SYSTEM IDENTIFICATION METHODS APPLIED TO MEASURED SEISMIC RESPONSE

J.L. BECK

Division of Engineering and Applied Science 104-44,
California Institute of Technology,
Pasadena, CA 91125, USA

ABSTRACT

A unified Bayesian statistical framework is described for system identification which can be used to extract important information for earthquake-resistant design from the measured seismic response of structures. In this approach, the "best" (optimal) models within a chosen class of models are those which are locally most probable based on the available data. Using the methodology, one can determine the prediction accuracy of the optimal structural models, the precision of the parameter estimates of these models as well as the precision of the optimal prediction-error probability models, the updated predictive probability density function for the structural response even when there are multiple optimal models, and a principle of parsimony for comparing different classes of models on the same data. The application of modal identification to measured seismic response from some buildings, a bridge and an off-shore platform is reviewed. An example is also given of how the methodology can be used to handle the nonuniqueness in structural model identification. This often arises in practice because of the low density of sensors which is typical of structural seismic instrumentation arrays.

KEYWORDS

System identification; statistical inference; Bayesian statistics; structural dynamics; earthquake records.

INTRODUCTION

Because of their large size and large inertia, the dynamic behavior of civil structures such as tall buildings, bridges and dams can not be studied as complete systems in the laboratory. Also, for similar reasons, these structures are difficult to excite to large response amplitudes in field tests. This leaves a great deal of uncertainty about their detailed dynamic behavior and their ultimate capacity to resist strong earthquake shaking. Therefore, earthquake engineers in many countries have made a concerted effort to instrument structures to record their motion during earthquakes in order to quantitatively assess earthquake-resistant design practices and to provide information to improve these practices. For example, one of the largest instrumentation programs is run by the Division of Mines and Geology, California Department of Conservation (Shakal *et al*, 1988), who have accelerometer arrays distributed over 135 buildings, 35 bridges and 20 dams, in addition to ground motion stations, throughout Cal-

ifornia. The worldwide efforts to capture the seismic response of structures have been rewarded by an increasing number of sets of structural strong-motion records. For example, during just one recent event, the 1994 Northridge earthquake in California, sets of acceleration records were obtained from the CDMG/CSMIP network from 57 buildings, 12 dams, 6 bridges, an airport control tower and a power plant, while the U.S. Geological Survey's National Strong-Motion Program produced records from 37 buildings, 12 dams, 1 bridge and 7 water-distribution facilities (Hall, 1994).

Despite the importance of structural strong-motion records to the earthquake engineering profession, analysis of these data has lagged far behind their rate of collection, with only the occasional example appearing in the literature. This situation may be partly a result of insufficient funds for these types of studies, but also because many researchers and engineers are not fully conversant with systematic computer-based techniques for extracting information. These techniques are known collectively as *system identification* methods in engineering, but a similar activity occurs in other fields such as science and economics under various names: *inverse problems, time series analysis, statistical inference*, etc. The principal difference between the applications in these fields is the amount and type of prior information used in the underlying models. In structural dynamics, for example, the theory of mechanics and other engineering information is available to be used in the system identification process.

There are three major steps in system identification:

1. *Specification* of the mathematical structure of a model for the system (e.g. linear dynamics with prescribed degrees of freedom). This can be done using physical laws and other information about the internal structure of the system, or by a "black-box" approach using some convenient general model class such as auto-regressive moving-average (ARMA) models or neural networks. This step of the process can be viewed as defining a class of possible models $\mathcal{M}$ where each member of $\mathcal{M}$ is given by assigning values to the "free" parameters in the underlying model. In theory, a "nonparametric" class of models can be used where each member is given by specifying functions, such as impulse response functions or transfer functions for linear models, but in practice it is more efficient to use a parameterized form of the functions to reduce the effects of "noise" during estimation.

2. *Estimation* of the model parameters from measurements made on the system. In effect, the objective is to use the system data to determine the "best" model in the class $\mathcal{M}$. A central aspect is the criterion for choosing the "best" model. Usually it involves a measure of fit between the measured system output and the corresponding model output, as in "least-squares output-error" methods, or the fit of the data to the mathematical equation of the model, as in "least-squares equation-error" methods. Any rational criterion should relate to model accuracy, although the prediction accuracy of the "best" model in $\mathcal{M}$ is typically not quantified. Also, if there are multiple "best" models in $\mathcal{M}$ based on the data, it is usually not clear how to proceed.

3. *Evaluation* of whether the model is adequate for its intended purpose. If it is not, then a new class of models can be selected, typically by altering the mathematical structure of the original underlying model, and all three steps repeated. This evaluation step is non-trivial and yet very important. If the identified "best" model is to be used to predict the response of the system, then it is a well-known pitfall to judge its prediction accuracy on the basis of how well it "fits" the available data. One strategy is to use part of the data for estimation and the remainder for evaluation, but this sacrifices data which could be used to improve the model identification. A useful model-building tool is to have a "Principle of Parsimony" or "Occam's Razor", which specifies how far to go in trading-off model simplicity against the goodness of fit to the available data (Jeffreys, 1961; Gull, 1988).

Some general texts on system identification are Eykhoff (1974), Goodwin and Payne (1977), and Ljung

(1987); while publications reviewing methods for structural dynamics applications include Pilkey and Cohen (1972), Hart and Yao (1976), Natke (1982), Shinozuka *et al* (1982) and Imai *et al* (1989). Early reviews of the application of system identification methods to strong motion records from structures can be found in Beck (1978, 1982). The principal goals of this latter application are to learn about the earthquake performance of structures, to assess seismic design codes, and to assess and improve modelling capabilites for dynamic design. There is also increasing interest in applying these methods to dynamic test data to identify high-fidelity models for the design of controllers for structural response control and to develop global methods for structural damage detection (Natke and Yao, 1988).

A survey of the system identification literature reveals a plethora of methods, making it difficult for the novice to the field to know which methods work well with real data. For structural identification, few methods pass the requirement that they be able to handle all of the following issues arising in practice:

1. The dynamic test data is incomplete relative to the level of detail desired to understand the system behavior. For example, the set of observed DOF (degrees of freedom) is a small subset of the set of DOF of interest, and the number of identifiable modes of vibration is much less than the number of DOF of interest.

2. The dynamic test data is contaminated by measurement noise.

3. The chosen class of models does not contain the actual structural system. Because any mathematical model is only an approximation of the real behaviour of a structure, model error always exists and there are no "true" values of the model parameters.

4. In order to reduce modelling uncertainties, the description of the class of structural models should be able to include as much prior information as possible about the structural system.

The underlying motivation for applying system identification to measured seismic response is that it is currently not possible to accurately predict this response using analytical (theoretical) structural models. This is because there are many sources of modelling errors which lead to uncertain accuracy in the predicted response: variations of the material properties during manufacture; inexact modelling of the material constitutive behavior; uncertainties introduced during the construction process; inexact modelling of the boundary conditions (e.g. there are no exact pinned or fixed joints); errors because of the spatial discretization of the distributed structural system; unmodelled features such as neglected "nonstructural" components; and so on.

Because of these modelling errors, system identification is best tackled as a statistical inference problem. This can be done by imbedding the "deterministic" structural models within a class of probability models so that the structural models give a predictable ("systematic") part and the prediction error is modelled as an uncertain ("random") part. Eykhoff (1974), Goodwin and Payne (1977) and Ljung (1987), for example, all emphasize a statistical approach to system identification using a "classical" statistical framework. Recently, powerful statistical techniques have been developed using a Bayesian statistical framework, including the Principle of Maximum (Information) Entropy (e.g. Box and Tiao, 1973; Jaynes, 1983; Bretthorst, 1988). Beck (1989) and Beck and Katafygiotis (1991) have used such a framework to develop a general system identification approach which is capable of handling linear and nonlinear models even when they are not uniquely identifiable based on the available data. In fact, the method addresses all the previous enumerated practical difficulties. This statistical system identification framework is presented here with some new extensions. Some applications to modal identification and structural model identification (model updating) are then reviewed.

