Probabilistic Models and Fragility Estimates for Bridge Components and Systems

Paolo Gardoni

Armen Der Kiureghian

and

Khalid M. Mosalam

Department of Civil and Environmental Engineering
University of California, Berkeley

Probabilistic Models and Fragility Estimates for Bridge Components and Systems

Paolo Gardoni

Armen Der Kiureghian

and

Khalid M. Mosalam

Structural Engineering, Mechanics and Materials
Department of Civil and Environmental Engineering
University of California, Berkeley

ABSTRACT

A comprehensive Bayesian methodology for developing probabilistic capacity and demand models for structural components and systems is formulated. The methodology is employed to develop probabilistic models for reinforced concrete (RC) columns and multi-bent bridges. The probabilistic models are used to objectively assess the seismic fragilities of RC structural components and systems, in particular, of highway bridge systems.

The approach seeks to properly account for both aleatory and epistemic uncertainties. The probabilistic models developed are similar to deterministic capacity models or demand procedures commonly used in practice, but have additional correction terms that explicitly describe the inherent systematic and random errors. Through the use of a set of "explanatory" functions, the terms that correct the bias in the existing models are identified. These functions provide means to gain insight into the underlying behavioral phenomena and to select ground motion parameters that are most relevant to the seismic demands. Systematic assessment of a measure of model quality can be made; thus, it is possible to differentiate between alternative candidate models. The approach takes into account information gained from scientific/engineering laws, observational data from laboratory experiments or field investigations, and engineering experience and subjective judgment. Methods for assessing the model parameters on the basis of the available information are described.

The probabilistic capacity models are combined with the probabilistic demand models to construct limit-state functions that are used to construct point and interval estimates of the fragilities of structural components and systems, with special attention given to the treatment and quantification of aleatory and epistemic uncertainties. First, the probabilistic capacity models are used to estimate the fragilities of a typical bridge column in terms of maximum deformation and shear demands. Next, the probabilistic demand models are used in conjunction with the component capacity models to objectively assess the seismic fragilities of an example RC bridge bent for a given set of ground motion parameters. Finally, the analysis is extended to the fragility assessment of bridge systems. Two configurations of typical new California highway overpass bridges are considered. Fragility estimates are computed both at the component level and at the system level.

ACKNOWLEDGMENT

CONTENTS

LIST OF FIGURES

x

LIST OF TABLES

1 Introduction

1.1 OBJECTIVES AND SCOPE

Recent earthquakes have emphasized the vulnerability of lifeline systems and the need to mitigate the risk consequent to the failure of these systems. Highway transportation systems are critical lifelines and their functionality after an earthquake is of primary importance for life safety and economic recovery of a community. The assessment and prediction of damage to highway systems from an earthquake and the estimation of consequent losses provide valuable information for pre-earthquake planning and risk mitigation, and for post-earthquake response and recovery purposes.

The reliability of a highway bridge for post-earthquake service depends on its damage state. Fragility is defined as the conditional probability of attaining or exceeding a specified damage state for a given set of input variables. More specifically, in earthquake engineering, fragility is defined as the conditional probability of failure of a structural component or system for given measures of ground motion intensity, e.g., peak ground acceleration or spectral acceleration, the frequency content of the ground motion, or duration of strong ground motion. The fragility of a bridge system is an important ingredient in assessing the seismic vulnerability of a highway transportation system.

1.1.1 Review of the Past Work

In the past, there have been several studies on seismic fragility of structures. Different authors have followed different strategies and approaches. Hwang and Huo (1994), Fukushima *et al*. (1996), Kai and Fukushima (1996), Shinozuka *et al*. (2000a), and Karim and Yamazaki (2001) have used Monte Carlo simulations for specific structural models. In particular, Hwang and Huo (1994) have presented an analytical method for generating fragility curves based on numerical simulations of the dynamic behavior of specific structures. The uncertainties in the earthquake-site-structure system are quantified by considering the parameters in the system as random. In order to save computational time in the Monte Carlo simulations, Fukushima *et al*. (1996) and Kai and Fukushima (1996) have proposed a fragility analysis method where random vibration theory in the frequency domain is used to evaluate the structural response. Shinozuka *et al*. (2000a) have examined the fragility curves of a bridge by Monte Carlo simulations, where the structural response is computed by two different approaches: the first uses the time-history analysis and the second uses the capacity spectrum method according to the ATC-40 (1996). Accounting for the uncertainties in the structure and ground motion have been made by considering an independent sample of 10 "nominally identical but statistically different" bridges and 80 ground motion time histories. The comparison between the two approaches has indicated that for the state of major damage and collapse the agreement is not as good as for the state of minor damage. This can be explained by the inaccuracy of the ATC-40 procedure when nonlinear effects play a crucial role. To overcome this problem, in Chapter 6 a method proposed by Chopra and Goel (1999) is considered that improves upon the ATC-40 and FEMA-273 capacity-demand diagram methods (ATC 1996; FEMA 1997). Finally, Karim and Yamazaki (2001) have developed an analytical approach to construct fragility curves for highway bridge piers of specific bridges. The simulation method makes use of the nonlinear dynamic response of an equivalent single-degree-of-freedom system of the pier obtained by static pushover analysis.

Fragility estimates have also been developed based on expert opinion. In the ATC-13 (1985) an advisor Project Engineering Panel has developed the damage probability matrices for 78 different facility types based on consensus estimates.

Another approach, pursued by Singhal and Kiremidjian (1998), has developed fragility estimates by Bayesian analysis of observed damage data for subclasses of structural systems.

Singhal and Kiremidjian (1998) have used the Park and Ang damage index (Park and Ang, 1985) to quantify damage to a structure as a function of structural capacity and demand. The fragility is then defined as the conditional probability that the damage index exceeds a certain threshold for a given ground motion. Singhal and Kiremidjian have assumed that the randomness in the damage index at a specified ground motion level can be represented by a lognormal distribution with unknown median and known constant standard deviation. Observed damage data from past earthquakes have been used to update the distribution of the median of the lognormal distribution of the damage index by using conjugate distributions.

Other authors have developed empirical fragility curves on the basis of the records of damage resulting from past earthquakes. Basoz and Kiremidjian (1997) have developed empirical fragility curves by logistic regression based on the bridge damage observations after the Northridge earthquake. After defining 11 bridge classes based on substructure material (e.g., concrete, steel, concrete/steel, timber, masonry, etc.) and on superstructure material and type (e.g., concrete girder, steel girder, concrete truss, suspension/cable stayed, arch, etc.), empirical fragility curves have been developed for bridges grouped by these structural characteristics. Mander and Basoz (1999) have estimated fragility curves with an approach similar to the one used by Singhal and Kiremidjian (1998). Each fragility curve is assumed to be a standard lognormal cumulative distribution function with unknown location parameter (e.g., mean) and known constant scale parameter (e.g., standard deviation), meant to incorporate epistemic and aleatory uncertainty of both capacity and demand. The unknown location parameter has been assessed by using ground motion data, geological maps, and the National Bridge Inventory (NBI) records that collect bridge attributes and geographical location. Shinozuka *et al*. (2000b) have developed empirical fragility curves for columns assuming that the curves can be expressed in the form of two-parameter lognormal distribution functions. The location and scale parameters of the distribution have been estimated by maximizing the likelihood of observing the damage data from the 1995 Hyogo-ken Nanbu (Kobe) earthquake. The same authors have also developed analytical fragility curves on the basis of nonlinear dynamic analysis. In this case, the location and scale parameters have been estimated by fitting a lognormal distribution to the failure/no failure data obtained from the numerical simulations of the nonlinear dynamic response of 2 bridges in the Memphis area having random material properties. Tanaka *et al*. (2000) have estimated the seismic fragility of highway systems by assuming that the fragility can be

expressed in the form of a two-parameter normal distribution function. The authors have estimated the unknown parameters by using the damage data after the 1995 Hyogo-ken Nanbu (Kobe) earthquake. A total of 3,683 bridges were grouped into 5 structure types and the damage level was ranked into 5 levels. Similarly, Yamazaki *et al*. (2000) have developed a set of empirical fragility curves based on the actual damage from the 1995 Hyogo-ken Nanbu (Kobe) earthquake considering 216 bridge structures and assuming a lognormal distribution for the fragility curves.

Several studies have focused on reviewing approaches of fragility analysis. Casciati and Faravelli (1991) have summarized several viable approaches; some have been already outlined above and some are more original, such as methods using artificial intelligence techniques. Williams and Sexsmith (1995) have given a review of local and global seismic damage indices with emphasis on their use in decision making under uncertainties.

A common characteristic of these approaches is that the modeling and estimation is carried out at the structural system level. Because of this, the fragility models developed are not transportable. That is, the fragility estimate for a specific structural system cannot be used to assess the fragility of another structure, except as a crude approximation when the two structures are of similar type or by an arbitrary combination of fragility curves developed for example bridges within a specified category (Shinozuka *et al*., 2000a). A further disadvantage of these approaches is that the fragility models cannot take advantage of experimental test data that are normally available at the structural component, not system, level. While ideally the fragility models should be derived from first principles, e.g., the rules of mechanics, these formulations generally assume an arbitrary distribution function (either normal or lognormal) to express the fragility curve and simply estimate its distribution parameters that have no direct physical interpretation. Finally, these approaches do not properly account for all the uncertainties that are involved, particularly the uncertainty in the idealized mathematical model used to describe structural systems and their behavior.

1.1.2 Proposed New Approach

In order to implement Performance-Based Earthquake Engineering (PBEE) within a probabilistic framework, the assessment methodology should seek to explicitly account for all the prevailing uncertainties, including uncertainties in structural properties and loading characteristics, statistical uncertainty, measurement errors, modeling errors arising from inaccurate model forms or missing variables, and inaccuracies in the methodology itself.

In this report, a methodology is developed for constructing component and system fragility estimates by solving reliability problems that involve the structural capacities at the component level, $\mathbf{C} = (C_1, \ldots, C_k, \ldots, C_q)$, and the corresponding demands, $\mathbf{D} = (D_1, \ldots, D_k, \ldots, D_q)$, due to an earthquake ground motion. Component capacities are defined as the set of forces and deformations that a component can carry without failing, e.g., maximum shear force or deformation that a column can sustain. The component demands are defined as the forces and deformations to which a component is subjected to for a given system demand, e.g., an earthquake ground motion characterized by its intensity, frequency content, and duration of strong motion. The probability of failure of a component is then defined as the probability that the demand measure D_k is larger than or equal to the corresponding capacity measure C_k, where k ranges over all the possible modes of failure (e.g., failure in shear or excessive deformation). The failure of the structural system is defined in terms of component failure events, and the corresponding probability is computed by use of the methods of structural system reliability.

Predictive capacity and demand models in current structural engineering practice are typically deterministic and on the conservative side. These models have been commonly developed by using simplified mechanics rules and conservatively fitting to available experimental data. As a result, they do not explicitly account for the uncertainty inherent in the model and they provide biased estimates of the quantities of interest. While these deterministic models have been successfully used to design safe structures, the demands of modern structural engineering practice and, in particular, the advent of the performance-based design concept require predictive capacity models that are unbiased and explicitly account for all the prevailing uncertainties. The same applied to demand models.

This report presents a Bayesian framework for the development of multi-variate probabilistic capacity and demand models that properly account for all the prevailing uncertainties and correct the conservatism inherent in the deterministic models. With the aim of facilitating their use in practice, rather than developing new models, correction terms to existing deterministic capacity and demand models in common use are developed. Methods for assessing the model parameters on the basis of observed experimental data are described. Through the use of a set of "explanatory" functions, terms that correct the bias in these existing models are identified. Moreover, these functions provide means to gain insight into the underlying behavioral phenomena and to select ground motion parameters that are most relevant to the seismic demands. Although the methodology described in this report is aimed at developing probabilistic capacity and demand models, the approach is general and can be applied to the assessment of models (i.e., model selection and parameter estimation) in many engineering problems.

The probabilistic capacity models are combined with the probabilistic demand models to construct limit-state functions that are used to estimate the fragility of structural components and systems, with special attention given to the treatment of aleatory and epistemic uncertainties. Different fragility estimates are developed for reinforced concrete (RC) bridge components and systems following this methodology. First, the probabilistic capacity models are used to estimate the fragility of a typical bridge column in terms of maximum deformation and shear demands. Next, the probabilistic demand models are used in conjunction with the capacity models to objectively assess the seismic fragility of a RC bridge bent for a given set of ground motion parameters. Finally, the analysis is extended to the fragility assessment of bridge systems. Two configurations of typical new California highway overpass bridges are considered and fragility estimates are computed both at the component and system levels.

1.2   Organization of the Report

Following the general introduction given in this chapter, Chapter 2 discusses a Bayesian approach for the statistical analysis along with a philosophical justification for selecting this among three alternative approaches (frequentist, Bayesian, and likelihood). The problem of

constructing a prior distribution that properly reflects the present state of knowledge is discussed and numerically efficient simulation methods for computing the posterior statistics are presented.

In Chapter 3, a framework is formulated for constructing probabilistic capacity models for structural components that are unbiased, i.e., correct the conservatism inherent in the deterministic capacity models, and explicitly account for the most relevant uncertainties, including model errors arising from an inaccurate model form or missing variables, measurement errors, and statistical uncertainty. With the aim of facilitating their use in practice, rather than developing new capacity models, we actually develop correction terms to existing commonly used deterministic models to account for the inherent bias and uncertainty in these models. Through the use of a set of "explanatory" functions, we are able to identify terms that give rise to the bias in an existing model form and gain insight into the underlying behavioral phenomena.

In Chapter 4, the methodology presented in Chapter 3 is used to develop multi-variate probabilistic deformation and shear capacity models for RC columns under cyclic loads. An objective assessment of the relative qualities of alternative models is made.

In Chapter 5 a Bayesian framework is developed for the formulation of demand models for structural components and systems that is consistent with the one used in Chapter 3 to construct the capacity models. The deterministic "model" used here is a set of procedures used in practice to assess demands. The "explanatory" functions in this case not only identify terms that are significant in correcting the bias in the existing deterministic procedures and provide insight into the underlying behavioral phenomena, but also are used to select ground motion parameters that are most relevant to the seismic demands.

In Chapter 6, the methodology presented in Chapter 5 is used to construct multi-variate probabilistic deformation and shear demand models for RC bridge bents and bridge systems by Bayesian statistical inference. Existing observational data and data generated by simulation are used.

In Chapter 7, the fragility for structural components and systems is defined along with alternative estimates that differ according to how the parameter uncertainties are treated. By using the capacity models described in Chapters 3 and the demand models described in Chapters 5, limit-state functions are constructed with special attention devoted to the treatment of aleatory and epistemic uncertainties.

Following the theory presented in Chapter 7, in Chapter 8 fragility estimates are developed for RC bridge components and systems. First, the fragility of a typical bridge column is estimated in terms of maximum deformation and shear demands by using the probabilistic capacity models developed in Chapter 4. Next, the seismic fragility of an RC bridge bent is objectively assessed for a given set of ground motion parameters by using limit-state functions that properly account for all the relevant uncertainties. The limit-state functions are constructed by combining the probabilistic capacity models (Chapter 4) and the demand models (Chapter 6). Finally, the analysis is extended to the fragility assessment of bridge systems. Two configurations of typical new California highway overpass bridges are considered and fragility estimates are computed both at the component level and at the system level.

Chapter 9 contains the overall summary and conclusions of the report.

In Appendix A, after a brief description of invertible transformations, a convenient form of the non-informative multi-variate Jeffreys's prior density function is derived. This form is applicable for estimating unknown standard deviations and correlation coefficients instead of variances and covariances.

2 Statistical Analysis

2.1 INTRODUCTION

In the context of data analysis based on probability models, three principal approaches are possible: frequentist, Bayesian, and likelihood. The frequentist approach is based on imagining repeated sampling from a particular model (the likelihood), which defines the probability distribution of the observed data conditional on unknown parameters. The Bayesian approach requires a sampling model and, in addition, a prior distribution for the unknown parameters. The prior and the likelihood are combined to construct the posterior distribution. In particular the empirical Bayes (EB) approach allows estimating the prior distribution on the basis of observed data. Finally, the "likelihood" (or "Fisherian") approach is based on a sampling model, as is the Bayesian approach, but without a prior distribution. The inferences are based only on the likelihood function.

The philosophical framework for the approach presented in this report is based on the Bayesian notion of probability. The eventual goal of developing probabilistic capacity and demand models and fragility estimates is seen in the context of making decisions with regard to the performance-based design of new structures or the retrofit and rehabilitation of existing structures. In this context, it is essential for the approach to be capable of incorporating all types of available information, including mathematical models of structural behavior, laboratory test data, field observations, past experience, and engineering judgment. It is equally important that the approach explicitly account for all the relevant uncertainties, including those that are aleatory in nature and those that are epistemic. The Bayesian framework employed in this work is ideally suited for this purpose. In this chapter a full Bayesian approach is presented.

2.2   BAYESIAN INFERENCE

This section presents the fundamental concepts of Bayesian inference closely following Box and Tiao (1992). We start our discussion by introducing the well-known updating rule that is used in Bayesian inference. Suppose that $\mathbf{y}' = (y_1,\ldots,y_n)$ is a vector of n observations, where the superscript $(\,)'$ indicates the transpose, and that its probability density function, $p(\mathbf{y}|\boldsymbol{\theta})$, depends on the values of k unknown parameters $\boldsymbol{\theta}' = (\theta_1,\ldots,\theta_k)$ having a probability distribution $p(\boldsymbol{\theta})$. Then,

$$p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\mathbf{y},\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\mathbf{y})p(\mathbf{y}) \tag{2.1}$$

From (2.1) the conditional distribution of $\boldsymbol{\theta}$ for given observed data $\mathbf{y}$ can be written as

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} \quad \text{with} \quad p(\mathbf{y}) \neq 0 \tag{2.2}$$

And

$$p(\mathbf{y}) = E_{\boldsymbol{\theta}}\big[p(\mathbf{y}|\boldsymbol{\theta})\big] = \kappa(\mathbf{y})^{-1} = \begin{cases} \int p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} & \boldsymbol{\theta} \text{ continuous} \\ \sum p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})\Delta\boldsymbol{\theta} & \boldsymbol{\theta} \text{ discrete} \end{cases} \tag{2.3}$$

can be written where the sum or the integral is taken over the admissible range of $\boldsymbol{\theta}$, and where $E_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})]$ is the mathematical expectation of $f(\boldsymbol{\theta})$ with respect to the distribution $p(\boldsymbol{\theta})$. By using (2.3), (2.2) can be written as

$$p(\boldsymbol{\theta}|\mathbf{y}) = \kappa\, p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.4}$$

Equations (2.2) and its equivalent (2.4) are expressions of the *Bayes theorem*, where $p(\boldsymbol{\theta})$ can be viewed as the *prior* distribution reflecting our state of knowledge about $\boldsymbol{\theta}$ prior to obtaining the observations and $p(\boldsymbol{\theta}|\mathbf{y})$ is the *posterior* distribution of $\boldsymbol{\theta}$ given $\mathbf{y}$, which in a sense tells us what is known about $\boldsymbol{\theta}$ given knowledge from the data. In practice, the prior might incorporate any subjective information about $\boldsymbol{\theta}$ that is based on our engineering experience and judgment.

The quantity $\kappa = \kappa(\mathbf{y})$ is a normalizing factor necessary to ensure that the posterior distribution $p(\boldsymbol{\theta}|\mathbf{y})$ integrates or sums to one.

Following Fisher (1922), $p(\mathbf{y}|\boldsymbol{\theta})$ in (2.4) may be viewed as a function of $\boldsymbol{\theta}$, for given data $\mathbf{y}$. In that case, it is called the *likelihood function* of $\boldsymbol{\theta}$ for given $\mathbf{y}$ and is written $L(\boldsymbol{\theta}|\mathbf{y})$.

After introducing the likelihood function, we can write Bayes's formula as

$$p(\boldsymbol{\theta}|\mathbf{y}) = \kappa L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta}) \qquad (2.5)$$

This formulation of the Bayes theorem says that our prior knowledge about $\boldsymbol{\theta}$, which is expressed in the prior distribution, is updated into the posterior probability distribution by multiplying the prior distribution by the likelihood function for $\boldsymbol{\theta}$ given $\mathbf{y}$. That is

posterior distribution $\propto$ likelihood $\times$ prior distribution

According to this interpretation, the likelihood function can be seen as representing the information about $\boldsymbol{\theta}$ coming from the new data.

Furthermore, application of the rule in (2.5) can be repeated to update our present state of knowledge every time new knowledge becomes available. For example, if an initial sample of observations, $\mathbf{y}_1$, is originally available, then application of the Bayes formula gives

$$p(\boldsymbol{\theta}|\mathbf{y}_1) \propto p(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{y}_1) \qquad (2.6)$$

If a second sample of observations, $\mathbf{y}_2$, distributed independently of the first sample, becomes available, $p(\boldsymbol{\theta}|\mathbf{y}_1)$ can be updated to account for the new information such that

$$p(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2) \propto p(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{y}_1)L(\boldsymbol{\theta}|\mathbf{y}_2)$$
$$\propto p(\boldsymbol{\theta}|\mathbf{y}_1)L(\boldsymbol{\theta}|\mathbf{y}_2) \qquad (2.7)$$

Expressions (2.6) and (2.7) are applications of (2.5) where the posterior distribution in (2.6) now plays the role of the prior distribution in (2.7).

Of course, the same updating process can be repeated any number of times. For example, if we have m independent samples of observations, the posterior distribution can be updated after each new sample becomes available; that is, the likelihood associated with the q-th sample

is combined with the posterior distribution of $\boldsymbol{\theta}$ that accounts for the information content of the previous $q-1$ samples. Mathematically, we can write

$$p\big(\boldsymbol{\theta}\big|\mathbf{y}_1,\ldots,\mathbf{y}_q\big) \propto p\big(\boldsymbol{\theta}\big|\mathbf{y}_1,\ldots,\mathbf{y}_{q-1}\big)L\big(\boldsymbol{\theta}\big|\mathbf{y}_q\big) \qquad q=2,\ldots,m \qquad\qquad (2.8)$$

where $p(\boldsymbol{\theta}\,|\,\mathbf{y}_1)$ is given as in (2.6). Repeated applications of Bayes's theorem can then be seen as a learning process, where our present knowledge about the unknown parameters $\boldsymbol{\theta}$ is updated, as new data become available.

2.2.1 Prior Distribution

Historically, selection of a prior distribution that properly reflects the present state of knowledge gave rise to several disputes. The Bayesian approach was often criticized as subjective and too fragile in its dependence on a specific prior. For this reason it became essential to be able to construct prior distributions that could reflect a situation where little is known a priori. Bayes himself made the first attempt to construct such a non-informative prior. He suggested that in case of lack of previous knowledge one could use a uniform distribution. This is usually referred to as "Bayes's postulate."

However, a simple example can show a fundamental inconsistency of Bayes's postulate: suppose that the distribution of a continuous parameter θ is taken locally uniform, then the distribution of a transformation of θ, e.g., $\ln\theta$ or θ^{-1}, would not be locally uniform. Different prior distributions would lead to inconsistent posteriors even for the same data depending on the choice of parameterization.

In general, formulation of any statistical model requires some degree of subjectivity, so this inconsistency does not mean that Bayes's postulate should not be used in practice. Furthermore, the logarithmic and reciprocal transformations are sometimes nearly linear over a range of uncertainty for θ that is not large compared to the mean value; therefore, approximate uniformity of θ would imply approximate uniformity for the transformed θ.

In selecting a prior distribution note that for a sample size that is even moderately large, fairly drastic changes in the prior distribution can lead to only minor changes in the posterior distribution. That is, in this case, the information content introduced by the likelihood tends to

overwhelm the information content of the prior. An illustration of the robustness of inference, under sensible modification of the prior, is provided by the study of Mosteller and Wallace (1964).

These considerations indicate that arbitrariness in the choice of the transformation in terms of which the prior is locally uniform is often acceptable. In light of this consideration, the next section of this report explores whether there is a way to better approximate the situation in which "little is known a priori."

Non-informative Prior Distributions

As remarked by Box and Tiao (1992), we can never be in a complete state of ignorance and the statement "knowing little a priori" has only a relative meaning compared to the information content of an experiment. In order to explore the possibility of providing a prior that has little information relative to an intended experiment, we start by considering an example for a single parameter.

The Normal Mean θ (σ^2 Known)

Let $\mathbf{y}' = (y_1, \ldots, y_n)$ be a random sample from a normal distribution $N(\theta, \sigma^2)$, where σ is known. The likelihood function for the mean θ for given $\mathbf{y}$ can be written as

$$L(\theta \mid \sigma, \mathbf{y}) \propto \exp\left[-\frac{1}{2\sigma^2}\sum(y_i - \theta)^2\right], \qquad -\infty < \theta < \infty \qquad (2.9)$$

Let $\bar{y}$ be the sample mean. Then

$$\sum(y_i - \theta)^2 = \sum(y_i - \bar{y})^2 + n(\theta - \bar{y})^2 \qquad (2.10)$$

and, since given the data $\sum(y_i - \bar{y})^2$ is a fixed constant, the likelihood reduces to

$$L(\theta \mid \sigma, \mathbf{y}) \propto \exp\left[-\frac{n}{2\sigma^2}(\theta - \bar{y})^2\right] \qquad (2.11)$$

The likelihood function of θ, standardized such that it integrates out to one, is graphically represented by a normal curve centered at $\bar{y}$, with standard deviation $\sigma/\sqrt{n}$. Figure 2.1.a shows

the standardized likelihood curves for three hypothetical experiments with sample means $\bar{y} = 6$, $\bar{y} = 9$, and $\bar{y} = 12$, $n = 10$, and $\sigma = 1$.

In case the quantity of immediate interest is not the mean θ itself, but, for example, its reciprocal $\xi = \theta^{-1}$, the likelihood is written as

$$L(\xi \mid \sigma, \mathbf{y}) \propto \exp\left[-\frac{n}{2\sigma^2} \left(\xi^{-1} - \bar{y} \right)^2 \right] \tag{2.12}$$

and the standardized likelihood curves appear as in Figure 2.1.b. Figure 2.1 also shows that a prior for the mean that is locally uniform in θ is not uniform in ξ.

The information content of the data is brought in by the likelihood function, and from (2.11) we see that the data enter only via $\bar{y}$. Figure 2.1.a shows that when the likelihood is expressed in terms of θ, the sample average $\bar{y}$ affects only the location of the likelihood curve but not its shape. That is, different $\bar{y}$ only translate the likelihood curve on the θ axis but leave it otherwise unchanged. Figure 2.1.b shows that, on the contrary, the likelihood in (2.12) is affected both for location and for spread by a change in $\bar{y}$.
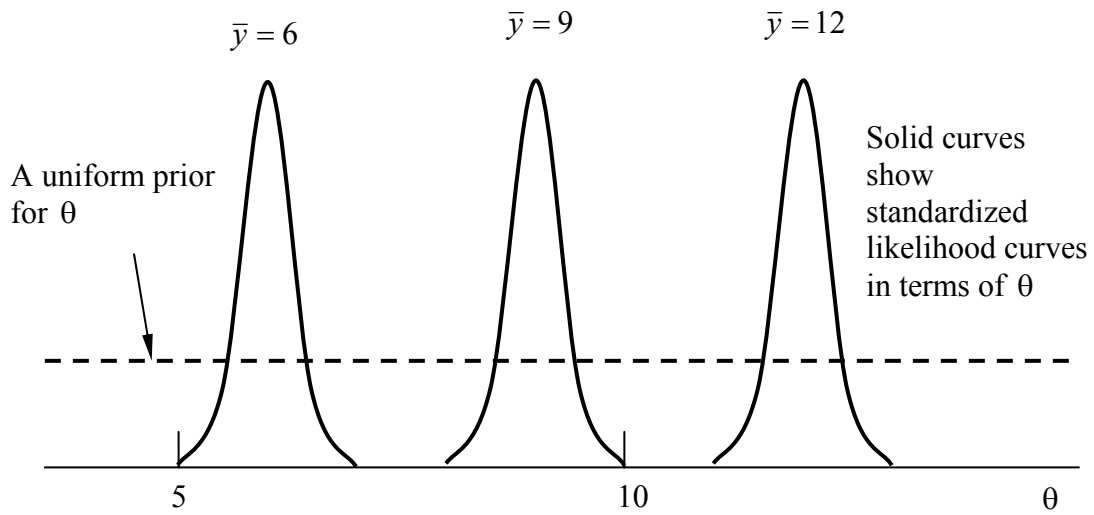
(a) The normal mean θ.
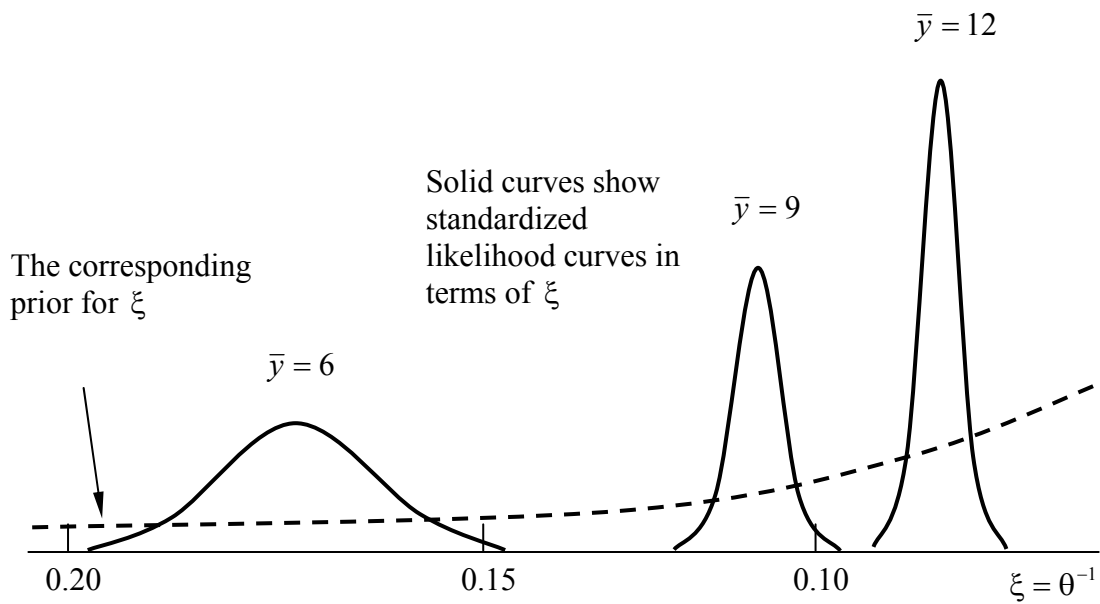
(b) Reciprocal of the normal mean $\xi = \theta^{-1}$

Figure 2.1. **Non-informative prior distributions and standardized likelihood curves: (a) for the normal mean θ, and (b) for $\xi = \theta^{-1}$.**

In light of the above observation, a hypothetical scale $\phi(\theta)$ is introduced such that the corresponding likelihood is data translated. That is, such that the shape of the likelihood curve does not depend on the data but only its location, as it was the case for the normal mean. In this

15

new scale knowing little a priori, relative to the information content of the data, is translated in our indifference for the value of $\phi(\theta)$. That is, a non-informative prior about $\phi(\theta)$ with respect to the data is simply a locally uniform distribution.

Going back to our example, the likelihood of θ is completely known a priori except for its location that is determined by $\bar{y}$. So in this case, the likelihood are data translated in the original scale θ. That is, $\phi(\theta) = \theta$ and a non-informative prior is locally uniform in θ itself. That is, locally,

$$p(\theta|\sigma) = \text{costant} \qquad (2.13)$$

Figure 2.1 shows this non-informative prior distribution with a dotted line. On the other hand, in case of ξ, where $\phi(\theta) = \theta^{-1}$, we can write that

$$p(\xi|\sigma) = p(\theta|\sigma)\left|\frac{d\theta}{d\xi}\right| = p(\theta|\sigma)\theta^2 \propto \xi^{-2} \qquad (2.14)$$

So the corresponding non-informative prior for ξ is not uniform but is locally proportional to θ^2, that is, to ξ^{-2}. From (2.14) we can draw a more general conclusion: if the non-informative prior for θ is locally uniform in $\phi(\theta)$, then the corresponding non-informative prior for θ is locally proportional to $|d\phi/d\theta|$, assuming the transformation is one to one.

In the new scale $\phi(\theta)$ we need to assume a uniform distribution only over the region of interest but not over the entire range of definition of θ. The dashed lines in Figure 2.2 indicate the proper distributions $p(\theta|\sigma)$, flat only over the region of interest, and $p(\xi|\sigma)$, obtained by a transformation that is proportional to ξ^{-2} only over the region of interest. However, it would be inappropriate mathematically and not necessary in practice to suppose, for example, that $p(\theta|\sigma)$ was uniform over an infinite range, or that $p(\xi|\sigma)$ was proportional to ξ^{-2} over an infinite range.

