#51

Updating of a Model and its Uncertainties Utilizing Dynamic Test Data

J.L. Beck and L.S. Katafygiotis
California Institute of Technology, Pasadena, CA 91125

ABSTRACT

The problem of updating a structural model and its associated uncertainties by utilizing structural response data is addressed. Using a Bayesian probabilistic formulation, the updated "posterior" probability distribution of the uncertain parameters is obtained and it is found that for a large number of data points it is very peaked at some "optimal" values of the parameters. These optimal parameters can be obtained by minimizing a positive-definite measure-of-fit function. This paper focuses on the identifiability of the optimal parameters. The problem of finding the whole set of optimal models that have the same output at the observed degrees of freedom for a given input is resolved for the first time, by presenting an algorithm which methodically and efficiently searches the parameter space. Also, a simplified expression is given for the weighting coefficients associated with each optimal model which are involved in the probability distribution for the predicted response.

INTRODUCTION

The uncertainties encountered when modeling a structure with a model out of a specified class can be divided into uncertainties of the "model parameters", concerned with which model out of the specified class is the most appropriate to describe the system, and uncertainties of the parameters describing the "model error", which is the error arising because any mathematical model is only an approximation of the real dynamic behavior of the structure. For example, if we choose the class of linear dynamic models, then there are uncertainties in the values of the parameters, such as Young's modulus E, which should be chosen for the structure. Furthermore, for each model in the class, we know that the corresponding predicted response will not be identical to the actual structural response because of model error.

In order to properly describe the modeling uncertainties, a probabilistic formulation can be followed, but from the Bayesian point of view, that is, probability is viewed as a multi-valued logic for plausible reasoning. Note that for most of the applications of interest in this study, the common interpretation of probability as a relative frequency of occurences in the long run does not make sense. Therefore, in order to quantify the uncertainty associated with a parameter, a probability distribution will be assigned describing how plausible each value is for the parameter, on the basis of the given information.

During the preliminary design of a structure, when no records of structural response are available, the modeling uncertainties must be estimated subjectively, on the basis of any available information and experience dealing with similar structures [Katafygiotis and Beck 1991]. When records from dynamic testing or earthquake records of structural response become available, the information contained in these records can be used to update the initial estimates of the modeling uncertainties by applying Bayes' Theorem. The updated probability model can be used for response predictions, for improving the performance of a control system, or for health monitoring of the structure by detecting changes in its stiffness distribution.

For a large number of available data points, the posterior distribution of the uncertain parameters resulting from Bayes' Theorem is very peaked at some optimal values of the parameters, therefore making these values much more plausible than the other values. Beck [1990] showed that it is then asymptotically correct for response predictions to use only the models corresponding to these optimal parameters, appropriately weighted. This result is very important, since the high-dimensional integrations which are required to calculate the uncertainties in the predictive response are computationally prohibitive [Katafygiotis and Beck, 1991], but the asymptotic result implies that they can be replaced by a weighted sum over all optimal parameters, assuming their number is finite. However, the implementation of these asymptotic results requires that the problem of finding the set of all optimal parameters be solved.

It is usually the case in structural model updating that there is more than one optimal value of the parameters for given input and output data, so the concepts of model and system identifiability are introduced to provide a framework to handle this nonuniqueness in the optimal parameters. Also, a new algorithm is presented to methodically and efficiently search the high-dimensional model parameter space by following only a finite set of one-dimensional curves. This algorithm finds all other optimal model parameters corresponding to models with identical model response at the observed degrees of freedom, thus resolving the problem of model identifiability. This important problem of finding all optimal parameters corresponding to models which are "output-equivalent" is solved for the first time.

FORMULATION OF THE PROBLEM

Let $\mathcal{D}_N$ denote the set of observed data for a structural system, consisting of a sampled observed input history $\hat{Z}_{1,N} = \{\hat{\underline{z}}(n) \in R^{N_I} : n = 1, 2, \ldots, N\}$ and output history $\hat{X}_{1,N} = \{\hat{\underline{x}}(n) \in R^{N_o} : n = 1, 2, \ldots, N\}$. Usually, the measured output consists of the acceleration histories at certain degrees of freedom (dof), which are referred to as observed or measured dof. Let $\underline{z}(n) \in R^{N_I}$ and $\underline{x}(n; Z_{1,N}) \in R^{N_o}$ denote the vector of the corresponding *system* input and output at time $t_n = n\Delta t$, where Δt is the sampling interval for the data in $\mathcal{D}_N$. Assuming that for modern instrumentation the measurement noise is negligible compared with the model error, it follows that $\underline{x}(n) = \hat{\underline{x}}(n)$ and $\underline{z}(n) = \hat{\underline{z}}(n)$ for $n \leq N$. Assume that an N_d-degree of freedom theoretical model $\mathcal{M}$ has been chosen to describe the input-output behavior of the system, and let $\underline{a} \in R^{N_a}$ be the vector of the uncertain model parameters. $\mathcal{M}$ provides a functional relationship between the *model* output vector $\underline{q}(n;\underline{a}) \in R^{N_o}$ at time $t_n = n\Delta t$ and the system input $Z_{1,n}$:

$$\underline{q}(n;\underline{a}) = \underline{q}(n;\underline{a}, Z_{1,n}, \mathcal{M}) \tag{1}$$

In the following, the dependence of $\underline{q}(n;\underline{a})$ on the input $Z_{1,n}$ and the theoretical model $\mathcal{M}$ as well as the dependence of $\underline{x}(n)$ on the input $Z_{1,n}$ will be suppressed in the notation.