STATISTICAL SYSTEM IDENTIFICATION

It is assumed that a general mathematical form has been chosen to specify a class of models $\mathcal{M}$ describing the input-output behavior of a structure, but that there are free parameters $\underline{a} \in R^{N_a}$ which need to be assigned values in order to choose a particular model $M(\underline{a}) \in \mathcal{M}$. The models can be linear or nonlinear, and static or dynamic. They can be expressed as Newtonian equations of motion, in state-space form or as ARMA models. An example is given later of models appropriate for modal identification.

Once the class of models is selected, modelling uncertainties of two general types need to be quantified by using probability models. The first type, *parameter uncertainty*, arises simply because the most appropriate values of the model parameters $\underline{a}$ to be used to describe the structure's behavior are uncertain, that is, it is not known *a priori* which model in the class of models $\mathcal{M}$ is the "best" to describe the structure's behavior. The second type of uncertainty, *prediction accuracy*, arises because modelling errors lead to an uncertain error in the response predictions given by any model in $\mathcal{M}$. For example, if the class of linear dynamic models is chosen, then there are uncertainties associated with the values of the various parameters that need to be chosen, such as Young's modulus E or the effective moment of inertia of a cracked concrete member. Furthermore, for any given model in the class, the corresponding predicted response will differ from the actual structural response because any mathematical model is only an approximation of the real behaviour of a structure, and uncertain modelling errors lead to uncertain accuracy for the model's predictions.

Probability models describing these two types of uncertainties must be set up, then statistical system identification consists of updating these probability models by applying Bayes Theorem to the structural data. This involves dealing with the probability of models, which are, of course, not repeatable events, so the common interpretation of probability as a relative frequency of occurrences of an event in the long run is not applicable. There is, however, another more useful interpretation of probability as a multi-valued logic for plausible reasoning under incomplete information. Cox (1961) demonstrated that to quantify the plausibility of a proposition (statement) denoted by b, based on information expressed in a proposition denoted by c, by using a function $P(b|c) \in [0,1]$, one is led to the usual calculus of probability merely by the requirement of consistency with classical mathematical logic. From this point of view, probability is a generalization of classical logic since the latter assumes that the information available is complete enough to deduce that either $c \Rightarrow b$ (i.e. $P(b|c) = 1$) or that $c \Rightarrow$ not b (i.e. $P(b|c) = 0$). Jaynes has written a nice introduction to probability as a logic (Jaynes, 1988), as well as a history of the development of the ideas (Jaynes, 1978).

Bayesian statistical methods using probability logic are built primarily on the earlier work of Laplace starting in 1774 (e.g. Laplace, 1812), Jeffreys from 1919 (e.g. Jeffreys, 1961), Cox in 1946 (e.g. Cox, 1961), and Jaynes starting in 1957 (e.g. Jaynes, 1983). Those studying the Bayesian statistical literature for the first time should be aware, however, that there are two principal lines of thought. One line is rooted in decision theory and allows subjective probability models as personal expressions of an individual's degree of belief (e.g. Bayes, 1763; Ramsey, 1931; De Finetti, 1974; Savage, 1972; Lindley, 1965). The other line (Laplace-Jeffreys-Cox-Jaynes) stresses that probability models should be uniquely determined by objective criteria, such as the Principle of Maximum Entropy, once the incomplete information available has been stated in a suitable explicit form.

Basic Probability Models

The specified class of models $\mathcal{M}$ provides a functional relationship between the model output vector

$\underline{x}(n;\underline{a}) \in R^{N_o}$ at time $t_n = n\Delta t$ and the system input $Z_1^n = \{\underline{z}(m) \in R^{N_I} : m = 1, 2, \ldots, n\}$:

$$\underline{x}(n;\underline{a}) = \underline{x}(n;\underline{a}, Z_1^n, \mathcal{M}) \tag{1}$$

The model output $\underline{x}$ here contains only the model output corresponding to observed response quantities. For example, if accelerations have been measured then it contains only the accelerations predicted by the model at the observed DOF. In the following, the dependence of $\underline{x}(n;\underline{a})$ and other model quantities on the input Z_1^n and the theoretical model $\mathcal{M}$ will be suppressed in the notation. Although the model output is only required at discrete times, the theoretical model which is the underlying basis of the class $\mathcal{M}$ can be in any form, such as a system of continuous-time differential equations. The assumption of discrete-time input is not critical but is typical of the data usually available.

The first step in imbedding the deterministic class of models $\mathcal{M}$ in a class of probability models is to define the *prediction error* $\underline{e}(n;\underline{a}) \in R^{N_o}$ as the difference between the model output and the system output, so that if $\underline{y}(n) \in R^{N_o}$ denotes the *system* output which would be measured at time t_n, then:

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

For technical reasons, $\underline{y}$ is the system output which would be *measured*. The prediction error is, therefore, actually a combined effect of measurement noise and model error. For modern instrumentation measuring structural response, however, the measurement noise is usually negligible compared with the model error, and so prediction of the system output which would be measured is essentially equivalent to predicting the actual system output.

In order to describe the uncertainty in the prediction error, a class of probability models $\mathcal{P}$ is chosen, parameterized by the *prediction-error parameters* $\underline{\sigma} \in R^{N_\sigma}$, which prescribes a function h_M giving the PDF (probability density function) of a sequence of M prediction errors for arbitrary M, that is,

$$p(E_1^M(\underline{a}) \mid \underline{\sigma}, \mathcal{P}) = h_M(\underline{e}(1;\underline{a}), \ldots, \underline{e}(M;\underline{a}); \underline{\sigma}) \tag{3}$$

where $E_1^M(\underline{a})$ denotes the proposition: "For model $M(\underline{a})$ in $\mathcal{M}$, the sequence of the first M prediction errors take on the values $\underline{e}(n;\underline{a}), n = 1, \ldots, M$," while $\underline{\sigma}, \mathcal{P}$ on the left side of this equation denote the proposition: "The probability model for these M prediction errors is given by the function $h_M(\underline{e}(1;\underline{a}), \ldots, \underline{e}(M;\underline{a}); \underline{\sigma})$ in the class $\mathcal{P}$." A similar interpretation will be used for other PDFs.

The selection of the classes $\mathcal{M}$ and $\mathcal{P}$ allows a class of probability models $\mathcal{M}_P$ to be defined, parameterized by $\underline{\alpha} = [\underline{a}^T, \underline{\sigma}^T]^T \in S(\underline{\alpha}) \subset R^{N_\alpha}$ where $N_\alpha = N_a + N_\sigma$, which prescribes a function f_M giving the PDF for the system output sequence $Y_1^M = \{\underline{y}(n) \in R^{N_o} : n = 1, 2, \ldots, M\}$:

$$\begin{aligned} p(Y_1^M \mid \underline{\alpha}, Z_1^M, \mathcal{M}_P) &= f_M(\underline{y}(1), \ldots, \underline{y}(M); \underline{\alpha}, Z_1^M) \\ &= h_M(\underline{y}(1) - \underline{x}(1;\underline{a}), \ldots, \underline{y}(M) - \underline{x}(M;\underline{a}); \underline{\sigma}) \end{aligned} \tag{4}$$

where f_M is determined by using equation (2) in equation (3). In order to account for the uncertainty in the values for the parameters $\underline{\alpha}$, the specification of $\mathcal{M}_P$ also involves choosing an initial ("prior") PDF $\pi(\underline{\alpha})$ over the set $S(\underline{\alpha})$ of possible parameter values, that is:

$$p(\underline{\alpha} \mid \mathcal{M}_P) = \pi(\underline{\alpha}) \tag{5}$$

The choice for $\pi(\underline{\alpha})$ allows engineering judgement about the plausibilities of the different models to be incorporated. It would typically be chosen as a smooth slowly-varying PDF which is mathematically convenient and roughly reflects the user's judgement. Alternatively, the initial PDF can be chosen by Principle of Maximum Entropy if the information which is desired to be utilized can be explicitly expressed as "testable" information (Jaynes, 1968). Dependence of the initial PDF on the input Z_1^M

can be introduced in equation (5). For example, this might be done when using linear models in order to account for the fact that the nonlinearities in structures typically cause them to become more flexible for stronger excitation, although the final results are often insensitive to the choice of the initial PDF.