$\bar{y} = 6$ $\qquad \bar{y} = 9$ $\qquad \bar{y} = 12$

A proper prior distribution for θ

5 $\qquad$ 10 $\qquad \theta$

(a) The normal mean θ

$\bar{y} = 12$

The corresponding proper prior distribution for ξ

$\bar{y} = 9$

$\bar{y} = 6$

0.20 $\qquad$ 0.15 $\qquad$ 0.10 $\qquad \xi = \theta^{-1}$
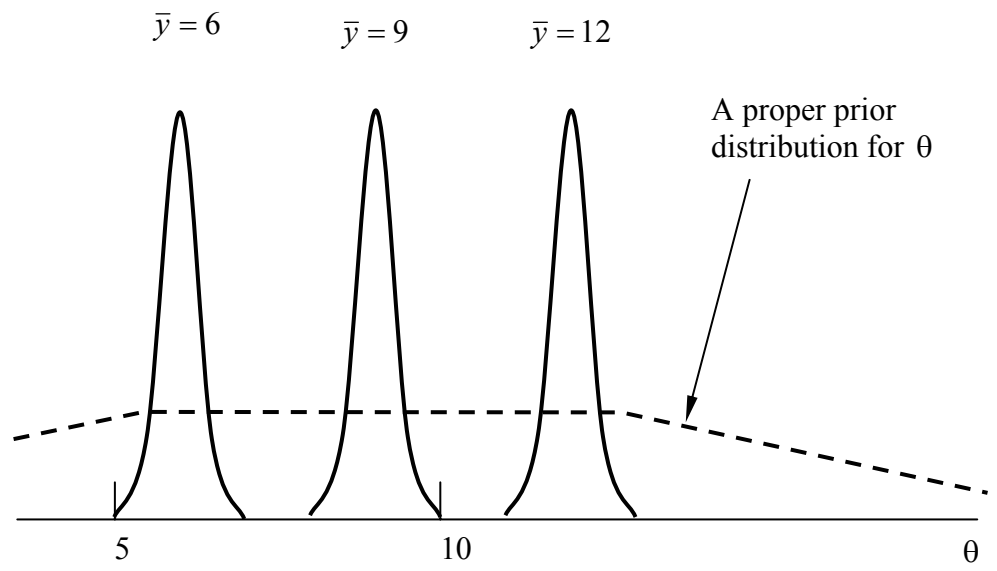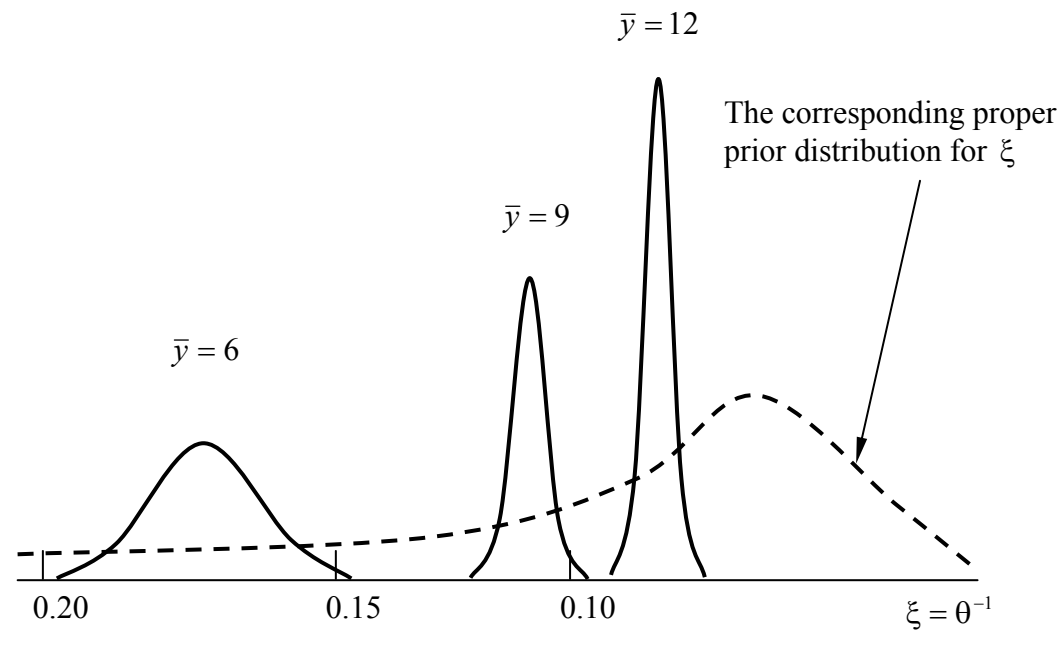
(b) Reciprocal of the normal mean $\xi = \theta^{-1}$

Figure 2.2. **Non-informative prior distributions (dashed lines) and standardized likelihood curves (solid lines) for normal distribution with known standard deviation.**

The Normal Standard Deviation σ (θ Known)

To gain more insight into the problem of constructing non-informative priors, we consider a second case by choosing a non-informative prior distribution for the standard deviation, σ, of a normal distribution that has known mean θ. In this case, the likelihood can be written as

$$L(\sigma|\theta, \mathbf{y}) \propto \sigma^{-n} \exp\left(-\frac{ns^2}{2\sigma^2}\right)$$ (2.15)

where

$$s^2 = \sum(y_i - \theta)^2 / n$$ (2.16)

Figure 2.3.a shows the standardized likelihood curves for σ with $s = 5$, $s = 10$, and $s = 20$ in case of $n = 10$ observations. It is evident that in the original scale, σ, the likelihood curves are not data translated, i.e., different values of s change both the location and the shape of the likelihood curves. So a locally uniform distribution in σ would not be non-informative.
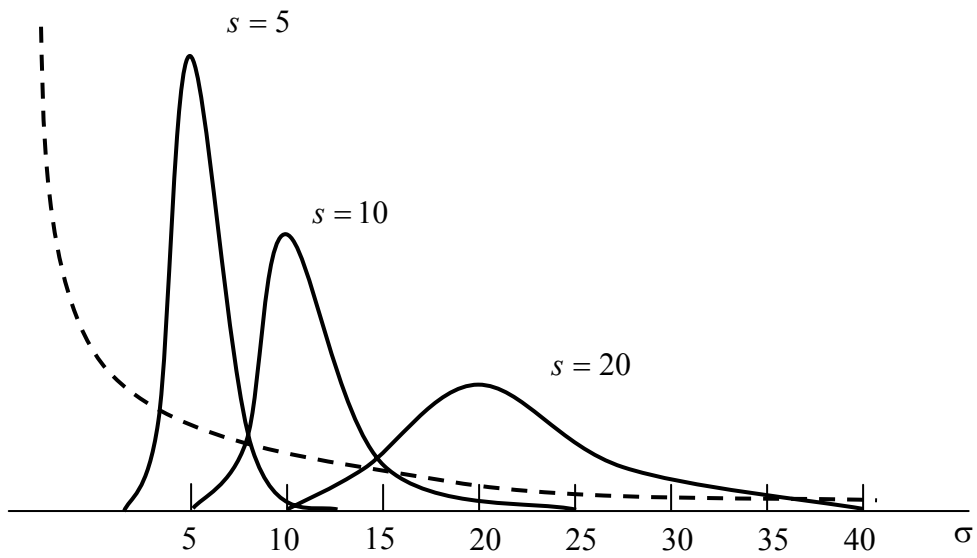
(a) Normal standard deviation σ
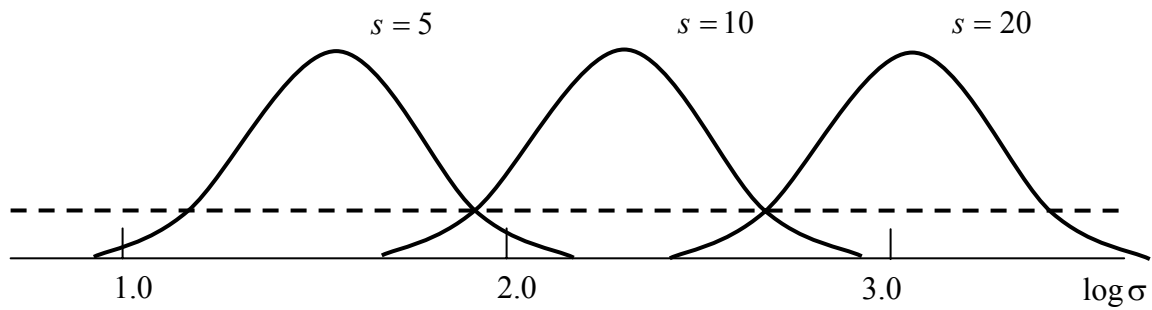
(b) Log of normal standard deviation, $\ln \sigma$

Figure 2.3. **Non-informative prior distributions (dashed lines) and standardized likelihood curves (solid lines) for normal distribution with known mean.**

However, the logarithmic transformation $\ln\sigma$ makes the likelihood curves exactly data translated as shown in Figure 2.3.b. To mathematically verify this, we note that multiplication by the constant s^n leaves the likelihood in (2.15) unchanged. Taking the logarithm we have

$$\ln\left\{\exp\left[s^n\sigma^{-n}\exp\left(-\frac{ns^2}{2\sigma^2}\right)\right]\right\} = \exp\left\{\ln\left[s^n\sigma^{-n}\exp\left(-\frac{ns^2}{2\sigma^2}\right)\right]\right\} \qquad (2.17)$$

Therefore, the likelihood of $\ln\sigma$ can be expressed as

$$L(\ln\sigma|\theta,\mathbf{y}) \propto \exp\left\{-n(\ln\sigma - \ln s) - \frac{n}{2}\exp[-2(\ln\sigma - \ln s)]\right\} \qquad (2.18)$$

In this new scale the data entering through s simply relocate the likelihood. We then conclude that a non-informative prior should be locally uniform in $\ln\sigma$. Going back to the original scale σ, the non-informative prior is thus locally proportional to σ^{-1},

$$p(\sigma|\theta) \propto \left|\frac{d\ln\sigma}{d\sigma}\right| = \sigma^{-1} \qquad (2.19)$$

Exact Data-Translated Likelihoods and Non-informative Priors

A general rule obtained from the above discussion is that given a one-to-one transformation of θ, $\phi(\theta)$, a prior distribution of θ that is locally proportional to $|d\phi/d\theta|$ is non-informative for the parameter θ if, in terms of ϕ, the likelihood curve is (data) translated, that is, the location but not the shape of the likelihood $L(\phi|\mathbf{y})$ changes with the data. Mathematically, a data-translated likelihood can be written as

$$L(\theta|\mathbf{y}) = g[\phi(\theta) - f(\mathbf{y})] \qquad (2.20)$$

where $g(\ )$ is a known function independent of the data $\mathbf{y}$ and $f(\mathbf{y})$ is a function of $\mathbf{y}$.

Note that the examples in the previous sections are in fact special cases of (2.20). For the normal mean, $\phi(\theta) = \theta$, $f(\mathbf{y}) = \bar{y}$, and for the normal standard deviation $\phi(\sigma) = \ln\sigma$, $f(\mathbf{y}) = \ln s$.

Derivation of Transformations Yielding Approximate Data-Translated Likelihoods

In general a transformation $\phi(\theta)$ that allows the likelihood to be expressed exactly in the form (2.20) may not be available. In these cases, the property of the posterior distribution of being insensitive to minor changes in the prior can be used even for a moderate-size sample to construct a transformation $\phi(\theta)$ in terms of which the likelihood is nearly independent of the data $\mathbf{y}$ except for its location.

Suppose $\mathbf{y}' = (y_1, \ldots, y_n)$ is a random sample from a distribution $p(y \mid \theta)$. According to Johnson (1967, 1970), under certain regularity conditions on $p(y \mid \theta)$, for sufficiently large n, the likelihood function of θ is approximately normal, and remains approximately normal under approximate one-to-one transformations of θ. The logarithm of the likelihood is then approximately quadratic:

$$l(\theta \mid \mathbf{y}) = \ln L(\theta \mid \mathbf{y}) = \ln \prod_i^n p(y_i \mid \theta)$$
$$\approx l(\hat{\theta} \mid \mathbf{y}) - \frac{n}{2}(\theta - \hat{\theta})^2 \left(-\frac{1}{n} \frac{\partial^2 l}{\partial \theta^2} \right)_{\hat{\theta}} \tag{2.21}$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ, defined as the point where the likelihood function attains its maximum value, so that

$$\left(\frac{\partial l}{\partial \theta} \right)_{\hat{\theta}} = 0 \tag{2.22}$$

In general, the quantity

$$J(\hat{\theta}) = \left(-\frac{1}{n} \frac{\partial^2 l}{\partial \theta^2} \right)_{\hat{\theta}} \tag{2.23}$$

is a positive function of $\mathbf{y}$ and assume for the moment that for given n, it can be expressed as a function of $\hat{\theta}$ only. On the other hand, the logarithm of a normal density function $p(y)$ is of the form

$$\ln p(y) \propto -\frac{1}{2}(y-\mu)^2 / \sigma^2 \qquad (2.24)$$

and, given the location parameter μ, it is completely determined by its standard deviation σ. A comparison of (2.21) and (2.24) shows that the standard deviation of the likelihood curve is approximately equal to $n^{-1/2} J^{-1/2}(\hat{\theta})$. If $\phi(\theta)$ is a one-to-one transformation we can write

$$J(\hat{\phi}) = \left(-\frac{1}{n}\frac{\partial^2 l}{\partial \phi^2}\right)_{\hat{\phi}} = \left(-\frac{1}{n}\frac{\partial^2 l}{\partial \theta^2}\right)_{\hat{\theta}}\left(\frac{d\theta}{d\phi}\right)_{\hat{\theta}}^2 = J(\hat{\theta})\left(\frac{d\theta}{d\phi}\right)_{\hat{\theta}}^2 \qquad (2.25)$$

and so if $\phi(\hat{\theta})$ is chosen such that

$$\left|\frac{d\theta}{d\phi}\right|_{\hat{\theta}} \propto J^{-1/2}(\hat{\theta}) \qquad (2.26)$$

$J(\hat{\phi})$ will be a constant independent of $\hat{\phi}$, and the likelihood will be approximately data translated in terms of ϕ. This means that a scale for which a locally uniform prior is approximately non-informative is obtained as

$$\frac{d\phi}{d\theta} \propto J^{1/2}(\theta) \quad \text{or} \quad \phi \propto \int^{\theta} J^{1/2}(t)\, dt \qquad (2.27)$$

where the integral is the improper integral evaluated at θ and that the corresponding non-informative prior for θ is

$$p(\theta) \propto \left|\frac{d\phi}{d\theta}\right| \propto J^{1/2}(\theta) \qquad (2.28)$$

In the above we have assumed that the quantity in (2.22) is a function of $\hat{\theta}$ only. More generally (Box and Tiao, 1992), this holds if the observations $\mathbf{y}$ are drawn from a distribution of the form

$$p(y|\theta) = h(y)w(\theta)\exp[c(\theta)u(y)] \qquad (2.29)$$

This is the form for which a single sufficient statistic for θ exists, where a sufficient statistic is defined according to Box and Tiao (1992) as follows:

22

Definition (2.1) Let $\mathbf{y}' = (y_1, \ldots, y_n)$ be a vector of observations whose distribution depends upon the k parameters $\mathbf{\theta}' = (\theta_1, \ldots, \theta_k)$. Let $\mathbf{t}' = (t_1, \ldots, t_q)$ be q functions of $\mathbf{y}$. Then the set of statistics $\mathbf{t}$ is said to be jointly sufficient for $\mathbf{\theta}$ if the likelihood function $L(\mathbf{\theta}|\mathbf{y})$ can be expressed in the form

$$L(\mathbf{\theta}|\mathbf{y}) \propto g(\mathbf{\theta}|\mathbf{t}) \tag{2.30}$$

and provided the ranges of $\mathbf{\theta}$, if dependent on the observations, can also be expressed as functions of $\mathbf{t}$.

Jeffreys's Rule, Information Measure, and Non-informative Priors

More generally, $p(y|\theta)$ is not of the form in (2.29) and the quantity in (2.22) is a function of all the data $\mathbf{y}$. In this case, we start by noting that, for given θ,

$$-\frac{1}{n}\frac{\partial^2 l}{\partial \theta^2} = -\frac{1}{n}\sum_{u=1}^{n} \frac{\partial^2 \ln p(y_u|\theta)}{\partial \theta^2} \tag{2.31}$$

is the average of n identical functions of $(y_1, \ldots, y_n)$, respectively. We also suppose that θ_o is the true value of θ so that $\mathbf{y}$ are drawn from $p(y|\theta_o)$. It follows that as $n \to \infty$ the average in (2.31) converges in probability to the expectation of the function, that is, to

$$\mathrm{E}_{y|\theta_o}\left[-\frac{\partial^2 \ln p(y|\theta)}{\partial \theta^2}\right] = -\int \frac{\partial^2 \ln p(y|\theta)}{\partial \theta^2} p(y|\theta_o) dy = a(\theta, \theta_o) \tag{2.32}$$

assuming that the expectation exists. Furthermore, as $n \to \infty$, the maximum likelihood estimate $\hat{\theta}$ converges in probability to θ_o. We can then write, approximately,

$$\left(-\frac{1}{n}\frac{\partial^2 l}{\partial \theta^2}\right)_{\hat{\theta}} \approx a(\hat{\theta}, \theta_o) \approx a(\hat{\theta}, \hat{\theta}) = J(\hat{\theta}) \tag{2.33}$$

where $J(\theta) = a(\theta, \theta)$ is defined as

$$J(\theta) = E_{y|\theta}\left[-\frac{\partial^2 \ln p(y|\theta)}{\partial \theta^2}\right] \qquad (2.34)$$

Now, using $J(\hat{\theta})$, which depends only on $\hat{\theta}$, to approximate (2.22) and arguing as before, we can find that the scale $\phi(\theta)$ for which a locally uniform prior is approximately non-informative is such that

$$\frac{d\phi}{d\theta} \propto J^{1/2}(\theta) \quad or \quad \phi \propto \int^{\theta} J^{1/2}(t)\, dt \qquad (2.35)$$

and that the corresponding non-informative prior for θ is

$$p(\theta) \propto J^{1/2}(\theta) \qquad (2.36)$$

It can be easily shown that, when $p(y|\theta)$ is of the form in (2.29), $J(\hat{\theta}) \equiv J(\hat{\theta})$. Thus the prior in (2.36) can be used generally.

The quantity $J(\theta)$ in (2.34) is known as Fisher's measure of information about θ in a single observation y (Fisher, 1922, 1925). More generally, Fisher's measure of information about θ in a sample $\mathbf{y}' = (y_1,\ldots,y_n)$ is defined as

$$J_n(\theta) = E_{y|\theta}\left[-\frac{\partial^2 l}{\partial \theta^2}\right] \qquad (2.37)$$

where the expectation is taken with respect to the distribution $p(\mathbf{y}|\theta)$. When $\mathbf{y}$ is a random sample we have $J_n(\theta) = nJ(\theta)$; thus, (2.36) can be expressed by *Jeffreys's rule* that says that the prior distribution for a single parameter θ is approximately non-informative if it is taken to be proportional to the square root of Fisher's information measure.

Non-informative Priors for Multiple Parameters

This section describes how to construct non-informative priors for multi-parameter models that are used in the applications presented in the following chapters. Consider a general multi-variate model with an observable vector $\mathbf{y}$ of q components, having the multi-normal distribution

$$\mathbf{y}|\boldsymbol{\mu}(\boldsymbol{\theta}),\boldsymbol{\Sigma} \sim \mathrm{N}[\boldsymbol{\mu}(\boldsymbol{\theta}),\boldsymbol{\Sigma}] \qquad (2.38)$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}) = [\mu_1(\boldsymbol{\theta}_1), \mu_2(\boldsymbol{\theta}_2), \ldots, \mu_q(\boldsymbol{\theta}_q)]$ is a vector-valued function of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_q)$, where $\boldsymbol{\theta}_k = (\theta_{ki}, i = 1, \ldots, p_k)$, $k = 1, \ldots, q$, and $\boldsymbol{\Sigma}$ is a $q \times q$ variance-covariance matrix, which is symmetric and positive definite. For constructing the prior distribution of the parameters $\boldsymbol{\Theta} = (\boldsymbol{\theta}, \boldsymbol{\Sigma})$, we assume that $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are approximately independent, so that

$$p(\boldsymbol{\Theta}) \approx p(\boldsymbol{\theta})p(\boldsymbol{\Sigma}) \qquad (2.39)$$

We also assume that the parameterization in terms of $\boldsymbol{\theta}$ is such that it is appropriate to take $\boldsymbol{\theta}$ as locally uniform (over the region $\mathbf{I}_{\boldsymbol{\theta}}$),

$$p(\boldsymbol{\theta}) = \text{constant}, \qquad \boldsymbol{\theta} \in \mathbf{I}_{\boldsymbol{\theta}} \qquad (2.40)$$

For the prior distribution of the $q(q+1)/2$ distinct elements of $\boldsymbol{\Sigma}$, application of similar arguments as in the previous section (Gelman *et al*., 1998, and Box and Tiao, 1992) lead to the non-informative multi-variate Jeffreys's prior density

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(q+1)/2} \qquad (2.41)$$

or equivalently (See Appendix A) to

$$p(\boldsymbol{\Sigma}) \propto |\mathbf{R}|^{-(q+1)/2} \prod_{i=1}^{q} \frac{1}{\sigma_i} \qquad (2.42)$$

where σ_i^2 denote the variances, $\mathbf{R} = [\rho_{ij}]$ denotes the $q \times q$ correlation matrix and $|\cdot|$ denotes the determinant.

The above results are valid only for multi-parameter models with a multi-normal observable vector. In case of other distributions, approximately non-informative priors can be derived extending Jeffreys's rule to general multi-parameter models (Box and Tiao, 1992).

2.2.2 Likelihood Function

As mentioned earlier, the likelihood is a function that is proportional to the conditional probability of the observations for given values of the model parameters. Formulation of the likelihood function depends on the type and form of the available information. We now consider the problem of formulating the likelihood function for a set of n q-variate observations with no censored data. We assume that for given $\boldsymbol{\theta} = (\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_q)$ and covariance matrix $\boldsymbol{\Sigma}$ the error vector

$$
\mathbf{e}_i = \begin{bmatrix} \sigma_1 \varepsilon_{1i} \\ \vdots \\ \sigma_k \varepsilon_{ki} \\ \vdots \\ \sigma_q \varepsilon_{qi} \end{bmatrix} = \begin{bmatrix} y_{1i} - \mu_{1i}(\boldsymbol{\theta}_1) \\ \vdots \\ y_{ki} - \mu_{ki}(\boldsymbol{\theta}_k) \\ \vdots \\ y_{qi} - \mu_{qi}(\boldsymbol{\theta}_q) \end{bmatrix}, \qquad i = 1,\ldots,n \tag{2.43}
$$

is distributed as the q-variate normal $N_q(\mathbf{0},\boldsymbol{\Sigma})$, and that the n q-variate observations are independent.

To construct the likelihood function, we start from the joint distribution of the n vectors of errors $\mathbf{e}' = (\mathbf{e}_1,\ldots,\mathbf{e}_i,\ldots,\mathbf{e}_n)$ with $\mathbf{e}_i' = (\mathbf{e}_{i1},\ldots,\mathbf{e}_{ik},\ldots,\mathbf{e}_{iq})$

$$
p(\mathbf{e}|\boldsymbol{\theta},\boldsymbol{\Sigma}) = \prod_{i=1}^{n} p(\mathbf{e}_i|\boldsymbol{\theta},\boldsymbol{\Sigma}) \qquad -\infty < e_{ki} < \infty, \; k=1,\ldots,q, \; i=1,\ldots,n
$$

$$
= (2\pi)^{-qn/2} |\boldsymbol{\Sigma}|^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \mathbf{e}_i'\boldsymbol{\Sigma}^{-1}\mathbf{e}_i\right) \tag{2.44}
$$

Introducing the quantity $\mathbf{S}(\boldsymbol{\theta})$ defined as the $q \times q$ symmetric matrix

$$
\mathbf{S}(\boldsymbol{\theta}) = \left[S_{kl}(\boldsymbol{\theta}_k,\boldsymbol{\theta}_l) \right] \tag{2.45}
$$

with

$$S_{kl}(\boldsymbol{\theta}_k, \boldsymbol{\theta}_l) = \sum_{i=1}^{n} e_{ki} e_{li} \tag{2.46}$$

and using (2.44), the likelihood can be written as

$$L(\boldsymbol{\theta}, \boldsymbol{\Sigma} | \mathbf{y}) \propto p(\mathbf{e} | \boldsymbol{\theta}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-n/2} \exp\left(-\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} \mathbf{S}(\boldsymbol{\theta}) \right) \tag{2.47}$$

2.2.3 Posterior Distribution

Combining the likelihood function in (2.47) and the prior distribution in (2.41) according to the Bayes updating rule, the posterior distribution for the parameters $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ of the multi-variate normal model can be written as

$$p(\boldsymbol{\theta}, \boldsymbol{\Sigma} | \mathbf{y}) \propto |\boldsymbol{\Sigma}|^{-(n+q+1)/2} \exp\left(-\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} \mathbf{S}(\boldsymbol{\theta}) \right) \tag{2.48}$$

where each parameter in $\boldsymbol{\theta}$ can vary from $-\infty$ to ∞ and $\boldsymbol{\Sigma}$ is positive definite. In particular, it can be shown (Box and Tiao, 1992) that for $n \geq q$, the marginal posterior distribution of $\boldsymbol{\theta}$ is given by

$$p(\boldsymbol{\theta} | \mathbf{y}) \propto |\mathbf{S}(\boldsymbol{\theta})|^{-(n)/2} \tag{2.49}$$

This simple expression is valid even when the expectation functions $\boldsymbol{\mu}(\boldsymbol{\theta}) = \left[\mu_1(\boldsymbol{\theta}_1), \mu_2(\boldsymbol{\theta}_2), \ldots, \mu_q(\boldsymbol{\theta}_q) \right]$ are not linear in the parameters.

For a general uni-variate model under the assumption of linearity of the expectation function in the unknown parameters $\boldsymbol{\theta}$, the model can be written as

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta} + \sigma\boldsymbol{\varepsilon} \tag{2.50}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations, $\mathbf{H}$ is a $n \times k$ matrix of known regressors, $\boldsymbol{\theta}$ denotes the set of unknown model parameters, $\boldsymbol{\varepsilon}$ is a $n \times 1$ vector of independent random variables having the normal distribution with zero mean and unit variance, and σ represents the standard deviation of the model errors. Expanding the matrices, (2.50) can be written as

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{k1} \\ \vdots & & \vdots \\ h_{1i} & \cdots & h_{ki} \\ \vdots & & \vdots \\ h_{1n} & \cdots & h_{kn} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_k \end{bmatrix} + \sigma \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.51}
$$

As shown by Box and Tiao (1992), for $n > k$, the posterior distribution of the unknown model parameters $(\boldsymbol{\theta}, \sigma^2)$ can be written as

$$
p(\boldsymbol{\theta}, \sigma^2 | \mathbf{y}) \propto p(\boldsymbol{\theta}, \sigma^2) p(s^2 | \sigma^2) p(\hat{\boldsymbol{\theta}} | \boldsymbol{\theta}, \sigma^2) \tag{2.52}
$$

where

$$
\hat{\boldsymbol{\theta}} = (\mathbf{H}'\mathbf{H})^{-1} \mathbf{H}'\mathbf{D}
$$

$$
s^2 = \frac{1}{\eta}(\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}})
$$

$$
\eta = n - k \tag{2.53}
$$

$$
\hat{\mathbf{y}} = \mathbf{H}\hat{\boldsymbol{\theta}}
$$

Box and Tiao (1992) show that, assuming a non-informative prior with $\boldsymbol{\theta}$ and $\ln(\sigma)$ approximately independent and locally uniform, i.e.,

$$
p(\boldsymbol{\theta}, \sigma^2) = p(\boldsymbol{\theta}) p(\sigma^2) \propto \sigma^{-2} \tag{2.54}
$$

one can rewrite the joint posterior distribution in (2.52) as

$$
p(\boldsymbol{\theta}, \sigma^2 | \mathbf{y}) \propto p(\sigma^2 | s^2) p(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}, \sigma^2) \tag{2.55}
$$

Furthermore, under the normality assumption on $\boldsymbol{\varepsilon}$, the marginal posterior distribution of σ^2 is $\eta s^2 \chi_\eta^{-2}$ and the marginal posterior distribution of $\boldsymbol{\theta}$ is

28

$$p(\mathbf{\theta}|\mathbf{y}) = \frac{\Gamma\left(\frac{\eta+k}{2}\right)|\mathbf{H}'\mathbf{H}|^{1/2} s^{-k}}{\left[\Gamma\left(\frac{1}{2}\right)\right]^k \Gamma\left(\frac{\eta}{2}\right)(\sqrt{\eta})^k}\left[1 + \frac{(\mathbf{\theta}-\hat{\mathbf{\theta}})'\mathbf{H}'\mathbf{H}(\mathbf{\theta}-\hat{\mathbf{\theta}})}{\eta s^2}\right]^{-(\eta+k)/2} \quad \begin{array}{l} -\infty < \theta_i < \infty \\ i = 1,\ldots,k \end{array} \quad (2.56)$$

which is the multi-variate t distribution, $t_k[\hat{\mathbf{\theta}}, s^2(\mathbf{H}'\mathbf{H})^{-1}, \eta]$. We note that $\hat{\mathbf{\theta}}$ is the mode and the mean of $\mathbf{\theta}$ and its covariance matrix is $\eta s^2(\mathbf{H}'\mathbf{H})^{-1}/(\eta-2)$, and the mean and variance of σ^2 are $\eta s^2/(\eta-2)$ and $2\eta^2 s^4/[(\eta-2)^2(\eta-4)]$, respectively.

A relevant property of the multi-variate t distribution that is used in this study is that the marginal distribution of an r-dimensional subset, $\mathbf{\theta}_1$, has the multi-variate t distribution, $t_r[\hat{\mathbf{\theta}}_1, s^2\mathbf{C}_{11}, \eta]$, that is,

$$p(\mathbf{\theta}_1|\mathbf{y}) = \frac{\Gamma\left(\frac{\eta+r}{2}\right)|\mathbf{C}_{11}^{-1}|^{1/2} s^{-r}}{\left[\Gamma\left(\frac{1}{2}\right)\right]^r \Gamma\left(\frac{\eta}{2}\right)(\sqrt{\eta})^r}\left[1 + \frac{(\mathbf{\theta}_1-\hat{\mathbf{\theta}}_1)'\mathbf{C}_{11}^{-1}(\mathbf{\theta}_1-\hat{\mathbf{\theta}}_1)}{\eta s^2}\right]^{-(\eta+r)/2} \quad \begin{array}{l} -\infty < \theta_{1i} < \infty \\ i = 1,\ldots,r \end{array} \quad (2.57)$$

where

$$\mathbf{\theta} = \begin{bmatrix} \mathbf{\theta}_1 \\ \mathbf{\theta}_2 \end{bmatrix}\begin{array}{l} r \\ k-r \end{array} \qquad \hat{\mathbf{\theta}} = \begin{bmatrix} \hat{\mathbf{\theta}}_1 \\ \hat{\mathbf{\theta}}_2 \end{bmatrix}\begin{array}{l} r \\ k-r \end{array}$$

$$(2.58)$$

$$(\mathbf{H}'\mathbf{H})^{-1} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}\begin{array}{l} r \\ k-r \end{array}$$
$$\phantom{(\mathbf{H}'\mathbf{H})^{-1} = \begin{bmatrix}}\begin{array}{cc} r & k-r \end{array}$$

In particular θ_{1i} has the distribution $t(\hat{\theta}_{1i}, s^2 c_{ii}, \eta)$, that is

$$t = \frac{\theta_{1i} - \hat{\theta}_{1i}}{s\sqrt{c_{ii}}} \qquad\qquad (2.59)$$

has the t distribution with $\eta = n - k$ degrees of freedom.

2.2.4 Computation of Posterior Statistics Using Importance Sampling

There are cases for which the closed form solutions presented in the previous sections are not valid. For example, if censored data are present, then the error vector is not distributed as the q-variate normal $N_q(\mathbf{0}, \mathbf{\Sigma})$ as we assumed before. In these cases, once the posterior distribution of $\mathbf{\Theta}$ is derived, one can still compute its mean vector $\mathbf{M_\Theta}$ and covariance matrix $\mathbf{\Sigma_{\Theta\Theta}}$. However, computation of these quantities, assuming they exist, as well as the normalizing constant κ in (2.5), may not be a simple matter as it requires multifold integration over the Bayesian kernel $L(\mathbf{\Theta})p(\mathbf{\Theta})$. An algorithm for computing these statistics is described in this section.

In the application of the Bayes formula, we need to compute integrals of the form

$$I = \int B(\mathbf{\Theta})\, d\mathbf{\Theta} \tag{2.60}$$

where $B(\mathbf{\Theta}) = w(\mathbf{\Theta})L(\mathbf{\Theta})p(\mathbf{\Theta})$ is the Bayesian integrand. The choice of the vector-valued function $w(\mathbf{\Theta})$ depends on the desired posterior statistic. If $w(\mathbf{\Theta}) \equiv 1$ is selected, the integral yields $I = 1/\kappa$ as the inverse of the normalizing constant in the Bayesian updating formula. If $w(\mathbf{\Theta}) = \kappa\mathbf{\Theta}$ is selected, the integral yields the posterior mean vector $\mathbf{M_\Theta}$ of the parameters. Finally, if $w(\mathbf{\Theta}) = \kappa\mathbf{\Theta\Theta}'$ is selected, I yields the mean square matrix $E[\mathbf{\Theta\Theta}']$, from which the covariance matrix can be computed as $\mathbf{\Sigma_{\Theta\Theta}} = E[\mathbf{\Theta\Theta}'] - \mathbf{M_\Theta}\mathbf{M_\Theta}'$.

An algorithm for computing the integral in (2.60) using multi-dimensional Gauss quadrature rules is developed by Geyskens *et al*. (1993). The approach works well for small a number of parameters, about up to 4. For a larger number of parameters, which is the case for the applications described in the following chapters, an alternative approach is needed. In the following, we describe an importance sampling method (Ditlevsen and Madsen, 1996) that we have successfully used for as many as 12 parameters. For the purpose of these applications, the algorithm was programmed in Matlab (1999).

By using an importance sampling density $S(\mathbf{\Theta})$, such that $S(\mathbf{\Theta}) \neq 0$ wherever $B(\mathbf{\Theta}) \neq 0$, the Bayesian integral (2.60) is modified to read

$$I = \int \left[\frac{B(\mathbf{\Theta})}{S(\mathbf{\Theta})}\right] S(\mathbf{\Theta})\, d\mathbf{\Theta} \tag{2.61}$$

It is seen that the value of the integral is equivalent to the expectation of the ratio $B(\mathbf{\Theta})/S(\mathbf{\Theta})$ relative to the sampling density. By using this property, the integral is estimated by randomly sampling realizations $\mathbf{\Theta}_i$, $i = 1, \ldots, N$, of $\mathbf{\Theta}$ according to the sampling density $S(\mathbf{\Theta})$, and computing the sample mean

$$\bar{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{B(\mathbf{\Theta}_i)}{S(\mathbf{\Theta}_i)} \tag{2.62}$$

The sample variance, divided by N, can be used as a measure of accuracy of the estimate. In particular, since κ is also unknown and has to be estimated along with $\mathbf{\Theta}$, the estimate of the coefficient of variation (c.o.v.) of $\mathrm{E}[\mathbf{\Theta}/\kappa]$ is used to formulate a criterion for terminating the simulation. The c.o.v. is estimated as

$$\mathrm{c.o.v.}\left(\frac{\mathbf{\Theta}}{\kappa}\right) = \frac{\sqrt{\sum_{i=1}^{N}\left[\mathbf{\Theta}_i L(\mathbf{\Theta}_i)p(\mathbf{\Theta}_i)/S(\mathbf{\Theta}_i)\right]^2 - \left\{\sum_{i=1}^{N}\left[\mathbf{\Theta}_i L(\mathbf{\Theta}_i)p(\mathbf{\Theta}_i)/S(\mathbf{\Theta}_i)\right]\right\}^2}}{\sum_{i=1}^{N}\left[\mathbf{\Theta}_i L(\mathbf{\Theta}_i)p(\mathbf{\Theta}_i)/S(\mathbf{\Theta}_i)\right]} \tag{2.63}$$

We terminate the simulation when the above c.o.v. is sufficiently small, about less than 0.10 or 0.05.

The key issue in this approach is the selection of an appropriate sampling density such that the ratio $B(\mathbf{\Theta})/S(\mathbf{\Theta})$ remains more or less constant at all sampling points. Obviously, the best choice is a sampling density that is proportional to the Bayesian integrand $B(\mathbf{\Theta})$. Of course this choice is not practical, but we must try to be as close to it as possible. An effective choice for the sampling density is described below.

It is well known (Richards, 1961) that under some mild conditions, the difference between the value of $\mathbf{\Theta}$ that maximizes the likelihood function, i.e., the so-called maximum-likelihood estimator, and the posterior mean $\mathbf{M}_{\mathbf{\Theta}}$ asymptotically approaches zero as the number of observations grow. Furthermore, the negative of the inverse of the Hessian of the log-likelihood function, $-[\nabla\nabla \ln L(\mathbf{\Theta})]^{-1}$, evaluated at the maximum-likelihood estimator, asymptotically approaches the posterior covariance matrix, $\mathbf{\Sigma}_{\mathbf{\Theta}\mathbf{\Theta}}$. We use these approximate second moments to construct the sampling density $S(\mathbf{\Theta})$. For this purpose, we make use of the Nataf joint distribution model developed by Liu and Der Kiureghian (1986), which is completely

defined by the second moments and marginal distributions of the random variables. In case of a q-variate model, we define $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_q, \boldsymbol{\sigma}, \boldsymbol{\rho})$, where $\boldsymbol{\theta}_k = (\theta_{ki}, i = 1, \ldots, p_k)$, $k = 1, \ldots, q$, $\boldsymbol{\sigma} = (\sigma_k, k = 1, \ldots, q)$ and $\boldsymbol{\rho}$ represents the set of $q(q-1)/2$ unique correlation coefficients $\boldsymbol{\rho} = (\rho_{kl}, \ k = 1, \ldots, q-1, \ l = k+1, \ldots, q)$. (See Chapter 3 for details.) Owing to the applicable ranges of the parameters, we select the marginal sampling distributions of θ_{ki}, $k = 1, \ldots, q$, $i = 1, \ldots, p_k$, to be normal, the marginal sampling distributions of σ_k, $k = 1, \ldots, q$, to be lognormal, and the marginal sampling distributions of ρ_{kl}, $k = 1, \ldots, q-1$, $l = k+1, \ldots, q$, to be beta within the interval $[-1, +1]$. In application, this choice of the sampling density has been found to be effective in reducing the required number of simulations for accurate estimation of the integral value in (2.61).

2.2.5   Computation of Posterior Statistics Using Bootstrap Methods

Another approach to estimate the statistics of the regression parameters is by resampling procedures, such as the jackknife and delta method that have been used starting from the late 1940s. These computer-intensive methods make use of extensive repeated calculations to explore the sampling distribution of a parameter estimator $\hat{\boldsymbol{\Theta}}$. In particular, bootstrap methods go back to Efron (1979), who unified the concepts of several resampling procedures and introduced the idea of resampling the data with replacement.

To introduce the general idea of bootstrap methods, suppose we have a random sample $y_1, \ldots, y_n$ drawn independently from one member of a parametric family $\{F_{\boldsymbol{\Theta}} | \boldsymbol{\Theta} \in I_{\boldsymbol{\Theta}}\}$ of distributions and suppose that $\boldsymbol{\Theta} = T(y)$ is a symmetric function of the sample, i.e., it does not depend on the order of the sample. Based on Efron's idea (Efron, 1982; Efron and Tibshirani, 1993; Davison and Hinkley, 1997), the bootstrap procedures assess the variability of $\hat{\boldsymbol{\Theta}}$ about the unknown true value $\boldsymbol{\Theta}$ by the variability of $\hat{\boldsymbol{\Theta}}_b$, $b = 1, \ldots, B$, about $\hat{\boldsymbol{\Theta}}$, where $\hat{\boldsymbol{\Theta}}_b$ is calculated based on the b-th of B samples from $\mathbf{y}$ with replacement.

Bootstrap methods are very general (Chernick, 1999). They can be applied to linear and nonlinear regression models and can be used for least-squares or for any other estimation

method. In this section a bootstrap method for a general uni-variate regression model is described. The extension to the multi-variate case is straightforward.

