ECCOMAS Thematic Conference on Computational Methods
in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, D.C. Charmpis, N.D. Lagaros, Y. Tsompanakis (eds.)
Rethymno, Crete, Greece, 13–16 June 2007

BAYESIAN UPDATING AND MODEL CLASS SELECTION OF DETERIORATING HYSTERETIC STRUCTURAL MODELS USING SEISMIC RESPONSE DATA

James L. Beck

California Institute of Technology
Mail Code 104-44, Pasadena, CA 91125, USA
e-mail: jimbeck@caltech.edu

Matthew Muto

California Institute of Technology
Mail Code 104-44, Pasadena, CA 91125, USA
e-mail: muto@caltech.edu

Keywords: Hysteretic Structural Models, Bayesian Updating, Model Class Selection

Abstract. *Identification of structural models from measured earthquake response can play a key role in structural health monitoring, structural control and improving performance-based design. System identification using data from strong seismic shaking is complicated by the nonlinear hysteretic response of structures where the restoring forces depend on the previous time history of the structural response rather than on an instantaneous finite-dimensional state. Furthermore, this inverse problem is ill-conditioned because even if some components in the structure show substantial yielding, others will exhibit nearly elastic response, producing no information about their yielding behavior. Classical least-squares or maximum likelihood estimation will not work with a realistic class of hysteretic models because it will be unidentifiable based on the data. On the other hand, Bayesian updating and model class selection provide a powerful and rigorous approach to tackle this problem when implemented using Markov Chain Monte Carlo simulation methods such as the Metropolis-Hastings, Gibbs Sampler and Hybrid Monte Carlo algorithms. The emergence of these stochastic simulation methods in recent years has led to a renaissance in Bayesian methods across all disciplines in science and engineering because the high-dimensional integrations that are involved can now be readily evaluated. The power of these methods to handle ill-conditioned or unidentifiable system identification problems is demonstrated by using a recently-developed stochastic simulation algorithm, Transitional Markov Chain Monte Carlo, to perform Bayesian updating and model class selection on a class of Masing hysteretic structural models that are relatively simple yet can give realistic responses to seismic loading. Examples will be given using deteriorating hysteretic building models with simulated seismic response data.*

1 INTRODUCTION

Current methods for developing finite-element models can produce structural responses that are consistent qualitatively with behavior observed during strong earthquake shaking, but there has long been an interest in using system identification methods for quantitative assessment of structural models using recorded seismic response. The objective may be to improve the predictive capabilities of structural models for dynamic design or for the design of structural control systems, or to implement structural health monitoring. System identification based on updating of finite-element models using measured seismic response is challenging, however, because the large number of uncertain parameters associated with realistic structural models makes the inverse problem extremely ill-conditioned.

Simplified models can be used in the identification procedure but the selection of an appropriate class of models to employ is complicated by the nonlinear response of structures under strong seismic loading; in particular, the structural restoring forces are hysteretic, depending on the previous time history of the structural response rather than on an instantaneous finite-dimensional state. Although some research into the identification of hysteretic systems has been carried out, this previous work [1,2,3,4] did not quantify the modeling uncertainties and did not properly deal with the ill-conditioning inherent in this inverse problem. However, the uncertainty associated with structural model predictions can have a significant impact on the decision-making process in structural design, control and health monitoring. Furthermore, classical estimation techniques such as least-squares and maximum likelihood do not usually work properly when applied to hysteretic model classes because they are nearly always unidentifiable based on the available data.

The Bayesian updating approach treats the probability of all models within a set of candidate models for a system, and consequently has the advantage of being able to quantify all of the uncertainties associated with modeling of a system and to handle ill-conditioned identification problems. Note that the probability of a model will not make sense if one interprets probability as a long-run frequency of an event, but it does when probability is interpreted as a multi-valued logic that expresses the degree of plausibility of a proposition conditioned on the given information [5]. Although Bayesian methods are widely used in many fields, their application to identification of dynamic hysteretic dynamic models seems to be very limited.

Many applications of Bayesian methods to model updating and model class selection for systems using dynamic response measurements have primarily used the Laplace asymptotic approximation. However, the approximation is most useful when there is a large amount of data and the model class is globally identifiable (described later); furthermore, in high-dimensional systems, optimization to find the required most probable parameter vectors can be computationally challenging. To avoid these difficult optimizations and to more readily treat cases where the model class is not globally identifiable, in recent years attention has been focused on stochastic simulation methods for Bayesian updating and prediction, especially Markov Chain Monte Carlo methods, such as the Metropolis-Hastings, Gibbs Sampler and Hybrid Monte Carlo algorithms [6]. The emergence of these stochastic simulation methods has led to a renaissance in Bayesian methods across all disciplines in science and engineering because the high-dimensional integrations involved can now be readily evaluated.