Summarizing, specification of the class $\mathcal{M}_P$ implies specification of two PDFs which are the basic probability models for system identification: f_M for the system output, which relates to the uncertainty in the prediction accuracy of the models in $\mathcal{M}$, and π for the model parameters $\underline{a}$ and $\underline{\sigma}$, which relates to the uncertainty in the deterministic models in $\mathcal{M}$ and the prediction-error probability models in $\mathcal{P}$.

Initial and Updated Predictive Probability Models

The PDF in equation (4) gives predictions for the system output for a particular probability model in class $\mathcal{M}_P$ which is specified by the value of the deterministic model and prediction-error parameters $\underline{\alpha}$. Prior to utilizing data, the best choice for making predictions of the system output using the assumed class of probability models is the initial predictive PDF:

$$\begin{aligned} p(Y_1^M \mid Z_1^M, \mathcal{M}_P) &= \int_{S(\underline{\alpha})} p(Y_1^M \mid \underline{\alpha}, Z_1^M, \mathcal{M}_P)\; p(\underline{\alpha} \mid \mathcal{M}_P)\; d\underline{\alpha} \\ &= \int_{S(\underline{\alpha})} f_M(Y_1^M; \underline{\alpha}, Z_1^M)\; \pi(\underline{\alpha})\; d\underline{\alpha} \end{aligned} \tag{6}$$

which is readily derived from the axioms of probability (the "Total Probability Theorem"). This gives a predictive PDF using the whole class $\mathcal{M}_P$ as a weighted average of the predictive PDFs for each model in $\mathcal{M}_P$, with weights given by their initial probabilities.

Let $\mathcal{D}_N$ denote a set of observed time history data from the structural system at discrete times $t_n = n\Delta t$, where Δt is the sampling interval. These data are assumed to consist of the sampled history $\hat{Z}_1^N = \{\hat{\underline{z}}(n) \in R^{N_I} : n = 1, 2, \ldots, N\}$ for N_I inputs and the sampled output history $\hat{Y}_1^N = \{\hat{\underline{y}}(n) \in R^{N_o} : n = 1, 2, \ldots, N\}$ which is the measured response at the N_o instrumented degrees of freedom of the structure. For simplicity here, it is assumed that the *observed* input gives a nearly complete description of the *system* input, otherwise a stochastic model of the system input must be constructed to cover unobserved or "noisy" inputs. For the applications of interest here, the system input consists of the accelerations induced by an earthquake at "support" degrees of freedom for the structure; the system output consists of the corresponding accelerations at the N_o observed degrees of freedom.

Based on the new information in the data $\mathcal{D}_N$, the predictive PDF in equation (6) can be replaced by an updated predictive PDF by using the Total Probability Theorem again:

$$\begin{aligned} p(Y_{N+1}^M \mid \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P) &= \int_{S(\underline{\alpha})} p(Y_{N+1}^M \mid \underline{\alpha}, \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P)\; p(\underline{\alpha} \mid \mathcal{D}_N, \mathcal{M}_P)\; d\underline{\alpha} \\ &= k \int_{S(\underline{\alpha})} p(\hat{Y}_1^N, Y_{N+1}^M \mid \underline{\alpha}, \hat{Z}_1^N, Z_{N+1}^M, \mathcal{M}_P)\; p(\underline{\alpha} \mid \mathcal{M}_P)\; d\underline{\alpha} \\ &= k \int_{S(\underline{\alpha})} f_M(\hat{Y}_1^N, Y_{N+1}^M; \underline{\alpha}, \hat{Z}_1^N, Z_{N+1}^M)\; \pi(\underline{\alpha})\; d\underline{\alpha} \end{aligned} \tag{7}$$

where now the system output is only predicted for the next $(M - N)$ sampling times for a prescribed future input Z_{N+1}^M, since the output is already specified for the first N samples by $\mathcal{D}_N$. In equation (7), Bayes Theorem is used to calculate the updated ("posterior") PDF from the initial PDF:

$$p(\underline{\alpha} \mid \mathcal{D}_N, \mathcal{M}_P) = k\, p(\hat{Y}_1^N \mid \underline{\alpha}, \hat{Z}_1^N, \mathcal{M}_P)\; p(\underline{\alpha} \mid \mathcal{M}_P)$$

$$= \quad k f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N)\ \pi(\underline{\alpha}) \tag{8}$$

where the normalizing constant k can be evaluated using equation (6), since it satisfies:

$$k^{-1} = p(\hat{Y}_1^N \mid \hat{Z}_1^N, \mathcal{M}_P) \tag{9}$$

The updated predictive PDF using the whole class $\mathcal{M}_P$ and the data $\mathcal{D}_N$ can be viewed as a weighted average of the predictive PDFs for each model in $\mathcal{M}_P$, like the initial case above, except that now the weights are given by the updated probabilities for each model.

The difficulty with the solution given in equation (6) for the initial predictive PDF for the system output is that the multi-dimensional integral can not be evaluated analytically, nor numerically if the dimension of the parameter space $S(\underline{\alpha})$ is too high (say, greater than 5). The situation is even worse for the numerical integration in equation (7) for the updated predictive PDF, since the function f_M is sharply peaked at one or more locations in $S(\underline{\alpha})$ if the number of sampling times N is not small. An asymptotic approach is presented next, which not only provides approximations for the required integrals, but also provides insight into the solution which is not revealed by the integrals themselves.

Asymptotic Approximations

Define the *optimal* parameters $\underline{\hat{\alpha}} = [\underline{\hat{a}}^T, \underline{\hat{\sigma}}^T]^T$ for $\mathcal{M}_P$ given data $\mathcal{D}_N$ to be values of the parameters $\underline{\alpha} \in S(\underline{\alpha})$ that *globally* maximize $f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N)$. The optimal parameters are equivalent to the maximum likelihood estimates used in "traditional" statistical theory based on the interpretation of probability as relative frequencies of events in the "long run". If there is a finite number K of optimal parameters, then the class of models $\mathcal{M}_P$ is said to be *system identifiable* under data $\mathcal{D}_N$ (Beck and Katafygiotis, 1991). Denote these optimal parameters by $\underline{\hat{\alpha}}^{(k)},\ k = 1, \ldots, K$. By expanding $\ln f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N)$ in a second-order Taylor series about an optimal parameter $\underline{\hat{\alpha}}$, the following *local* approximation is derived:

$$f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N) \simeq f_N(\hat{Y}_1^N; \underline{\hat{\alpha}}, \hat{Z}_1^N) \exp\left(-\frac{1}{2}[\underline{\alpha} - \underline{\hat{\alpha}}]^T A_N(\underline{\hat{\alpha}})[\underline{\alpha} - \underline{\hat{\alpha}}]\right) \tag{10}$$

where the elements of the $N_\alpha \times N_\alpha$ Hessian matrix $A_N(\underline{\alpha})$ are given by:

$$[A_N(\underline{\alpha})]_{ij} = -\frac{\partial^2 \ln f_N\left(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N\right)}{\partial \alpha_i \partial \alpha_j} \tag{11}$$