Similar to (2.50), a general regression model can be written as

$$\mathbf{y} = \mathbf{H}(\mathbf{\theta}) + \sigma\mathbf{\varepsilon} \tag{2.64}$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations, $\mathbf{H}$ is a $n \times 1$ vector of functions h_i of known form and may depend on a fixed vector of covariates, $\mathbf{\theta}$ denotes the set of unknown model parameters, $\mathbf{\varepsilon}$ is a $n \times 1$ vector of random variables having zero mean and unit variance, and σ represents the standard deviation of the model errors. Expanding the matrices (2.64) can be written as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} h_1(\mathbf{\theta}) \\ \vdots \\ h_i(\mathbf{\theta}) \\ \vdots \\ h_n(\mathbf{\theta}) \end{bmatrix} + \sigma \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.65}$$

The unknown parameters $\mathbf{\Theta} = (\mathbf{\theta}, \sigma)$ can be estimated by various procedures, e.g., by minimizing some measure of distance, by maximum likelihood, or as the mean or mode of the posterior distribution (Cano, 1992). If we call the general estimate $\hat{\mathbf{\Theta}}$, the residuals are obtained as

$$\hat{\mathbf{\varepsilon}} = \frac{\mathbf{y} - \mathbf{H}(\hat{\mathbf{\theta}})}{\hat{\sigma}} \tag{2.66}$$

One bootstrap procedure consists then in bootstrapping the residuals, that is, we construct the distribution F_n placing probability $1/n$ on each $\hat{\varepsilon}_i$, and we generate bootstrap residuals ε_i^* for $i = 1, \ldots, n$ by sampling independently from F_n (i.e., we sample with replacement from $\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n$). We then generate a bootstrap sample data set as

$$\mathbf{y}^* = \mathbf{H}(\hat{\mathbf{\theta}}) + \hat{\sigma}\mathbf{\varepsilon}^* \tag{2.67}$$

For each generated bootstrap data set $\mathbf{y}^*$, we obtain the estimate $\hat{\boldsymbol{\Theta}}^*$, with the same technique as before. This procedure is repeated B times, obtaining the bootstrap replications $\hat{\boldsymbol{\Theta}}_1^*,\ldots,\ \hat{\boldsymbol{\Theta}}_B^*$. Then, according to Efron (1982), we can take as an estimate of the covariance matrix of $\hat{\boldsymbol{\Theta}}$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\boldsymbol{\Theta}}_b^* - \hat{\boldsymbol{\Theta}}^*\right)\left(\hat{\boldsymbol{\Theta}}_b^* - \hat{\boldsymbol{\Theta}}^*\right)^T \qquad (2.68)$$

where

$$\hat{\boldsymbol{\Theta}}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\boldsymbol{\Theta}}_b^* \qquad (2.69)$$

3   Probabilistic Capacity Models

3.1   INTRODUCTION

Predictive capacity models in current structural engineering practice are typically deterministic and on the conservative side. These models were developed using simplified mechanics rules and conservatively fitting to available experimental data. As a result, they do not explicitly account for the uncertainty inherent in the model and they provide biased estimates of the capacity. While these deterministic models have been successfully used to design safe structures, the needs of modern structural engineering practice, and especially the advent of the performance-based design concept, require predictive capacity models that are unbiased, that is, on the average correctly predict the mean, and which account for all the prevailing uncertainties.

This chapter presents a Bayesian framework for the development of multi-variate probabilistic capacity models for structural components that account for the most relevant uncertainties, including model errors arising from an inaccurate model form or missing variables, measurement errors, and statistical uncertainty. With the aim of facilitating their use in practice, rather than developing new capacity models, we employ existing deterministic capacity models in common use but add correction terms to properly account for the inherent bias and uncertainty in these models. Through the use of a set of "explanatory" functions, we are able to identify terms that correct the bias in an existing model and provide insight into the underlying behavioral phenomena. Although the methodology described in this chapter is aimed at developing probabilistic capacity models, the approach is general and can be applied to the assessment (i.e., model selection and parameter estimation) of models in many engineering

problems. In Chapter 4, the methodology presented here is applied to develop probabilistic shear and deformation capacity models for RC columns under cyclic loading.

3.2 CAPACITY MODELS

In the context of this work, a "model" is a mathematical expression relating one or more quantities of interest, e.g., the capacities of a structural component, to a set of measurable variables $\mathbf{x} = (x_1, x_2, \ldots)$, e.g., material property constants, member dimensions, or imposed boundary conditions. The main purpose of the model is to provide a means for predicting the quantities of interest for given deterministic or random values of the variables $\mathbf{x}$. The model is said to be uni-variate when only one quantity is to be predicted and multi-variate when several quantities are to be predicted. We begin our discussion with the uni-variate form of the model and then generalize to the multi-variate case.

A uni-variate capacity model has the general form

$$C = C(\mathbf{x}, \mathbf{\Theta}) \tag{3.1}$$

where $\mathbf{\Theta}$ denotes a set of parameters introduced into the model to "fit" the model to observed data and C is the capacity quantity of interest. The function $C(\mathbf{x}, \mathbf{\Theta})$ can have a general form involving algebraic expressions, integrals, or derivatives. Ideally, it should be derived from first principles, e.g., the rules of mechanics. For the applications described in Chapter 4, rather than developing new models, we adopt commonly used deterministic models, to which we add correction terms. We believe this approach will facilitate the use of the resulting probabilistic models in practice. With this in mind, we adopt the general uni-variate model form

$$C(\mathbf{x}, \mathbf{\Theta}) = \hat{c}(\mathbf{x}) + \gamma(\mathbf{x}, \mathbf{\theta}) + \sigma\varepsilon \tag{3.2}$$

where $\hat{c}(\mathbf{x})$ is a selected deterministic model, $\gamma(\mathbf{x}, \mathbf{\theta})$ is a correction term for the bias inherent in the deterministic model that is expressed as a function of the variables $\mathbf{x}$ and parameters $\mathbf{\theta} = (\theta_1, \theta_2, \ldots)$, ε is a normal random variable with zero mean and unit variance, σ represents the standard deviation of the model error, and $\mathbf{\Theta} = (\mathbf{\theta}, \sigma)$ denotes the set of unknown model

parameters. Note that for given $\mathbf{x}$, $\boldsymbol{\theta}$, and σ, we have $\text{Var}\big[C(\mathbf{x},\boldsymbol{\Theta})\big] = \sigma^2$ as the variance of the model.

The above additive model correction form is valid under the following assumptions: (a) the model standard deviation is independent of $\mathbf{x}$ (homoskedasticity assumption) and (b) the model error has the normal distribution (normality assumption). Employing a suitable transformation of each capacity measure approximately satisfies these assumptions. For a positive-valued quantity Y, Box and Cox (1964) have suggested a parameterized family of transformations of the form

$$C = \frac{Y^\lambda - 1}{\lambda} \qquad \lambda \neq 0$$

$$= \ln Y \qquad \lambda = 0$$

(3.3)

where Y denotes the quantity of interest in the original space and λ is a parameter that defines a particular transformation. As special cases, $\lambda = 0$ specifies the logarithmic transformation, $\lambda = 1/2$ specifies the square-root transformation, $\lambda = 1$ is the linear transformation, and $\lambda = 2$ specifies the quadratic transformation. Under the assumptions of homoskedasticity and normality, one can formulate the posterior distribution of λ by use of Bayes's theorem and estimate its value for given data. However, in many practical situations, the model formulation itself often suggests the most suitable transformation. Diagnostic plots of the data or the residuals against model predictions or individual regressors can be used to verify the suitability of an assumed transformation (Rao and Toutenburg, 1997).

As defined earlier, the function $\gamma(\mathbf{x},\boldsymbol{\theta})$ corrects the bias in the deterministic model $\hat{c}(\mathbf{x})$. Since the deterministic model usually involves approximations, the true form of $\gamma(\mathbf{x},\boldsymbol{\theta})$ is unknown. In order to explore the sources of bias in the deterministic model, we select a suitable set of p "explanatory" basis functions $h_i(\mathbf{x})$, $i = 1, \ldots, p$, and express the bias correction term in the form

$$\gamma(\mathbf{x},\boldsymbol{\theta}) = \sum_{i=1}^{p} \theta_i h_i(\mathbf{x})$$

(3.4)

By examining the posterior statistics of the unknown parameters θ_i, we are able to identify those explanatory functions that are significant in describing the bias in the deterministic model. Note that while the bias correction term is linear in the parameters θ_i, it is not necessarily linear in the basic variables $\mathbf{x}$.

A structural component may have several capacity measures with respect to the demands placed on it. For example, a RC column has different capacities relative to failure in shear, bending, reinforcing bar pullout or buckling, or excessive deformation. For the analysis of such a component, a q-dimensional multi-variate capacity model is formulated in the form

$$C_k(\mathbf{x},\boldsymbol{\theta}_k,\boldsymbol{\Sigma}) = \hat{c}_k(\mathbf{x}) + \gamma_k(\mathbf{x},\boldsymbol{\theta}_k) + \sigma_k \varepsilon_k, \qquad k = 1,\ldots,q \qquad (3.5)$$

where

$$\gamma_k(\mathbf{x},\boldsymbol{\theta}_k) = \sum_{i=1}^{p_k} \theta_{ki} h_{ki}(\mathbf{x}), \qquad k = 1,\ldots,q \qquad (3.6)$$

With the exception of the new term $\boldsymbol{\Sigma}$, all entries in the above expressions have definitions analogous to those of the uni-variate model. $\boldsymbol{\Sigma}$ denotes the covariance matrix of the variables $\sigma_k \varepsilon_k$, $k = 1,\ldots,q$, with its (k,l) element being $\rho_{kl}\sigma_k\sigma_l$, where ρ_{kl} denotes the correlation coefficient between ε_k and ε_l. The set of unknown parameters of the model in (3.5) is $\boldsymbol{\Theta} = (\boldsymbol{\theta},\boldsymbol{\Sigma})$, where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_q)$ and $\boldsymbol{\theta}_k = (\theta_{k1},\ldots,\theta_{kp_k})$. Considering symmetry, $\boldsymbol{\Sigma}$ includes q unknown variances σ_k^2, $k = 1,\ldots,q$, and $q(q-1)/2$ unknown correlation coefficients ρ_{kl}, $k = 1,\ldots,q-1$, $l = k+1,\ldots,q$.

3.3   UNCERTAINTIES IN MODEL ASSESSMENT AND PREDICTION

In assessing a model, or in using a model for prediction purposes, one has to deal with two broad types of uncertainties: aleatory uncertainties (also known as inherent variability or randomness) and epistemic uncertainties. The former are those inherent in nature; they cannot be influenced by the observer or the manner of the observation. Referring to the model formulations in the preceding section, this kind of uncertainty is present in the variables $\mathbf{x}$ and partly in the error

terms ε_k. The epistemic uncertainties are those that arise from our lack of knowledge, our deliberate choice to simplify the model, from errors that arise in measuring observations, and from the finite size of observation samples. This kind of uncertainty is present in the model parameters $\boldsymbol{\Theta}$ and partly in the error terms ε_k. The fundamental difference between the two types of uncertainties is that whereas aleatory uncertainties are irreducible, epistemic uncertainties are reducible, e.g., by use of higher-order models, more accurate measurements, and collection of additional samples. The specific types of uncertainties that arise in assessing capacity models are described below. For simplicity in the notation, we use the formulation of a uni-variate model.

Model inexactness: Mathematical models are used in all aspects of performance-based earthquake engineering (PBEE), starting from the modeling of the input motion, characterization of the ground effects and structural response, and assessment of decision variables. For example, an attenuation law used to predict the spectral displacement at a site for a given earthquake magnitude and location is a mathematical model. Similarly, a finite element model of the site and the structure, including the employed material laws, geometric configurations, and simplified mechanics rules, is a mathematical model. So is a rule describing the down time of a structure for a given level of structural damage.

Without exception, all mathematical models are idealizations of reality and, therefore, implicit with error and uncertainty. The uncertainty associated with some of the models used in PBEE is well known and quantified. For example, the error inherent in an attenuation law is quantified in the process of fitting the model to observed data. In other cases, the model uncertainty remains completely unknown.

An important area where the model uncertainty remains largely unknown is where models are used to predict the response of soils and structures. If the soil/structure remains within the linear elastic range, the existing models of linear theory are fairly accurate. These models have been validated against numerous laboratory and field observations. However, when the soil or structure behavior is in the inelastic range, the behavior is a lot more complex and our mathematical models are unable to capture that complexity. As a general rule, model uncertainty tends to increase with the severity of structural response. This is the subject of most interest in PBEE. Models used in the current practice to assess damage and collapse of structures are likely

to have large errors relative to real structures. This uncertainty is likely to be far greater than the uncertainty arising from the natural variability in materials and some loads.

The only way to assess model uncertainty is to compare model predictions with real-world observations, either in the field or in the laboratory (with proper account of the departure of laboratory specimens from field reality). Observations of building response after major earthquakes, including the occurrence or non-occurrence of damage or collapse, provide valuable information for model assessment. Although detailed measurements are most informative, observations without measurements can also be used. Note that observations of no damage or no collapse after an earthquake can be as informative as the observation of damage or collapse. Laboratory observations can be used to assess models at component level, as most laboratory tests are conducted for structural components.

With our notation, this type of uncertainty arises when approximations are introduced in the derivation of the deterministic model $\hat{c}(\mathbf{x})$. It has two essential components: error in the form of the model, e.g., a linear expression is used when the actual relation is nonlinear, and missing variables, i.e., $\mathbf{x}$ contains only a subset of the variables that influence the quantity of interest. In (3.2), the term $\gamma(\mathbf{x}, \boldsymbol{\theta})$ provides a correction to the form of the deterministic model, whereas the error term $\sigma\varepsilon$ represents the influence of the missing variables as well as that of the remaining error due to the inexact model form. Since the effect of missing variables are inherently random, that component of ε that represents the influence of the missing variables has aleatory uncertainty, whereas the component representing inexact model form has epistemic uncertainty. In practice, it is difficult to distinguish the two uncertainty components of ε. However, after correction of the model form with the term $\gamma(\mathbf{x}, \boldsymbol{\theta})$, one can usually assume that most of the uncertainty inherent in ε is of aleatory nature. The coefficient σ represents the standard deviation of the model error arising from model inexactness.

Measurement error: Uncertainty arises from errors inherent in our laboratory or field measurements. This kind of uncertainty is also present when certain variables in a model remain unknown, such as in the case of assessing the capacity of an existing building where the material strength cannot be directly measured. Measurement uncertainty can be reduced by use of more accurate measurement devices and procedures.

As we shall shortly see, the parameters of the model are assessed by use of a sample of observations C_i of the dependent variable (corresponding to observed values Y_i in the original space) for observed values $\mathbf{x}_i$, $i = 1, \ldots, n$, of the independent variables. These observed values, however, could be inexact due to errors in the measurement devices or procedures. To model these errors, let $C_i = \hat{C}_i + e_{Ci}$ and $\mathbf{x}_i = \hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}$ be the true values for the i-th observation, where $\hat{C}_i$ and $\hat{\mathbf{x}}_i$ are the measured values and e_{Ci} and $\mathbf{e}_{\mathbf{x}i}$ are the respective measurement errors. The statistics of the measurement errors can be obtained through calibration of measurement devices and procedures. The mean values of these errors represent biases in the measurements, whereas their variances represent the uncertainties inherent in the measurements. In most engineering problems the random variables e_{Ci} and $\mathbf{e}_{\mathbf{x}i}$ can be assumed to be statistically independent and normally distributed. The uncertainty arising from measurement errors is epistemic in nature, since improving the measurement devices or procedures can reduce it.

Statistical uncertainty: Statistical uncertainty arises from the sparseness of data. Gathering more data can reduce it. If additional data cannot be gathered, then one must properly account for the effect of this uncertainty in all predictions.

In particular, the accuracy of estimation of the model parameters $\boldsymbol{\Theta}$ depends on the observation sample size, n, among other things. The smaller the sample size, the larger the uncertainty in the estimated values of the parameters. This uncertainty can be measured in terms of the estimated variances of the parameter. Statistical uncertainty is epistemic in nature, as it can be reduced by further collection of data.

3.4 LIKELIHOOD FUNCTION

As mentioned earlier, the likelihood is a function that is proportional to the conditional probability of the observations for given values of the model parameters. Formulation of the likelihood function depends on the type and form of the available information. Here, we start by considering the uni-variate model with exact measurements. Next, the effect of measurement

error is then incorporated in an approximate manner. The formulation is extended to multi-variate models.

In observing the state of a structural component in a laboratory test or in the field with respect to a specific mode of failure, one of three possible outcomes may be realized: (a) the demand is measured at the instant of failure, in which case the measured demand represents the component capacity; (b) the component does not fail, in which case the measured demand represents a lower bound to the component capacity; and (c) the component has failed under a lower demand than measured, in which case the measured demand represents an upper bound to the component capacity. These observations are categorized as three types of data, as described below.

Failure datum — observed value of the capacity C_i for a given $\mathbf{x}_i$, measured at the instant when the component fails. By using (3.2), we have $C_i = \hat{c}(\mathbf{x}_i) + \gamma(\mathbf{x}_i, \boldsymbol{\theta}) + \sigma\varepsilon_i$ or $\sigma\varepsilon_i = r_i(\boldsymbol{\theta})$, where

$$r_i(\boldsymbol{\theta}) = C_i - \hat{c}(\mathbf{x}_i) - \gamma(\mathbf{x}_i, \boldsymbol{\theta}) \tag{3.7}$$

denotes the outcome of the model error term at the i-th observation.

Lower-bound datum — observed value of a lower-bound C_i to the capacity for a given $\mathbf{x}_i$, when the component does not fail. In this case we have $C_i < \hat{c}(\mathbf{x}_i) + \gamma(\mathbf{x}_i, \boldsymbol{\theta}) + \sigma\varepsilon_i$ or $\sigma\varepsilon_i > r_i(\boldsymbol{\theta})$.

Upper-bound datum — observed value of an upper-bound C_i to the capacity for a given $\mathbf{x}_i$, when the component is known to have failed at a lower demand level. In this case we have $C_i > \hat{c}(\mathbf{x}_i) + \gamma(\mathbf{x}_i, \boldsymbol{\theta}) + \sigma\varepsilon_i$ or $\sigma\varepsilon_i < r_i(\boldsymbol{\theta})$. Lower- and upper-bounded data are often referred to as "censored data."

With exact measurements, and under the assumption of statistically independent observations, the likelihood function for the uni-variate model has the general form

$$L(\boldsymbol{\theta}, \sigma) \propto \prod_{\text{failure data}} p[\sigma\varepsilon_i = r_i(\boldsymbol{\theta})] \times \prod_{\substack{\text{lower bound} \\ \text{data}}} P[\sigma\varepsilon_i > r_i(\boldsymbol{\theta})] \times \prod_{\substack{\text{upper bound} \\ \text{data}}} P[\sigma\varepsilon_i < r_i(\boldsymbol{\theta})] \tag{3.8}$$

Since ε has the standard normal distribution, we can write

$$L(\mathbf{\theta},\sigma) \propto \prod_{\text{failure data}} \left\{ \frac{1}{\sigma} \varphi \left[\frac{r_i(\mathbf{\theta})}{\sigma} \right] \right\} \times \prod_{\text{lower bound data}} \Phi \left[-\frac{r_i(\mathbf{\theta})}{\sigma} \right] \times \prod_{\text{upper bound data}} \Phi \left[\frac{r_i(\mathbf{\theta})}{\sigma} \right] \quad (3.9)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal probability density function and the cumulative distribution function, respectively.

Now consider the case where measurement errors are present. Denote $\hat{C}_i$ and $\hat{\mathbf{x}}_i$ as the measured values in the i-th observation and e_{Ci} and $\mathbf{e}_{\mathbf{x}i}$ as the corresponding measurement errors. Without loss of generality, we assume the measurements have been corrected for any systematic error, so that the means of e_{Ci} and $\mathbf{e}_{\mathbf{x}i}$ are zeros. Let s_i^2 and $\mathbf{\Sigma}_i$ denote the variance of e_{Ci} and the covariance matrix of $\mathbf{e}_{\mathbf{x}i}$, respectively. As should be evident, we allow dependence between the measurement errors for different variables at each observation; however, we assume independence between the measurement errors at different observations, i.e., for different i. We also assume that the error terms are normally distributed. For the failure data we have $\hat{C}_i + e_{Ci} = \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \mathbf{\theta}) + \sigma\varepsilon_i$, for the lower-bound data $\hat{C}_i + e_{Ci} < \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \mathbf{\theta}) + \sigma\varepsilon_i$ and for the upper-bound data $\hat{C}_i + e_{Ci} > \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \mathbf{\theta}) + \sigma\varepsilon_i$. Defining

$$r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i}) = \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) - \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \mathbf{\theta}) \quad (3.10)$$

the conditions for the three types of data can now be written as $\sigma\varepsilon_i - e_{Ci} = r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i})$, $\sigma\varepsilon_i - e_{Ci} > r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i})$, and $\sigma\varepsilon_i - e_{Ci} < r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i})$, respectively. Unfortunately $r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i})$ in general is a nonlinear function of the random variables $\mathbf{e}_{\mathbf{x}i}$, which makes the computation of the likelihood function enormously more difficult. To overcome this difficulty, under the assumption that the errors $\mathbf{e}_{\mathbf{x}i}$ are small in relation to the measurements $\hat{\mathbf{x}}_i$, a first-order approximation is used to express $r_i(\mathbf{\theta}, \mathbf{e}_{\mathbf{x}i})$ as a linear function of $\mathbf{e}_{\mathbf{x}i}$. Using a Maclaurin series expansion around $\mathbf{e}_{\mathbf{x}i} = \mathbf{0}$, we have

$$r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i}) \cong \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i) - \gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) - \left[\nabla_{\hat{\mathbf{x}}_i} \hat{c}(\hat{\mathbf{x}}_i) + \nabla_{\hat{\mathbf{x}}_i} \gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) \right] \mathbf{e}_{\mathbf{x}i}$$
$$= \hat{r}_i(\boldsymbol{\theta}) + \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta}) \mathbf{e}_{\mathbf{x}i} \tag{3.11}$$

where $\nabla_{\mathbf{x}}$ denotes the gradient row vector with respect to $\mathbf{x}$ and

$$\hat{r}_i(\boldsymbol{\theta}) = \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i) - \gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) \tag{3.12}$$

The conditions for the three types of data can now be written as $\sigma\varepsilon_i - e_{Ci} - \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta}) \mathbf{e}_{\mathbf{x}i} = \hat{r}_i(\boldsymbol{\theta})$, $\sigma\varepsilon_i - e_{Ci} - \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta}) \mathbf{e}_{\mathbf{x}i} > \hat{r}_i(\boldsymbol{\theta})$, and $\sigma\varepsilon_i - e_{Ci} - \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta}) \mathbf{e}_{\mathbf{x}i} < \hat{r}_i(\boldsymbol{\theta})$, respectively. The left-hand sides of these expressions are a normal random variable with zero mean and variance $\hat{\sigma}^2(\boldsymbol{\theta}, \sigma) = \sigma^2 + s_i^2 + \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta}) \boldsymbol{\Sigma}_i \nabla_{\hat{\mathbf{x}}_i} \hat{r}_i(\boldsymbol{\theta})'$. Hence, in presence of measurement errors, the likelihood function approximate takes the form

$$L(\boldsymbol{\theta}, \sigma) \propto \prod_{\substack{\text{failure} \\ \text{data}}} \left\{ \frac{1}{\hat{\sigma}(\boldsymbol{\theta}, \sigma)} \varphi\left[\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta}, \sigma)} \right] \right\} \times \prod_{\substack{\text{lower bound} \\ \text{data}}} \Phi\left[-\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta}, \sigma)} \right] \times \prod_{\substack{\text{upper bound} \\ \text{data}}} \Phi\left[\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta}, \sigma)} \right] \tag{3.13}$$

We now consider the multi-variate model in (3.5) under exact measurements. For the i-th observation, any of the q capacity measures can be either directly observed, observed from below (lower-bound data), or observed from above (upper-bound data). However, these observations in general are dependent because of the correlation between the model error terms ε_k. As mentioned earlier, this correlation arises from the missing variables that may be common to all the capacity models. Figure 3.1 illustrates a conceptual representation of the various ways that the data for a bi-variate capacity model may appear. The curved lines indicate the limit states for the two failure modes and the areas with varying intensities of shading indicate regions of failure and non-failure with respect to each mode. The dots indicate hypothetical data points. It can be seen that the data points can be in $3^2 = 9$ different categories (i.e., lower-bound lower-bound data, lower-bound failure data, lower-bound upper-bound data, etc.). More generally, the data points for a q-variate model can be of at most 3^q different types.
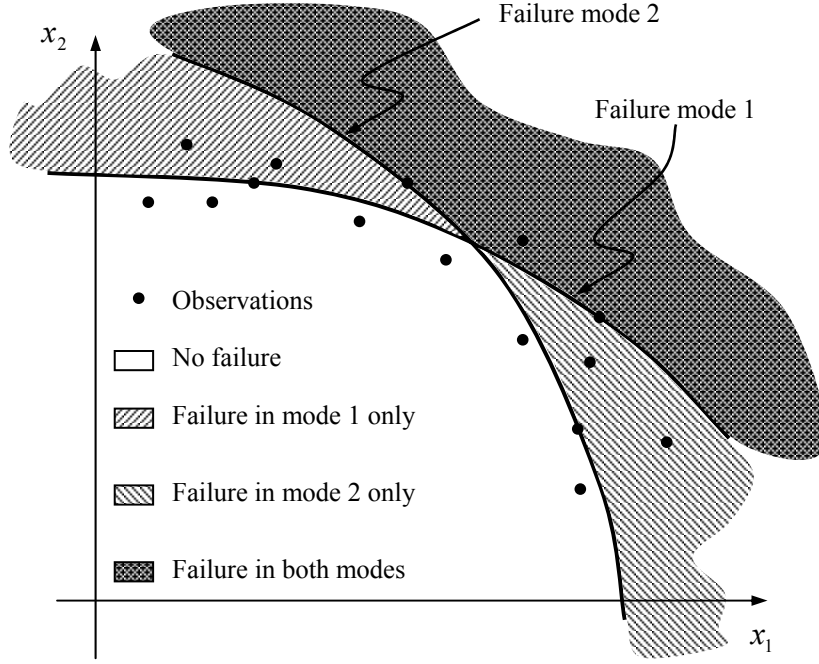
Figure 3.1.   Representation of data types.

For the i-th observation of the k-th capacity model, define

$$r_{ki}(\boldsymbol{\theta}_k) = C_{ki} - \hat{c}_k(\mathbf{x}_i) - \gamma_k(\mathbf{x}_i, \boldsymbol{\theta}_k) \qquad (3.14)$$

where C_{ki} is the measured value of the k-th capacity or its lower or upper bound. Also, let ε_{ki} be the outcome of the error term for the k-th capacity model in the i-th observation. Noting that any of the q capacity terms can be measured as a failure datum, lower-bound datum or upper-bound datum, the likelihood function takes the form

$$L(\boldsymbol{\Theta}) \propto \prod_{\substack{\text{observation} \\ i}} \mathrm{P}\left\{ \bigcap_{\substack{\text{failure} \\ \text{data } k}} [\sigma_k \varepsilon_{ki} = r_{ki}(\boldsymbol{\theta}_k)] \bigcap_{\substack{\text{lower bound} \\ \text{data } k}} [\sigma_k \varepsilon_{ki} > r_{ki}(\boldsymbol{\theta}_k)] \bigcap_{\substack{\text{upper bound} \\ \text{data } k}} [\sigma_k \varepsilon_{ki} < r_{ki}(\boldsymbol{\theta}_k)] \right\} \qquad (3.15)$$

45

where $\boldsymbol{\Theta}=(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_q,\boldsymbol{\Sigma})$. The events in the above expression in general are dependent because of the correlation between ε_{ki} for different indices k. The probability term for each observation can be computed by using the multi-normal probability density and cumulative distribution functions. As an example, Table 3.1 lists the expressions for these terms for a bi-variate model with only lower-bound and failure data.

Table 3.1.   Probability terms for a bi-variate capacity model with lower-bound and failure data.

		Capacity model 1	
		Failure	Lower bound
Capacity model 2 — Failure		$\dfrac{1}{\sigma_{1\|2}}\varphi\left[\dfrac{r_{1i}(\boldsymbol{\theta})-\mu_{1\|2}}{\sigma_{1\|2}}\right]\dfrac{1}{\sigma_2}\varphi\left[\dfrac{r_{2i}(\boldsymbol{\theta})}{\sigma_2}\right]$	$\Phi\left[-\dfrac{r_{1i}(\boldsymbol{\theta})-\mu_{1\|2}}{\sigma_{1\|2}}\right]\dfrac{1}{\sigma_2}\varphi\left[\dfrac{r_{2i}(\boldsymbol{\theta})}{\sigma_2}\right]$
Capacity model 2 — Lower bound		$\Phi\left[-\dfrac{r_{2i}(\boldsymbol{\theta})-\mu_{2\|1}}{\sigma_{2\|1}}\right]\dfrac{1}{\sigma_1}\varphi\left[\dfrac{r_{1i}(\boldsymbol{\theta})}{\sigma_1}\right]$	$\displaystyle\int_{r_{2i}}^{\infty}\Phi\left[-\dfrac{r_{1i}(\boldsymbol{\theta})-\mu_{1\|\zeta}}{\sigma_{1\|2}}\right]\dfrac{1}{\sigma_2}\varphi\left(\dfrac{\zeta}{\sigma_2}\right)d\zeta$

Note: $\mu_{k|l}=\rho_{kl}(\sigma_k/\sigma_l)r_{li}$, $\sigma_{k|l}=\sigma_k\sqrt{1-\rho_{kl}^2}$, $k,l=1,2$, and $\mu_{1|\zeta}=\rho_{12}(\sigma_1/\sigma_2)\zeta$.

Chapter 4 describes an application that uses these formulations for assessing a bi-variate capacity model for RC columns. Naturally, the required computational effort for evaluating the likelihood function grows with increasing dimension of the model.

In the presence of measurement error, the likelihood function for the multi-variate model remains similar to that in (3.15) with the term $\sigma_k\varepsilon_{ki}$ for the i-th observation of the k-th model replaced by $\sigma_k\varepsilon_{ki}-e_{Cki}-\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\mathbf{e}_{\mathbf{x}i}$ and $r_{ki}(\boldsymbol{\theta}_k)$ replaced by $\hat{r}_{ki}(\boldsymbol{\theta}_k)$, which is equivalent to (3.12) for the k-th model. This former term is a zero mean normal random variable with variance $\hat{\sigma}_k^2(\boldsymbol{\theta}_k,\sigma_k)=\sigma_k^2+s_{ki}^2+\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\boldsymbol{\Sigma}_i\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)'$ for the i-th observation. Furthermore, the terms for the k-th and l-th models have the covariance

$\hat{\rho}_{kl}\hat{\sigma}_k\hat{\sigma}_l = \rho_{kl}\sigma_k\sigma_l + \nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\boldsymbol{\Sigma}_i\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{li}(\boldsymbol{\theta}_l)'$, where $\hat{\rho}_{kl}$ denotes the correlation coefficient. For the bi-variate model described in Table 3.1, the formulation remains the same with σ_k, ρ_{kl}, and $r_{ki}(\boldsymbol{\theta}_k)$ replaced by $\hat{\sigma}_k$, $\hat{\rho}_{kl}$, and $\hat{r}_{ki}(\boldsymbol{\theta}_k)$, respectively.

3.5   MODEL SELECTION

For the sake of simplicity of notation, the discussion in this section is focused on the uni-variate model. However, the concepts discussed are equally applicable to a multi-variate model.

The probabilistic model in (3.2) and (3.4) requires the selection of the deterministic model $\hat{c}(\mathbf{x})$ and a set of explanatory functions $h_i(\mathbf{x})$, $i = 1, \ldots, p$. For practical prediction purposes, the selection process should aim at a model that is unbiased, accurate, and can be easily adopted in practice. Furthermore, from a statistical standpoint, it is desirable that the correction term $\gamma(\mathbf{x},\boldsymbol{\theta})$ has a parsimonious parameterization (i.e., has as few parameters θ_i as possible) in order to avoid loss of precision of the estimates and of the model due to inclusion of unimportant predictors and to avoid over-fit of the data.

The model form in (3.2) is unbiased by formulation. Furthermore, a good measure of its accuracy is represented by the standard deviation σ. Specifically, among a set of parsimonious candidate models (in terms of the selected forms of $\hat{c}(\mathbf{x})$ and $h_i(\mathbf{x})$), the one that has the smallest σ can be considered to be the most accurate. Therefore, an estimate of the parameter σ and of its standard deviation, e.g., its posterior mean and standard deviation, can be used to select the most accurate model among several viable candidates. This procedure is followed in Chapter 4 to compare two existing models for the shear capacity of RC columns.

The explanatory functions $h_i(\mathbf{x})$ should be selected to enhance the predictive capability of the deterministic model $\hat{c}(\mathbf{x})$. It is appropriate to select terms that are thought to be missing in $\hat{c}(\mathbf{x})$. Ideally, rules of mechanics should be used in formulating the explanatory functions. However, in many cases reliance on engineering judgment is necessary. It is also desirable that $h_i(\mathbf{x})$ have the same dimension as $\hat{c}(\mathbf{x})$ so that θ_i are dimensionless. It is best to start the model

assessment process with a comprehensive candidate form of $\gamma(\mathbf{x},\boldsymbol{\theta})$ and then simplify it by deleting unimportant terms or combining terms that are closely correlated. A step-wise deletion process may proceed as follows:

1. Compute the posterior statistics of the model parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ and σ.

2. Identify the term $h_i(\mathbf{x})$ whose coefficient θ_i has the largest posterior coefficient of variation. The term $h_i(\mathbf{x})$ is the least informative among all the explanatory functions, so one may select to drop it from $\gamma(\mathbf{x},\boldsymbol{\theta})$.

3. If $\theta_i h_i(\mathbf{x})$ is retained, determine the largest absolute value correlation coefficient $\left| \rho_{\theta_i \theta_j} \right| = \max_{k \neq i} \left| \rho_{\theta_i \theta_k} \right|$ between θ_i and the remaining parameters θ_k, $k \neq i$. A value of $\left| \rho_{\theta_i \theta_j} \right|$ close to 1, say $0.7 \leq \left| \rho_{\theta_i \theta_j} \right|$, is an indication that the information contents in $h_i(\mathbf{x})$ and $h_j(\mathbf{x})$ are closely related and that these two explanatory functions can be combined. On the other hand, a value of $\left| \rho_{\theta_i \theta_j} \right|$ small in relation to 1, say $\left| \rho_{\theta_i \theta_j} \right| \leq 0.5$, is an indication that the information content in $h_i(\mathbf{x})$ is not closely related to that in the remaining terms. If $0.7 \leq \left| \rho_{\theta_i \theta_j} \right|$, one can choose to replace θ_i by

$$\hat{\theta}_i = \mu_{\theta_i} + \rho_{\theta_i \theta_j} \frac{\sigma_{\theta_i}}{\sigma_{\theta_j}} \left(\theta_j - \mu_{\theta_j} \right) \tag{3.16}$$

where μ_{θ_i} and σ_{θ_i} are the posterior mean and standard deviation of θ_i, respectively. The above expression provides the best linear predictor of θ_i as a function of θ_j (Stone, 1996). This reduces one parameter in $\gamma(\mathbf{x},\boldsymbol{\theta})$.

4. Assess the reduced model of step 2 or 3 by estimating its parameters. If the posterior mean of σ has not increased by an unacceptable amount, accept the reduced model and return to step 2 or 3 for possible further reduction of the model. Otherwise, the reduction is not desirable and the model form before the reduction is as parsimonious as possible.

There is considerable room for judgment in the above procedure. This is a part of the art of model building. Applications in Chapter 4 demonstrate this step-wise model reduction procedure.

4 Applications of Probabilistic Capacity Models

4.1 INTRODUCTION

A Bayesian framework for the development of probabilistic capacity models for structural components was presented in Chapters 2 and 3. This approach seeks to explicitly account for the most relevant uncertainties, including errors arising from an inaccurate model form or missing variables, measurement errors, and statistical uncertainty. Moreover, systematic assessment of a measure of model quality can be made, thus allowing a comparison of the accuracy of alternative candidate models. Through "explanatory" functions in a model correction term, insight into the underlying behavioral phenomena is gained.

Large uncertainty is inherent in predicting the capacity of RC structural components under repeated cyclic loading (Park and Ang, 1985). At the same time, a large body of valuable experimental data is available that has not been fully utilized. These facts have motivated us to employ the methodology presented in the previous chapters to develop probabilistic deformation and shear capacity models for RC circular columns (Gardoni *et al*., 2002). This specific class of structural components is selected because of their predominant use for bridge structures in many seismically active regions of the world.

As described in Chapter 3, the probabilistic models can be built upon existing deterministic models. The deformation capacity model used in this study is based on the notion of decomposing the total displacement of the RC column into its basic components. Specifically, the column displacement is considered to be composed of elastic and inelastic components, with the elastic component itself consisting of contributions from the flexural and shear deformations and from the slip of the longitudinal reinforcing bars. For the shear capacity model, owing to the complex nature of the underlying load transfer mechanisms, a unique consensus model does not

exist. Here, we consider two alternative deterministic models used in practice and assess objective measures of their relative qualities. The more accurate shear capacity model is subsequently used together with the deformation capacity model to formulate a bi-variate deformation-shear capacity model.

4.2   EXPERIMENTAL DATA

The behavior of RC columns under the effect of repeated cyclic loading has been the focus of experimental research by a large number of investigators for many years. A large collection of these experimental data is organized at the World Wide Web site http://maximus.ce.washington.edu/~peera1/ where references to the original publications for each tested column are listed. At the time of this writing, the database contained the results of cyclic lateral load tests on 134 circular or octagonal columns, 188 rectangular columns, 11 retrofitted columns, and 4 spliced columns from 74 different experimental studies conducted by 115 investigators. During these experiments, all columns were subjected to constant axial loads. For the purpose of this study, out of the 134 tested columns with circular or octagonal cross sections, we originally considered the first 117 that were available at the time (June 2000). In this database, columns 38, 69, 75, and 82 did not include spiral reinforcement; column 106 had missing data; and columns 108 and 111 were subjected to tensile axial load. These columns were excluded in the present study. Furthermore, column 53 was selected as a sample column for the subsequent fragility analysis and, hence, was also excluded. Following an initial analysis, the data for columns 50, 51, and 52 were identified as outliers because the reported test data appeared to be inconsistent with reasonable predictions. These columns were also excluded from further consideration. Thus, the analysis reported in this study is based on the data from the remaining 106 columns.

Reported in the database are the material properties and geometry of each test column. The ranges of the important variables for the considered columns are listed in Table 4.1.

