unit normal vector of the yield surface in the deviatoric stress space. λ can be found based on the consistency condition $F = \dot{F} = 0$ [4].

Isotropic and kinematic hardening models are commonly used for describing the strain hardening process of structural steels. In the isotropic hardening model, the yield surface increases its size during the evolution of plastic deformation without changing its shape and the location of its center. The back-stress tensor α, therefore, does not appear in Eq. (3), leading the yield surface to be an isotropic function of the stress that cannot capture the Bauschinger effect [34].

As structural steels exhibit a saturation point of stress at large deformation, the increment of the yield surface can be described by the following monotonically increasing nonlinear function of isotropic hardening model [35]:

$$\sigma_{\mathrm{y}} = \sigma_{\mathrm{y},0} + Q_{\infty}[1 - \exp(-b\epsilon_{\mathrm{eq}}^{\mathrm{p}})] \tag{5}$$

where $\sigma_{\mathrm{y},0}$ denotes the initial yield stress; Q_{∞} the difference of the stress saturation and $\sigma_{\mathrm{y},0}$; b the isotropic saturation rate; and $\epsilon_{\mathrm{eq}}^{\mathrm{p}}$ the current equivalent plastic strain.

The equivalent plastic strain $\epsilon_{\mathrm{eq}}^{\mathrm{p}}$ in Eq. (5) is determined using its previous state and the current rate $\dot{\epsilon}_{\mathrm{eq}}^{\mathrm{p}}$ that reads

$$\dot{\epsilon}_{\mathrm{eq}}^{\mathrm{p}} = \sqrt{\frac{2}{3}\dot{\epsilon}^{\mathrm{p}} : \dot{\epsilon}^{\mathrm{p}}} = \sqrt{\frac{2}{3}}\lambda \tag{6}$$

In contrast to the isotropic hardening model, the kinematic hardening model does not change the size and shape of the yield surface during the evolution of plastic strain. Instead, the center of the yield surface changes its location by performing a rigid translation in the evolution direction of the plastic strain. Thus, this hardening model can capture the Bauschinger effect. The simplest linear kinematic hardening [36] models the evolution of α by

$$\dot{\alpha} = \sqrt{\frac{2}{3}}C\dot{\epsilon}_{\mathrm{eq}}^{\mathrm{p}}\mathbf{n} \tag{7}$$

where C is the hardening ratio.

Armstrong and Frederick [37] proposed a nonlinear kinematic hardening rule for modeling $\dot{\alpha}$ by introducing a recall term, such that

$$\dot{\alpha} = \sqrt{\frac{2}{3}}C\dot{\epsilon}_{\mathrm{eq}}^{\mathrm{p}}\mathbf{n} - \gamma\dot{\epsilon}_{\mathrm{eq}}^{\mathrm{p}}\alpha \tag{8}$$

where γ is the rate of α.

For a better approximation, α can be defined as a superposition of n_{k} components as [38]

$$\alpha = \sum_{k=1}^{n_{\mathrm{k}}} \alpha_k \tag{9}$$

The evolution of α_k can be modeled using the nonlinear kinematic hardening rule in Eq. (8).

The nonlinear combined isotropic/kinematic model was developed for simultaneous use of the isotropic and nonlinear kinematic hardening models. Let $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbb{R}^{+n}$ denote the vector of n positive material parameters for this hardening model. Thus, $\mathbf{x} = [E, Q_{\infty}, b, \sigma_{\mathrm{y},0}, C_1, \gamma_1]$ if we use one back-stress component and a fixed Poisson's ratio, where E denotes Young's modulus of the material incorporated in the isotropic elastic tensor $\mathbf{D}$.
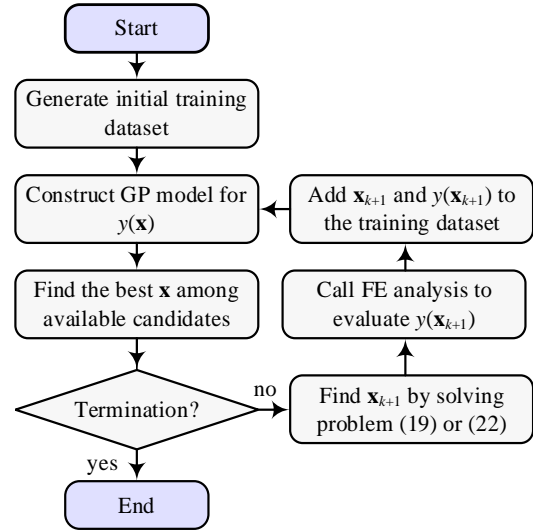

Figure 1: BO for parameter identification.

3. Bayesian optimization for inverse parameter identification

3.1. Inverse problem for parameter identification

The inverse problem in this study is to identify the vector of material parameters $\mathbf{x}$ based on experimental measures from cyclic loading tests of a structural steel or a structural component. We minimize an error function $f(\mathbf{x}): \mathbb{R}^{+n} \rightarrow \mathbb{R}$ that represents the difference between the structural responses simulated from a numerical model of the experiment, characterized by $\mathbf{x}$, and the corresponding responses measured
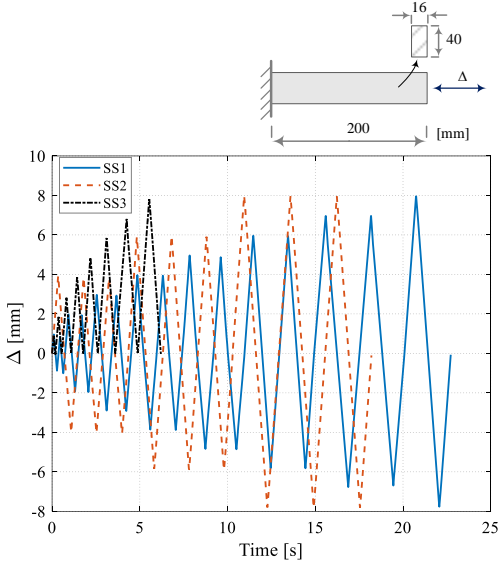
Figure 2: Steel specimen and three loading histories for cyclic tests [39].

experimentally. The optimal $\mathbf{x}$ is identified by solving the following minimization problem:

$$\begin{array}{cl} \underset{\mathbf{x}}{\text{minimize}} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in [\mathbf{x}_l, \mathbf{x}_u] \end{array} \tag{10}$$

where $\mathbf{x}_l$ and $\mathbf{x}_u$ are the pre-specified lower and upper bounds of $\mathbf{x}$, respectively.

Let d_i^s and d_i^m denote the simulated and measured values of the response of interest at the ith time step of a cyclic loading history of N_1 steps. Following Ohsaki et al. [6], the error function for this loading history reads

$$f(\mathbf{x}) = \sqrt{\frac{1}{N_1} \sum_{i=1}^{N_1} \left(d_i^s - d_i^m \right)^2} \tag{11}$$

If the material or structural component is tested under a total of T different cyclic loading histories, $f(\mathbf{x})$ is formulated to incorporate the experimental results with N_t ($t = 1, \dots, T$) steps from these loading histories, such that

$$f(\mathbf{x}) = \sum_{t=1}^{T} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} \left(d_{i,t}^s - d_{i,t}^m \right)^2} \tag{12}$$

where $d_{i,t}^s$ and $d_{i,t}^m$ represent d_i^s and d_i^m associated with the tth loading history, respectively.

As previously mentioned, the parameters identified from a single loading history may inaccurately predict the material behavior under other loading histories. The error function in Eq. (12) is expected to reduce the bias toward a set of parameters that offers the best fit to a specific experimental dataset.

It is also important to take into account the effect of uncertainty in the experimental measures on the identified parameters because a reliable $\mathbf{x}$ solely relies on the quality of

experimental datasets. In addition to noise-free measures, we consider another case in which the experimental datasets are contaminated by noise. In this case, the real value of d_i^m for a single loading history is fluctuated around a noisy measure h_i^m as

$$h_i^m = d_i^m + \omega_i, \quad \omega_i \sim \mathcal{N}\left(0, \tau_d^2\right) \tag{13}$$

where ω_i is additive Gaussian noise with zero mean and standard deviation τ_d that is specified or calibrated for each experiment by the user. Consequently, the actual value $f(\mathbf{x})$ of the error function is contaminated by noise ω_f as

$$y(\mathbf{x}) = f(\mathbf{x}) + \omega_f, \quad \omega_f \sim \mathcal{N}\left(0, \tau_f^2\right) \tag{14}$$

where $y(\mathbf{x})$ is the noisy value of $f(\mathbf{x})$. ω_f is assumed to be a Gaussian with zero mean and standard deviation τ_f, which is empirically found by generating random samples of d_i^m for the specified value of τ_d.

It is difficult to solve problem (10) effectively since the calculation of $f(\mathbf{x})$ requires an expensive simulation. This makes the common use of population-based optimization algorithms [5, 7, 18, 40–42] practically infeasible because a large number of simulations is needed for obtaining a good solution. The use of gradient-based algorithms [19, 43] is also hindered by the requirement of calculating the gradient of the error function. Moreover, the problem considering noisy experimental measures has not been fully explored by the conventional optimization algorithms. Thus, BO is used for solving problem (10) with replacing $f(\mathbf{x})$ by $y(\mathbf{x})$ as it can provide a global-optimization framework and address observational noise while keeping the number of simulations as low as possible.

3.2. Bayesian optimization for inverse problem with noise-free experimental measures

BO constructs a GP model to approximate the error function from a training dataset $\mathcal{D} = \{\mathbf{X}, \mathbf{y}\} = \{\mathbf{x}_i, y_i\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbb{R}^{+n}$ represents a set of the parameters and $y_i \in \mathbb{R}$ the observed value of the error function at $\mathbf{x}_i$. The samples $\mathbf{x}_i$ are randomly generated using Latin-hypercube sampling. The number of initial sample points N is problem-dependent and only serves as a guide. As a recommendation, $N = 15n$ for a fixed training dataset [44]. By adopting the adaptive sampling strategy of BO, we use $N = 10n$ in this work. For each sample $\mathbf{x}_i$ and a specific loading history, the nonlinear combined isotropic/kinematic hardening incorporated in the FE model of the experiment evaluates d_i^s in Eq. (11) or $d_{i,t}^s$ in Eq. (12), and y_i can be found accordingly. The GP model describing the relationship between y and $\mathbf{x}$, denoted as $\hat{y}(\mathbf{x})$, is developed by conditioning a Gaussian probability distribution over possible regression functions on $\mathcal{D}$, thereby producing a probabilistic regression model that is a conditional Gaussian, for which the detailed derivations are provided in Appendix. The GP model at a particular $\mathbf{x}$ reads

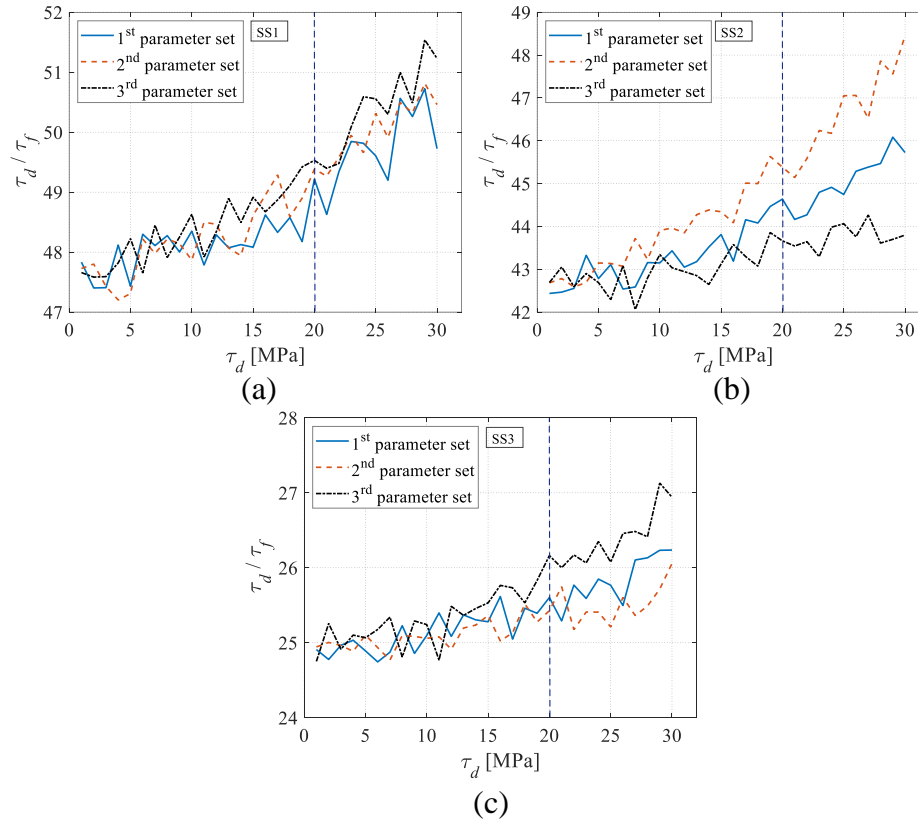$$\hat{y}(\mathbf{x}) \sim \mathcal{N}\left(\mu_y(\mathbf{x}), \tau_y^2(\mathbf{x})\right) \tag{15}$$

