Pai S.G.S., M. Sanayei, and I. F. C. Smith. 2021. "Model class selection using clustering and classification for structural identification and reserve capacity prediction" *J. Comput. Civ. Eng.* 35 (1): 1–15.
https://doi.org/10.1061/(ASCE)CP.1943-5487.0000932.

Model class selection using clustering and classification for structural identification and prediction

Sai G.S. Pai[1,*], Masoud Sanayei[2], and Ian F.C. Smith[3,4]

[1]Project Coordinator, Cyber Civil Infrastructure research group, ETH Zurich, Future Cities Laboratory, Singapore-ETH Centre, 1 Create Way, CREATE Tower, Singapore 138602
[*]Corresponding author email: sai.pai@sec.ethz.ch
[2]Professor, Department of Civil and Environmental Engineering, Tufts University, Medford, USA
[3]Professor and Head of laboratory, Applied Computing and Mechanics Laboratory (IMAC), School of Architecture, Civil and Environmental Engineering (ENAC), Swiss Federal Institute of Technology (EPFL), Station 18, Lausanne, 1015, Switzerland
[4]Principal Inverstigator, Cyber Civil Infrastructure research group, ETH Zurich, Future Cities Laboratory, Singapore-ETH Centre, 1 Create Way, CREATE Tower, Singapore 138602

ABSTRACT

Structural identification using physics-based models and subsequent prediction have much potential to enhance civil infrastructure asset-management decision-making. Interpreting monitoring information in the presence of multiple uncertainty sources and systematic bias using a physics-based model is a computationally expensive task. The computational cost of this task is exponentially proportional to the number of model parameters updated using monitoring data. In this paper, a novel model-class selection

method is proposed to obtain computationally optimal and identifiable model classes. Unlike traditional sensitivity methods for model-class selection, in the proposed method, model responses at sensor locations are clustered to identify underlying trends in model response datasets. K-means clustering is used to determine relevant clusters in the data. Cluster indices are then used as labels for classification. Support-vector machine classification using forward variable selection with sequential search is used to select model parameters that help classify trends in data. The result of the sequential search is a trade-off curve comparing classification error with number of parameters in the model class. This curve helps select a practical and near optimal model class. The model-class selection method proposed in this paper is compared with linear regression-based sensitivity analysis using a full-scale bridge. Identification with model classes obtained using both methods for two sensor configurations suggests that the model-based clustering method helps select an identifiable and computationally efficient model class. The minimum remaining fatigue life of the bridge predicted using the updated model classes is 720 years and this represents fatigue-life extension of ten times compared with design predictions prior to measurements. This approach provides good support for asset managers when they interpret measurement data.

INTRODUCTION

Structural identification of civil infrastructure involves use of physics-based models to interpret measurement data. The objective of interpreting measurement data in the context of structural identification is to improve knowledge of model parameters. Measurements carried out seldom contain information pertaining to all parameters of the physics-based model. The task of selecting the right parameters for identification is important to carry out structural identification in an efficient manner. In this paper, a novel method to select parameters for identification using measurement data is presented.

Structural identification enables better management of existing civil infrastructure assets. The task of asset management is one of the foremost challenges today. The construction industry is one of the largest consumers of raw materials. With deteriorating supply of existing infrastructure and increasing demand, a large part of future global investment will be made towards maintaining existing infrastructure to ensure a safe and good standard of living (World Economic Forum 2014). Good asset-management decision-making supports timely inspection and intervention, thus minimizing cost of unnecessary and ill-informed management actions. Structural identification, using physics-based models to interpret measurement data, leads to a better understanding

of real structural behavior. This enables engineers to make informed decisions about repair, retrofit and replacement of existing civil infrastructure. With the advent of fast computing technologies (Frangopol and Soliman 2016) and cheap sensing tools (Lynch and Loh 2006; Taylor et al. 2016), use of structural identification in practice is feasible in an increasing number of situations.

Structural identification of civil infrastructure is an ill-conditioned inverse problem due to the presence of large modeling uncertainty from use of justifiably conservative design models. Due to the nature of the inverse problem, unique solutions of model parameters cannot be identified. Rigorous evaluation of uncertainties associated with the system can help obtain accurate solutions that are robust in the presence of model bias. For structural identification, probabilistic methods such as Bayesian model updating (Beck and Katafygiotis 1998) and error-domain model falsification (EDMF) (Goulet and Smith 2013) are useful when updating knowledge of structural behavior. Use of probabilistic methods for data-interpretation is computationally expensive as many combinations of model parameters have to be evaluated using detailed physics-based models. Probabilistic methods provide solutions as sets of model parameters, which provide model responses that are compatible with measurements observed during a load test.

To obtain solutions using probabilistic methods, a large model-parameter space has to be explored. The larger the parameter space to be explored, the greater is the computational cost of solving the inverse problem. Hence, to limit the computational cost, only a small subset of parameters are chosen from a much larger set of parameters that define the physics-based model. For a given physics-based model, from a large set of model parameters, many smaller subsets of parameters can be selected for identification. Each of these subsets of parameters defines a model class for identification. The task of selecting an appropriate model class for identification and subsequent predictions of structural behavior is called as model class selection (or feature selection) (Liu and Motoda 1998; Bennani and Cakmakov 2002; Guyon and Elisseeff 2006). Selection of a model-class for identification without utilizing information from measurements is called as a-priori model class selection and this is the focus of this paper.

Traditionally, selection of a model-class has been carried out using sensitivity analysis based on linear-regression models (Friedman 1991). Linear-regression models are developed for model response at sensor locations. Each linear-regression model provides information about sensitivity of model response to model parameters at a sensor location. Using these sensitivities as heuristics, parameters are selected based on their

importance to response at various sensor locations. Matos et al. (2016) suggested using the coefficient of variation as a metric to determine importance of model parameters to structural response. Another common approach that has been suggested is analysis of variance (ANOVA) to determine the model parameters that govern structural response (Van Buren et al. 2013; Van Buren et al. 2015). These methods do not implicitly contain any penalty for over-fitting and may lead to selection of unimportant parameters for identification.

Criteria in addition to sensitivity have been developed for parameter selection, such as Akaike information criterion (AIC) (Akaike 1974), Bayesian information criterion (BIC) (Schwarz et al. 1978) and Mallows C_p (Mallows 1973). These criteria have been utilized with forward or backward parameter selection strategies (Draper and Smith 2014; Jain et al. 2000) using adaptive strategies such as greedy heuristics to reduce computational cost (Caruana and Freitag 1994; Rao et al. 2015). These methods take into account over-fitting of models to training data. However, the applicability of such information criteria are limited to when the data are defined by Gaussian distributions. Such assumptions typically are not valid for responses of civil infrastructure.

Regularization (Tikhonov 1963) methods have also been proposed to select model parameters that are important to model response. Shrinkage methods reduce the number of parameters in a model to avoid over-fitting to training data. Using regularization methods that have been proposed, such as lasso regression (Tibshirani 1996), parameters that are not important to the response are eliminated. The elimination of parameters is determined by a shrinkage parameter that has to be tuned to obtain optimal model-class selection. Tuning of the shrinkage parameter has been shown to be a challenging task (Fan and Tang 2013).

Moreover, the a-priori methods available in literature based on linear regression assume linear relationships between model parameters and responses, which is typically not valid for civil infrastructure. Also, all a-priori methods focus on finding a good subset of parameters that influence model response at one sensor location. For civil infrastructure, model response at various sensor locations may not be governed by the same set of parameters. Typically, the importance of parameters to model responses have been averaged (Matos et al. 2016) or an intersection of parameters important to response at all sensor locations (Van Buren et al. 2015) have been assumed to provide an optimal model class. However, novel sensor placement strategies (Papadopoulou et al. 2015; Bertola et al. 2017; Argyris et al. 2017) have been developed with the aim of minimizing the number of sensors and maximizing information from each sensor.

Ideally, these strategies result in each sensor providing new information about model parameters. Under such conditions, use of averaged sensitivities is not an appropriate metric for model-class selection.

In this paper, a model-class selection method is presented to overcome many challenges related to structural identification of civil infrastructure. The selection of parameters is not carried out by evaluating the importance of parameters to response at each sensor location. Instead, a global understanding of structural behavior is evaluated using clustering. Clustering is a form of unsupervised learning that helps understand data by grouping together data-points that are similar. Clustering is inherently a subjective task as grouping of data-points is carried out using user-selected distance metrics (for example Euclidean distance) (Jain et al. 1999). Data-points grouped together form clusters, each with its own cluster index or label.

K-means clustering (Fisher and Ness 1971; Jain 2010) has been the most commonly used method for clustering. Other clustering methods that have been developed are Gaussian-mixture models (Biernacki et al. 2000; Friedman et al. 2001) and hierarchical clustering (Johnson 1967; Friedman et al. 2001). A key input required for most clustering methods is the number of clusters. To estimate the number of clusters in a dataset, many methods have been developed, such as the Hubert statistic (Halkidi et al. 2002), the Davies-Bouldin index (Davies and Bouldin 1979), the silhouette index (Kaufmann and Rousseeuw), the Dunn index (Dunn 1974) and the score function (Saitta et al. 2007; Saitta et al. 2008).

The gap method (Tibshirani et al. 2001) is another method that has been developed to determine the number of clusters. In this method, the number of clusters present in a given dataset is evaluated in a probabilistic manner. A gap value metric is introduced in this method, which compares the possibility that the datset contains c clusters with the hypothesis that the dataset contains no clusters. This gap value is calculated for various values of c ($c \in [1, k]$). Let K be the real number of clusters in the dataset. For $c = K$, the gap vlaue calculated is maximum compared with gap values calculated for other values of c. Therefore, the value c that maximizes the gap value is estimated as the optimal number of clusters in the dataset.

Utilizing the appropriate k values obtained using the gap method, k-means clustering helps obtain a global understanding of structural behavior at sensor locations. Further on, the parameters governing the trends in structural behavior can be ascertained using classification methods (Cristianini et al. 2000). This reveals the parameters that govern the structural behavior globally, rather than evaluating response at each sensor location

individually. The strategy of using clustering for parameter selection is similar to model-based clustering methods (Fraley and Raftery 2002) that have been developed for use with Gaussian-mixture models. Methods have also been developed to select parameters that affect model-based clustering (Scrucca and Raftery 2014). Most developments in parameter selection for model-based clustering methods have been to reduce the computational complexity (Scrucca 2016; Celeux et al. 2018). However, the extension of these methods to the context of structural identification have not yet been evaluated.

In Bayesian applications of structural identification, after updating (a-posteriori), assessment of identifiability (Ljung 2010) of model class has been carried out using Bayesian model class selection (Muto and Beck 2008; Yuen and Kuok 2015). Similarly, an a-posteriori methodology has been suggested by Saitta et al. (2009) to assess inclusion of parameters in a model class for identification for EDMF. However, no research has so far been conducted to quantify identifiability of model classes utilized for EDMF. In this paper, a cross-validation based identifiability assessment of identification in the form of a precision metric is presented. Moreover, the assessment of precision helps compare identifiability of model classes obtained using model-based clustering and other model-class selection methods.

A new methodology for model-class selection is explained in detail in the following section. After this, the application and comparison of this methodology with traditional sensitivity analysis for structural identification of a full-scale bridge are presented.

METHODOLOGY

In a-priori model-class selection, a subset of parameters, which define a physics-based model, is selected for identification. The objective of model-class selection methods is to select parameters that can be identified using measurements available from a load-test. Generally a-priori model-class selection is conducted using a simulated dataset, $\mathbf{D}$, similar to one shown in Figure 1.