Table 4.1.   **Ranges of the variables from the database**.

Variable	Symbol	Range
Compressive strength of concrete [MPa]	f'_c	18.9-42.2
Yield stress of longitudinal reinforcement [MPa]	f_y	207-607
Ultimate strength of longitudinal reinforcement [MPa]	f_{su}	396-758
Yield stress of transverse reinforcement [MPa]	f_{yh}	207-607
Longitudinal reinforcement ratio [%]	ρ_l	0.53-5.50
Volumetric transverse reinforcement ratio [%]	ρ_s	0.17-3.00
Slenderness ratio	H/D_g	1.09-10.00
Ratio of gross to core diameters	D_g/D_c	1.05-1.31
Axial load ratio	$4P/\pi D_g^2 f'_c$	0.00-0.87

In this table, H represents the equivalent cantilever length (clear column height) and D_g and D_c are the gross and core column diameters, respectively. For octagonal cross sections, the largest circle that can be included in the cross section is used. The database reports the applied constant axial load P, the cyclic lateral load-deformation relationships, and the mode of failure (shear, flexure, or combined shear-flexure) for all the tested columns. Since these are all laboratory experiments, the measurement errors were judged to be small in relation to the uncertainties in the models and were neglected.

4.3   DEFORMATION CAPACITY MODEL

Following common practice (Park and Paulay, 1975; Lynn *et al.*, 1996), the deformation capacity of a column is defined as the displacement Δ corresponding to a drop in the lateral force re-

sistance equal to 20% of its peak value. Figure 4.1a illustrates this definition for a cyclically loaded structural component.

The lateral load-deformation relationships from the experimental database were examined to determine Δ for each of the 106 tested columns. Three categories of observations were identified. First are columns for which the lateral force resistance is reached and followed by strength degradation up to the threshold drop of 20%. Figure 4.1b illustrates a representative case. This type of observation is identified as "failure" datum. Second are columns whose lateral force resistance is not reached because of premature load reversal. Figure 4.1c illustrates such a case, where stiffness deterioration occurs without reaching the lateral force resistance. The measured displacement in this case (i.e., that corresponding to 80% of the peak lateral load) is obviously a lower bound to the deformation capacity Δ. This type of observation is identified as "lower-bound" datum. Third are columns whose deformation capacity is not reached because of possible limitation in the applied maximum displacement. Figure 4.1d illustrates such a case. Obviously, the measured maximum displacement provides a lower bound to the deformation capacity Δ. This kind of observation is also a "lower-bound" datum, just described.
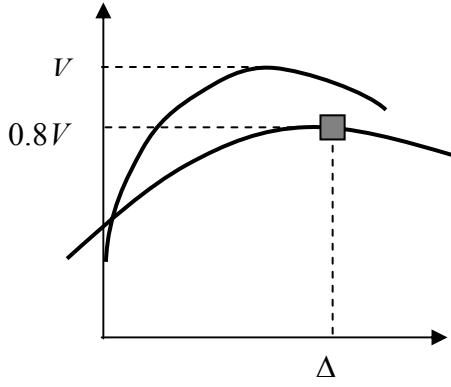
(a) Definition of deformation capacity Δ
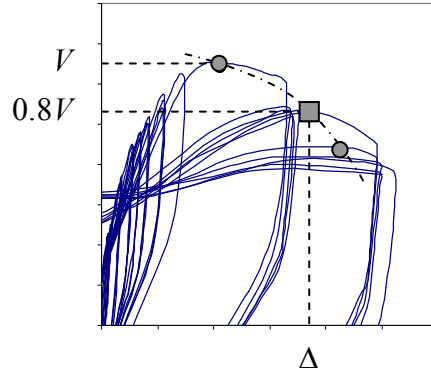
(b) Typical failure datum
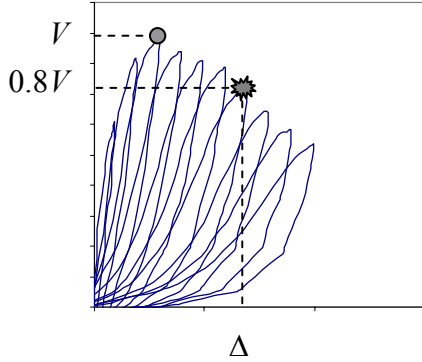
(c) Censored datum due to premature load reversal
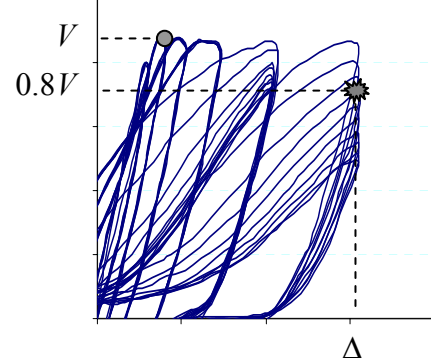
(d) Censored datum due to limited applied displacement

Figure 4.1.   Deformation capacity definition and data types.

To develop the deformation capacity model, we employ the drift ratio capacity $\delta = \Delta / H$, where H is the column height from its base to the centerline of the bridge deck. This is a dimensionless quantity, convenient for model formulation. Let $\hat{\delta}(\mathbf{x})$ be an existing deterministic model for predicting δ, where $\mathbf{x} = (f'_c, f_y, \ldots, H, D_g, \ldots, P)$ is the set of constituent material, geometry, and load variables. Considering the non-negative nature of the deformation capacity, the logarithmic variance-stabilizing transformation is selected among other possible transformations to formulate a homoskedastic model. Thus, we adopt the model form

$$\ln[\delta(\mathbf{x}, \mathbf{\Theta})] = \ln[\hat{\delta}(\mathbf{x})] + \gamma_\delta(\mathbf{x}, \mathbf{\theta}) + \sigma\varepsilon \qquad (4.1)$$

where $\mathbf{\Theta} = (\mathbf{\theta}, \sigma)$ is the set of unknown model parameters, $\gamma_\delta(\mathbf{x}, \mathbf{\theta})$ is a correction term for the bias inherent in the deterministic model, and ε is a standard normal random variable (i.e., with zero mean and unit variance). As described in Chapter 3, $\sigma\varepsilon$ represents the random component of the model error. Thus, σ denotes the standard deviation of the model error and is a measure of the quality of the model. Because of the employed logarithmic transformation, one can show that σ is approximately equal to the coefficient of variation (c.o.v.) of the drift ratio. What remains to be defined in the model are $\hat{\delta}(\mathbf{x})$ and $\gamma_\delta(\mathbf{x}, \mathbf{\theta})$. These are developed in the following two sections.

4.3.1 Deterministic Model

As stated earlier, $\hat{\delta}(\mathbf{x})$ on the right-hand side of (4.1) represents the drift ratio capacity predicted by a deterministic model. It is a common practice to decompose the ultimate displacement capacity of a column into two components: the elastic component $\hat{\Delta}_y$ due to the onset of yield, and the inelastic component $\hat{\Delta}_p$ due to the plastic flow, as illustrated in Figure 4.2 for a single RC column bridge bent. Accordingly,

$$\hat{\delta}(\mathbf{x}) = \frac{1}{H}\left(\hat{\Delta}_y + \hat{\Delta}_p\right) \qquad (4.2)$$
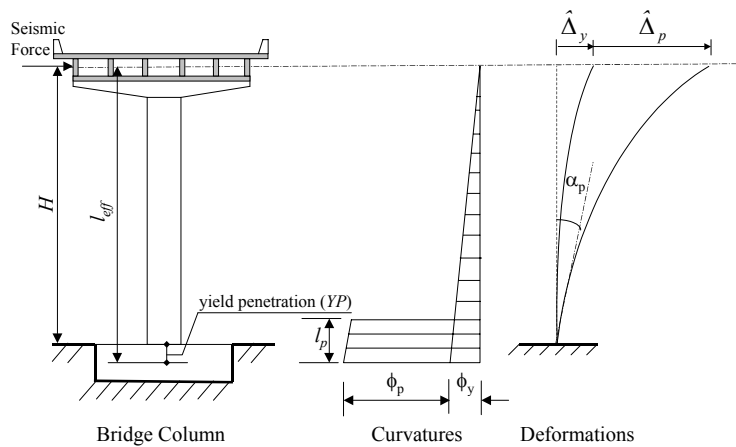
Figure 4.2.   Decomposition of lateral displacement of a single-column bridge bent.

For an RC column responding as a cantilever (i.e., with a single curvature) with fixed base, the yield displacement is comprises a flexural component $\hat{\Delta}_f$ based on a linear curvature distribution along the full column height (Figure 4.2), a shear component $\hat{\Delta}_{sh}$ due to shear distortion, and a slip component $\hat{\Delta}_{sl}$ due to the local rotation at the base caused by slipping of the longitudinal reinforcing bars. These three components are illustrated in Figure 4.3.
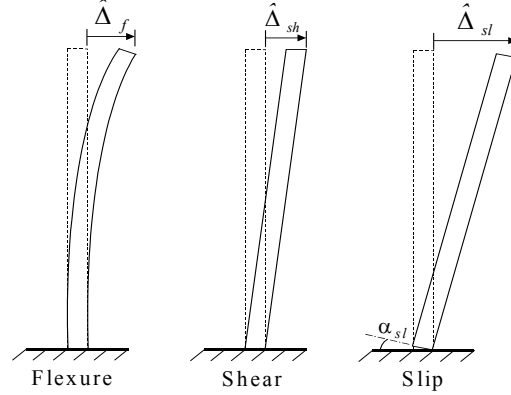
Figure 4.3. **Components of yield displacement $\hat{\Delta}_y$ for RC column.**

Thus, we have

$$\hat{\Delta}_y = \hat{\Delta}_f + \hat{\Delta}_{sh} + \hat{\Delta}_{sl} \tag{4.3}$$

Given the curvature ϕ_y at yield, the flexural component of the displacement is given by

$$\hat{\Delta}_f = \frac{1}{3}\phi_y l_{eff}^2 \tag{4.4}$$

where $l_{eff} = H + YP$ is the effective length of the column, in which YP denotes the depth of the yield penetration into the column base (Figure 4.2). The latter term accounts for the additional rotation of the critical section resulting from yield penetration of the longitudinal reinforcement into the column footing. According to Priestley *et al.* (1996), YP is estimated as

$$YP = 0.022 f_y d_b \tag{4.5}$$

where d_b is the diameter of the longitudinal reinforcement having yield stress f_y which must be expressed in units of MPa.

The shear deformation is obtained from the well-known expression

$$\hat{\Delta}_{sh} = \frac{V_y H}{GA_{ve}} \qquad (4.6)$$

where V_y is the shear force at yield, G is the shear modulus of concrete, and A_{ve} is the effective shear area. The latter is computed as $A_{ve} = k_I k_s A_g$, where A_g is the gross cross-sectional area, $k_s = 0.9$ is the shape factor for a circular cross section, and the factor k_I reflects the increased shear deformation in a flexurally cracked RC column. Owing to lack of specific research data, it is usually assumed that the reduction in shear stiffness is proportional to the reduction in flexural stiffness (Priestley *et al.*, 1996) such that $k_I = I_e / I_g$, where I_e is the effective moment of inertia determined from the moment-curvature relationship of the column cross section, as demonstrated in Figure 4.4, and I_g is the gross moment of inertia of the cross section.
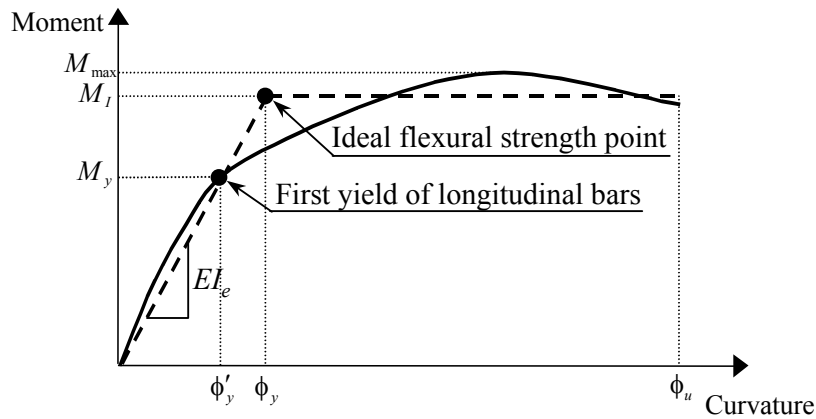
Figure 4.4.   Generic moment-curvature diagram.

The contribution to the yield deformation due to slippage of the column longitudinal reinforcing bars is related to the local rotation at the base of the column α_{sl}, (Figure 4.3). We adopt the assumptions by Pujol *et al.* (1999) whereby the bond stress at yielding is uniformly distributed and is given by $\mu = 1.08\sqrt{f_c'}$ when MPa units are used, and the column rotates about the

neutral axis of the flexurally critical section when slip takes place. These assumptions lead to $\alpha_{sl} = (\phi_y\, f_y\, d_b)/(8\mu)$. Accordingly,

$$\hat{\Delta}_{sl} = \frac{\phi_y\, f_y\, d_b\, H}{8.64\sqrt{f_c'}} \qquad\qquad (4.7)$$

The contributions from post-elastic flexural behavior, diagonal tension cracking, and yield penetration are manifested in the so-called plastic hinge rotation, α_p, as shown in Figure 4.2. An equivalent rectangular plastic curvature is commonly assumed (Park and Paulay, 1975). Accordingly, the plastic deformation $\hat{\Delta}_p$ in (4.2) is obtained from

$$\hat{\Delta}_p = \alpha_p\, H = \phi_p\, l_p\, H \qquad\qquad (4.8)$$

where $l_p = 0.08H + YP \geq 0.044\, f_y\, d_b$ is the equivalent plastic hinge length (Priestley *et al.*, 1996), in which f_y in the lower-bound limit must be expressed in units of MPa, and $\phi_p = \phi_u - \phi_y$ is the plastic curvature where ϕ_u denotes the ultimate curvature.

A generic moment-curvature relationship with an elastic-perfectly plastic idealization is illustrated in Figure 4.4 showing the definitions of ϕ_y and ϕ_u. In this figure, M_y is the moment that induces the first yielding in the column longitudinal reinforcement, ϕ_y' is the corresponding curvature, and M_I is the ideal (theoretical) moment capacity corresponding to the idealized yield curvature ϕ_y. In experiments on RC columns conducted by Priestley and Park (1987) and Watson and Park (1994), spalling of the concrete cover, which normally precedes the yielding of the longitudinal reinforcement, was found to take place when $\varepsilon_c \geq 0.005$, where ε_c is the longitudinal compressive strain of the extreme concrete fiber. In the present study M_I and the corresponding ϕ_y are conservatively determined by using $\varepsilon_c = 0.005$. For each tested RC column, all the moments and the corresponding curvatures defined above are computed by using fiber-element section analysis (Thewalt and Stojadinović, 1994).

From the elastic-perfectly plastic idealization in Figure 4.4, ϕ_y is determined from the linear extrapolation

$$EI_e = \frac{M_y}{\phi_y'} = \frac{M_I}{\phi_y} \tag{4.9}$$

On the other hand, the ultimate curvature ϕ_u corresponds to $\varepsilon_c = \varepsilon_{cu}$, where ε_{cu} accounts for the confining effects of the transverse reinforcement. This is conducted using the energy balance argument of Mander *et al.* (1988) leading to the conservative estimate

$$\varepsilon_{cu} = 0.004 + \frac{1.4\rho_s f_{yh}\varepsilon_{su}}{f_{cc}'} \tag{4.10}$$

where $\rho_s = 4A_h/D_cS$ is the volumetric ratio of the confining steel, in which A_h is the cross-sectional area of the transverse reinforcement and S is the longitudinal spacing of the hoops or spirals, ε_{su} is the strain at the maximum tensile stress of the transverse steel, which is commonly taken as 0.12 (Priestley *et al.*, 1996), and f_{cc}' is the compressive strength of the confined concrete, defined according to Mander *et al.* (1988) as

$$f_{cc}' = f_c'\left(2.254\sqrt{1 + \frac{7.94f_l'}{f_c'}} - 2\frac{f_l'}{f_c'} - 1.254\right) \tag{4.11}$$

where $f_l' = K_e f_{yh}\rho_s/2$ is the effective lateral confining stress, in which $K_e = 0.95$ is the confinement effectiveness coefficient for circular sections.

4.3.2 Model Correction

The term $\gamma_\delta(\mathbf{x},\boldsymbol{\theta})$ on the right-hand side of (4.1) is intended to correct the bias inherent in the deterministic model $\ln[\hat{\delta}(\mathbf{x})]$. As described in Chapter 3, we use the form

$$\gamma_\delta(\mathbf{x},\boldsymbol{\theta}) = \sum_{i=1}^{p}\theta_i h_i(\mathbf{x}) \tag{4.12}$$

where $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_p)$ is a vector of unknown model parameters and $h_1(\mathbf{x}),\ldots,h_p(\mathbf{x})$ are selected "explanatory" functions. To capture a potential bias in the model that is independent of the vari-

ables $\mathbf{x}$, we select $h_1(\mathbf{x}) = 1$. To detect any possible under- or overestimation of the individual contributions defined in the deterministic models (4.2) and (4.3) to the total deformation, we select the next four explanatory functions as $h_2(\mathbf{x}) = \hat{\Delta}_f / H$, $h_3(\mathbf{x}) = \hat{\Delta}_p / H$, $h_4(\mathbf{x}) = \hat{\Delta}_{sh} / H$ and $h_5(\mathbf{x}) = \hat{\Delta}_{sl} / H$. Additional explanatory functions are selected to capture the possible dependencies of the bias in $\ln[\hat{\delta}(\mathbf{x})]$ on different factors characterizing the behavior of the column. We select $h_6(\mathbf{x}) = D_g / H$ to account for the possible effect of the aspect (slenderness) ratio. To capture the possible effect of the idealized elastic-perfectly plastic shear force $V_I = M_I / H$, we introduce $h_7(\mathbf{x}) = 4V_I /(\pi D_g^2 f_t')$, where $f_t' = 0.5\sqrt{f_c'}$ in MPa units is the tensile strength of concrete. To account for the possible influences of the confining transverse reinforcement and the core size, $h_8(\mathbf{x}) = \rho_s (f_{yh} / f_c')(D_c / D_g)$ are selected. To explore the effect of the longitudinal reinforcement, we choose $h_9(\mathbf{x}) = \rho_l f_y / f_c'$. Finally, to capture the effects of the material properties, $h_{10}(\mathbf{x}) = f_y / f_c'$ and $h_{11}(\mathbf{x}) = \varepsilon_{cu}$ are employed. Note that these explanatory functions are all dimensionless. As a result, the parameters $\boldsymbol{\theta}$ are also dimensionless. While additional explanatory functions or different forms of these functions could be selected, we believe that the selected ones are sufficiently broad to capture all the factors that may significantly influence the deformation capacity of the column.

4.3.3 Parameter Estimation

Having defined the deterministic model $\hat{\delta}(\mathbf{x})$ and the correction term $\gamma_\delta(\mathbf{x}, \boldsymbol{\theta})$, we are now ready to assess the probabilistic model in (4.1), i.e., estimate its parameters $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_{11}, \sigma)$ by the Bayesian updating formula described in Chapter 2. For this purpose, the statistical algorithms are programmed in Matlab (1999). Having no prior information on these parameters, we select a non-informative prior probability density function $p(\boldsymbol{\Theta})$. As described in Section 2.2, this implies a diffuse prior for $\boldsymbol{\theta}$ and a prior for σ that is proportional to σ^{-1}. Hence, we use $p(\boldsymbol{\Theta}) \propto \sigma^{-1}$. We note that given the large amount of observed data, any reasonable choice of the prior has practically no influence on the posterior estimates of the parameters.

In Chapter 3, a step-wise deletion procedure was described for reducing the number of terms in $\gamma_\delta(\mathbf{x},\boldsymbol{\theta})$ to achieve a compromise between model simplicity (few correction terms) and model accuracy (small σ). In essence, we eliminate each term $\theta_i h_i(\mathbf{x})$ when the coefficient of variation of θ_i is large in comparison to σ. Because of the logarithmic transformation in (4.1), σ is approximately equal to the c.o.v. of the predicted drift ratio. In general, the accuracy of the model is not expected to improve by including a term that has a c.o.v. much greater than σ.

Figure 4.5 summarizes the step-wise term deletion procedure for the deformation model. For each step, the figure shows the posterior c.o.v. of the model parameters θ_i (solid dots) and the posterior mean of the model standard deviation σ (open squares).
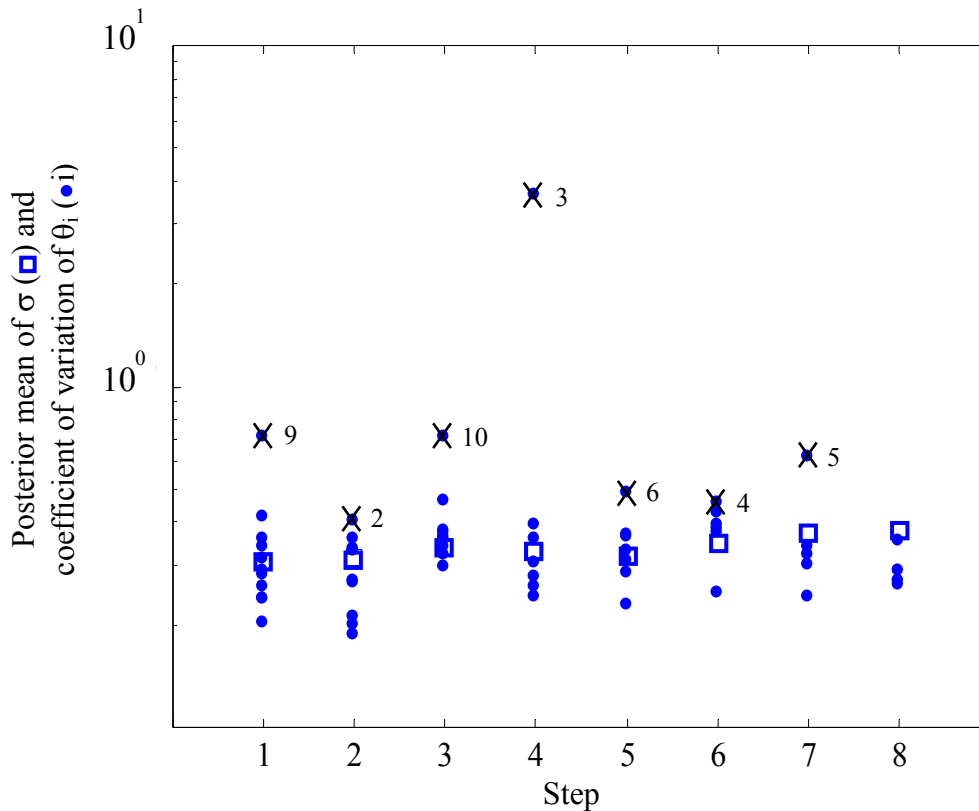
Figure 4.5. **Step-wise deletion process for the deformation capacity model. A superposed cross (×) indicates the term to be removed in the subsequent step.**

60

At Step 1 with the complete 12-parameter model, the posterior mean of σ is 0.306 and the parameter with the largest c.o.v. ($\cong 0.70$) is θ_9. To simplify the model, we drop the term $\theta_9 h_9(\mathbf{x})$. This is indicated by a cross symbol in Figure 4.5. In Step 2, we assess the reduced 11-parameter model. The posterior mean of σ now is 0.314, which indicates no appreciable deterioration of the model, and the parameter with the highest c.o.v. ($\cong 0.42$) now is θ_2. At the next step, we remove the term $\theta_2 h_2(\mathbf{x})$ and continue the same procedure. After 8 steps, we find the largest c.o.v. (for parameter θ_7) to be nearly equal to σ. This is an indication that further reduction may deteriorate the quality of the model. Stopping at this step, we are left with the terms $\theta_1 h_1(\mathbf{x})$, $\theta_7 h_7(\mathbf{x})$, $\theta_8 h_8(\mathbf{x})$ and $\theta_{11} h_{11}(\mathbf{x})$. At this stage the mean of σ is 0.379.

Analysis with the above reduced model reveals that θ_8 and θ_{11} are strongly correlated ($\rho = -0.85$) with the posterior statistics $\mu_{\theta_8} = 0.687$ and $\mu_{\theta_9} = 13.9$, respectively. Considering the definitions of $h_8(\mathbf{x})$ and $h_{11}(\mathbf{x})$ and the expression given in (4.10), a strong correlation between these parameters is not surprising. As a further simplification, by using the above posterior estimates of the 4-parameter model, θ_8 is expressed by its linear regression in θ_{11} (see (3.16)) as

$$\hat{\theta}_8 = -6.035 - 1.034 \theta_{11} \tag{4.13}$$

Thus, the reduced correction term takes the form

$$\gamma_\delta(\mathbf{x}, \boldsymbol{\theta}) = \theta_1 + \theta_7 \frac{4 V_I}{\pi D_g^2 f_t'} + \left(-6.035 + 1.034 \theta_{11}\right) \frac{\rho_s f_{yh} D_c}{f_c' D_g} + \theta_{11} \varepsilon_{cu} \tag{4.14}$$

with only three unknown parameters.

Table 4.2 lists the posterior statistics of the remaining parameters $\boldsymbol{\Theta} = (\theta_1, \theta_7, \theta_{11}, \sigma)$ of the reduced model.

Table 4.2.   Posterior statistics of the parameters in the deformation model.

Parameter	Mean	St. dev.	Correlation coefficient			
			θ_1	θ_7	θ_{11}	σ
θ_1	0.531	0.119	1.0			
θ_7	0.701	0.204	−0.37	1.0		
θ_{11}	−48.4	13.6	−0.59	−0.38	1.0	
σ	0.383	0.050	−0.04	0.14	0.20	1.0

The following observations derived from the parameter estimates in Table 4.2 are note-worthy: (a) The positive mean of θ_1 indicates that, independent of the variables $\mathbf{x}$, the determi-nistic model $\hat{\delta}(\mathbf{x})$ tends to underestimate the deformation capacity of the column. (b) The posi-tive estimate of θ_7 indicates that the deterministic model tends to underestimate the effect of the idealized shear force V_I (corresponding to M_I). This is expected in view of the conservative assumption regarding the concrete strain in determining M_I and ϕ_y, as described in Section 4.3.1. (c) The negative estimate of θ_{11} indicates that the deterministic model tends to underesti-mate the contribution of the transverse reinforcement and overestimate the contribution of the ultimate concrete strain.

Figure 4.6 shows a comparison between the measured and predicted values of the drift ra-tio capacities for the test columns based on the deterministic (top chart) and the probabilistic (bottom chart) models. For the probabilistic model, median predictions ($\varepsilon = 0$) are shown. The failure data are shown as solid dots and the censored data (lower bounds) are shown as open tri-angles. For a perfect model, the failure data should line up along the 1:1 dashed line and the cen-sored data should lie above it. The deterministic model on the top is strongly biased on the con-servative side, since most of the failure and many of the censored data lie below the 1:1 line. The probabilistic model on the bottom clearly corrects this bias. The dotted lines in the bottom figure delimit the region within one standard deviation of the model. We note that a majority of the

failure data points fall within the one-standard deviation limits and that most of the censored data are above the 1:1 line.

While the conservatism inherent in the deterministic deformation capacity model might be appropriate for a traditional design approach, for a performance-based design methodology, unbiased estimates of the capacity are essential. The constructed probabilistic model is unbiased and properly accounts for all the underlying uncertainties.
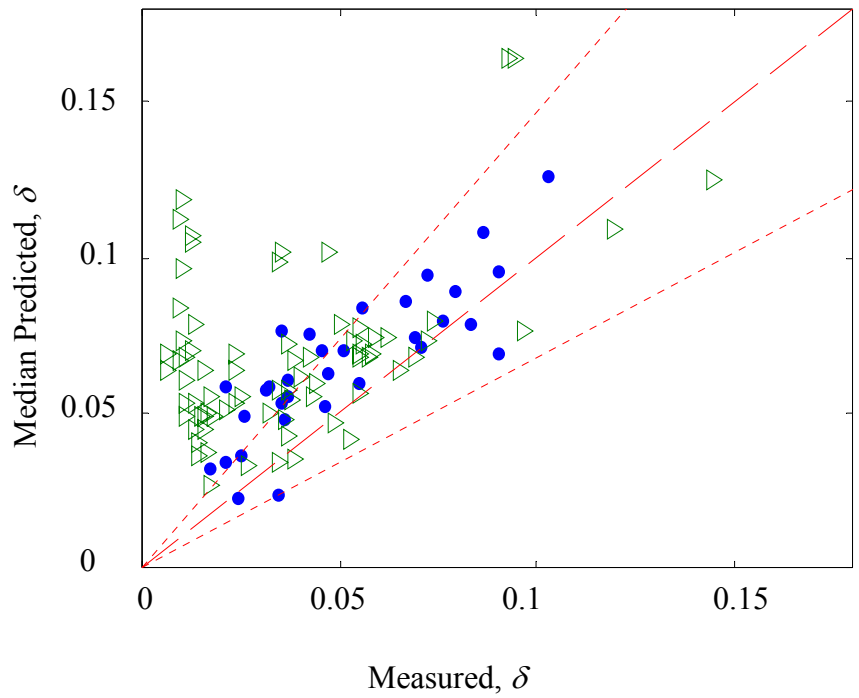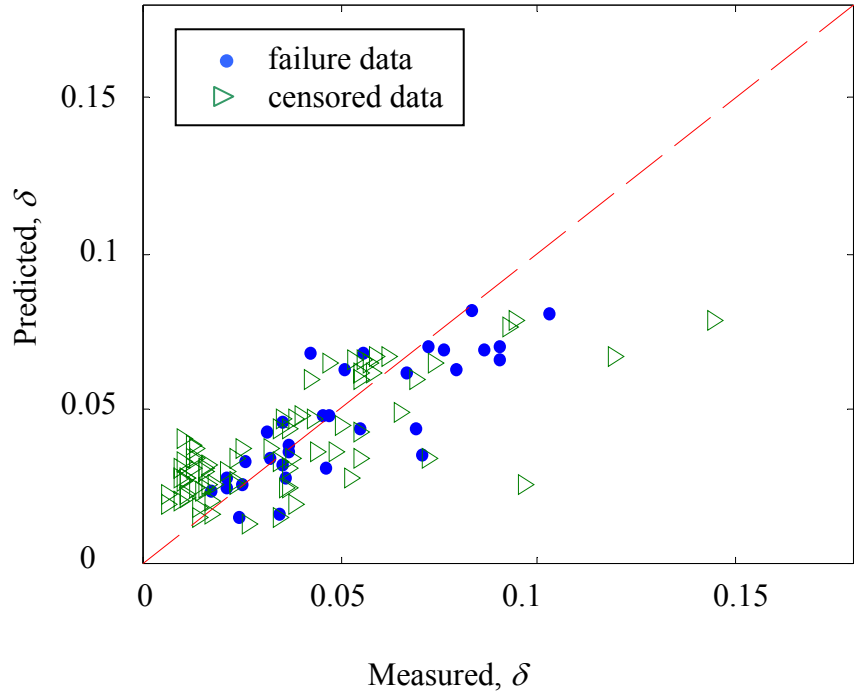
Figure 4.6. Comparison between measured and median predicted drift ratio capacities based on the deterministic (top) and probabilistic (bottom) models.
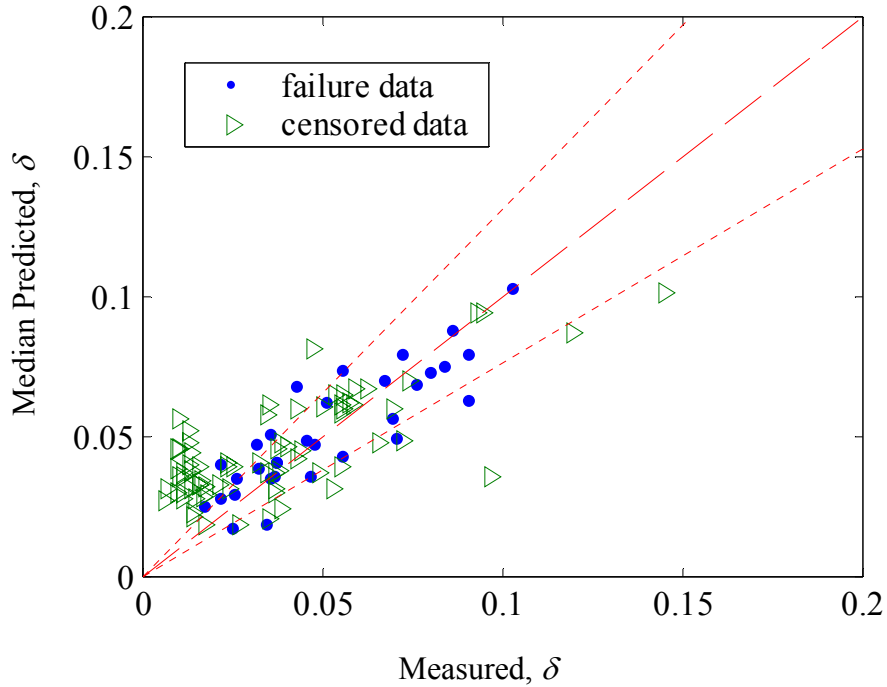
Figure 4.7. **Comparison between measured and median predicted drift ratio capacities based on the probabilistic model assessed with only failure data.**

Figure 4.7 explores the information content of the censored data by showing the median predicted drift ratio capacities versus the measured capacities for the probabilistic model assessed using only the failure data. We see that the predictions for the failure data alone are unbiased; however, there is no significant improvement on the prediction of the censored data compared to the deterministic model. This shows that the censored data have a relevant information content that is not included in the failure data. The region within one standard deviation of the model is now narrower ($\sigma = 0.272$) than in Figure 4.6 ($\sigma = 0.383$). This is because in combining the information content of both samples of observations, we have more data points but the same flexibility of the model (i.e., the number of free parameters) in fitting the data as before.

4.4   SHEAR CAPACITY MODELS

In this section we construct and compare two probabilistic shear capacity models for RC circular columns. For this purpose, the maximum lateral load measured in each experiment is classified as a "failure" datum if the tested column failed in shear, and as a lower-bound "censored" datum if the tested column failed in flexure or in a combined flexural-shear failure mode. In the database used for this analysis, 57 out of the 106 tested columns are in the latter category.

To develop a dimensionless model for the shear force capacity V, we consider the normalized quantity $v = V/(A_g \, f'_t)$, where A_g is the gross cross-sectional area and $f'_t = 0.5\sqrt{f'_c}$ is the tensile strength of concrete in MPa units. Owing to the non-negative nature of the shear capacity, we choose the logarithmic variance-stabilizing transformation to make the model homoskedastic. Thus, we adopt the model form

$$\ln[v(\mathbf{x},\mathbf{\Theta})] = \ln[\hat{v}(\mathbf{x})] + \gamma_v(\mathbf{x},\mathbf{\theta}) + \sigma\varepsilon \qquad (4.15)$$

where $\mathbf{\Theta} = (\mathbf{\theta},\sigma)$ is the set of unknown model parameters, $\hat{v}(\mathbf{x})$ is an existing deterministic model for predicting v, $\gamma_v(\mathbf{x},\mathbf{\theta})$ is a correction term for the bias inherent in the deterministic model, ε is a standard normal random variable and $\sigma\varepsilon$ represents the random component of the model error. Owing to the logarithmic transformation used, σ is approximately equal to the coefficient of variation of v. Furthermore, it is a measure of the quality of the model and can be used for selection among the competing models. The following sections describe the formulations of $\hat{v}(\mathbf{x})$ and $\gamma_v(\mathbf{x},\mathbf{\theta})$.

4.4.1   Deterministic Models

We consider and compare two predictive models for the shear capacity of RC columns. The first model was proposed by the ASCE-ACI Joint Task Committee 426 (1973) and is widely used in practice. The second model, proposed by Moehle *et al.* (1999, 2000), is a refinement of the FEMA 273 (1997) model.

The ASCE-ACI model is based on the well-known approach of considering the shear capacity as the sum of a contribution from the concrete, $\hat{V}_c$, and a contribution from the transverse reinforcement, $\hat{V}_s$, i.e.,

$$\hat{V} = \hat{V}_c + \hat{V}_s \tag{4.16}$$

where $\hat{V}$ denotes the deterministic prediction of the shear capacity (often denoted nominal shear capacity V_n). According to Priestly *et al.* (1996), in circular bridge columns, the contribution from the concrete is governed by the shear force required to initiate flexure-shear cracking and can be expressed as

$$\hat{V}_c = v_b A_e + \frac{M_d}{a} \tag{4.17}$$

where $v_b = (0.067 + 10\rho_t)\sqrt{f'_c} \leq 0.2\sqrt{f'_c}$ with units of MPa is the "basic" shear strength of concrete, in which $\rho_t = 0.5\rho_l$ is the longitudinal tension reinforcement ratio, A_e is the effective shear area taken as $0.8A_g$ for circular sections (Priestly *et al.*, 1996), $M_d = P I_g /(A_g y_t) = P D_g /8$ is the decompression moment with the axial load P and $y_t = D_g/2$, and $a = M/V$ is the shear span expressed as the ratio of the moment to shear at the critical section, which is taken to be equal to H for a cantilever column. The contribution from the transverse steel in (4.16) is based on the well-known truss analogy and is given by

$$\hat{V}_s = \frac{A_v f_{yh} D_e}{S} \tag{4.18}$$

where $A_v = 2A_h$ is the total area in a layer of the transverse reinforcement in the direction of the shear force, D_e is the effective depth commonly taken as $0.8D_g$ for circular cross sections, and S is the spacing of transverse reinforcement. Substituting (4.17) and (4.18) in (4.16), we have

$$\hat{V} = 0.8v_b A_g + 0.125\frac{P D_g}{H} + 1.6\frac{A_h f_{yh} D_g}{S} \tag{4.19}$$

The second deterministic model, proposed by Moehle *et al.* (1999, 2000), is a refinement of the FEMA 273 (1997) model. This model accounts for the reduction in the shear strength due to the effects of flexural stress and redistribution of internal forces as cracking develops. Here the

concrete contribution $\hat{V}_c$ is obtained by setting the principal tensile stress in the column equal to f_t'. Considering the stress transformation in Figure 4.8a, interaction between flexural and shear stresses, and strength degradation within the plastic hinge, the final form of this model is

$$\hat{V}_c = k\left(\frac{f_t'}{a/D_e}\sqrt{1+\frac{P}{f_t'A_g}}\right)A_e \qquad (4.20)$$

where the aspect ratio a/D_e is limited to the range 1.7–3.9, (Moehle *et al.*, 2000) and k is a factor included to account for the strength degradation within the plastic hinge region as a function of the displacement ductility $\mu_\Delta = \Delta/\hat{\Delta}_y = \delta/\hat{\delta}_y$, as defined in Figure 4.8b.