Figure 3: Empirical relations between τ_f and τ_d corresponding to three loading histories for cyclic tests of the specimen. (a) SS1; (b) SS2; (c) SS3.
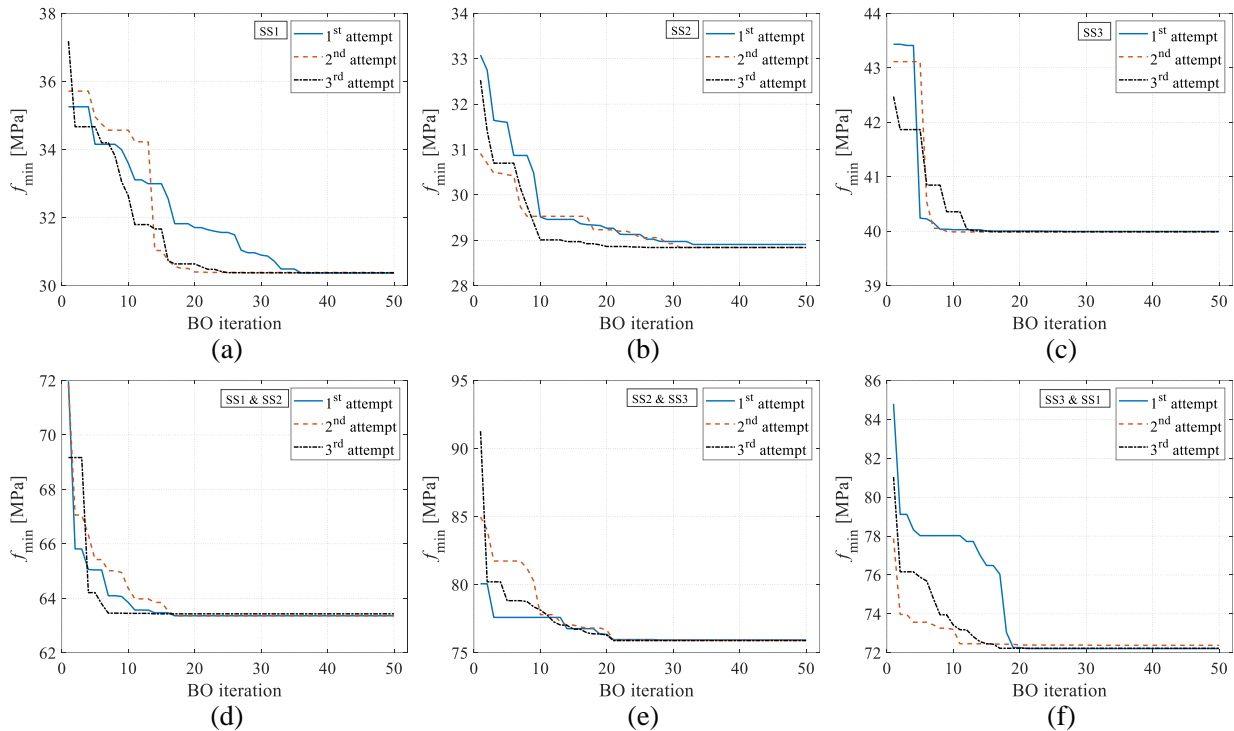

Figure 4: Histories of BO for different noise-free experimental datasets from cyclic tests of the specimen. (a)–(f) obtained from groups 1–6, respectively.
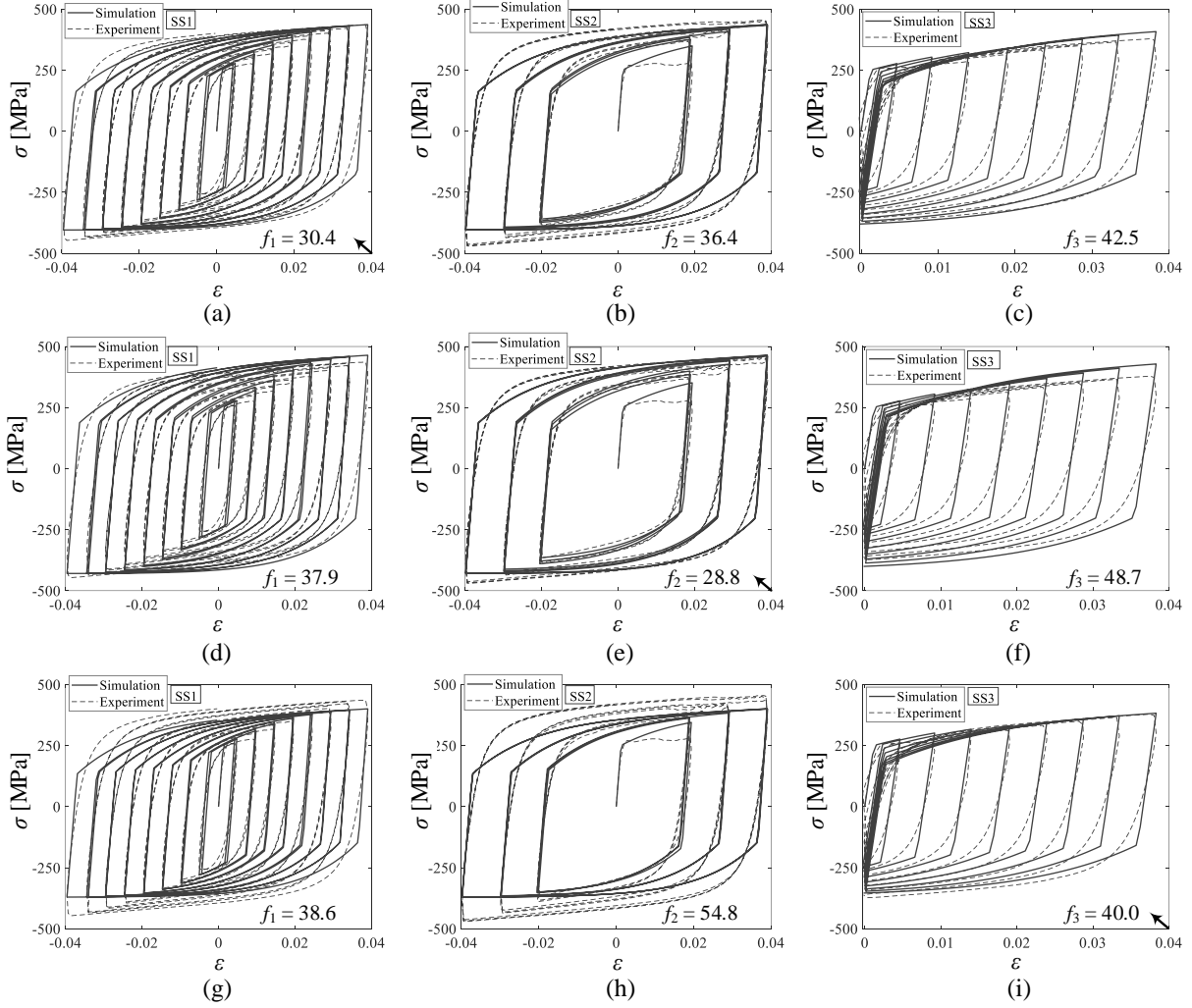
Figure 5: Comparison of test data and model predictions for cyclic tests of the specimen with parameters identified from a specific noise-free experimental dataset. (a), (b), (c) Parameters from group 1; (d), (e), (f) Parameters from group 2; (g), (h), (i) Parameters from group 3.

where $\mu_y(\mathbf{x})$ and $\tau_y(\mathbf{x})$ denote the mean and standard deviation of the Gaussian $\hat{y}(\mathbf{x})$, respectively.

With Eq. (15), the remainder of this section goes over the remaining steps of using the conventional BO for solving problem (10) without observational noise, i.e., $\tau_d = 0$ and $y(\mathbf{x}) = f(\mathbf{x})$. A modification of the conventional BO to solve the problem with observational noise, i.e., $\tau_d > 0$ and $y(\mathbf{x}) = f(\mathbf{x}) + \omega_f$, is the subject of the next section.

Suppose BO has completed its kth iteration at which the training dataset $\mathcal{D}_k = \{\mathbf{X}_k, \mathbf{y}_k\}$ has a total of K samples and the corresponding GP model $\hat{y}_k(\mathbf{x})$ has been constructed from $\mathcal{D}_k$. The next step is to specify a new sampling point $\mathbf{x}_{k+1}$ as input to a new simulation that evaluates $\mathbf{y}(\mathbf{x}_{k+1})$ for updating $\mathcal{D}_k$ and $\hat{y}_k(\mathbf{x})$. Since we wish to reduce the number of simulations as much as possible, $\mathbf{x}_{k+1}$ is likely to be ideal in the parameter space and should be specified based on the information we have so far, i.e., $\mathcal{D}_k$ and $\hat{y}_k(\mathbf{x})$. In the conventional BO, how promising each point in the parameter space is can be transformed to a measure of our belief about an improvement in the best-observed solution to problem (10) by

means of the acquisition function. Thus, the new sampling point $\mathbf{x}_{k+1}$ is the maximizer of the acquisition function. A popular acquisition function proposed by Jones et al. [26] reads

$$I(\mathbf{x}) = \left[f_{\min} - f(\mathbf{x}) \right]^+ \tag{16}$$

where $[\cdot]^+ = \max(0, \cdot)$ and $f_{\min} = \min\left(\mathbf{f}(\mathbf{X}_k)\right)$ with $\mathbf{f}(\mathbf{X}_k) = \left[f(\mathbf{x}_1), \dots, f(\mathbf{x}_K) \right]$. Maximizing $I(\mathbf{x})$ gives a value of $\mathbf{x}_{k+1}$ at which the error function is equal to or less than the best-observed error function value $f_{\min}$. By further conditioning $I(\mathbf{x})$ on $\mathcal{D}_k$, the expectation of $I(\mathbf{x})$ reads

$$EI(\mathbf{x}) = \mathbb{E}\left[I(\mathbf{x}) | \mathbf{f}(\mathbf{X}_k) \right] \tag{17}$$

Without observational noise, $f_{\min}$ and $\mathbf{f}(\mathbf{X}_k)$ are identical to $y_{\min}$ and $\mathbf{y}(\mathbf{X}_k)$, respectively, which enables the derivation of an analytical form for $EI(\mathbf{x})$ using integration by parts [26].
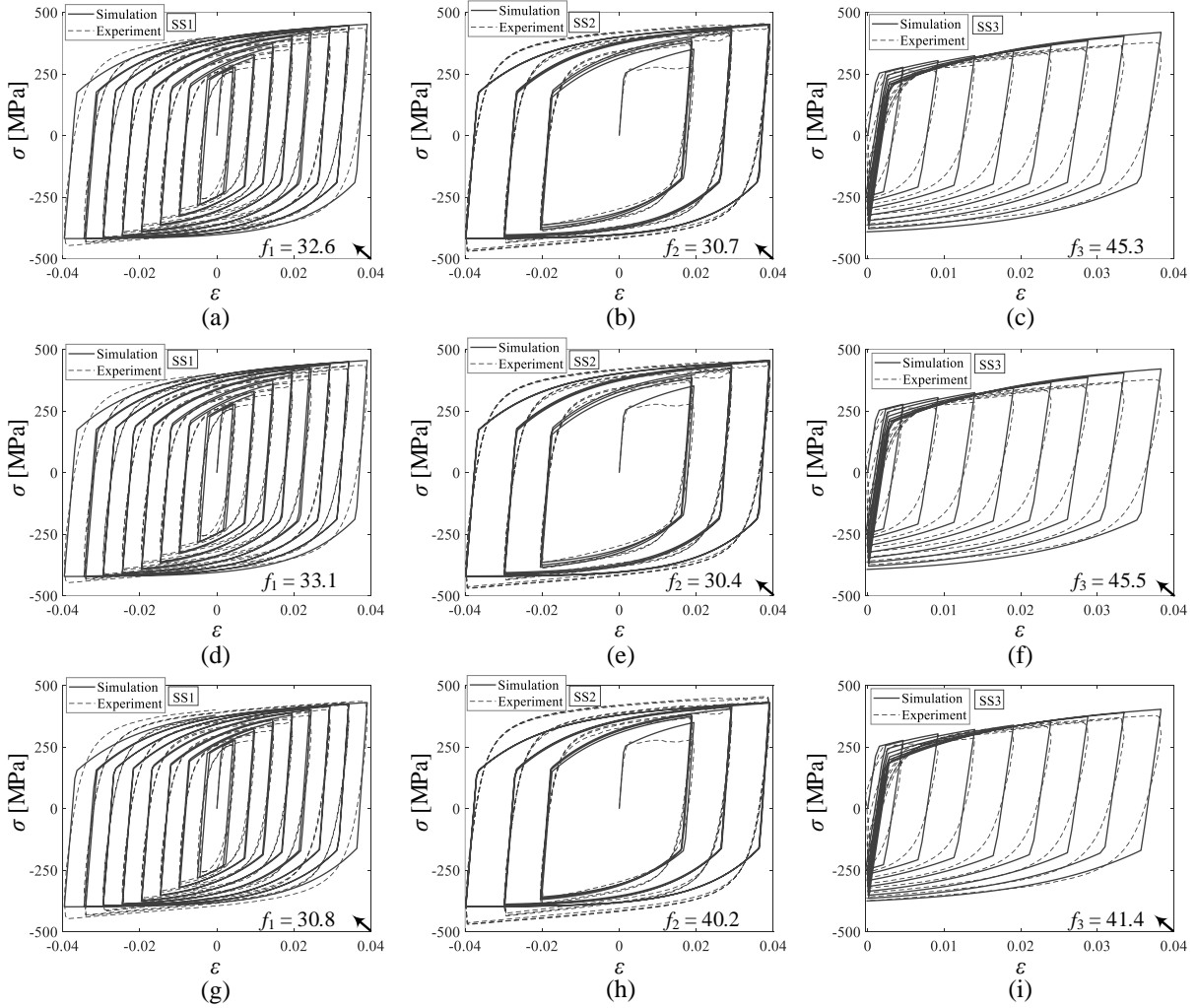
Figure 6: Comparison of test data and model predictions for cyclic tests of the specimen with parameters identified from a pair of two noise-free experimental datasets. (a), (b), (c) Parameters from group 4; (d), (e), (f) Parameters from group 5; (g), (h), (i) Parameters from group 6.