The model error $e(\underline{n})$ is defined to be the difference between the system output and the model output, so:

$$\underline{x}(n) = \underline{q}(n;\underline{a}) + \underline{e}(n) \tag{2}$$

In order to account for the model error, a class of probability models $\mathcal{P}$ is chosen which prescribes a function h_M giving the probability density function of the sequence $E_{1,M} = \{\underline{e}(\underline{n}); n = 1, \ldots, M\}$. The class $\mathcal{P}$ selected here assumes that $E_{1,M}$ is a zero-mean stationary Gaussian white-noise sequence. In addition, it assumes that the variance of the model error is the same at all observed degrees of freedom, and it is denoted by σ^2. Therefore, σ is the only model-error parameter required to specify a particular probability model out of the class $\mathcal{P}$ and h_M is:

$$\begin{aligned} p(E_{1,M}|\sigma,\mathcal{P}) &= h_M(E_{1,M};\sigma) \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{MN_o}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{M}\sum_{i=1}^{N_o} e_i(n)^2\right) \end{aligned} \tag{3}$$

The selection of the classes $\mathcal{M}$ and $\mathcal{P}$ defines the class $\mathcal{M}_P$ parameterized by $\underline{\tilde{a}} = [\underline{a}^T, \sigma]^T$ prescribing the pdf:

$$\begin{aligned} p(X_{1,M}|\underline{\tilde{a}}, Z_{1,M}, \mathcal{M}_P) &= f_M(X_{1,M};\underline{\tilde{a}}, Z_{1,M}) \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{MN_o}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{M}\sum_{i=1}^{N_o}(x_i(n) - q_i(n;\underline{a}))^2\right) \end{aligned} \tag{4}$$

In order to account for the uncertainty in the parameters $\underline{a}$ and σ, it is assumed that $\mathcal{M}_P$ also specifies a function $\pi_{\underline{a},\sigma}$ corresponding to their prior pdf:

$$p(\underline{a},\sigma|\mathcal{M}_P) = \pi_{\underline{a},\sigma}(\underline{a},\sigma) \equiv \pi_{\underline{\tilde{a}}}(\underline{\tilde{a}}) \tag{5}$$

The updated, or "posterior", joint pdf of $\underline{\tilde{a}}$ is given by Bayes' Theorem:

$$\begin{aligned} p(\underline{\tilde{a}}|\mathcal{D}_N, \mathcal{M}_P) &= \frac{p(\hat{X}_{1,N}|\underline{\tilde{a}}, \hat{Z}_{1,N}, \mathcal{M}_P)p(\underline{\tilde{a}}|\mathcal{M}_P)}{p(\hat{X}_{1,N}|\hat{Z}_{1,N}, \mathcal{M}_P)} \\ &= kp(\hat{X}_{1,N}|\underline{\tilde{a}}, \hat{Z}_{1,N}, \mathcal{M}_P)p(\underline{\tilde{a}}|\mathcal{M}_P) \\ &= kf_N(\hat{X}_{1,N};\underline{\tilde{a}}, \hat{Z}_{1,N})\pi_{\underline{\tilde{a}}}(\underline{\tilde{a}}) \end{aligned} \tag{6}$$

where

$$k^{-1} = p(\hat{X}_{1,N}|\hat{Z}_{1,N}, \mathcal{M}_P) = \int_{S(\underline{\tilde{a}})} p(\hat{X}_{1,N}|\underline{\tilde{a}}, \hat{Z}_{1,N}, \mathcal{M}_P)p(\underline{\tilde{a}}|\mathcal{M}_P)d\underline{\tilde{a}} \tag{7}$$

As can be seen from Equation (6), the effect of utilizing the available records to update the pdf of the model parameters $\underline{a}$ and the model-error parameter σ is contained in the term $k f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N})$, where k serves as a normalizing constant. The effect of the function f_N on the updated distribution is much more drastic than that of the prior distribution, since it can be shown that f_N is very peaked at some optimal values of the parameters, making these values much more probable than the other ones.

Define the *optimal* parameters $\underline{\hat{\tilde{a}}} = [\underline{\hat{a}}^T, \hat{\sigma}]^T$ to be the values of the parameters $\underline{\tilde{a}}$ that globally maximize $f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N})$. Maximizing $f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N})$ with respect to $\underline{\tilde{a}}$ is equivalent to maximizing:

$$\ln f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N}) = -c - N N_o \ln\sigma - \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{i=1}^{N_o} (\hat{x}_i(n) - q_i(n; \underline{a}))^2 \quad (8)$$

At $\underline{\tilde{a}} = \underline{\hat{\tilde{a}}}$ the following conditions hold:

$$\left. \frac{\partial \ln f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N})}{\partial \underline{\tilde{a}}} \right|_{\underline{\tilde{a}} = \underline{\hat{\tilde{a}}}} = 0 \quad (9)$$

For fixed $\underline{a}$, maximizing $\ln f_N(\hat{X}_{1,N}; \underline{\tilde{a}}, \hat{Z}_{1,N})$ with respect to σ requires:

$$\hat{\sigma}(\underline{a})^2 = \frac{1}{N N_o} \sum_{n=1}^{N} \sum_{i=1}^{N_o} (\hat{x}_i(n) - q_i(n; \underline{a}))^2 \quad (10)$$

This shows how the most probable variance $\hat{\sigma}(\underline{a})^2$, for given $\underline{a}$, depends on the choice of the model parameters $\underline{a}$. Obviously, the condition for the overall most probable variance $\hat{\sigma}^2$ is given by (10) when $\underline{a} = \underline{\hat{a}}$. Substituting Equation (10) into Equation (8):

$$\ln f_N(\hat{X}_{1,N}; \underline{a}, \hat{\sigma}(\underline{a}), \hat{Z}_{1,N}) = -c - N N_o \ln\hat{\sigma}(\underline{a}) - \frac{N N_o}{2} \quad (11)$$

Thus $\underline{\hat{a}}$ is given by minimizing $\hat{\sigma}(\underline{a})$ or, equivalently, minimizing:

$$J(\underline{a}) = N N_o \hat{\sigma}(\underline{a})^2 = \sum_{n=1}^{N} \sum_{i=1}^{N_o} (\hat{x}_i(n) - q_i(n; \underline{a}))^2 \quad (12)$$