Two alternatives are considered for the drift ratio to be used in the expression for μ_Δ. The first is simply $\delta = \hat{\delta}(\mathbf{x})$, the deterministic deformation model described in Section 4.3.1. The second is the median of the probabilistic deformation capacity model $\delta = \hat{\delta}(\mathbf{x})\exp[\gamma_\delta(\mathbf{x}, \mathbf{M_\Theta})]$, where the parameters are fixed at the posterior mean estimates, $\mathbf{M_\Theta}$. Note that because the probabilistic deformation model was developed making the normality assumption on $\ln(\delta)$, its mean (obtained by setting $\varepsilon = 0$ and fixing the parameters at their posterior mean estimates $\mathbf{M_\Theta}$) corresponds to the median of δ.
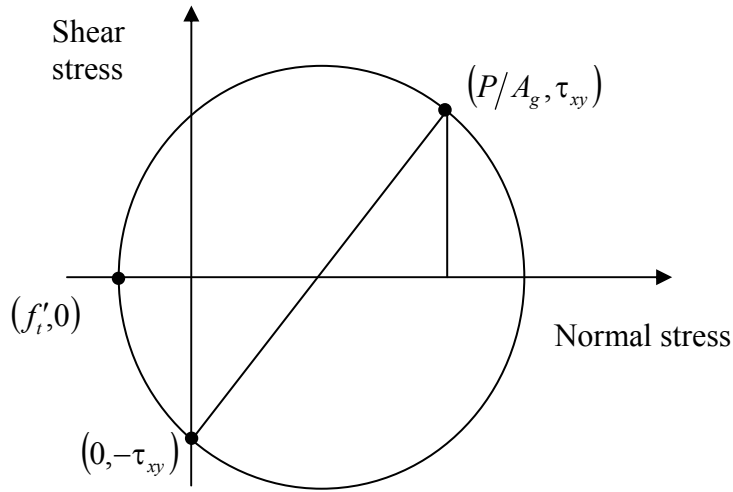
(a) Stress state in the column at shear failure.
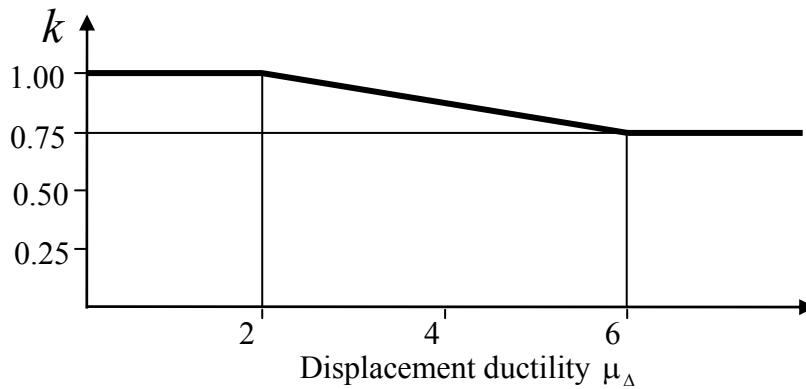
(b) Degradation of V_c with displacement ductility.

Figure 4.8.   Shear failure model by Moehle *et al.* (1999, 2000)

In the two papers by Moehle *et al.* (1999, 2000), different approaches are proposed to in-
troduce the effect of strength degradation on, $\hat{V}_s$. In Moehle *et al.* (1999), a reduction factor of
0.5 is applied to $\hat{V}_s$, whereas in Moehle *et al.* (2000) $\hat{V}_s$ is reduced by the factor k. In this study,
although we do not introduce any reduction factor for $\hat{V}_s$ in the deterministic model, we do as-

sess possible modification of $\hat{V}_s$ due to strength degradation through a properly selected explanatory function in the model correction term, as discussed in the following section.

4.4.2   Model Correction

As in the deformation capacity model, the model correction term $\gamma_v(\mathbf{x}, \boldsymbol{\theta})$ in (4.15) is intended to capture the inherent bias in $\ln[\hat{v}(\mathbf{x})]$. As described in Chapter 3, we use the linear form

$$\gamma_v(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{p} \theta_i h_i(\mathbf{x}) \tag{4.21}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ is a vector of unknown parameters of the model and $h_1(\mathbf{x}), \ldots, h_p(\mathbf{x})$ are selected "explanatory" functions. To capture a potential constant bias in $\ln[\hat{v}(\mathbf{x})]$, we select $h_1(\mathbf{x}) = 1$. To account for a possible correction in the contribution of the longitudinal column reinforcement we select $h_2(\mathbf{x}) = \rho_l$, and to account for any correction in the effect of the axial load we select $h_3(\mathbf{x}) = PD_g / A_g f_t' H$. Furthermore, to account for any needed modification in the contribution from the transverse reinforcement and, in particular, to investigate the effects of strength degradation on $\hat{V}_s$, we select $h_4(\mathbf{x}) = A_v f_{yh} D_g / (A_g f_t' S)$. Note that the explanatory functions are again dimensionless, making the parameters θ_i also dimensionless. While additional or different explanatory functions could be selected, we believe that the above functions capture the most significant factors that may influence the shear capacity of RC columns.

4.4.3   Parameter Estimation

We are now prepared to estimate the parameters $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_4, \sigma)$ for the shear capacity model by using Bayesian inference. Owing to the lack of prior information, the non-informative prior $p(\boldsymbol{\Theta}) \propto \sigma^{-1}$ is selected.

Following the step-wise deletion procedure described in Chapter 3 and in Section 4.3.2 for the deformation capacity model, we use the posterior statistics of the parameters to detect the superfluous explanatory functions, which are then dropped to simplify each model. The reduced

form of $\gamma_v(\mathbf{x}, \boldsymbol{\theta})$ is different for the two deterministic shear capacity models. Furthermore, the parameter estimates depend on whether $\delta = \hat{\delta}(\mathbf{x})$ or $\delta = \hat{\delta}(\mathbf{x}) \exp[\gamma_\delta(\mathbf{x}, \mathbf{M_\Theta})]$ is used to compute the factor k. Table 4.3 lists the expression of $\gamma_v(\mathbf{x}, \boldsymbol{\theta})$ and the posterior mean and standard deviation of σ for each model, computed according to Section 2.2.4.

As mentioned earlier, a measure of the predictive accuracy of each model is the posterior mean estimate of σ. The last two columns of Table 4.3 show that the deformation-dependent shear capacity model with $\delta = \hat{\delta}(\mathbf{x}) \exp[\gamma_\delta(\mathbf{x}; \mathbf{M_\Theta})]$ used in computing the factor k has the smallest error standard deviation and, therefore, is the most accurate model. In the remainder of this chapter, we present results only for this model.

Table 4.3. **Reduced model correction terms and posterior means and standard deviations of σ for the selected shear models**.

Deterministic model	$\gamma_v(\mathbf{x}; \boldsymbol{\theta})$	Mean of σ	St. dev. of σ
ASCE-ACI 426 model	$\theta_1 + \theta_2 h_2$	0.189	0.019
Deformation-dependent model with $\delta = \hat{\delta}(\mathbf{x})$ used in computing k	$\theta_2 h_2 + \theta_4 h_4$	0.179	0.019
Deformation-dependent model with $\delta = \hat{\delta}(\mathbf{x}) \exp[\gamma_\delta(\mathbf{x}; \mathbf{M_\Theta})]$ used in computing k	$\theta_2 h_2 + \theta_4 h_4$	0.153	0.013

Figure 4.9 summarizes the step-wise deletion process for the selected shear capacity model. For each step, the figure shows the c.o.v. of the model parameters θ_i (solid dots) and the mean of the model standard deviation σ (open square). At Step 1 with the complete 5-parameter model, the c.o.v. of θ_3 is 0.340 and the mean of σ is 0.136. To simplify the model, we drop the term $\theta_3 h_3(\mathbf{x})$ from the model correction term. This is indicated by a cross symbol at Step 1 in Figure 4.9. In Step 2 for the reduced 4-parameter model, the mean of σ is 0.144, which indicates

an insignificant deterioration of the model. The parameter with the highest c.o.v. (=0.348) now is θ_1. We remove the term $\theta_1 h_1(\mathbf{x})$ and continue the same procedure. At Step 3 we find the largest c.o.v. (for parameter θ_4) to be of the same order of magnitude as the mean of σ, which indicates that further simplification is not justified. Thus, the reduced model correction term is

$$\gamma_v(\mathbf{x},\boldsymbol{\theta}) = \theta_2 \rho_l + \theta_4 A_v f_{yh} D_g / A_g f_t' S \tag{4.22}$$

Table 4.4 lists the posterior statistics of the parameters θ_2, θ_4, and σ for the reduced model.
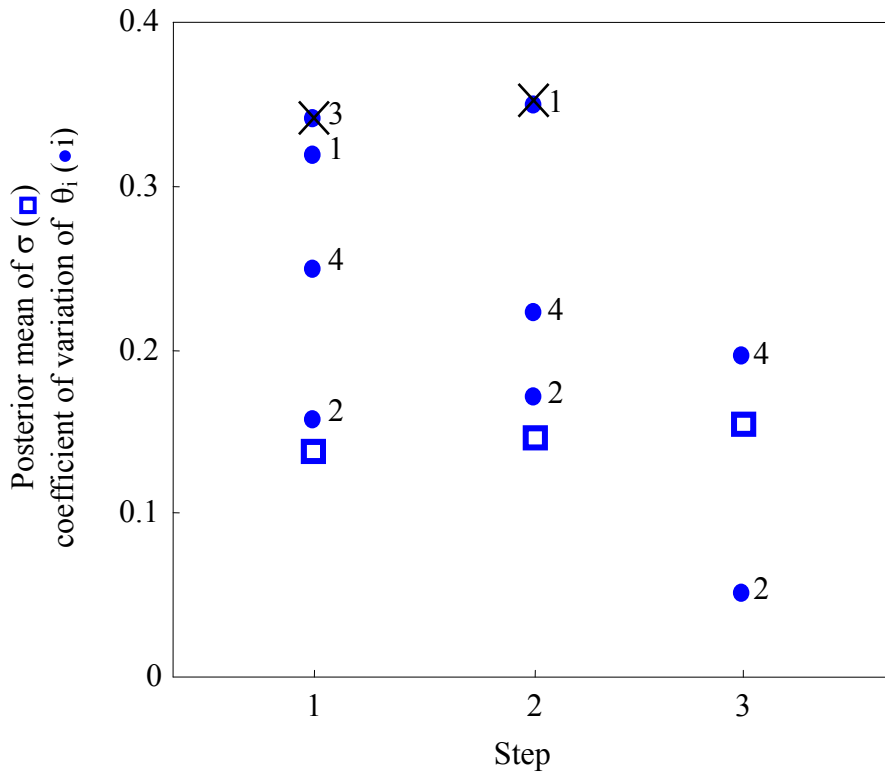
Figure 4.9. **Step-wise deletion process for the shear capacity model. A superposed cross (×) indicates the term to be removed in the subsequent step.**

Table 4.4.   Posterior statistics of the parameters in the selected shear model.

| Parameter	Mean	St. dev.	Correlation coefficient		
			θ_2	θ_4	σ
θ_2	23.1	1.2	1		
θ_4	−0.614	0.120	−0.87	1	
σ	0.153	0.013	−0.13	0.06	1

The following noteworthy observations can be made from the preceding results: (a) The fact that the explanatory function $h_1 = 1$ is not informative suggests that there is no constant bias in the deterministic model. (b) The presence of $h_2(\mathbf{x}) = \rho_l$ with a positive coefficient in (4.22), equal to 23.1, is an indication that the contribution of the longitudinal reinforcement to the shear capacity is underestimated in the deterministic model. This could be because the deformation-dependent model was calibrated using rectangular column data (Moehle *et al.*, 1999), for which the contribution of longitudinal reinforcement is known to be less important than that for circular columns, where longitudinal reinforcement is uniformly distributed around the circumference. Interestingly, for the ASCE-ACI 426 model, which includes a term representing the contribution of ρ_l, the posterior mean of θ_2 is 4.43, which is far smaller than that for the selected model, i.e., 23.1. (c) The fact that the explanatory function $h_3(\mathbf{x}) = PD_g / A_g f_t'H$ appears not informative is an indication that the effect of the axial force is accurately accounted for in the deterministic model through the transformation of stresses by the Mohr's circle (Figure 4.8a). (d) The presence of the explanatory function $h_4(\mathbf{x}) = A_v f_{yh} D_g / A_g f_t'S$ with a negative coefficient in (4.22) might represent the effect of strength degradation in the contribution from the transverse steel that is needed for a more accurate prediction. However, we should note that there are substantial difficulties that arise in interpreting the numerical values of empirical regression coefficients in case of high positive or negative correlation between the parameters. Owing to the high negative correlation between θ_2 and θ_4, observations (b) and (d) suggest the need for further experimental investigations.

Figure 4.10 shows a comparison between the measured and predicted values of the normalized shear capacities for the test columns based on the deterministic (top chart) and probabilistic (bottom chart) deformation-dependent shear capacity models. The same definitions as in Figure 4.6 apply. It is seen that the deterministic model on the top is strongly biased on the conservative side. The probabilistic model on the bottom clearly corrects this bias. The dotted lines in this figure delimit the region within one standard deviation of the model.
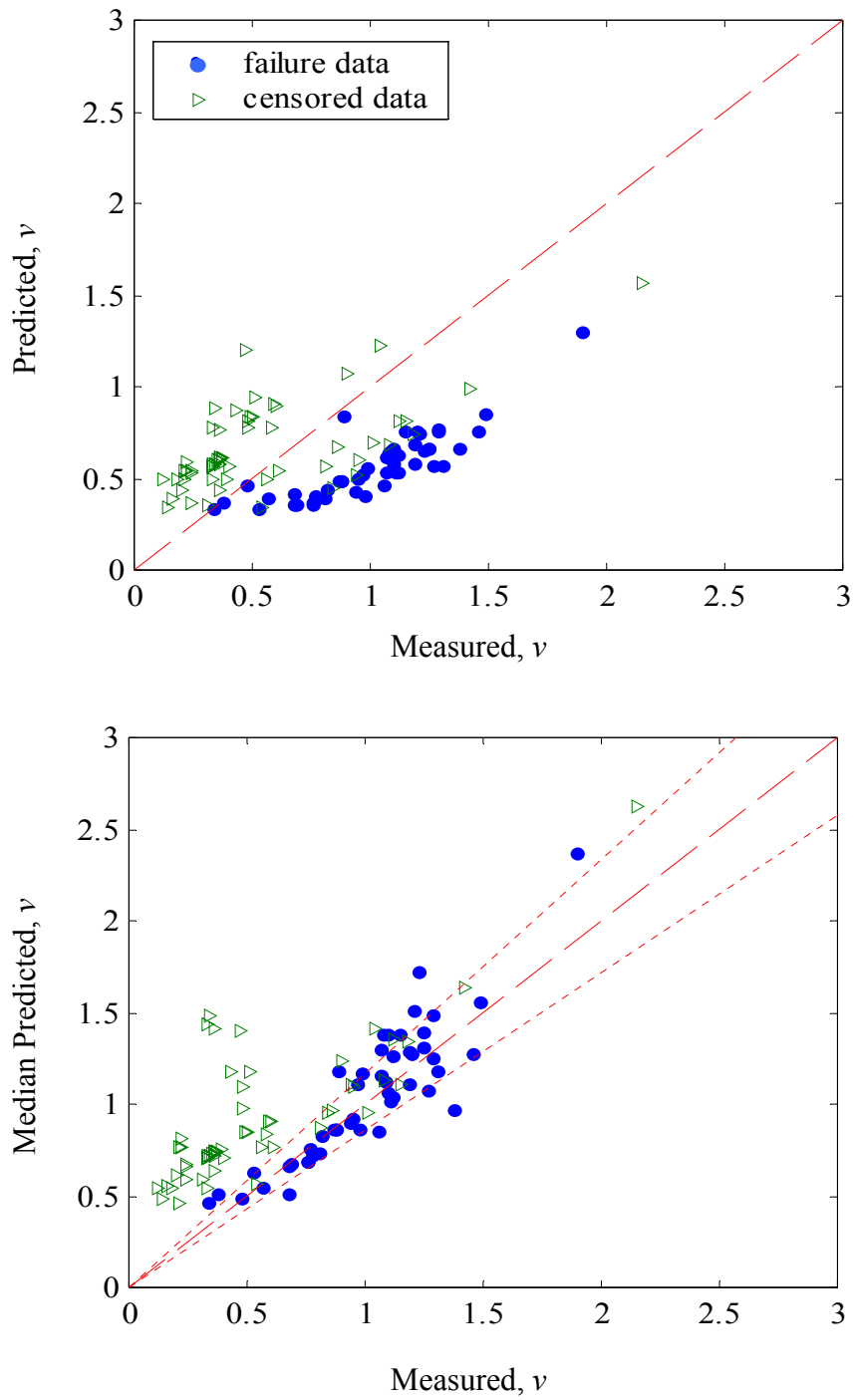
Figure 4.10.   Comparison between measured and median predicted shear capacities based on the deterministic (top) and probabilistic (bottom) models.

As stated for the deformation capacity, while the conservatism inherent in the mechanical model might be appropriate for a traditional design approach, for a performance-based design methodology unbiased estimates of the capacity are essential. As for the deformation model, the Bayesian estimates for the shear model are unbiased and properly account for the underlying uncertainties.

Figure 4.11 explores the information content of the censored data by showing the median predicted shear capacities versus the measured capacities for the probabilistic model assessed using only the failure data. The predictions for the failure data now are unbiased, but a significant improvement can be seen in the prediction of the censored data compared to the deterministic model in Figure 4.10 (top). This improvement should not be surprising. As we can see in Figure 4.10 (top) the censored data lie above an imaginary line going through the failure data, so correction for the bias for the failure data also corrects for bias in the censored data. It appears that, in this case, the censored data do not provide a significant information content that is not already included in the failure data. For the same reason the regions within one standard deviation of the two models are essentially the same in Figures 4.10 (bottom) ($\sigma = 0.153$) and 4.11 ($\sigma = 0.152$).
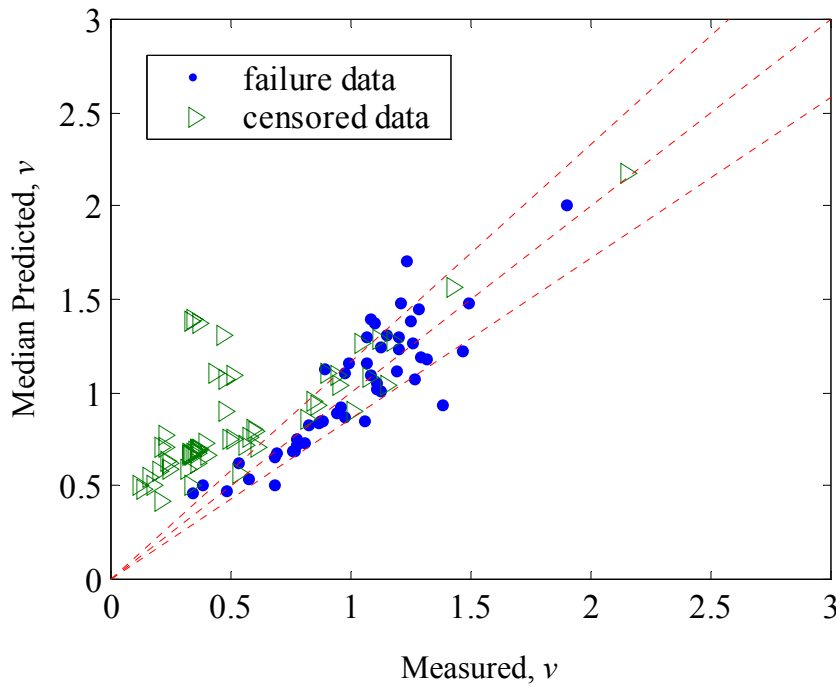
Figure 4.11. Comparison between measured and median predicted shear capacities based on the probabilistic model assessed with only failure data.

4.5   BI-VARIATE DEFORMATION–SHEAR CAPACITY MODEL

In this section a bi-variate deformation-shear capacity model is constructed that accounts for the correlation between the two models. The subscripts δ and v are used to indicate quantities related to the deformation and shear capacity models, respectively. By using (4.1) and (4.15), the bi-variate capacity model is written as

$$\ln[\delta(\mathbf{x},\boldsymbol{\theta}_\delta,\sigma_\delta)] = \ln[\hat{\delta}(\mathbf{x})] + \gamma_\delta(\mathbf{x},\boldsymbol{\theta}_\delta) + \sigma_\delta\varepsilon_\delta \qquad (4.23a)$$

$$\ln[v(\mathbf{x},\boldsymbol{\theta}_v,\sigma_v)] = \ln[\hat{v}(\mathbf{x})] + \gamma_v(\mathbf{x},\boldsymbol{\theta}_v) + \sigma_v\varepsilon_v \qquad (4.23b)$$

where $\delta(\mathbf{x},\boldsymbol{\theta}_\delta,\sigma_\delta)$, $\hat{\delta}(\mathbf{x})$, $\gamma_\delta(\mathbf{x},\boldsymbol{\theta}_\delta)$, $v(\mathbf{x},\boldsymbol{\theta}_v,\sigma_v)$, $\hat{v}(\mathbf{x})$, and $\gamma_v(\mathbf{x},\boldsymbol{\theta}_v)$ are as defined in Sections 4.3 and 4.4. Also let ρ be the unknown correlation coefficient between the model errors ε_δ and ε_v. The unknown model parameters $\boldsymbol{\Theta} = (\boldsymbol{\theta}_\delta,\sigma_\delta,\boldsymbol{\theta}_v,\sigma_v,\rho)$ are estimated by Bayesian inference. Having already determined the informative explanatory functions for each model, the reduced model correction terms in (4.14) and (4.22) are used. Owing to the lack of prior information, we select the non-informative prior $p(\boldsymbol{\Theta}) \propto (1-\rho^2)^{-3/2}/(\sigma_\delta\sigma_v)$ (see (2.42) specialized for $q=2$). Table 4.5 shows the posterior statistics of the parameters $\boldsymbol{\Theta}$. As expected, the estimates of $\boldsymbol{\theta}_\delta$, $\boldsymbol{\theta}_v$, σ_δ and σ_v are nearly the same as the estimates based on the marginal models. The negative sign of the posterior mean of the correlation coefficient shows that the deformation and shear capacities are negatively correlated. This indicates that, relative to their median values, a column with high deformation capacity is likely to have a low shear capacity, and vice versa.

Table 4.5.   Posterior statistics of the parameters in the bi-variate deformation–shear model.

	$\theta_{\delta,1}$	$\theta_{\delta,7}$	$\theta_{\delta,11}$	σ_δ	$\theta_{v,2}$	$\theta_{v,4}$	σ_v	ρ
Mean	0.512	0.828	−50.8	0.383	43.2	−0.584	0.189	−0.535
St.dev.	0.0537	0.198	11.5	0.0485	2.73	0.180	0.0187	0.166

Correlation coefficients

	$\theta_{\delta,1}$	$\theta_{\delta,7}$	$\theta_{\delta,11}$	σ_δ	$\theta_{v,2}$	$\theta_{v,4}$	σ_v
$\theta_{\delta,7}$	−0.38						
$\theta_{\delta,11}$	−0.60	−0.51					
σ_δ	−0.41	0.36	0.065				
$\theta_{v,2}$	0.021	−0.10	0.066	−0.058			
$\theta_{v,4}$	0.12	0.0019	−0.11	−0.0077	−0.84		
σ_v	0.059	0.045	−0.095	0.061	0.017	0.031	
ρ	0.14	−0.29	0.12	−0.26	0.12	0.006	−0.15

4.6   SUMMARY

Uni-variate and bi-variate probabilistic models for the deformation and shear capacity of RC circular columns subjected to cyclic loading are developed by a Bayesian approach using existing experimental data for cyclically tested columns. The models are unbiased and explicitly account for all the relevant uncertainties, including errors arising from an inaccurate model form or missing variables, and statistical uncertainty. With the aim of facilitating their use in practice, the models are constructed by developing correction terms to existing deterministic models. Through a model selection process that makes use of a set of "explanatory" functions, the terms that effec-

tively correct the bias in the existing model are identified and insight into the underlying behavioral phenomena is gained. However, we should note that there are substantial difficulties that arise in interpreting the numerical values of empirical regression coefficients in case of high positive or negative correlation between the parameters. In this case, further experimental investigations may be necessary to gain insight into the physical phenomenon.

5 Probabilistic Demand Models

5.1 INTRODUCTION

Predictive demand models in current structural engineering practice are deterministic. Typically, they provide biased estimates (i.e., the mean over many trials is different from the true mean) and do not account for uncertainties inherent in the modeling process. The advent of a performance-based design approach requires unbiased predictive capacity and demand models that explicitly account for all the relevant uncertainties. This chapter presents a comprehensive Bayesian methodology to construct probabilistic demand models that explicitly account for all the relevant uncertainties, including model errors arising from an inaccurate model form or missing variables, measurement errors and statistical uncertainty.

Similar to the formulation of the capacity models (Chapter 3), the probabilistic demand models are developed by use of deterministic demand models or procedures used in practice, with additional terms that explicitly describe the inherent systematic and random errors. In contrast to the deterministic capacity models, the deterministic demand models are not analytical expressions. Rather, they are "procedural" models. Two possible practical alternatives are considered: (a) Pushover analysis following a prescribed set of rules (e.g., how to select the magnitude and distribution of the equivalent loads, what material models to use to approximately account for cyclic effects, and how to account for vertical accelerations) followed by nonlinear response spectrum analysis; (b) nonlinear time-history dynamic analysis using an "equivalent" SDOF hysteretic model, again following prescribed rules (e.g., how to determine the equivalent mass, stiffness, damping, and hysteretic properties). Of course these estimates of the demand are

typically biased and in error. After some initial investigations, option (a) was selected as the "deterministic model" for probabilistic analysis.

Through the use of a set of "explanatory" functions, we are able to identify terms in the selected model that are significant in correcting the bias in the existing deterministic model. Moreover, these functions provide means to gain insight into the underlying behavioral phenomena and to select ground motion parameters that are most relevant to the seismic demand of interest.

5.2 PROBABILISTIC DEMAND MODELS FOR COMPONENTS

A demand "model" is a mathematical expression relating the structural demands at the component level, e.g., shear and deformation demands of each column of a bridge bent, to the demand at the system level, i.e., measures of intensity of the earthquake ground motion. The main purpose of the model is to provide a means for predicting the demand on each component for given deterministic or random values of a set of basic variables $\mathbf{x} = (x_1, x_2, \ldots)$ representing, e.g., material property constants, member dimensions, imposed boundary conditions and measures of intensity of the earthquake ground motion. As asserted in Chapters 3 and 4, for practical implementation, it is desirable that the model form be based on an existing deterministic model. With this in mind, the following form is adopted for a q-dimensional demand model

$$D_k(\mathbf{x}, \boldsymbol{\theta}_k, \boldsymbol{\Sigma}) = \hat{d}_k(\mathbf{x}) + \gamma_k(\mathbf{x}, \boldsymbol{\theta}_k) + \sigma_k \varepsilon_k, \qquad k = 1, \ldots, q \qquad (5.1)$$

In the above expression, D_k is the k-th demand measure (or a suitable transformation of it, see Section 3.2), $\hat{d}_k(\mathbf{x})$ is the selected deterministic demand model (or the corresponding transformation), and $\gamma_k(\mathbf{x}, \boldsymbol{\theta}_k)$ and $\sigma_k \varepsilon_k$ are terms that respectively correct for the bias and the random error in $\hat{d}_k(\mathbf{x})$, in which ε_k is a standard normal random variable and σ_k is the standard deviation of the model error. The above additive model correction form is valid under the following assumptions: (a) the model standard deviation is independent of $\mathbf{x}$ (homoskedasticity assumption) and (b) the model error has the normal distribution (normality assumption).

Employing a suitable transformation of each demand measure approximately satisfies these assumptions (see Section 3.2). In order to explore the sources of bias in the deterministic model, the bias correction term is written in the form

$$\gamma_k(\mathbf{x}, \boldsymbol{\theta}_k) = \sum_{j=1}^{p_k} \theta_{kj} h_{kj}(\mathbf{x}), \qquad k = 1, \ldots, q \qquad (5.2)$$

where θ_{kj} and $h_{kj}(\mathbf{x})$, $j = 1, \ldots, p_k$, are a set of p_k model parameters and "explanatory" basis functions, respectively, for the k-th demand model. By examining the posterior statistics of the unknown parameters θ_{kj}, we are able to identify those explanatory functions that are significant in describing the bias in the deterministic model. Ideally, the "explanatory" basis functions should be selected based on first principles, i.e., laws of structural dynamics, making use of the most appropriate parameters that are essential for describing the important characteristics in a compact form.

The eventual goal of developing probabilistic demand models is to construct fragility estimates for structural systems. In this context, it is essential to select variables for constructing the demand models that reduce the overall aleatory uncertainty both in the demand models and in the predictive relationships that are used to estimate the variables within the model, e.g., an attenuation relationship for computing peak ground acceleration or spectral ordinate for an earthquake with given magnitude and location. Selection of variables that are strongly correlated with the component demand but that are difficult to predict for a given earthquake are useless, since all the uncertainty would be lumped into the predictive relationship. Similarly, selection of variables that can be predicted easily but that are not significant in estimating the component demands would shift all the uncertainty to the demand models. The selection of the model variables should be based on our engineering judgment guided by statistical analyses.

Finally, the random error terms ε_k, $k = 1, \ldots, q$, for the different models in general can be correlated. Let $\boldsymbol{\Sigma}$ denote the covariance matrix of the variables $\sigma_k \varepsilon_k$, $k = 1, \ldots, q$. The set of unknown parameters of the model in (5.1) then is $\boldsymbol{\Theta} = (\boldsymbol{\theta}, \boldsymbol{\Sigma})$, where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_q)$ and $\boldsymbol{\theta}_k = (\theta_{k1}, \ldots, \theta_{kp_k})$. Considering symmetry, $\boldsymbol{\Sigma}$ includes q unknown variances σ_k^2, $k = 1, \ldots, q$, and $q(q-1)/2$ unknown correlation coefficients ρ_{kl}, $k = 1, \ldots, q-1$, $l = k+1, \ldots, q$.

Having defined the deterministic models $\hat{d}_k(\mathbf{x})$ and explanatory functions $h_{kj}(\mathbf{x})$, we assess the probabilistic model in (5.1) by estimating its parameters $\boldsymbol{\Theta}$ through Bayesian inference as described in Chapter 2.

5.3 PROBABILISTIC DEMAND MODELS FOR SYSTEMS

A structural system may consist of several structural components and, as seen in Section 5.2, each structural component may have different demands (e.g., relative to failure in shear, bending, reinforcing bar pullout or buckling, or excessive deformation). For the analysis of such a system consisting of s components each having q different demands, we formulate a $s \times q$-dimensional multi-variate demand model in the form

$$D_{ik}\left(\mathbf{x}_i, \boldsymbol{\theta}_k, \boldsymbol{\Sigma}\right) = \hat{d}_{ik}\left(\mathbf{x}_i\right) + \gamma_{ik}\left(\mathbf{x}_i, \boldsymbol{\theta}_k\right) + \sigma_k \varepsilon_{ik}, \quad i = 1, \ldots, s \quad k = 1, \ldots, q \qquad (5.3)$$

where

$$\gamma_{ik}\left(\mathbf{x}_i, \boldsymbol{\theta}_k\right) = \sum_{j=1}^{p_k} \theta_{kj} h_{kj}\left(\mathbf{x}_i\right), \qquad i = 1, \ldots, s \quad k = 1, \ldots, q \qquad (5.4)$$

and σ_k represents the standard deviation of the model error for the k-th demand, which is assumed to be the same for different components. This notation assumes that component demand modes are an ordered set, so that the k-th demands for all components correspond to the same mode.

All entries in the above expressions have definitions analogous to those of the q-dimensional demand model presented in the previous section. The term $\boldsymbol{\Sigma}$ denotes the covariance matrix of the variables $\sigma_k \varepsilon_{ik}$, $i = 1, \ldots, s$, $k = 1, \ldots, q$, and it accounts for the possible correlation between the demands on various components. Let $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j)$ be the correlation coefficient between demand k of component i and demand l of component j, then the equality $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j) = \rho_{lk}(\mathbf{x}_j, \mathbf{x}_i)$ holds for any structural configuration. In the particular case of symmetric structural systems (in terms of geometry and material properties), the equality $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j) = \rho_{lk}(\mathbf{x}_i, \mathbf{x}_j)$ also holds. In the following analysis we assume that this equality is

approximately true for non-symmetric structures. Hence, $\boldsymbol{\Sigma}$ includes q unknown variances σ_k^2, $k = 1, \ldots, q$, and, considering the symmetry assumption, $s \times q (s + s \times q + q - 3)/4$ unknown correlation coefficients $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j)$.

In order to explore the sources of correlation in the multi-variate model, we select a suitable set of p_R "explanatory" basis functions $h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)$, $w = 1, \ldots, p_R$, $i, j = 1, \ldots, s$, and express the correlation coefficient between demand k of component i and demand l of component j in the form

$$\rho_{kl}\left(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}_{kl}\right) = \frac{\sum\limits_{w=1}^{p_R} \theta_{klw} h_{Rw}\left(\mathbf{x}_i, \mathbf{x}_j\right)}{1 + \left|\sum\limits_{w=1}^{p_R} \theta_{klw} h_{Rw}\left(\mathbf{x}_i, \mathbf{x}_j\right)\right|}, \quad k, l = 1, \ldots, q \ \ i, j = 1, \ldots, s \tag{5.5}$$

This form was selected because, while $\sum\limits_{w=1}^{p_R} \theta_{klw} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)$ can be greater than 1 or smaller than -1, the resulting model for the correlation coefficient, $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta})$, has the property of ranging over $[-1, 1]$. Also note that the dependence of the correlation coefficients on the component properties (e.g., geometry and material properties) is lumped in the known explanatory functions, while the unknown parameters $\boldsymbol{\theta}_{kl} = (\theta_{klw}, w = 1, \ldots, p_R)$ vary with the demands, k and l, under consideration. This formulation is convenient because, since $\boldsymbol{\theta}_{kl}$ are independent from the geometry, once they have been estimated, they can be used to construct the correlation coefficients for any other structure with given geometry. By examining the posterior statistics of the unknown parameters θ_{klw}, we are able to identify those explanatory functions that are significant in describing the correlation between the model error terms ε_{ik}, $i = 1, \ldots, s$, $k = 1, \ldots, q$. The set of unknown parameters of the multi-variate model is then $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_q, \boldsymbol{\sigma}, \boldsymbol{\theta}_{11}, \ldots, \boldsymbol{\theta}_{1q}, \ldots, \boldsymbol{\theta}_{qq})$, where $\boldsymbol{\theta}_k = (\theta_{kj}, \quad j = 1, \ldots, p_k)$, $k = 1, \ldots, q$, $\boldsymbol{\sigma} = (\sigma_k, k = 1, \ldots, q)$ and $\boldsymbol{\theta}_{kl} = (\theta_{klw}, w = 1, \ldots, p_R)$. Applications in Chapter 6 demonstrate specific formulations of the probabilistic demand models presented here.

6   Applications of Probabilistic Demand Models

6.1   INTRODUCTION

In Chapter 5 a Bayesian framework was developed for the formulation of demand models for structural components and systems. In this chapter probabilistic deformation and shear demand models for RC bridge bents and bridge systems are developed by use of existing observational or simulated data. Explicit account of all the prevailing aleatory and epistemic uncertainties is made.

As described in Chapter 5, the probabilistic models are akin to deterministic demand models or procedures used in practice, but they have additional terms that explicitly describe the inherent systematic and random errors. Through the use of a set of "explanatory" functions, we are able to identify terms in the model that are significant in correcting the bias in the existing deterministic model. Moreover, these functions provide means to gain insight into the underlying behavioral phenomena and to select ground motion parameters that are most relevant to the seismic deformation and shear demands.

