Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization

Published article available online: Mechanical Systems and Signal Processing 114 (2019) 328–345

Daniel T. Bartilson[a,*], Jinwoo Jang[b], Andrew W. Smyth[a]

[a] *Department of Civil Engineering and Engineering Mechanics, Columbia University, NY 10027, USA*
[b] *Department of Civil, Environmental & Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA*

Abstract

Finite element model updating seeks to modify a structural model to reduce discrepancies between predicted and measured data, often from vibration studies. An updated model provides more accurate prediction of structural behavior in future analyses. Sensitivity-based parameter clustering and regularization are two techniques used to improve model updating solutions, particularly for high-dimensional parameter spaces and ill-posed updating problems. In this paper, a novel parameter clustering scheme is proposed which considers the structure of the objective function to facilitate simultaneous updating of disparate data, such as natural frequencies and mode shapes. In a small-scale updating example with simulated data, the proposed clustering scheme is shown to provide moderate to excellent improvement over existing parameter clustering methods, depending on the accuracy of initial model. A full-scale updating example on a large suspension bridge shows similar improvement using the proposed parametrization scheme. Levenberg–Marquardt minimization with Bayesian regularization is also implemented, providing an optimal regularized solution and insight into parametrization efficiency.

Keywords: Finite element model updating, Sensitivity-based model updating, Sensitivity-based clustering, Parametrization, Regularization

1. Introduction

Modern structural analysis generally depends on finite element (FE) models to predict dynamic behavior and understand the current state of a system. Though these models are often developed from detailed design drawings, discrepancies always exist between measured (observed) and model-output behavior [1]. Typical sources of discrepancy are model idealization, FE discretization errors, and uncertain model parameters such as material properties, section properties, geometry, and boundary conditions [1, 2]. Discrepancies indicate that a model cannot reliably predict the behavior of its corresponding physical structure, limiting the utility of the model for future analysis.

Model updating is the process which seeks to reduce discrepancies between measured data and model-output data by adjusting parameters of an FE model [1–3]. Model updating has been successfully applied to a wide variety of aerospace, mechanical, and civil structures. Examples include a helicopter airframe [2, 4], an aluminum space-frame [5], a prestressed single-span highway bridge [6], a prestressed multi-span highway bridge [7], a concrete-filled steel tubular arch bridge [8], an actively-damped high-rise structure [9], and a residential reinforced concrete frame [10].

Model updating techniques may be divided into two categories: stochastic methods and deterministic methods. Stochastic methods estimate probability distribution functions for parameters through repeated sampling in the parameter space, giving insights into uncertainties of updating parameters and predicted

*Corresponding author
Email addresses: dtb2121@columbia.edu (Daniel T. Bartilson), jangj@fau.edu (Jinwoo Jang), smyth@civil.columbia.edu (Andrew W. Smyth)
URL: http://www.columbia.edu/cu/civileng/smyth/ (Andrew W. Smyth)

responses. However, probabilistic methods are orders-of-magnitude more computationally expensive than deterministic methods. An excellent review of probabilistic model updating can be found in [11].

Deterministic model updating produces a unique optimal solution and encompasses direct methods and iterative methods [3]. Direct methods, also known as global or one-step methods, directly modify the FE mass and stiffness matrices to reduce discrepancy. Unfortunately, the resulting mass and stiffness matrices often lose physical relevance, appropriate connectivity, and sparseness. Iterative methods, also known as local methods, incrementally modify the FE model parameters and have the capability to be physically relevant while maintaining desirable properties of the FE matrices. Of course, as these schemes generally involve minimizing a non-linear function, they are possibly subject to convergence problems.

Among iterative methods, the sensitivity method [2] is one of the most intuitive and popular techniques for model updating. The sensitivity method approaches model updating as a non-linear least-squares minimization problem which is solved by iterations of linear approximations. The objective function is a sum of squared differences between measured and model-output data, making it easy to incorporate various data. The use of linear approximations also makes this method physically-intuitive and efficient, as the Jacobian matrix is directly relatable to model parameter sensitivities. However, the sensitivity method often results in an ill-posed problem, necessitating a reduction in the number of updating parameters and/or the inclusion of a side-constraint.

Shahverdi *et al* [4] presented sensitivity-based parameter clustering as a viable method for reducing the number of updating parameters. By observing the sensitivities of model outputs to changes in model parameters, sensitivity-based parameter clustering generates clusters of model parameters which have similar effects on targeted model outputs. Then, each cluster of model parameters is updated by a single parameter. This gives a reduced-order model, generally with better condition, while retaining the physical relevance of clustered model parameters. This technique was successfully applied to the updating of a helicopter airframe [4]. Jang and Smyth [12, 13] applied this method for the updating of a large-scale suspension bridge.

Regularization is another technique used to solve ill-posed and noisy problems which often occur in FE model updating [2, 14–16]. Generally, regularization adds equations which help constrain the updating solution. This can help produce a unique solution to an underdetermined problem (fewer measurements than parameters), though this situation should be avoided. Regularization is often used to give a minimum-norm solution, but it may also be used to enforce user-specified constraints between parameters [2].

While sensitivity-based parameter clustering is very promising, it is difficult to utilize disparate sources of data, such as natural frequencies and mode shapes, due to differences in scale. Previous work with parameter clustering only used one type of data [1, 4], or used only natural frequency sensitivities for clustering despite the inclusion of mode shapes in the objective function [12, 13]. To alleviate scaling issues during parameter clustering, it is necessary to develop a weighting technique which is efficient and reflective of the problem structure. The presented research details an objective-consistent weighting technique based on the residual. This paper also implements Bayesian regularization [17, 18] in model updating, which gives a statistically optimal regularized solution. Bayesian regularization also provides insight into the effective number of updating parameters, related to model selection and efficiency.

The paper begins with the definition of residual between measurements and corresponding model outputs, along with analytical sensitivities of model outputs to model parameters (Section 2). Model parametrization, clustering, and the objective-consistent weighting scheme are discussed in Section 3. The Levenberg–Marquardt minimization method, with the accompanying Bayesian regularization technique, are detailed in Section 4. Two model updating exercises are then performed to exhibit the efficiency of the objective-consistent clustering scheme for simultaneous updating of natural frequency and mode shape data. The first exercise uses a small-scale 2-dimensional truss with simulated measurements (Section 5), while the second uses a full-scale large suspension bridge with real data (Section 6). The findings are then discussed and concluding remarks are made in Section 7.

2. Residual definition and analytical sensitivity of model parameters

The sensitivity method for FE model updating [2] begins with the definition of a discrepancy, or residual, to be minimized by modifying a set of updating parameters. Traditionally, the residual $\boldsymbol{r}$ is defined as the difference between the column vector of m measured outputs $\tilde{\boldsymbol{z}}$ and the column vector of m analytical model

outputs $\boldsymbol{z}(\boldsymbol{\theta})$ which is a function of the p updating parameters $\boldsymbol{\theta}$. The relationship between $\boldsymbol{r}$ and $\boldsymbol{\theta}$ is generally non-linear, but can be linearized by truncating the Taylor series after the linear term:

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \approx \boldsymbol{r}(\boldsymbol{\theta}_i) + \boldsymbol{J}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i) \tag{1}$$

At iteration i, $\boldsymbol{\theta}_i$ is the updating parameter vector and $\boldsymbol{J}_i \in \mathbb{R}^{m \times p}$ is the Jacobian matrix of $\boldsymbol{r}$ with respect to $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta}_i$:

$$\boldsymbol{J}_i = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i} \tag{2}$$

The evaluation of $\boldsymbol{J}$, also known as the sensitivity matrix, forms the basis of Gauss–Newton updating methods [19].

In any model updating problem, one must choose data to set as a target for updating. Common choices are natural frequencies or eigenvalues, mode shapes, and frequency-response functions. For the purposes of this paper, the chosen residual is a concatenation of the natural frequency residual vector $\boldsymbol{r}_f$ (Section 2.1) and the mode shape residual vector $\boldsymbol{r}_s$ (Section 2.2), leading to $\boldsymbol{r} = [\boldsymbol{r}_f^T \;\; \boldsymbol{r}_s^T]^T$. The weighted sum-of-squared residual E_r can be written

$$E_r = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_f^T & \boldsymbol{r}_s^T \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_r^f & \\ & \boldsymbol{W}_r^s \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_f \\ \boldsymbol{r}_s \end{bmatrix} \tag{3}$$

$$= \underbrace{\boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f}_{E_r^f} + \underbrace{\boldsymbol{r}_s^T \boldsymbol{W}_r^s \boldsymbol{r}_s}_{E_r^s} \tag{4}$$

which includes a term for the weighted sum-of-squared frequency residual (E_r^f) and a term for the weighted sum-of-squared mode shape residual (E_r^s). $\boldsymbol{W}_r$ is the residual weighting matrix, which is discussed in more detail in Section 4, and reflects the importance of each residual term.

The natural frequencies and mode shapes are assumed to be obtained from a model of an undamped structure, giving only real frequencies and mode shapes. A generic FE model of an undamped structure with N degrees-of-freedom (DoFs) consists of a stiffness matrix $\boldsymbol{K}$ and a mass matrix $\boldsymbol{M}$, where $\boldsymbol{K}, \boldsymbol{M} \in \mathbb{R}^{N \times N}$. These matrices are used to solve the generalized eigenvalue problem $\boldsymbol{K}\boldsymbol{\phi}_j = \omega_j^2 \boldsymbol{M}\boldsymbol{\phi}_j \quad \forall j = 1, ..., N$, where ω_j represents the j^{th} angular natural frequency with an equivalent temporal natural frequency f_j. Each corresponding mode shape $\boldsymbol{\phi}_j$ is mass-normalized such that $\boldsymbol{\phi}_j^T \boldsymbol{M} \boldsymbol{\phi}_j = 1$.

For computing $\boldsymbol{J}$, note that each column is a concatenation of frequency residual gradients and mode shape residual gradients, i.e. $\partial \boldsymbol{r}/\partial\theta_k = [(\partial \boldsymbol{r}_f/\partial\theta_k)^T \; (\partial \boldsymbol{r}_s/\partial\theta_k)^T]^T$. Analytical methods for calculating the sensitivities are given in Sections 2.1 and 2.2. Alternatively (or for the purposes of verification), sensitivities can be estimated numerically using finite differences.