2 BAYESIAN MODEL UPDATING

A Bayesian statistical framework for model updating and predictions for linear or nonlinear dynamic systems that explicitly treats prediction-error and other model uncertainties has been presented [7,8,9]. A basic concept in this framework is that any set of possible determi-

nistic dynamic models for a system can be embedded in a set of predictive probability models for the system by specifying a probability distribution for the uncertain prediction error, which is the difference between the actual system output and the deterministic model output. In particular, modeling the prediction error as a zero-mean, stationary, white-noise Gaussian stochastic process is supported by the principle of maximum differential entropy [5]. Each predictive probability model is assumed to be uniquely specified by assigning a value to a model parameter vector. Therefore, a probability distribution over the set of possible predictive models that specifies the plausibility of each such model is equivalent to a probability distribution over a corresponding set of possible values for the model parameter vector. When dynamic data is available from the system, a chosen initial (prior) probability distribution over the parameters can be updated using Bayes' Theorem to give a posterior probability distribution, as follows.

Consider a *Bayesian model class* $\mathcal{M}$, which is characterized by: (i) a set of *predictive* PDFs, $p(\mathcal{D} \mid \theta, \mathcal{M})$, for system response $\mathcal{D}$ that is parameterized by N_p model parameters $\theta \in \Theta \in \mathbb{R}^{N_p}$; and (ii) a chosen *prior* PDF $p(\theta \mid \mathcal{M})$ that can incorporate existing knowledge of the system. The prior PDF is chosen to express the initial plausibility of each model in the class $\mathcal{M}$ defined by the value of the parameter vector θ.

Now suppose a set of data $\mathcal{D}$ from the system is available. The goal of Bayesian updating is to use $\mathcal{D}$ to update the probability distribution over the parameters to give the posterior PDF $p(\theta \mid \mathcal{D}, \mathcal{M})$ based on Bayes' Theorem:

$$p(\theta \mid \mathcal{D}, \mathcal{M}) \propto p(\mathcal{D} \mid \theta, \mathcal{M})\, p(\theta \mid \mathcal{M}) \tag{1}$$

Here, $p(\mathcal{D} \mid \theta, \mathcal{M})$ as a function of θ is called the *likelihood function*. The constant of the proportionality is the reciprocal of $p(\mathcal{D} \mid \mathcal{M})$, the *evidence* for model class $\mathcal{M}$, and it is discussed later. The posterior PDF gives the updated plausibility of each model in $\mathcal{M}$ when the information in the data $\mathcal{D}$ is incorporated.

For a given model class $\mathcal{M}$ and data $\mathcal{D}$, it is useful to characterize the topology of the posterior PDF as a function of the model parameter vector by whether it has a global maximum at a single most probable parameter value, at a finite number of them, or at a continuum of most probable parameter values lying on some manifold in the parameter vector space. These three cases may be described as globally identifiable, locally identifiable, and unidentifiable model classes based on given dynamic data from the system.

3 BAYESIAN MODEL CLASS SELECTION

Bayesian model class selection (or model comparison) is essentially Bayesian updating at the model class level to make comparisons between alternative candidate model classes for predicting the response of a system. It has long been recognized that comparisons between model classes should factor in not only the quality of the data fit, but also the complexity of the model. Jeffreys referred to the need for a "simplicity postulate," that is, simpler models that are consistent with the data should be preferred over more complex models which offer only slight improvements in the fit to the data [10]. Early quantitative forms for a Principle of Model Parsimony utilized a penalty against using a larger number of uncertain (adjustable) parameters in combination with a quantification of the model data-fit based on the log likelihood of the optimal model in the model class; however, the form of these penalty term did not have a very rigorous basis. Subsequent work made it clear that Bayes' Theorem at the model class level automatically enforces model parsimony without ad-hoc penalty terms [11,12].

Consider a set $\mathbf{M} \equiv \{\mathcal{M}_1,\mathcal{M}_2,\ldots,\mathcal{M}_{NM}\}$ of N_M candidate model classes for representing a system. Given data $\mathcal{D}$, the posterior probability of each model class $P(\mathcal{M}_j | \mathcal{D},\mathbf{M})$, j=1,…, N_M, is:

$$P\left(\mathcal{M}_j \mid \mathcal{D},\mathbf{M}\right) = \frac{p\left(\mathcal{D} \mid \mathcal{M}_j\right) P\left(\mathcal{M}_j \mid \mathbf{M}\right)}{\sum_{i=1}^{N_M} p\left(\mathcal{D} \mid \mathcal{M}_i\right) P\left(\mathcal{M}_i \mid \mathbf{M}\right)} \tag{2}$$

If all model classes are treated as equally plausible a priori, then the probability of model class $\mathcal{M}_j$ is proportional to its evidence,

$$p\left(\mathcal{D} \mid \mathcal{M}_j\right) = \int p\left(\mathcal{D} \mid \theta_j,\mathcal{M}_j\right) p\left(\theta_j \mid \mathcal{M}_j\right) d\theta_j \tag{3}$$