The resulting expression reads

$$EI(\mathbf{x}) = \left(f_{\min} - \mu_f(\mathbf{x}) \right) \Phi \left(\frac{f_{\min} - \mu_f(\mathbf{x})}{\tau_f(\mathbf{x})} \right) \\ + \tau_f(\mathbf{x}) \phi \left(\frac{f_{\min} - \mu_f(\mathbf{x})}{\tau_f(\mathbf{x})} \right) \quad (18)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution and probability density functions, respectively; $\mu_f(\mathbf{x}) = \mu_y(\mathbf{x})$; and $\tau_f(\mathbf{x}) = \tau_y(\mathbf{x})$.

The first and second terms of $EI(\mathbf{x})$ are responsible for exploitation and exploration, respectively. In other words, $EI(\mathbf{x})$ simultaneously incorporates the best error function value $f_{\min}$ observed so far (exploitation) and large amount of prediction uncertainty $\tau_f(\mathbf{x})$ (exploration) into finding $\mathbf{x}_{k+1}$. Moreover, as values of $EI(\mathbf{x})$ associated with the members of $\mathcal{D}_k$ are always non-positive, maximizing $EI(\mathbf{x})$ does not reselect the members of $\mathcal{D}_k$.

Without observational noise, $\mathbf{x}_{k+1}$ is found by solving

$$\begin{aligned} \text{Find} \quad & \mathbf{x}_{k+1} = \underset{\mathbf{x}}{\text{argmax}} \quad EI(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in [\mathbf{x}_l, \mathbf{x}_u] \end{aligned} \quad (19)$$

3.3. Bayesian optimization for inverse problem with noisy experimental measures

Maximizing $EI(\mathbf{x})$ is more difficult in the presence of observational noise as $f_{\min}$ in Eq. (16) is not known if $\tau_d > 0$ and the closed form in Eq. (18) is no longer available. This issue has been addressed by various BO variants [45]. The idea is to replace $f_{\min}$ by an evaluable and efficient representative so that BO is still able to use Eq. (18) for guiding the optimization process, termed "plug-in" method.

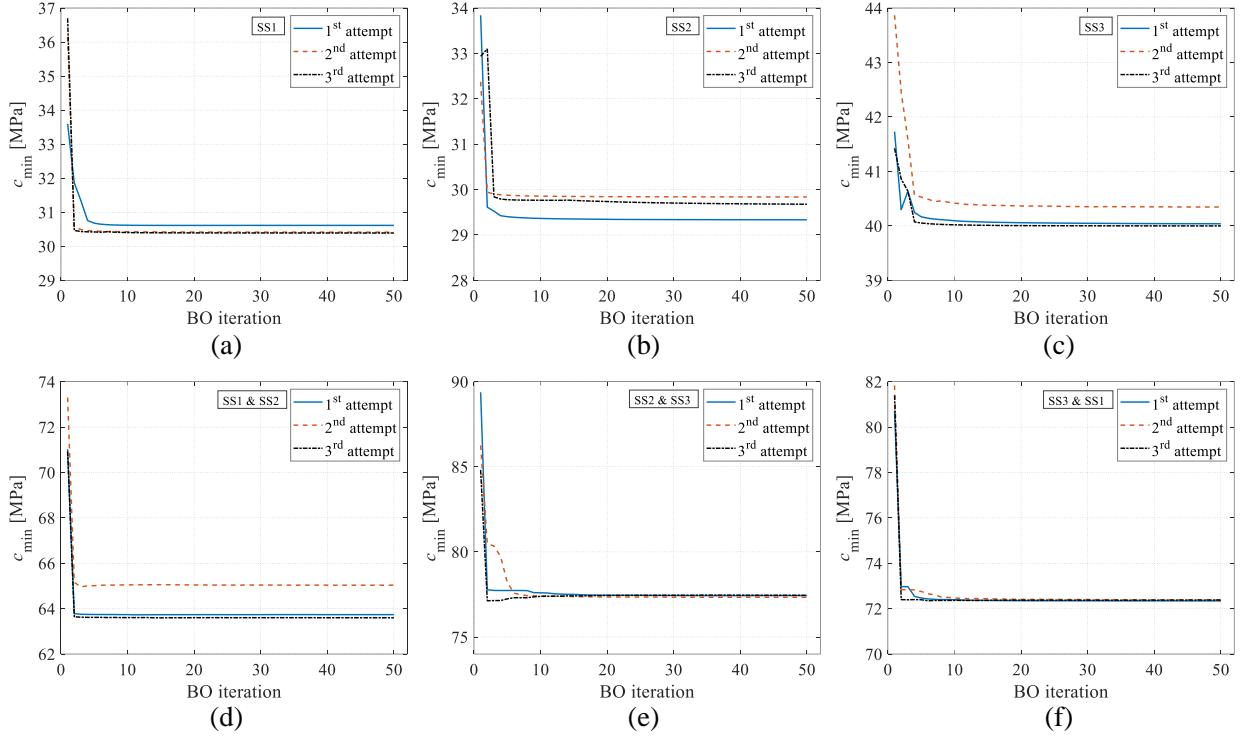Let $c_{\min}$ denote the representative of $f_{\min}$, the expected

Figure 7: Histories of BO for different noisy experimental datasets from cyclic tests of the specimen. (a)–(f) obtained from groups 1–6, respectively.

improvement in the presence of noise reads

$$
\begin{aligned}
EI_n(\mathbf{x}) = & \left(c_{\min} - \mu_y(\mathbf{x})\right) \Phi\left(\frac{c_{\min} - \mu_y(\mathbf{x})}{\tau_y(\mathbf{x})}\right) \\
& + \tau_y(\mathbf{x})\phi\left(\frac{c_{\min} - \mu_y(\mathbf{x})}{\tau_y(\mathbf{x})}\right)
\end{aligned}
\tag{20}
$$

Huang et al. [46] used the GP mean in Eq. (A.8) at a so-called effective best solution $\mathbf{x}^*$ for $c_{\min}$, i.e., $c_{\min} = \mu_y(\mathbf{x}^*)$. $\mathbf{x}^*$ is selected from $\mathbf{X}_k$ so that it minimizes $\mu_y(\mathbf{x}) + \alpha\tau_y(\mathbf{x})$, where $\alpha = 1$ is recommended by Huang et al. [46].

To further enhance exploration, an augmented EI [46], denoted as AEI, is formulated by adding a heuristic multiplier to $EI_n(\mathbf{x})$, such that

$$
AEI(\mathbf{x}) = EI_n(\mathbf{x})\left(1 - \frac{\tau_f}{\sqrt{\tau_y^2(\mathbf{x}) + \tau_f^2}}\right)
\tag{21}
$$

$AEI(\mathbf{x})$ strengthens exploration by penalizing points with small standard deviation $\tau_y(\mathbf{x})$. It also returns to $EI_n(\mathbf{x})$ if $\tau_f = 0$. As the use of $AEI(\mathbf{x})$ has been justified by empirical performance [45, 46], we exclusively use it as the BO guide to solving problem (10) when considering observational noise. Thus, the next sampling point $\mathbf{x}_{k+1}$ is found by solving

$$
\begin{aligned}
\text{Find} \quad & \mathbf{x}_{k+1} = \underset{\mathbf{x}}{\arg\max} \quad AEI(\mathbf{x}) \\
\text{subject to} \quad & \mathbf{x} \in [\mathbf{x}_l, \mathbf{x}_u]
\end{aligned}
\tag{22}
$$

3.4. Identification procedure

Fig. 1 summarizes the identification procedure using BO. The following five steps are sequentially implemented:

- Step 1: Specify τ_d, determine τ_f based on τ_d, and randomly generate initial samples of $\mathbf{x}$ using Latin-hypercube sampling. Then, create the training dataset $\mathcal{D}$ by performing FE analyses for the generated samples.

- Step 2: Construct GP model for $y(\mathbf{x})$ from $\mathcal{D}$; see Appendix.

- Step 3: Find $f_{\min}$ or $c_{\min}$ and the associated parameter vector among the members of $\mathcal{D}$. Terminate the identification process and output the parameters if the number of BO iterations reaches a pre-specified upper limit. Otherwise, proceed to Step 4.

- Step 4: Find $\mathbf{x}_{k+1}$ by solving problem (19) or (22) corresponding to $\tau_d = 0$ or $\tau_d > 0$, respectively.

- Step 5: Evaluate $y(\mathbf{x}_{k+1})$ using the FE analysis, update $\mathcal{D}$ and the current GP model for $y(\mathbf{x})$, and reiterate from Step 3.

4. Illustrative examples

4.1. Parameters for a steel specimen

This section identifies the elastoplastic parameters to model the uniaxial cyclic behavior of a steel specimen as shown in
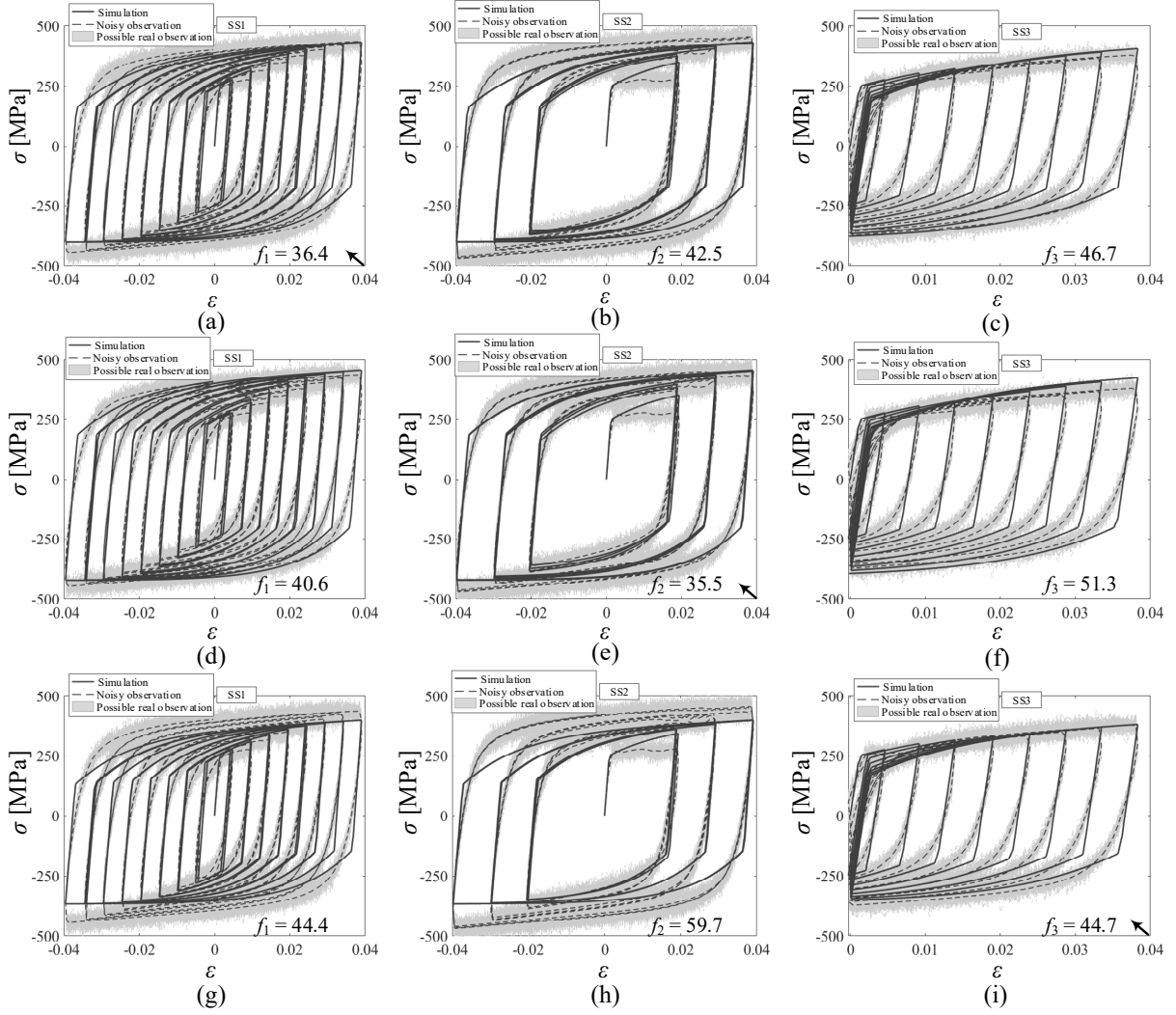
Figure 8: Comparison of test data and model predictions for cyclic tests of the specimen with parameters identified from a specific noisy experimental dataset. (a), (b), (c) Parameters from group 1; (d), (e), (f) Parameters from group 2; (g), (h), (i) Parameters from group 3.

Fig. 2. The specimen was tested under three different static cyclic loading histories SS1, SS2, and SS3 by Yamada and Jiao [39]. The pseudo time durations considered as load path parameters for SS1, SS2, and SS3 are 22.72, 18.20, and 6.20 s, respectively.

The test results were three experimental datasets of the true stress σ and true strain ϵ corresponding to the three loading histories. The values of σ and ϵ were derived from the engineering stress σ_e and engineering strain ϵ_e as

$$\sigma = (1 + \epsilon_e)\sigma_e \tag{23}$$

$$\epsilon = \ln(1 + \epsilon_e) \tag{24}$$

where σ_e at a specific increment was obtained by dividing the measured axial force by the initial cross-sectional area of the specimen, and ϵ_e calculated by dividing the axial deformation of the specimen by its initial length.