2.1. Undamped frequency residual

The frequency residual vector $\boldsymbol{r}_f$ is defined as the difference between the column vector of l measured natural frequencies $\tilde{\boldsymbol{f}}$ and the corresponding column vector of l model-output natural frequencies $\boldsymbol{f}(\boldsymbol{\theta})$:

$$\boldsymbol{r}_f = \tilde{\boldsymbol{z}}_f - \boldsymbol{z}_f(\boldsymbol{\theta}) = \tilde{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{\theta}) \tag{5}$$

It is essential to perform mode pairing [3] to ensure that measured and model-output data refer to the same modes when calculating residual. When $\boldsymbol{W}_r^f$ is diagonal, the weighted sum-of-squared frequency residual E_r^f is

$$E_r^f = \boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j(\boldsymbol{\theta}))^2 \tag{6}$$

The frequency residual sensitivity vector (for use in $\boldsymbol{J}$) is

$$\frac{\partial \boldsymbol{r}_f}{\partial \theta_k} = -\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \theta_k} \tag{7}$$

The sensitivity of model natural frequency f_j to a change in updating parameter θ_k can be calculated analytically using the results of [20, 21]

$$2\pi\frac{\partial f_j}{\partial \theta_k} = \frac{\partial \omega_j}{\partial \theta_k} = \frac{1}{2\omega_j}\boldsymbol{\phi}_j^T\left[\frac{\partial \boldsymbol{K}}{\partial \theta_k} - \omega_j^2\frac{\partial \boldsymbol{M}}{\partial \theta_k}\right]\boldsymbol{\phi}_j \tag{8}$$

where $\boldsymbol{\phi}_j$ is the j^{th} mass-normalized mode shape, i.e. $\boldsymbol{\phi}_j^T\boldsymbol{M}\boldsymbol{\phi}_j = 1$.

2.2. Undamped mode shape residual

The mode shape residual $\boldsymbol{r}_s$ is defined as the difference between the concatenated set of l measured mode shapes $\tilde{\boldsymbol{z}}_s$ and the corresponding concatenated model-output mode shapes $\boldsymbol{z}_s(\boldsymbol{\theta})$:

$$\boldsymbol{r}_s = \tilde{\boldsymbol{z}}_s - \boldsymbol{z}_s(\boldsymbol{\theta}) \tag{9}$$

The concatenated mode shape measurement vector is written as $\tilde{\boldsymbol{z}}_s = [\tilde{\boldsymbol{v}}_1^T \ \cdots \ \tilde{\boldsymbol{v}}_l^T]^T$ where $\tilde{\boldsymbol{v}}_j$ is the j^{th} unit-normalized measured mode shape, i.e. $\tilde{\boldsymbol{v}}_j = \tilde{\boldsymbol{\phi}}_j/(\tilde{\boldsymbol{\phi}}_j^T\tilde{\boldsymbol{\phi}}_j)^{1/2}$. Similarly, the concatenated model-output mode shapes are $\boldsymbol{z}_s(\boldsymbol{\theta}) = [\mu_1\boldsymbol{v}_1^T(\boldsymbol{\theta}) \ \cdots \ \mu_l\boldsymbol{v}_l^T(\boldsymbol{\theta})]^T$, where $\boldsymbol{v}_j(\boldsymbol{\theta})$ is the j^{th} unit-normalized model-output mode shape. The model-output modes should be sorted according to the mode pairing results. μ_j is the modal scale factor between corresponding measured and model-output mode shapes, and minimizes the difference between $\tilde{\boldsymbol{v}}_j$ and $\boldsymbol{v}_j(\boldsymbol{\theta})$ in the least-squares sense [22]:

$$\mu_j = \tilde{\boldsymbol{v}}_j^T\boldsymbol{v}_j(\boldsymbol{\theta}) \tag{10}$$

Note that $\tilde{\boldsymbol{v}}_j$ and $\boldsymbol{v}_j(\boldsymbol{\theta})$ must be measured at the same n points on the physical and model structures. This means that $\boldsymbol{v}_j(\boldsymbol{\theta}), \tilde{\boldsymbol{v}}_j \in \mathbb{R}^n \quad \forall j = 1, ..., l$ and thus $\boldsymbol{r}_s \in \mathbb{R}^{nl}$.

Assuming that $\boldsymbol{W}_r^s$ is diagonal and can be decomposed into a scalar multiple of $\boldsymbol{I}_n$ for each mode ($[\boldsymbol{W}_r^s]_j = w_{rj}^s\boldsymbol{I}_n$), then the sum-of-squared mode shape residual E_r^s can be written

$$E_r^s = \boldsymbol{r}_s^T\boldsymbol{W}_r^s\boldsymbol{r}_s = \sum_{j=1}^{l} w_{rj}^s\boldsymbol{r}_{sj}^T\boldsymbol{r}_{sj} = \sum_{j=1}^{l} w_{rj}^s\|\tilde{\boldsymbol{v}}_j - \mu_j\boldsymbol{v}_j(\boldsymbol{\theta})\|_2^2 \tag{11}$$

where $\boldsymbol{r}_{sj} = \tilde{\boldsymbol{v}}_j - \mu_j\boldsymbol{v}_j(\boldsymbol{\theta})$ is the residual for mode shape j and $\| \ \|_2$ is the L_2 norm. The Modal Assurance Criterion (MAC) [22] is defined as

$$\mathrm{MAC}(\boldsymbol{a}, \boldsymbol{b}) = \frac{(\boldsymbol{a}^T\boldsymbol{b})^2}{\boldsymbol{a}^T\boldsymbol{a}\cdot\boldsymbol{b}^T\boldsymbol{b}} \tag{12}$$

which allows Eq. (11) to be rewritten as

$$E_r^s = \sum_{j=1}^{l} w_{rj}^s\left[1 - \mathrm{MAC}(\tilde{\boldsymbol{v}}_j, \boldsymbol{v}_j(\boldsymbol{\theta}))\right] \tag{13}$$

The derivative of the residual for mode shape j, $\boldsymbol{r}_{sj}$, (for use in $\boldsymbol{J}$), is

$$\frac{\partial \boldsymbol{r}_{sj}}{\partial \theta_k} = -\frac{\partial(\mu_j\boldsymbol{v}_j)}{\partial \boldsymbol{v}_j}\frac{\partial \boldsymbol{v}_j}{\partial \theta_k} = -\left[\boldsymbol{v}_j\tilde{\boldsymbol{v}}_j^T + \mu_j\boldsymbol{I}\right]\frac{\partial \boldsymbol{v}_j}{\partial \theta_k} \tag{14}$$

The residual sensitivities for each mode are concatenated, such that $\partial\boldsymbol{r}_s/\partial\theta_k = [(\partial\boldsymbol{r}_{s1}/\partial\theta_k)^T \ \cdots \ (\partial\boldsymbol{r}_{sl}/\partial\theta_k)^T]^T$. where the sensitivity of the j^{th} mode shape can be analytically calculated by the results of [20, 21]:

$$\frac{\partial \boldsymbol{v}_j}{\partial \theta_k} = \sum_{h=1}^{H}\frac{a_{jkh}}{\boldsymbol{v}_h^T\boldsymbol{M}\boldsymbol{v}_h}\boldsymbol{v}_h \tag{15}$$

The upper limit of summation H is the number of dynamic modes used to estimate the sensitivities, where $H \leq N$. When $H = N$, the results are exact. The factors a_{jkh} are given by

$$a_{jkh} = \frac{1}{\omega_j^2 - \omega_h^2}\boldsymbol{v}_h^T\left[\frac{\partial \boldsymbol{K}}{\partial \theta_k} - \omega_j^2\frac{\partial \boldsymbol{M}}{\partial \theta_k}\right]\boldsymbol{v}_j \ ; \quad j \neq h \tag{16}$$

$$a_{jkj} = -\frac{1}{2}\boldsymbol{v}_j^T \left[\frac{\partial \boldsymbol{M}}{\partial \theta_k}\right] \boldsymbol{v}_j \tag{17}$$

Combining the results of Eqs. (6) and (13), the weighted sum-of-squared residual from Eq. (3) is

$$E_r = E_r^f + E_r^s = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j(\boldsymbol{\theta}))^2 + \sum_{j=1}^{l} w_{rj}^s \left[1 - \mathrm{MAC}(\tilde{\boldsymbol{v}}_j, \boldsymbol{v}_j(\boldsymbol{\theta}))\right] \tag{18}$$

Note that when $\boldsymbol{W}_r$ is equal to the reciprocal of the covariance of $\tilde{\boldsymbol{z}}$ (for uncorrelated measurements), then Eq. (18) implies that each frequency has its own measurement variance, while the components of each mode shape vector have equal variance to other components of the same mode shape.

3. Selection of updating parameters

3.1. Model parametrization

When updating an FE model, the overarching approach is to modify a set of model parameters $\boldsymbol{x}$ to reduce error. Following the parametrization in [12, 13], the j^{th} element of the modified physical parameter vector $\boldsymbol{x}$ is defined by:

$$x_j = x_j^0(1 - \delta_j) \tag{19}$$

where x_j^0 is the initial value of x_j and δ_j is the physical parameter modification term. Since $\boldsymbol{x}$ may contain disparate parameters, such as stiffnesses and densities, updating $\boldsymbol{x}$ directly may result in poor scaling. The use of $\boldsymbol{\delta}$ results in a far better scaled set of model parameters. Perhaps the most natural implementation of this parametrization in an FE model is to modify each element or substructure e (out of a total number d) stiffness ($\boldsymbol{K}_e$) and mass ($\boldsymbol{M}_e$) matrices prior to summation into the global stiffness ($\boldsymbol{K}$) and mass ($\boldsymbol{M}$) matrices:

$$\boldsymbol{K} = \sum_{e=1}^{d} \boldsymbol{K}_e(1 - \delta_e^k) = \boldsymbol{K}_0 - \sum_{e=1}^{d} \boldsymbol{K}_e \delta_e^k \tag{20}$$

$$\boldsymbol{M} = \sum_{e=1}^{d} \boldsymbol{M}_e(1 - \delta_e^m) = \boldsymbol{M}_0 - \sum_{e=1}^{d} \boldsymbol{M}_e \delta_e^m \tag{21}$$

$\boldsymbol{K}_0$ and $\boldsymbol{M}_0$ are the initial global stiffness and mass matrices prior to modification or updating. δ_e^k and δ_e^m are the stiffness and mass physical parameter modifications, respectively, for element or substructure e. Note that $\boldsymbol{\delta}$ for this problem may be viewed as a concatenation of $\boldsymbol{\delta}^k$ and $\boldsymbol{\delta}^m$, therefore $\boldsymbol{\delta} \in \mathbb{R}^{2d}$. In this form, calculating partial derivatives for use in analytical sensitivity calculations is simple, as $\partial \boldsymbol{K}/\partial \delta_e^k = -\boldsymbol{K}_e$ and $\partial \boldsymbol{M}/\partial \delta_e^m = -\boldsymbol{M}_e$.