3.1 Information-theoretic interpretation

Further insight into the form of this penalty against complexity can be obtained by considering the evidence from an information-theoretic point of view [13]. Consider the log of the evidence:

$$\begin{aligned} \ln\left[p\left(\mathcal{D} \mid \mathcal{M}_j\right)\right] &= \int \ln\left[\frac{p\left(\mathcal{D} \mid \theta_j,\mathcal{M}_j\right) p\left(\theta_j \mid \mathcal{M}_j\right)}{p\left(\theta_j \mid \mathcal{D},\mathcal{M}_j\right)} \right] p\left(\theta_j \mid \mathcal{D},\mathcal{M}_j\right) d\theta_j \\ &= \int \ln\left[p\left(\mathcal{D} \mid \theta_j,\mathcal{M}_j\right)\right] p\left(\theta_j \mid \mathcal{D},\mathcal{M}_j\right) d\theta_j - \int \ln\left[\frac{p\left(\theta_j \mid \mathcal{D},\mathcal{M}_j\right)}{p\left(\theta_j \mid \mathcal{M}_j\right)} \right] p\left(\theta_j \mid \mathcal{D},\mathcal{M}_j\right) d\theta_j \end{aligned} \tag{4}$$

This formulation for the log evidence for model class $\mathcal{M}_j$ shows that it is the difference between two terms: the first term is the posterior mean of the log-likelihood function, which is a measure of the average data fit for model class $\mathcal{M}_j$, while the second term is the relative entropy between the prior and posterior distributions, which is a measure of the information gained about the parameters θ_j from the data $\mathcal{D}$. Therefore, the log evidence is comprised of a data-fit term and a term which provides a penalty against more "complex" models that extract more information from the data. This gives an intuitive understanding of why the application of Bayes' Theorem at the model class level automatically enforces Ockham's razor. Although this information-theoretic interpretation was initially presented in [12], it was derived there using a large-sample Laplace asymptotic approximation that depended on global identifiability of the model classes.

4 APPLYING BAYESIAN METHODS USING STOCHASTIC SIMULATION

The goal of the stochastic simulation methods is to generate samples which are distributed according to the posterior probability density function (PDF) described in Equation 1. In this work, we focus specifically on Markov Chain Monte Carlo (MCMC) methods that are most useful for Bayesian updating. One advantage of these methods is that non-normalized PDFs can be sampled, so that samples may be drawn from the posterior PDF without evaluating the normalizing constant (the evidence) that usually requires evaluating a high-dimensional integral over the parameter space. A remaining challenge associated with model updating by stochastic simulation is the fact that, unless the data is very sparse, the posterior PDF occupies a much smaller volume in the parameter space than the prior PDF over the parameters. This fact makes it difficult to draw samples from the posterior PDF.

Commonly-implemented MCMC methods, such as the Metropolis-Hastings (M-H) algorithm, are difficulty to apply in higher-dimensional parameter spaces since it is often difficult to draw samples that cover all regions of high-probability content. An alternative sampling algorithm [14] proposed gradual updating of the model, using the M-H algorithm to sample from a sequence of target PDFs, each target PDF being the posterior PDF based on an increasing fraction of the available data. In this manner, the target PDF gradually converges from the broad prior PDF to the final concentrated posterior PDF. The Transitional Markov Chain Monte Carlo (TMCMC) method used in this study is a modified version of this approach [15]. This technique also uses a sequence of intermediate PDFs. The novel feature of this algorithm is that, rather than applying updating with part of the available data, the entire data set is used but its full effect is diluted by taking the target PDF for the m^{th} level of the sampler to be proportional to $p(\mathcal{D}|\theta,\mathcal{M})^{\beta_m}\, p(\theta|\mathcal{M})$, where $0 \le \beta_m \le 1$; here, $\beta_0 = 0$ gives the initial target distribution proportional to the prior PDF and $\beta_M = 1$ for the final level of the sampler gives a target distribution proportional to the posterior PDF. The TMCMC algorithm can also be used to estimate the evidence for a model class [15].

5 MASING HYSTERETIC MODELS

Modeling hysteretic force-deformation relations for structural members and assemblages of members from constitutive equations ("plasticity models") is a difficult task, due to factors such as complex stress distributions, material inhomogeneities and the large number of structural elements. An alternative approach is to develop simplified models that capture the essential features of the hysteretic force-deformation relationship but then, lacking a fundamental theoretical basis, these models should be validated against the observed behavior of structures. An example of this type of model is the well-known Bouc-Wen model. While these models are mathematically convenient, especially for random vibration studies using equivalent linearization, when subjected to asymmetric cyclic loading, they can exhibit an unphysical "drifting" behavior [1]. This makes them unsuitable as a class of identification models for strong seismic response where this type of irregular loading occurs.