6.2   THE PROBLEM

The structure of interest is a general RC highway bridge with single-column bents (Figure 6.1). For this purpose, we want to construct probabilistic models to predict the deformation demand,

$D_{i\delta}$, and the shear demand, D_{iv}, on each column i, $i = 1,\ldots,s$, of the bridge during an earthquake event, where s denotes the number of bents.
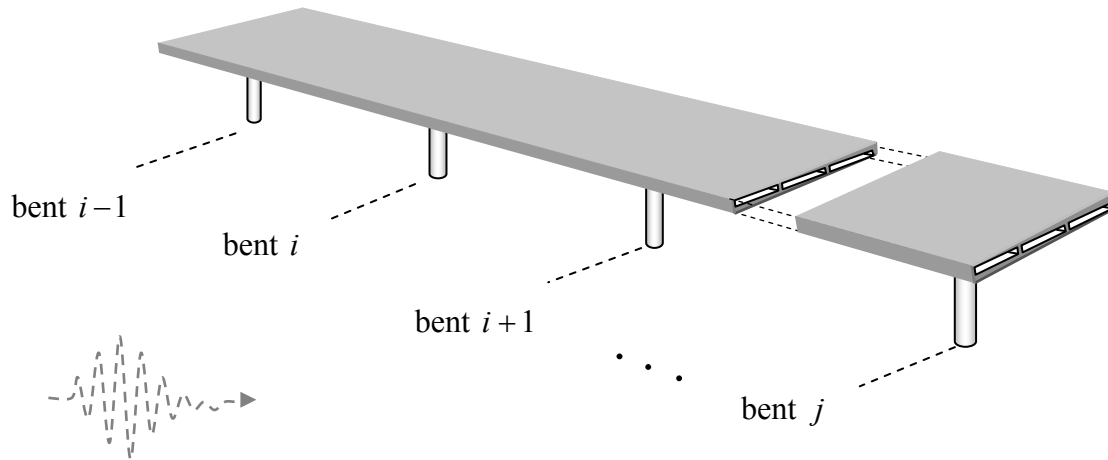
Figure 6.1.   RC highway bridge with single-column bents.

Figure 6.2 shows the quantities of interest for the highway bridge in Figure 6.1. In general, the demands D_{ik} and D_{jl}, for failure modes $k,l = v,\delta$ on columns $i,j = 1,\ldots,s$, are correlated due to the presence of the deck. The correlation coefficients $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j)$ are unknown and need to be estimated along with the model parameters.
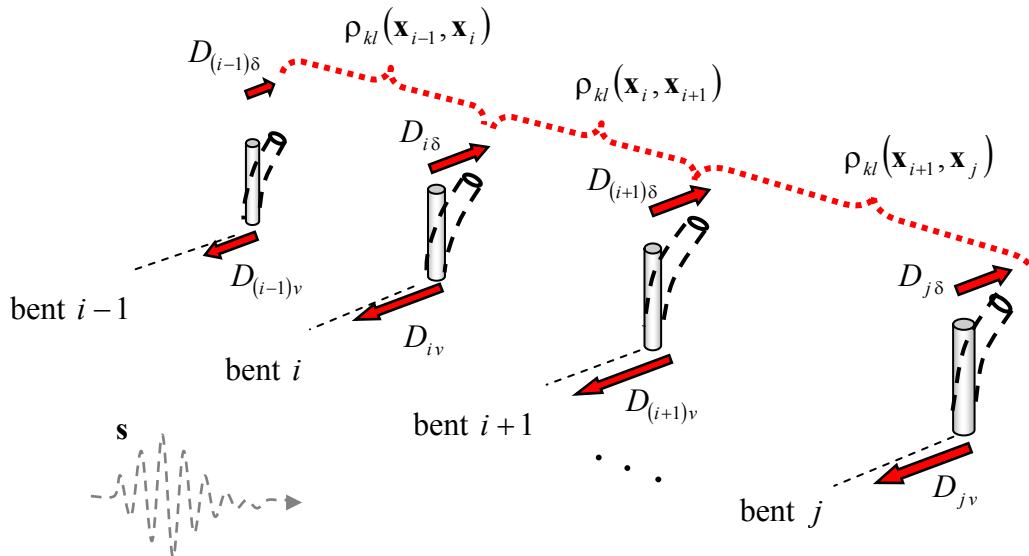
Figure 6.2.   Representation of unknown quantities to be estimated.

The structure in Figure 6.1 can be idealized as a series system of bents with correlated dynamic responses. In common engineering practice, each bent is idealized as a *single-degree-*

of-freedom (SDOF) system with nonlinear force-displacement relation. An idealized SDOF system is shown schematically in Figure 6.3. It consists of a mass *m* concentrated at the deck level and a massless equivalent column that provides stiffness and damping to the system. The column is usually assumed to be inextensible in the axial direction. This system may be considered as an idealization of a bridge bent, where each structural member (beam, column, etc.) of the actual structure contributes to the inertia (mass), stiffness (or flexibility), and energy dissipation (damping) properties of the system. The equivalent SDOF system and the manner in which its properties are determined represent the "deterministic model" in this case.
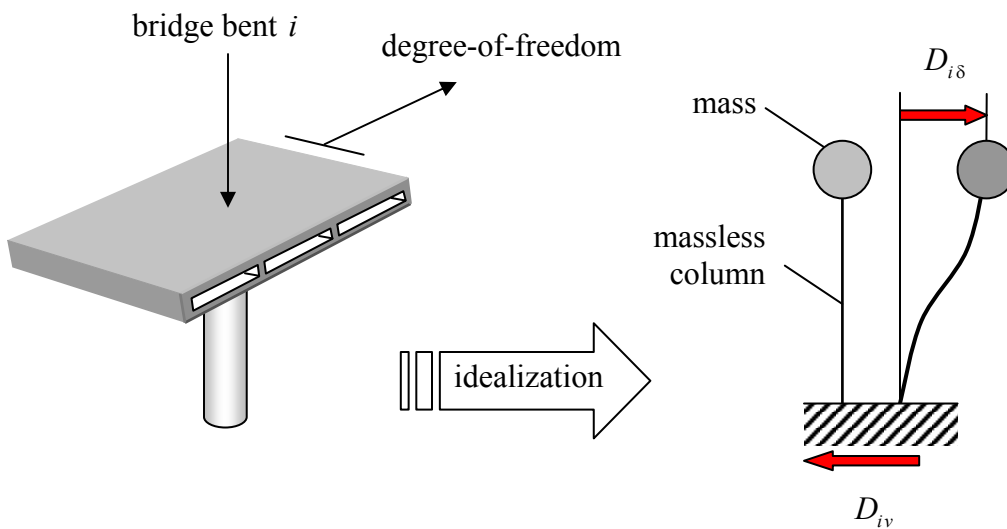
Figure 6.3.   Idealized single-degree-of-freedom system.

Figure 6.4 summarizes the topic of this chapter. Given an earthquake ground motion characterized by a set of parameters, we want to predict D_{iv} and $D_{i\delta}$ on each column i, $i = 1,\ldots,s$, along with their correlation structure, by use of the equivalent SDOF model.
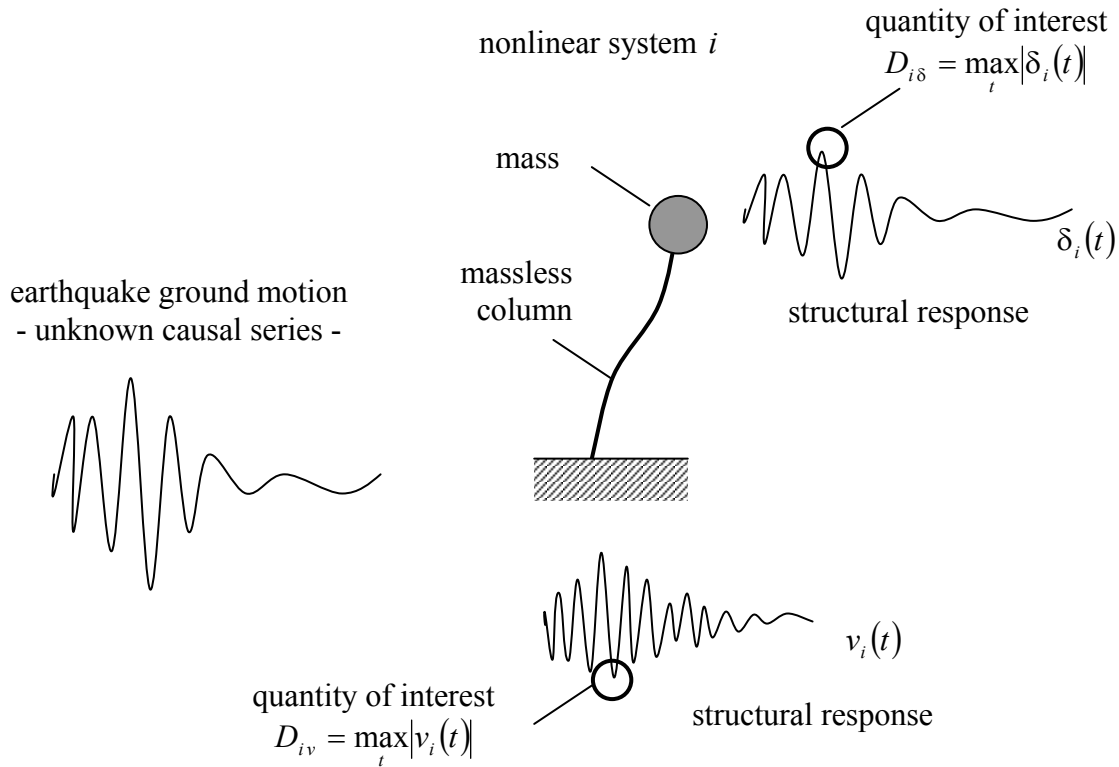
Figure 6.4. **Illustration of the quantities of interest (maximum deformation and shear demands) on an equivalent SDOF system subjected to an unknown earthquake ground motion with specified characteristics.**

6.3   DETERMINISTIC DEMAND MODELS

Ideally, the selected deterministic model for predicting the demand on each component of the system should be simple, yet accurate in estimating the quantities of interest. Moreover, it should account for the interaction of the components that constitute the structural system. In the present study the method proposed by Chopra and Goel (1999) is proposed. The procedure is an improvement of the ATC-40 (ATC, 1996) and FEMA-273/FEMA-274 (FEMA, 1997) capacity-demand diagram methods, which use the well-known constant-ductility spectrum for the demand diagram. In this procedure, first, a nonlinear static analysis of the structure subjected to a monotonically increasing lateral load is performed (pushover analysis). The distribution of the lateral forces corresponds to an assumed displacement shape weighted by tributary masses. Then, an equivalent single-degree-of-freedom (SDOF) system with a bilinear force-displacement

relationship is derived from the pushover curve of the structure. The deformation demand of the equivalent SDOF system is estimated by response spectrum analysis using inelastic spectra. Finally, the local seismic demands are determined by pushing the original structure to the maximum displacement determined in the previous step. In our study, this procedure is implemented numerically using the $R_y - \mu - T_n$ equations proposed by Krawinkler and Nassar (1992). For the sake of completeness, a summary of this procedure is given below:

1. Consider a bilinear hysteretic system with known properties: T_n, natural vibration period, ζ, viscous damping ratio of linear elastic system, V_y, yield strength, α, strain hardening ratio, and $w = mg$, weight of the system, where m is the mass and g is the acceleration due to gravity.

2. Determine the pseudo-acceleration spectrum ordinate (in units of g) from the given elastic response spectrum (following the standard pseudo-acceleration format) for a linear-elastic system with period T_n and damping ratio ζ.

3. Determine the pseudo-acceleration (in units of g) corresponding to the yield deformation $A_y = V_y / w$.

4. Determine the yield deformation $\Delta_y = (2\pi / T_n)^2 A_y \, g$.

5. Determine the yield reduction factor $R_y = A / A_y$.

6. Compute the ductility factor by using the Krawinkler and Nassar (1992) $R_y - \mu - T_n$ equation $\mu = 1 + \left[(R_y)^c - 1 \right] / c$, where $c(T_n, \alpha) = \dfrac{(T_n)^a}{1 + (T_n)^a} + \dfrac{b}{T_n}$ with the numerical coefficients a and b depending on the hardening slope αk as follows: $a = 1.0$ and $b = 0.42$ for $\alpha = 0.0\%$, $a = 1.0$ and $b = 0.37$ for $\alpha = 2.0\%$, and $a = 0.8$ and $b = 0.29$ for $\alpha = 10.0\%$.

7. Estimate the required deformation demand as $\Delta = \mu \, \Delta_y$.

This procedure can be numerically implemented and the results are equivalent to those obtained from the graphical approach by using the well-known constant-ductility design spectrum (Chopra and Goel, 1999).

Chopra and Goel's work focuses on buildings. For buildings, the pushover curve is simply the base shear force V versus the top displacement Δ (Figure 6.5), where the employed distribution of the lateral forces P_i corresponds to the fundamental mode shape φ_1 weighted by the tributary mass of each floor. As demonstrated above, the backbone of the method is the development of the idealized force-deformation relationship of the SDOF system based on utilization of a nonlinear static procedure to develop what is commonly known as the pushover curve. This approach can subsequently be combined with the capacity spectrum method originally developed by Freeman *et al*. (1975). In this approach, the following approximations or assumptions are implied:

1. A fixed lateral force distribution derived from the fundamental vibration mode is used.

2. The earthquake-induced deformation of the inelastic SDOF can be estimated by a nonlinear static procedure without the need for dynamic analysis of the inelastic SDOF system.

Several attempts have been made to include the effect of higher modes, e.g., (Paret *et al*., 1996) and rationalize the nonlinear static procedure, e.g., (Chopra and Goel, 1999).
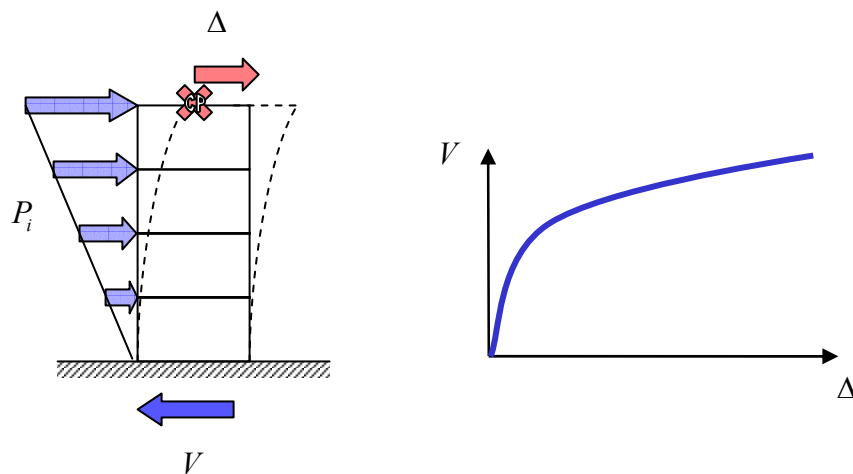
**Figure 6.5   Development of the pushover curve for buildings.**

Fajfar *et al*. (1997) extended a similar methodology (the N2 method, Fajfar and Fischinger, 1987, 1989; Fajfar 2000) to bridges. In the case of bridges, the properties of the equivalent SDOF system are determined based on a characteristic force-displacement relationship of the bridge system in the transverse direction. The force is the sum of all the lateral

forces (total force) and the displacement is monitored at a "characteristic point" at the deck level, where the largest lateral displacement is expected. Note that these selections are rational extensions to bridges of base shear (sum of all the lateral forces) and top displacement used for buildings. According to Fajfar (2000), the distribution of lateral loading P_i should correspond to an assumed displacement shape (not necessarily the first mode shape) weighted by the tributary masses. Figure 6.6a shows the displacement shapes φ suggested by Fajfar for different bridge configurations, together with the location of the characteristic point (solid cross). In the case of a three-bent bridge with columns close to the abutments shorter than the center column (top figure), Fajfar suggests using a triangular displacement shape. In the case of a bridge with more than three columns (center figure), the displacement shape is taken as trapezoidal, so that the deformations at all the degrees of freedom are equal. A similar trapezoidal displacement shape is suggested for a three-bent bridge with columns close to the abutments taller than the center column (lower figure).

In our study, we follow Fajfar's extension of the capacity-demand diagram method to bridges but with one fundamental modification: since we are interested in the shear forces in the columns, the force V is taken as the sum of all the shear forces in the columns, but not including the forces taken by the abutments (Figure 6.6b). This approach provides more accurate estimates of the shear force demands on the bridge columns.
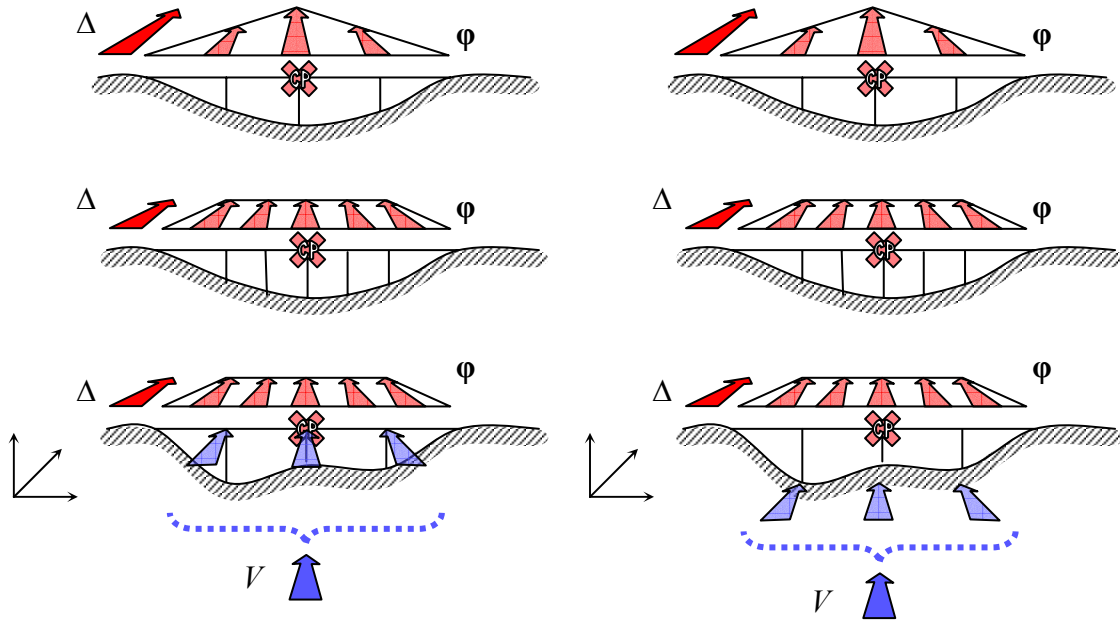
(a) Forces used in the pushover analysis according to Fajfar (2000).

(b) Forces used in the pushover analysis in the present study.

Figure 6.6 **Displacement shapes for different bridge configurations and locations of the characteristic point (solid cross).**

6.4 MODEL CORRECTION

The term $\gamma_{ik}(\mathbf{x}_i, \boldsymbol{\theta}_k)$ on the right-hand side of (5.3) is intended to correct for the bias inherent in the deterministic model $\hat{d}_{ik}(\mathbf{x}_i)$. We select the linear form in (5.4) for this function, where $\boldsymbol{\theta}_k = (\theta_{kj}, j = 1, \ldots, p_k)$, $k = 1, \ldots, q$, is a vector of unknown parameters and $h_{k1}(\mathbf{x}_i), \ldots,$ $h_{k p_k}(\mathbf{x}_i)$, $i = 1, \ldots, s$, $k = 1, \ldots, q$, are selected "explanatory" functions. To capture a potential constant bias in the model that is independent of the variables $\mathbf{x}$, we select $h_{k1}(\mathbf{x}_i) = 1$. To detect any possible under- or overestimation of the deterministic model, we select $h_{k2}(\mathbf{x}_i) = \hat{d}_{ik}(\mathbf{x}_i)$, $k = v$, or δ for shear or deformation demands. Additional explanatory functions are selected to capture the possible dependence of the residuals on ground motion parameters, which may not be properly included or accounted for in the deterministic model.

From an earthquake-engineering standpoint, the most important characteristics of a strong ground motion are the amplitude, the frequency content, and the duration (Kramer, 1996). These characteristics can significantly influence earthquake demand. Nevertheless, knowledge of these quantities alone is not sufficient to accurately describe the damage potential of a ground motion.

With this in mind, we select the following candidate explanatory functions: (a) To provide information on the amplitude and frequency content of an earthquake ground motion we select $h_{k3}(\mathbf{x}_i) = S_a$, where S_a, in units of g, is the spectral acceleration ordinate at the natural period T_n of the system vibrating within its linear elastic range. (b) For a simple harmonic oscillation with peak velocity v_{max} and peak acceleration a_{max}, we know that $v_{max}/a_{max} = T/2\pi$, where T is the period of oscillations. For an earthquake ground motion that includes many frequencies, the quantity $2\pi(v_{max}/a_{max})$ can be interpreted as the period of vibration of an equivalent harmonic wave. This measure can be used as an indication of the predominant period of the ground motion. For this reason we select $h_{k4}(\mathbf{x}_i) = 2\pi(v_{max}/a_{max})/T_n$. Finally, many physical processes, such as stiffness and strength degradation are related to the number of load reversals that occur during an earthquake. Since the number of load reversals is related to the duration of an earthquake as well as to the period of the structure, we select $h_{k5}(\mathbf{x}_i) = t_D/T_n$, where t_D is the ground motion duration defined by Trifunac and Brady (1975) as the time interval between the points at which 5% and 95% of the total energy has been recorded. Note that these explanatory functions are all dimensionless. While additional or different explanatory functions could be selected, we believe that the above functions capture the most significant factors that may influence the demands of RC bridge systems.
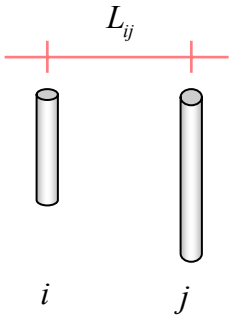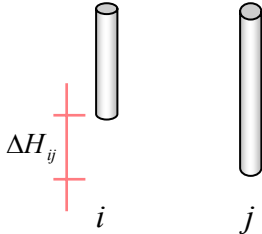
6.5   CORRELATION MATRIX

In order to explore the sources of correlation in the multi-variate model, we select a suitable set of p_R "explanatory" basis functions $h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)$, $w = 1, \ldots, p_R$ and $i, j = 1, \ldots, s$, and express the correlation coefficients as in (5.5). To capture a potential correlation that is independent of the variables $\mathbf{x}$, we select $h_{R1}(\mathbf{x}_i, \mathbf{x}_j) = 1$. To detect any possible dependence of the correlation on

the distance between two bents, we select $h_{R2}(\mathbf{x}_i,\mathbf{x}_j)=L_{ij}/L_{ave}$, where L_{ij} is the distance between bent i and bent j, and L_{ave} is the average span length along the longitudinal axis of the bridge. We select $h_{R3}(\mathbf{x}_i,\mathbf{x}_j)=\Delta H_{ij}/H_{ave}$ to capture the possible dependence of the correlation on the difference in height between bent i and bent j, ΔH_{ij}, where H_{ave} is the average bridge bent height. Finally, to capture the influence of the spectral acceleration S_a, we employ $h_{R4}(\mathbf{x}_i,\mathbf{x}_j)=S_a$. Note that these explanatory functions are all dimensionless. As a result, the parameters $\boldsymbol{\theta}_{kl}=(\boldsymbol{\theta}_{kl1},\ldots,\boldsymbol{\theta}_{klp_R})$, $k,l=1,\ldots,q$, are also dimensionless. While one could select additional explanatory functions or different forms of these functions, we believe the selected ones are sufficiently broad to capture all the factors that may significantly influence the correlation between the quantities of interest. Table 6.1 summarizes our selection of the explanatory function for the correlation matrix.

Table 6.1.   Selected explanatory functions for the correlation matrix.

Explanatory functions			
$h_{R1}(\mathbf{x}_i,\mathbf{x}_j)=1$	$h_{R2}(\mathbf{x}_i,\mathbf{x}_j)=\dfrac{L_{ij}}{L_{ave}}$	$h_{R3}(\mathbf{x}_i,\mathbf{x}_j)=\dfrac{\Delta H_{ij}}{H_{ave}}$	$h_{R4}(\mathbf{x}_i,\mathbf{x}_j)=S_a$
Constant			

With the above selection of explanatory functions, the correlation coefficient between deformation and shear demands of a single column i reduces to

$$\rho_{v\delta}(\mathbf{x}_i) = \frac{\theta_{v\delta1}^{i=j} + \theta_{v\delta4}^{i=j} S_a}{1 + \left|\theta_{v\delta1}^{i=j} + \theta_{v\delta4}^{i=j} S_a\right|} \qquad\qquad i = 1,\ldots,s \qquad\qquad (6.1)$$

and the correlation coefficients between different columns i and j reduce to

$$\rho_{\delta\delta}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{w=1}^{4} \theta_{\delta\delta w} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)}{1 + \left|\sum_{w=1}^{4} \theta_{\delta\delta w} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)\right|} \qquad\qquad i,j = 1,\ldots,s, \quad i \neq j$$

$$\rho_{vv}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{w=1}^{4} \theta_{vvw} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)}{1 + \left|\sum_{w=1}^{4} \theta_{vvw} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)\right|} \qquad\qquad i,j = 1,\ldots,s, \quad i \neq j$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.2)$$

$$\rho_{v\delta}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{w=1}^{4} \theta_{v\delta w} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)}{1 + \left|\sum_{w=1}^{4} \theta_{v\delta w} h_{Rw}(\mathbf{x}_i, \mathbf{x}_j)\right|} \qquad\qquad i,j = 1,\ldots,s \quad i \neq j$$

6.6   EXPERIMENTAL DATA

The predictive demand models are assessed empirically by using "observed" values of the dependent or endogenous variable D_{ik} and of the independent, exogenous, or explanatory variables $\mathbf{x}_i$, for an observed sample size n. Unfortunately, data for full-scale bridge systems are not available except for ground motions with small intensity (e.g., see Arici and Mosalam, 2000), which are not relevant for our study. For this reason, we make use of shake-table tests on single-column bridge bents and "virtual experiments" on full bridge systems. By "virtual experiments" we mean nonlinear dynamic analysis performed on detailed bridge models subjected to selected ground motions.

6.6.1 Experimental Observations

Existing data on tests of bridge bents are organized at the World Wide Web site http://www.ce.berkeley.edu/~gardoni/, where references to the original publications for the tests are given. The database currently contains the results of 51 shake-table tests on single-column bridge bents with circular cross section. Included in the database are the material properties and geometry of each bent, the table motion parameters, and the deformation and shear demands defined, respectively, as the maximum deformation to which the column is subjected and the maximum applied shear force. Unfortunately no experiments on multi-column bents have been conducted.

It is well known that in short-period structures the displacement ductility factor μ rapidly varies with the natural period T_n (Krawinkler and Nassar, 1992). As a result, a small error in estimating T_n could make a significant error in the ductility estimate. Preliminary analysis of the bridge bent data revealed that the test data for columns with $T_n \leq 0.14$ seconds were largely inconsistent with predictions. These data (6 data points) were considered as unreliable and were excluded from consideration in this study. Thus, the probabilistic model assessed using the remaining 45 data points is not appropriate for very short-period ($T_n \leq 0.14$ seconds) structures. Since all tests were conducted under careful laboratory conditions, measurement errors were judged to be small in relation to other sources of model uncertainty and were accordingly neglected.

The nonlinear pushover analyses of the bents needed to compute the deterministic deformation and shear demands are performed using a nonlinear finite element model implemented in the PEER's OpenSees platform (McKenna and Fenves, 2000). In this model, the column is modeled using a two-dimensional nonlinear beam-column fiber element with a circular cross section having one layer of steel evenly distributed around the perimeter of the confined core. The concrete is modeled using the Kent-Scott-Park stress-strain relation (Kent and Park, 1971) with degraded linear unloading/reloading and no strength in tension, and a uni-axial bilinear steel model with kinematic hardening with a post-yield stiffness equal to 5% of the pre-yield stiffness.

6.6.2 Virtual Experiments

To assess the unknown parameters $\boldsymbol{\theta}_{kl}$, $k,l = 1,\ldots,q$, in the correlation coefficients $\rho_{kl}(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}_{kl})$ defined in (5.5) and to include the overall system behavior in the estimates of the parameters $\boldsymbol{\theta}_k$ in $\gamma_{ik}(\mathbf{x}_i, \boldsymbol{\theta}_k)$, $i = 1,\ldots,s$, $k = 1,\ldots,q$, defined in (5.4), we consider detailed nonlinear finite element dynamic analyses performed on a selected bridge system for a suite of earthquakes.

For this purpose, we considered the analyses performed by Fenves and Ellery (1998) on a three-dimensional nonlinear model of the Route 14/Interstate 5 separation and overhead bridge. As described in Fenves and Ellery (1998), the structure is a curved, ten-span, 483 m long bridge with cast-in-place structural concrete box girder superstructure, which partially collapsed in the 1994 Northridge earthquake. Figure 6.7 shows the elevation and plan of this bridge system. The bridge has five frames with single-column piers, connected at four intermediate hinges. The column heights vary over the bridge. Pier 7 has the tallest column (37 m) and Pier 2 has the shortest (8.7 m). The upper 4.3 m of all the columns are tapered at the soffit of the bridge deck. Piers 6 through 9 have internal voids for their entire length. All the columns continue into the ground with a 3.7 m diameter cast-in-place drilled shaft.

The nominal compressive strength of Piers 1 through 7 and 10 is 28 MPa and the one of Piers 8 and 9 is 21 MPa. The longitudinal reinforcement is grade 60 and assumed to have an actual yield stress of 460 MPa. The column transverse reinforcement is assumed to have an actual yield stress of 310 MPa.

The superstructure is a five-cell box girder, 16.2 m wide and 2.1 m deep. The two end frames and the central frame have prestressed box girder superstructures with the deck, soffit, and web thickness of 180 mm, 150 mm, and 300 mm, respectively. Six tendons per web prestress the section; each of the tendons consists of ten 13 mm nominal diameter strands. The conventionally reinforced box girder has a deck, soffit, and web thickness of 180 mm, 160 mm, and 200 mm, respectively.

Fenves and Ellery used the DRAIN-3DX computer program to carry out nonlinear time history analyses for a suite of four recorded and two simulated ground motions. The recorded motions are from the 1994 Northridge earthquake. They were recorded at Arleta Nordhoff Avenue Fire Station (located 10 km from the epicenter), at Jensen Filter Plant (located above the

fault rupture zone with an epicentral distance of 12 km), at Newhall Los Angeles County Fire Station (located 20 km from the epicenter) and at the Sylmar County Hospital Parking lot (approximately 16 km from the epicenter). These ground motions have the largest strong motion acceleration recorded near the interchange site during the Northridge earthquake and are generally characteristic of the earthquake motion in the epicentral region. The simulated ground motions were developed by Horton *et al*. (1995) at bent 2 and by Hutchings *et al*. (1996) at ICN Station. The motions were assumed to be free-field and uniform for all supports.

The earthquake analyses provided estimates of the force and deformation demands at the component level. The finite element models used by Fenves and Ellery (1998) for the time history analyses and for the pushover analyses in the present study are identical. This model is illustrated in Figures 6.8 and 6.9.
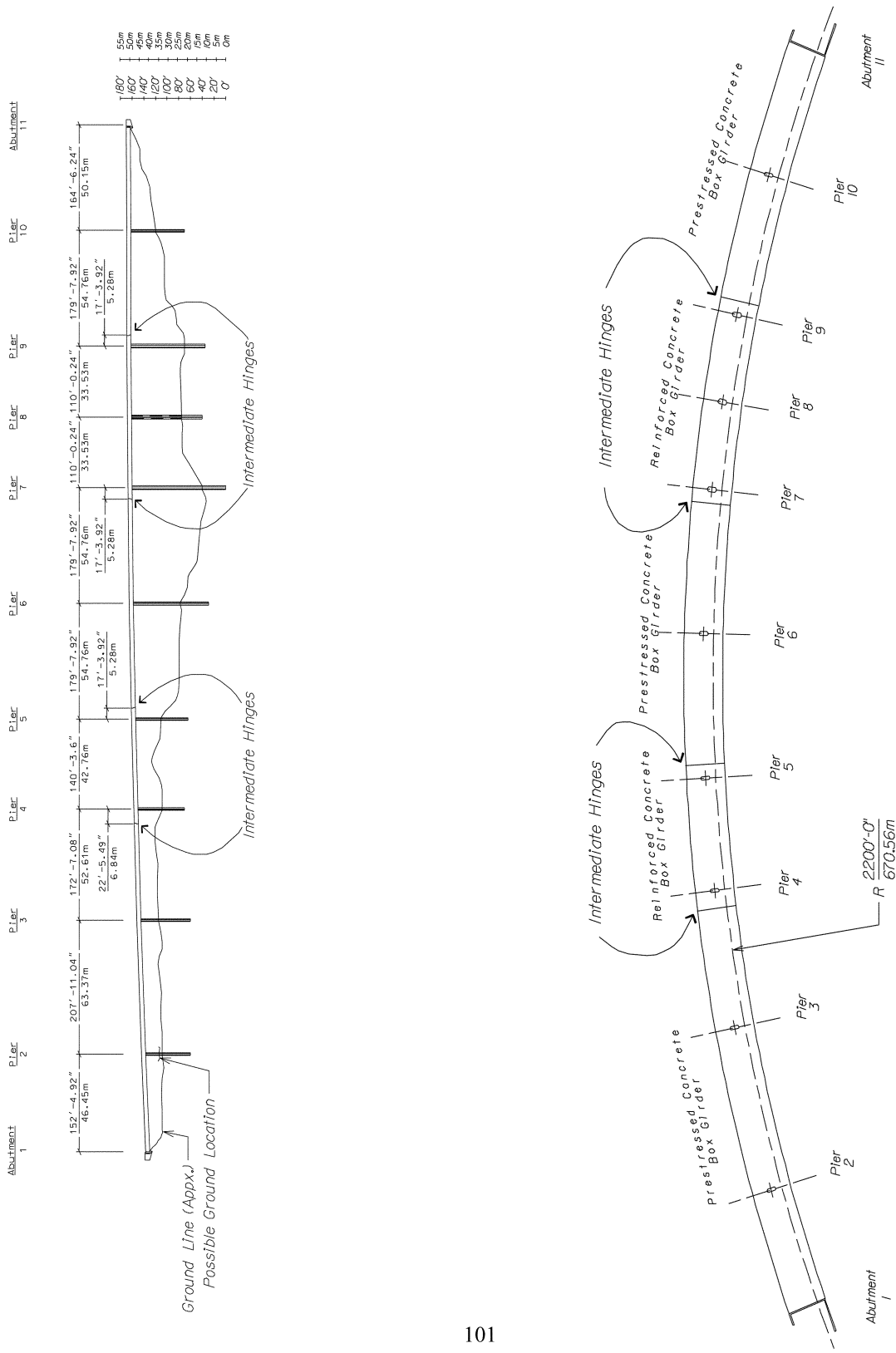
Figure 6.7   Elevation and plan of the I5/14 Interchange bridge (Fenves and Ellery, 1998).
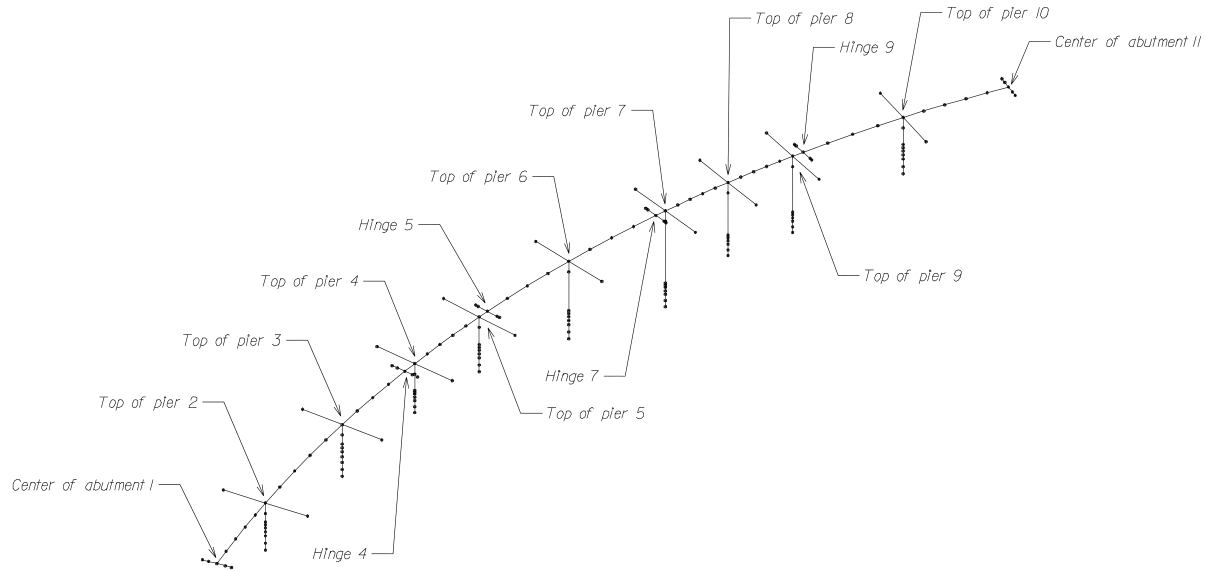
Figure 6.8.   Model of separation and overhead bridge (Fenves and Ellery, 1998).
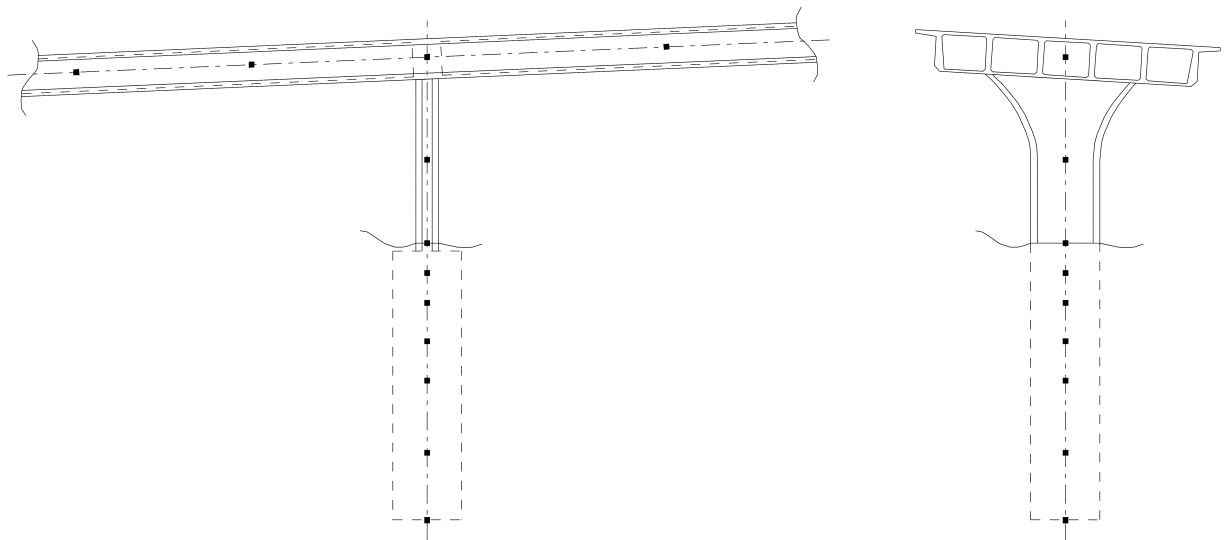
Figure 6.9.   Typical pier model showing node location in column and shaft (Fenves and Ellery, 1998).