While FE model updating based on $\boldsymbol{\delta}$ is natural and straightforward, it can quickly become intractable or ill-posed in an FE model with thousands of elements, possibly with several parameters per element. Thus, a smaller set of updating parameters are sought through cluster analysis. The set of FE model parameters $\boldsymbol{\delta}$ are not directly updated, but instead $\boldsymbol{\theta}$ are updated and $\boldsymbol{\delta} = g(\boldsymbol{\theta})$. In general, it is desired that the clustering is "hard" (i.e. each physical parameter belongs to one and only one cluster). Thus, g may be written as a linear transformation:

$$\boldsymbol{\delta} = g(\boldsymbol{\theta}) = \boldsymbol{C}\boldsymbol{\theta} \tag{22}$$

where $\boldsymbol{C} \in \mathbb{R}^{2d \times p}$ such that

$$C_{jk} = \mathbb{I}_k(j) = \begin{cases} 1 & \delta_j \text{ is included in cluster } k, \text{ updated by } \theta_k \\ 0 & \text{else} \end{cases} \tag{23}$$

where $\mathbb{I}_k(j)$ is the indicator function.

The FE model elements can be summed into stiffness and mass substructures, $\boldsymbol{K}_j^\star$ and $\boldsymbol{M}_h^\star$, respectively, using the indicator function

$$\boldsymbol{K}_j^\star = \sum_{e=1}^{d} \mathbb{I}_j(e)\boldsymbol{K}_e; \qquad \boldsymbol{M}_h^\star = \sum_{e=1}^{d} \mathbb{I}_h(e)\boldsymbol{M}_e \tag{24}$$

which are updated by a corresponding θ_j^k and θ_h^m:

$$\boldsymbol{K} = \sum_{j=1}^{p^k} \boldsymbol{K}_j^\star(1-\theta_j^k) = \boldsymbol{K}_0 - \sum_{j=1}^{p^k} \boldsymbol{K}_j^\star\theta_j^k \tag{25}$$

$$\boldsymbol{M} = \sum_{h=1}^{p^m} \boldsymbol{M}_h^\star(1-\theta_h^m) = \boldsymbol{M}_0 - \sum_{h=1}^{p^m} \boldsymbol{M}_h^\star\theta_h^m \tag{26}$$

Similar to $\boldsymbol{\delta}$, $\boldsymbol{\theta}$ is a concatenation of stiffness updating parameters $\boldsymbol{\theta}^k \in \mathbb{R}^{p^k}$ and mass updating parameters $\boldsymbol{\theta}^m \in \mathbb{R}^{p^m}$, with the total number of updating parameters $p = p^k + p^m$.

3.2. Sensitivity-based cluster analysis of updating parameters

Selection of updating parameters is an integral step in the model updating process, where three general conditions should be satisfied: (1) parameters should be chosen to avoid an ill-posed problem, (2) the choice of parameters should reflect the objective of reducing modeling error, and (3) the model-outputs should be sensitive to the selected parameters [3]. Satisfying these three conditions is non-trivial. Ill-conditioning tends to occur in larger problems [23], in which columns of the Jacobian are increasingly likely to exhibit linear dependence [4]. One solution is to use a subset of model parameters, a so-called 'subset selection' [24], which is chosen to be linearly-independent, sensitive, and representative of the residual.

An alternative approach retains all of the model parameters, but updates groups or clusters of model parameters, thereby reducing model order. This idea was presented in [25] as the 'best subspace approach', where clusters were chosen based on the angle between subspaces. This idea was improved upon and validated in Shahverdi *et al*'s work [4], where clusters were selected based on hierarchical cluster analysis of parameter sensitivities. A major advantage of sensitivity-based clustering, as opposed to subset selection methods, is that clusters have a physical relevance: model parameters in the same cluster exhibit similar effect on model outputs. In the examples presented in [2, 4, 12, 13], this meant that clustered parameters exhibited similar effects on the model natural frequencies.

Following the work of Shahverdi *et al* [4] and Jang and Smyth [12, 13], hierarchical clustering was selected for use as the grouping method in this paper. For a brief overview of hierarchical clustering, see [4], while a more thorough coverage is provided in [26]. When using hierarchical clustering, it is necessary to select a distance measure and a linkage method. The selected distance measure for this study is cosine distance, defined between column vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ as

$$d_{cos}(\boldsymbol{a},\boldsymbol{b}) = 1 - \frac{\boldsymbol{a}^T\boldsymbol{b}}{\sqrt{\boldsymbol{a}^T\boldsymbol{a}\cdot\boldsymbol{b}^T\boldsymbol{b}}} \tag{27}$$

Cosine distance is a measure of dissimilarity in shape between two vectors and does not consider the relative magnitudes. Furthermore, one can note that similar vectors are near-parallel. When near-parallel columns of the Jacobian are reduced through clustering, the condition of the problem tends to improve [4].

With a chosen clustering method and distance measure, the only remaining choice is the linkage method, which determines how vectors are combined to form clusters. The chosen method in this paper is the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [27]. This is an agglomerative method which combines the two nearest clusters at each step, evaluating the distance between the unweighted means of each cluster. This agglomerative process begins at the 'branches' of the dendrogram with single elements, then combines clusters at each step until it reaches the 'root' with only a single cluster encompassing all

elements. When the dendrogram is built, the user is able to retrieve clusters by choosing to 'cut' the tree at a certain distance level, guaranteeing that each cluster is more than some chosen distance from any other. Alternatively, the user can input a desired number of clusters and the tree will be 'cut' to yield the desired number.

3.3. Objective-consistent scaling of cosine distance

While the use of cosine distance for sensitivity-based clustering was an excellent development, it is susceptible to becoming a skewed metric. When evaluating d_{cos} for two sensitivity vectors, skewing occurs when the sensitivity vectors comprise disparate sources, such as natural frequency and mode shape sensitivities. If the natural frequency sensitivities are considerably larger in magnitude than the mode shape sensitivities, the natural frequency sensitivities tend to dominate the resulting distance calculation. Indeed, there is no natural scale for mode shape sensitivities, as the mode shapes themselves can have arbitrary normalization. To mitigate the difference in magnitude for different sources of sensitivity, it is necessary to scale the residual sensitivity vectors.

To the authors' knowledge, no systematic method has been proposed for incorporating more than one type of data in cluster selection for FE model updating. Previous uses of parameter clustering in model updating only updated natural frequencies [2, 4], or didn't incorporate mode shapes into cluster selection despite their inclusion in the objective function [12]. The proposed scaling for the residual sensitivity vectors is found by observing the gradient of the sum-of-squared residual from Eq. (3):

$$\frac{\partial E_r}{\partial \delta_j} = 2\boldsymbol{r}^T \boldsymbol{W}_r \frac{\partial \boldsymbol{r}}{\partial \delta_j} \tag{28}$$

where the sensitivity vector $\partial \boldsymbol{r}/\partial \delta_j$ is scaled by the weighted residual, $\boldsymbol{r}^T \boldsymbol{W}_r$. Given a diagonal $\boldsymbol{W}_r$, then each term of the sensitivity vector is multiplied by the corresponding residual and weighting term. Decomposing the inner product of $\boldsymbol{r}^T \boldsymbol{W}_r$ and $\partial \boldsymbol{r}/\partial \delta_j$ results in a sensitivity vector which is scaled by a matrix $\boldsymbol{W}_{\text{OC}}$:

$$\boldsymbol{s}_j = \boldsymbol{W}_{\text{OC}} \frac{\partial \boldsymbol{r}}{\partial \delta_j}; \qquad \boldsymbol{W}_{\text{OC}} = \text{diag}(\boldsymbol{r}^T \boldsymbol{W}_r) \tag{29}$$

Accordingly, the calculation of the cosine distance between scaled sensitivities can be written as

$$d_{cos}(\boldsymbol{s}_j, \boldsymbol{s}_k) = 1 - \left(\frac{\partial \boldsymbol{r}^T}{\partial \delta_j} \boldsymbol{W}_{\text{OC}}^2 \frac{\partial \boldsymbol{r}}{\partial \delta_k} \right) \Bigg/ \sqrt{\left(\frac{\partial \boldsymbol{r}^T}{\partial \delta_j} \boldsymbol{W}_{\text{OC}}^2 \frac{\partial \boldsymbol{r}}{\partial \delta_j} \right) \cdot \left(\frac{\partial \boldsymbol{r}^T}{\partial \delta_k} \boldsymbol{W}_{\text{OC}}^2 \frac{\partial \boldsymbol{r}}{\partial \delta_k} \right)} \tag{30}$$

which clearly indicates the role of the objective-consistent (OC) sensitivity scaling matrix, $\boldsymbol{W}_{\text{OC}}$.

This choice of scaling based on weighted residual confers several benefits. First, the cosine distance and corresponding clustering give more weight to sensitivity vector components which have a large corresponding residual, and less weight to sensitivity components with a small corresponding residual. This ameliorates scenarios where, say, a natural frequency sensitivity is large, but that particular natural frequency is already well-matched and has a residual component near zero. Second, the clustering reflects the scaling of the weighting matrix $\boldsymbol{W}_r$ and its effect on the objective value. Finally, this does not compromise the physical relevance of the clustering. Clusters are still selected based on the similarity of residual sensitivity vectors (which is related to model behavior), while the weighting only dictates the importance of each term in the cosine distance calculation.

4. Levenberg–Marquardt minimization with Bayesian regularization

With a well-defined residual and parametrized model, model updating proceeds with selection of an error minimization algorithm. The objective function is first modified to include regularization, then the minimization procedure is described. The regularized objective function, F, includes both a residual term E_r (see Eq. (3)) as well as an updating parameter size (penalty) term E_θ:

$$F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta = \beta \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} + \alpha \boldsymbol{\theta}^T \boldsymbol{W}_\theta \boldsymbol{\theta} \tag{31}$$

where $\{E_r, E_\theta\}$ are analogous to $\{E_D, E_W\}$ in [17, 18]. α and β are scalar regularization parameters which influence the relative importance of E_r and E_θ during minimization. The process for calculating $\{\alpha, \beta\}$ is covered in Section 4.1.

The residual weighting matrix $\boldsymbol{W}_r$ and the parameter weighting matrix $\boldsymbol{W}_\theta$ must generally be symmetric positive semi-definite and should reflect the uncertainty in $\boldsymbol{r}$ and $\boldsymbol{\theta}$, respectively. The optimum choices (in a Bayesian sense) for $\boldsymbol{W}_r$ and $\boldsymbol{W}_\theta$ are the inverse of the measurement covariance matrix and the parameter covariance matrix, respectively [28–30]. The measurement covariance matrix is diagonal for the case of uncorrelated uncertainties, and is usually simple to estimate from data. The parameter covariance matrix is difficult to estimate, though [31] provides a method for relating $\boldsymbol{W}_\theta$ to the sensitivity matrix. The work presented in this paper uses a simple choice of $\boldsymbol{W}_\theta = \boldsymbol{I}$, similar to Tikhonov regularization [32].