The dataset, as shown in Figure 1, is simulated using a physics-based model, $\boldsymbol{g}(\boldsymbol{\theta})$. This model is defined by parameters $\boldsymbol{\theta} = \left[\theta_1, ..., \theta_p\right]$, where p is the total number of parameters in the model. Values of these parameters are not known and they are included in the model as random variables. Using the physics-based model, predictions of structural response are made at sensor locations to simulate the dataset, $\mathbf{D}$. Let the number of sensor locations where measurements are recorded during a load-test be m. Using the physics-based model, n evaluations are performed for various combinations of model parameters. Thus the dataset for parameter selection is composed of two subsets,

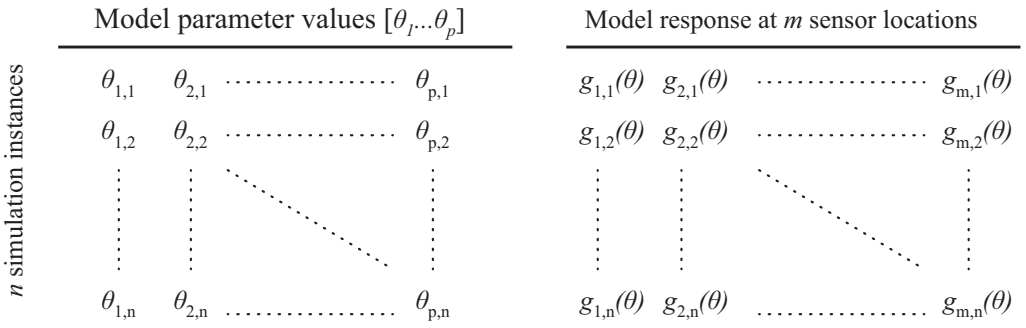


Fig. 1. Dataset simulated using a physics-based model, $\boldsymbol{g}(\boldsymbol{\theta})$, of size $n \times (m + p)$.

the model response set of size $n \times m$ and a set of corresponding model parameters of size $n \times p$ that serves as input into the model class. Thus, the size of the entire dataset is $n \times (m + p)$. These datasets are traditionally used for parameter selection using linear-regression-based sensitivity methods.

Dataset, **D**, simulated using the physics-based model, propagates uncertainty from model parameters to model response (behavior). Uncertainty in model parameters that govern structural response contribute more significantly to variability in model response. In typical sensitivity analysis, variability in model response is evaluated at each sensor location to determine the most important model parameters at individual sensor locations. However, parameters that are important at a sensor location are not necessarily important at other locations due to the global structural behavior.

Instead of evaluating the importance of model parameters to responses at each sensor location independently, clustering is utilized to obtain insight into changes in structural behavior from changes in parameter values. For example, a simple beam when simply supported, behaves in a manner different from the same beam under fixed-fixed end conditions. Comparing the two behaviors would indicate differences in deflection response predicted at sensor locations (measuring deflection) along the beam span. The parameters of this beam that are governing this variation in deflection response are the stiffnesses of the boundary conditions. Similarly, for full-scale structures, varying parameters in a physics-based model changes the structural behavior, which is reflected in response at sensor locations. These underlying changes in structural behavior are inferred from responses at sensor locations using k-means clustering.

The model response at sensor locations, as shown in Figure 2, is clustered to better understand structural behavior. A key input for k-means clustering is the number of clusters in the dataset. This is determined with the help of the gap method (Tibshirani et al. 2001). Subsequently, the model responses obtained at various sensor locations for different combinations of parameter values are clustered, according to similar values for quantities that sensors measure.

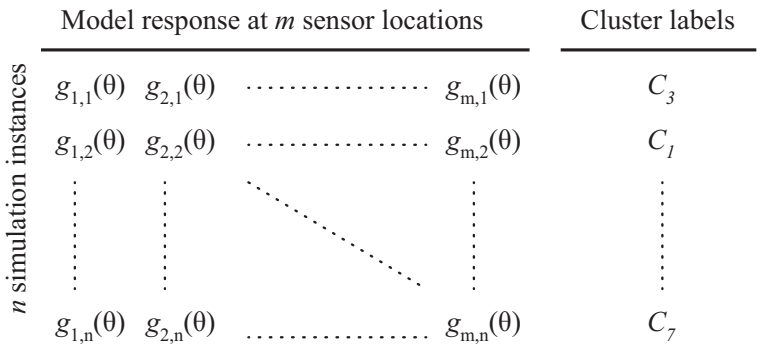


Fig. 2. Reducing the dimension of the model response dataset using k-means clustering.

In Figure 2, model response at sensor locations for each combination of parameter values is labeled belonging to a cluster. Each cluster in this dataset represents a different group of response-value sets of quantities that sensors measure (strains, deflections, etc.). Changes in values of important parameters leads to different groups of response-value sets. On the contrary, changes in values of unimportant parameters does not lead to any changes in the groups of value sets. While these clusters in model response represent changes in structural behavior they provide no information regarding the parameters that are governing these changes.

Inspection of parameters that are governing changes to structural behavior and thereby the clustering is carried out using support-vector machine (SVM) classification. Cluster labels, as shown in Figure 2, are utilized to train a SVM classifier and help determine the parameters that govern the underlying trends in the dataset, **D**. The dataset used for classification after clustering is shown in Figure 3 for n sets of parameter values (p parameters) and seven clusters. In this dataset, the cluster labels are responses and model parameters are the features to train the SVM classifier. The cluster

labels attributed to each set of model parameter value sets are obtained using k-means clustering. The cluster labels are not ordered, rather they correspond to the set of model parameter value sets provided as input for simulation using the physics-based model class.

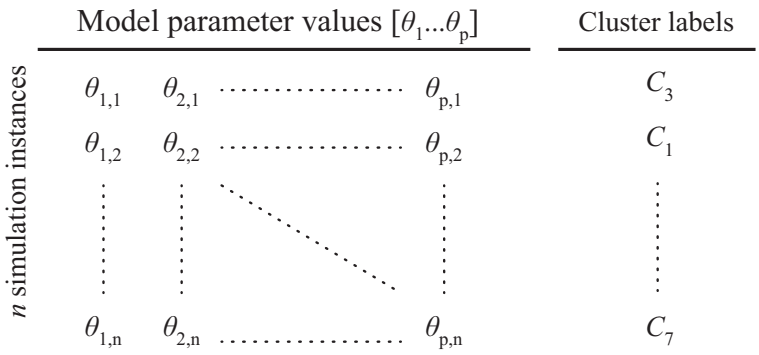


Fig. 3. Dataset for forward variable-search using SVM classification.

Let the SVM classifier be h_{svm}, which is defined by a model class $\mathcal{M}_j$, as shown in Eq. 1.

$$C_r = h_{svm,\mathcal{M}_j}(\boldsymbol{\theta}) + \epsilon_{class,\mathcal{M}_j} \qquad r \in [1, k] \tag{1}$$

In Eq. 1, the classifier, h_{svm}, is trained to predict the cluster labels, C_r for $r \in [1, k]$. The parameters included in training the classifier are defined by the model class, $\mathcal{M}_j$. In model class, $\mathcal{M}_j$, j denotes the size of the model class, *i.e.*, number of parameters in the model class, which varies from from 1 to p. For example, model class $\mathcal{M}_1$ contains one parameter from p parameters used to define the physics-based model $\boldsymbol{g}(\boldsymbol{\theta})$. The model class, $\mathcal{M}_j$, that provides least classification error, $\epsilon_{class,\mathcal{M}_j}$, is the model class that best explains underlying trends in the model response. Therefore, this model class is most suitable for use in structural identification. A key criteria is thus estimation of the classification error. In this study, the classification error is estimated using k-fold validation.

In k-fold validation, the dataset $\mathbf{D}$ of n simulation instances is separated (or folded) into r sub-datasets, where r is the number of folds provided by the user, as shown in

Figure 4. Each sub-dataset obtained from folding contains n/r samples. Each fold (subset of data) is held out, while the classifier is trained using samples from the other folds. The fold not included for training is used to validate the classifier. This process is then repeated with leaving out other folds and the average error over all folds is calculated to train the classifier by tuning the hyper-parameters. Moreover, this classification error is also used to select features to be used for training while performing forward-variable selection.

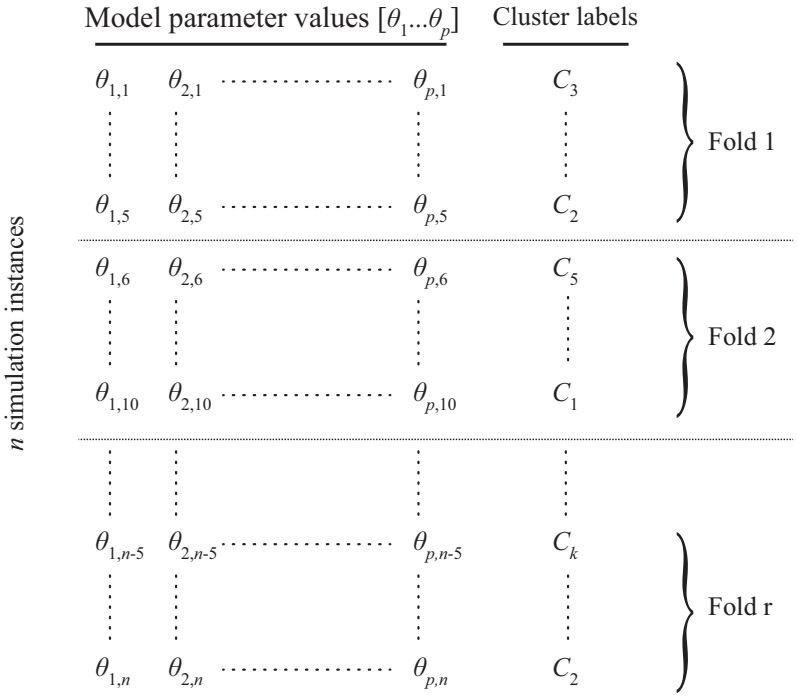


Fig. 4. Subsets for k-fold cross-validation for classification using dataset **D**. These subsets are used for calculating the classification error.

In Figure 4, if $r = 3$, then the dataset is separated into three sets. Two sets (folds) are used in training the classifier and the third set is used to validate predictions obtained using the classifier. This process is iterated with each set being considered as the validation set while the classifier is trained with the remaining two sets. Few of the

model parameters, $[\theta_1, ..., \theta_2]$, are features used to train the classifier to predict the labels, C_r, for $r \in [1, k]$. The mean of classification error obtained using the three iterations is the k-fold loss or classification error. For r folds, the classification error is calculated as shown in Eq. 2.

$$\epsilon_{svm,\mathcal{M}_j} = \sum_{i=1}^{r} \frac{N_{err,i}}{rN_i} \tag{2}$$

In Eq. 2, $\epsilon_{svm,\mathcal{M}_j}$ is the classification error obtained using k-fold validation for model class $\mathcal{M}_j$. For r folds, N_i is the number of instances in the validation fold, i and $N_{err,i}$ is the number of instances in fold i that are mis-classified.

Selection of an optimal model class based on the k-fold classification error is carried out using forward variable selection. In forward variable search, initially no parameters are included in the model class. At each iteration of the search a new parameter is included in the model class in order to reduce the classification error. So in forward variable selection, the selection starts with a model class of size 1, *i.e.*, $\mathcal{M}_1$ and concludes when all parameters are included in the model class, *i.e.*, $\mathcal{M}_p$. Initially, a model class with few parameters under-fits the data leading to a large classification error that reduces gradually as more parameters are included in the model class until an optimal model class is reached. Subsequent addition of model parameters leads to over-fitting and increase in classification error.