We fix Young's modulus of the steel at $E = 205.94$ GPa as it can be directly determined from the experimental re-

sults [39]. Therefore, a total of five parameters are identified for the specimen, i.e., $\mathbf{x} = [Q_\infty, b, \sigma_{y,0}, C_1, \gamma_1]$. The intervals associated with the material parameters are provided in Table 1. Poisson's ratio of the steel is 0.3.

Table 1
Material parameter intervals for the specimen.

Parameter	Lower bound	Upper bound
E [GPa]	205.94	–
$\sigma_{y,0}$ [MPa]	250	260
Q_∞ [MPa]	10	100
b	5	25
C_1 [MPa]	2000	8000
γ_1	10	100

We set $\tau_d = 0$ and 20 MPa as the standard deviation values for noise-free and noisy true stress measures, respectively. Exact observational noise can be calibrated experi-
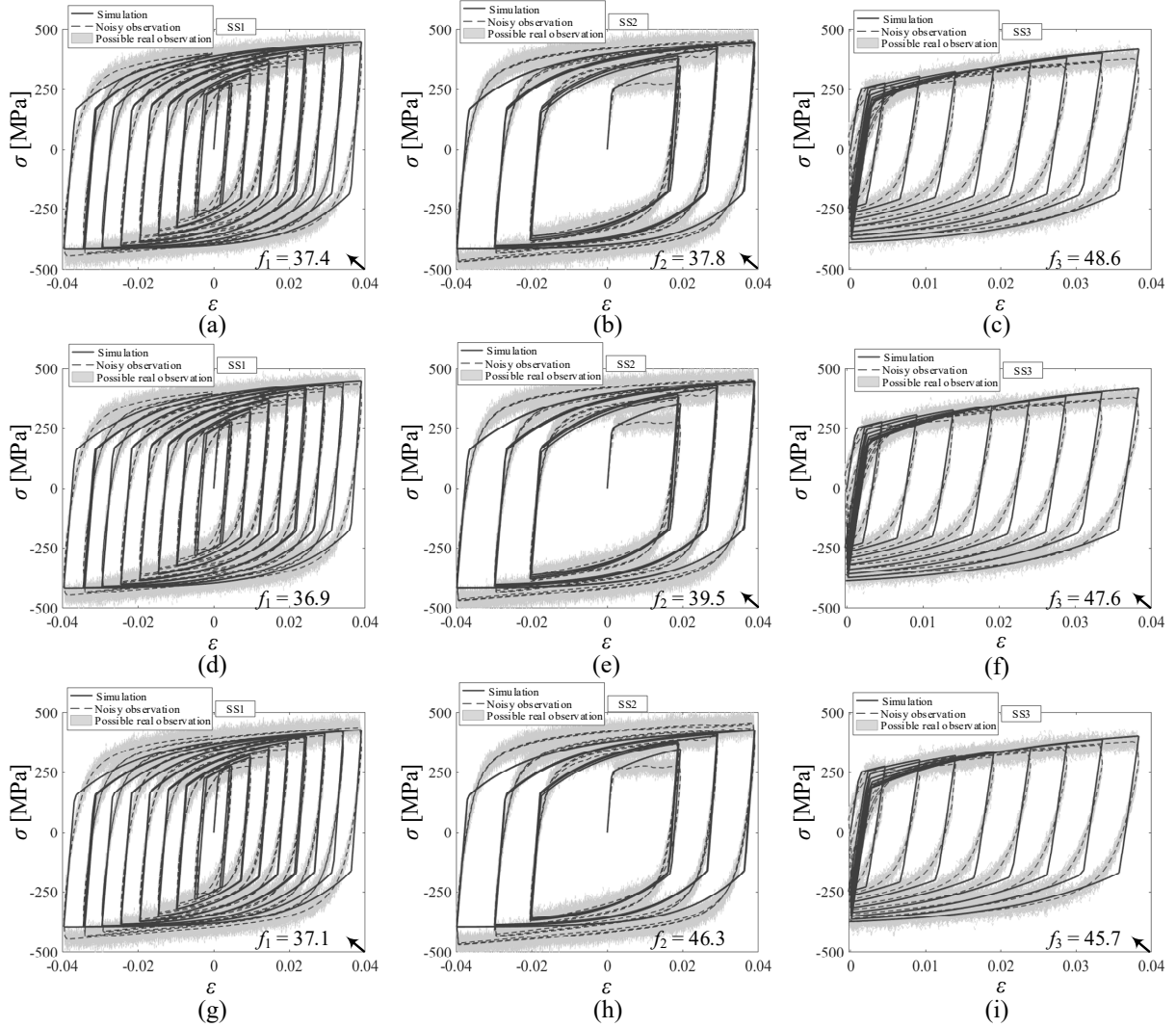
Figure 9: Comparison of test data and model predictions for cyclic tests of the specimen with parameters identified from a pair of two noisy experimental datasets. (a), (b), (c) Parameters from group 4; (d), (e), (f) Parameters from group 5; (g), (h), (i) Parameters from group 6.

Table 2

Parameters for GA.

Parameter	Value
Population size	4000
Maximum number of generations	100
Crossover fraction	65%
Elite transfer	2
Fitness function tolerance	10^{-6}
Constraint tolerance	10^{-6}
Parallel computing	'false'

mentally by obtaining measurement data without use of the specimen [12]. However, this is not the focus of the present study. The value of τ_f associated with $\tau_d = 20$ MPa is empirically determined for each loading history by the following three steps. First, we select three arbitrary sets of parameters within their bounds, and perform FE analysis to obtain the simulated $\sigma - \epsilon$ curve for each set. Second, a total of 10^4 samples are randomly generated surrounding each of the measured true stresses with $\tau_d = 20$ MPa, thereby producing a total of 10^4 values of the error function for each simulated $\sigma - \epsilon$ curve. These values are then used to evaluate the sampled standard deviation τ_f for each $\sigma - \epsilon$ curve. τ_f is further represented by the ratio τ_d/τ_f. Thus, three values of τ_d/τ_f are found for each loading history. Finally, τ_f is derived from the mean value of the three values of τ_d/τ_f. The empirical relation between τ_f and τ_d for each loading history is established and illustrated in Fig. 3. With $\tau_d = 20$ MPa, the values of τ_f associated with SS1, SS2, and SS3 are 0.40, 0.45, and 0.78 MPa, respectively.

To investigate how experimental datasets used for parameter identification affect the resulting parameters as well as their prediction ability, we classify the experimental datasets obtained from the three loading histories into a total of six groups indexed as 1, 2, 3, 4, 5, and 6, which correspond to

Table 3
Comparison of identification results obtained from different noise-free datasets of the specimen.

Group	$\sigma_{y,0}$ [MPa]	Q_∞ [MPa]	b	C_1 [MPa]	γ_1	f_1 [MPa]	f_2 [MPa]	f_3 [MPa]
1	250.004	42.105	5.001	7999.997	69.188	**30.372**	36.370	42.499
2	250.010	11.939	5.015	7999.769	70.308	37.898	**28.840**	48.692
3	250.008	11.596	5.000	8000.000	72.634	38.635	54.834	**39.986**
4	250.011	56.630	5.005	7999.996	69.096	**32.627**	**30.732**	45.261
5	250.056	57.661	5.000	8000.000	67.594	33.092	**30.397**	**45.497**
6	250.002	34.289	5.003	8000.000	67.985	**30.827**	40.150	**41.389**

Table 4
Comparison of identification results for the specimen obtained from BO, GA, and PSO.

Group	$\sigma_{y,0}$ [MPa]	Q_∞ [MPa]	b	C_1 [MPa]	γ_1	f_1 [MPa]	f_2 [MPa]	f_3 [MPa]	Simulation calls
1 [BO]	250.004	42.105	5.001	7999.997	69.188	**30.372**	36.370	42.499	100
1 [GA-1]	250.123	29.700	12.616	7397.003	61.045	**34.021**	43.787	43.110	100
1 [GA-2]	256.562	32.546	6.265	7409.248	65.632	**32.253**	44.282	39.747	200
1 [PSO-1]	251.419	41.499	7.598	6940.747	54.853	**33.010**	39.311	43.911	100
1 [PSO-2]	251.540	37.018	6.139	8000.000	64.631	**31.260**	37.088	43.217	200
4 [BO]	250.011	56.630	5.005	7999.996	69.096	**32.627**	**30.732**	45.261	200
4 [GA-1]	250.815	63.098	8.169	6440.088	59.086	**35.640**	**32.788**	49.728	200
4 [GA-2]	252.821	51.212	6.885	7536.697	65.882	**33.324**	**32.419**	47.323	400
4 [PSO-1]	250.000	55.367	6.836	7979.444	64.027	**34.491**	**30.918**	47.514	200
4 [PSO-2]	250.000	55.113	5.000	8000.000	72.489	**32.068**	**31.608**	44.957	400

SS1, SS2, SS3, SS1 & SS2, SS2 & SS3, and SS3 & SS1, respectively. The training dataset for each group is created by randomly generating 50 samples of the parameters and performing FE analysis for these samples to evaluate the corresponding error functions. As the specimen was axially loaded during the cyclic tests, we simply model it using one Abaqus linear hexahedral element of type C3D8 [47]. The maximum increment size for each loading history is set as 0.01 s. For each value of τ_d, each of groups 1, 2, and 3 is sequentially used as the training dataset and Eq. (11) as the error function. We do so for each of groups 4, 5, and 6 with Eq. (12). Note that the standard deviation values for groups 4, 5, and 6 are evaluated based on the fact that the error function values for SS1, SS2, and SS3 are independent. The identified parameters from each group are then used to predict $\sigma - \epsilon$ curves corresponding to the loading histories that are not used for parameter identification.

To examine the robustness of BO, we randomly generate three different training datasets for each group. Thus, three

sets of parameters are found from each group corresponding to the three BO attempts. The limit of number of BO iterations is set as 50. Therefore, the total number of simulations required for BO working on one of groups 1, 2, and 3 is 100 and for that working on one of groups 4, 5, and 6 is 200. In each iteration, BO uses genetic algorithm (GA) for maximizing $EI(\mathbf{x})$ or $AEI(\mathbf{x})$. The parameters characterizing GA are given in Table 2. A large population size of 4000 is to increase the chance of finding the global optimizer of the acquisition function in each BO iteration, which can mitigate the effect of the randomness of GA on the performance of BO, and this value has a minor effect on the computational cost of 50 BO iterations because the acquisition function is given analytically. Alternatively, any global optimization or multi-start local search algorithms can be used.

Fig. 4 shows the histories of BO for three different training datasets from each group with $\tau_d = 0$. The three BO attempts from each group tend to converge to an error function value after 50 iterations even though they start from different

Table 5
Comparison of identification results obtained from different noisy datasets of the specimen.

Group	$\sigma_{y,0}$ [MPa]	Q_∞ [MPa]	b	C_1 [MPa]	γ_1	f_1 [MPa]	f_2 [MPa]	f_3 [MPa]
1	250.000	40.133	5.000	8000.000	70.221	**36.376**	42.513	48.665
2	250.000	64.899	5.000	8000.000	72.468	40.577	**35.493**	51.314
3	250.000	10.000	5.000	8000.000	73.057	44.443	59.697	**44.695**
4	250.000	52.343	5.000	8000.000	68.855	**37.351**	**37.778**	48.611
5	250.000	46.445	5.000	8000.000	64.642	36.884	**39.526**	**47.615**
6	250.000	31.653	5.000	8000.000	69.114	**37.107**	46.342	**45.667**

Table 6

Statistical estimates of identified parameters for the specimen.

Parameter	Noise-free case		Noisy case	
	mean	std	mean	std
$\sigma_{y,0}$ [MPa]	250.059	0.192	251.714	3.824
Q_∞ [MPa]	45.687	19.277	39.036	16.465
b	5.023	0.057	5.003	0.015
C_1 [MPa]	7997.540	7.595	7999.120	5.026
γ_1	69.447	1.835	71.099	3.4362

initial training datasets.

Figs. 5 and 6 compare the predicted $\sigma - \epsilon$ curves associated with the best set of identified parameters from each group and the experimental ones, where f_1, f_2, and f_3 denote the error function values at the identified parameters corresponding to SS1, SS2, and SS3, respectively, and the arrow at the lower right corner of the figure indicates the experimental dataset used for identification. All identified parameters well reproduce the $\sigma - \epsilon$ curves from the loading histories not used for identification. Fig. 5 (see along each column) indicates the bias toward a specific loading history as the error function corresponding to the loading history used for identification is smaller than those predicted by the parameters identified from other loading histories. The parameters identified from a specific loading history (e.g., SS3) may lead to large errors in prediction of the $\sigma - \epsilon$ curves associated with other loading histories (e.g., SS1 and SS2). These errors can be reduced by using the experimental datasets from two loading histories for identification as observed in Fig. 6.

Table 3 lists the identified parameters from each group without observational noise, where the boldface value indicates the error function corresponding to the experimental dataset used for identification. There is no major difference in values of $\sigma_{y,0}$, b, C_1, and γ_1 among the groups. However, Q_∞ is affected by the group. We also compare the identification results by BO working on groups 1 and 4 with those by minimizing the corresponding error functions using GA and particle swarm optimization (PSO) algorithm. For each group (i.e., 1 and 4), we perform GA and PSO two times, namely, GA-1, GA-2, PSO-1, and PSO-2. Each algorithm has a population of 20 individuals (or particles). Other parameters for PSO are set as default using the MATLAB Global Optimization Toolbox. The numbers of iterations for GA(PSO)-1 and GA(PSO)-2 are limited at five and ten, respectively. Thus, the numbers of simulations for GA(PSO)-1 and GA(PSO)-2 are 100 and 200, which are the same as those for BO when working on group 1, and 200 and 400 when working on group 4, respectively. Comparison results in Table 4 show that BO outperforms GA and PSO in terms of the optimized error function value as well as the prediction ability of identified parameters when expending the same number of simulation calls. Interestingly, GA and PSO should double the number of simulations to offer performance approaching that of BO.