A simplified hysteretic model with a physical basis was presented by Masing [16], which is based on the hypothesis that a one-dimensional hysteretic system may be viewed as a collection of ideal elasto-plastic elements (a linear spring in series with a Coulomb damper) with the same elastic stiffness but with a distribution of different yield strengths. This idea was used in structural dynamics by Iwan to form the Distributed Element Model (DEM), which consists of a collection of N ideal elasto-plastic elements connected in parallel [17] with a common stiffness k/N for the springs but different yield strengths $r_i^*/N, i=1,\ldots,N$, as shown in Figure 1. The restoring force r of a single-degree of freedom DEM subjected to a displacement x under intial loading is given by:

$$r = \sum_{i=1}^{n} \frac{r_i^*}{N} + kx\frac{N-n}{N} \tag{5}$$

where n is the number of yielded elements. Infinite collections of elasto-plastic elements can be considered by introducing a yield strength distribution function $\varphi(r^*)$ such that the restoring force $r(x)$ is given by:

$$r(x) = \int_0^{kx} r^* \phi(r^*)\, dr^* + kx \int_{kx}^{\infty} \phi(r^*)\, dr^* \tag{6}$$

Because there is an underlying physical basis for the model, DEMs with a finite number of elements have been shown to give good representations of the hysteretic behavior of some structures, and do not exhibit the previously-discussed drifting behavior. However, DEMs with an infinite number of elements are difficult to implement directly, in contrast to the finite case where the state of each element is tracked, although there have been recent advances in this area [4]. Fortunately, the class of Masing hysteretic models exactly describe the behavior of DEMs without needing to keep track of the internal behavior of the elements. Jayakumar showed that the hysteretic behavior under arbitrary loading is completely described by the initial loading curve, described by the function $f(x,r)$ and a relatively simple set of rules [1]. Chiang [18] later demonstrated the inverse relationship, that is, given an initial loading curve for a Masing model, one can find the yield strength distribution, $\varphi(r^*)$ in Equation 6, for the equivalent DEM.

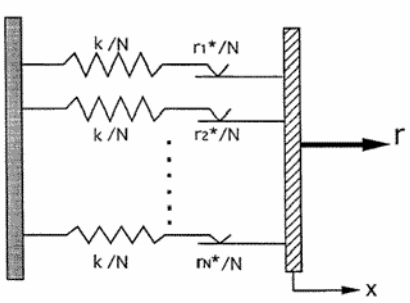

Figure 1: Conceptual sketch for the Distributed Element Model (DEM). Taken from [18].

5.1 Masing Shear Building Model

Jayakumar and Beck [2] modeled an n-story shear building in 2-D by applying the Masing model to the relationship between story shear forces and the inter-story drifts. Consider a structural model where the vector of relative floor displacements $x(t)$ is related to the ground acceleration $\ddot{y}(t)$ as follows:

$$M\ddot{x} + C\dot{x} + R = -Mb\ddot{y}(t) \tag{7}$$

where M is the mass matrix, C is the viscous damping matrix, and R is the vector of restoring forces. The inter-story shear force at the i^{th} story is given by:

$$R_i = r_i - r_{i+1} \tag{8}$$

In this work, the initial loading curve relating story shear forces and inter-story drifts is specified by choosing a generalized Rayleigh distribution for the yield strength distribution function [18]. The resulting backbone curve is defined by following differential equation:

$$\dot{r}_i = K_i\left(\dot{x}_i - \dot{x}_{i-1}\right)\exp\left[-\left(\Gamma\left(1+\frac{\eta_i+1}{\eta_i}\right)\frac{K_i\left(x_i - x_{i-1}\right)}{r_{u,i}}\right)^{\eta_i}\right] \tag{9}$$

where K_i is the small-amplitude inter-story stiffness, $r_{u,i}$ is the story ultimate strength, η_i controls the smoothness of the transition from elastic to plastic and $\Gamma(.)$ is the Gamma function. Figure 2 shows how the shape of the initial loading curve is influenced by η_i. Note that for $i = n$ in Equation 8, r_{n+1}=0, and for $i = 1$ in Equation 9, x_0=0.

A potentially important advantage of the Masing shear building model is that all model parameters, except η_i, correspond to actual physical properties (initial stiffness, ultimate strength) and initial estimates can be calculated from material properties and structural drawings.

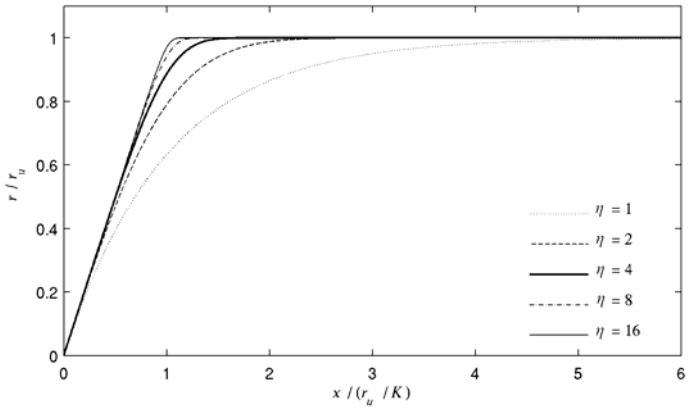

Figure 2: Initial loading curves for different values of the elastic-to-plastic transition parameter η.