Since $J(\underline{a})$ might attain its minimum at more than one value $\underline{\hat{a}}$, the identifiability of the optimal model parameters $\underline{\hat{a}}$ must be resolved. On the other hand, the optimal model error parameter $\hat{\sigma}$ given by:

$$\hat{\sigma} = \left(\frac{1}{N N_o} \min_{\underline{a} \in S(\underline{a})} J(\underline{a})\right)^{\frac{1}{2}} \quad (13)$$

is uniquely determined and is therefore said to be globally identifiable. Here $S(\underline{a})$ denotes the space of permissible values of $\underline{a}$. It can be shown that:

$$\frac{p(\underline{a},\hat{\sigma}(\underline{a})|\mathcal{D}_N,\mathcal{M}_P)}{p(\hat{\underline{a}},\hat{\sigma}(\hat{\underline{a}})|\mathcal{D}_N,\mathcal{M}_P)} = \frac{\pi_{\underline{a},\sigma}(\underline{a},\hat{\sigma}(\underline{a}))}{\pi_{\underline{a},\sigma}(\hat{\underline{a}},\hat{\sigma}(\hat{\underline{a}}))}\left(\frac{\hat{\sigma}(\hat{\underline{a}})}{\hat{\sigma}(\underline{a})}\right)^{NN_o} = \frac{\pi_{\underline{a},\sigma}(\underline{a},\hat{\sigma}(\underline{a}))}{\pi_{\underline{a},\sigma}(\hat{\underline{a}},\hat{\sigma}(\hat{\underline{a}}))}\left(\frac{J(\hat{\underline{a}})}{J(\underline{a})}\right)^{\frac{NN_o}{2}} \tag{14}$$

The posterior pdf $p(\underline{\tilde{a}}|\mathcal{D}_N,\mathcal{M}_P)$ can be approximated locally, in the neighborhood $\mathcal{H}(\underline{\tilde{a}};\hat{\underline{\tilde{a}}})$ of an optimal parameter $\hat{\underline{\tilde{a}}}$, with a multi-dimensional Gaussian distribution with mean $\hat{\underline{\tilde{a}}}$ and an $(N_o+1)\times(N_o+1)$ covariance matrix $A_N^{-1}(\hat{\underline{\tilde{a}}})$ [Beck 1990]:

$$p(\underline{\tilde{a}}|\mathcal{D}_N,\mathcal{M}_P) \simeq p(\hat{\underline{\tilde{a}}}|\mathcal{D}_N,\mathcal{M}_P)\exp(-\frac{1}{2}[\underline{\tilde{a}}-\hat{\underline{\tilde{a}}}]^T A_N(\hat{\underline{\tilde{a}}})[\underline{\tilde{a}}-\hat{\underline{\tilde{a}}}]) \; ; \; \underline{\tilde{a}} \in \mathcal{H}(\underline{\tilde{a}};\hat{\underline{\tilde{a}}}) \tag{15}$$

where the elements $[A_N(\hat{\underline{\tilde{a}}})]_{ij}$ are given by:

$$[A_N(\hat{\underline{\tilde{a}}})]_{ij} = -\left.\frac{\partial^2 \ln f_N(\hat{X}_{1,N};\underline{\tilde{a}},\hat{Z}_{1,N})}{\partial\tilde{a}_i\partial\tilde{a}_j}\right|_{\underline{\tilde{a}}=\hat{\underline{\tilde{a}}}} - \left.\frac{\partial^2 \ln \pi_{\underline{\tilde{a}}}(\underline{\tilde{a}})}{\partial\tilde{a}_i\partial\tilde{a}_j}\right|_{\underline{\tilde{a}}=\hat{\underline{\tilde{a}}}} \cong -\left.\frac{\partial^2 \ln f_N(\hat{X}_{1,N};\underline{\tilde{a}},\hat{Z}_{1,N})}{\partial\tilde{a}_i\partial\tilde{a}_j}\right|_{\underline{\tilde{a}}=\hat{\underline{\tilde{a}}}} \tag{16}$$

The elements of A_N are $\mathcal{O}(N)$ and, therefore, for a large number N of available data points, which is usually the case with dynamic tests or earthquake records of structural response, the pdf $p(\underline{\tilde{a}}|\mathcal{D}_N,\mathcal{M}_P)$ becomes very peaked at the optimal parameters $\hat{\underline{\tilde{a}}}$; this result can also be concluded by viewing Equation (14). It has been shown [Beck 1990], that if the number of optimal parameters is finite, it is asymptotically correct for prediction purposes to use out of the class $\mathcal{M}_P$ only the probability models corresponding to the optimal parameters $\hat{\underline{\tilde{a}}}_k$, $k=1,\ldots,K$, since for large N, the updated predictive distribution is given by:

$$p(X_{1,M}|\mathcal{D}_N,Z_{N+1,M},\mathcal{M}_P) \simeq \sum_{k=1}^{K} w_k p(X_{1,M}|\hat{\underline{\tilde{a}}}_k,Z_{1,M},\mathcal{M}_P) \tag{17}$$

The predictive distribution for each of the optimal models is weighted proportionally to the volume of the posterior pdf $p(\underline{\tilde{a}}|\mathcal{D}_N,\mathcal{M}_P)$ under its Gaussian-shaped peak positioned at the corresponding optimal parameters. The mathematical expression for the weighting coefficient w_k corresponding to the k^{th} vector of optimal parameters $\hat{\underline{\tilde{a}}}_k$ is [Beck 1990]:

$$w_k = \frac{w'_k}{\sum_{k=1}^{K} w'_k} \tag{18}$$

where

$$w'_k = \pi_{\underline{\tilde{a}}}(\hat{\underline{\tilde{a}}}_k)|A_N^{-1}(\hat{\underline{\tilde{a}}}_k)|^{1/2} \tag{19}$$

Notice that the prior pdf $\pi_{\hat{a}}(\hat{\underline{a}})$ does not need to be specified over the whole domain $S(\hat{\underline{a}})$. Instead, only the relative values for the optimal parameters $\hat{\hat{\underline{a}}}_k$ need to be specified. The elements of $A_N(\hat{\hat{\underline{a}}}_k)$ can be evaluated numerically through Equation (16). Numerical examples have shown that these calculations can be very sensitive to roundoff errors if the vector of the observed model parameters $\underline{a}$ does not consist of modal quantities exclusively. In addition, independent of the choice of $\underline{a}$, the matrix $A_N(\hat{\hat{\underline{a}}}_k)$ is often ill-conditioned, which results in numerical errors when calculating $|A_N^{-1}(\hat{\hat{\underline{a}}}_k)|$. Thus, in general, the weighting effect of $|A_N^{-1}(\hat{\hat{\underline{a}}}_k)|^{\frac{1}{2}}$ cannot be estimated reliably by calculating it directly. A reliable alternative expression is presented later to overcome this difficulty.