For p parameters that define the physics-based model, there are many possible combinations of model parameters that could possibly be used to define the classifier. To reduce the possible combinations and obtain an optimal model class efficiently, a sequential search is employed. A heuristic employed to speed-up the search is that parameters which are individually informative (independent, without information from other parameters) for classification are more likely to be a part of the optimal solution. To aid in development of this heuristic, k-fold classification error is used as a metric to first rank individual parameters on their importance in developing the classifier. Each parameter is used independently to train a classifier and the classification error for each classifier is recorded as shown in Eq.3, which is then used to determine the optimal model class with one parameter.

$$C_r = h_{svm,\mathcal{M}_{1,i}}(\boldsymbol{\theta}) + \epsilon_{class,\mathcal{M}_{1,i}} \qquad r \in [1, k] \qquad i \in [1, p] \tag{3}$$

In Eq. 3, C_r are the cluster labels predicted using the classifier $h_{svm,\mathcal{M}_{1,i}}(\theta_i)$, where

the classifier is trained using only one parameter θ_i. For the p classifiers developed, the classification error $\epsilon_{class,\mathcal{M}_{1,i}}$ is used as a metric to determine the relevance of the parameter θ_i to classify and differentiate between the forms of structural behavior defined by cluster labels, C_r. Parameters are then ranked based on the k-fold classification error. An example for such parameter ranking based on k-fold validation error is shown in Figure 5.

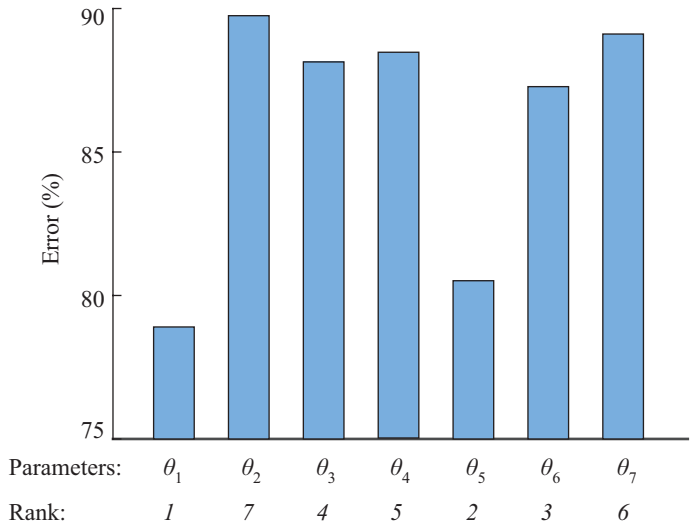


Fig. 5. The ranking of parameters based on their importance to classify trends in simulated dataset. The larger the classification error for a parameter, the less information the parameter contains about the clusters in the model-response data. In the figure, parameter θ_1 has the lowest classification error obtained using k-fold validation and thus contains the most information about clusters in the model-response data.

In Figure 5, the classification error associated with only using parameter θ_1 is the lowest among the six parameters evaluated. Thus, parameter, θ_1, is the most informative and ranked as 1. Similarly, parameters θ_2 to θ_7 are ranked based on the classification error obtained using k-fold validation.

For forward variable selection, the first parameter included in the model class is the parameter ranked 1, which is θ_1. Subsequently, parameters are added to the model class

to train the classifier, h_{svm} in an incremental manner. For example, while increasing model class size from 1 to 2 parameters, the parameter θ_1 is retained. One among parameters θ_2 to θ_7 is included to minimize the classification error. The order of evaluation of parameters is determined by their ranking as shown in Figure 5.

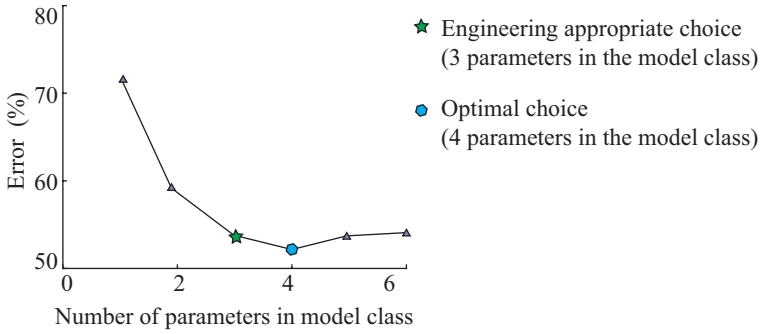


Fig. 6. Schematic representation of a trade-off curve obtained using forward variable selection for selecting parameters for identification. Based on this curve, the optimal model class is the one with four parameters. An engineering appropriate choice would be the model class with three parameters because fewer parameters for identification leads to lower computational cost with a negligible increase in error. The trade-off curve as shown provides engineers with guidance in selecting an appropriate model class for structural identification.

Sequentially adding parameters to the model class and evaluating the classification error at each iteration leads to a trade-off curve as shown in Figure 6. The variable selection procedure is continued till all parameters are included in the model class. In Figure 6, a schematic representation of such a trade-off curve is shown. Using forward-variable search, effect of parameter compensation is overcome to select all parameters important for classification. The trade-off curve shows decrease in classification error as model class size is increased till an optimum is reached ($\mathcal{M}_4$). Further addition of parameters leads to over-fitting and thus increase in classification error (or no change in certain cases).

It is critical to tune the classifiers developed using the k-fold cross-validation method during the process of forward-variable selection. SVM classifiers have hyper-parameters

such as box constraints and kernel scaling. Tuning the hyper-parameters results in an optimal classifier for each model-class instance while employing forward-variable search. For the case study, described in the next section, the hyper-parameters have been tuned using Bayesian optimization (Snoek et al. 2012) involving k-fold cross-validation.

An optimal model class is not necessarily the best choice from an engineering perspective. While the trade-off curve can be used directly to obtain the optimal model class, it is important to make a decision based on practical limitations. For example, in Figure 6, while $\mathcal{M}_4$ is the optimal model class, $\mathcal{M}_3$ provides similar classification error with a smaller model class size. As increase in model class size increases the computational cost of structural identification, a practical near-optimal choice is $\mathcal{M}_3$. Therefore, the trade-off curve provides the engineer with support to select a model class considering available computational resources.

The development of the model class selection method is independent of the structural identification methodology that is chosen. A key consideration in determining useful parameters is the prior distribution of model parameters. Appropriate estimations of prior distributions of model parameters is necessary for successful model-class selection and subsequent structural identification. This is not only relevant for the proposed model-class selection methodology, but for all other a-priori model-class selection and structural identification methodologies available in literature. A methodology for structural identification of civil infrastructure typically has requirements of accuracy and ease-of-use. The case-study that is described later in the paper is evaluated using EDMF for structural identification. In the next section, a brief explanation of this methodology is provided.

Background - Error-domain model falsification

EDMF is a data-interpretation methodology that was developed by Goulet and Smith (2013) for identification tasks where modeling uncertainties are significant. It has been shown to provide accurate identification compared with Bayesian model updating and residual minimization due to its robustness to correlation assumptions and explicit estimation of model bias based on engineering heuristics (Goulet and Smith 2013; Pasquier and Smith 2015; Reuland et al. 2017a; Pai et al. 2018). It is compatible with the assertion by Popper (1959) that models cannot be validated by data; they can only be falsified. Parameter values are provided as input to the physics-based model, which simulates structural response. Parameter instances that provide responses, $\boldsymbol{g}(\boldsymbol{\theta})$, incompatible with measurements, $\boldsymbol{y}$, are rejected by EDMF. Parameter values that lead to

model responses that are close to measurements are accepted into a candidate model set (CMS). Probabilistically-defined thresholds are used to determine closeness between model responses and measurements. Parameter instances within CMS are further utilized to make predictions under other load scenarios to assist in asset management.

The criteria for closeness is determined using the uncertainty associated with the model and measurements. Let $\epsilon_{mod,q}$ be the modeling uncertainty and $\epsilon_{meas,q}$ the measurement uncertainty, both at a measurement location q. The true response of the structure, R_q, at a measurement location, q, is given by Eq. 4.

$$y_q + \epsilon_{meas,q} = R_q = g_q\left(\boldsymbol{\theta}^*\right) + \epsilon_{mod,q}, \tag{4}$$

In Eq. 4, $g_q\left(\boldsymbol{\theta}^*\right)$ is the model response at a measurement location q for the real values of the model parameters, $\boldsymbol{\theta}^*$. y_q is the measured response of the structure at measurement location q. Structural response, measured using sensors, y_q, may be affected from environmental and operation conditions as well. When these factors are not explicitly included in the model, model predictions, $g_q\left(\boldsymbol{\theta}^*\right)$, may be biased from real structural response. This bias should be included in the modeling uncertainty, $\epsilon_{mod,q}$, in addition to other sources of modeling uncertainty.

Rearranging the terms in Eq. 4, the residual between model response, $g_q(\theta)$ and measurement, y_q, is equal to a combination of uncertainties, $\epsilon_{meas,q}$ and $\epsilon_{mod,q}$ at location q, as shown by Eq. 5.

$$g_q\left(\boldsymbol{\theta}^*\right) - y_q = \epsilon_{meas,q} - \epsilon_{mod,q}, \tag{5}$$

The criteria for compatibility are thresholds, $T_{high,q}$ and $T_{low,q}$, which are calculated using the combined uncertainty. To calculate these thresholds, in line with engineering practice, first a target reliability of identification, ϕ, is chosen. Based on this ϕ, the thresholds $T_{high,q}$ and $T_{low,q}$, are computed using Eq. 6.

$$\phi^{1/m} = \int_{T_{low,q}}^{T_{high,q}} f_{U_{c,q}}\left(\epsilon_{c,q}\right) d\epsilon_{c,q}. \tag{6}$$

In Eq. 6, $f_{U_{c,i}}\left(\epsilon_{c,q}\right)$ is the PDF of combined uncertainty at measurement location q and ϕ is the target reliability of identification. Thresholds, $T_{high,q}$ and $T_{low,q}$, correspond to the shortest interval providing a probability equal to target reliability, ϕ. In Eq. 6, the term $1/m$ is the Šidák correction (Šidák 1967), which accounts for m independent

measurements used in identification of model parameters. The Šidák correction controls the error rate such that the possibility of rejecting the true parameter values is lower than $1 - \phi$.

Residuals between model responses, $\boldsymbol{g}(\boldsymbol{\theta})$, and measurements, $\boldsymbol{y}$, are compared with the thresholds, $T_{low,q}$ and $T_{high,q}$. If this residual between model response and measurements lies within the thresholds for all measurement locations, then the model instance is accepted. This criteria for falsification is shown in Eq. 7.

$$T_{low,q} \leq g_q(\boldsymbol{\theta}) - y_q \leq T_{high,q} \qquad q \in \{1 \ldots m\}. \tag{7}$$

If predictions for a model instance, $\boldsymbol{\theta}_i$, does not satisfy Eq. 7 for any measurement location, then that model instance is falsified. All candidate model instances are considered equally likely and thus, assigned a uniform probability density (Goulet 2012; Pasquier 2015; Goulet and Smith 2013; Pasquier and Smith 2015). EDMF is traditionally carried out using grid sampling. In grid sampling, samples from prior distribution of model parameters, $\boldsymbol{\theta}$, are drawn. If n_s samples are drawn from the prior distribution of each parameter, then all possible combinations of these samples constitute a grid, which is called the initial model set (IMS). For n_s samples drawn from n_p parameters, the total number of model instances in IMS is $n_s^{n_p}$. Therefore, increase in number of parameters to be identified (n_p) exponentially increases the computational cost associated with identification.

Candidate models identified are used for making further predictions using the physics-based model with reduced parametric uncertainty (Pasquier and Smith 2015). The EDMF methodology has been applied to more than 20 full-scale systems since 1998 (Smith 2016). Recent applications include: model identification (Pai et al. 2018); leak detection in water supply (Moser et al. 2015); wind simulation (Vernay et al. 2015); fatigue life evaluation (Pasquier et al. 2014; Pasquier et al. 2016; Pai and Smith 2017); measurement-system design (Papadopoulou et al. 2015; Papadopoulou et al. 2016); post-earthquake assessment (Reuland et al. 2019); damage localization in tensegrity structures (Sychterz and Smith 2018); and occupant localization (Reuland et al. 2017b; Drira et al. 2019).