The elements of A_N are $\mathcal{O}(N)$ and, therefore, for a large number N of sampling times, which is usually the case with measured seismic response, $f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N)$ is very peaked at each optimal parameter $\underline{\hat{\alpha}}^{(k)}$. Therefore, Laplace's method for asymptotic expansions can be applied to obtain the following approximation for the integral in equation (7) (Beck and Katafygiotis, 1996):

$$p(Y_{N+1}^M \mid \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P) = \sum_{k=1}^{K} w_k\, p(Y_{N+1}^M \mid \underline{\hat{\alpha}}^{(k)}, \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P)[1 + \mathcal{O}(N^{-1})] \tag{12}$$

where

$$\begin{aligned} p(Y_{N+1}^M \mid \underline{\hat{\alpha}}^{(k)}, \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P) &= \frac{p(\hat{Y}_1^N, Y_{N+1}^M \mid \underline{\hat{\alpha}}^{(k)}, \hat{Z}_1^N, Z_{N+1}^M, \mathcal{M}_P)}{p(\hat{Y}_1^N \mid \underline{\hat{\alpha}}^{(k)}, \hat{Z}_1^N, \mathcal{M}_P)} \\ &= \frac{f_M(\hat{Y}_1^N, Y_{N+1}^M; \underline{\hat{\alpha}}^{(k)}, \hat{Z}_1^N, Z_{N+1}^M)}{f_N(\hat{Y}_1^N; \underline{\hat{\alpha}}^{(k)}, \hat{Z}_1^N)} \end{aligned} \tag{13}$$

and the weighting coefficient w_k corresponding to optimal parameter $\underline{\hat{\alpha}}^{(k)}$ is given by:

$$w_k = \frac{w'_k}{\sum_{j=1}^{K} w'_j}, \quad w'_k = \pi(\underline{\hat{\alpha}}^{(k)}) \mid A_N(\underline{\hat{\alpha}}^{(k)}) \mid^{-1/2} \tag{14}$$

The continuous weighted average for the updated predictive PDF in equation (7) is therefore approximated by a discrete weighted average of the predictive PDFs for each optimal model in class $\mathcal{M}_P$.

Since $f_N(\hat{Y}_1^N; \underline{\alpha}, \hat{Z}_1^N)$ is very peaked at each optimal parameter $\underline{\hat{\alpha}}$, the updated PDF $p(\underline{\alpha} \mid \mathcal{D}_N, \mathcal{M}_P)$ is too if the initial PDF is a smooth and slowly-varying function of $\underline{\alpha}$. It can be deduced from equations (8) and (10) that in this case $p(\underline{\alpha} \mid \mathcal{D}_N, \mathcal{M}_P)$ behaves locally about each optimal parameter $\underline{\hat{\alpha}}$ like a multi-dimensional Gaussian distribution with mean $\underline{\hat{\alpha}}$ and an $N_\alpha \times N_\alpha$ covariance matrix $A_N^{-1}(\underline{\hat{\alpha}})$. The optimal models may therefore be interpreted as locally most probable models within the class $\mathcal{M}_P$ based on the data $\mathcal{D}_N$, and the covariance matrix can be used to examine how precisely the optimal model parameters are identified by the data. Furthermore, it can be deduced that the predictive PDF in equation (12) for each of the optimal models is weighted in proportion to the volume of the updated PDF $p(\underline{\alpha}|\mathcal{D}_N, \mathcal{M}_P)$ under its Gaussian-shaped peak positioned at the corresponding optimal parameter. In fact, a more accurate version of equation (12) would include all probability models for which $p(\underline{\alpha}|\mathcal{D}_N, \mathcal{M}_P)$ has a local maximum, but as N increases, the contributions of the optimal models eventually dominate.

Equation (14) shows that for the asymptotic approximation of the updated predictive PDF given by equation (12), the initial PDF $\pi(\underline{\alpha})$ is not required over the whole domain $S(\underline{\alpha})$. Instead, only the relative values at the optimal parameters $\underline{\hat{\alpha}}^{(k)}$ need be specified. This allows engineering judgement to be incorporated since some of the optimal models based on the data may not be very plausible from an engineering point of view. If $\mathcal{M}_P$ is *globally system identifiable* under $\mathcal{D}_N$, that is, $K = 1$ in equation (14), the initial PDF does not enter at all. This is the well-known Bayesian result that a large amount of data will overwhelm the effect of the chosen initial PDF and one can proceed using the predictive PDF for the unique optimal model (Lindley, 1965). The above result shows, however, that this is *only* true in the case of global system identifiability.

It is of interest to note that the exact updated predictive PDF in equation (7) does not require parameter estimation in the usual sense, but that the asymptotic approximation does. The problem is converted from an integration over all possible probability models in the class $\mathcal{M}_P$, which is computationally prohibitive when there are more than just a handful of parameters, to a nontrivial optimization problem which can at least be tackled numerically (Beck and Katafygiotis, 1991; Yang and Beck, 1996). Since the optimal parameters are also Fisher's MLEs (maximum likelihood estimates), the well-known MLE approach to parameter estimation can be justified in terms of an asymptotic approximation to the exact Bayesian solution for response predictions, but, in addition, the asymptotic result in equation (12) shows how to proceed when there are multiple MLEs, while it is not clear how to proceed in this situation in "traditional" statistics.

Least-Squares Output-Error Method

By making appropriate choices for the class $\mathcal{P}$ of prediction-error probability models, various "standard" methods for system identification can be viewed as special cases of the general Bayesian statistical framework presented here. For example, least-squares output-error or equation-error methods can be derived as approximations to the exact Bayesian solutions. Alternatively, recursive procedures can be formulated if sequential calculation of the most probable model parameters is desired, leading to forms of the extended Kalman filter (West and Harrison, 1989), which has been applied to system identification using earthquake records in Beck (1978) and Hoshiya and Saito (1984), for example.

The least-squares output-error approach has been the dominant method for the application of system identification to measured structural motions, as summarized later. One reason is that, unlike the least-squares equation-error approach, it can be applied when the observed output of the structure does not correspond to the complete state vector of the structural model. This situation arises, for example, in updating finite-element models where the number of model DOF is typically much larger than the number of observed DOF. The output-error case is therefore examined here. Although least-squares parameter-estimation methods can be accepted in their own right, by viewing them within the Bayesian statistical framework one can determine the prediction accuracy of the most probable models, the precision of the parameter estimates of the structural models and prediction-error probability models, the updated predictive PDF even when there are multiple optimal models, and a principle of parsimony for comparing different classes of models on the same data.

Consider a choice for the class $\mathcal{P}$ in equation (3) which makes $\underline{e}(n;\underline{a})$ a zero-mean stationary Gaussian white-noise stochastic process with independent components. This is equivalent to a class of probability models for expressing the engineer's incomplete knowledge which asserts that knowing the prediction errors at other times, or at other locations within the structure, does not influence the engineer's uncertainty concerning the value of the prediction error at a specified time and location. This choice for the class $\mathcal{P}$ also corresponds to a probability distribution h_M in equation (3) which is given by the Principle of Maximum Entropy under the condition of zero means and finite variances (Jaynes, 1978). This means that under the latter conditions, h_M gives the maximal uncertainty that E_1^M in (3) can have, just as the uniform distribution gives the maximal uncertainty for an uncertain variable with a finite range. Therefore, if any other choice for h_M is made, there should be strong grounds for the implied reduction in uncertainty in the predictions.