Fig. 7 shows the histories of BO for three different training datasets from each group with $\tau_d = 20$ MPa. In the presence of observational noise, BO attempts from different datasets result in different error function values after 50 iterations. Still, there is a good agreement between the predicted $\sigma - \epsilon$ curves at the best set of the parameters identified from each group and the experimental ones, as shown in Figs. 8 and 9, where f_1, f_2, and f_3 are the mean values of the error function at the identified parameters corresponding to SS1, SS2, and SS3, respectively. Each of these mean values is evaluated from a total of 10^4 samples generated surrounding each of the corresponding experimental measures using $\tau_d = 20$ MPa (i.e., possible real observations in Figs. 8 and 9). Observational noise does not affect the bias toward a set of experimental measures even though it considerably increases the minimized error function as compared with that from the noise-free case. Table 5 reports the best identified material parameters from each group for $\tau_d = 20$ MPa. The values of $\sigma_{y,0}$, b, C_1, and γ_1 among the groups are similar, while Q_∞ varies across the groups as it plays an important role in mitigating the bias toward a set of experimental results.

To provide statistical estimates of each identified parameter, additional three BO trials are carried out for each group of both noise-free and noisy cases. Six groups of experimental datasets, therefore, provide a total of 36 sets of parameters, which are used to estimate the mean and standard deviation (std) of each parameter, as listed in Table 6. There is a major difference in the std values of $\sigma_{y,0}$ obtained from the noise-free and noisy cases. The std of Q_∞ is very large in both cases, which again confirms that Q_∞ is group-dependent.

Table 7

Material parameter intervals for the cantilever.

	Parameter	Lower bound	Upper bound
Web	E [GPa]	175.05	–
	$\sigma_{y,0}$ [MPa]	300	340
	Q_∞ [MPa]	10	100
	b	5	25
	C_1 [MPa]	2000	8000
	γ_1	10	100
Flange	E [GPa]	175.05	–
	$\sigma_{y,0}$ [MPa]	270	290
	Q_∞ [MPa]	10	100
	b	5	25
	C_1 [MPa]	2000	8000
	γ_1	10	100

4.2. Parameters for a bi-material cantilever

This section applies BO to identification of the parameters for a steel cantilever tested under three different cyclic loading histories RH1, RH2, and RH3, as shown in Fig. 10. The pseudo time durations considered as load path parameters for RH1, RH2, and RH3 are 5.55, 3.62, and 4.43 s, respectively. The cantilever is a built-up wide-flange beam H-244 × 175 ×
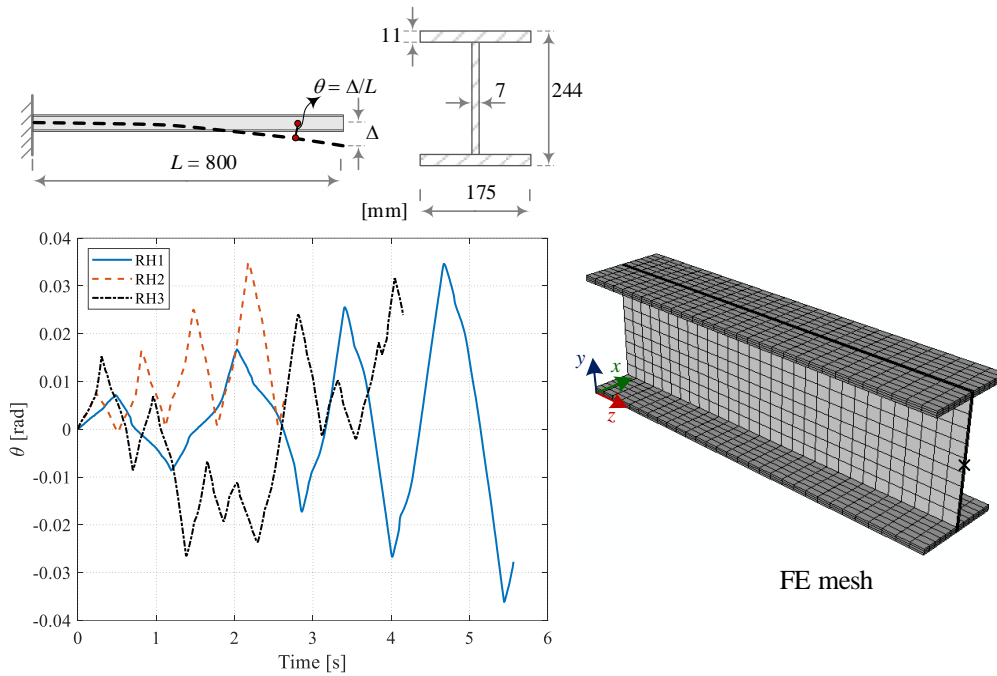
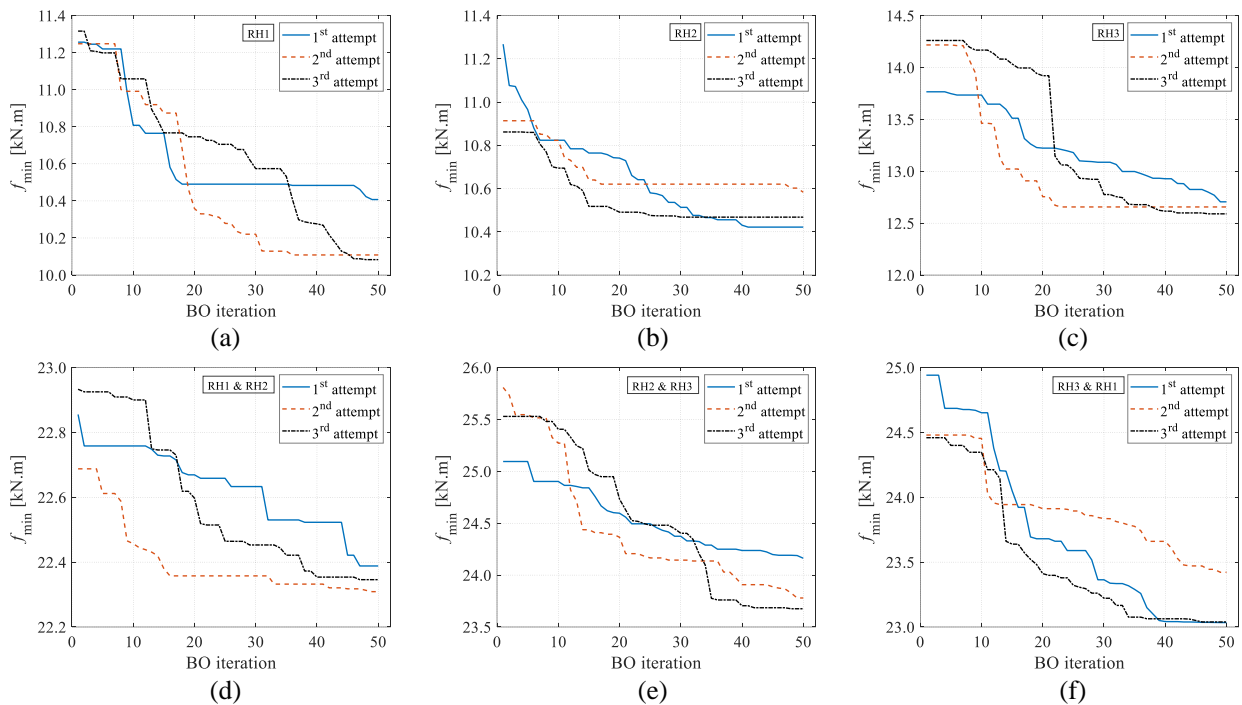Figure 10: The cantilever, its FE mesh, and three loading histories for cyclic tests [6, 39].

Figure 11: Histories of BO for different noise-free experimental datasets from cyclic tests of the cantilever. (a)–(f) obtained from groups 1–6, respectively.

7×11 in Japanese specification. The web and flange have the same Young's modulus but different plastic material parameters. Thus, the cantilever can be viewed as a structural component of two different materials. During the cyclic tests conducted by Yamada and Jiao [39], the left end of the cantilever was fixed, while forced vertical displacement was applied at the right end. The deflection angle θ of the cantilever is defined as the ratio of the vertical tip displacement Δ mm to the cantilever length $L = 800$ mm, i.e., $\theta = \Delta/L$.

The test results associated with the three loading histories consist of three datasets of the bending moment M at the cantilever support and the deflection angle θ [39]. Young's

Table 8

Comparison of identification results obtained from different noise-free datasets of the cantilever.

	Parameter	Group					
		1	2	3	4	5	6
Web	$\sigma_{y,0}$ [MPa]	339.957	300.049	339.924	336.672	330.990	339.880
	Q_∞ [MPa]	99.890	99.827	99.923	99.874	99.972	99.974
	b	24.946	24.910	24.952	24.761	24.809	24.869
	C_1 [MPa]	7160.521	7968.280	3596.024	7766.642	7999.973	7682.000
	γ_1	99.622	10.178	99.967	99.999	69.775	99.577
Flange	$\sigma_{y,0}$ [MPa]	271.262	270.004	270.047	270.008	270.001	270.061
	Q_∞ [MPa]	10.065	10.135	10.000	10.140	10.107	10.001
	b	5.008	5.120	5.047	5.534	5.092	5.041
	C_1 [MPa]	7999.840	5170.983	6636.794	6537.440	5140.712	7998.996
	γ_1	66.632	62.092	71.565	60.903	69.258	84.763
Error	f_1 [kNm]	**10.083**	14.166	11.469	**10.781**	12.760	**10.107**
	f_2 [kNm]	12.836	**10.422**	11.377	**11.491**	**10.794**	12.380
	f_3 [kNm]	13.601	14.060	**12.590**	12.959	**12.881**	**12.961**

modulus and Poisson's ratio for the web and flange are fixed at $E = 175.05$ GPa and 0.3, respectively. Thus, a total of ten parameters (i.e., five for the web and five for the flange) are identified for the cantilever. The interval for each parameter is provided in Table 7.

The cantilever is modeled using Abaqus [47]. A fine mesh, as shown in Fig. 10, consisting of 4960 nodes and 3510 linear hexahedral elements of type C3D8 is generated. The maximum increment size for each loading history is set as 0.01 s.

We set $\tau_d = 0$ and 15 kNm as the standard deviation values for noise-free and noisy moment measures, respectively. For $\tau_d = 15$ kNm, the same procedure as described in Section 4.1 is carried out to determine τ_f corresponding to τ_d for each loading history. As a result, values of τ_f associated with RH1, RH2, and RH3 are 0.55, 0.81, and 0.65 kNm, respectively.

We also classify the experimental datasets from the three loading histories into a total of six groups indexed as 1, 2, 3, 4, 5, and 6, corresponding to RH1, RH2, RH3, RH1 & RH2, RH2 & RH3, and RH3 & RH1, respectively. For each value of τ_d, there different training datasets are created for each group to investigate the performance of BO. Each training dataset is constructed by randomly generating 100 samples of the material parameters and performing elastoplastic analysis for each generated sample to evaluate the corresponding error function. The remaining steps of the identification process using the training datasets from each group are identical to those performed in Section 4.1.

Fig. 11 shows the histories of BO attempts for three different training datasets from each group with $\tau_d = 0$. Although BO with different initial training datasets cannot provide a unique set of the parameters after 50 iterations, it considerably reduces the error function as it terminates.

Table 8 provides the best set of identified parameters that gives the smallest error function value among three values obtained from each group. Figs. 12 and 13 compare the pre-

dicted $M - \theta$ curves at the parameters identified from each group listed in Table 8 and the experimental ones, where f_1, f_2, and f_3 are the error function values corresponding to RH1, RH2, and RH3, respectively. As expected, all predicted $M - \theta$ curves associated with the loading histories not used for identification are in good agreement with the experimental ones. Results in Table 8 and Fig. 12 show that the dataset-specific bias is considerable when the experimental dataset from a single loading history is used for identification. This bias can be mitigated when using the experimental datasets from two loading histories; see Fig. 13.

Fig. 14 shows the histories of BO attempts for three different training datasets from each group with $\tau_d = 15$ kNm. Under the effect of observational noise, $c_{\min}$ considerably fluctuates in the very first iterations of BO and becomes stable when approaching the 50th iteration. Regardless of the difference in BO histories from each group, the identified parameters from all groups still arrive at a good agreement between the predicted $M - \theta$ curves and the experimental ones, as shown in Figs. 15 and 16, where f_1, f_2, and f_3 denote the mean values of the error function (estimated from a total of 10^4 samples generated surrounding each of moment measures with $\tau_d = 15$ kNm) corresponding to RH1, RH2, and RH3, respectively. Table 9 lists the best set of material parameters from each group with $\tau_d = 15$ kNm. It is observed that $\sigma_{y,0}$, Q_∞, and b tend to concentrate to their bounds under the effect of observational noise.