In order to compare structural identification results obtained with different model classes and data-interpretation methodologies a validation strategy is required. Pai et al. (2019) suggested the use of leave-one-out cross-validation strategy to assess accuracy and precision of structural identification. For the purpose of this paper, accuracy

indicates whether or not solutions obtained using structural identification are correct. Accuracy of solutions depends upon factors such as the selection of an appropriate model class for identification, accurate estimation of uncertainties affecting the task of identification and detection of outlier measurements. While this paper focuses on a method for selecting an appropriate model class for identification, other aspects such as estimation of uncertainties (Goulet et al. 2010), outlier detection (Proverbio et al. 2018) etc. have been studied by other researchers for EDMF applications.

Precision, on the other hand, is a measure of variability (uncertainty) in solutions. While accuracy is a necessary condition, precision of structural identification depends upon the model class, the uncertainty associated with the model and measurements and the information gained from measurements. Therefore, the precision metric is a useful measure for assessing the identifiability of a model classes after structural identification using EDMF.

In leave-one-out cross validation, observation from one sensor is omitted and structural identification is carried out using all remaining measurements. Updated model-parameter values, $\boldsymbol{\theta}_{\mathbf{CMS}}$ are then used to predict the model response at the omitted sensor. If the omitted measurement, y_j lies within the updated bounds of model predictions, $\boldsymbol{g_j}(\boldsymbol{\theta}_{\mathbf{CMS}})$, then structural identification is deemed to be accurate for that sensor location. This procedure is repeated by omitting each sensor separately in order to assess accuracy at all measurement locations. Structural identification carried out is accurate when the condition shown in Eq. 8 is satisfied.

$$\min\left(\boldsymbol{g_j}(\boldsymbol{\theta}_{\mathbf{CMS}})\right) \leq y_j \leq \max\left(\boldsymbol{g_j}(\boldsymbol{\theta}_{\mathbf{CMS}})\right) \quad \forall j \in [1, ..., m] \tag{8}$$

In Eq. 8, $\min\left(\boldsymbol{g_j}(\boldsymbol{\theta}_{\mathbf{CMS}})\right)$ and $\max\left(\boldsymbol{g_j}(\boldsymbol{\theta}_{\mathbf{CMS}})\right)$ are the bounds of updated model-predictions. The criteria for accuracy is a check whether the bounds of updated model-predictions distribution includes the measurement value, y_j. While Eq. 8 presents the condition for accuracy, it does not provide a metric to assess the information gained from measurements, which is precision.

Precision, using leave-one-out cross-validation, is evaluated by comparing the reduction in prediction-error range after identification. Prediction error is the difference between measurement and the predicted range of model response at a measurement location. Let j be the measurement omitted during structural identification. Taking into account the prior variability in model parameters, the range for model-prediction error at sensor location j is $\Re_j$. After including information from measurements, the

model parameter values are updated to reduce their variability. Utilizing these update parameter values, the range for model-prediction error at sensor location j is $\Re''_j$. Similar values of prediction-error ranges before and after identification are calculated by omitting other measurements. Precision, φ, calculated as shown in Eq. 9.

$$\varphi = \frac{(\mu_{\Re} - \mu_{\Re''})}{\mu_{\Re}} \tag{9}$$

In Eq. 9, $\mu_{\Re}$ and $\mu_{\Re''}$ are the mean values of prediction-error ranges, before and after structural identification ($\Re_j$ and $\Re''_j$ for $j \in [1, ..., m]$). Precision, φ, represents the proportion of reduction in prediction error after model updating. Precision, φ varies from 0 to 1. Precision, φ, equal to zero implies no gain in information from model updating, while φ equal to one implies perfect model updating wherein updated parameter distributions have zero variability. Therefore, precision is a metric that helps assess information gained from measurements in identifying a model class, *i.e.*, identifiability. In this way, using leave-one-out cross-validation, accuracy and precision of structural identification can be evaluated to enable comparisons between model classes and data-interpretation methodologies. In the next section, model-class selection and EDMF is evaluated on a full-scale case study.

CASE STUDY: POWDER MILL BRIDGE

Structure description

The model-based clustering methodology explained earlier is compared with traditional sensitivity analysis for structural identification using EDMF. The case-study used for comparison is the Powder Mill Bridge (PMB) in USA. The PMB, shown in Figure 7, is a steel-concrete bridge located over Ware river near the Powder Mill pond in Barre, Massachusetts, USA. The bridge was built in 2009 and connects the state highway 122 with a depot road that services mainly truck traffic for a waste management site. Using the PMB case-study, the model-class selection methodology described in this paper is compared with traditional sensitivity analysis for structural identification using two sensor configurations.

PMB, shown in Figure 7, is a three-span continuous steel-girder bridge in composite action with a reinforced concrete deck. A schematic drawing of this bridge is shown in Figure 8. The bridge has a total span of 47m. The reinforced concrete deck is supported by six I-section steel girders.

(a) Truck over PMB (b) Side view of PMB

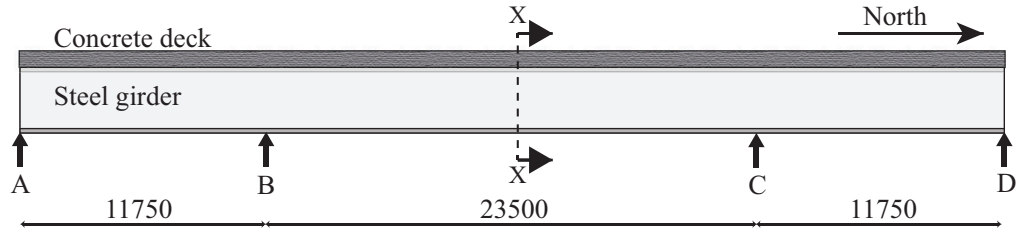
Fig. 7. Powder Mill Bridge (PMB) located in Massachusetts, USA.

The bridge has been instrumented to better understand its behavior and subsequently assess its remaining fatigue life (RFL). Prediction of RFL requires estimates of future traffic on the bridge related to loading and frequency. Damage incurred at a critical detail on the bridge for each cycle of loading can be evaluated with physics-based modeling. Cumulating this damage over possible future loads leads to prediction of the RFL. As knowledge of traffic on bridges is not typically available, AASHTO (2016) provides design loads and frequencies of traffic for fatigue life assessments.

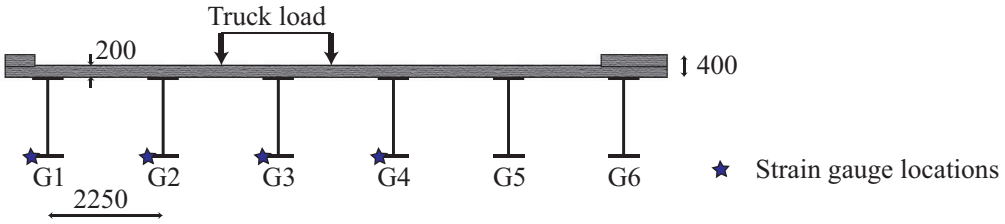
Interpreting monitoring data using EDMF helps improve understanding of structural behavior of the bridge through updating a physics-based model of the bridge. This physics-based model can then used to predict response of the bridge to design loads as specified in AASHTO (2016) for fatigue assessments. Using the response of the bridge, the fatigue damage incurred by the bridge at the critical detail is assessed and used to predict the RFL of the bridge.

Load test and monitoring

The bridge response to truck loading is measured with strain gauges during a load test. The sensors used are wired foil strain gauges, type KFG-5-120-C1-11L3M3R from the manufacturer Omega (Omega). The position of the load on the bridge, transversally, is shown in Figure 8 (b). The strain gauges on the bridge are located on the bottom flange of the steel girders. Location of all strain gauges is shown in Figure 9 (a). In



(a) Longitudinal schematic of Powder Mill Bridge

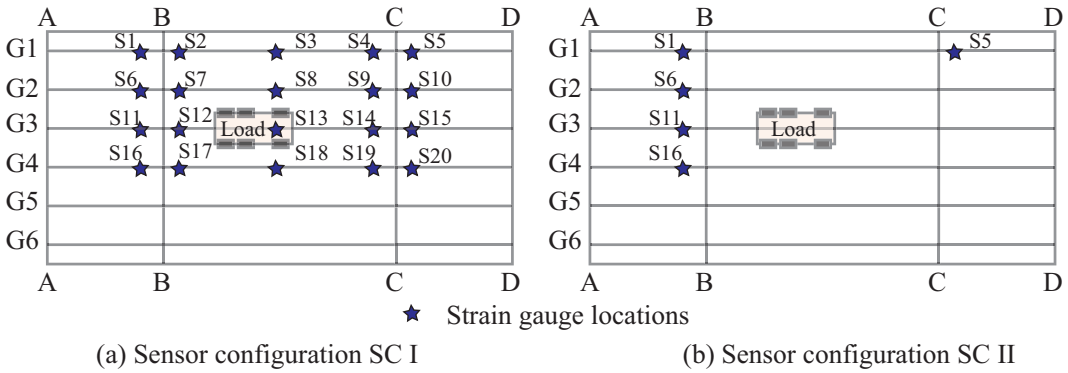


(b) Transverse schematic of Powder Mill Bridge at X-X (viewed South to North)

Fig. 8. Schematic drawing of Powder Mill Bridge. All dimensions are in millimeters (mm).

total, 20 strain gauges record structural response during the load test. This sensor configuration with 20 measurements is referred to as sensor configuration one (SC-I). To evaluate the utility of model-based clustering for model-class selection, a second sensor configuration, SC-II, which is a subset of SC-I, see Figure 9 (a), is utilized for structural identification. This sensor configuration is shown in 9 (b).

During the load test, a truck weighing 33 tonnes is driven across the bridge slowly (1.5 km/hr to 3 km/hr) to avoid dynamic amplification effects. Data recorded in the sensors is stored in data loggers installed on the bridge, which are shown in the inset of Figure 7 (b). Strain from all gauges recorded when the truck position leads to maximum strain in sensor SG13 (see Figure 9) is utilized for identification.



(a) Sensor configuration SC I (b) Sensor configuration SC II

Fig. 9. Sensor configurations, (a) Sensor configuration SC-I (20 sensors), (b) Sensor configuration SC-II (5 sensors)

Model development

The strain measurements are interpreted using a finite element (FE) model developed in ANSYS (ANSYS 2012). In the FE model, the concrete deck is modeled as homogeneous using SHELL182 elements. The sidewalk on the bridge and the railings contribute to bridge structural behavior (Sanayei et al. 2011). However, stiffness of the connection between the concrete deck and railings is not known. Thus, the deck slab thickness and thickness of the deck and railing at the edge of the bridge are parameterized in the FE model. The steel girders are modeled using SHELL182 elements. The composite action between the steel girders and concrete deck (in transversal and longitudinal directions) are modeled using zero-length spring elements (COMBIN14). The end supports of the bridge (abutments A and D) and intermediate supports (piers support B and C) are modeled with zero-length spring elements (COMBIN14) with parameterized stiffness in longitudinal and vertical directions. Springs of different supports are attributed different stiffness values. Table 1 shows the parameters included in the FE model and the prior distributions assumed for these parameters based on engineering heuristics.