The basic probability model f_M of equation (4) can now be derived as:

$$f_M(Y_1^M;\underline{\alpha},Z_1^M)=\prod_{i=1}^{N_0}\left[\frac{1}{(\sqrt{2\pi}\sigma_i)^M}\exp\left(-\frac{1}{2\sigma_i^2}\sum_{n=1}^{M}(y_i(n)-x_i(n;\underline{a}))^2\right)\right] \tag{15}$$

The prediction-error parameters $\underline{\sigma}$ are the unknown standard deviations, that is, $\underline{\sigma}=[\sigma_1,\sigma_2,...,\sigma_{N_o}]^T$. In the special case where the variances of the prediction errors at all observed DOF (degrees of freedom) are assumed to be equal, the vector $\underline{\sigma}$ reduces to a single parameter σ. The optimal parameters $\hat{\underline{\alpha}}=[\hat{\underline{a}}^T,\hat{\underline{\sigma}}^T]^T$ are determined by maximizing f_N with respect to $\underline{\alpha}$. For fixed model parameters $\underline{a}$, maximizing $\ln f_N(\hat{Y}_1^N;\underline{\alpha},\hat{Z}_1^N)$ with respect to the σ_i leads to:

$$\hat{\sigma}_i^2(\underline{a})=\frac{1}{N}\sum_{n=1}^{N}(\hat{y}_i(n)-x_i(n;\underline{a}))^2 \tag{16}$$

Therefore, in the case of independent variances, the most probable standard deviation $\hat{\sigma}_i(\underline{a})$, for given model parameters $\underline{a}$, is equal to the RMS of the prediction errors at the i^{th} observed DOF. In the special case where all the variances are assumed equal, it can be easily shown that $\hat{\sigma}(\underline{a})$ is equal to the RMS of the prediction errors at all observed DOF, that is:

$$\hat{\sigma}^2(\underline{a})=\frac{1}{N_oN}\sum_{i=1}^{N_o}\sum_{n=1}^{N}(\hat{y}_i(n)-x_i(n;\underline{a}))^2 \tag{17}$$

Obviously, the condition for the overall most probable variance is given by equations (16) or (17) when $\underline{a}=\hat{\underline{a}}$. Substituting equations (16) or (17) into equation (15) with $M=N$:

$$f_N(\hat{Y}_1^N;\underline{a},\hat{\underline{\sigma}}(\underline{a}),\hat{Z}_1^N)=[cJ(\underline{a})]^{-N_oN/2} \tag{18}$$

where constant $c = 2\pi e$ and $J(\underline{a})$ is either the geometric or arithmetic mean of the RMS prediction-errors $\hat{\sigma}_i^2(\underline{a})$ at the observed DOF, depending on whether independent or equal variances are assumed:

$$\begin{aligned} J(\underline{a}) &= \left[\prod_{i=1}^{N_o} \hat{\sigma}_i^2(\underline{a})\right]^{\frac{1}{N_o}} \quad \text{(independent variances)} \\ &= \left[\frac{1}{N_o}\sum_{i=1}^{N_o} \hat{\sigma}_i^2(\underline{a})\right] = \hat{\sigma}^2(\underline{a}) \quad \text{(equal variances)} \end{aligned} \tag{19}$$

The optimal structural model parameters $\underline{\hat{a}}$ are determined by minimizing $J(\underline{a})$ in equation (19), and the equal-variance case corresponds to the usual least-squares output-error method for estimation of model parameters. In structural applications, the model output $x_i(n;\underline{a})$ involved in equation (19) is a nonlinear function of the parameters, even if the model has linear dynamics, and so the minimization must be done numerically by an iterative optimization algorithm. As pointed out earlier, there may be multiple optimal model parameters because $J(\underline{a})$ might attain its minimum at more than one value $\underline{\hat{a}}$ in $S(\underline{a})$, the set of permissible values of $\underline{a}$. The task of finding all the global minima of the nonconvex function $J(\underline{a})$ is nontrivial. Powerful methods have been developed to treat this problem, however (Beck and Katafygiotis, 1991; Yang and Beck, 1996). In the independent-variance case, for each optimal model parameter $\underline{\hat{a}}$ there are unique corresponding optimal prediction-error variances $\hat{\sigma}_i^2 = \hat{\sigma}_i^2(\underline{\hat{a}}), i = 1, \ldots, N_o$, given by equation (16) with $\underline{a} = \underline{\hat{a}}$. In the case of equal variances, the optimal prediction-error variance $\hat{\sigma}^2 = \hat{\sigma}^2(\underline{\hat{a}}) = J(\underline{\hat{a}})$ is uniquely determined even though $\underline{\hat{a}}$ may not be, since by definition $J(\underline{\hat{a}})$ is the global minimum of $J(\underline{a})$, and so each optimal structural model has the same prediction-error variance.

For the asymptotic approximation in equation (12), the PDF in equation (13) is now:

$$p(Y_{N+1}^M \mid \underline{\hat{\alpha}}, \mathcal{D}_N, Z_{N+1}^M, \mathcal{M}_P) = \prod_{i=1}^{N_0}\left(\frac{1}{(\sqrt{2\pi}\hat{\sigma}_i)^{M-N}} \exp\left[-\frac{1}{2\hat{\sigma}_i^2}\sum_{n=N+1}^{M}(y_i(n) - x_i(n;\underline{\hat{a}}))^2\right]\right) \tag{20}$$

The Hessian matrix $A_N(\underline{\hat{\alpha}})$ needed to evaluate the weights in equation (14) is block diagonal with one block being an N_axN_a matrix $B_N(\underline{\hat{\alpha}})$ corresponding to the structural model parameters $\underline{a}$, and the other block matrix $C_N(\underline{\hat{\alpha}})$ (an N_oxN_o matrix for independent variances and a scalar for equal variances) corresponding to the prediction-error parameters $\underline{\sigma}$. It can be shown that the determinant of $C_N(\underline{\hat{\alpha}}^{(k)})$ is the same for each optimal parameter $\underline{\hat{\alpha}}^{(k)}$, so the weights depend only on the determinant of the matrix B_N with elements given by:

$$\begin{aligned} [B_N(\underline{\hat{a}})]_{jk} &= \frac{N}{2}\sum_{i=1}^{N_o}\left[\frac{1}{\hat{\sigma}_i^2(\underline{a})}\frac{\partial^2\hat{\sigma}_i^2(\underline{a})}{\partial a_j \partial a_k}\right]_{\underline{\hat{a}}} \quad \text{(independent variances)} \\ &= \frac{NN_o}{2}\left[\frac{1}{\hat{\sigma}^2(\underline{a})}\frac{\partial^2\hat{\sigma}^2(\underline{a})}{\partial a_j \partial a_k}\right]_{\underline{\hat{a}}} \quad \text{(equal variances)} \end{aligned} \tag{21}$$

where $\hat{\sigma}_i^2(\underline{a})$ and $\hat{\sigma}^2(\underline{a})$ are given by equations (16) and (17) respectively. Because of the block diagonal structure of the Hessian matrix $A_N(\underline{\hat{\alpha}}^{(k)})$, the inverse of $B_N(\underline{\hat{a}}^{(k)})$ is the covariance matrix for the local Gaussian behavior of the updated PDF $p(\underline{a} \mid \mathcal{D}_N, \mathcal{M}_P)$ centered at $\underline{\hat{a}}^{(k)}$ in the structural model parameter space.

From equations (12) and (20), the mean and variance of the predicted system output $\underline{y}(n)$ at times $t_n > t_N$ are given by:

$$\begin{aligned} E[y_i(n)] &\simeq \sum_{k=1}^{K} w_k x_i(n;\underline{\hat{a}}^{(k)}) \\ \text{Var}[y_i(n)] &\simeq \sum_{k=1}^{K} w_k \hat{\sigma}_i^2(\underline{\hat{a}}^{(k)}) + \sum_{k=1}^{K} w_k x_i^2(n;\underline{\hat{a}}^{(k)}) - E[y_i(n)]^2 \end{aligned} \tag{22}$$

Therefore, the mean system output is a weighted sum of the outputs of each optimal structural model and the variance for the system output is the sum of the prediction-error variances for each optimal model, and a contribution due to the lack of uniqueness in these models. If $K = 1$ (global system identifiability), the last two terms in the variance expression cancel and the predicted system output is Gaussian with mean and variance given in equation (22). If $K > 1$ (local system identifiability), the asymptotic predictive PDF is not Gaussian, even though each term in the sum in equation (12) is a Gaussian PDF (see equation (20)).