MODEL AND SYSTEM IDENTIFIABILITY OF THE OPTIMAL PARAMETERS

Let $Q_{1,N}(\underline{a};\hat{Z}_{1,N}) = \{\underline{q}(n;\underline{a},\hat{Z}_{1,N},\mathcal{M}) \in R^{N_o} : n = 1,2,\ldots,N\}$ denote the *model* output history, which corresponds to the observed quantities, for the given input $\hat{Z}_{1,N}$ and for a model $M(\underline{a}) \in \mathcal{M}$, and let $S(Q_{1,N};\hat{Z}_{1,N})$ denote the space formed by the range of $Q_{1,N}(\underline{a};\hat{Z}_{1,N})$ as $\underline{a}$ ranges over $S(\underline{a})$. There is a natural mapping of the models in the class $\mathcal{M}$ onto $S(Q_{1,N};\hat{Z}_{1,N})$, but it may happen that several models in $\mathcal{M}$ are "output-equivalent", that is, they get mapped into the same output under the specified input, making the inverse problem non-unique for that input and output. Investigating the uniqueness of this inverse problem constitutes the problem of *model* identifiability of the model parameters.

Let $S_{opt}(M(\hat{\underline{a}});\hat{Z}_{1,N}) \subset \mathcal{M}$ denote the set of all optimal models which are output-equivalent to model $M(\hat{\underline{a}})$ under input $\hat{Z}_{1,N}$. Let $S_{opt}(\hat{\underline{a}};\hat{Z}_{1,N}) \subset S(\underline{a})$ denote the set of all corresponding optimal model parameters. The following definitions are introduced:

M1. A parameter a_j of $\underline{a}$ is *globally M-identifiable* ("model identifiable") at $\hat{\underline{a}}$ for the input $\hat{Z}_{1,N}$ if $S_{opt}(\hat{\underline{a}};\hat{Z}_{1,N})$ contains only one optimal parameter or, if not, then:

$$\hat{\hat{\underline{a}}}^{(1)}, \hat{\hat{\underline{a}}}^{(2)} \in S_{opt}(\hat{\underline{a}};\hat{Z}_{1,N}) \Rightarrow \hat{\hat{a}}_j^{(1)} = \hat{\hat{a}}_j^{(2)} \tag{20}$$

Definition M1 implies that a_j is uniquely specified by $\hat{Z}_{1,N}$ and $Q_{1,N}(\hat{\underline{a}};\hat{Z}_{1,N})$.

M2. A parameter a_j of $\underline{a}$ is *locally identifiable* at $\hat{\underline{a}}$ for the input $\hat{Z}_{1,N}$ if there exists a positive number ϵ_j such that:

$$\hat{\hat{\underline{a}}}^{(1)}, \hat{\hat{\underline{a}}}^{(2)} \in S_{opt}(\hat{\underline{a}};\hat{Z}_{1,N}) \Rightarrow |\hat{\hat{a}}_j^{(1)} - \hat{\hat{a}}_j^{(2)}| > \epsilon_j \text{ or } \hat{\hat{a}}_j^{(1)} = \hat{\hat{a}}_j^{(2)} \tag{21}$$

Definition M2 implies that a_j is uniquely specified within a neighborhood of each of its possible values by $\hat{Z}_{1,N}$ and $Q_{1,N}(\hat{\underline{a}};\hat{Z}_{1,N})$, and that if $S(\underline{a})$ is a closed-bounded parameter set, there are only a finite number of possible values for a_j under the given input and model output. Note that if a_j is globally M-identifiable at $\hat{\underline{a}}$, then it is also locally M-identifiable at $\hat{\underline{a}}$.

M3. A parameter a_j of $\underline{a}$ is M-identifiable at $\hat{\underline{a}}$ for the input $\hat{Z}_{1,N}$ if it is either locally or globally M-identifiable.

The above definitions can be extended as follows: The parameter vector $\underline{a}$, or a portion of it, is globally (locally) M-identifiable at $\hat{\underline{a}}$ if all its elements are globally (locally) M-identifiable at $\underline{\hat{a}}$. The parameter vector $\underline{a}$ is not M-identifiable at $\underline{\hat{a}}$ if at least one of its elements is not M-identifiable at $\underline{\hat{a}}$.