The residual vector is a column of m elements given by

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \tag{32}$$

where $\tilde{\boldsymbol{z}}$ represents a vector of m measurements and $\boldsymbol{z}(\boldsymbol{\theta})$ represents a vector of m model outputs, given the column vector of p updating parameters $\boldsymbol{\theta}$. The parameter values and residual at iteration i are written $\boldsymbol{\theta}_i$ and $\boldsymbol{r}_i = \boldsymbol{r}(\boldsymbol{\theta}_i)$, respectively. At each iteration, the model parameters are updated such that

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i \tag{33}$$

To begin the solution algorithm, the residual at $\boldsymbol{\theta}_{i+1}$ is approximated using the truncated Taylor series of Eq. (1), giving

$$\boldsymbol{r}(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i) \approx \boldsymbol{r}_i + \boldsymbol{J}_i\,\Delta\boldsymbol{\theta}_i \tag{34}$$

where $\boldsymbol{J}_i$ is the Jacobian of $\boldsymbol{r}$ evaluated at $\boldsymbol{\theta}_i$, as in Eq. (2). Using this estimate for $\boldsymbol{r}_{i+1}$ in Eq. (31) and minimizing $F(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i)$ with respect to $\Delta\boldsymbol{\theta}_i$ yields

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i]^{-1}[\beta\boldsymbol{J}_i^T\boldsymbol{W}_r\boldsymbol{r}_i + \alpha\boldsymbol{W}_\theta\boldsymbol{\theta}_i] \tag{35}$$

which is the Gauss–Newton algorithm [19]. The Hessian at iteration i, $\boldsymbol{H}_i$, is estimated by the Gauss–Newton (GN) approximation

$$\boldsymbol{H}_i = \nabla\nabla F \approx 2[\beta\boldsymbol{J}_i^T\boldsymbol{W}_r\boldsymbol{J}_i + \alpha\boldsymbol{W}_\theta] \tag{36}$$

Eq. (35) can be improved into the more robust Levenberg–Marquardt (LM) algorithm [33, 34], which modifies the Hessian with a scalar damping term λ (unrelated to damping in mechanical vibrations), giving a trust-region solution:

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda\boldsymbol{I}]^{-1}[\beta\boldsymbol{J}_i^T\boldsymbol{W}_r\boldsymbol{r}_i + \alpha\boldsymbol{W}_\theta\boldsymbol{\theta}_i] \tag{37}$$

As $\lambda \to 0$, the LM algorithm becomes the GN algorithm, while $\lambda \to \infty$ leads to the gradient-descent algorithm. λ is adjusted using the multiplication scheme originally described by Marquardt [34], such that if the current value of λ results in $\Delta\boldsymbol{\theta}_i$ with $F_{i+1} < F_i$, λ is divided by a factor v for the next iteration. Otherwise λ is multiplied a factor v, then $\Delta\boldsymbol{\theta}_i$, F_{i+1} are recomputed until $F_{i+1} < F_i$ or convergence. A reasonable set of initial values are $\lambda_0 = 0.01$ and $v = 10$, which fulfills the requirement that $\lambda \geq 0$.

4.1. Bayesian selection of regularization parameters

The regularization parameters α and β are used to produce a model updating solution that balances parameter values (E_θ) and residual (E_r). Within the Bayesian framework, α and β are treated as random variables. If α and β are known, the posterior probability of the model parameters $\boldsymbol{\theta}$ is given by Bayes' rule:

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M})P(\boldsymbol{\theta}|\alpha, \mathcal{M})}{P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M})} \tag{38}$$

where $\tilde{\boldsymbol{z}}$ are the measurements and $\mathcal{M}$ represents the chosen parametrization or model. $P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M})$ is the likelihood of the measurements given $\boldsymbol{\theta}$, $P(\boldsymbol{\theta}|\alpha, \mathcal{M})$ is the prior density of $\boldsymbol{\theta}$, and $P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M})$ is a

normalization factor corresponding to the evidence for α and β. If it is assumed that noise in $\tilde{\boldsymbol{z}}$ is Gaussian, and that the prior distribution of $\boldsymbol{\theta}$ is also Gaussian, then the likelihood and the prior are

$$P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M}) = \frac{e^{-\beta E_r}}{Z_{\tilde{z}}} \; ; \qquad P(\boldsymbol{\theta}|\alpha,\mathcal{M}) = \frac{e^{-\alpha E_\theta}}{Z_\theta} \tag{39}$$

where the normalization factors are $Z_{\tilde{z}}(\beta) = (\pi/\beta)^{m/2}/\det(\boldsymbol{W}_r)^{1/2}$ and $Z_\theta(\alpha) = (\pi/\alpha)^{p/2}/\det(\boldsymbol{W}_\theta)^{1/2}$. Thus, the posterior is

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}},\alpha,\beta,\mathcal{M}) = \frac{e^{-F(\boldsymbol{\theta})}}{Z_F(\alpha,\beta)} \tag{40}$$

where $Z_F(\alpha,\beta)$ is a normalization factor which can be estimated by using a truncated Taylor series expansion of F from Eq. (31) about the local minimum point $\boldsymbol{\theta}_{\mathrm{MP}}$, which corresponds to the maximum probability point of the posterior distribution. This approximation gives

$$Z_F(\alpha,\beta) \approx e^{-F(\boldsymbol{\theta}_{\mathrm{MP}})}(2\pi)^{p/2}/\det(\boldsymbol{H}_{\mathrm{MP}})^{1/2} \tag{41}$$

where $\boldsymbol{H}_{\mathrm{MP}} = \boldsymbol{H}(\boldsymbol{\theta}_{\mathrm{MP}})$ is the Hessian estimated from Eq. (36). Rewriting the posterior in Eq. (38) to find the evidence and substituting the results of Eqs. (39) and (40) gives

$$P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M}) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M})P(\boldsymbol{\theta}|\alpha,\mathcal{M})}{P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}},\alpha,\beta,\mathcal{M})} = \frac{Z_F(\alpha,\beta)}{Z_{\tilde{z}}(\beta)Z_\theta(\alpha)} \tag{42}$$

Substituting the expressions for the normalization terms and taking the logarithm of the evidence yields

$$\log P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M}) = -\beta E_r^{\mathrm{MP}} - \alpha E_\theta^{\mathrm{MP}} - \frac{1}{2}\log\det(\boldsymbol{H}_{\mathrm{MP}}) + \frac{m}{2}\log\beta + \frac{p}{2}\log\alpha + c \tag{43}$$

Note that c is not a function of α or β. Maximizing the log-evidence with respect to α and β yields

$$\alpha = \frac{\gamma}{2E_\theta^{\mathrm{MP}}} \qquad \beta = \frac{m-\gamma}{2E_r^{\mathrm{MP}}} \tag{44}$$

where γ is the number of effective updating parameters, ranging from 0 to the number of parameters p:

$$\gamma = p - 2\alpha\,\mathrm{tr}(\boldsymbol{H}_{\mathrm{MP}}^{-1}\boldsymbol{W}_\theta) \tag{45}$$

γ provides an interesting insight into model efficiency and model selection [18]. If γ is close to p, then the problem might be 'saturated' and need more parameters to improve the solution. If increasing p results in the same γ as before, then the original p was sufficient (without changing the parameter selection scheme). $\gamma \ll p$ implies the solution is not sensitive to most of the chosen parameters, suggesting that p is too large or the selected parameters are inefficient.

Using knowledge of the relationships for α, β, and γ, the Levenberg–Marquardt minimization algorithm can then be modified to include iterative recalculation of the Bayesian regularization parameters, as done in [18]. The resulting pseudocode is provided in Algorithm 1. Note that γ is recalculated in each iteration, but is only meaningful at a local minimum point (i.e. at a converged solution).

5. Model updating of a small-scale FE model: 29-element truss

5.1. Model description and cluster analysis

The first exercise of the proposed model updating scheme is the 29-element, 28-DoF, 2-dimensional truss depicted in Fig. 1. This truss was modified from the 29-DoF truss presented in Papadimitriou *et al* [35]. The boundary conditions were changed to pin-pin (from pin-roller), thus making the structure symmetric and statically indeterminate. All truss elements utilized identical material properties, with Young's modulus (E) of 200 GPa, mass density (ρ) of 7800 kg/m^3, and area (A) 0.25 m^3. These values were essentially arbitrary, being chosen only to have a rough physical relevance.

Algorithm 1: Levenberg–Marquardt minimization with Bayesian regularization

Input: Objective function $F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta$ to be minimized
Output: Optimal parameters $\boldsymbol{\theta}_{\text{MP}}$, effective number of parameters γ
initialization: Set $\boldsymbol{\theta}_0$, $\alpha = 0.1$, $\beta = 1$, $\lambda = 0.01$, $v = 10$, $i = 0$;
while *not converged* **do**
  Compute $\boldsymbol{r}_i$, $\boldsymbol{J}_i$, and $\boldsymbol{H}_i = 2[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta]$;
  Compute parameter update $\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda \boldsymbol{I}]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i]$;
  Update $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i$;
  if $F(\boldsymbol{\theta}_{i+1}) > F(\boldsymbol{\theta}_i)$ **then**
    $\lambda \leftarrow \lambda \cdot v$;
    Go back to parameter update computation step ;
  else
    $\lambda \leftarrow \lambda / v$;
  end
  Compute effective number of parameters $\gamma = p - 2\alpha \operatorname{tr}(\boldsymbol{H}_i^{-1} \boldsymbol{W}_\theta)$;
  Compute new estimates $\alpha = \gamma / (2E_\theta(\boldsymbol{\theta}_{i+1}))$ and $\beta = (m - \gamma)/(2E_r(\boldsymbol{\theta}_{i+1}))$;
  $i = i + 1$;
end

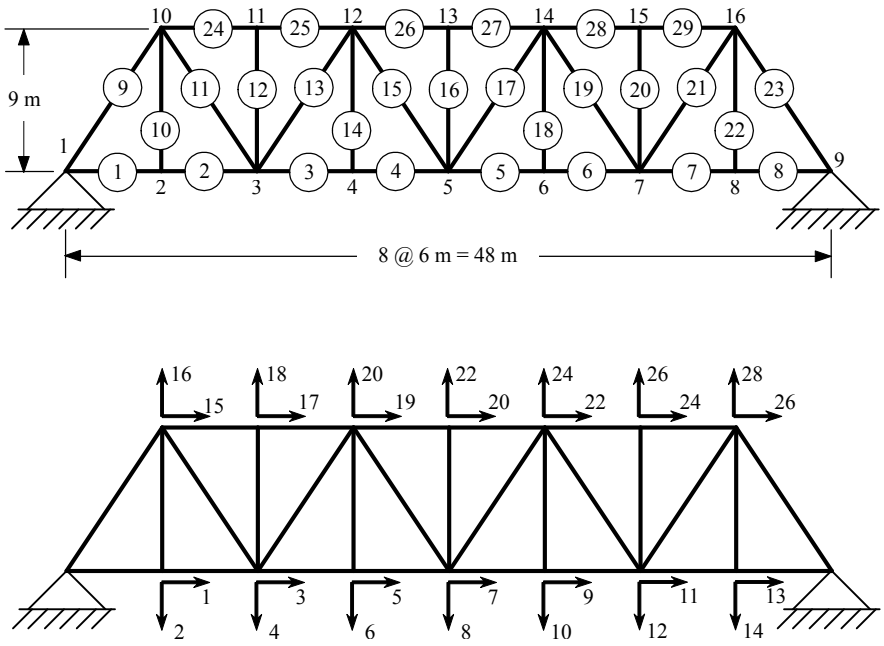


Figure 1: 28-DoF truss structure, modified from Papadimitriou *et al* [35]

The purpose of this small-scale model is not to demonstrate the power of parameter clustering for FE model updating, as a 29-element updating problem is easily tractable without any reduction in parameters. Instead, the small scale of the truss model allows a computationally-feasible comparison between different clustering methods across numerous realizations of measured modal properties. A full-scale study only offers one realized state for measurement, limiting the ability to deduce if one updating method is generally superior to another method, or simply superior for that measured state.