Figure 6.10 shows the values of the "measured" normalized shear and deformation demands (solid dots) for two example piers (Piers 7 and 10) when the structure is subjected to the 1994 Northridge earthquake recorded at Sylmar Hospital. These are the values estimated by Fenves and Ellery (1998). The solid lines represent the force-displacement relationships obtained from the pushover analysis of the same structural model and the open circles represent the normalized shear force and deformation demands as predicted by the deterministic procedure

described in Section 6.3. The disparity between the "measured" and predicted demands is due to the error in the deterministic model, i.e., the procedure described in Section 6.3. Similar results are obtained for the other piers of the bridge and other ground motions. In the next section we construct correction terms that explicitly describe the inherent systematic and random errors in the deterministic procedure relative to the "measured" demands.
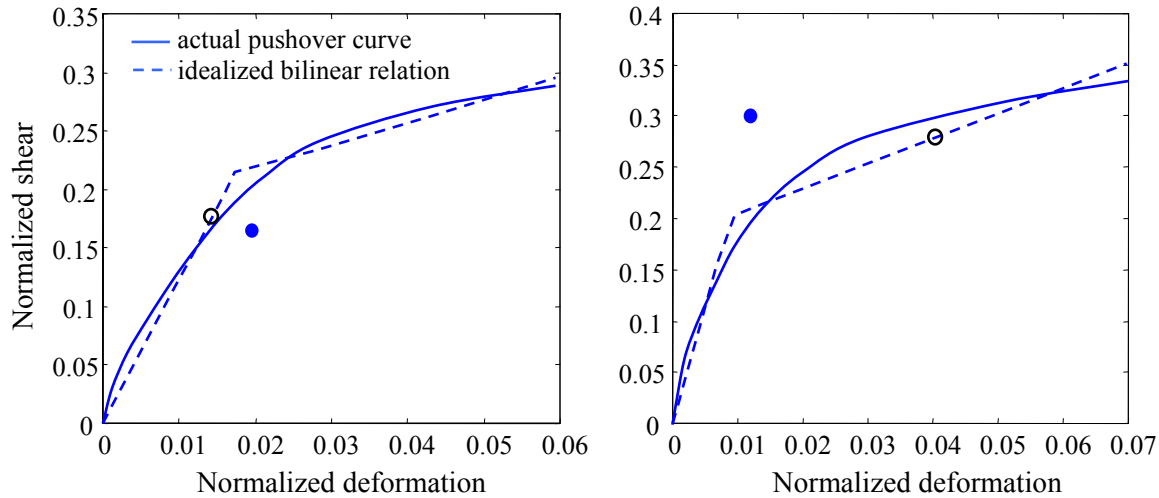
Figure 6.10. **Force-displacement relationship for Piers 7 (left) and 10 (right) and comparison between measured (•) and predicted demands (○) based on the deterministic models.**

6.7 PROBABILISTIC DEMAND MODELS FOR COMPONENTS

The experimental observations on single-column bridge bents and the virtual experiments on the three-dimensional nonlinear model of the Route 14/Interstate 5 separation and overhead bridge are treated as two independent samples of observations. They were used in two stages to assess the probabilistic deformation and shear demand models for RC single-column multi-bent bridge systems with circular cross section subjected to earthquake ground motion. In this section, uni-variate and bi-variate probabilistic models are developed for single bents by using only the first sample of observations, i.e., the experimental observations described in Section 6.6.1. Then, in

the next section, the overall system behavior is included in the estimates of the unknown parameters by updating their posterior distribution based on the first sample of observations by the information content of the second sample of observations, i.e., the virtual experiments described in Section 6.6.2. This application is an example of the versatility of the Bayesian approach in combining information coming from different sources in a consistent manner.

6.7.1　Deformation Demand Model

To develop the deformation demand model we employ the drift demand ratio, δ, defined as the deformation demand, Δ, normalized by the equivalent cantilever length (clear column height), H, of the column. This is a dimensionless quantity, thus convenient for model formulation. Diagnostic plots show that the variance is increasing with the demand level, which is a sign of heteroskedasticity, so a simple linear regression model is not appropriate. Consequently, considering the non-negative nature of the deformation demand, the logarithmic variance-stabilizing transformation is selected among other possible transformations to formulate a homoskedastic model. Diagnostic plots show that the variance in this case can be considered approximately independent of the demand level. Note that this transformation is consistent with the one used in Chapter 3 in the formulation of the deformation capacity model.

Diagnostic plots of the residuals from the model versus the explanatory functions $h_{\delta 2}(\mathbf{x}) = \hat{d}_{i\delta}(\mathbf{x})$, $h_{\delta 3}(\mathbf{x}) = S_a$, $h_{\delta 4}(\mathbf{x}) = 2\pi (v_{\max} / a_{\max}) / T_n$, and $h_{\delta 5}(\mathbf{x}) = t_D / T_n$ show no evident trend and there is no suggestion of dependence of the residuals on these candidate explanatory functions. Evidently, none of the selected ground motion parameters add information to the model beyond that already present in the model. The lack of dependence of the residuals on these explanatory functions lead us to consider the following simple form for the probabilistic model without including terms explicitly involving the ground motion parameters:

$$D_{\delta}(\mathbf{x}, \mathbf{\Theta}_{\delta}) = \hat{d}_{\delta}(\mathbf{x}) + \theta_{\delta 1} h_{\delta 1}(\mathbf{x}) + \theta_{\delta 2} h_{\delta 2}(\mathbf{x}) + \sigma_{\delta} \varepsilon \qquad (6.3)$$

In the above, $D_{\delta}(\mathbf{x}, \mathbf{\Theta}_{\delta})$ is the natural logarithm of the predicted deformation demand, $\hat{d}_{\delta}(\mathbf{x})$ is the natural logarithm of the deterministic demand estimate, and $\mathbf{\Theta}_{\delta} = (\mathbf{\theta}_{\delta}, \sigma_{\delta})$ is the set of unknown model parameters. Substituting the expressions for $h_{\delta j}(\mathbf{x})$ in (6.3), one can write

$$D_\delta(\mathbf{x}, \mathbf{\Theta}_\delta) = \hat{d}_\delta(\mathbf{x}) + \theta_{\delta 1} + \theta_{\delta 2}\hat{d}_\delta(\mathbf{x}) + \sigma_\delta\varepsilon$$

$$= (1 + \theta_{\delta 2})\hat{d}_\delta(\mathbf{x}) + \theta_{\delta 1} + \sigma_\delta\varepsilon \tag{6.4}$$

The algorithms needed for the statistical analyses are programmed in Matlab (1999). In this statistical analysis, no prior information on the parameters is available. Therefore, we select a non-informative prior probability density function for $p(\mathbf{\Theta}_\delta)$, constructed assuming $\mathbf{\theta}_\delta$ and σ_δ are approximately independent, so that $p(\mathbf{\theta}_\delta, \sigma_\delta) \cong p(\mathbf{\theta}_\delta)p(\sigma_\delta)$, $\mathbf{\theta}_\delta$ locally uniform, $p(\mathbf{\theta}_\delta) = \text{constant}$, and taking $p(\sigma_\delta) \propto \sigma_\delta^{-1}$. According to Section 2.2.3, the posterior distribution of σ_δ^2 is then $\eta s_\delta^2 \chi_\eta^{-2}$ where $\eta = 45 - 2 = 43$, and the posterior distribution of $\mathbf{\theta}_\delta = (\theta_{\delta 1}, \theta_{\delta 2})$ is $t_2[\hat{\mathbf{\theta}}_\delta, s_\delta^2 (\mathbf{H}'\mathbf{H})^{-1}, \eta]$. Table 6.2 lists the posterior statistics of $\mathbf{\theta}_\delta$ and σ_δ^2 based on the experimental observations.

Table 6.2. **Posterior statistics of the parameters in the component deformation model based on the experimental observations**.

Parameter	$\theta_{\delta 1}$	$\theta_{\delta 2}$	σ_δ
Mean	0.012	−0.153	0.216
St. dev.	0.116	0.028	0.022
Correlation coefficients			
$\theta_{\delta 2}$	0.96		
σ_δ	0.05	0.00	

From Table 6.2, we note that $\theta_{\delta 1}$ and $\theta_{\delta 2}$ are strongly correlated, $\rho = 0.96$; such that, according to Section 3.5, one can use the approximation

$$\hat{\theta}_{\delta 1} = 0.61 + 3.90\theta_{\delta 2} \tag{6.5}$$

By using this relation, the demand model in (6.4) simplifies to

$$D_\delta(\mathbf{x}, \boldsymbol{\Theta}_\delta) = 0.61 + 3.90\theta_{\delta 2} + (1 + \theta_{\delta 2})\hat{d}_\delta(\mathbf{x}) + \sigma_\delta \varepsilon \qquad (6.6)$$

Figure 6.11 shows a comparison between the measured versus the predicted demands for the test single-column bents based on the deterministic (top chart) and median probabilistic models ($\varepsilon = 0$) (bottom chart). The dotted lines in the bottom chart delimit the region within one standard deviation of the model. The constructed probabilistic model is unbiased and properly accounts for all the underlying uncertainties.
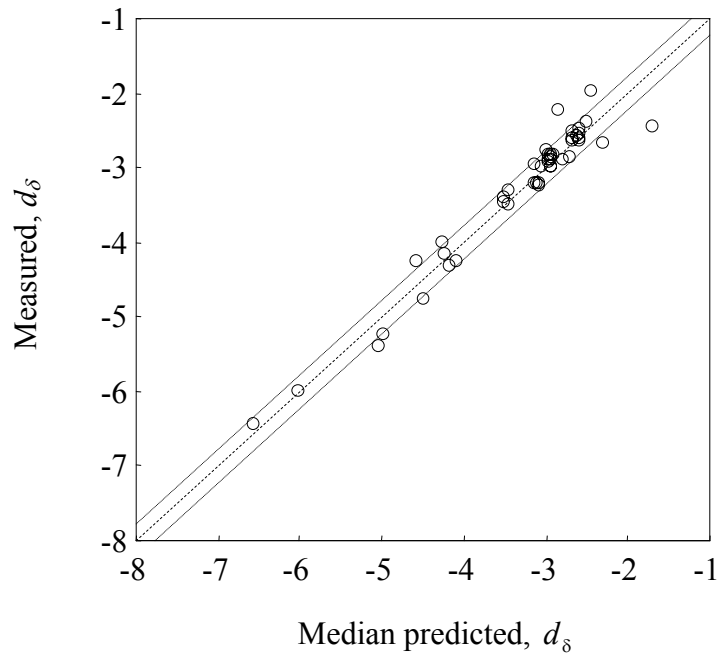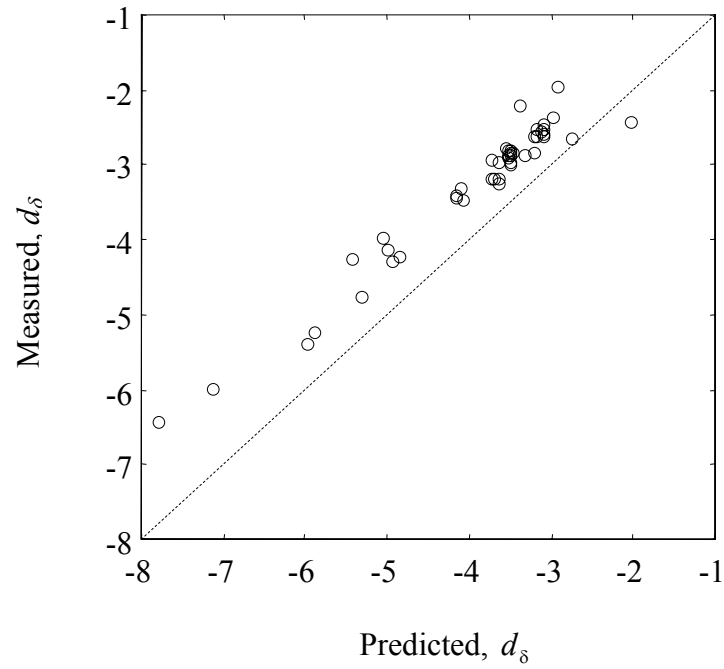
Figure 6.11. Comparison between the measured versus the median predicted (logarithmic) deformation demands on the deterministic (top) and probabilistic (bottom) models.

6.7.2   Shear Demand Model

To develop the shear demand model we employ the normalized quantity $d_v = V/(A_g f_t')$, where V is the shear demand, A_g is the gross cross-sectional area, $f_t' = 0.5\sqrt{f_c'}$ is the nominal tensile strength of concrete in MPa units, and f_c' is the compressive strength of concrete in MPa units. This is a dimensionless quantity, thus convenient for model formulation. The deterministic shear demand is taken as the shear force that corresponds to the deformation demand in the approximate bilinear force-displacement relationship used in the deterministic analysis procedure proposed by Chopra and Goel (1999).

After stabilizing the variance of the model by the logarithmic transformation, we investigate the dependence of the residuals on the explanatory functions $h_{v2}(\mathbf{x}) = \hat{d}_v(\mathbf{x})$, $h_{\delta2}(\mathbf{x}) = \hat{d}_\delta(\mathbf{x})$, $h_{v3}(\mathbf{x}) = S_a$, $h_{v4}(\mathbf{x}) = 2\pi(v_{max}/a_{max})/T_n$, and $h_{v5}(\mathbf{x}) = t_D/T$. Starting from the complete first-order model in the explanatory functions, a step-wise deletion process is performed while keeping the hierarchy of terms. Owing to the lack of prior information, the non-informative prior $p(\mathbf{\Theta}_v) \propto \sigma_v^{-1}$ is selected. The reduced model after deleting insignificant terms has the form

$$D_v(\mathbf{x},\mathbf{\Theta}_v) = \hat{d}_v(\mathbf{x}) + \theta_{v1} h_{v1}(\mathbf{x}) + \theta_{v2} h_{v2}(\mathbf{x}) + \theta_{v3} h_{\delta2}(\mathbf{x}) + \sigma_v\varepsilon \qquad (6.7)$$

where $D_v(\mathbf{x},\mathbf{\Theta}_v)$ is the natural logarithm of the predicted shear demand, $\hat{d}_v(\mathbf{x})$ is the natural logarithm of the deterministic demand estimate, and $\mathbf{\Theta}_v = (\mathbf{\theta}_v,\sigma_v)$ is the set of unknown model parameters. Substituting the expressions for the explanatory functions in (6.7), one can write

$$D_v(\mathbf{x},\mathbf{\Theta}_v) = \hat{d}_v(\mathbf{x}) + \theta_{v1} + \theta_{v2}\,\hat{d}_v(\mathbf{x}) + \theta_{v3}\,\hat{d}_\delta(\mathbf{x}) + \sigma_v\varepsilon \qquad (6.8)$$

Again it can be seen that none of the selected ground motion variables provide additional information to correct the model beyond that already included in the model. According to Section 2.2.3, the posterior distribution of σ_v^2 for the reduced model is $\eta s_v^2 \chi_\eta^{-2}$ where $\eta = 45 - 3 = 42$ and the posterior distribution of $\mathbf{\theta}_v = (\theta_{v1},\theta_{v2},\theta_{v3})$ is $t_3[\hat{\mathbf{\theta}}_v, s_v^2(\mathbf{H}'\mathbf{H})^{-1},\eta]$.

Table 6.3 lists the values of the posterior statistics of $\boldsymbol{\theta}_v$ and σ_v based on the experimental observations.

Table 6.3. **Posterior statistics of the parameters in the shear demand model based on the experimental observations**.

Parameter	θ_{v1}	θ_{v2}	θ_{v3}	σ_v
Mean	−0.526	−1.161	0.375	0.278
St. dev.	0.193	0.152	0.052	0.029
Correlation coefficients				
θ_{v2}	0.58			
θ_{v3}	0.19	−0.66		
σ_v	−0.01	0.03	−0.06	

In Chapter 4 we considered a deterministic model from Moehle *et al.* (1999, 2000) for the shear capacity of an RC circular column that accounted for the reduction in the shear strength due to the effects of flexural stress and redistribution of internal forces as cracking developed. A factor was included in the model (see (4.20)) to account for the strength degradation within the plastic hinge region as a function of the displacement ductility, $\mu_\Delta = \Delta / \hat{\Delta}_y = \delta / \hat{\delta}_y$. The presence of the explanatory function $h_{\delta 2}(\mathbf{x}) = \hat{d}_\delta(\mathbf{x})$ with a positive coefficient in the shear demand model (6.8) may similarly represent the dependence of the error in the deterministic model, $D_v(\mathbf{x}, \boldsymbol{\Theta}_v) - \hat{d}_v(\mathbf{x})$, on the deformation demand.

Figure 6.12 shows a comparison between the measured and predicted values of the shear demands based on the deterministic (top chart) and the probabilistic (bottom chart) models. For the probabilistic model, median predictions are shown ($\varepsilon = 0$). The dotted lines in the bottom chart delimit the region within one standard deviation of the model. We observe that the

Bayesian estimates are unbiased and properly account for the underlying uncertainties in the model.
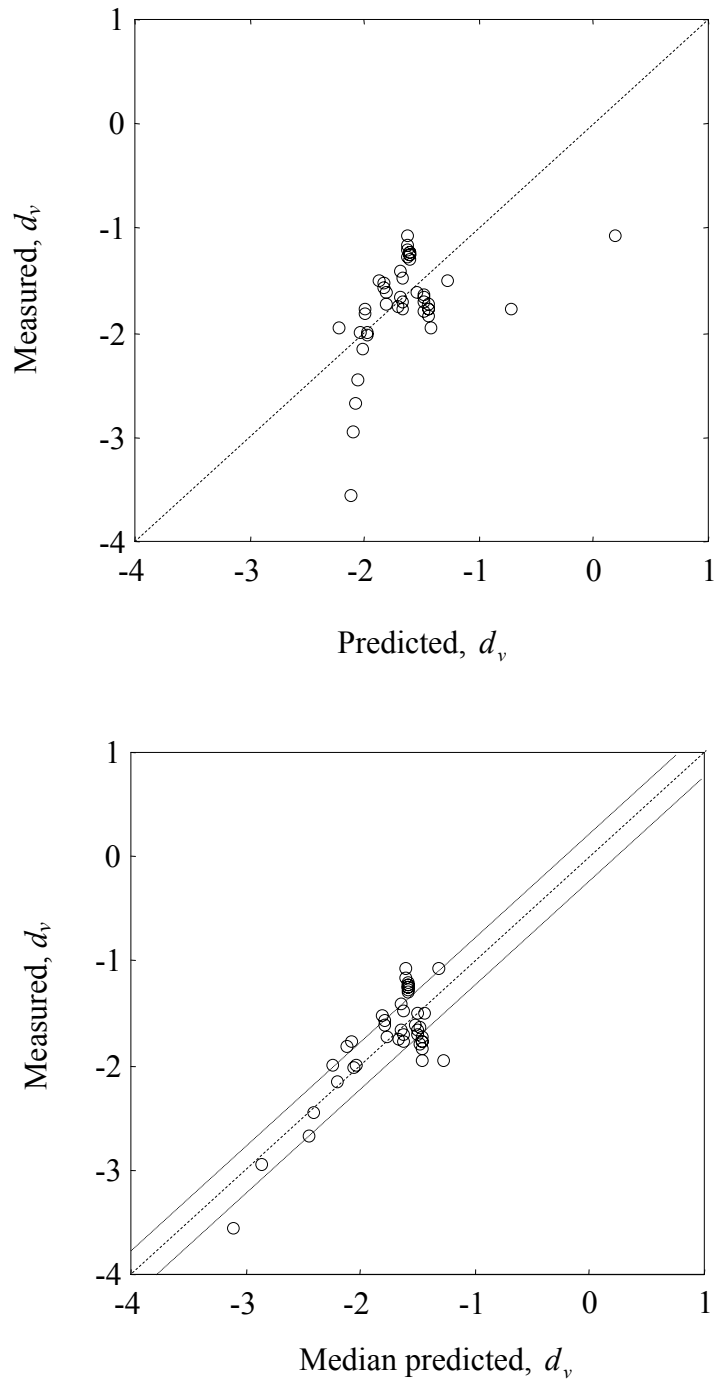
Figure 6.12. **Comparison between the measured versus the median predicted (logarithmic) shear demands based on the deterministic (top) and probabilistic (bottom) models.**

6.7.3 Bi-variate Deformation–Shear Demand Model

In this section a bi-variate deformation-shear demand model is constructed that accounts for the correlation between the two models. By using (6.4) and (6.8), the bi-variate demand model is written as

$$D_\delta\left(\mathbf{x},\boldsymbol{\theta}_\delta,\sigma_\delta,\rho\right)=\hat{d}_\delta\left(\mathbf{x}\right)+\theta_{\delta 1}+\theta_{\delta 2}\hat{d}_\delta\left(\mathbf{x}\right)+\sigma_\delta\varepsilon_\delta$$

$$\quad\quad (6.9)$$

$$D_v\left(\mathbf{x},\boldsymbol{\theta}_v,\sigma_v,\rho\right)=\hat{d}_v\left(\mathbf{x}\right)+\theta_{v1}+\theta_{v2}\hat{d}_v\left(\mathbf{x}\right)+\theta_{v3}\hat{d}_\delta\left(\mathbf{x}\right)+\sigma_v\varepsilon_v$$

The unknown parameters are $\boldsymbol{\theta}=(\boldsymbol{\theta}_\delta,\boldsymbol{\theta}_v)$, where $\boldsymbol{\theta}_\delta=(\theta_{\delta 1},\theta_{\delta 2})$ and $\boldsymbol{\theta}_v=(\theta_{v1},\theta_{v2},\theta_{v3})$, the standard deviations are σ_δ and σ_v, and the correlation coefficient ρ is between ε_δ and ε_v. The posterior distribution of the unknown parameters is given by (2.48) where $n=45$ and $q=2$. Table 6.4 shows the posterior statistics of $(\boldsymbol{\theta},\sigma_\delta,\sigma_v,\rho)$. Since the parameter estimation is based on the same set of data as in Sections 6.7.1 and 6.7.2, the estimates of $\boldsymbol{\Theta}_\delta=(\boldsymbol{\theta}_\delta,\sigma_\delta)$ and $\boldsymbol{\Theta}_v=(\boldsymbol{\theta}_v,\sigma_v)$ are nearly the same as the estimates based on the individual models. No significant correlation between the errors in the deformation and shear demand models is estimated (see the last column of Table 6.4).

Table 6.4. **Posterior statistics of the parameters in the component bi-variate model based on the experimental observations**.

	$\theta_{\delta 1}$	$\theta_{\delta 2}$	θ_{v1}	θ_{v2}	θ_{v3}	σ_{δ}	σ_{v}	ρ
Mean	0.012	−0.152	−0.510	−1.138	0.369	0.213	0.275	0.033
St. dev.	0.118	0.029	0.207	0.196	0.060	0.021	0.029	0.185
Correlation coefficients								
$\theta_{\delta 2}$	0.96							
θ_{v1}	0.04	0.04						
θ_{v2}	0.02	0.007	0.67					
θ_{v3}	0.02	0.03	−0.07	−0.77				
σ_{δ}	0.08	0.07	0.01	0.01	−0.02			
σ_{v}	−0.07	−0.07	0.00	0.01	−0.02	0.03		
ρ	0.04	0.00	0.48	0.69	−0.51	0.00	0.00	

6.8 PROBABILISTIC DEMAND MODELS FOR BRIDGE SYSTEMS

In this section the bi-variate probabilistic demand model introduced in the previous section (see (6.9)) is extended to multi-bent bridge systems subjected to earthquake ground motion. To assess the unknown parameters $\boldsymbol{\theta}_{kl}$, $k,l = 1,\ldots,q$, in the correlation coefficient $\rho_{kl}(\mathbf{x}_i,\mathbf{x}_j,\boldsymbol{\theta}_{kl})$ defined in (5.5) and to include the overall system behavior in the estimation of the parameters $\boldsymbol{\theta}_k$ in $\gamma_{ik}(\mathbf{x}_i,\boldsymbol{\theta}_k)$ and σ_k, $i = 1,\ldots,s$, $k = 1,\ldots,q$, defined in (5.4), we consider the virtual experiments described in Section 6.6.2. The nonlinear pushover analyses of the Route 14/Interstate 5 separation and overhead bridge needed to compute the deterministic deformation and shear demands are performed using the same nonlinear finite element model as the one used for the

virtual experiments. The algorithms needed for the statistical analyses are programmed in Matlab (1999).

Following the updating process illustrated in Section 2.2, we use as prior distribution for the parameters $\boldsymbol{\theta}_k$ the posterior distribution derived in the previous section. The updated posterior distribution of the unknown parameters is proportional to the likelihood function for the virtual experiments multiplied by the previous posterior distribution. This application is an example of the versatility of the Bayesian approach in combining information coming from different sources in a consistent way.

For a bridge system with s single-column bents, (6.9) can be generalized as

$$D_{i\delta}(\mathbf{x}_i,\boldsymbol{\theta}_\delta,\sigma_\delta,\boldsymbol{\rho})=\hat{d}_{i\delta}(\mathbf{x}_i)+\theta_{\delta1}+\theta_{\delta2}\hat{d}_{i\delta}(\mathbf{x}_i)+\sigma_\delta\varepsilon_{i\delta} \qquad i=1,\ldots,s \qquad (6.10\text{a})$$

$$D_{iv}(\mathbf{x}_i,\boldsymbol{\theta}_v,\sigma_v,\boldsymbol{\rho})=\hat{d}_{iv}(\mathbf{x}_i)+\theta_{v1}+\theta_{v2}\hat{d}_{iv}(\mathbf{x}_i)+\theta_{v3}\hat{d}_{i\delta}(\mathbf{x}_i)+\sigma_v\varepsilon_{iv} \qquad i=1,\ldots,s \qquad (6.10\text{b})$$

After removing the non-informative terms, the correlation coefficients for the deformation-shear model in (6.1) and (6.2) reduce to

$$\rho_{v\delta}=\frac{\theta_{v\delta1}^{i=j}}{1+\left|\theta_{v\delta1}^{i=j}\right|} \qquad i=j \qquad (6.11\text{a})$$

$$\rho_{\delta\delta}=\frac{\theta_{\delta\delta1}}{1+\left|\theta_{\delta\delta1}\right|} \qquad i\neq j \qquad (6.11\text{b})$$

$$\rho_{vv}=\frac{\theta_{vv1}}{1+\left|\theta_{vv1}\right|} \qquad i\neq j \qquad (6.11\text{c})$$

$$\rho_{v\delta}=\frac{\theta_{v\delta1}}{1+\left|\theta_{v\delta1}\right|} \qquad i\neq j \qquad (6.11\text{d})$$

The unknown parameters are then $\boldsymbol{\theta}=(\boldsymbol{\theta}_\delta,\boldsymbol{\theta}_v)$, where $\boldsymbol{\theta}_\delta=(\theta_{\delta1},\theta_{\delta2})$ and $\boldsymbol{\theta}_v=(\theta_{v1},\theta_{v2},\theta_{v3})$, and the entries of the covariance matrix $(\sigma_\delta,\sigma_v,\theta_{v\delta1}^{i=j},\theta_{\delta\delta1},\theta_{vv1},\theta_{v\delta1})$. Table 6.5 shows the posterior statistics of $(\boldsymbol{\theta},\sigma_\delta,\sigma_v,\theta_{v\delta1}^{i=j},\theta_{\delta\delta1},\theta_{vv1},\theta_{v\delta1})$. The posterior means were estimated by maximizing the posterior distribution of the parameters and the standard deviations and

correlation coefficients were estimated by bootstrapping the residuals with 300 repetitions (see Section 2.2.5). For the purpose of these applications, the bootstrapping algorithms are programmed in Matlab (1999).

Table 6.5.   **Posterior statistics of the parameters in the multi-variate system model based on the experimental observations and the virtual experiments**.

	$\theta_{\delta 1}$	$\theta_{\delta 2}$	θ_{v1}	θ_{v2}	θ_{v3}	σ_{δ}	σ_{v}	$\theta_{v\delta 1}^{i=j}$	$\theta_{\delta\delta 1}$	θ_{vv1}	$\theta_{v\delta 1}$
Mean	−1.512	−0.388	−0.076	−0.764	0.278	0.639	0.598	−0.005	0.008	−0.008	0.002
St. dev.[1]	0.041	0.025	0.027	0.077	0.045	0.054	0.043	0.137	0.206	0.224	0.179

<div align="center">Correlation coefficients[1]</div>

	$\theta_{\delta 1}$	$\theta_{\delta 2}$	θ_{v1}	θ_{v2}	θ_{v3}	σ_{δ}	σ_{v}	$\theta_{v\delta 1}^{i=j}$	$\theta_{\delta\delta 1}$	θ_{vv1}
$\theta_{\delta 2}$	−0.96									
θ_{v1}	0.64	−0.68								
θ_{v2}	−0.74	0.77	−0.90							
θ_{v3}	−0.17	0.19	−0.05	−0.02						
σ_{δ}	−0.42	0.45	−0.62	0.74	−0.05					
σ_{v}	−0.45	0.47	−0.47	0.60	0.29	0.81				
$\theta_{v\delta 1}^{i=j}$	−0.74	0.74	−0.40	0.53	0.16	0.16	0.31			
$\theta_{\delta\delta 1}$	−0.61	0.65	−0.62	0.75	0.10	0.87	0.90	0.38		
θ_{vv1}	−0.56	0.59	−0.65	0.77	0.00	0.93	0.86	0.29	0.97	
$\theta_{v\delta 1}$	−0.59	0.60	−0.63	0.76	0.02	0.87	0.84	0.49	0.90	0.89

1. By bootstrapping the residuals.

If there were no information content in the virtual experiments, the posterior statistics in Table 6.4 and Table 6.5 would be exactly the same. The numerical differences in the estimates then reflect the added information content of the virtual experiments that was not included in Table 6.5. In particular, the following observations are noteworthy: (a) The posterior means of $\theta = (\theta_{\delta}, \theta_{v})$ are vastly different in the two cases. This indicates that it is important to include the overall system behavior in the model assessment and that the information content of the virtual experiments is quite relevant. (b) Combining the information content of both samples, we

increase the number of data points while maintaining the number of degrees of freedom of the model (the number of parameters) in fitting the data. This increases the values of σ_δ and σ_v in Table 6.5 with respect to the values in Table 6.4. (c) Using the expressions of the correlation coefficients in (6.11) and the posterior estimates of $(\theta_{v\delta1}^{i=j}, \theta_{\delta\delta1}, \theta_{vv1}, \theta_{v\delta1})$ in Table 6.5, we have $\rho_{v\delta} = -0.005$ for $i = j$, and $\rho_{\delta\delta} = 0.008$, $\rho_{vv} = -0.008$ and $\rho_{v\delta} = 0.002$ for $i \neq j$. The corresponding standard deviations are relatively small, i.e., of order 0.2. These indicate that the error terms ε_{ik} and ε_{jl}, $i, j = 1, \ldots, s$, $k, l = 1, \ldots, q$, are practically uncorrelated. The next section further explores the effects of the information content of the two samples of observations by examining different predictions of the demands.

6.9   EFFECTS OF THE TWO INDEPENDENT SAMPLES OF OBSERVATIONS

This section explores the effects of the information content of the first sample of observations, the experimental observations described in Section 6.6.1, and of the second sample of observations, the virtual experiments described in Section 6.6.2.

Figure 6.13 shows a comparison between the measured versus the predicted demands for the tested single-column bents (open circles) and the results from the virtual experiments (solid dots), based on the deterministic and median probabilistic models. For a perfect model, the data would line up along the 1:1 dashed line. The dotted lines delimit the region within one standard deviation of the probabilistic model. Figure 6.13a shows a comparison between the measured and predicted values of the deformation demands based on the deterministic model. In Figure 6.13b the measured values are plotted versus the median predicted demands based on the probabilistic model assessed using only the first sample (single-column experiments) of observations. It can be seen that the probabilistic prediction for the data on single- column bents are unbiased, as already pointed out in Section 6.7.1; however, for the virtual experiments the median of the probabilistic model overestimates the deformation demand. Figure 6.13c shows the measured values plotted versus the median predicted demands based on the probabilistic model estimated now using only the second sample (virtual experiments) of observations. Clearly, this is the opposite case of the one just described: the predictions for the virtual experiments are unbiased, while for the experimental data on single-column bents the median of

the probabilistic model underestimates the deformation demand. Finally, Figure 6.13d shows the measured values plotted versus the median predicted demands estimated using both samples of observations. The latter are based on the posterior statistics in Table 6.5. In this case, if we do not distinguish between solid dots and open circles, the predictions are overall unbiased. If we just look at the estimates for the data on single-column bents or at the estimates for the data from the virtual experiments, we see a worsening from their best estimates shown respectively in Figures 6.13b and 6.13c. This is because in combining the information content of both samples of observations we have more data points but we have used the same degrees of freedom of the model (number of parameters) in fitting the data as before. While one would expect the region within one standard deviation of the model to be narrower in Figures 6.13b and 6.13c than in Figure 6.13d, we note that, since most of the uncertainty is coming from the virtual experiments, they are essentially the same in Figures 6.13c ($\sigma_\delta = 0.627$) and 6.13d ($\sigma_\delta = 0.639$).