5.2 Deteriorating Masing Models

Cifuentes and Iwan introduced a modified version of the DEM for modeling deteriorating systems [3]. The model again consists of a collection of linear springs and slip elements, however, in this case an element is allowed to "break" if a certain maximum displacement is exceeded, defined as $\mu x_{y,i}$, where $x_{y,i}$ is the yield displacement of the i^{th} element and μ is the breaking ductility ratio, which for simplicity was assumed to be the same for all elements.

The deteriorating DEM was successfully applied to system identification and damage detection of real structures using earthquake data [3]. However, as in the case of non-deteriorating DEMs, there are restrictions in terms of the number of parameters that limit the applicability of the model. Chiang developed a general formulation for deteriorating Masing, determined the specific form for a Masing model equivalent to the displacement-controlled DEM, and developed expressions for the initial loading, unloading and reloading curves given a backbone curve $f(x,r)$ and the breaking ductility ratio μ. Figure 3 shows how the monotonic loading curve for a deteriorating Masing model, with a non-deteriorating form shown in Figure 2, is influenced by μ.

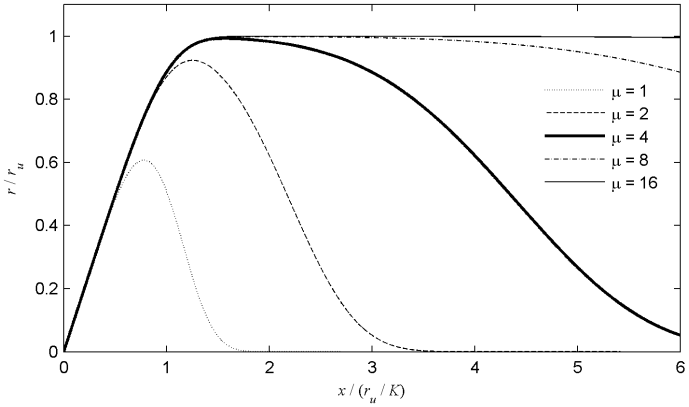

Figure 3: Monotonic loading curves for deteriorating Masing hysteretic model for different values of the breaking ductility ratio, μ.

6 EXAMPLE

To illustrate the application of Bayesian methods to ill-conditioned systems, we consider a three-story deteriorating Masing shear building model, with parameters as given in Table 1. The system is subjected to two different ground motion excitations, both recorded during the 1994 Northridge, California earthquake (M_w=6.7). The first record was obtained at the Coldwater Canyon School in North Hollywood, 12.5 km from the fault, with a peak acceleration of 3.13 m/s^2. The second was recorded at Olive View Hospital in Sylmar, 9.9 km from the fault, with a peak acceleration of 5.92 m/s^2. The first ten seconds of each record are used to generate acceleration responses for the system. The data sets generated from the Coldwater Canyon and Olive View records will be referred to as $\mathcal{D}_{CC}$ and $\mathcal{D}_{OV}$, respectively. The viscous damping matrix C in Equation 5 is omitted, as this is study is primarily concerned with identifying properties associated with large structural deformations, while viscous damping is generally used to model small-amplitude energy dissipation.

Story	Mass (kg)	K (N/m)	r_u (N)	η	μ
1st	1.25×10^5	2.50×10^8	1.75×10^6	2	6.5
2nd	1.25×10^5	2.50×10^8	1.75×10^6	2	6.5
3rd	1.00×10^5	2.00×10^8	1.40×10^6	2	6.5

Table 1: Parameters for Masing shear building system used to generate data.

For each ground-motion record, 500 time-steps of simulated acceleration are "measured" at each floor. To provide a realistic level of prediction error, Gaussian discrete white noise with a standard deviation of 1 m/s^2 , corresponding to approximately 40% of the RMS value of the acceleration data, is added to each channel of data.

Two identification model classes are considered, both of which are Masing hysteretic shear-building models as defined by Equation 7-9. These model classes are used to generate vectors of predicted floor accelerations $a_t^{(i)}$, i = 1,2,3, t = 1,…500. The prediction-error for the system output is assumed to be Gaussian with the variance of σ^2 for each measurement. Therefore, the form of the likelihood function $p(\mathcal{D} \mid \theta,\mathcal{M})$ is:

$$p\left(\mathcal{D} \mid \theta, \mathcal{M}\right) = \frac{1}{\left(2\pi\sigma^2\right)^{\frac{1500}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{3}\sum_{t=1}^{500}\left(a_t^{(i)}\left(\theta\right) - \hat{a}_t^{(i)}\right)^2\right] \tag{10}$$

where $\hat{a}_t^{(i)}$ is the measurement for channel i at time-point t and θ is the vector of parameters to be updated. Model class $\mathcal{M}_1$ is a non-deteriorating model with ten free parameters, the small-amplitude stiffness K_i, ultimate strength $r_{u,i}$ and elastic-to-plastic transition parameter η_i for each story, i = 1,2,3, and the prediction-error variance σ. Model class $\mathcal{M}_2$ is a deteriorating model with the same ten free parameters included in $\mathcal{M}_1$ and one additional parameter, the breaking ductility ratio μ, which is constrained such that $\mu_i = \mu$, i = 1,2,3. For both model classes, the mass matrix M is assumed to be known, which is a reasonable assumption, given that masses can be accurately computed from structural drawings.