The FE model of the truss was developed in MATLAB [36] using consistent (non-lumped) element mass matrices. For consistency with [35], the first five vibrational modes were analyzed, having natural frequencies of {9.31, 19.8, 26.9, 37.3, 51.2} Hz with no structural modification. It was assumed that the full field of 28

DoFs were measured.

Random realized states were generated by modifying the Young's modulus (E) and density (ρ) for each element, e, by an uncorrelated, uniform random number δ_e^k or δ_e^m, as in Eqs. (20) and (21). Since the truss element stiffness ($\boldsymbol{K}_e$) and element mass ($\boldsymbol{M}_e$) are linearly dependent on the element E and ρ, respectively, this could also be seen as a direct modification of the element matrices. Two different modification levels were analyzed, $\delta \in [-0.5, 0.5]$ and $\delta \in [-0.05, 0.05]$, representing a poor and a good initial model state, respectively. 1000 random realized states were generated for each modification level.

Noise was then added to the natural frequency and mode shape measurements from each of the 2000 structural realizations. Each natural frequency measurement was corrupted with a Gaussian white noise sample with standard deviation equal to 0.5% of the measured natural frequency. Each mode shape was corrupted with a Gaussian white noise vector with standard deviation equal to 5% of the mode shape's standard deviation. This noise model was considered to be consistent with typical measurement conditions, in which natural frequencies are often reliable to within 1%, but mode shape measurements may exhibit an order-of-magnitude lower precision [1, 37].

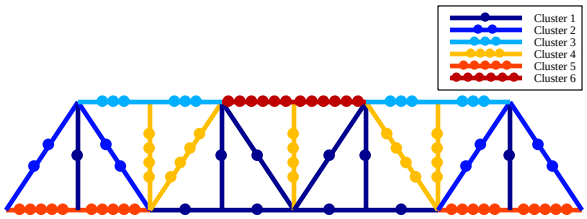
The objective function took the form of Eq. (18), with $\boldsymbol{W}_r$ equal to the inverse of measurement covariance matrix. Three different methods, or sets of sensitivity vectors, were used to generate six stiffness clusters and one mass cluster ($p = 7$) for each method using the techniques described in Section 3.2. The first method (f cluster) used only natural frequency sensitivities $\partial \boldsymbol{f}/\partial \delta_k$, as in [2, 4, 12, 13]. Of course, the authors in [2, 4] only sought to update frequencies in their full-scale tests, so utilizing mode shape sensitivities would have been superfluous. The second method ($f + \phi$ cluster) used concatenated natural frequency and mode shape sensitivities $\partial \boldsymbol{f}/\partial \delta_k$ and $\partial \boldsymbol{v}_j/\partial \delta_k$, similar to $-\partial \boldsymbol{r}/\partial \delta_k$. The third method (OC cluster), which is the proposed method in this paper, utilized the objective-consistent weighting in Eq. (29) for each residual sensitivity vector.

Representative stiffness clusters for the three clustering methods, with the corresponding sensitivities, are shown in Fig. 2. Clustered elements are indicated by color and the number of dots on the element in Figs. 2a, c, and e. Note that f clustering and $f + \phi$ clustering were based only on initial model sensitivities and were therefore constant for all realizations. Conversely, OC clustering had weighting which depended on the realized residual and was not constant for all realizations. f clustering necessarily results in symmetric clusters for a symmetric structure, i.e. perturbing one element's stiffness or mass will have the same effect on natural frequencies as perturbing the symmetric element's stiffness or mass. Mode shape sensitivities are not generally symmetric or anti-symmetric across structural elements, even in a symmetric structure. Incorporating mode shape sensitivities into clustering resulted in asymmetric clusters, as in the $f+\phi$ and OC clustering in Figs. 2c and e. This asymmetry may not be a general result for OC clustering, e.g. a weighting matrix which strongly emphasizes natural frequency sensitivities may result in symmetric clustering. The absolute value of sensitivities are presented as 3-d bar plots in Figs. 2b, d, and f. The cosine distance calculation utilized the full set of mode shape sensitivities (5 modes with 28 DoFs per mode), but are represented by MAC sensitivity in the bar plots for clarity.

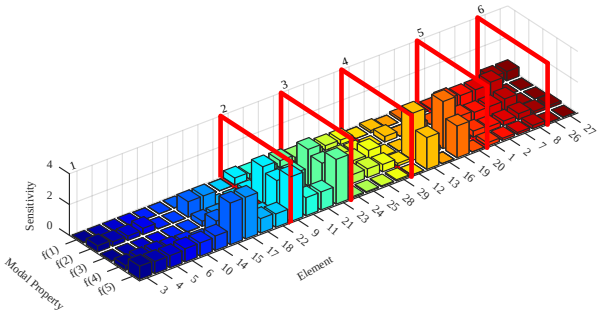
5.2. Model updating results

The objective function was minimized using the Levenberg–Marquardt method with Bayesian regularization described in Section 4 and Algorithm 1. Model updating was performed on each of the 2000 measurement realizations (i.e. for each randomized structural state, after the addition of noise) using the three different methods to select clusters of FE model parameters. The average results of model updating with a poor initial model, i.e. $\delta \in [-0.5, 0.5]$, are shown in Table 1. The L_2 norm of the relative frequency error showed significant improvement with each clustering method, with slightly better results for f clustering over OC clustering. This situation was reversed for average MAC (across all 5 mode shapes and 1000 realizations), where OC clustering slightly outperformed the other methods. However, OC clustering clearly outperformed the other methods in objective value (E_r), with approximately 17% lower objective function value relative to the initial value. OC clustering also resulted in slightly better γ, indicating that the objective function was more sensitive to those updating parameters. Each method had γ adequately close to $p = 7$, suggesting that most of the parameters were effective in updating.

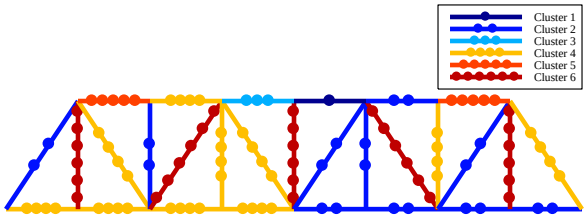
This set of observations is even more clear with a good initial model, $\delta \in [-0.05, 0.05]$ (Table 2). OC clustering was again outperformed by f and $f + \phi$ clustering for reducing frequency error, but it produced an average MAC value of 0.997 while f clustering essentially failed to improve the average MAC from



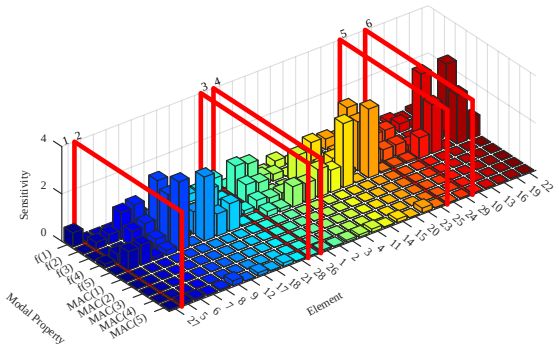
(a) Frequency (f) clustering



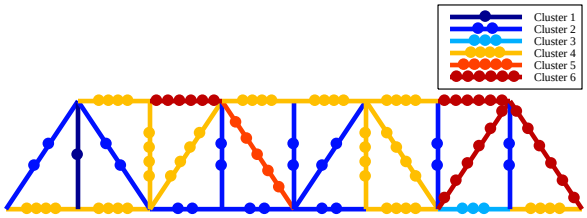
(b) Frequency (f) sensitivities



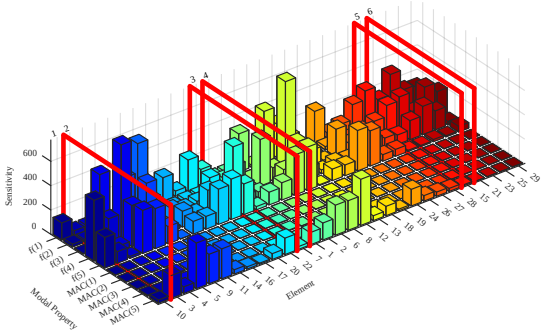
(c) Freq. and mode shape ($f + \phi$) clustering



(d) Freq. and mode shape ($f + \phi$) sensitivities



(e) Objective-consistent (OC) clustering



(f) Objective-consistent (OC) sensitivities

Figure 2: Selected truss clustering results and corresponding sensitivities

Table 1: Truss model updating results, $\delta \in [-0.5, 0.5]$

	Initial	Updated		
Average result		f cluster	$f + \phi$ cluster	OC cluster
$\left\|(\tilde{\boldsymbol{f}} - \boldsymbol{f})/\tilde{\boldsymbol{f}}\right\|_2$	15.5%	4.36%	9.20%	7.95%
MAC	0.920	0.933	0.944	0.956
Obj. value, E_r ($\times 10^3$)	6.47	4.43	5.05	3.33
Eff. no. params., γ		4.67	4.42	4.86

0.977. Ultimately, this means that f clustering was ineffective in reducing the objective function value E_r, while $f + \phi$ and OC clustering were extremely effective, with 80% and 86% reduction in E_r. The updating

parameters selected by f clustering were largely ineffective, suggesting that fewer clusters could have been used to achieve a similar result. OC clustering was considerably more effective, with approximately $\gamma = 5$ effective parameters.

Table 2: Truss model updating results, $\delta \in [-0.05, 0.05]$

	Initial	Updated		
Average result		f cluster	$f + \phi$ cluster	OC cluster
$\left\|(\tilde{\boldsymbol{f}} - \boldsymbol{f})/\tilde{\boldsymbol{f}}\right\|_2$	1.55%	1.01%	1.51%	1.25%
MAC	0.977	0.977	0.996	0.997
Obj. value, E_r ($\times 10^3$)	1.48	1.46	0.29	0.21
Eff. no. params., γ		1.62	4.52	4.81

6. Model updating of a full-scale FE model: large suspension bridge

6.1. System identification and model description

The second model updating exercise involves a full-scale FE model of a suspension bridge, as used in [12, 13]. The bridge is a double-deck steel structure with two towers and four suspension cables, each 982 m long. The three spans (two side-spans, one main-span) total 2089 m of length, with the main span comprising 451 m. A 2009 vibration study was performed to identify natural frequencies, damping ratios, and mode shapes from ambient excitation (traffic and wind). Dynamic responses were recorded at 9 locations using tri-axial force-balance accelerometers, totaling 27 measured DoFs at a sampling frequency of 200 Hz. The first seven dynamic modes were selected for model updating, totaling $m = 196$ modal measurements between 1 natural frequency and 27 DoFs for each mode.