As discussed earlier, it is of particular importance to investigate the identifiability of the optimal parameter vector $\underline{\hat{\tilde{a}}}$ based on the input and output data $\mathcal{D}_N$ from a structural system. An *optimal model* $M_P(\underline{\hat{\tilde{a}}})$ for given data $\mathcal{D}_N$ is defined to be any model in $\mathcal{M}_P$ such that:

$$f_N(\hat{X}_{1,N};\underline{\hat{\tilde{a}}},\hat{Z}_{1,N}) = \max_{\underline{\tilde{a}}\in S(\underline{\tilde{a}})} f_N(\hat{X}_{1,N};\underline{\tilde{a}},\hat{Z}_{1,N}) \tag{22}$$

where the parameters $\underline{\hat{\tilde{a}}} = [\underline{\hat{a}}^T,\hat{\sigma}]^T$ are called optimal parameters. Let $S_{opt}(M_P(\underline{\tilde{a}});\mathcal{D}_N) \subset \mathcal{M}_P$ denote the set of all optimal models in the class $\mathcal{M}_P$ and $S_{opt}(\underline{\tilde{a}};\mathcal{D}_N) \subset S(\underline{\tilde{a}})$ denote the set of all corresponding optimal parameters. The following definitions are introduced:

S1. A parameter $\tilde{a}_j$ of $\underline{\tilde{a}}$ is ***globally S-identifiable*** ("system identifiable") at $\underline{\hat{\tilde{a}}}$ for the input and output data $\mathcal{D}_N$ if $S_{opt}(\underline{\hat{\tilde{a}}};\mathcal{D}_N)$ contains only one optimal parameter, or, if not, then:

$$\underline{\hat{\tilde{a}}}^{(1)},\underline{\hat{\tilde{a}}}^{(2)} \in S_{opt}(\underline{\hat{\tilde{a}}};\mathcal{D}_N) \Rightarrow \hat{\tilde{a}}_j^{(1)} = \hat{\tilde{a}}_j^{(2)} \tag{23}$$

S2. A parameter $\tilde{a}_j$ of $\underline{\tilde{a}}$ is *locally S-identifiable* at $\underline{\hat{\tilde{a}}}$ for the input and output data $\mathcal{D}_N$ if there exists a positive number ϵ_j such that:

$$\underline{\hat{\tilde{a}}}^{(1)},\underline{\hat{\tilde{a}}}^{(2)} \in S_{opt}(\underline{\hat{\tilde{a}}};\mathcal{D}_N) \Rightarrow |\hat{\tilde{a}}_j^{(1)} - \hat{\tilde{a}}_j^{(2)}| > \epsilon_j \text{ or } \hat{\tilde{a}}_j^{(1)} = \hat{\tilde{a}}_j^{(2)} \tag{24}$$

S3. A parameter $\tilde{a}_j$ of $\underline{\tilde{a}}$ is S-identifiable at $\underline{\hat{\tilde{a}}}$ for the input and output data $\mathcal{D}_N$ if it is either locally or globally S-identifiable.

Given an optimal model $M_P(\underline{\hat{a}},\hat{\sigma})$ in the class $\mathcal{M}_P$, all other models $M(\underline{a}^*) \in S_{opt}(M(\underline{\hat{a}});\hat{Z}_{1,N})$, if any, having the same observed model output as $M(\underline{\hat{a}})$, correspond to an optimal model $M_P(\underline{a}^*,\hat{\sigma}) \in S_{opt}(M_P(\underline{\hat{a}},\hat{\sigma});\mathcal{D}_N)$ in the class $\mathcal{M}_P$. Another way of looking at this result is that if the parameter vector $\underline{a}$ is not globally M-identifiable at $\underline{\hat{a}}$, then $\underline{\tilde{a}}$ cannot be globally S-identifiable at $[\underline{\hat{a}}^T,\hat{\sigma}]^T$. Furthermore, the number of optimal probability models in $S_{opt}(M_P(\underline{\hat{a}},\hat{\sigma});\mathcal{D}_N) \subset \mathcal{M}_P$ must be at least as large as the number of optimal models in $S_{opt}(M(\underline{\hat{a}});\hat{Z}_{1,N}) \subset \mathcal{M}$. However, there can be models in $S_{opt}(\mathcal{M}_P(\underline{\hat{a}},\hat{\sigma});\mathcal{D}_N)$ which do not have the same model response $Q_{1,N}(\underline{\hat{a}};\hat{Z}_{1,N})$ but still give an equally good fit to the data.

This paper addresses M-identifiability for linear MDOF (multi-degree-of-freedom) models and therefore makes an important first step in solving the very difficult problem of S-identifiability for such a class of models.

MDOF LINEAR STRUCTURAL MODELS

Consider the class $\mathcal{M}_{N_d}$ of N_d-degree of freedom linear structural models starting from rest, and subject to a base excitation:

$$M\underline{\ddot{q}} + C\underline{\dot{q}} + K\underline{q} = -M\underline{b}\ddot{z}(t) \quad ; \quad \underline{q}(0) = \underline{\dot{q}}(0) = 0 \tag{25}$$

The $N_d \times N_d$ matrices M, C, and K are the mass, the damping and the stiffness matrix, respectively. It is assumed that classically damped modes exist. The vector

$\underline{q} = [q_1, q_2, \ldots, q_{N_d}]^T$ consists of the generalized displacements relative to the base of each degree of freedom. The components of the vector $\underline{b} = [b_1, b_2, \ldots, b_{N_d}]^T$ are called pseudo-static influence coefficients, and they are known from the prescribed geometry of the structural model.

The response at the i^{th} dof can be expressed as a superposition of the first $N_m \leq N_d$ modal contributions, where the higher modal contributions are neglected:

$$q_i(t) \cong \sum_{r=1}^{N_m} q_i^{(r)}(t) \quad ; \quad N_m \leq N_d \tag{26}$$

The equation of motion for the contribution of the r^{th} mode to the response at the i^{th} degree of freedom (dof) is:

$$\ddot{q}_i^{(r)} + 2\zeta_r\omega_r\dot{q}_i^{(r)} + \omega_r^2 q_i^{(r)} = -\beta_i^{(r)}\ddot{z}(t) \quad ; \quad q_i^{(r)}(0) = \dot{q}_i^{(r)}(0) = 0 \tag{27}$$

where ω_r is the r^{th} modal frequency, ζ_r is the damping ratio of the r^{th} mode, and $\beta_i^{(r)}$ is the effective participation factor of the r^{th} mode at the i^{th} dof [Beck 1978]. The elements of the mass matrix M are assumed to be determinstically known, since they can be estimated accurately enough from the structural drawings. The uncertainties concerning the damping of the structural model is accounted for by uncertainties of the modal damping ratios ζ_r. Finally, the uncertainty in the stiffness distribution is parameterized through a set of nondimensional positive parameters $\theta_i, i = 1, \ldots, N_\theta$, so that:

$$K = K_0 + \sum_{i=1}^{N_\theta} \theta_i K_i \quad ; \quad \theta_i > 0\ ,\ i = 1, \ldots, N_\theta \tag{28}$$

Each parameter θ_i scales the stiffness contribution K_i of a certain substructure to the total stiffness matrix and K_0 accounts for the stiffness contributions of those substructures with deterministic stiffnesses.