Not all parameters shown in Table 1 have significant impact on the structural response at sensor locations. Only parameters that govern structural response at sensor locations can be identified using measurements. In the next few sections, traditional sensitivity

Table 1. Parametric sources of uncertainty in the model and their range

Index	Parameter	Variable	Range
1	Modulus of elasticity of concrete (GPa)	E_c	20-55
2	Modulus of elasticity of steel (GPa)	E_s	195-210
3	Thickness of deck slab (mm)	H_d	200-210
4	Height of concrete slab, sidewalk and railing (mm)	H_r	300-500
5	Deck-girder connection stiffness, transversal (log N/mm)	$K_{dg,x}$	2-6
6	Deck-girder connection stiffness, longitudinal (log N/mm)	$K_{dg,z}$	4-10
7	Vertical stiffness of abutment A (log N/mm)	$K_{1,y}$	4-7
8	Horizontal stiffness of abutment A (log N/mm)	$K_{1,z}$	2-5
9	Vertical stiffness of pier B (log N/mm)	$K_{2,y}$	4-7
10	Horizontal stiffness pier B (log N/mm)	$K_{2,z}$	2-5
11	Vertical stiffness of pier C (log N/mm)	$K_{3,y}$	4-7
12	Horizontal stiffness of pier C (log N/mm)	$K_{3,z}$	2-5
13	Vertical stiffness of abutment D (log N/mm)	$K_{4,y}$	4-7
14	Horizontal stiffness of abutment D (log N/mm)	$K_{4,z}$	2-5

analysis and model-based clustering are used to select a subset of parameters listed in Table 1 for structural identification using EDMF.

Model-class selection using sensitivity analysis

Traditionally, sensitivity analysis is conducted using linear regression. Dataset, **D**, is simulated using the FE model of the bridge, while accounting for variability in model parameter values as listed in Table 1. Using Latin-Hypercube sampling, 300 combinations of model-parameter values are generated to provide as input to the FE model of the bridge. Simulations using the FE model result in responses at the sensor locations corresponding to configurations, SC-I and SC-II. These simulations include uncertainty only from model parameters. The linear-regression model developed for each sensor location provides the sensitivity of model response to various model-parameter combinations. Sets of parameters that govern response may differ between sensor locations. The sensitivities obtained are generally averaged over all sensor locations. An arbitrary cut-off for the average relative importance, for instance 5%, is

then used to select parameters to be included in the model class for identification.

The averaged importance of model parameters over all sensor locations in sensor configuration, SC-I, is shown in Figure 10 (a). A cut-off of 5% is utilized and parameters whose importance is greater than 5% are included in the model class for identification. The parameters included in the model class for identification based on traditional sensitivity analysis for sensor configuration, SC-I, are E_c, E_s, H_r, $K_{dg,z}$, $K_{2,y}$ and $K_{3,y}$ (see Table 1). Similarly, the parameters to be included in the model class for identification for sensor configuration, SC-II, are E_c, E_s, H_r, $K_{dg,z}$, $K_{2,y}$ and $K_{3,y}$, as shown in Figure 10 (b). A change in sensor configuration, passing from 20 to 5 sensors, does not change the model class and has little impact on the sensitivities of model responses to model parameters. This implies that the six unknown parameters may be identified using 5 measurements from sensor configuration SC-II.

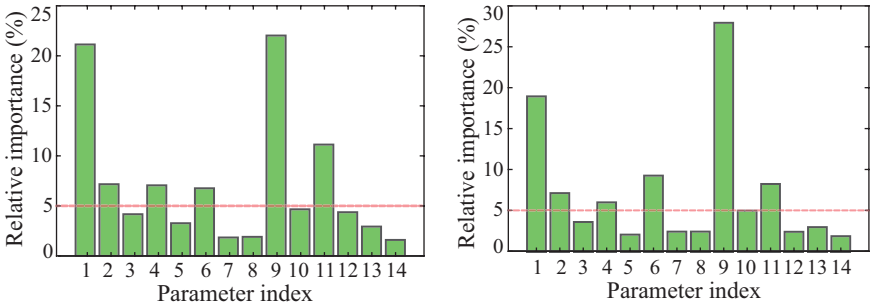


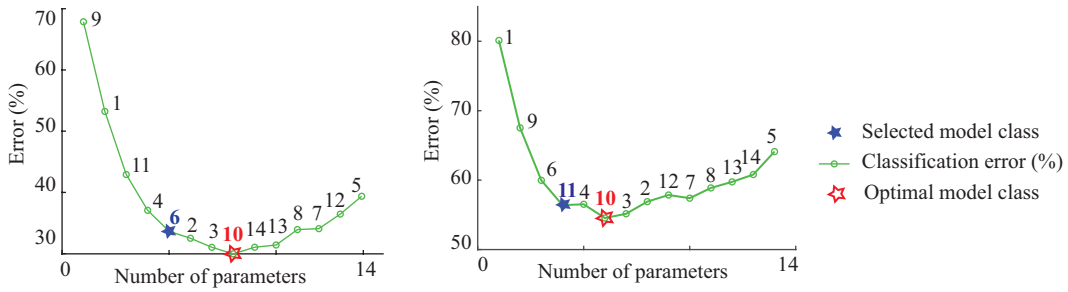
Fig. 10. Traditional sensitivity analysis using linear regression. (a) For sensor configuration, SC I (see Figure 9 (a)). The sensitivities have been averaged over the 20 sensors. (b) For sensor configuration, SC II (see Figure 9 (b)). The sensitivities have been averaged over the 5 sensors.

Model-class selection using the model-based clustering method

Instead of evaluating the importance of model parameters to model response at each sensor location, in model-based clustering method, an understanding of structural behavior is obtained with k-means clustering. The number of clusters in the dataset is evaluated using the gap method. In the gap method, the number of clusters is evaluated by comparing the probability that there are r clusters in the dataset compared with no

clusters in the dataset. The value of r is varied for the case study from 1 to 20 for model-class selection using both sensor configurations shown in Figure 9. For sensor configuration, SC-I, the optimal number of clusters is determined based on gap method to be 17, while for sensor configuration, SC-II, the optimal number clusters is estimated to be 10. These clusters represent significant changes in response of the model due to changes in input parameter values. In the next step of model-class selection, parameters that are important and cause these changes in model response are evaluated using classification.

Parameters that help distinguish between clusters that represent variations in structural behavior are evaluated using forward variable search with SVM classification as described in the previous section. For this case study, the SVM classifier is trained using a linear kernel after comparing classification error performance with other kernel types such as Gaussian and polynomial kernels. The hyper-parameters related to this kernel have been tuned using Bayesian optimization with a maximum of 30 iterations. The loss value (error) utilized for optimization is calculated with k-fold cross-validation. For this case study, 5 folds have been used to calculate the error. Tuning of the hyper-parameters leads to an optimal SVM model at each iteration within forward-variable selection. Subsequently, parameters are added to the SVM model to obtain a trade-off curve that shows the relationship between classification error and parameters within the model class. The trade-off curves obtained to assist in parameter selection for identification for sensor configurations, SC-I and SC-II, are shown in Figure 11.



(a) Sensor configuration SCI (b) Sensor configuration SCII

Fig. 11. Model selection for (a) for sensor configuration, SC I and (b) sensor configuration, SC II. See Figure 9 for layouts of sensor configurations.

In Figure 11 (a), the trade-off curve obtained for model-class selection using responses at sensor locations in SC-I is shown. The trade-off curve suggests the optimal model class includes 8 parameters, which is a computationally expensive model class. It can be observed that the gain in increasing model class size from 5 to 8 is not significant (less than 5% reduction in classification error). Therefore, an engineering near-optimal choice of model class for structural identification is the one with five parameters, namely E_c, H_r, $K_{dg,z}$, $K_{2,y}$ and $K_{3,y}$ (see Table 1). Similarly, for sensor configuration SC-II, a trade-off curve is obtained that helps select an engineering optimal model class with four parameters. These four parameters are E_c, $K_{dg,z}$, $K_{2,y}$ and $K_{3,y}$ (see Table 1).

For each of the two sensor configurations, SC-I and SC-II, two model classes, one from traditional sensitivity analysis and another from model-based clustering method are determined for identification using EDMF. These model classes are listed in Table 2.

Table 2. Model classes selected using linear regression and model-based clustering for sensor configurations, SC I and SC II (see Figure 9). Each model class represents a new case for structural identification using EDMF.

Case	SC	Method	Parameters included
C_1	I	Linear regression	E_c, E_s, H_r, $K_{dg,z}$, $K_{2,y}$, $K_{3,y}$
C_2	I	Model-based clustering	E_c, H_r, $K_{dg,z}$, $K_{2,y}$, $K_{3,y}$
C_3	II	Linear regression	E_c, E_s, H_r, $K_{dg,z}$, $K_{2,y}$, $K_{3,y}$
C_4	II	Model-based clustering	E_c, $K_{dg,z}$, $K_{2,y}$, $K_{3,y}$

Structural Identification

In Table 2, four cases of identification for PMB are shown. Utilizing a FE model to make the comparison between these cases, while feasible, is computationally expensive. Therefore, surrogate models are developed to replace the FE model, for the purpose of reducing computational cost in comparing precision of identification using the model classes shown in Table 2. Surrogate models, in this context, are regression models trained to simulate behavior of complex physics-based models. Good surrogate models replicate behavior of physics-based models within the domain of training data and quickly provide simulation results.

The surrogate-modeling strategy adopted for the PMB is Gaussian-process regression (Worden and Cross 2018) with an exponential kernel. Using the FE model, Gaussian-process regression models are developed to emulate structural responses at sensor locations. For each model class, as described in Table 2, a training dataset and validation dataset are simulated using the FE model. In the datasets, samples of parameters in the model class are drawn based on the prior parameter ranges. These samples are provided as inputs to the FE model to predict response at sensor locations. With the structural response as output and the parameters as input, Gaussian-process regression models are trained for each sensor location for all model classes. The error between the surrogate models and FE model response is quantified with the help of the validation dataset. The error obtained from validation is reported in Table 3. The choice of using Gaussian-process regression with an exponential kernel to develop the surrogate models was made after comparing the surrogate-modeling error with those obtained using other surrogate-modeling strategies such as support-vector regression (Chou and Pham 2017), neural networks (Chang et al. 2000) and linear regression (Friedman 1991).

The uncertainties associated with identification are shown in Table 3. Measurement uncertainty is estimated based on knowledge of sensors. Also, during the load-test, two strain gauges were placed close to one-another at all sensor locations shown in Figure 9 (a). Ideally, pairs of strain gauges should record the same reading due to their proximity. However, due to sensor noise and imperfections in sensor placement, the recordings are not necessarily the same. Accounting for this variability helps estimate the uncertainty associated with measurements. As two sensors in close proximity provide correlated and redundant information of structural behavior, only one of the two sensors is used for identification.

Load uncertainty includes uncertainty from magnitude of the truck load and uncertainty in its position on the bridge. This uncertainty has been estimated as uniform probability distribution whose bounds are calculated by propagating variability in loading conditions to uncertainty structural response at sensor locations. Model bias from simplifications is estimated to be larger close to supports than at mid-span of the bridge as assumptions made in modeling the supports and deck affect the sensors close to them to a larger magnitude. Sanayei et al. (2011) found that the deck close to supports may have hairline cracks in concrete in the negative bending moment zone, thus changing the moment of inertia close to the supports, which has not been included in the model developed for PMB in this paper. This is further justification for increasing modeling uncertainty at supports. Due to lack of knowledge of distributions related to modeling

uncertainty and bias, this uncertainty is represented as uniform probability distributions. Surrogate modeling uncertainty is calculated using hold-out cross-validation, **i.e.**, by comparing predictions of Gaussian process regression models with predictions obtained using the FE model of the bridge for input data used during training.

Table 3. Uncertainty sources and their distribution (%). Uncertainty from sources other than measurement are quantified as uniform probability distributions.