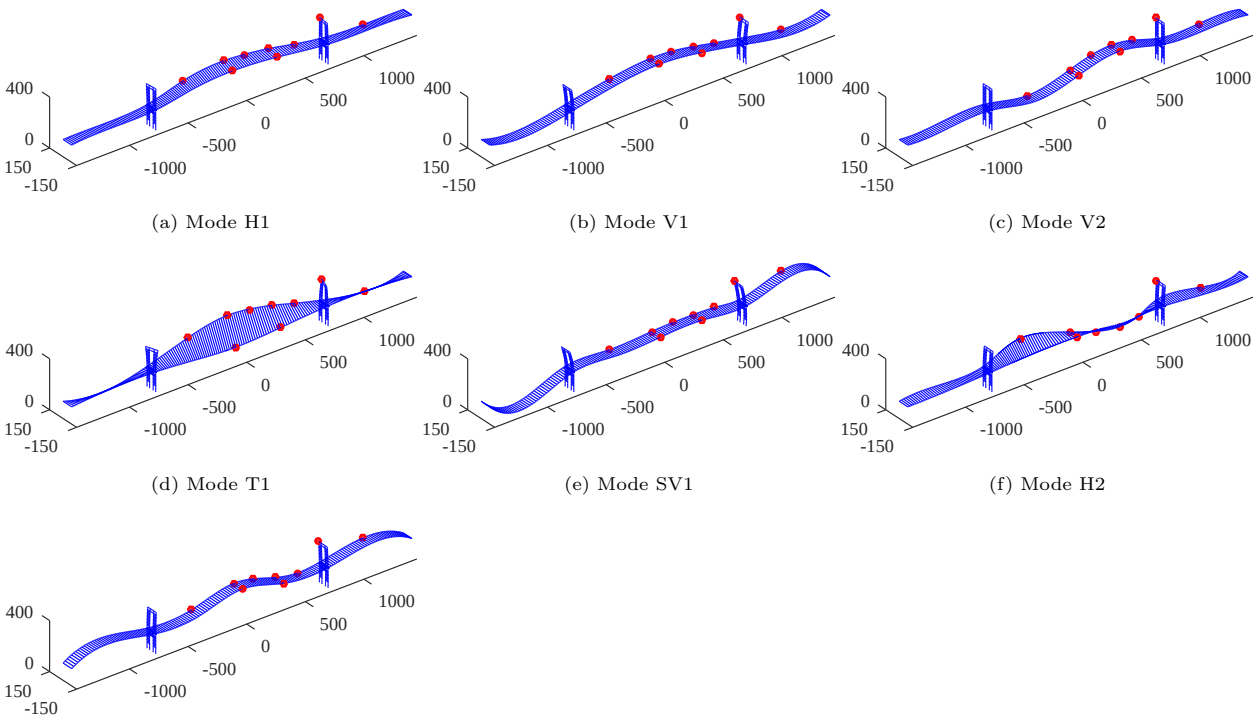
Responses were recorded during four 1-hour sessions in a single day, at 3 a.m., 8 a.m., 1 p.m., and 8 p.m. The responses during these time periods were assumed to be approximately stationary, as the environmental and operating conditions were unlikely to vary significantly. Output-only modal identification was performed on each of the 1-hour data sets using the enhanced frequency domain decomposition (EFDD) method [38]. Power spectral density matrices were constructed using Welch's method with a Hamming window. The identified modal properties for all four 1-hour periods were then averaged to give a set of modal properties representative of an average daily behavior. As shown in previous work , the measured modes had negligible imaginary components [12, 13], so only the real components of the mode shapes were used. The average measured natural frequencies are given in Table 3.

Table 3: Suspension bridge measured and initial modal properties

		Measured	Initial		
Mode	Description	$\tilde{f}$ (Hz)	f (Hz)	$(\tilde{f} - f)/\tilde{f}$	MAC
H1	First lateral	0.194	0.236	-21.7%	0.984
V1	First vertical	0.227	0.294	-29.8%	0.969
V2	Second vertical	0.303	0.356	-17.5%	0.986
T1	First torsional	0.373	0.384	-3.0%	0.741
SV1	First side-span vertical	0.337	0.453	-34.2%	0.879
H2	Second lateral	0.450	0.540	-19.8%	0.845
V3	Third vertical	0.500	0.596	-19.4%	0.743

Fig. 3 depicts the measured mode shapes, with sensor locations indicated by dots. The mode shape magnitudes at the sensor locations are representative of the average measured modal data, while the unmeasured parts of the mode shape were interpolated to fit projected data and given boundary conditions. For more information on the interpolated mode shapes, please refer to [12, 13]. Note that interpolation was only used for purposes of depiction, while any discussion of measured modes or MAC only utilizes data at the 27

measured DoFs. Fig. 4a shows the AutoMAC matrix, showing strong orthogonality between the measured mode shapes.

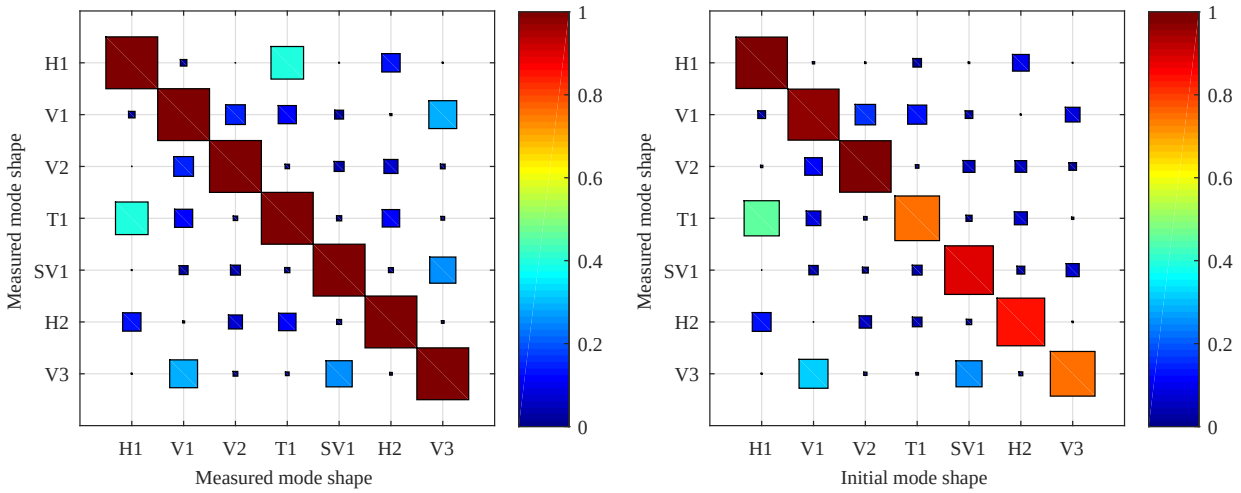


(a) Mode H1 (b) Mode V1 (c) Mode V2

(d) Mode T1 (e) Mode SV1 (f) Mode H2

(g) Mode V3

Figure 3: Suspension bridge measured mode shapes (measurement locations indicated by red dots)



(a) AutoMAC (b) Initial MAC

Figure 4: Suspension bridge AutoMAC (a) and initial MAC (b) matrices

The FE model of the suspension bridge was implemented in ABAQUS [39] with 19,632 beam elements, 1,464 truss elements, and 18,614 nodes. The node locations and connectivity were carefully defined based on partial technical drawings and photographs. Element properties were calculated from indicated cross-section data when available, while reasonable assumptions based on photographs were made for uninformed cases.

Soil springs were included at the bottom of the pylons, the end of the deck, and the end of the suspension cables. To account for temperature expansion joints, hinge springs were implemented between the deck and towers. For a more thorough discussion of element types, boundary conditions, connections, and initial model material properties, please refer to Jang and Smyth's description of the same FE model [12].

The natural frequencies, relative frequency error, and MAC values of the initial FE model are given in Table 3. The initial MAC values are visually represented Fig. 4b. The initial MAC values are slightly different from those in Jang and Smyth's previous work [12, 13] due to FE model modifications to account for more realistic structural behavior, particularly in the boundary conditions and interactions between cable and deck components. For each mode, the natural frequencies of the initial FE model were higher than their measured counterparts. The first torsional (T1) mode exhibited the lowest initial error at -3.0%, while the first side-span vertical (SV1) mode exhibited the highest initial error at -34.2%. Initial MAC values were excellent (greater than 0.95) for the first two vertical modes (V1 and V2) as well as the first lateral mode (H1). Conversely, the first torsional (T1) and the third vertical mode (V3) exhibited the most unsatisfactory MAC values at 0.741 and 0.743, respectively. Every mode exhibited a low MAC and/or a high frequency error, suggesting that every mode would be important in the updating process.

6.2. Parameter clustering

The selection of clusters and updating parameters for the full-scale suspension bridge model proceeded using similar methodology to the small-scale truss model in Section 5. The large scale of the suspension bridge eliminated the possibility of directly updating every element mass and stiffness. The FE model had well over 42,000 physical parameters to update, among more than 21,000 elements with separate Young's moduli and mass densities, plus several soil and hinge spring constants. Note that the geometry of the suspension bridge was not included within the updating parameters.

Prior to clustering, the structural components of the FE model were arranged into 132 substructures based on location and element type, which could be viewed as an expert-informed pre-clustering. The span was longitudinally divided into 8 main span and 8 side span segments, which were further partitioned based on element type. Each tower was divided into three vertical sections, which was further divided into bracing and pylon groups. The 132 resulting substructures were assigned mass density (δ^m) and Young's modulus (δ^k) modifications as model parameters. The soil and hinge springs, comprising 15 spring coefficients, were also assigned spring constant modification parameters (δ^k). This totaled 147 stiffness model parameters and 132 mass model parameters, for 279 total model parameters to be updated.

Two sets of clusters were selected using the methodology of Section 3.2, one based on frequency sensitivity (f cluster) and the other based on the objective-consistent weighted residual sensitivity (OC cluster). The residual weighting matrix $\boldsymbol{W}_r$ was formed based on an assumed noise model, equivalent to that in Section 5. It was assumed that each frequency measurement had a standard deviation equal to 0.5% of the measured frequency value. For each mode, it was assumed that every component had a standard deviation equal to 5% of the measured mode shape's standard deviation. Therefore, the covariance matrix of $\tilde{\boldsymbol{z}}$ was assumed to be diagonal, with $\boldsymbol{W}_r$ as the reciprocal of the measurement covariance matrix.