Prior PDFs for the inter-story stiffness, strength, and breaking ductility ratios, were taken to be independent lognormal distributions with logarithmic means of $\ln(2.50\times10^8)$, $\ln(1.75\times10^8)$ and $\ln(8)$, respectively, and a lognormal standard deviation of 0.5. The prior PDFs for each of the elastic-to-plastic transition parameters were also lognormal, with a logarithmic mean of $\ln(2)$ and a logarithmic standard deviation of 1. The prior PDF for the prediction-error variance was taken to be a uniform distribution between 0 and 3.

Samples from the posterior PDF were generated using the Transitional Markov Chain Monte Carlo algorithm. Three runs were performed for updating with each data set, with 600 samples generated per run. Tables 2 and 3 show the sample means for each parameter for updating with data sets $\mathcal{D}_{CC}$ and $\mathcal{D}_{OV}$, respectively, compared with the parameter values obtained by direct numerical optimization of the posterior PDF. Note that convergence of the optimization algorithm was slow and only achieved when the initial parameter estimates were based on the stochastic simulation results.

	Mdl	K_1 10^8 N/m	K_2 10^8 N/m	K_3 10^8 N/m	$r_{u,1}$ 10^6 N	$r_{u,2}$ 10^6 N	$r_{u,3}$ 10^6 N	η_1	η_2	η_3	μ	σ m/s^2
Sim	$\mathcal{M}_1$	2.54	2.45	2.00	1.78	2.00	1.60	1.90	1.81	2.15	-	1.00
		(0.03)	(0.04)	(0.02)	(0.04)	(0.22)	(0.51)	(0.10)	(0.21)	(0.59)		(0.02)
Opt	$\mathcal{M}_1$	2.53	2.46	2.00	1.75	1.87	1.64	1.96	1.87	1.86	-	1.00
Sim	$\mathcal{M}_2$	2.58	2.47	1.98	1.80	2.04	1.79	1.79	1.69	2.18	8.04	1.00
		(0.04)	(0.05)	(0.02)	(0.04)	(0.20)	(0.81)	(0.10)	(0.19)	(0.66)	(1.68)	(0.02)
Opt	$\mathcal{M}_2$	2.57	2.48	1.99	1.80	1.97	1.47	1.81	1.71	2.03	7.53	1.00

Table 2: Sample means for posterior samples generated by updating with $\mathcal{D}_{CC}$ (standard deviations shown in parentheses) compared to values obtained by optimization of posterior PDF.

	Mdl	K_1 10^8 N/m	K_2 10^8 N/m	K_3 10^8 N/m	$r_{u,1}$ 10^6 N	$r_{u,2}$ 10^6 N	$r_{u,3}$ 10^6 N	η_1	η_2	η_3	μ	σ m/s^2
Sim	$\mathcal{M}_1$	2.24	2.77	2.07	1.75	2.51	2.45	2.22	1.00	1.32	-	1.08
		(0.02)	(0.07)	(0.02)	(0.02)	(0.23)	(0.62)	(0.10)	(0.01)	(0.21)		(0.02)
Opt	$\mathcal{M}_1$	2.23	2.79	2.08	1.75	2.47	2.72	2.20	0.98	1.20	-	1.08
Sim	$\mathcal{M}_2$	2.46	2.54	2.01	1.76	1.81	1.60	2.01	1.79	1.98	6.56	0.98
		(0.03)	(0.04)	(0.03)	(0.02)	(0.05)	(0.37)	(0.07)	(0.12)	(0.67)	(0.13)	(0.02)
Opt	$\mathcal{M}_2$	2.46	2.53	2.01	1.74	1.75	1.67	2.08	1.88	1.72	6.53	0.97

Table 3: Sample means for posterior samples generated by updating with $\mathcal{D}_{OV}$ (standard deviations shown in parentheses) compared to values obtained by optimization of posterior PDF.

For updating with the smaller-amplitude data set, $\mathcal{D}_{CC}$, estimates of inter-story stiffnesses are fairly well-constrained, and close to the actual values. However, since there is relatively little non-linear behavior, there are larger uncertainties associated with the strength and elastic-to-plastic transition parameters, and especially the breaking ductility ratio in model class $\mathcal{M}_2$, since there is very little deterioration in the actual system. Figure 5 shows the posterior samples obtained by updating model class $\mathcal{M}_2$ with data set $\mathcal{D}_{CC}$, projected on the $\{r_{u,i}, \eta_i\}$ subspace for each story. Note that the parameters for the first story, where the greatest inter-story shear forces and displacements occur, are fairly well-constrained, despite the 40% RMS noise. However, for the parameters for the second and third stories, where there is much less inelastic response, the samples are distributed over a broad region of the parameter space.