It follows from the above that the vector of uncertain model parameters is:

$$\underline{a} = [\theta_1, \theta_2, \ldots, \theta_{N_\theta}, \zeta_1, \zeta_2, \ldots, \zeta_{N_M}]^T \tag{29}$$

M-IDENTIFIABILITY OF THE MODEL PARAMETERS

Assume that a set of optimal parameters $\hat{\underline{a}}$, where $\underline{a}$ is given by Equation (29), has been found by globally minimizing $J(\underline{a})$, given by Equation (12). The issue here is the M-identifiability of the optimal parameters $\hat{\underline{\theta}} = [\hat{\theta}_1, \ldots, \hat{\theta}_{N_\theta}]^T$ and $\hat{\underline{\zeta}} = [\hat{\zeta}_1, \ldots, \hat{\zeta}_{N_M}]^T$

Let $\mathcal{L}^o$ and $\mathcal{L}^u$ denote the set of integers corresponding to the observed and unobserved degrees of freedom, respectively. The two sets are related as follows:

$$\mathcal{L}^u = \{1, 2, \ldots, N_d\} - \mathcal{L}^o \tag{30}$$

It has been shown [Beck 1978] that the parameters $\{\omega_r, \zeta_r, \beta_i^{(r)}, r = 1, 2, \ldots, N_d, i \in \mathcal{L}^o\}$ which comprise the elements of $\underline{a}$ are globally M-identifiable from the input and output if the following conditions are met: (a) the model has no repeated modes, that is, no two modes have the same modal frequencies and damping ratios, (b) there are no modes with a zero participation factor, and (c) no mode has a node at each coordinate at which the response is measured. Conditions (b) and (c) can be stated as follows: for each mode $r = 1, 2, \ldots, N_d$, there exists at least one $i \in \mathcal{L}^o$, such that $\beta_i^{(r)} \neq 0$. Notice that if this condition is not satisfied, that is, if $\beta_i^{(r)} = 0$ for each $i \in \mathcal{L}^o$, the rth mode will be missing from the output and hence ω_r and ζ_r will not be able to be determined from the input and output.

Utilizing the previous result, it follows that out of the elements of $\hat{\underline{a}}$, the vector $\hat{\underline{\zeta}}$ is globally M-identifiable and therefore, only the M-identifiability of the optimal stiffness parameters $\hat{\underline{\theta}}$ needs to be resolved. Udwadia et al [1978] showed that the problem of identifying the stiffness distribution of an N_d-story shear building from its base input and its response at the roof is a non-unique problem with an upper bound on the number of stiffness distributions corresponding to "output-equivalent" models of $N_d!$. However, the exact number of these "output-equivalent" solutions and an algorithm for obtaining these solutions was not given.

An algorithm is presented here to resolve the M-identifiability of the optimal parameters $\hat{\underline{\theta}}$.

ALGORITHM FOR M-IDENTIFIABILITY

Assume that there exists a finite number of "output-equivalent" models $M(\hat{\underline{\theta}}^{(k)}, \hat{\underline{\zeta}})$; $k = 1, 2, \ldots, K$. Let $S_{opt}(\hat{\underline{\theta}}; \hat{Z}_{1,N})$ denote the set of all corresponding optimal stiffness parameters $\hat{\underline{\theta}}^{(k)}$; $k = 1, 2, \ldots, K$. According to the earlier stated result by Beck [1978], all these stiffness parameters have the following corresponding modal quantities identical:

$$\omega_r(\hat{\underline{\theta}}^{(k)}) = \hat{\omega}_r \quad ; \quad r = 1, \ldots, N_m \quad , \quad k \in \{1, 2, \ldots, K\} \tag{31}$$

$$\beta_i^{(r)}(\hat{\underline{\theta}}^{(k)}) = \hat{\beta}_i^{(r)} \quad ; \quad r = 1, \ldots, N_m \quad , \quad i \in \mathcal{L}^o \quad , \quad k \in \{1, 2, \ldots, K\} \tag{32}$$

Let $\Theta_{\hat{\Omega}}$ denote the set of all parameters $\tilde{\underline{\theta}}^{(k)} \in S(\underline{\theta})$ with corresponding set of modal frequencies $\hat{\Omega} = \{\hat{\omega}_r, r = 1, \ldots, N_m\}$, that is:

$$\Theta_{\hat{\Omega}} = \{\tilde{\underline{\theta}} \in S(\underline{\theta}) : \omega_r(\tilde{\underline{\theta}}) = \hat{\omega}_r ; r = 1, \ldots, N_m\} \tag{33}$$

It is obvious from the definition of $\Theta_{\hat{\Omega}}$ that it is a superset of $S_{opt}(\hat{\underline{\theta}}; \hat{Z}_{1,N})$:

$$S_{opt}(\hat{\underline{\theta}}; \hat{Z}_{1,N}) \subseteq \Theta_{\hat{\Omega}} \tag{34}$$

The methodology for finding the set $S_{opt}(\hat{\underline{\theta}}; \hat{Z}_{1,N})$ consists of two steps. First, the parameter space $S(\underline{\theta})$ is searched methodically, using a new proposed algorithm, to find all elements of $\Theta_{\hat{\Omega}}$. It can be shown that an upper bound on the elements of $\Theta_{\hat{\Omega}}$ is $N_d!$. After $\Theta_{\hat{\Omega}}$ has been found, the second step is taken, consisting of an elimination process, to determine which elements of $\Theta_{\hat{\Omega}}$ satisfy (32), belonging, therefore, in the desired set $S_{opt}(\hat{\underline{\theta}}; \hat{Z}_{1,N})$.