For equivalence with previous work on this bridge [12, 13], 5 mass clusters and 17 stiffness clusters were used for both the f clustering and OC clustering results. The sensitivities of the model outputs to the 22 OC cluster updating parameters are represented in Fig. 5. A more detailed representation of one mass cluster (3) and one stiffness cluster (19) are presented in Fig. 6. Similar to Section 5, OC clustering resulted in several asymmetric clusters, while frequency clustering as done in previous work [12] resulted in symmetric clusters. Cluster 3 is appreciably asymmetric with all clustered model parameters exhibiting a significant effect on the second horizontal mode (H2). Cluster 19 includes many structural elements in the mid-span area, ends of the side-spans, and the cable-soil springs. Physical intuition would indicate that these elements would be most impactful on the vertical modes (V1-V3), which is confirmed by the sensitivities in Fig. 6d.

6.3. Model updating results

With the models parametrized using the two clustering schemes, model updating proceeded using Levenberg–Marquardt minimization with Bayesian regularization, detailed in Section 4 and Algorithm 1. The updating parameters of both clustering schemes converged in 3 iterations, with results shown in Table 4.

Natural frequency updating results are similar for f and OC clustering for modes V2, T1, SV1, and V3. Both clustering schemes struggled to update the first torsional mode (T1), with neither achieving less than

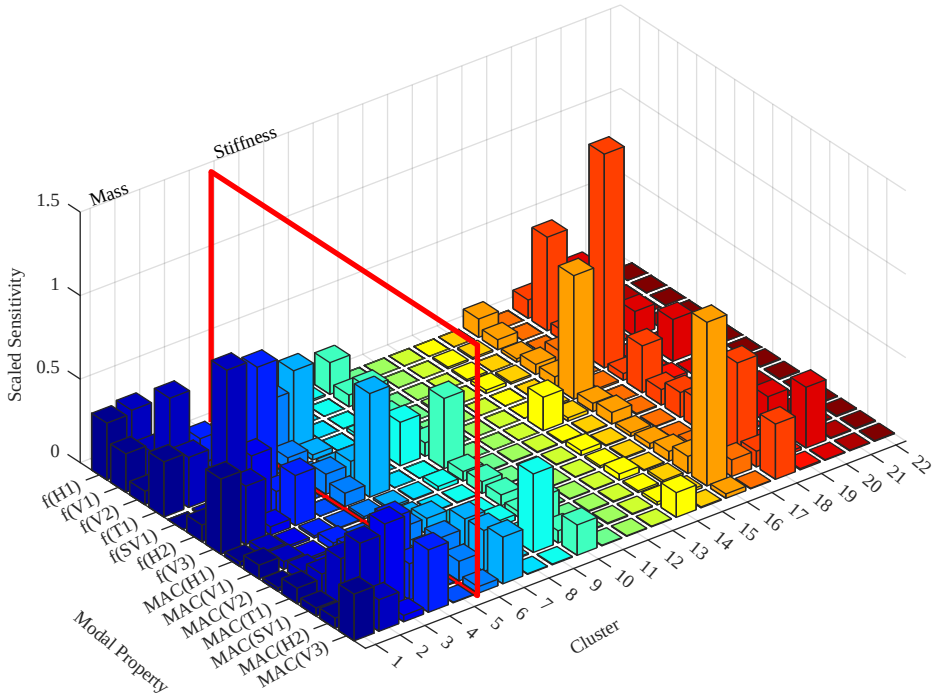


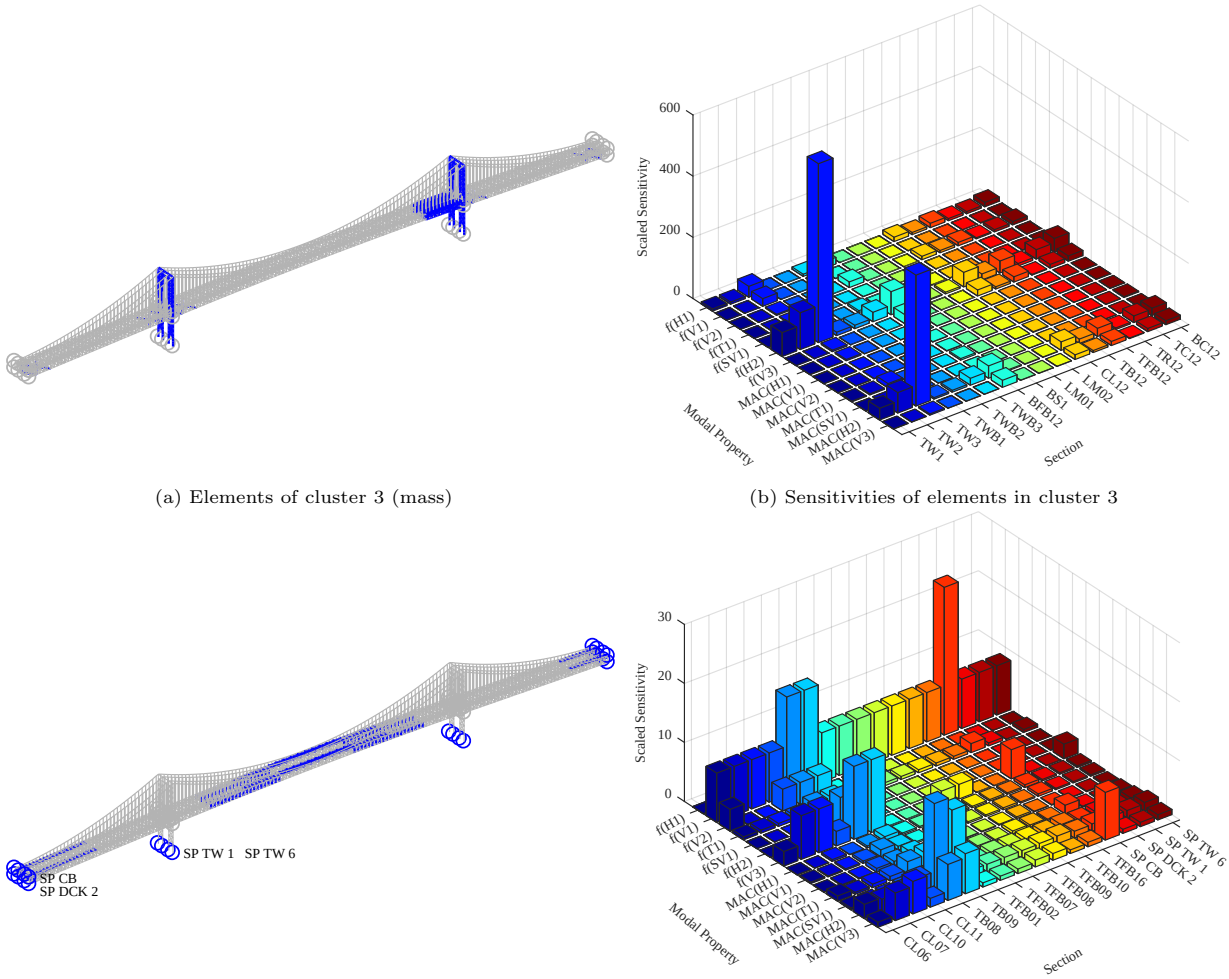
Figure 5: Suspension bridge sensitivities to the 22 objective-consistent cluster updating parameters

Table 4: Suspension bridge model updating results

	Initial		Updated			
			f cluster		OC cluster	
Mode	$(\tilde{f}-f)/\tilde{f}$	MAC	$(\tilde{f}-f)/\tilde{f}$	MAC	$(\tilde{f}-f)/\tilde{f}$	MAC
H1	-21.7%	0.984	1.0%	0.989	4.4%	0.993
V1	-29.8%	0.969	-6.9%	0.966	-11.0%	0.973
V2	-17.5%	0.986	2.2%	0.977	2.2%	0.984
T1	-3.0%	0.741	12.9%	0.766	12.4%	0.832
SV1	-34.2%	0.879	-4.7%	0.861	-5.4%	0.869
H2	-19.8%	0.845	-3.6%	0.876	2.1%	0.970
V3	-19.4%	0.743	-5.3%	0.812	-5.6%	0.983
$\left\|(\tilde{\boldsymbol{f}}-\boldsymbol{f})/\tilde{\boldsymbol{f}}\right\|_2$	60.0%		16.8%		19.0%	
Av. MAC		0.878		0.893		0.928
Obj. value, E_r ($\times 10^4$)	2.39		0.95		0.70	
Eff. no. params., γ			6.4		10.1	

12% relative frequency error. The difficulty with mode T1 was noted in previous work [12]. Generally, f clustering achieved better natural frequency results, achieving a 72% reduction in L_2 norm natural frequency error while OC clustering achieved 68% reduction.

However, the converse occurred for mode shape updating, as indicated by OC clustering increasing the average MAC to 0.928 from an initial value of 0.878 while f clustering only achieved a mild improvement to 0.893. OC clustering produced better MAC values than f clustering for every mode, with mild improvements noted for both methods on modes H1, V1, and V2, which had high initial MAC values. OC clustering had the most notable impact on the MAC of modes H2 and V3, which were boosted from less than 0.850 to over 0.970, while f clustering showed only a minor improvement. Similarly, OC clustering was able to raise the MAC of mode T1 to 0.832 from 0.741 which is an appreciable increase and clearly outperformed f cluster's



(a) Elements of cluster 3 (mass)

(b) Sensitivities of elements in cluster 3

(c) Elements of cluster 19 (stiffness)

(d) Sensitivities of elements in cluster 19

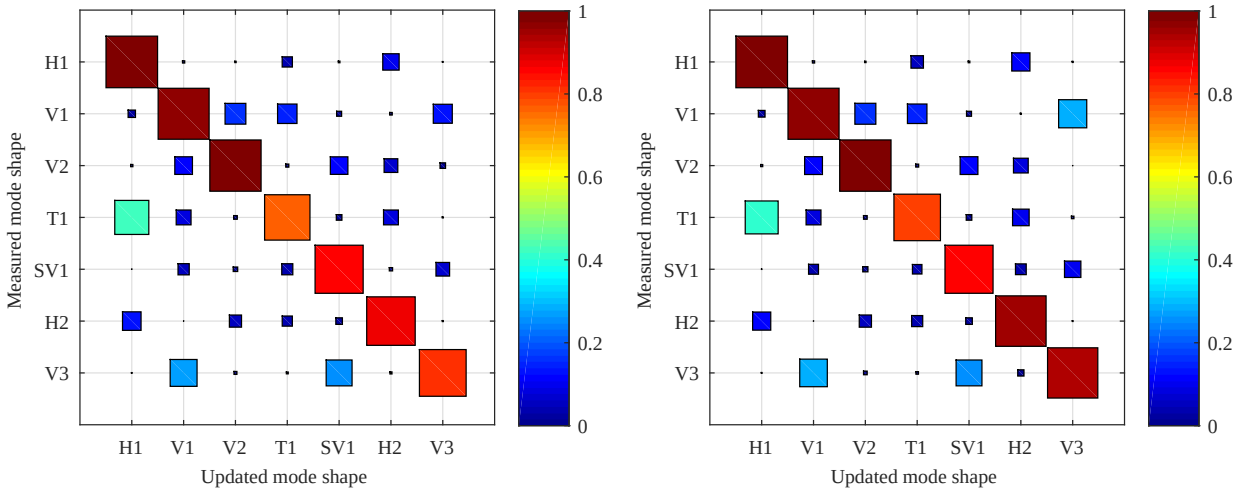
Figure 6: Suspension bridge objective-consistent clustering results: selected clusters and sensitivities

improvement to 0.766. Both clustering methods struggled to update the mode shape of the side-span vertical mode (SV1), actually reporting slight decreases in MAC value. The updated MAC values are also presented diagrammatically in Fig. 7.