The identified parameters in Tables 8 and 9 are further used to reproduce the true stress-true strain relation obtained from a monotonic tensile coupon test of the cantilever's flange [39]. This is feasible because the flange is deemed to be subjected to cyclic axial stress during the cyclic tests. This also suggests a potential approach to identification of material parameters from structural tests, which is different from the conventional way of parameter identification from material tests. It would be more interesting if cyclic axial behavior of the flange can be simulated using the parameters identi-
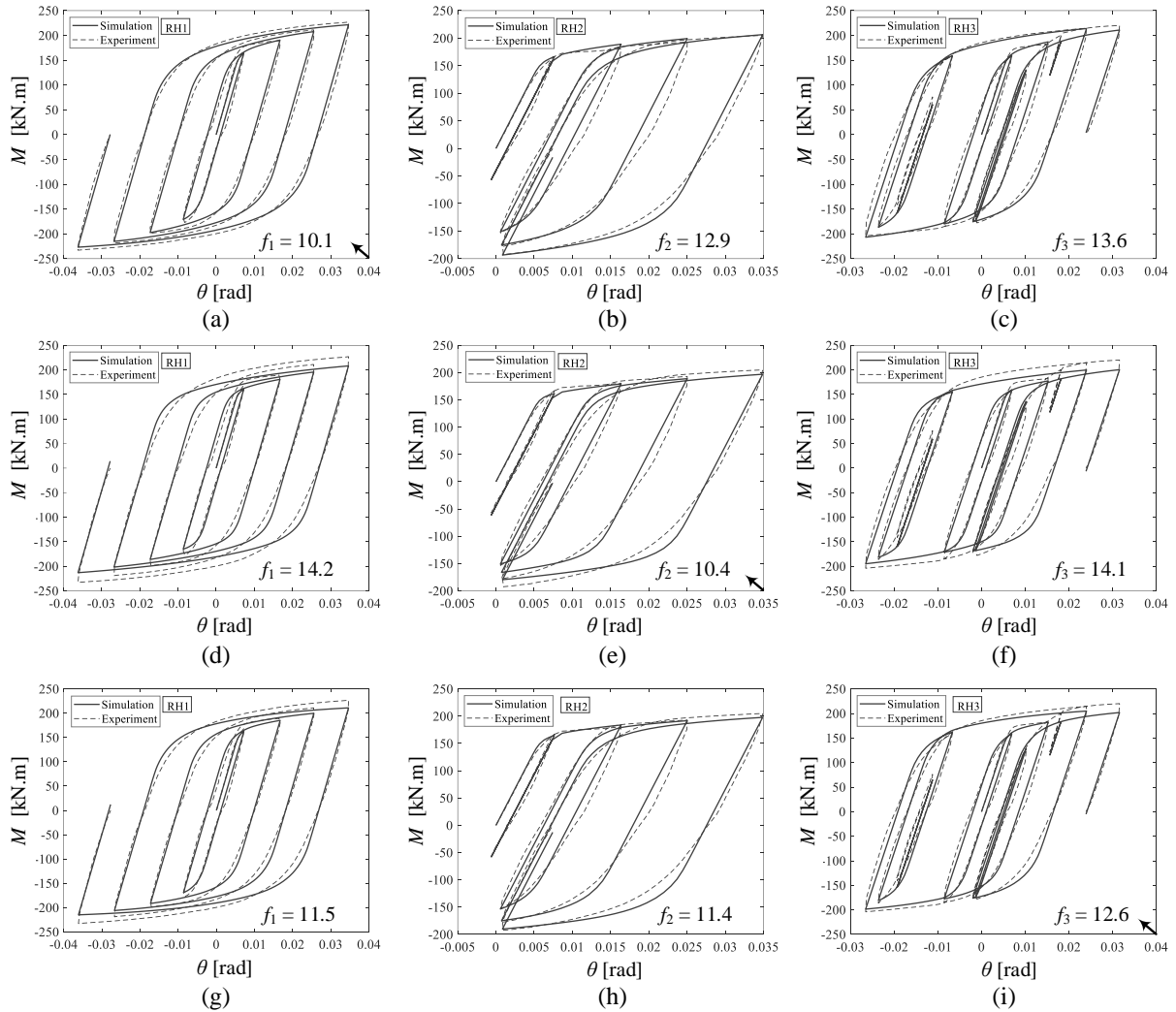
Figure 12: Comparison of test data and model predictions for cyclic tests of the cantilever with parameters identified from a specific noise-free experimental dataset. (a), (b), (c) Parameters from group 1; (d), (e), (f) Parameters from group 2; (g), (h), (i) Parameters from group 3.

Table 9

Comparison of identification results obtained from different noisy datasets of the cantilever.

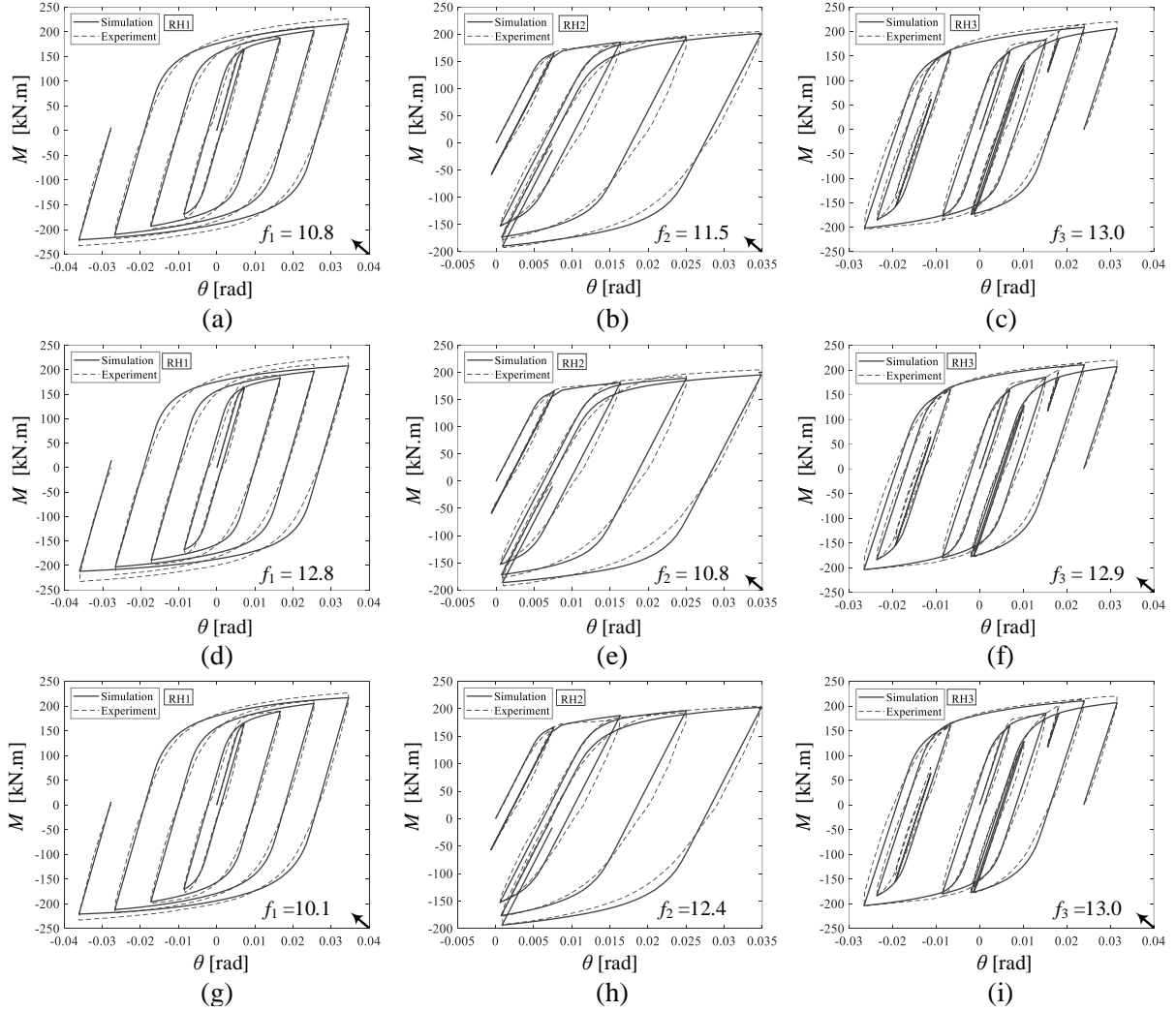| | Parameter | Group | | | | | |
		1	2	3	4	5	6
	$\sigma_{y,0}$ [MPa]	339.966	300.000	339.998	339.993	339.985	339.999
	Q_∞ [MPa]	99.945	10.000	100.000	10.082	10.013	99.995
Web	b	24.999	5.000	5.003	5.198	5.013	24.997
	C_1 [MPa]	6813.514	7999.978	2000.006	7382.899	7999.971	7999.997
	γ_1	10.026	10.000	99.999	43.900	99.999	99.999
	$\sigma_{y,0}$ [MPa]	271.518	270.000	270.001	270.008	270.013	270.001
	Q_∞ [MPa]	10.002	10.000	10.003	23.168	10.007	10.002
Flange	b	19.290	5.000	5.000	5.077	12.427	5.003
	C_1 [MPa]	7999.942	7041.196	7999.997	7999.855	5017.701	7999.999
	γ_1	86.961	99.988	99.997	86.555	74.038	99.997
	f_1 [kNm]	**18.174**	21.549	18.935	**18.734**	21.389	**18.263**
Error	f_2 [kNm]	19.815	**18.286**	19.201	**19.052**	**18.578**	19.060
	f_3 [kNm]	20.442	21.016	**19.922**	20.281	**20.535**	**19.708**

Figure 13: Comparison of test data and model predictions for cyclic tests of the cantilever with parameters identified from a pair of two noise-free experimental datasets. (a), (b), (c) Parameters from group 4; (d), (e), (f) Parameters from group 5; (g), (h), (i) Parameters from group 6.

fied from structural tests. However, cyclic material tests are not available for the flange. Figs. 17(a) and (b) compare the true stress-strain relations simulated by the identified parameters for the flange from noise-free and noisy experimental datasets with the experimental one, respectively. All sets of the identified parameters of the flange can capture the monotonic hardening process of the material. The yield plateau cannot be simulated because of the nature of the nonlinear combined isotropic/kinematic hardening model in Eqs. (5) and (8).

5. Conclusions

The cyclic response of a structure depends on the constitutive laws of its materials. If this response can be measured experimentally, an inverse problem can be formulated for identification of the underlying material parameters. We have successfully applied BO to solving such an inverse problem considering noise-free and noisy experimental measures.

Starting at a small number of material parameter sets, a constructed GP model probabilistically describes the relationship between the parameters and the error function. An acquisition function addressing the trade-off between exploitation and exploration of BO is formulated in each BO iteration for intelligently selecting a new promising parameter vector. The acquisition function allows BO to select a good parameter vector in the next iteration without calling any numerical simulation. The parameters are found as BO terminates and their prediction ability is confirmed through reproducing the material behavior under different cyclic loadings.

Through two identification examples of a steel specimen and a bi-material cantilever, BO demonstrates its good ability to identify the parameters for the nonlinear combined isotropic/kinematic hardening model for structural steels. The material parameters obtained by BO with and without considering observational noise can reliably predict the cyclic behavior of steel subjected to different loading conditions. It is also found that BO outperforms GA and PSO in terms
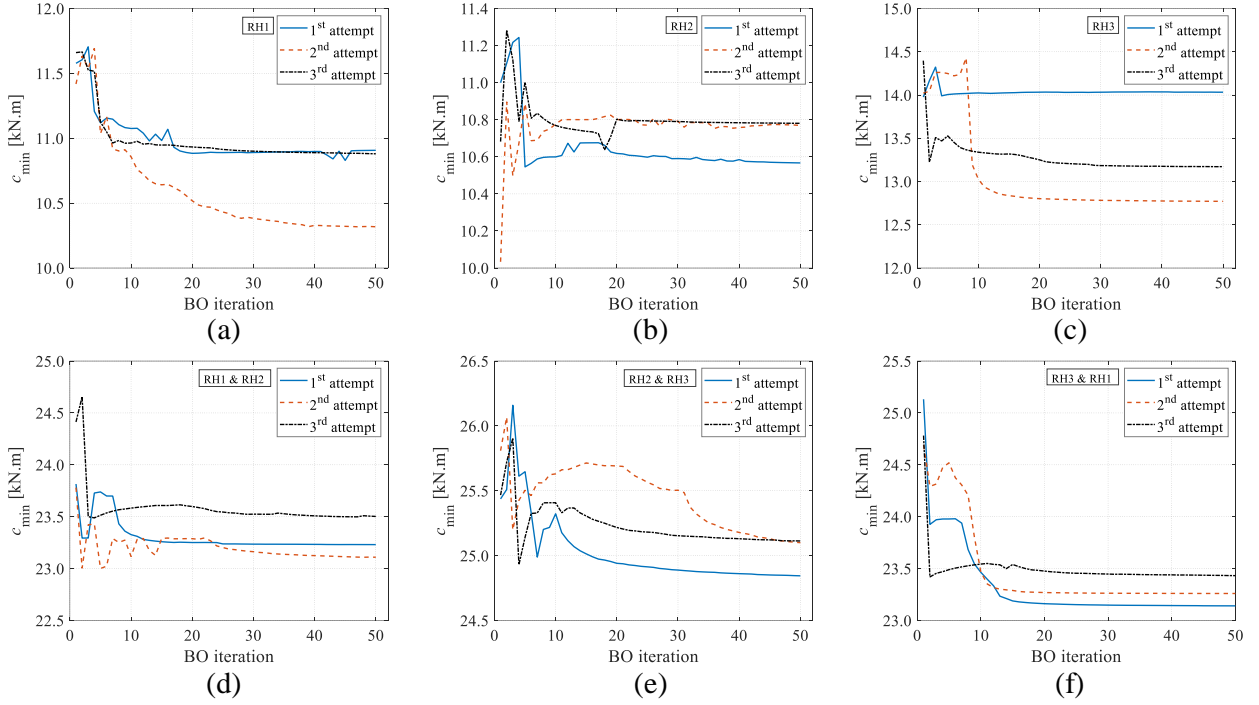
Figure 14: Histories of BO for different noisy experimental datasets from cyclic tests of the cantilever. (a)–(f) obtained from groups 1–6, respectively.

of the prediction performance of identified parameters when expending the same number of simulation calls. Moreover, the identification results for the cantilever suggest a possibility of identifying the material parameters from structural tests even for the structure with two different materials. This is remarkable because cyclic material tests are difficult and usually not carried out before structural tests.