In this paper, we will assume that the number N_θ of uncertain stiffness parameters is equal to the number of contributing modes N_m. If $\Theta_{\hat{\omega}_i} = \{\underline{\theta} \in S(\underline{\theta}) : \omega_i(\underline{\theta}) = \hat{\omega}_i\}$ is the $(N_\theta - 1)$-dimensional hypersurface in the N_θ-dimensional space $S(\underline{\theta})$ where the i^{th} modal frequency remains constant and equal to $\hat{\omega}_i$, then the required set $\Theta_{\hat{\Omega}}$ is given by:

$$\Theta_{\hat{\Omega}} = \cap_{i=1}^{N_\theta} \Theta_{\hat{\omega}_i} \tag{35}$$

Let $c_k(\underline{\theta};\underline{\theta}^*)$ denote a one-dimensional curve in the space $S(\underline{\theta})$, passing through a point $\underline{\theta}^*$, with the property that along this curve all of the first N_m modal frequencies remain fixed except for the k^{th} modal frequency, which is allowed to vary, that is:

$$\begin{aligned} c_k(\underline{\theta};\underline{\theta}^*) = &\text{the largest connected subset of} \\ &\{\underline{\theta} \in S(\underline{\theta}) : \omega_r(\underline{\theta}) = \omega_r(\underline{\theta}^*); r = 1,\dots,k-1,k+1,\dots,N_m\} \\ &\text{containing } \underline{\theta}^* \end{aligned} \tag{36}$$

Following all the different curves $c_k(\underline{\theta};\underline{\tilde{\theta}}^{(1)})$, passing through the known optimal point $\underline{\tilde{\theta}}^{(1)} = \underline{\hat{\theta}}^{(1)}$, and monitoring when, if ever, the "released" frequency ω_k corresponding to each $\underline{\theta} \in c_k(\underline{\theta};\underline{\tilde{\theta}}^{(1)})$ becomes equal to $\hat{\omega}_k$, the "frequency-equivalent" set $\Theta_{\hat{\Omega}}$ is built up. The algorithm has a systematic way of following all possible curves $c_k(\underline{\theta};\underline{\tilde{\theta}}^{(l)}), k = 1,\dots,K$, for all $\underline{\tilde{\theta}}^{(l)} \in \Theta_{\hat{\Omega}}$ that are found. Each curve $c_k(\underline{\theta};\underline{\tilde{\theta}}^{(l)})$ is followed solving a system of linear equations involving the gradient of the function $\underline{\omega}(\underline{\theta})$. The following simple analytical expression can be used to calculate $\frac{\partial \omega_r}{\partial \theta_i}$:

$$\frac{\partial \omega_r}{\partial \theta_i} = \frac{1}{2\omega_r} \underline{\phi}^{(r)^T} K_i \underline{\phi}^{(r)} \tag{37}$$

Figure 1 illustrates schematically the concepts of the proposed algorithm for the case of a uniform two-degree of freedom shear building with parameters θ_1 and θ_2 scaling the interstory stiffnesses of the first and second floor respectively.

It can be shown [Katafygiotis 1991] that the earlier Equation (19) can be replaced by the following:

$$w_k' = \pi_{\underline{\theta}}(\underline{\hat{\theta}}^{(k)}) \mathcal{J}^{-1}(\underline{\hat{\theta}}^{(k)}) \tag{38}$$

where $\mathcal{J}(\underline{\hat{\theta}}^{(k)}) = |\nabla \underline{\omega}(\underline{\hat{\theta}}^{(k)})|$ is the Jacobian of the transformation $\underline{\theta} \to \underline{\omega}(\underline{\theta})$ calculated at $\underline{\theta} = \underline{\hat{\theta}}^{(k)}$. It is interesting to notice that the weighting coefficient w_k, given by the expressions (18) and (38), does not depend explicitly on the measured output. This is surprising at first, since the term $|A_N^{-1}(\underline{\hat{a}}_k)|^{\frac{1}{2}}$ in the earlier expression (19) for w_k' clearly depended on the measured output.

SOME NUMERICAL RESULTS

Consider all possible stiffness distribution solutions $\underline{\hat{\theta}}^{(k)}$ that are obtained for a linear planar shear building with a number of degrees of freedom N_d when the observed degree of freedom is the one corresponding to the roof. The mass distribution was assumed to be uniform and known. The number K of "output-equivalent" solutions found by our algorithm is much smaller than the upper bound of ($N_d!$) derived by Udwadia et al [1978], and for the tested cases, where N_d ranged from two to ten dof, is given by $K = 2^{\mathrm{INT}(\frac{N_d}{2})}$. .

Table 1 shows the eight "output-equivalent" solutions for a six-story ($N_d = 6$) uniform shear building, when the observed degree of freedom is the one corresponding to the roof. Figure 2 shows the effective participation factors of the first three modes at the different floor levels corresponding to all the different optimal solutions $\hat{\underline{\theta}}^{(i)}, i = 1, \ldots, 8$, shown in Table 1. It can be seen that while all these different solutions have exactly the same effective participation factors at the observed degree of freedom, their values at the lower degrees of freedom become increasingly scattered. It can be concluded that if predictions are to be made at the roof, then any of these optimal solutions is going to give the same results, while if the response at a lower degree of freedom is to be predicted, the predictions of all optimal models must be included, with their probabilities appropriately weighted through the coefficients w_k. The weighting factors w_k for each model are given in the last column of Table 1, based on (18) and (38), and under the assumption that the models are equally plausible *a priori*, so that the factors $\pi_{\underline{\theta}}(\hat{\underline{\theta}}^{(k)})$ can be omitted.

CONCLUSIONS

An algorithm to investigate model identifiability of the optimal model parameters in structural dynamics is presented. It is shown that choosing just a single model, as usually done by estimating the model parameters through optimization of the model and measured responses at certain degrees of freedom, can lead to unreliable response predictions at the unobserved degrees of freedom, when the model used is not globally identifiable.