Despite mildly poorer results with natural frequencies, the enhanced improvement in mode shapes provided OC clustering with a significant edge over f clustering with respect to the objective function value E_r (Table 4). OC clustering reduced E_r by 70% from its initial value, while f clustering only reduced E_r by 60%. This result is partly explained by OC clustering having a higher effective number of parameters, γ, during the updating process. Out of $p = 22$ updating parameters, $\gamma = 10.1$ were effective for OC clustering while only $\gamma = 6.4$ were effective for f clustering.

7. Conclusions

Sensitivity-based parameter clustering presents a viable method for improving the condition and efficiency of FE model updating problems. The presented OC parametrization scheme allows for compatible use of disparate data (e.g. natural frequencies and mode shapes) in the selection of parameter clusters, thereby improving the efficiency and quality of results. The proposed clustering scheme retains the physical relevance of previous clustering schemes, which only used natural frequency sensitivities. Laid atop this foundation, OC clustering also considers the measurement residual and residual weighting inherent in the objective



(a) f cluster updated MAC

(b) OC cluster updated MAC

Figure 7: Suspension bridge frequency cluster (a) and OC cluster (b) updated MAC matrices

function. Indeed, the presented weighting scheme is extremely generic, as it helps produce clusters which have the most similar effect on the weighted residual.

The proposed OC clustering scheme was shown to be highly effective in both a small-scale exercise with simulated data and a full-scale exercise with real data. When compared to a clustering scheme based only on natural frequency sensitivity, OC clustering resulted in significantly better updating of mode shapes at the cost of slightly more frequency error. Incorporation of mode shape data produced asymmetric clusters, while clustering based on frequency produced symmetric clusters for the symmetric structures of study. For the small-scale model with a poor initial model (far from measured state), OC clustering achieved 17% better error reduction than frequency clustering. With a good initial model, OC clustering reduced the objective value by 84% more than frequency clustering. These results were more tempered on the full-scale model, where OC clustering reduced the objective value by 10% more than frequency clustering, but it also showed significantly better agreement with measured mode shapes.

Bayesian regularization was proposed for use in model updating. Built atop the Levenberg–Marquardt minimization algorithm, Bayesian regularization delivers an optimal set of regularization parameters with minimal computational overhead. Additionally, it provides key insight into the effective number of updating parameters, which can be used for model selection and model comparison. Observing the effective number of parameters further confirmed the improved efficiency of OC cluster analysis, which consistently showed a greater number of effective parameters compared to frequency clustering.

The two contributions of this paper, OC clustering and Bayesian regularization, present simple and effective developments on existing methods. However, these contributions are not without drawbacks. On well-posed problems, parameter clustering and regularization are unnecessary. While OC clustering is intuitive and showed improvement over existing schemes, it is not guaranteed to be the optimal technique for every structure or realization. Further work is required to analyze the effect of clustering on problem condition. OC clustering also suffers by not considering physical proximity of elements, resulting in clustered elements which are generally not physically adjacent. This lack of proximity poses problems for damage localization [12]. It may be noted that clusters are only selected once, starting at the initial model. As the model is updated, however, the model sensitivity matrix will change. Depending on the amount of change, it may be such that the initial clustering is no longer efficient. This problem was not considered, and warrants further research.

Regularization has the drawback of altering the objective, adding a side-constraint which may not be important to the user. While Bayesian regularization provides a key insight, it is not a refined tool for model selection. While it can be used to compare the efficiency of different parametrizations, it does not provide strong suggestions for alternative parametrizations other than to increase or decrease the number of parameters. Further work is needed to understand the limitations of this approach.

Acknowledgments

The authors gratefully acknowledge Columbia University's Graduate School of Arts and Sciences in support of the first author through the Guggenheim Fellowship and Presidential Fellowship. This work was partially supported by the U.S. National Science Foundation (Grant No. CMMI-1563364).

References

References

[1] J. E. Mottershead, M. I. Friswell, Model updating in structural dynamics: a survey, Journal of Sound and Vibration 167 (2) (1993) 347–375.

[2] J. E. Mottershead, M. Link, M. I. Friswell, The sensitivity method in finite element model updating: a tutorial, Mechanical Systems and Signal Processing 25 (7) (2011) 2275–2296.

[3] M. I. Friswell, J. E. Mottershead, Finite Element Model Updating in Structural Dynamics, Vol. 38, Springer Science & Business Media, 1995.

[4] H. Shahverdi, C. Mares, W. Wang, J. E. Mottershead, Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise, Shock and Vibration 16 (1) (2009) 75–87.

[5] J. E. Mottershead, C. Mares, M. I. Friswell, S. James, Selection and updating of parameters for an aluminium space-frame model, Mechanical Systems and Signal Processing 14 (6) (2000) 923–944.

[6] J. M. W. Brownjohn, P. Moyo, P. Omenzetter, Y. Lu, Assessment of highway bridge upgrading by dynamic testing and finite-element model updating, Journal of Bridge Engineering 8 (3) (2003) 162–172.

[7] A. Teughels, G. De Roeck, Structural damage identification of the highway bridge Z24 by FE model updating, Journal of Sound and Vibration 278 (3) (2004) 589–610.

[8] B. Jaishi, W.-X. Ren, Structural finite element model updating using ambient vibration test results, Journal of Structural Engineering 131 (4) (2005) 617–628.

[9] J. R. Wu, Q. S. Li, Finite element model updating for a high-rise structure based on ambient vibration measurements, Engineering Structures 26 (7) (2004) 979–990.

[10] P. G. Bakir, E. Reynders, G. De Roeck, Sensitivity-based finite element model updating using constrained optimization with a trust region algorithm, Journal of Sound and Vibration 305 (1) (2007) 211–225.

[11] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, Mechanical Systems and Signal Processing 56 (2015) 123–149.

[12] J. Jang, A. W. Smyth, Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters, Mechanical Systems and Signal Processing 83 (2017) 337–355.

[13] J. Jang, A. W. Smyth, Bayesian model updating of a full-scale finite element model with sensitivity-based clustering, Structural Control and Health Monitoring. In press.

[14] H. Ahmadian, J. E. Mottershead, M. I. Friswell, Regularisation methods for finite element model updating, Mechanical Systems and Signal Processing 12 (1) (1998) 47–64.

[15] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Finite–element model updating using experimental test data: parametrization and regularization, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 359 (1778) (2001) 169–186.

[16] B. Titurus, M. I. Friswell, Regularization in model updating, International Journal for Numerical Methods in Engineering 75 (4) (2008) 440–478.

[17] D. J. C. MacKay, Bayesian interpolation, Neural Computation 4 (3) (1992) 415–447.

[18] F. D. Foresee, M. T. Hagan, Gauss-Newton approximation to Bayesian learning, in: International Conference on Neural Networks, 1997, Vol. 3, IEEE, 1997, pp. 1930–1935.

[19] Å. Björck, Numerical methods for least squares problems, SIAM, Philadelphia, PA, 1996.

[20] R. L. Fox, M. P. Kapoor, Rates of change of eigenvalues and eigenvectors, AIAA Journal 6 (12) (1968) 2426–2429.

[21] S. Adhiakri, Rates of change of eigenvalues and eigenvectors in damped dynamic system, AIAA Journal 37 (11) (1999) 1452–1458.

[22] R. J. Allemang, D. L. Brown, A correlation coefficient for modal vector analysis, in: Proceedings of the 1st International Modal Analysis Conference, Vol. 1, 1982, pp. 110–116.

[23] K. D. Hjelmstad, M. O. Banan, M. A. Banan, On building finite element models of structures from modal response, Earthquake Engineering & Structural Dynamics 24 (1) (1995) 53–67.

[24] G. Lallement, J. Piranda, Localization methods for parametric updating of finite element models in elastodynamics, in: International Modal Analysis Conference, 8th, 1990, pp. 579–585.

[25] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Combining subset selection and parameter constraints in model updating, Journal of Vibration and Acoustics 120 (4) (1998) 854–859.

[26] L. Rokach, O. Maimon, Clustering Methods, Springer US, Boston, MA, 2005, Ch. 15, pp. 321–352.

[27] R. R. Sokal, C. D. Michener, A statistical method for evaluating systematic relationships, University of Kansas Science Bulletin 38 (1958) 1409–1438.

[28] J. D. Collins, G. C. Hart, T. Haselman, B. Kennedy, Statistical identification of structures, AIAA Journal 12 (2) (1974) 185–190.

[29] M. I. Friswell, The adjustment of structural parameters using a minimum variance estimator, Mechanical Systems and Signal Processing 3 (2) (1989) 143–155.

[30] T. Strutz, Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond, Vieweg and Teubner, Germany, 2010.

[31] M. Link, Updating of analytical models–procedures and experience, in: Proceedings of the Conference on Modern Practice in Stress and Vibration Analysis, Sheffield Academic Press, 1993, pp. 35–52.

[32] A. N. Tikhonov, V. I. Arsenin, Solutions of ill-posed problems, Vol. 14, V. H. Winston and Sons (distributed by Wiley, New York), 1977.

[33] K. Levenberg, A method for the solution of certain problems in least squares, Quarterly of Applied Mathematics 2 (1944) 164–168.

[34] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics 11 (2) (1963) 431–441.

[35] C. Papadimitriou, J. L. Beck, S.-K. Au, Entropy-based optimal sensor location for structural model updating, Journal of Vibration and Control 6 (5) (2000) 781–800.

[36] MATLAB, version 9.1.0 (R2016b), The MathWorks Inc., Natick, MA, 2016.

[37] S. V. Modak, T. K. Kundra, B. C. Nakra, Comparative study of model updating methods using simulated experimental data, Computers & Structures 80 (5) (2002) 437–447.

[38] R. Brincker, C. Ventura, P. Andersen, Damping estimation by frequency domain decomposition, in: 19th International Modal Analysis Conference, 2001, pp. 698–703.

[39] ABAQUS/CAE, User's Guide: Version 6.14, Dassault Systèmes Simulia Corp., Providence, RI, 2014.