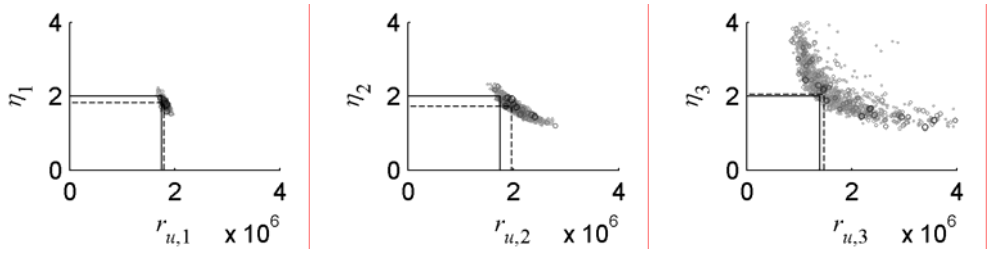

Figure 5: Samples generated by updating the deteriorating model class ,$\mathcal{M}_2$, with $\mathcal{D}_{OV}$, projected onto the $\{r_{u,i}, \eta_i\}$ sub-spaces. Dashed lines indicate results of numerical optimization of posterior and solid lines indicate the actual values used to generate the data.

For updating model class $\mathcal{M}_1$ with the large-amplitude data set, $\mathcal{D}_{OV}$, estimates of many of the parameters, including some inter-story stiffnesses, are substantially different from the actual values, because there is significant deterioration involved in the response of the structure, which cannot be included in $\mathcal{M}_1$. As expected, the identified parameter values for model class $\mathcal{M}_2$, which contains the system used to generate the data, are much closer to the actual values. However, there is still some uncertainty associated with the third-story yielding parameters, as shown in Figure 6, which shows the posterior samples obtained by updating model class $\mathcal{M}_2$ with data set $\mathcal{D}_{OV}$, projected on the $\{r_{u,i}, \eta_i\}$ sub-space for each story. Note that the samples are distributed in a more concentrated region than those obtained by updating with $\mathcal{D}_{CC}$ (shown in Figure 5).

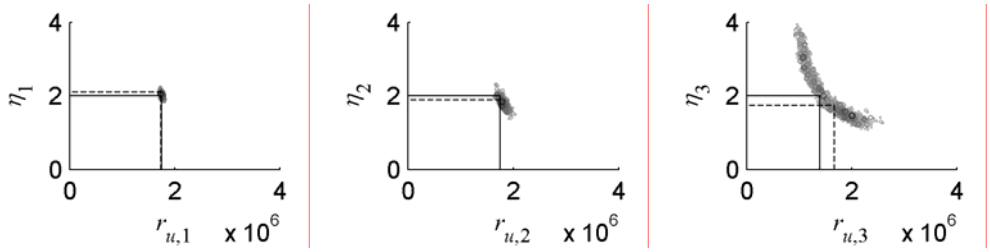

Figure 6: Samples generated by updating the deteriorating model class ,$\mathcal{M}_2$, with $\mathcal{D}_{OV}$, projected onto the $\{r_{u,i}, \eta_i\}$ sub-spaces. Dashed lines indicate results of numerical optimization of posterior and solid lines indicate the actual values used to generate the data.

Table 4 summarizes the results for Bayesian model class selection. The log-evidence and average log-likelihood for each model class are estimated using the samples generated from stochastic simulation. The information gain is then calculated from these quantities using Equation 9. For data set $\mathcal{D}_{CC}$, the deteriorating model class is preferred, but there is significant probability for the non-deteriorating model class. However, for data set $\mathcal{D}_{OV}$, model class $\mathcal{M}_2$ is overwhelmingly preferred. The difference in the accuracy of the two model classes, as expressed by the average log-likelihood, is large enough to more than compensate for the extra information extracted by model class $\mathcal{M}_2$.

Data	Model Class	Log Evidence	Log-Likelihood	Information Gain	$P(\mathcal{M} \mid \mathcal{D})$
$\mathcal{D}_{CC}$	$\mathcal{M}_1$	-2158.8	-2132.7	26.1	0.162
$\mathcal{D}_{CC}$	$\mathcal{M}_2$	-2157.2	-2130.9	26.3	0.838
$\mathcal{D}_{OV}$	$\mathcal{M}_1$	-2272.2	-2242.7	29.5	0.000
$\mathcal{D}_{OV}$	$\mathcal{M}_2$	-2137.9	-2098.9	39.0	1.000

Table 4: Bayesian model class selection result.

The difference in the information gained for the two data sets is illustrated by Figure 7, which shows normalized histograms of the posterior samples for the breaking ductility ratio μ in model class $\mathcal{M}_2$, plotted with the prior PDF. Updating with data set $\mathcal{D}_{CC}$, which features little deterioration, results in a broad posterior PDF that is fairly similar to the prior PDF. However, updating with $\mathcal{D}_{OV}$ results in a very peaked posterior PDF for μ, that indicates more information has been extracted from the data.