Source	**Distribution**
Measurement	N (0, 5)
Load	U (-5, 5)
Model bias (at sensors near supports)	U (-15, 5)
Model bias (at sensor near mid-span)	U (-7, 5)
Surrogate model uncertainty	U(-1,1)

Apart from the uncertainties mentioned in Table 3, parameters not included in the model class for identification contribute to the combined uncertainty. This uncertainty is called as the secondary-parameter uncertainty. This uncertainty is quantified as a uniform random variable. Bounds of this uniform distribution are calculated by evaluating variability in structural response at sensor locations when secondary-parameter values are varied. The secondary-parameter uncertainty associated with each model class is shown in Table 4. Uncertainties in Table 3 and Table 4 are quantified as uniform due to lack of complete knowledge of their distributions. As bounds of the distribution can be estimated, based on the principle of maximum entropy (Jaynes 1957), uniform distributions are appropriate for quantification. Uncertainty from all sources mentioned in Table 3 and 4 are combined together to determine falsification thresholds using Eq. 6.

Uncertainty from multiple sources (see Table 3 and Table 4) are combined using Monte Carlo sampling to obtain the combined uncertainty PDF. Using this combined uncertainty PDF, the falsification thresholds are determined using Eq. 6. The target reliability of identification used for EDMF in this paper is 0.95. As uncertainty is not the same for the four model classes (see Table 2 and Table 4), the falsification thresholds change as well. For each case of identification, an initial population of model instances

Table 4. Secondary-parameter uncertainty ($\mu\epsilon$) for the four model classes shown in Table 2. Uncertainty from secondary parameters is estimated to be uniform (U), whose bounds are provided in the table.

Model class	**Distribution**
C_1	U(-17, 20)
C_2	U(-25, 10)
C_3	U(-17, 20)
C_4	U(-25, 10)

is generated. 5 equidistant samples are drawn from the prior distribution of each model parameter to generate the initial grid of model instances. Thus, the number of model instances in the initial grid for identification cases C_1 to C_4 are 15625 (6 parameters, 5^6 model instances), 3125 (5 parameters, 5^5 model instances), 15625 (6 parameters, 5^6 model instances) and 625 (4 parameters, 5^4 model instances).

The population of initial model instances are evaluated with EDMF using falsification thresholds determined for each identification case. Candidate model instances are used for further predictions, such as RFL. The bounds of model parameters obtained with EDMF are shown in Table 5. Prior model-parameter ranges are also listed for comparison.

Table 5. Range of updated model parameters for the four cases (see Table 2)

Case	E_c (GPa)	E_s (GPa)	H_r (mm)	$K_{dg,z}$ (log N/mm)	$K_{2,y}$ (log N/mm)	$K_{3,y}$ (log N/mm)
Prior	20-55	195-210	300-500	4-10	4-7	4-7
C_1	46.25-55	195-210	350-500	5.5-10	4.75-7	4.75-7
C_2	46.25-55		300-500	5.5-10	4.75-7	4-7
C_3	20-55	195-210	300-500	4-10	4-7	4-7
C_4	20-55			5.5-10	4-7	4-7

In Table 5, the updated parameter ranges are similar to the prior distribution of model parameters. With EDMF, the bounds do not adequately represent the updated information acquired. Many model instances within these bounds are falsified as shown in Figure 12. This figure shows the candidate and falsified model instances obtained for identification case C_4. In the parallel axes plot, the first four vertical axes correspond to the model parameters and the next five correspond to the predictions at the five sensor locations utilized in identification. Each value on the first four vertical axis represent possible model parameter values. Thus, a line connecting values on the first four parallel axes represents a model parameter instance, which is provided as an input to surrogate models. The predictions obtained from the surrogate models for that model-parameter instance correspond to values on the next five vertical axes, one for each sensor. Therefore, a model instance and its predictions at sensor locations is represented by a line connecting values on vertical axis 1 to 9.

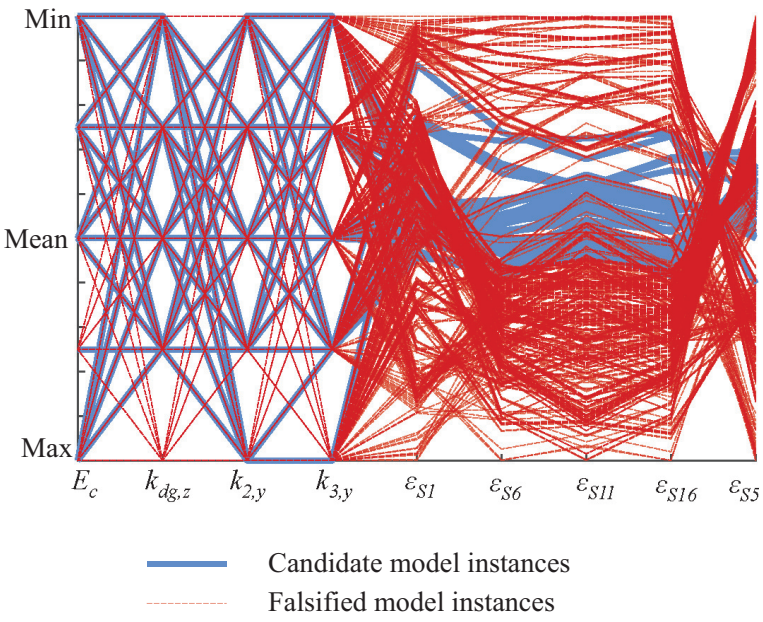


Fig. 12. Parallel-axes plot showing candidate and falsified model instances obtained by model falsification for case C_4 (see Table 2).

Before employing the candidate models obtained using EDMF for further predictions, leave-one-out cross-validation is conducted to assess accuracy and precision of identification. As an example, the estimation of accuracy for case C_4 is shown in Figure 13.

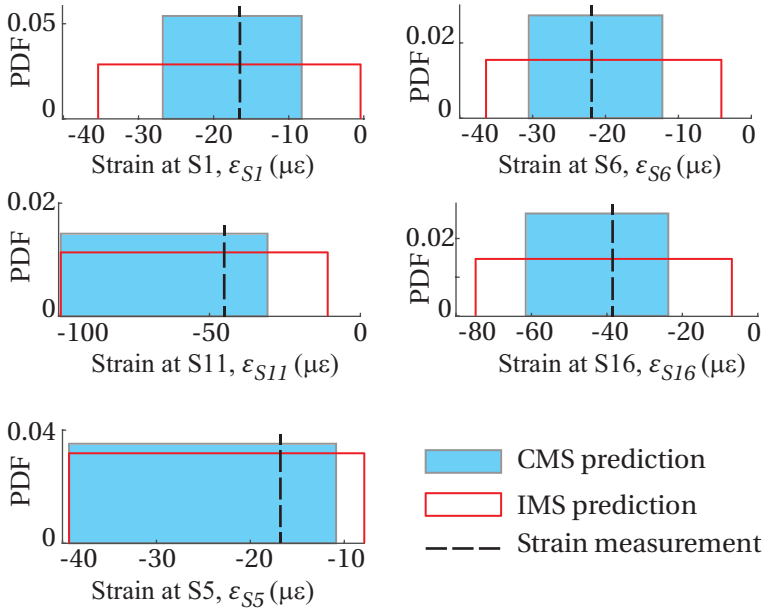


Fig. 13. Leave one out cross-validation prediction for case C_4 (see Table 2).

In Figure 13, updated prediction bounds include the measured strain response for all five sensor locations, which suggests identification using EDMF is accurate. Precision of identification is estimated using Eq. 9. Accuracy of structural identification is determined using the condition shown in Eq. 8. Results of leave-one-out cross-validation obtained for all cases of identification for PMB are summarised in Table 6. In Table 6, a checkmark for accuracy indicates that structural identification was deemed to be accurate at all sensor locations using Eq. 8.

In Table 6, precision of identification for C_1 is higher than for C_2. Thus, the model class obtained using traditional sensitivity analysis provides more precise model updating than the model class obtained using model-based clustering. However, this

Table 6. Accuracy and precision of structural identification for the four cases of PMB (see Table 2).

Case	Accuracy	Precision	Computational cost (relative to case C_4)	Comments
C_1	✓	0.54	25	Optimal
C_2	✓	0.43	5	Engineering choice
C_3	✓	0.29	25	Inefficient
C_4	✓	0.29	1	Fastest

improvement in precision is obtained with a more complex model class that requires four times more computation time. Using the model-based clustering approach, as shown in Figure 11 (a), an "engineering" near-optimal model class, C_2, is chosen to reduce computational cost. Thus, the model class obtained using model-based clustering method, while less precise, is more computationally efficient.

For sensor configuration SC-II, which is an irregular configuration of sensors, model class obtained with model-based clustering method, C_4, performs as well as the model class obtained using traditional sensitivity analysis, C_3. In case C_4, identification involves four parameters instead of six, thus reducing computational cost by 96% compared with case C_3.

Utilizing trade-off curves as shown in Figure 6 to select engineering near-optimal model classes significantly reduces computational cost, with little to no loss in precision of identification. Moreover, the use of leave-one–out cross validation helps assess accuracy of structural identification. As shown in Table 6, structural identification using EDMF is accurate with all four model classes. A comparison of precision of structural identification provides an a-posteriori estimate of the utility of chosen model-class for structural identification.

Remaining fatigue life prediction

Using the CMS obtained for the four cases of structural identification, reserve capacity of PMB is predicted with respect to its RFL. The critical detail evaluated for fatigue is a welded connection located on girder G2, close to north pier (near sensor S10,

see Figure 9), shown in Figure 8 (a). The category of this detail is 'C', which has a detail constant, A, of 44 ksi^3. This detail has also been evaluated by Saberi et al. (2016). The vehicular traffic on this bridge is mostly comprised of trucks, as the bridge connects the highway to a depot for waste management. Initial design for this bridge did not account for such heavy-load traffic and thus the fatigue life of this bridge has been investigated with in-service measurements as well. Based on in-service measurements carried out on the bridge, the average daily truck traffic (ADTT) is 255 vehicles/day. The RFL of PMB is predicted using the updated knowledge of structural behavior using the equation provided by AASHTO LFRD reference manual (AASHTO 2016), as shown in Eq. 10.

$$RFL = \frac{log\left[\frac{R_R \cdot A}{365 \cdot n \cdot ADTT \cdot [\Delta\sigma]^3} \cdot g(1+g)^{a-1} + 1\right]}{log(1+g)} \tag{10}$$

In Eq. 10, R_R is the resistance factor, which is equal to 1, A is the detail constant and n is the number of cycles per truck passage, equal to 2. In the equation, g is the annual growth of traffic in percentage, which is assumed to be 1% and a is the present age of the bridge, which during measurements was 11 years. $\Delta\sigma$ in the equation is the effective stress range. The effective stress range for PMB is computed using the FE model with the fatigue load as specified by the design code. Based on Eq. 10, the predictions of RFL are shown in Table 7. The reserve capacity in Table 7 is calculated as the percentage of RFL above design value, which is 64 years.

Table 7. Remaining fatigue life prediction. Design value of RFL is 64 years. All cases predict a reserve capacity of around 900%. Case C_4, with information from only 5 measurements, is able to provide this reserve capacity assessment with least computational cost compared with other model classes utilized for identification.

Case	RFL prediction using EDMF (years)	Reserve capacity (%)	Computational cost (relative to case C_4)
C_1	750-940	1070	25
C_2	750-1050	950	5
C_3	700-1000	915	25
C_4	730-980	960	1

The RFL calculated for PMB is at least 900 % above the design prediction. The results obtained are comparable with observations made using in-service measurements. Saberi et al. (2016) predicted the RFL of PMB bridge without accounting for growth in traffic as 1621 years. The in-service RFL prediction obtained is greater than the predictions obtained using the FE model. This is due to the fact that the sensor used to evaluate stress-ranges during in-service measurements is at a location further away from the critical detail. At this location, the stress-ranges obtained are lower than at the critical detail. Moreover, the RFL prediction with the FE model carried out for the PMB uses the design truck for fatigue evaluation, while in-service strain measurements take into account variability in truck weight, that may be lower than the design load.