APPLICATION TO MODAL IDENTIFICATION

Method for Time-Domain Identification

An important application of system identification in structural dynamics is to *modal identification* where modal parameters based on a model with linear dynamics are determined using dynamic data from a structure. For measured seismic response from tall buildings, Beck (1978) and Beck and Jennings (1980) showed that this can be done in the time domain, without the need to develop a structural model involving mass, stiffness and damping matrices. Beck extended the methodology to multiple inputs and implemented it in a computer program MODE-ID, in order to find the modal parameters from seismic motions recorded on a bridge (Werner *et al*, 1987). MODE-ID is based on the least-squares output-error method described earlier, together with a class of models defined as follows.

Structural motion at the N_o observed DOF is modelled as a superposition of N_m dominant modes:

$$x_i(t) = \sum_{r=1}^{N_m} x_{ir}(t), \quad i = 1, \ldots, N_o \tag{23}$$

The model for the $(N_m - 1)$ dynamic modes of vibration is an accurate discrete approximation (Beck and Dowling, 1988) of the well-known equation of motion:

$$\ddot{x}_{ir} + 2\zeta_r\omega_r\dot{x}_{ir} + \omega_r^2 x_{ir} = \phi_{ir}\sum_{k=1}^{N_I} p_{rk}f_k(t) \quad x_{ir}(0) = \phi_{ir}c_r, \quad \dot{x}_{ir}(0) = \phi_{ir}d_r, \quad \sum_{i=1}^{N_o}\phi_{ir}^2 = 1 \tag{24}$$

One pseudostatic "mode" is also necessary:

$$\ddot{x}_{ir}(t) = \sum_{k=1}^{N_I} r_{ik}f_k(t) \tag{25}$$

This accounts for the quasi-static contributions to the structural motions induced by the support motions during the earthquake, ignoring inertial and damping effects since these are accounted for in the dynamic response contributions (Werner *et al*, 1987). The simplest pseudostatic mode is rigid-body motion as in the direct effects of rocking and translation of the base of a building. In the governing equations, $\{f_k(t) : k = 1, \ldots, N_I\}$ are linearly independent support accelerations defining the seismic input to the structure, but they could also be exciting forces. An extension of the result proved by Beck (1978) shows that the class of models is globally identifiable, at least if the measurement noise and model error are not too large.

The model parameters $\underline{a}$ defining the class of models are the modal parameters for each of the identified $(N_m - 1)$ dynamic modes, that is, the natural frequencies and damping ratios, ω_r and ζ_r, the initial modal displacement and velocity, c_r and d_r, the modeshape components $\{\phi_{ir}, i = 1, \ldots, N_o\}$, and the input participation factors $\{p_{rk}, k = 1, \ldots, N_I\}$; together with the pseudostatic influence coefficients

$\{r_{ik}, i = 1, \ldots, N_o,\ k = 1, \ldots, N_I\}$. The latter parameters may be fixed on a theoretical basis in some situations (e.g. for the pseudostatic response due to rocking and translation of the base of a building). Only the modeshape components at the observed DOF can be identified since the "missing" modeshape components at the unobserved DOF cannot be identified directly without introducing a structural model as a basis for the "interpolation".

Typically, the system output which gets least-squares matched by the model output during the parameter estimation stage (equations (17) and (19)) consists of measured accelerations at N_o =10 to 20 DOF. The optimization algorithm used in the MODE-ID program is a robust one exploiting the linearity of the model dynamics (Beck, 1978; Werner *et al*, 1987). Although MODE-ID does the output matching in the time domain, it is also possible to use the frequency domain by matching the complex Fourier transforms of the system and model outputs (McVerry, 1980). The theory in the previous section still applies because the time-domain $J(\underline{a})$ in equation (19) is essentially equivalent to the frequency-domain version by Parseval's identity.

The above uncoupled equations of motion assume classical normal modes of vibration. The method has been extended recently to a model with non-classical modes of vibration (Tan and Cheng, 1993). It should be noted, however, that the linear dynamics forming the basis of these models assumes that the structural damping is linear viscous damping. This is known to be a poor model of structural damping, which in tests on members and components exhibits a rate-independent hysteretic behavior, even at small deformation amplitudes. Still, at amplitude levels where there is no damage and no substantial yielding, so that the damping is small, the application of modal identification to seismic response data has demonstrated that the simple classical linear viscous damping model is an acceptable approximation.

Previous Applications to Seismic Structural Response

The modal identification methodology described above has been one of the few system identification techniques applied to recorded seismic response, and a significant number of structures of different types have been studied by several researchers. Most of the data studied were recorded in tall buildings during earthquakes in California (Beck, 1978; Beck and Jennings, 1980; McVerry, 1980; McVerry and Beck, 1983; Papageorgiou and Lin, 1989a,b; Li and Mau, 1991; Nisar *et al*, 1992; Hashimoto *et al*, 1993; Durrani *et al*, 1994; Mau and Aruna, 1994). These studies confirm that a small number of "linear" modes (5 to 10), can capture the translational and torsional behavior of tall buildings during earthquakes. For nondamaging response, identified modal periods are typically within 10 % to 15 % of the corresponding values calculated from theoretical (finite-element) models. These periods differ from values given by simple code-based formulas by much larger amounts, and they greatly exceed the periods identified from small-amplitude ambient vibration tests. For nondamaging response, identified equivalent viscous damping ratios for each mode are typically in the range 3 % to 8 % of critical damping.

This summary of results refers to equivalent linear modal parameters for a "time-invariant" model of a structure where the entire duration of recorded earthquake motions is used. In addition, the time variation of these parameters can be studied by using smaller time windows of data of duration two fundamental periods or more (Beck, 1978). This procedure can provide insight into the extent and nature of the nonlinearities in the structural behavior. In particular, many of the above studies have shown that substantial stiffness losses of the order of 50 % or more can occur as the strong-motion amplitudes build up during the earthquake, even if the structure is undamaged and has not undergone significant plastic yielding. This stiffness loss may be due to one or more sources: micro-cracking of concrete, loosening of connections between nonstructural components and the principal structural load-

carrying system, changes in boundary conditions such as softening in the soil foundations, and other effects. Most of the stiffness loss is recovered as the motion dies out. Studies to localize these stiffness losses are of much interest but appear to require more dense instrumentation arrays than the 10 to 20 or so accelerometers that are typically distributed over buildings at present.

The most probable prediction-error standard deviations, given by the square root of the global minimum of $J(\underline{a})$ in (19), range from about 15 % to 70 % of the RMS of the measured accelerations, with the low end corresponding to small nondamaging seismic motions and the high end corresponding to damaging response where "time-invariant" linear dynamic models do not capture well the enormous changes in stiffness that occur. A better class of models could be developed which include a dependence of the prediction-error variance on the strength of shaking. Identification of such models has become more practical in the last few years, since data for different strengths of shaking have been recorded in the same structure in several cases. For example, some buildings in San Jose, California, have had their motion measured during three earthquakes which produced substantially different lateral peak roof accelerations; in one case, the range was from 8 % to 38 % g (e.g. Nisar *et al*, 1992).