The dataset-specific bias in both noise-free and noisy cases is notable if only one experimental dataset obtained from a specific loading history is used for parameter identification. Such bias can be mitigated when using the experimental datasets from different loading histories. As a recommendation, experimental measures from various loading histories should be used for parameter identification simultaneously.

A nontrivial extension of this work is to apply the BO approach to identifying the parameters for steel structural systems under cyclic loading where different responses, such as story drifts, accelerations, and stresses, can be experimentally measured. These responses and their correlation may be used for formulation of a multi-objective parameter identification problem. The BO should be modified accordingly, and this is under our investigation. Another extension of this work is to handle modeling errors and uncertainty in material parameters. The former may combine a sophisticated isotropic hardening model with a kinematic hardening model of more back-stress components. The latter may find a posterior PDF describing the parameters using the probabilistic approach with a constant prior PDF obtained from BO.

Acknowledgments

Appendix: Gaussian process (GP)

Consider the training dataset $\mathcal{D} = \{\mathbf{X}, \mathbf{y}\} = \{\mathbf{x}_i, y_i\}_{i=1}^{N}$, where $\mathbf{x}_i \in \mathbb{R}^{+n}$ are n-dimensional vectors of the material parameters and $y_i \in \mathbb{R}$ are the corresponding (noise-free or noisy) error function values. We establish the relationship between the input vector $\mathbf{x}$ and output variable y using the mapping $y = \hat{g}(\mathbf{x}) + \omega_f : \mathbb{R}^{+n} \to \mathbb{R}$, where $\hat{g}(\mathbf{x})$ is a Gaussian conditioned on $\mathcal{D}$ and $\omega_f \sim \mathcal{N}\left(0, \tau_f^2\right)$ is additive Gaussian noise.

A GP assumes that any finite subset of an infinite set of the error function values has a joint Gaussian distribution [29]. For the set of N parameter vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, the corresponding error function values $\{y_1, \ldots, y_N\}$ are distributed according to

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \sim \mathcal{N}_N\left(\begin{bmatrix} m(\mathbf{x}_1) \\ \vdots \\ m(\mathbf{x}_N) \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} + \tau_f^2 \mathbf{I}_N\right)$$

(A.1)

where $\mathcal{N}_N$ denotes an N-variate Gaussian; $\mathbf{I}_N$ the N-by-N identity matrix; and $m(\mathbf{x}) = \mathbb{E}[\hat{g}(\mathbf{x})]$ and $k(\mathbf{x}, \mathbf{x}')$ the mean
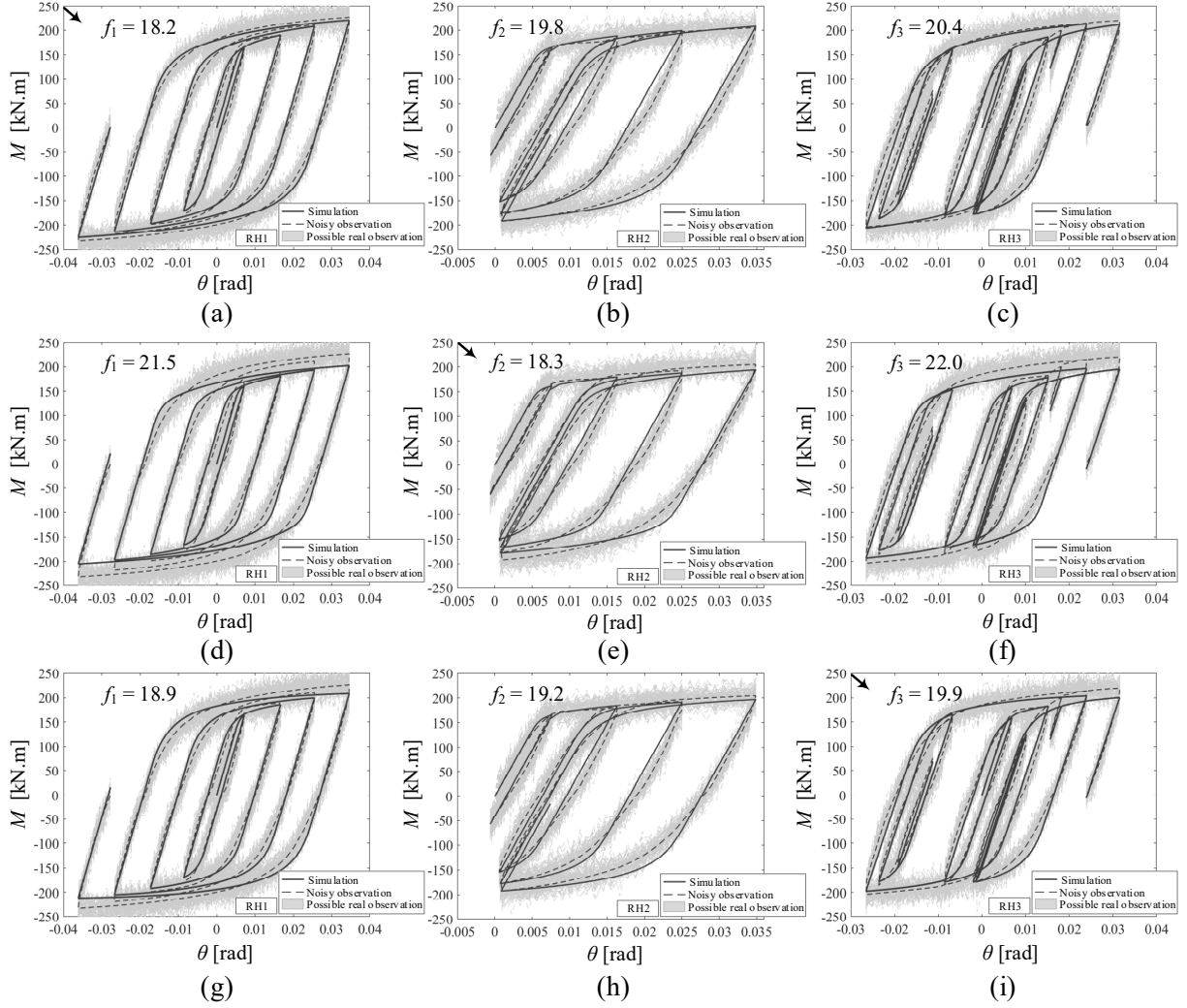
Figure 15: Comparison of test data and model predictions for cyclic tests of the cantilever with parameters identified from a specific noisy experimental dataset. (a), (b), (c) Parameters from group 1; (d), (e), (f) Parameters from group 2; (g), (h), (i) Parameters from group 3.

and covariance kernel functions, respectively. The mean function in this study is set as $m(\mathbf{x}) = 0$ because the covariance kernel function is flexible enough to handle the role of $m(\mathbf{x})$ [29]. The covariance kernel function is defined for any pair of the parameter vectors $\mathbf{x}$ and $\mathbf{x}'$ to measure the similarity between two corresponding error function values $y = \hat{g}(\mathbf{x})$ and $y = \hat{g}(\mathbf{x}')$, such that

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[\left(\hat{g}(\mathbf{x}) - m(\mathbf{x})\right)\left(\hat{g}(\mathbf{x}') - m(\mathbf{x}')\right)\right] \quad \text{(A.2)}$$

In this work, we use Gaussian kernel as

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^T(\mathbf{x} - \mathbf{x}')}{2l^2}\right) \quad \text{(A.3)}$$

where l denotes the characteristic length-scale parameter determined by maximizing the likelihood estimation of $\mathcal{D}$. A 5-fold cross-validation strategy incorporated in the built-in MATLAB function *fitrgp* is used to tune the kernel parameter.

Once l has been determined, the information in Eq. (A.1) is used for predicting the error function value y^* at a new parameter vector $\mathbf{x}^*$, i.e., $y^*|\mathbf{y} = \hat{g}(\mathbf{x}^*)$. Because of the GP nature, the joint PDF of y^* and $\mathbf{y}$ is also a Gaussian. Let $\mathbf{m}(\mathbf{X}) = [m(\mathbf{x}_1), \ldots, m(\mathbf{x}_N)]^T$, we have

$$\begin{bmatrix} y^* \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N}_{N+1}\left(\begin{bmatrix} m(\mathbf{x}^*) \\ \mathbf{m}(\mathbf{X}) \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}^*, \mathbf{x}^*) & \mathbf{K}(\mathbf{x}^*, \mathbf{X}) \\ \mathbf{K}(\mathbf{x}^*, \mathbf{X})^T & \mathbf{K}(\mathbf{X}, \mathbf{X}) \end{bmatrix}\right) \quad \text{(A.4)}$$

where

$$\mathbf{K}(\mathbf{x}^*, \mathbf{X}) = \left[k(\mathbf{x}^*, \mathbf{x}_1), \ldots, k(\mathbf{x}^*, \mathbf{x}_N)\right] \quad \text{(A.5)}$$

$$\mathbf{K}(\mathbf{X}, \mathbf{X}) = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} + \tau_f^2 \mathbf{I}_N \quad \text{(A.6)}$$

The conditional Gaussian $y^*|\mathbf{y} = \hat{g}(\mathbf{x}^*)$ can be derived from Eq. (A.4) using the standard conditioning rule [29],
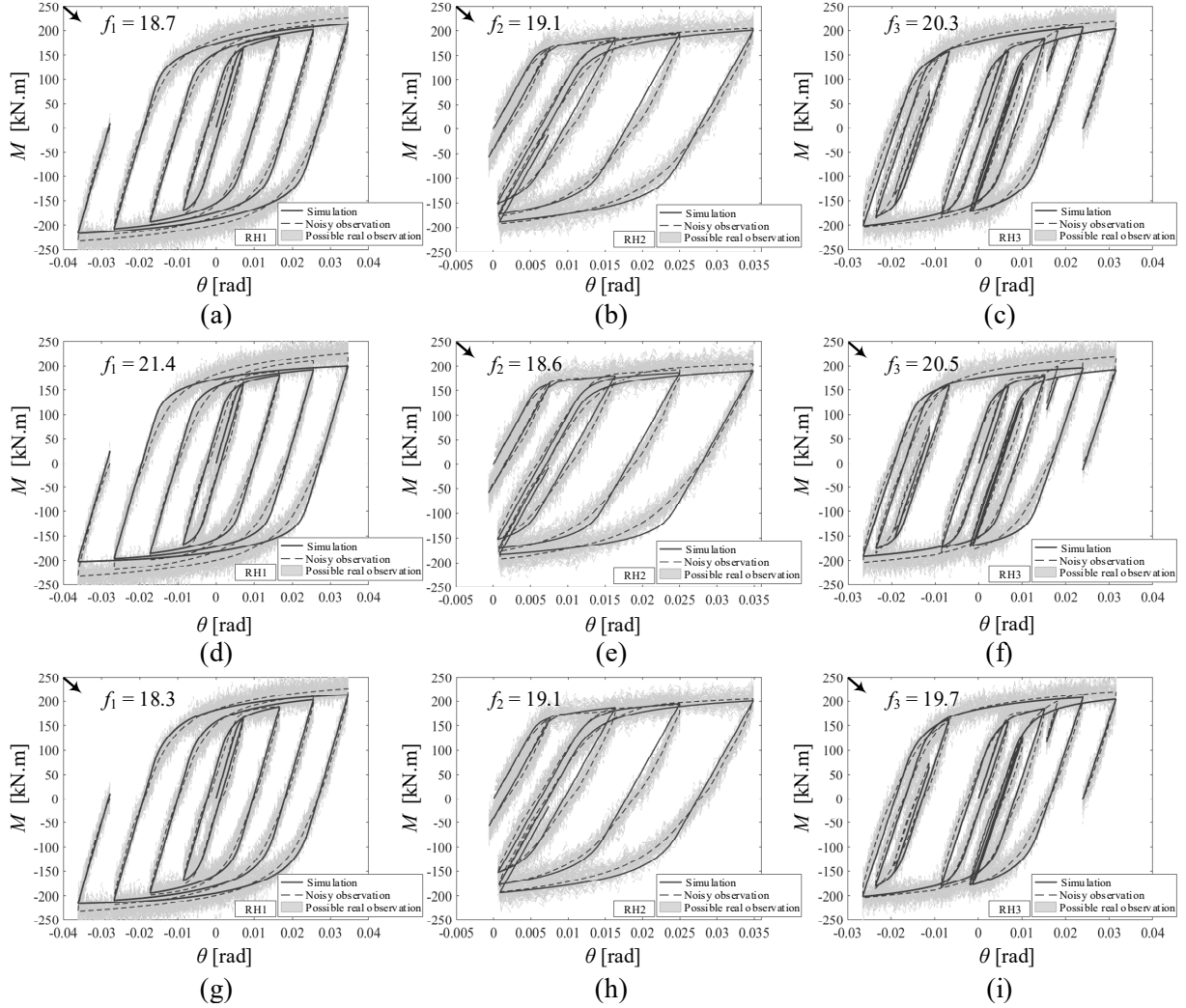
Figure 16: Comparison of test data and model predictions for cyclic tests of the cantilever with parameters identified from a pair of two noisy experimental datasets. (a), (b), (c) Parameters from group 4; (d), (e), (f) Parameters from group 5; (g), (h), (i) Parameters from group 6.

such that

$$y^*|\mathbf{y} \sim \mathcal{N}\left(\mu_{y^*}(\mathbf{x}^*), \tau_{y^*}^2(\mathbf{x}^*)\right) \tag{A.7}$$

where

$$\mu_{y^*}(\mathbf{x}^*) = m(\mathbf{x}^*) + \mathbf{K}(\mathbf{x}^*, \mathbf{X})\mathbf{K}(\mathbf{X}, \mathbf{X})^{-1}(\mathbf{y} - \mathbf{m}(\mathbf{X})) \tag{A.8}$$

$$\tau_{y^*}^2(\mathbf{x}^*) = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{K}(\mathbf{x}^*, \mathbf{X})\mathbf{K}(\mathbf{X}, \mathbf{X})^{-1}\mathbf{K}(\mathbf{x}^*, \mathbf{X})^T \tag{A.9}$$

References

[1] M. Ohsaki, T. Miyamura, M. Kohiyama, M. Hori, H. Noguchi, H. Akiba, K. Kajiwara, T. Ine, High-precision finite element analysis of elastoplastic dynamic responses of super-high-rise steel frames, Earthquake Engineering and Structural Dynamics 38 (5) (2009) 635–654. doi:10.1002/eqe.900.