The presence of significant RFL over design calculations is due to the conservative nature of modeling during the design phase followed by conservative construction practices. Over-design of civil infrastructure for safety is a common observation that has been made. Moreover, the activity of construction as well as processing of raw materials accounts for significant embodied energy in these over-designed structures. Better decisions can be made regarding existing structures once the reserve capacity due to over-design is quantified. Moreover, in the future, due to sustainability considerations and lack of availability of raw materials, better design guidelines may be necessary to minimize wastage of raw materials and reduce energy consumption. Model-based data-interpretation improves knowledge of behavior of existing structures and provides support in developing future design guidelines.

DISCUSSION

In this paper, a methodology for selecting parameters for structural identification has been presented. In this methodology, forward-variable selection is used to select parameters that are important to discriminate (classify) between forms of structural behavior estimated with the help of clustering This procedure for parameter selection for structural identification has been explained in detail in the methodology section of the paper.

The objective of clustering is to determine changes in model behavior as model-parameter values are varied. The data utilized for clustering is the model response at sensor locations obtained using either static or dynamics simulations, depending upon type of monitoring carried out. The model response is the same quantity that is being measured at each sensor location. This quantity could be strains, displacements, accelerations, mode shapes and natural frequencies. While clustering in this paper has

been carried out using the gap method (Tibshirani et al. 2001), other methods such as Hubert statistic (Halkidi et al. 2002), the Davis-Bouldin index (Davies and Bouldin 1979), score function index (Saitta et al. 2008) etc. may also be used.

Subsequent to clustering, in this paper, the parameters governing structural behavior are determined using forward-variable selection. Each cluster determined using k-means clustering is a different form of structural behavior. SVM classifiers are trained to discriminate between these clusters of structural behavior. Parameters (or features) that are most important in training this classifier are determined using forward-variable search. There are other classifier methods such as neural networks, k-nearest neighbors, logistic regression etc. that may also be employed to determine parameters governing structural behavior. Depending upon the case study, some methodologies may be more suitable than others. However, for parameter selection, the relative improvement in classification accuracy between model classes is a more important criteria than selecting the most appropriate strategy for classification. While making comparisons between model classes, the classifier used must be trained appropriately including any hyper-parameters and validated using either k-fold or hold-out validation methods. While using these validation methods, the sensitivity of results to data chosen for validation (such as number of folds in k-fold validation) should be evaluated. This is future work.

Also, the use of forward-variable selection to search for parameters that govern structural behavior makes it important to train the SVM classifier. However, forward-variable selection using classification error (Weston et al. 2001; Noori et al. 2011) is not the only method available. Variable selection has also been carried out using parameter weights (Guyon et al. 2002; Chang and Lin 2008) and backward searches (Balakrishnan et al. 2008).Other methods such as AIC (Akaike 1974), BIC (Schwarz et al. 1978) and regularization (Tikhonov 1963) may also be used for feature selection. Any of the aforementioned methods and other sensitivity-based methods may be used for selecting features of the classifier. The objective is to select select parameters that govern changes in structural behavior and therefore may be identifiable using measurements.

In this paper, forward-variable selection has been employed for feature selection because using this method a trade-off curve as shown in Figure 6 can be obtained. Rather than providing a single model class as a solution, a trade-off curve that is obtained using forward-variable selection provides engineers with support and options in selecting an appropriate model class for structural identification. This ensures that model-class selection is not carried out in an opaque manner, for example, using a black-box. In this paper, for forward-variable selection, the parameters are chosen

based on cross-validation error. Sensitivity-based methods could also be utilized to provide necessary heuristics to guide the search during forward-variable selection.

An aspect of this methodology that has not been addressed is its suitability for selecting an appropriate model class when measurements from multiple load-tests are available. Further research to extend this model-class selection methodology for structural identification using multiple load-tests is needed. One solution is to carry out feature selection for training the classifier with load tests as a feature that is already included. A set of features is thus selected that are identifiable using data from all load-tests.

Selection of model classes and subsequent structural identification are dependent upon the choice of prior distribution of model parameters. Quantifying uncertainties related to the model parameters requires prior understanding of the identification based on physical principles (Papaioannou and Straub 2015) and in-service conditions of the structure (Pasquier and Smith 2016). This is crucial for structural identification of civil infrastructure as each structure is unique in its form, function and utility. Moreover, using EDMF for structural identification efficiently falsifies wrong model classes when prior distributions estimates are inaccurate (Pasquier and Smith 2016). Therefore, any structural identification task has to be carried out in a transparent manner, taking into account the engineering knowledge that is available of the particular structure. Moreover, the task of structural identification should not be treated as a black box.

Use of data to enhance decision making helps ensure sustainable engineering. However, interpreting data in the presence of uncertainties is a complex and challenging task. Assessing accuracy of solutions obtained using structural identification is critical to ensure that decisions are made based on correct predictions. In this paper, a leave-one-out cross-validation method has been employed to assess accuracy of structural identification. Assessment of accuracy using leave-one-out cross-validation is necessary but not sufficient. More comprehensive strategies involving hold-out cross-validation are necessary to assess accuracy of structural identification. Moreover, to implement these strategies, measurement systems need to be designed so that the need for validation is part of the criteria.

In this paper, a model-class selection methodology has been developed and used to select an appropriate model class for structure identification of the Powder Mill Bridge. This novel methodology, compared with traditional sensitivity analysis, helps select identifiable and efficient model classes. Validation of structural identification results obtained using EDMF has been carried out using leave-one-out cross-validation. Using

leave-one-out cross-validation, the precision metric may also be calculated (see Eq. 9). The precision metric enables a-posteriori assessment of model classes for structural identification using EDMF.

After structural identification, using updated knowledge of structural behavior, the minimum remaining fatigue life of a critical detail on the PMB is estimated to be 700 years. This remaining fatigue-life is at least 900 % greater than the life estimated at the design stage. This shows that the bridge possesses significant reserve capacity compared with design calculations. This adds to the growing body of evidence that most structures are safe and possess significant reserve capacity above safety factors (Pasquier et al. 2014; Pasquier et al. 2016; Pasquier 2015; Pai et al. 2018; Pai et al. 2019; Brühwiler 2012) that may be utilized to enhance management actions when quantified (Smith 2016). Appropriate selection of model-classes for structural identification, as outlined in this paper, supports interpretation of monitoring data to improve knowledge of structural behavior and enhance decision making.

In the future, the model-class selection methodology described in this paper may be compared with other feature selection methodologies and other options for classification, clustering and model-class search. Also, the methodology may be extended to account for parameters important to predict the reserve capacity. Also, combining the task of model-class selection with sensor placement in an iterative framework improves the value of monitoring to support asset-management decision making.

CONCLUSION

In this paper, a new model-based clustering method for selecting parameters for structural identification is presented along with its comparison with traditional sensitivity analysis. The comparison has been made with application of both methods for parameter selection to a full-scale case study. Conclusions are as follows:

- The model-based clustering method helps select a computationally more efficient and identifiable model class compared with those obtained using traditional sensitivity analysis. Utilizing the trade-off curve to select an engineering optimal model class significantly reduces computational cost of two cases by 82% to 96%, with little to no loss in precision of identification.
- Use of leave-one-out cross-validation helps assess the accuracy of structural identification and enables a-posteriori assessment of information gain based on the precision metric. This helps compare the identifiability of model classes

determined using traditional sensitivity analysis and model-based clustering method.

- Subsequent to accurate structural identification, the reserve capacity of the bridge with respect to its minimum remaining-fatigue life is at least 900 % beyond design requirements. This result contributes to a growing body of evidence that bridge structures are over-designed with respect to code prescriptions for critical limit states. Model-based data interpretation helps quantify this over-design (reserve capacity), thereby enabling engineers to make better decisions regarding repair, retrofit and replacement actions.

ACKNOWLEDGEMENTS

This work was funded by the Swiss National Science Foundation under contract no. 200020-169026 and Singapore-ETH Centre (SEC) under contract no. FI 370074011-370074016. The authors would like to acknowledge B. Raphael for fruitful discussions related to development of the methodology in this paper.

DATA AVAILABILITY STATEMENT

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

AASHTO (2016). "Load and Resistance Factor Design (LRFD) for Highway Bridge Superstructures." *Report no.*, Federal Highway Administration, U.S. Department of transportation.

Akaike, H. (1974). "A new look at the statistical model identification." *IEEE transactions on automatic control*, 19(6), 716–723.

ANSYS (2012). "ANSYS Mechanical APDL Element Reference." *Report no.*

Argyris, C., Papadimitriou, C., and Panetsos, P. (2017). "Bayesian optimal sensor placement for modal identification of civil infrastructures." *Journal of Smart Cities*, 2(2), 69–86.

Balakrishnan, S., Narayanaswamy, R., Savarimuthu, N., and Samikannu, R. (2008). "Svm ranking with backward search for feature selection in type ii diabetes databases." *2008 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2628–2633.

Beck, J. L. and Katafygiotis, L. S. (1998). "Updating models and their uncertainties. I: Bayesian statistical framework." *Journal of Engineering Mechanics*, 124(4), 455–461.

Bennani, Y. and Cakmakov, D. (2002). *Feature selection for pattern recognition*. Informa Press, Bayreuth, Germany.

Bertola, N. J., Papadopoulou, M., Vernay, D., and Smith, I. F. (2017). "Optimal multi-type sensor placement for structural identification by static-load testing." *Sensors*, 17(12), 2904.

Biernacki, C., Celeux, G., and Govaert, G. (2000). "Assessing a mixture model for clustering with the integrated completed likelihood." *IEEE transactions on pattern analysis and machine intelligence*, 22(7), 719–725.

Brühwiler, E. (2012). "Extending the service life of swiss bridges of cultural value." *Proceedings of the Institution of Civil Engineers - Engineering History and Heritage*, 165(4), 235–240.

Caruana, R. and Freitag, D. (1994). "Greedy attribute selection." *Machine Learning Proceedings 1994*, Elsevier, 28–36.

Celeux, G., Maugis-Rabusseau, C., and Sedki, M. (2018). "Variable selection in model-based clustering and discriminant analysis with a regularization approach." *Advances in Data Analysis and Classification*, 1–20.

Chang, C.-C., Chang, T., and Xu, Y. (2000). "Adaptive neural networks for model updating of structures." *Smart materials and structures*, 9(1), 59.

Chang, Y.-W. and Lin, C.-J. (2008). "Feature ranking using linear svm." *Causation and Prediction Challenge*, 53–64.

Chou, J.-S. and Pham, A.-D. (2017). "Nature-inspired metaheuristic optimization in least squares support vector regression for obtaining bridge scour information." *Information Sciences*, 399, 64–80.

Cristianini, N., Shawe-Taylor, J., et al. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

Davies, D. L. and Bouldin, D. W. (1979). "A cluster separation measure." *IEEE transactions on pattern analysis and machine intelligence*, (2), 224–227.

Draper, N. R. and Smith, H. (2014). *Applied regression analysis*, Vol. 326. John Wiley & Sons.

Drira, S., Reuland, Y., Pai, S. G., Noh, H. Y., and Smith, I. F. (2019). "Model-based occupant tracking using slab-vibration measurements." *Frontiers in Built Environment*, 5, 63.

Dunn, J. C. (1974). "Well-separated clusters and optimal fuzzy partitions." *Journal of cybernetics*, 4(1), 95–104.