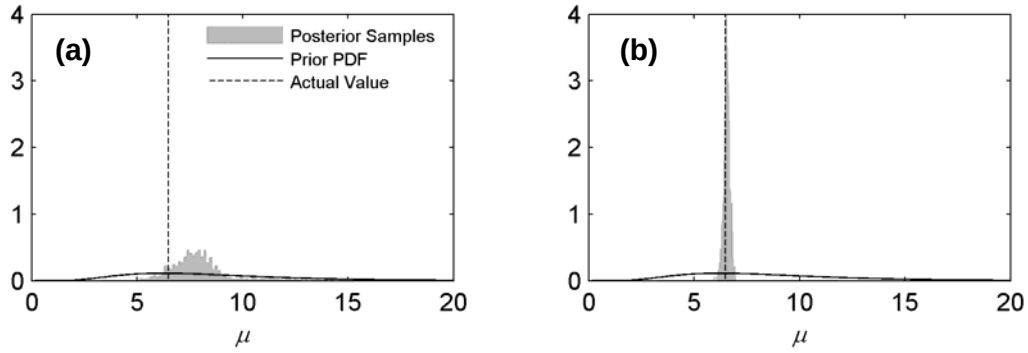

Figure 7: Prior PDFs for the breaking ductility ratio plotted against normalized histograms of posterior samples for updating model class $\mathcal{M}_2$ with (a) data set $\mathcal{D}_{CC}$ and (b) data set $\mathcal{D}_{OV}$.

7 CONCLUSIONS

Bayesian methods for model updating and model class selection can be used to study systems which are essentially unidentifiable using classical system identification approaches. Additionally, viewing the problem of model class selection in a Bayesian context allows for an information-theory interpretation of model complexity. Stochastic simulation is an effective tool for the application of Bayesian methods and in the presented example, it is used to generate samples for a posterior PDF with a complex geometry in the parameter space, using very noisy data.

REFERENCES

[1] P. Jayakumar, *Modeling and identification in structural dynamics*, Technical Report. EERL 87-01, California Institute of Technology, Pasadena, CA, 1987.

[2] P. Jayakumar, J.L. Beck, System identification using nonlinear structural models. H.G. Natke, J.T.P. Yao eds. *Structural safety evaluation based on system identification approaches*, Vieweg and Sons, 1988.

[3] A.O. Cifuentes, W.D. Iwan, Nonlinear system identification based on modeling of restoring force behavior, *Soil Dynamics and Earthquake Engineering*, **8**, 2-8, 1989.

[4] S.A. Ashrafi, A.W. Smyth, R. Betti, A parametric identification scheme for non-deteriorating and deteriorating non-linear hysteretic behavior, *Structural Control and Monitoring*, **13**, 108-131, 2005.

[5] E.T. Jaynes, *Probability theory: the logic of science*. Cambridge University Press, 2003.

[6] C.P. Robert, G. Casella, *Monte Carlo statistical methods*. Springer, 1999.

[7] J.L. Beck, Statistical system identification of structures, *5th International Conference on Structural Safety and Reliability*, San Francisco, California, 1989.

[8] J.L. Beck, L.S. Katafygiotis, Updating of a model and its uncertainties utilizing dynamic test data, *1st International Conference on Computational Stochastic Mechanics*, Corfu, Greece, 1991.

[9] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties – Bayesian statistical framework, *Journal of Engineering Mechanics*, **124**, 455-461, 1998.

[10] H. Jeffreys, *Theory of probability*. Clarendon Press, 1939.

[11] S.F. Gull, Bayesian inductive inference and maximum entropy. J. Skilling ed. *Maximum entropy and Bayesian methods*. Kluwer, 1989.

[12] J.L. Beck, K.-Y. Yuen, Model selection using response measurements: Bayesian probabilistic approach, *Journal of Engineering Mechanics*, **130**, 192-203, 2004.

[13] J. Ching, M. Muto, J.L. Beck, Bayesian linear structural model updating using Gibbs sampler with modal data, *9th International Conference on Structural Safety and Reliability*, Rome, Italy, 2005.

[14] J.L. Beck, S.-K. Au, Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation, *Journal of Engineering Mechanics*, **128**, 380-391, 2002.

[15] J. Ching, Y.-J. Chen, Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection and model averaging, *Journal of Engineering Mechanics*, in print, 2007.

[16] G. Masing, Eigenspannungen und verfestigung beim messing (Self-stretching and hardening for brass), *2nd International Congress on Applied Mechanics*, Zurich, Switzerland, 1926.

[17] W.D. Iwan, A distributed-element model for hysteresis and its steady-state dynamic response, *Journal of Applied Mechanics*, **33**, 893-900, 1966.

[18] D.-Y. Chiang, *Parsimonious modeling of inelastic systems*, Technical Report EERL 92-02, California Institute of Technology, Pasadena, CA, 1992.