The modal identification approach has also been used to determine the level of earthquake-induced forces, base shears, overturning moments and deformations compared with code design values. This is done by producing complete modeshapes from the identified components through some form of simple "interpolation"; for example, fitting a curve through the known floor modeshape components to estimate the components at the uninstrumented floors (e.g. Nisar *et al*, 1992). A conventional modal time-history analysis can then be performed because a complete set of modal parameters is available for those modes contributing significantly to the seismic response. These studies have shown that earthquake-induced base shears and overturning moments can exceed design levels by factors of 2 to 3 without any structural damage, confirmng that structures designed to code levels often have considerable "overstrength". The same studies of undamaged buildings confirm that the interstory drift ratios correlate better with observed seismic performance since the earthquake-induced drift ratios were no more than a few tenths of a percent, even when one building experienced a lateral peak roof acceleration of 38% g during the 1989 Loma Prieta earthquake in California (Nisar *et al*, 1992). The same type of system identification analyses have yet to be published for the few buildings that have been instrumented and suffered structural damage during an earthquake (e.g. CDMG/CSMIP station number 24386, a 7-story reinforced-concrete hotel in Van Nuys which suffered severe shear damage in some of the columns during the 1994 Northridge earthquake in California). A "time-varying" modal model based on identification using a moving time window would be more appropriate in these cases.

The modal identification methodology has also been applied to recorded seismic response from a highway bridge (Werner *et al*, 1987) and an off-shore oil platform (Mason *et al*, 1989). These studies have shown the levels of structural damping exhibited during strong shaking. This type of information is valuable since it cannot be derived on theoretical basis. In addition, to examine the nonlinear effects of the earth embankments on the response of the highway bridge, quick-release (snap-back) tests were performed and the identified modal parameters were compared with those identified from much stronger earthquake motions (Werner *et al*, 1990, 1994). This study showed that softening of the earth approach embankments at large amplitudes has a significant effect on the dynamics of the bridge, which should be considered during design.

APPLICATION TO STRUCTURAL MODEL UPDATING

Over the last decade, the challenging problem of *structural model updating* has gained much interest as finite-element modelling capabilities and modal testing have become more mature areas of structural

dynamics; see, for example, the survey articles by Natke (1988) and Mottershead and Friswell (1993). The motivation to solve the model updating problem comes from the desire to reconcile a theoretically-based finite-element model of a structure with dynamic test data from the structure, since there is never complete agreement between these. Usually linearity is assumed and the reconciliation is done by adjusting the finite-element model so that either the calculated response time histories, frequency response functions or modal parameters "best" match the corresponding quantities measured or identified from the test data. This is therefore a system identification problem. There is, however, no well-accepted solution to the model updating problem and this is primarily because of an inherent difficulty: there is a mismatch between the level of information in the detailed theoretical finite-element model, which is of uncertain accuracy, and the sparse information in the "incomplete" set of "noisy" test data. This produces an ill-conditioned and often nonunique inverse problem in updating the theoretical model.

This nonuniqueness problem is particularly acute in model updating using measured seismic response because of the low density of sensors typical of seismic instrumentation arrays installed on structures. In fact, this application is only in the early stages of development, and there are apparently no applications to actual earthquake data. It is important, however, because it offers a systematic way of determining and correcting the weaknesses in a finite-element model of a structure so that more accurate and complete response predictions can be made. Also, once a finite-element model has been "calibrated" for a structure using earthquake or dynamic test data, updating the stiffness matrix with subsequent data provides a tool for rapidly detecting, locating and assessing structural damage. For this application, it is assumed that local loss of stiffness corresponds to damage at the location. If a viable technology with this capability had been installed in buildings at the time of the 1994 Northridge earthquake, along with near real-time data retrieval, it would not have taken several weeks to detect the serious problem of failed connections in many steel-frame buildings (SAC, 1995).

An important feature of the general Bayesian statistical framework for system identification presented here is that it can deal with a class of models which are not globally identifiable, and so it can deal with the nonuniqueness in model updating; at least if there is sufficient data to produce system identifiability, that is, to produce only a finite number of optimal models. Beck and Katafygiotis (1991, 1992) have used the methodology to investigate the model-updating problem for substructured linear finite-element models, although only with simulated dynamic data. They showed how to find all the optimal models consistent with the data and how to use the asymptotic approach described here to make statistical predictions of the structural response based on all of these models. An example they give is briefly described here.

Consider a planar-frame building of N stories which is modelled as a "shear" building with N unknown interstory stiffnesses but known mass distribution. If the horizontal seismic motion is measured at the base and at the roof only, Udwadia (1985) showed that the inverse problem of estimating the model stiffnesses from these data is nonunique and he proved that $N!$ is an upper bound on the number of solutions. Beck and Katafygiotis (1991) presented a new algorithm to find all the stiffness solutions. They showed that for a uniform six-story building, for example, there are a total of eight stiffness solutions (optimal models) which give identical response at the roof for any base excitation. Each model, however, gives distinct predictions at the lower floors where, by assumption, the motion was not measured. Beck and Katafygiotis (1992) calculated the weighting coefficients in equation (14) for this example and showed how they were influenced by the choice of the initial PDF $\pi(\underline{\alpha})$ in equation (5), which was chosen with the damage detection application in mind. They showed that despite the multiple optimal models, the probability of damage (decreased stiffness) could be computed and used to assess where damage had occurred. The example is an idealized one, but it does serve to demonstrate the usefulness of the general framework presented here.

CONCLUSION

A unified Bayesian statistical framework is described for system identification which can be used to extract important information for earthquake-resistant design from the measured seismic response of structures. In this approach, the "best" (optimal) models within a chosen class of models are those which are locally most probable based on the available data. Prediction using only these optimal models is asymptotically correct for large samples compared with the full Bayesian predictive solution for the class of models. In the case of multiple optimal models, response predictions are given by a weighted average of the predictive probability models corresponding to each such model. The methodology allows a more comprehensive interpretation of the least-squares output-error approach to system identification which is the dominant method for modal identification from measured seismic response. In particular, it allows the uncertain prediction accuracy of the identified models to be quantified. It can also handle the nonuniqueness expected in updating a finite-element model using measured seismic response.

ACKNOWLEDGEMENTS

The author is currently on sabbatical leave at HKUST (Hong Kong University of Science and Technology) and the kind hospitality of Prof. J.C. Chen, Director of Research Center, is greatly appreciated. Discussions with Drs. Lambros Katafygiotis and Costas Papadimitriou at HKUST have also been most stimulating.

REFERENCES

Bayes, T. (1763). An essay toward solving a problem in the doctrine of chances. *Phil. Trans. Royal Society*, **53**, 330-418. (Reprinted 1958 in *Biometrika*, **45**, 293-315).

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

Beck, J.L. and P.C. Jennings (1980). Structural identification using linear models and earthquake records. *Earthq. Engng and Struct. Dyn.*, **8**, 145-160.

Beck, J.L. (1982). System identification applied to strong motion records from structures. In: *Earthquake Ground Motion and its Effects on Structures* (S.K. Datta, Ed.), AMD-53, ASME, New York.

Beck, J.L. and M. Dowling (1988). Quick algorithms for computing either displacement, velocity or acceleration of an oscillator. *Earthq. Engng and Struct. Dyn.*, **16**, 245-253.

Beck, J.L. (1989). Statistical system identification of structures. *Proc. 5th Int. Conf. on Structural Safety and Reliability*, 1395-1402, ASCE, New York.

Beck, J.L. and L.S. Katafygiotis (1991). Updating of a model and its uncertainties utilizing dynamic test data. *Proc. 1st Int. Conf. on Computational Stochastic Mechanics*, 125-136, Computational Mechanics Publications, Boston.

Beck, J.L. and L.S. Katafygiotis (1992). Probabilistic system identification and health monitoring of structures. *Proc. 10th World Conf. on Earthquake Engineering*, Madrid, Spain.

Beck, J.L. and L.S. Katafygiotis (1996). Updating structural dynamic models and their uncertainities: Part I - Statistical system identification. *J. Engng Mech.*, submitted.

Box, G.E. and G.C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Addison-Wesley, Reading.