Fan, Y. and Tang, C. Y. (2013). "Tuning parameter selection in high dimensional penalized likelihood." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(3), 531–552.

Fisher, L. and Ness, J. W. V. (1971). "Admissible clustering procedures." *Biometrika*, 58(1), 91–104.

Fraley, C. and Raftery, A. E. (2002). "Model-based clustering, discriminant analysis, and density estimation." *Journal of the American statistical Association*, 97(458), 611–631.

Frangopol, D. M. and Soliman, M. (2016). "Life-cycle of structural systems: Recent achievements and future directions." *Structure and infrastructure engineering*, 12(1), 1–20.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, Vol. 1. Springer series in statistics New York, NY, USA:.

Friedman, J. H. (1991). "Multivariate Adaptive Regression Splines." *The Annals of Statistics*, 19(1), 1–67.

Goulet, J.-A. (2012). "Probabilistic Model Falsification for Infrastructure Diagnosis." Ph.D. thesis, Ph.D. thesis.

Goulet, J.-A., Kripakaran, P., and Smith, I. F. C. (2010). "Multimodel structural performance monitoring." *Journal of Structural Engineering*, 136(10), 1309–1318.

Goulet, J.-A. and Smith, I. F. C. (2013). "Structural identification with systematic errors and unknown uncertainty dependencies." *Computers & Structures*, 128, 251–258.

Guyon, I. and Elisseeff, A. (2006). "An introduction to feature extraction." *Feature extraction*, Springer, 1–25.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). "Gene selection for cancer classification using support vector machines." *Machine learning*, 46(1-3), 389–422.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002). "Clustering validity checking methods: part ii." *ACM Sigmod Record*, 31(3), 19–27.

Jain, A. K. (2010). "Data clustering: 50 years beyond k-means." *Pattern recognition letters*, 31(8), 651–666.

Jain, A. K., Duin, R. P., and Mao, J. (2000). "Statistical pattern recognition: A review." *IEEE Transactions on pattern analysis and machine intelligence*, 22(1), 4–37.

Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). "Data clustering: a review." *ACM computing surveys (CSUR)*, 31(3), 264–323.

Jaynes, E. T. (1957). “Information theory and statistical mechanics.” *Physical review*, 106(4), 620.

Johnson, S. C. (1967). “Hierarchical clustering schemes.” *Psychometrika*, 32(3), 241–254.

Kaufmann, L. and Rousseeuw, P. “Finding groups in data: An introduction to cluster analysis, 1990.

Liu, H. and Motoda, H. (1998). *Feature extraction, construction and selection: A data mining perspective*, Vol. 453. Springer Science & Business Media.

Ljung, L. (2010). “Perspectives on system identification.” *Annual Reviews in Control*, 34(1), 1–12.

Lynch, J. P. and Loh, K. J. (2006). “A summary review of wireless sensors and sensor networks for structural health monitoring.” *Shock and Vibration Digest*, 38(2), 91–130.

Mallows, C. L. (1973). “Some comments on c p.” *Technometrics*, 15(4), 661–675.

Matos, J. C., Cruz, P. J. S., Valente, I. B., Neves, L. C., and Moreira, V. N. (2016). “An innovative framework for probabilistic-based structural assessment with an application to existing reinforced concrete structures.” *Engineering Structures*, 111, 552–564.

Moser, G., Paal, S. G., and Smith, I. F. (2015). “Performance comparison of reduced models for leak detection in water distribution networks.” *Advanced Engineering Informatics*, 29(3), 714 – 726.

Muto, M. and Beck, J. L. (2008). “Bayesian updating and model class selection for hysteretic structural models using stochastic simulation.” *Journal of Vibration and Control*, 14(1-2), 7–34.

Noori, R., Karbassi, A., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M., Farokhnia, A., and Gousheh, M. G. (2011). “Assessment of input variables determination on the svm model performance using pca, gamma test, and forward selection techniques for monthly stream flow prediction.” *Journal of Hydrology*, 401(3-4), 177–189.

Omega. “Strain gage technical data, <https://www.omega.com/en-us/resources/strain-gauge-technical-data>. [2020.03.12].

Pai, S. G., Nussbaumer, A., and Smith, I. F. (2018). “Comparing structural identification methodologies for fatigue life prediction of a highway bridge.” *Frontiers in Built Environment*, 3, 73.

Pai, S. G., Reuland, Y., and Smith, I. F. (2019). “Data-interpretation methodologies for practical asset-management.” *Journal of Sensor and Actuator Networks*, 8(2), 36.

Pai, S. G. S. and Smith, I. F. C. (2017). "Comparing three methodologies for system identification and prediction." *14th International Probabilistic Workshop*, R. Caspeele, L. Taerwe, and D. Proske, eds., Springer International Publishing, 81–95.

Papadopoulou, M., Raphael, B., Smith, I. F., and Sekhar, C. (2016). "Evaluating predictive performance of sensor configurations in wind studies around buildings." *Advanced Engineering Informatics*, 30(2), 127–142.

Papadopoulou, M., Raphael, B., Smith, I. F. C., and Sekhar, C. (2015). "Optimal sensor placement for time-dependent systems: Application to wind studies around buildings." *Journal of Computing in Civil Engineering*, 4015024.

Papaioannou, I. and Straub, D. (2015). "Computing the reliability of shallow foundations with spatially distributed measurements." *Proc. 5th International Symposium on Geotechnical Safety and Risk, Rotterdam, Netherlands*.

Pasquier, R. (2015). "Performance Assessment and Prognosis for Civil Infrastructure Based on Model Falsification Reasoning." Ph.D. thesis, Ph.D. thesis.

Pasquier, R., D. Angelo, L., Goulet, J.-A., Acevedo, C., Nussbaumer, A., and Smith, I. F. C. (2016). "Measurement, Data Interpretation, and Uncertainty Propagation for Fatigue Assessments of Structures." *Journal of Bridge Engineering*, 21(5).

Pasquier, R., Goulet, J.-A., Acevedo, C., and Smith, I. F. C. (2014). "Improving Fatigue Evaluations of Structures Using In-Service Behavior Measurement Data." *Journal of Bridge Engineering*, 19(11), 4014045.

Pasquier, R. and Smith, I. F. C. (2015). "Robust system identification and model predictions in the presence of systematic uncertainty." *Advanced Engineering Informatics*, 29(4).

Pasquier, R. and Smith, I. F. C. (2016). "Iterative structural identification framework for evaluation of existing structures." *Engineering Structures*, 106, 179–194.

Popper, K. (1959). *The logic of scientific discovery*. Routledge.

Proverbio, M., Bertola, N. J., and Smith, I. F. (2018). "Outlier-detection methodology for structural identification using sparse static measurements." *Sensors*, 18(6), 1702.

Rao, N., Shah, P., and Wright, S. (2015). "Forward–backward greedy algorithms for atomic norm regularization." *IEEE Transactions on Signal Processing*, 63(21), 5798–5811.

Reuland, Y., Lestuzzi, P., and Smith, I. F. (2017a). "Data-interpretation methodologies for non-linear earthquake response predictions of damaged structures." *Frontiers in Built Environment*, 3, 43.

Reuland, Y., Lestuzzi, P., and Smith, I. F. (2019). "Measurement-based support for

post-earthquake assessment of buildings." *Structure and Infrastructure Engineering*, 1–16.

Reuland, Y., Pai, S. G., Drira, S., and Smith, I. F. (2017b). "Vibration-based occupant detection using a multiple-model approach." *IMAC XXXV -Structural Dynamics Challenges in Next Generation Aerospace Systems, Garden Grove, California*, Bethel, Connecticut, Society for Experimental Mechanics (SEM).

Saberi, M. R., Rahai, A. R., Sanayei, M., and Vogel, R. M. (2016). "Bridge fatigue service-life estimation using operational strain measurements." *Journal of Bridge Engineering*, 21(5), 1–11.

Saitta, S., Kripakaran, P., Raphael, B., and Smith, I. F. C. (2009). "Feature selection using stochastic search: An application to system identification." *Journal of Computing in Civil Engineering*, 24(1), 3–10.

Saitta, S., Raphael, B., and Smith, I. F. (2007). "A bounded index for cluster validity." *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, Springer, 174–187.

Saitta, S., Raphael, B., and Smith, I. F. C. (2008). "A comprehensive validity index for clustering." *Intelligent Data Analysis*, 12(6), 529–548.

Sanayei, M., Phelps, J. E., Sipple, J. D., Bell, E. S., and Brenner, B. R. (2011). "Instrumentation, nondestructive testing, and finite-element model updating for bridge evaluation using strain measurements." *Journal of bridge engineering*, 17(1), 130–138.

Schwarz, G. et al. (1978). "Estimating the dimension of a model." *The annals of statistics*, 6(2), 461–464.

Scrucca, L. (2016). "Genetic algorithms for subset selection in model-based clustering." *Unsupervised Learning Algorithms*, Springer, 55–70.

Scrucca, L. and Raftery, A. E. (2014). "clustvarsel: A package implementing variable selection for model-based clustering in r." *arXiv preprint arXiv:1411.0606*.

Šidák, Z. (1967). "Rectangular confidence regions for the means of multivariate normal distributions." *Journal of the American Statistical Association*, 62(318), 626–633.

Smith, I. F. C. (2016). "Studies of Sensor-Data Interpretation for Asset Management of the Built Environment." *Frontiers in Built Environment*, 2, 8.

Snoek, J., Larochelle, H., and Adams, R. P. (2012). "Practical bayesian optimization of machine learning algorithms." *Advances in neural information processing systems*, 2951–2959.

Sychterz, A. C. and Smith, I. F. (2018). "Using dynamic measurements to detect and

locate ruptured cables on a tensegrity structure." *Engineering Structures*, 173, 631–642.

Taylor, S. G., Raby, E. Y., Farinholt, K. M., Park, G., and Todd, M. D. (2016). "Active-sensing platform for structural health monitoring: Development and deployment." *Structural Health Monitoring*, 15(4), 413–422.

Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tibshirani, R., Walther, G., and Hastie, T. (2001). "Estimating the number of clusters in a data set via the gap statistic." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411–423.

Tikhonov, A. N. (1963). "On the solution of ill-posed problems and the method of regularization." *Doklady Akademii Nauk*, Vol. 151, Russian Academy of Sciences, 501–504.

Van Buren, K. L., Hall, T. M., Gonzales, L. M., Hemez, F. M., and Anton, S. R. (2015). "A case study to quantify prediction bounds caused by model-form uncertainty of a portal frame." *Mechanical Systems and Signal Processing*, 50, 11–26.

Van Buren, K. L., Mollineaux, M. G., Hemez, F. M., and Atamturktur, S. (2013). "Simulating the dynamics of wind turbine blades: part ii, model validation and uncertainty quantification." *Wind Energy*, 16(5), 741–758.

Vernay, D. G., Raphael, B., and Smith, I. F. C. (2015). "A model-based data-interpretation framework for improving wind predictions around buildings." *Journal of Wind Engineering and Industrial Aerodynamics*, 145, 219–228.

Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., and Vapnik, V. (2001). "Feature selection for svms." *Advances in neural information processing systems*, 668–674.

Worden, K. and Cross, E. J. (2018). "On switching response surface models, with applications to the structural health monitoring of bridges." *Mechanical Systems and Signal Processing*, 98, 139–156.

World Economic Forum (2014). "Strategic Infrastructure, Steps to operate and Maintain Infrastructure Efficiently and Effectively, <http://reports.weforum.org/strategic-infrastructure-2014/>.

Yuen, K.-V. and Kuok, S.-C. (2015). "Efficient Bayesian sensor placement algorithm for structural identification: a general approach for multi-type sensory systems." *Earthquake Engineering & Structural Dynamics*, 44(5), 757–774.