REFERENCES

Beck, J.L., "Determining Models of Structures from Earthquake Records," EERL Report No, 78-01, California Institute of Technology, 1978.

Beck, J.L., "Statistical System Identification of Structures," *Structural Safety and Reliability*, ASCE, II, 1395-1402, 1990.

Katafygiotis, L.S., "Treatment of Model Uncertainties in Structural Dynamics," EERL Report No. 91-01, California Institure of Technology, 1991.

Katafygiotis, L.S. and Beck, J.L., "An Efficient Treatment of Model Uncertainties for the Dynamic Response of Structures," *Proceedings 1st International Conference on Computational Stochastic Mechanics*, Corfu, Greece, 1991.

Udwadia, F.E. and Shah, P.C., "Some Uniqueness Results Related to Building Structural Identification," *SIAM Journal of Applied Mathematics*, Vol. 24, No. 1, 1978.

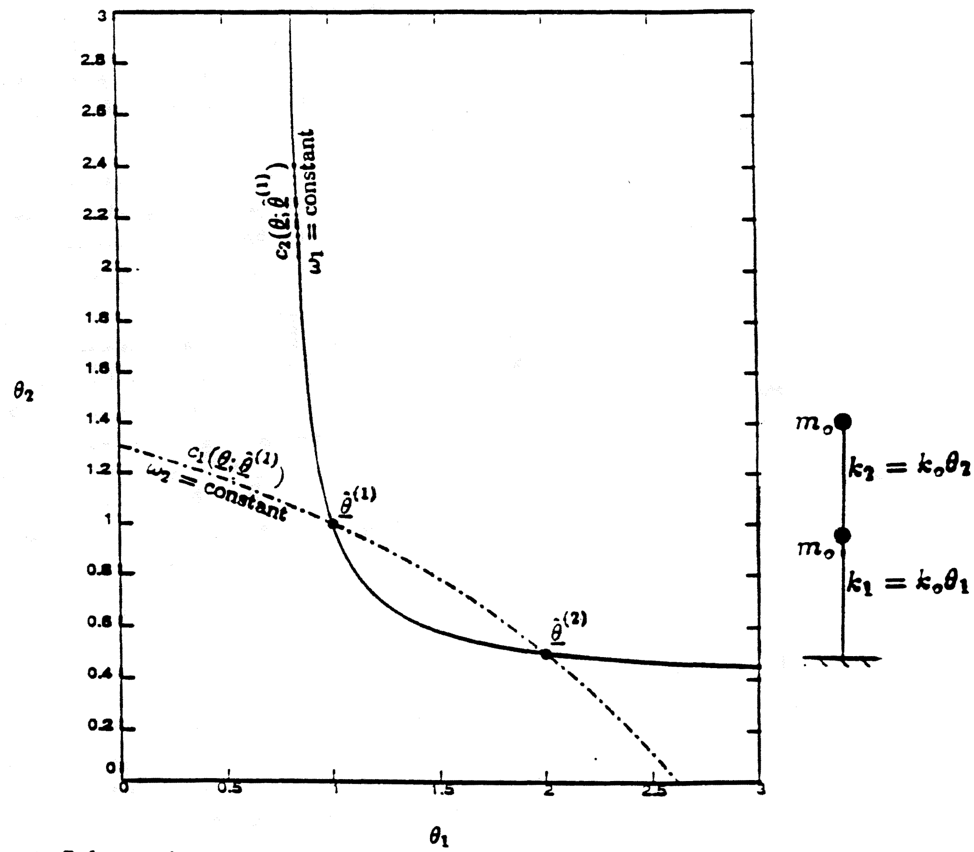

Figure 1 Schematic representation of the proposed algorithm investigating the model indentifiability of the stiffness parameters $\underline{\theta}$ for the case of a two-story uniform shear building.

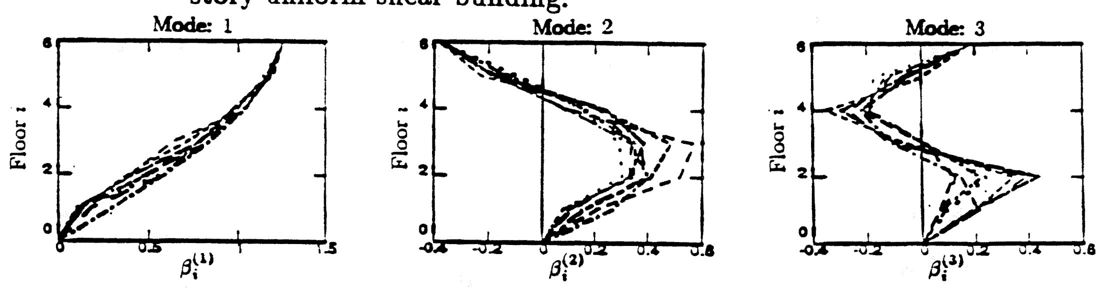

Figure 2 Effective participation factors of the first three modes corresponding to the "output-equivalent" stiffness parameters of Table 1.

No.	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$w_k(\%)$
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	21.35
2	1.5848	0.6963	1.2875	0.7574	1.1766	0.7898	13.49
3	1.9970	0.7980	0.7095	1.3848	0.7113	0.8980	4.91
4	2.0000	1.0000	1.0000	0.5000	1.0000	1.0000	21.35
5	2.0932	1.0476	0.7240	0.7374	0.6705	1.2738	17.07
6	2.2911	0.6304	0.9321	1.1774	0.9515	0.6631	6.46
7	2.4913	0.8777	0.6514	1.1106	0.6672	0.9475	7.40
8	2.8252	0.6753	0.8826	0.9021	0.8753	0.7520	7.97

Table 1 "Output-equivalent" stiffness distributions for a six-story uniform shear building, when the only observed degree of freedom is the one corresponding to the roof.