[2] M. Wang, Y. Shi, Y. Wang, G. Shi, Numerical study on seismic behaviors of steel frame end-plate connections, Journal of Constructional Steel Research 90 (2013) 140–152. doi:10.1016/j.jcsr.2013.07.033.

[3] F. Yoshida, T. Uemori, K. Fujiwara, Elastic-plastic behavior of steel sheets under in-plane cyclic tension-compression at large strain, International Journal of Plasticity 18 (5) (2002) 633–659. doi:10.1016/S0749-6419(01)00049-3.

[4] J. L. Chaboche, A review of some plasticity and viscoplasticity constitutive theories, International Journal of Plasticity 24 (10) (2008) 1642–1693. doi:10.1016/j.ijplas.2008.03.009.

[5] A. H. Mahmoudi, S. M. Pezeshki-Najafabadi, H. Badnava, Parameter determination of Chaboche kinematic hardening model using a multi objective genetic algorithm, Computational Materials Science 50 (3) (2011) 1114–1122. doi:10.1016/j.commatsci.2010.11.010.

[6] M. Ohsaki, T. Miyamura, J. Y. Zhang, A piecewise linear isotropic-kinematic hardening model with semi-implicit rules for cyclic loading and its parameter identification, Computer Modeling in Engineering & Sciences 111 (4) (2016) 303–333. doi:10.3970/cmes.2016.111.303.

[7] A. Nath, K. K. Ray, S. V. Barai, Evaluation of ratcheting behaviour in cyclically stable steels through use of a combined kinematic-isotropic hardening rule and a genetic algorithm optimization technique, International Journal of Mechanical Sciences 152 (2019) 138–150. doi:10.1016/j.ijmecsci.2018.12.047.

[8] R. Carreño, H. Lotfizadeh K., P. Conte J., I. Restrepo J., Material model parameters for the Giuffrè-Menegotto-Pinto uniaxial steel stress-strain model, Journal of Structural Engineering 146 (2) (2020)
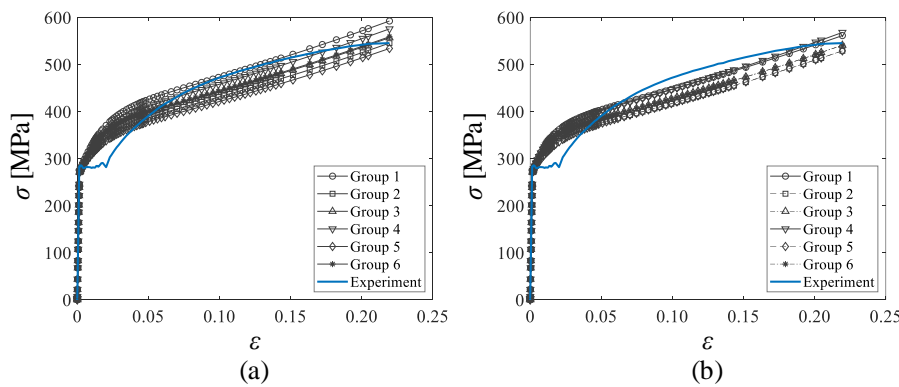
Figure 17: Comparison of test data and model predictions for coupon test of the cantilever's flange. (a) Parameters from noise-free experimental datasets; (b) Parameters from noisy experimental datasets.

04019205. doi:10.1061/(ASCE)ST.1943-541X.0002505.

[9] R. Hartloper Alexander, de Castro e Sousa Albano, G. Lignos Dimitrios, Constitutive modeling of structural steels: nonlinear isotropic/kinematic hardening material model and its calibration, Journal of Structural Engineering 147 (4) (2021) 04021031. doi:10.1061/(ASCE)ST.1943-541X.0002964.

[10] T. Marwala, S. Sibisi, Finite element model updating using Bayesian framework and modal properties, Journal of Aircraft 42 (1) (2005) 275–278. doi:10.2514/1.11841.

[11] B. V. Rosić, A. Kučerová, J. Sýkora, O. Pajonk, A. Litvinenko, H. G. Matthies, Parameter identification in a probabilistic setting, Engineering Structures 50 (2013) 179–196. doi:10.1016/j.engstruct.2012.12.029.

[12] H. Rappel, L. A. A. Beex, L. Noels, S. P. A. Bordas, Identifying elasto-plastic parameters with Bayes' theorem considering output error, input error and model uncertainty, Probabilistic Engineering Mechanics 55 (2019) 28–41. doi:10.1016/j.probengmech.2018.08.004.

[13] J. L. Beck, L. S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, Journal of Engineering Mechanics 124 (4) (1998) 455–461. doi:10.1061/(ASCE)0733-9399(1998)124:4(455).

[14] A. Tarantola, Inverse problem theory and methods for model parameter estimation, SIAM, 2005. doi:10.1137/1.9780898717921.

[15] H. Wang, Y. Zeng, X. Yu, G. Li, E. Li, Surrogate-assisted Bayesian inference inverse material identification method and application to advanced high strength steel, Inverse Problems in Science and Engineering 24 (7) (2016) 1133–1161. doi:10.1080/17415977.2015.1113960.

[16] U. Alper, T. Panos, Constitutive model for cyclic response of structural steels with yield plateau, Journal of Structural Engineering 137 (2) (2011) 195–206. doi:10.1061/(ASCE)ST.1943-541X.0000287.

[17] F. Hu, G. Shi, Y. Shi, Constitutive model for full-range elasto-plastic behavior of structural steels with yield plateau: Formulation and implementation, Engineering Structures 171 (2018) 1059–1070. doi:10.1016/j.engstruct.2016.02.037.

[18] S. Talatahari, A. Kaveh, N. Mohajer Rahbari, Parameter identification of Bouc-Wen model for MR fluid dampers using adaptive charged system search optimization, Journal of Mechanical Science and Technology 26 (8) (2012) 2523–2534. doi:10.1007/s12206-012-0625-y.

[19] B. M. Chaparro, S. Thuillier, L. F. Menezes, P. Y. Manach, J. V. Fernandes, Material parameters identification: Gradient-based, genetic and hybrid optimization algorithms, Computational Materials Science 44 (2) (2008) 339–346. doi:10.1016/j.commatsci.2008.03.028.

[20] S. Shirgir, B. F. Azar, A. Hadidi, Reliability-based simplification of Bouc-Wen model and parameter identification using a new hybrid algorithm, Structures 27 (2020) 297–308. doi:10.1016/j.istruc.2020.05.024.

[21] R. Ghajar, N. Naserifar, H. Sadati, J. Alizadeh K., A neural network approach for predicting steel properties characterizing cyclic ramberg-osgood equation, Fatigue & Fracture of Engineering Materials & Structures 34 (7) (2011) 534–544. doi:10.1111/j.1460-2695.2010.01545.x.

[22] T. A. Horton, I. Hajirasouliha, B. Davison, Z. Ozdemir, Accurate prediction of cyclic hysteresis behaviour of rbs connections using deep learning neural networks, Engineering Structures 247 (2021) 113156. doi:10.1016/j.engstruct.2021.113156.

[23] M. C. Kennedy, A. O'Hagan, Bayesian calibration of computer models, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63 (3) (2001) 425–464. doi:10.1111/1467-9868.00294.

[24] S. Arridge, P. Maass, O. Öktem, C.-B. Schönlieb, Solving inverse problems using data-driven models, Acta Numerica 28 (2019) 1–174. doi:10.1017/S0962492919000059.

[25] S. K. Paul, S. Sivaprasad, S. Dhar, S. Tarafder, Key issues in cyclic plastic deformation: Experimentation, Mechanics of Materials 43 (11) (2011) 705–720. doi:10.1016/j.mechmat.2011.07.011.

[26] D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions, Journal of Global Optimization 13 (4) (1998) 455–492. doi:10.1023/A:1008306431147.

[27] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proceedings of the IEEE 104 (1) (2016) 148–175. doi:10.1109/JPROC.2015.2494218.

[28] P. Feliot, J. Bect, E. Vazquez, A Bayesian approach to constrained single- and multi-objective optimization, Journal of Global Optimization 67 (1) (2017) 97–133. doi:10.1007/s10898-016-0427-3.

[29] C. E. Rasmussen, C. K. I. Williams, Gaussian processes for machine learning, The MIT Press, Cambridge, Massachusetts, 2006. doi:10.7551/mitpress/3206.001.0001.

[30] Y. Zhang, D. W. Apley, W. Chen, Bayesian optimization for materials design with mixed quantitative and qualitative variables, Scientific Reports 10 (1) (2020) 4924. doi:10.1038/s41598-020-60652-9.

[31] A. Mathern, O. S. Steinholtz, A. Sjöberg, M. Önnheim, K. Ek, R. Rempling, E. Gustavsson, M. Jirstrand, Multi-objective constrained Bayesian optimization for structural design, Structural and Multidisciplinary Optimization 63 (2) (2021) 689–701. doi:10.1007/s00158-020-02720-2.

[32] B. Do, M. Ohsaki, M. Yamakawa, Bayesian optimization for robust design of steel frames with joint and individual probabilistic constraints, Engineering Structures 245 (2021) 112859. doi:10.1016/j.engstruct.2021.112859.

[33] J. Kuhn, J. Spitz, P. Sonnweber-Ribic, M. Schneider, T. Böhlke, Identifying material parameters in crystal plasticity by Bayesian optimization, Optimization and Engineering (2021). doi:10.1007/s11081-021-09663-7.

[34] J. Lemaitre, J. L. Chaboche, Mechanics of solid materials, Cambridge University Press, 1994.

[35] E. Voce, The relationship between stress and strain for homogeneous

deformation, Journal of the Institute of Metals 74 (1948) 537–562.

[36] W. Prager, Recent developments in the mathematical theory of plasticity, Journal of Applied Physics 20 (3) (1949) 235–241. `doi:10.1063/1.1698348`.

[37] P. J. Armstrong, C. O. Frederick, A mathematical representation of the multiaxial Bauschinger effect, Report RD/B/N731, Berkeley, UK (1966).

[38] J. L. Chaboche, G. Rousselier, On the plastic and viscoplastic constitutive equations–Part I: Rules developed with internal variable concept, Journal of Pressure Vessel Technology 105 (2) (1983) 153–158. `doi:10.1115/1.3264257`.

[39] S. Yamada, Y. Jiao, A concise hysteretic model of structural steel considering the Bauschinger effect, International Journal of Steel Structures 16 (3) (2016) 671–683. `doi:10.1007/s13296-015-0134-9`.

[40] S. K. Azad, O. Hasançebi, An elitist self-adaptive step-size search for structural design optimization, Applied Soft Computing 19 (2014) 226–235. `doi:10.1016/j.asoc.2014.02.017`.

[41] A. Kaveh, A. Dadras, Structural damage identification using an enhanced thermal exchange optimization algorithm, Engineering Optimization 50 (3) (2018) 430–451. `doi:10.1080/0305215X.2017.1318872`.

[42] S. K. Azad, Monitored convergence curve: a new framework for metaheuristic structural optimization algorithms, Structural and Multidisciplinary Optimization 60 (2) (2019) 481–499. `doi:10.1007/s00158-019-02219-5`.

[43] A. Andrade-Campos, S. Thuillier, P. Pilvin, F. Teixeira-Dias, On the determination of material parameters for internal variable thermoelastic-viscoplastic constitutive models, International Journal of Plasticity 23 (8) (2007) 1349–1379. `doi:10.1016/j.ijplas.2006.09.002`.

[44] A. Afzal, K.-Y. Kim, J.-W. Seo, Effects of Latin hypercube sampling on surrogate modeling and optimization, International Journal of Fluid Machinery and Systems 10 (3) (2017) 240–253. `doi:10.5293/IJFMS.2017.10.3.240`.

[45] V. Picheny, T. Wagner, D. Ginsbourger, A benchmark of Kriging-based infill criteria for noisy optimization, Structural and Multidisciplinary Optimization 48 (3) (2013) 607–626. `doi:10.1007/s00158-013-0919-4`.

[46] D. Huang, T. T. Allen, W. I. Notz, N. Zeng, Global optimization of stochastic black-box systems via sequential Kriging meta-models, Journal of Global Optimization 34 (3) (2006) 441–466. `doi:10.1007/s10898-005-2454-3`.

[47] D. Systèmes, Abaqus user's manual ver. 2017 (2017).