Bretthorst, G.L. (1988). *Bayesian Spectrum Analysis and Parameter Estimation.* Springer Lecture Notes on Physics No. 48, Springer-Verlag, Berlin.

Cox, R. T. (1961). *The Algebra of Probable Inference.* Johns Hopkins Press, Baltimore.

De Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment.* Wiley, New York.

Durrani, A.J., S.T. Mau, A.A. AbouHashish and Y. Li (1994). Earthquake response of flat-slab build-

ings. *J. Struct. Engng*, **120**, 947-964.

Eykhoff, P. (1974). *System Identification: Parameter and State Estimation*. Wiley, New York.

Goodwin, G.C. and R.L. Payne (1977). *Dynamic System Identification, Experiment Design and Data Analysis*. Academic Press, New York.

Gull, S.F. (1988). Bayesian inductive inference and maximum entropy. In: *Maximum-Entropy and Bayesian Methods in Science and Engineering, Vol. 1: Foundations* (G.J. Erickson and C.R. Smith, Eds), Kluwer Academic, Boston.

Hall, J.F. (Ed.) (1994). *EERI Northridge Earthquake Reconnaissance Report*, Earthquake Engineering Research Institute, Oakland, CA. Also in *Earthquake Spectra*, **11**, Supplement C, April, 1995.

Hart, G.C. and J.T.P. Yao (1976). System identification in structural dynamics. *J. Engng Mech.*, **103**, 1089-1104.

Hashimoto, P.S., L.W. Tiong, L.K. Steele, J.J. Johnson and J.L. Beck (1993). *Stiffness and Damping Properties of a Low Aspect Ratio Shear Wall Building Based on Recorded Earthquake Responses*. Report NUREG/CR-6012, U.S. Nuclear Regulatory Commission, Washington, D.C.

Hoshiya, M. and E. Saito (1984). Structural identification by extended Kalman filter. *J. Engng Mech.*, **110**, 1757-1770.

Imai, H., C.B. Yun, O. Maruyama and M. Shinozuka (1989). Fundamentals of system identification in structural dynamics. *Prob. Engng Mech.*, **4**, 162-173.

Jaynes, E.T. (1968). Prior probabilities. *IEEE Trans. Syst. Sci. and Cybernetics*, SSC-4, 227-241.

Jaynes, E.T. (1978). Where do we stand on maximum entropy? In: *The Maximum Entropy Formalism* (R.D. Levine and M. Tribus, Eds), MIT Press, Cambridge, MA.

Jaynes, E.T. (1983). *E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics* (R.D. Rosenkrantz, Ed.), D. Reidel Publishing Co., Dordrecht, Holland. (Reprinted 1989, Kluwer Academic, Boston).

Jaynes, E.T. (1988). How does the brain do plausible reasoning? In: *Maximum-Entropy and Bayesian Methods in Science and Engineering, Vol. 1: Foundations* (G.J. Erickson and C.R. Smith, Eds), Kluwer Academic, Boston. (Reprint of 1957 Microwave Laboratory Rept. No. 421, Stanford University).

Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press, Oxford, 3rd ed. (1st ed., 1939).

Laplace, P.S. (1812). *Théorie Analytique des Probabilités*. Reprinted in Vol. 7, *Oeuvres Complètes*, Gauthier-Villars, Paris.

Li, Y. and S.T. Mau (1991). A case study of MIMO system identification applied to building seismic records. *Earthq. Engng and Struct. Dyn.*, **20**, 1045-1064.

Lindley, D.V. (1965). *Introduction to Probability and Statistics from a Bayesian Viewpoint*. Cambridge University Press, Cambridge, England.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, New Jersey.

Mason, A.B., J.L. Beck, J. Chen and R.R. Ullmann (1989). Modal parameter identification of an off-shore platform from earthquake response records. *Proc. Sessions Related to Seismic Engng at Struct. Congress*, 271-226, ASCE, New York.

Mau, S.T. and V. Aruna (1994). Story drift, shear and OTM estimation from building seismic records. *J. Struct. Engng*, **120**, 3366-3385.

McVerry, G.H. (1980). Structural identification in the frequency domain from earthquake records. *Earthq. Engng and Struct. Dyn.*, **8**, 161-180.

McVerry, G.H. and J.L. Beck (1983). *Structural identification of JPL Building 180 Using Optimally Synchronized Earthquake Records*. EERL Rept No. 83-01, California Institute of Technology, Pasadena.

Mottershead, J.E. and M.I. Friswell (1993). Model updating in structural dynamics: a survey. *J. Sound and Vibration*, **167**, 347-375.

Natke, H.G. (Ed.) (1982). *Identification of Vibrating Structures*. Springer-Verlag, New York.

Natke, H.G. (1988). Updating computational models in the frequency domain based on measured data: a survey. *Prob. Engng Mech.*, **8**, 28-35.

Natke, H.G. and J.T.P. Yao (Eds) (1988). *Proc. Int. Workshop on Structural Safety Evaluation based on System Identification Approaches*. Vieweg and Sons, Wiesbaden.

Nisar, A., S.D. Werner and J.L. Beck (1992). Assessment of UBC seismic design provisions using recorded building motions. *Proc. 10th World Conf. on Earthquake Engineering*, Madrid, Spain.

Papageorgiou, A.S. and B.C. Lin (1989a). Influence of lateral-load-resisting system on the earthquake response of structures - a system identification study. *Earthq. Engng and Struct. Dyn.*, **18**, 799-814.

Papageorgiou, A.S. and B.C. Lin (1989b). Study of the earthquake response of the base-isolated Law and Justice Center in Rancho Cucamonga. *Earthq. Engng and Struct. Dyn.*, **18**, 1189-1200.

Pilkey, W.D. and R. Cohen (Eds) (1972). *System Identification of Vibrating Structures*. ASME, N.Y.

Ramsey, F.P. (1931). *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, London.

SAC (1995). *Interim Guidelines: Evaluation, Repair, Modification and Design of Welded Steel Moment Frame Structures*. Rept. No. SAC-95-02, SAC Joint Venture, Sacramento, CA.

Savage, L.J. (1972). *The Foundations of Statistics*. 2nd rev. ed., Dover, New York (1st ed., 1954).

Shakal, A., M. Huang and C. Ventura (1988). The California strong motion instrumentation program: objectives, status and recent data. *Proc. 9th World Conf. on Earthquake Engineering*, Tokyo, Japan.

Shinozuka, M., C.B. Yun and H. Imai (1982). Identification of linear structural dynamics systems. *J. Struct. Engrg, ASCE*, **108**, 1371-1390.

Tan, R.Y. and W.M. Cheng (1993). System identification of a non-classically damped linear system. *Computers and Structures*, **46**, 67-75.

Udwadia, F.E. (1985). Some uniqueness results related to soil and building structural identification. *SIAM J. of Applied Math.*, **45**, 674-685.

Werner, S.D., J.L. Beck and M.B. Levine (1987). Seismic response evaluation of Meloland Road Overpass using 1979 Imperial Valley earthquake records. *Earthq. Engng and Struct. Dyn.*, **15**, 249-274.

Werner, S.D., J.L. Beck and A. Nisar (1990). Dynamic tests and seismic excitation of a bridge structure. *Proc. 4th U.S. National Conf. on Earthq. Engng*, Palm Springs, EERI, California.

Werner, S.D., C.B. Crouse, L.S. Katafygiotis and J.L. Beck (1994). Use of strong motion records for model evaluation and seismic analysis of a bridge structure. *Proc. 5th U.S. National Conf. on Earthq. Engng*, Chicago, EERI, California.

West, M. and J. Harrison (1989). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, N.Y.

Yang, C.M. and J.L. Beck (1996). Generalized trajectory methods for finding multiple extrema and roots of functions. *Journal of Optimization Theory and Applications*, submitted.