Figure 6.14 shows the same comparison between measured and predicted demands as Figure 6.13 but now for the shear demands. Similar remarks on the effects of the two samples of observations can be made.
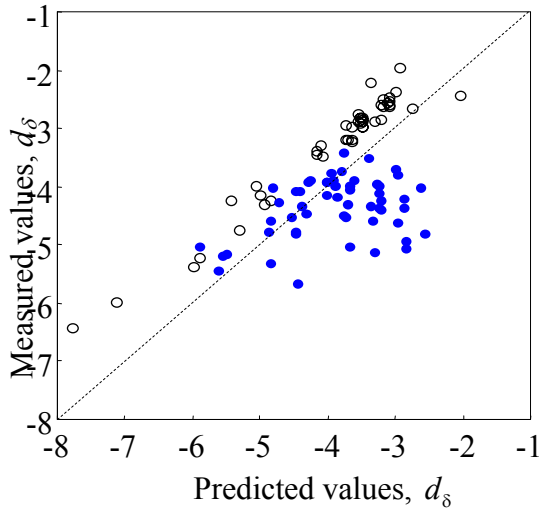
(a) Measured values versus deterministic predictions
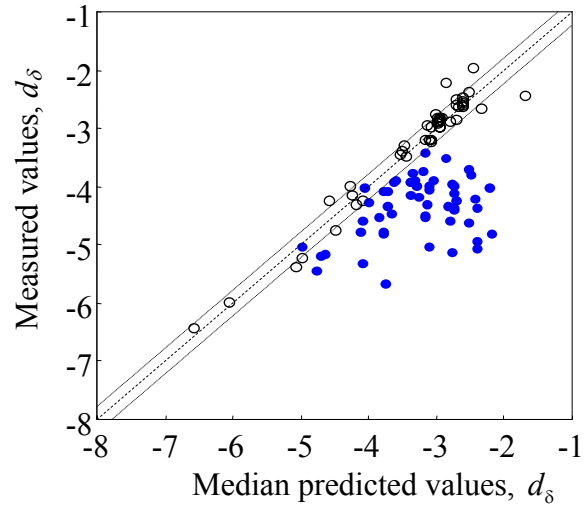
(b) Measured values versus median probabilistic predictions using single-column experiments
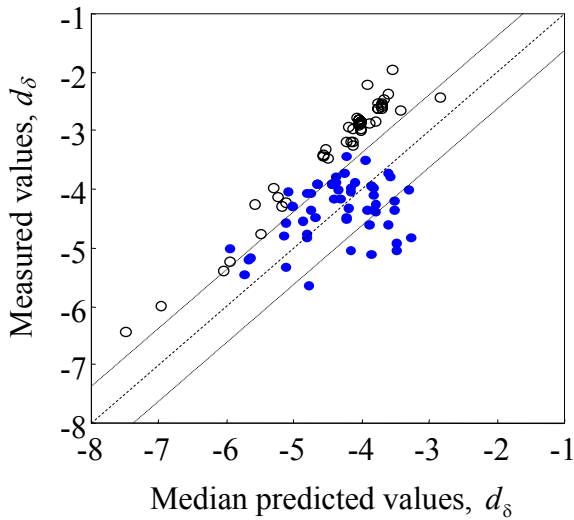
(c) Measured values versus median probabilistic predictions using virtual experiments
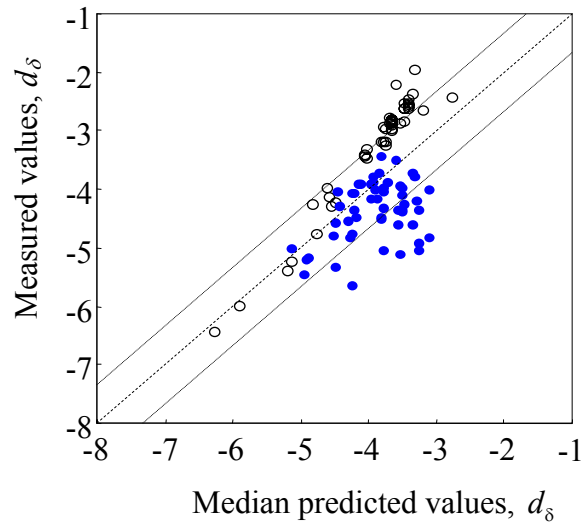
(d) Measured values versus median probabilistic predictions using all data

Figure 6.13. **Comparison between measured and median predicted (logarithmic) deformation demands for the tested single-column bents (o) and the virtual experiments (•), based on the deterministic and median probabilistic models.**

(a) Measured values versus deter-ministic predictions
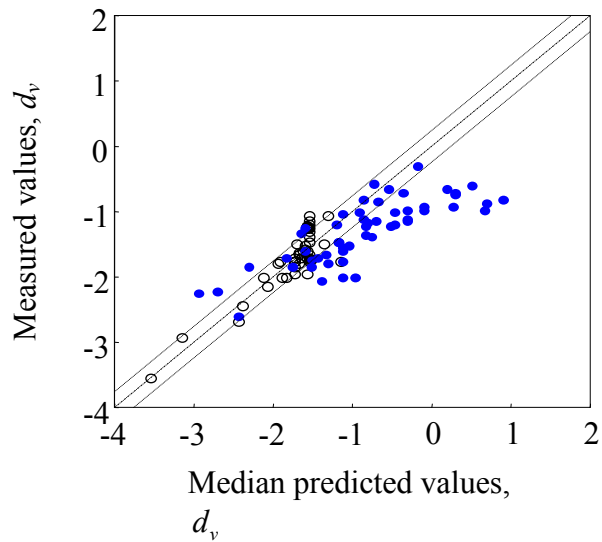
(b) Measured values versus median probabilistic predictions using single-column experiments
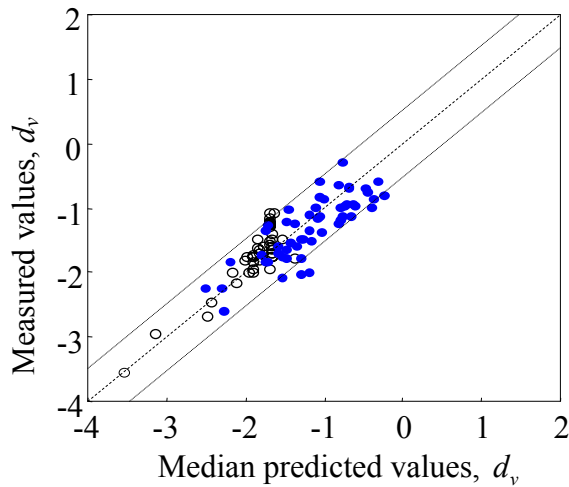
(c) Measured values versus median probabilistic predictions using virtual experiments
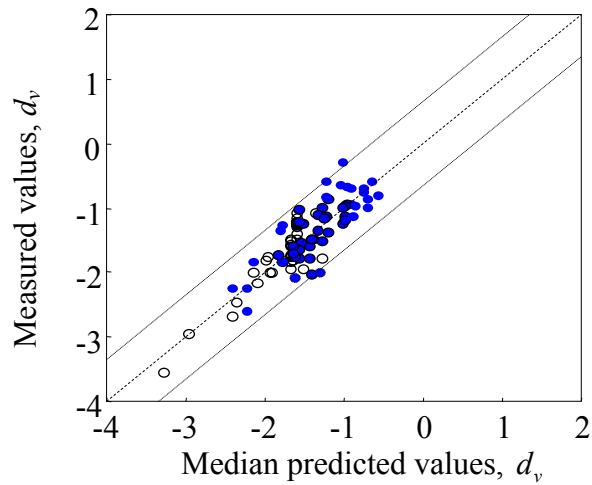
(d) Measured values versus median probabilistic predictions using all data

Figure 6.14. **Comparison between measured and median predicted (logarithmic) shear demands for the tested single-column bents (∘) and the virtual experiments (•), based on the deterministic and median probabilistic models.**

Figure 6.15 shows the deformation demands in the transverse direction for each bent for each of the virtual experiments. The solid dots represent the observed demands while the open circles are the median predicted values. The dotted lines delimit the region within one standard deviation of the model. Each row of charts corresponds to a different excitation among a suite of four recorded and two simulated ground motions. From top to bottom, the rows correspond to the free-field ground motions recorded at Arleta, Sylmar, Jensen, and Newhall stations and to the ground motions simulated by Horton *et al.* (1995) at bent 2 and by Hutchings *et al.* (1996) at ICN Station. The first column to the left of the chart shows the observed deformation demands and the predicted demands based on the deterministic model. For a perfect model the solid dots should nail the open circles. We see that the choice of the deterministic method was appropriate because, despite its simplicity, it produces fairly accurate results. The second column shows the observed deformation demands and the median predicted demands based on the probabilistic model assessed using only the first sample of observations (single-column experiments). The third column shows the observed deformation demands and the median predicted demands based on the probabilistic model assessed using only the second sample of observations (virtual experiments). Since the model is assessed based only on the data that are shown here, the model corrects the inaccuracies of the deterministic model over the entire length of the bridge. Finally, the fourth column shows the observed deformation demands and the median predicted demands for the model assessed using both independent samples of observations. In this case, a worsening of the predictions can be seen with respect to column 3. As we previously noted, this is because in combining the information content of both samples of observations, we have more data points but the same degrees of freedom (i.e., number of parameters) of the model as before. A similar remark as for Figure 6.13 can be made about the variation of the width of the region within one standard deviation of the models.

For all the ground motions, we note that the closer the pier is to the abutments the larger the discrepancy between its observed deformation demands and its median predicted demands. This may be due to the use of a trapezoidal displacement shape function, which implies constant displacement for each pier. In reality, the short piers close to the abutments tend to displace less than the taller piers closer to the center-span of the bridge. The assumed displacement shape is more reasonable on the right side of the bridge, where the pier close to the abutment (Pier 10) is almost as tall as the tallest pier (Pier 7), than it is on the left side of the bridge where Piers 2, 3, and 4 are all significantly shorter than Pier 7. As a result, while the deterministic predictions of

the deformation demands on the piers close to the center span are quite accurate, the predictions close to the abutments overestimate the actual demands, especially for the piers close to the left abutment. A deformation function tapered to the left could give better estimates. These discrepancies are also indications that the deterministic procedure employed for the analysis provides a poor model of the influence of the abutments.

Figure 6.16 shows the same comparison between measured and (median) predicted demands as Figure 6.15 but now for the shear demands. Similar remarks can be made for Figure 6.16 as for Figure 6.15. Note, however, that in this case the discrepancies between the observed shear demands and the median predicted demands do not increase when approaching the abutments as much as before and that these discrepancies are almost completely removed after assessing the model by using the virtual experiments (see the last two columns).
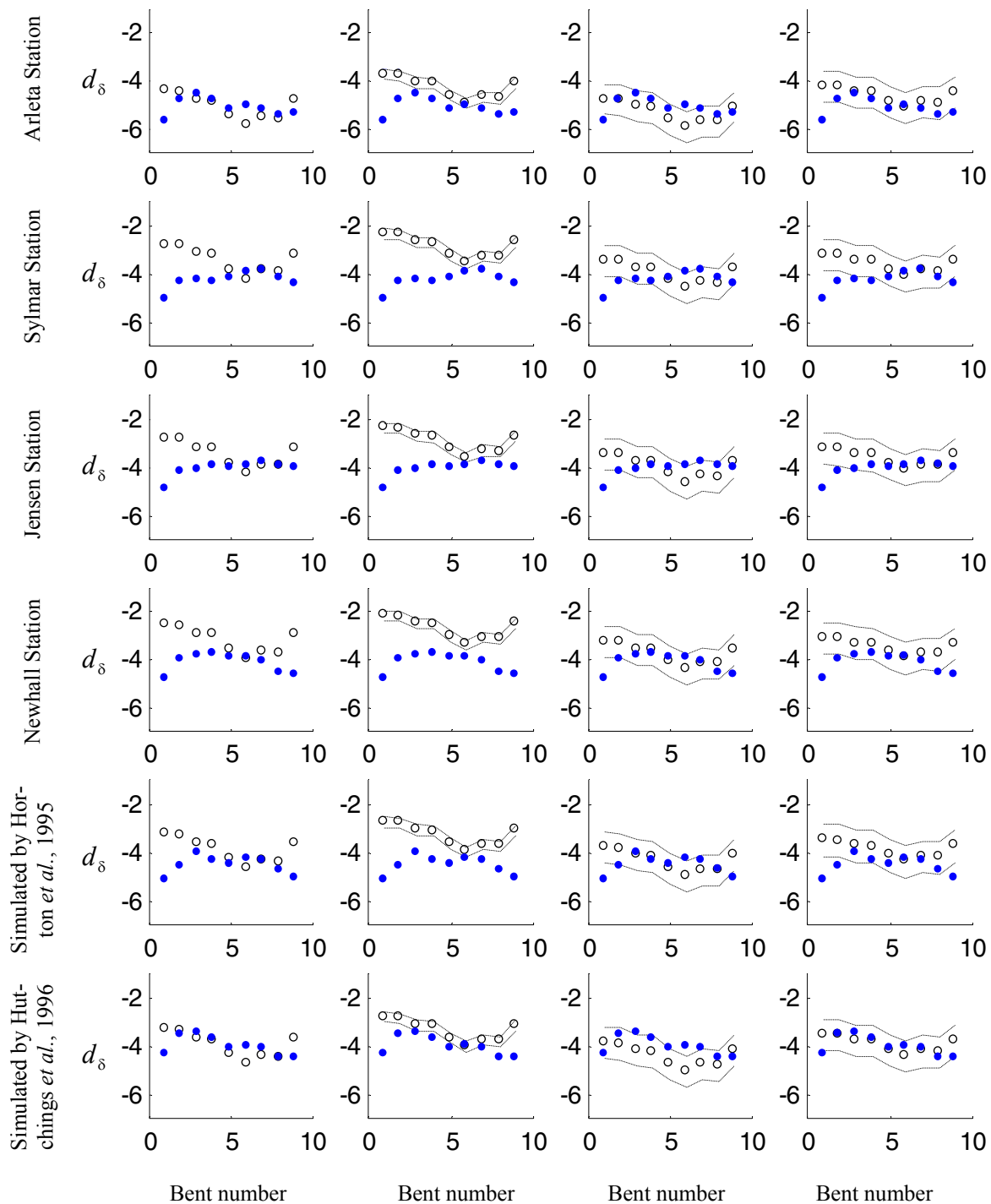
Figure 6.15. Observed (•) and median predicted (o) deformation demands for the virtual experiments.
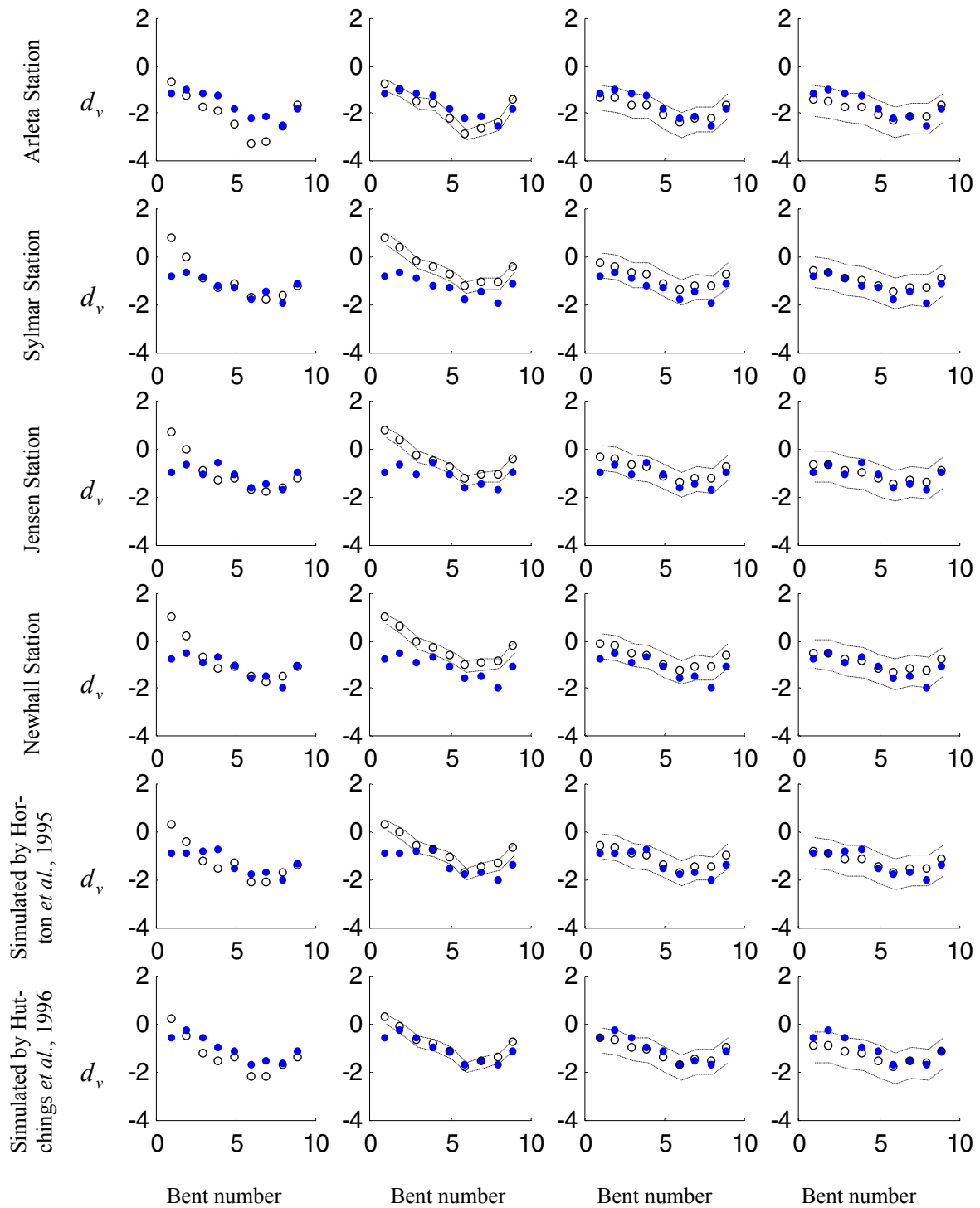
Figure 6.16. Observed (•) and median predicted (○) shear demands for the virtual experiments.

6.10   SUMMARY

Probabilistic models for the deformation and shear demands of RC bridge components and systems subjected to earthquake ground motion are developed by a Bayesian approach using existing experimental observations on single-column bridge bents and virtual experiments on a three-dimensional nonlinear model of the Route 14/Interstate 5 separation and overhead bridge. First, uni-variate and bi-variate probabilistic models for single bents are developed using only the experimental observations on single-column bridge bents. Then, the overall system behavior is included in the estimates of the unknown parameters by updating their posterior distribution based on the experimental observations by the information content of the virtual experiments. This application is an example of the versatility of the Bayesian approach in combining information coming from different sources in a consistent manner.

The probabilistic models are unbiased and explicitly account for all the relevant uncertainties, including errors arising from an inaccurate model form or missing variables, and statistical uncertainty. The correlation coefficients between demand measures are also assessed. With the aim of facilitating their use in practice, the models are constructed by developing correction terms to an existing deterministic procedure. Through the use of a set of "explanatory" functions, we are able to identify the terms that effectively correct the bias in the existing models and to gain insight into the underlying behavioral phenomena.

7 Fragility Estimates for Structural Components and Systems

7.1 INTRODUCTION

In this chapter we define the fragility of structural components and systems along with providing alternative estimates that differ according to how parameter uncertainties are treated. By using the capacity models described in Chapters 3 and the demand models described in Chapters 5, limit-state functions are constructed and fragility estimates are given with special attention given to the treatment of aleatory and epistemic uncertainties.

7.2 FRAGILITY ASSESSMENT

For structural components and systems, fragility is defined as the conditional probability of attaining or exceeding prescribed limit states for a given set of demand variables. Following the conventional notation in structural reliability theory (Ditlevsen and Madsen, 1996), let $g_{ik}(\mathbf{x}, \boldsymbol{\Theta})$ be a mathematical model for the i-th structural component ($i = 1, \ldots, s$) of a general system, describing its k-th limit state of interest ($k = 1, \ldots, q$). The limit-state function is defined such that the event $\{g_{ik}(\mathbf{x}, \boldsymbol{\Theta}) \leq 0\}$ denotes the attainment or exceedance of the k-th limit state by the i-th structural component. As in the previous chapters, $\mathbf{x}$ denotes a vector of measurable variables and $\boldsymbol{\Theta}$ denotes a vector of model parameters. Usually $\mathbf{x}$ can be partitioned in the form $\mathbf{x} = (\mathbf{r}, \mathbf{s})$, where $\mathbf{r}$ is a vector of material and geometrical variables, and $\mathbf{s}$ is a vector of demand variables such as boundary forces or deformations.

By using the capacity models described in Chapter 3 and the demand models described in Chapter 5, the limit-state functions for a structural component can be formulated as

$$g_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}) = C_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_C) - D_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_D) \qquad i=1,\dots,s, \quad k=1,\dots,q \qquad (7.1)$$

where $D_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_D)$ denotes the demand value relative to the capacity $C_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_C)$ for the k-th failure mode of the i-th structural component. For example, for the failure in shear of an RC column i, $C_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_C)$ may be expressed in terms of the maximum shear force that the column can sustain, whereas $D_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_D)$ may denote the maximum applied shear force. Note that both quantities can be functions of demand variables, e.g., the applied axial force. Therefore, both functions $C_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_C)$ and $D_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}_D)$ generally could include $\mathbf{s}$ as an argument.

Most generally, the fragility of a series structural system (e.g., a bridge constituted by single-column bents) can be stated as

$$F(\mathbf{s},\mathbf{\Theta}) = \mathrm{P}\left[\bigcup_{i=1}^{s}\bigcup_{k=1}^{q} \{ g_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta}) \le 0 \} \,\Big|\, \mathbf{s},\mathbf{\Theta} \right] \qquad (7.2)$$

where $\mathrm{P}[A\,|\,\mathbf{b}]$ denotes the conditional probability of event A for the given values of variable(s) $\mathbf{b}$. The uncertainty in the event for the given $\mathbf{s}$ arises from the inherent randomness in the capacity variables $\mathbf{r}$, variability in the actual demand for the given $\mathbf{s}$, which is caused by the inexact nature of the limit-state model $g_{ik}(\mathbf{r},\mathbf{s},\mathbf{\Theta})$ (or its sub-models), and the uncertainty inherent in the model parameters $\mathbf{\Theta}$. We have expressed the fragility as a function of the parameters to emphasize that an estimate of the fragility depends on how the model parameters are treated.

Various estimates of fragility can be developed depending on how the parameter uncertainties are treated (Der Kiureghian, 1999). These are described in the following sections.

7.2.1   Point Estimates of Fragility

A point estimate of the fragility is obtained by ignoring the uncertainty in the model parameters and by a point estimate $\hat{\mathbf{\Theta}}$ in place of $\mathbf{\Theta}$. Most commonly the posterior mean, $\mathbf{M}_{\mathbf{\Theta}}$, or the

maximum likelihood estimate, $\mathbf{\Theta}_{\text{MLE}}$, is used. The corresponding point estimate of the fragility is denoted as

$$\hat{F}(\mathbf{s}) = F(\mathbf{s}, \hat{\mathbf{\Theta}}) \qquad (7.3)$$

The uncertainty in this estimation arises from the intrinsic variability in $\mathbf{r}$ and from the random model correction term $\varepsilon_{g\,ik}$, arising from the correction terms of the capacity and the demand models, which is essentially aleatory in nature. In the special case when variables $\mathbf{r}$ are deterministically known, $\hat{F}(\mathbf{s})$ can be computed in terms of the multi-normal probability distribution of $\varepsilon_{g\,ik}$. More generally, a multifold integral involving the joint distribution of $\mathbf{r}$ and $\varepsilon_{g\,ik}$ over the failure domain must be computed. Methods for the numerical computation of such probability terms have been well developed in the field of structural reliability (Ditlevsen and Madsen, 1996).

7.2.2 Predictive Estimate of Fragility

The point estimate of fragility in (7.3) does not incorporate the epistemic uncertainties inherent in the model parameters $\mathbf{\Theta}$. To incorporate these uncertainties, $\mathbf{\Theta}$ must be considered as random variables. The predictive estimate of fragility, $\tilde{F}(\mathbf{s})$, is the expected value of $F(\mathbf{s}, \mathbf{\Theta})$ over the posterior distribution of $\mathbf{\Theta}$, i.e.,

$$\tilde{F}(\mathbf{s}) = \int F(\mathbf{s}, \mathbf{\Theta}) f(\mathbf{\Theta}) d\mathbf{\Theta} \qquad (7.4)$$

where $f(\mathbf{\Theta})$ is the posterior density of $\mathbf{\Theta}$. This estimate of fragility incorporates the epistemic uncertainties in an average sense, but it does not distinguish between the fundamentally different natures of the aleatory and epistemic uncertainties.

7.2.3   Bounds on Fragility

In some applications, it is desirable to determine the uncertainty inherent in the fragility estimate due to the epistemic uncertainties. This uncertainty is reflected in the probability distribution of $F(\mathbf{s},\Theta)$ relative to the parameters Θ, as shown in Figure 7.1. As stated above and shown in the figure, $\widetilde{F}(\mathbf{s})$ is the mean of this distribution. Exact evaluation of this distribution unfortunately requires nested reliability calculations (Der Kiureghian, 1989). Approximate confidence bounds can be obtained by first-order analysis in the manner described below.
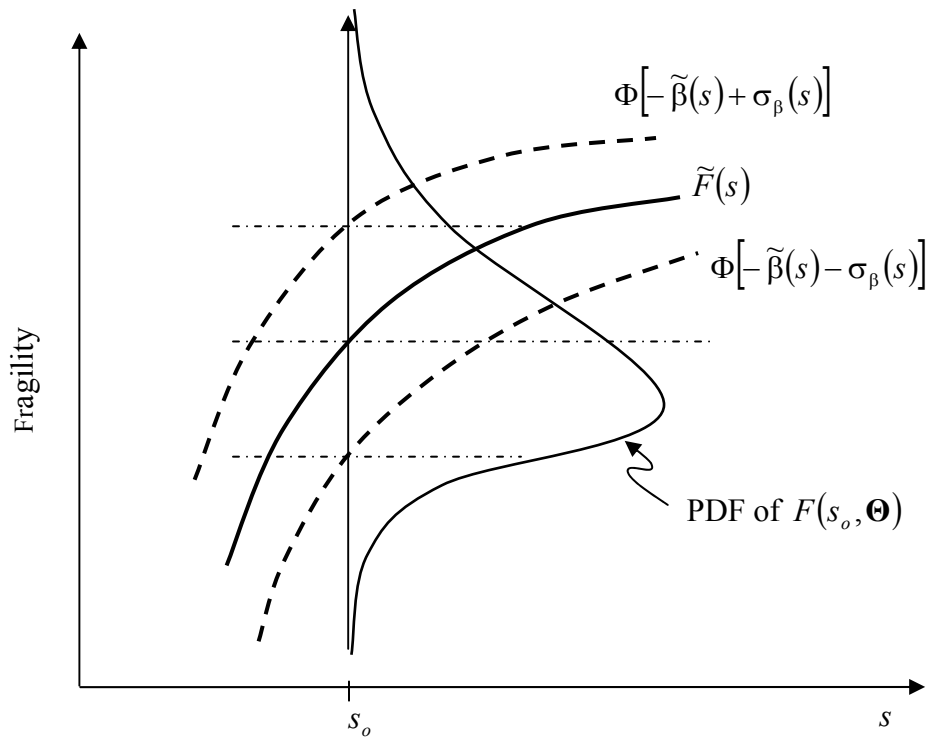
Figure 7.1.   Fragility estimates incorporating epistemic uncertainties.

Consider the reliability index corresponding to the conditional fragility in (7.2), which is defined as

$$\beta(\mathbf{s},\Theta) = \Phi^{-1}\left[1 - F(\mathbf{s},\Theta)\right] \tag{7.5}$$

where $\Phi^{-1}[\,\cdot\,]$ denotes the inverse of the standard normal cumulative probability. In general $\beta(\mathbf{s},\boldsymbol{\Theta})$ is less strongly nonlinear in $\boldsymbol{\Theta}$ than $F(\mathbf{s},\boldsymbol{\Theta})$. By using a first-order Taylor series expansion around the mean point, the variance of $\beta(\mathbf{s},\boldsymbol{\Theta})$ is approximately given by

$$\sigma_\beta^2(\mathbf{s}) \approx \nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})\,\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}}\,\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})^{\mathrm{T}} \qquad\qquad (7.6)$$

where $\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})$ is the gradient row vector of $\beta(\mathbf{s},\boldsymbol{\Theta})$ at the mean point $\mathbf{M}_{\boldsymbol{\Theta}}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}}$ denotes the posterior covariance matrix of $\boldsymbol{\Theta}$. The gradient vector $\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})$ is easily computed by first-order reliability analysis (see Ditlevsen and Madsen, 1996). Bounds on the reliability index can now be expressed in terms of a specified number of standard deviations away from the mean. For example, $\widetilde{\beta}(\mathbf{s}) \pm \sigma_\beta(\mathbf{s})$, where $\widetilde{\beta}(\mathbf{s}) = \Phi^{-1}[1 - \widetilde{F}(\mathbf{s})]$, denotes the mean plus/minus one standard deviation bounds of the reliability index. Transforming these back into the probability space, one obtains

$$\left\{ \Phi\!\left[-\widetilde{\beta}(\mathbf{s})-\sigma_\beta(\mathbf{s})\right],\ \Phi\!\left[-\widetilde{\beta}(\mathbf{s})+\sigma_\beta(\mathbf{s})\right] \right\} \qquad\qquad (7.7)$$

as the "one standard-deviation" bounds of the fragility estimate, as illustrated in Figure 7.1. These bounds approximately correspond to 15% and 85% probability levels. Applications in Chapter 8 demonstrate the theoretical formulation presented here.

8 Fragilities of Reinforced Concrete Bridge Components and Systems

8.1 INTRODUCTION

In this chapter fragility estimates are developed for RC bridge components and systems following the theory presented in Chapter 7. First, the probabilistic capacity models developed in Chapter 4 are used to estimate the fragilities of a typical bridge column in terms of maximum deformation and shear demands. Point and interval estimates of the fragilities are computed that implicitly or explicitly reflect the influence of epistemic uncertainties. The fragility estimates account for the effect of cyclic loading, since the experimental data used in developing the probabilistic capacity models (Chapter 4) are obtained from tests on RC columns subjected to cyclic loading.

Next, the probabilistic demand models developed in Chapter 6 are used in conjunction with the capacity models (Chapter 4) to construct limit-state functions that properly account for all the relevant uncertainties and to objectively assess the seismic fragilities of an example RC bridge bent for a given set of ground motion parameters. Predictive estimates of the fragilities are computed, where the influences of aleatory and epistemic uncertainties are explicitly reflected. For a typical bridge bent in California, the deformation failure mode dominates the fragility estimate.

Finally, this analysis is extended to the fragility assessment of bridge systems. Two configurations of typical new California highway overpass bridges are considered and fragility estimates are computed both at the component level and at the system level.

8.2   FRAGILITY ESTIMATES FOR RC BRIDGE COLUMNS

The probabilistic deformation and shear capacity models developed in Chapter 4 can be used to assess the fragility for any circular column with specified geometry, material properties and applied compressive axial force. In this section we estimate two uni-variate fragility curves and a bi-variate fragility surface for an example column with geometry and material properties that are representative of currently constructed RC highway bridge columns in California (Naito *et al.*, 2000). The considered column has the longitudinal reinforcement ratio $\rho_l = 1.99\%$, gross diameter $D_g = 1{,}520$ mm with the ratio of gross to core diameters $D_g / D_c = 1.07$, clear height $H = 9{,}140$ mm, and volumetric transverse reinforcement ratio $\rho_s = 0.65\%$ with yield stress of transverse reinforcement $f_{yh} = 493$ MPa. Furthermore, to account for material variability we assume the compressive strength of concrete, f_c', to be described by a lognormal distribution with mean 35.8 MPa and 10% coefficient of variation, and the yield stress of longitudinal reinforcement, f_y, to be lognormally distributed with mean 475 MPa and 5% coefficient of variation. To account for variability in the axial load, we assume P to be normally distributed with mean 4,450 kN (corresponding to 7% of the axial capacity based on the gross cross-sectional area) and 25% coefficient of variation. Finally, to account for variability in construction, we assume that the effective moment of inertia, I_e, is lognormally distributed with mean 2.126×10^{11} mm^4 and 10% coefficient of variation.

As defined in the Chapter 7, fragility is the conditional probability of failure given one or more measures of demand $\mathbf{s} = (s_1, s_2, \ldots)$. The predictive fragility estimate $\widetilde{F}(\mathbf{s})$ is the expected fragility estimate with respect to the distribution of the model parameters $\mathbf{\Theta}$. This estimate accounts for the effect of epistemic uncertainties (uncertainty in the model parameters) in an average sense. Explicit account of the variability in the fragility estimate due to epistemic uncertainties is provided by confidence bounds at specified probability levels. For the purpose of these applications, the probabilities of failure are computed with CalREL (Liu *et al.*, 1989) by Monte Carlo simulation and the confidence bounds are computed according to Section 7.2.3 by using first-order reliability analysis. In the following, these estimates are presented for the example column.

Figures 8.1 and 8.2, respectively, show the uni-variate fragility curves with respect to drift demand, δ, and normalized shear demand, v, for the example column. The solid lines represent the predictive estimates $\widetilde{F}(\delta)$ and $\widetilde{F}(v)$ and the dashed lines indicate the approximate 15% and 85% confidence bounds. The dispersion indicated by the slope of the solid curve represents the effect of aleatory uncertainties (those present in f_c', f_y, P, I_e, ε_d, and ε_v) and the dispersion indicated by the confidence bounds represents the influence of the epistemic uncertainties (those present in the model parameters Θ).
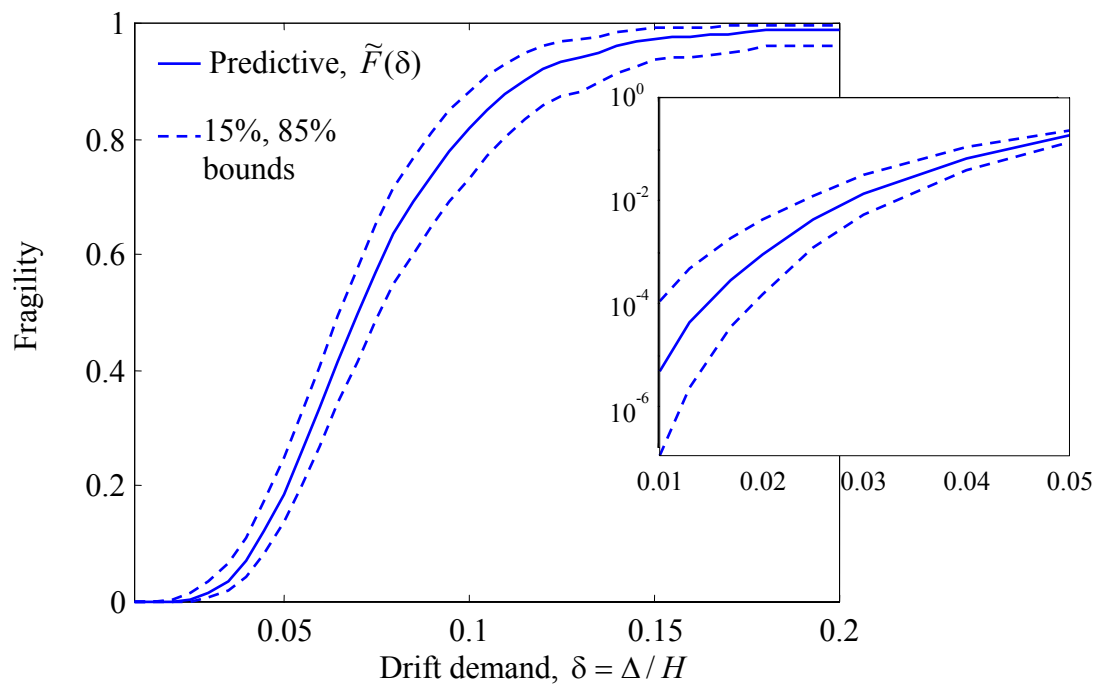
Figure 8.1.   **Fragility estimate for deformation failure of the example circular RC column.**
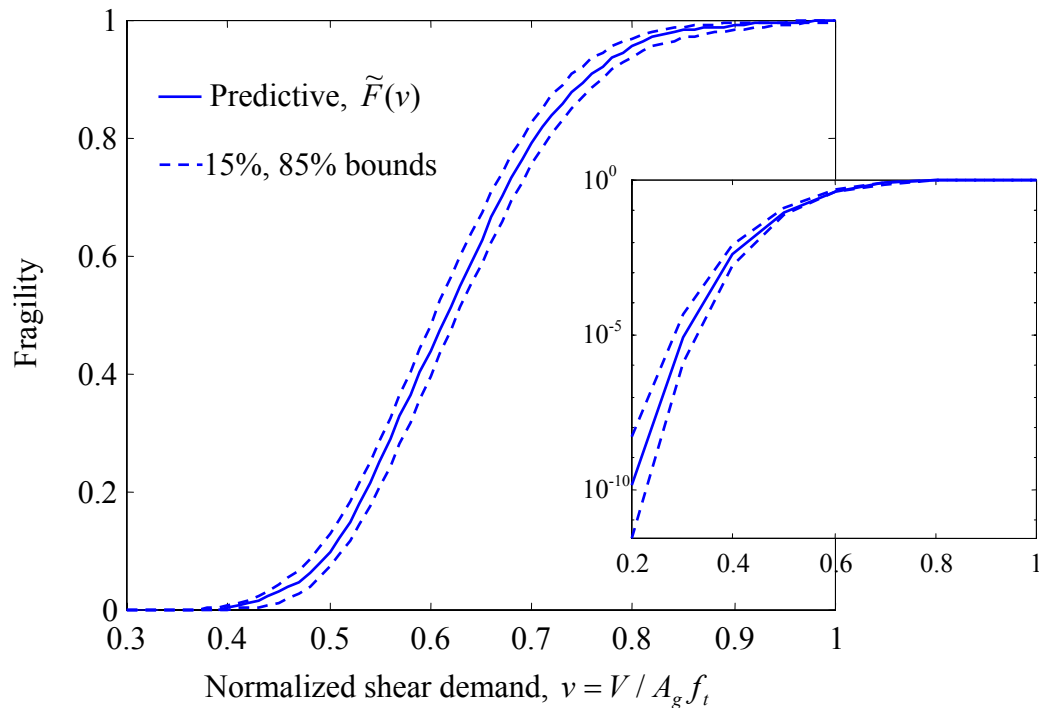
Figure 8.2. **Fragility estimate for shear failure of the example circular RC column.**

The bi-variate deformation-shear fragility estimates for the example column are obtained using the bi-variate capacity model developed in Chapter 4. The fragility in this case is defined as failure of the column, in either deformation or shear mode, for a given pair of deformation and shear demands. Figure 8.3 shows contour plots of the predictive fragility surface $\widetilde{F}(\delta, v)$ in terms of the drift ratio demand δ and the normalized shear demand v. Each contour in this plot connects pairs of values of the demands δ and v that give rise to a given level of fragility in the range 0.1–0.9. Significant interaction between the two failure modes, particularly at high demand levels, is observed.
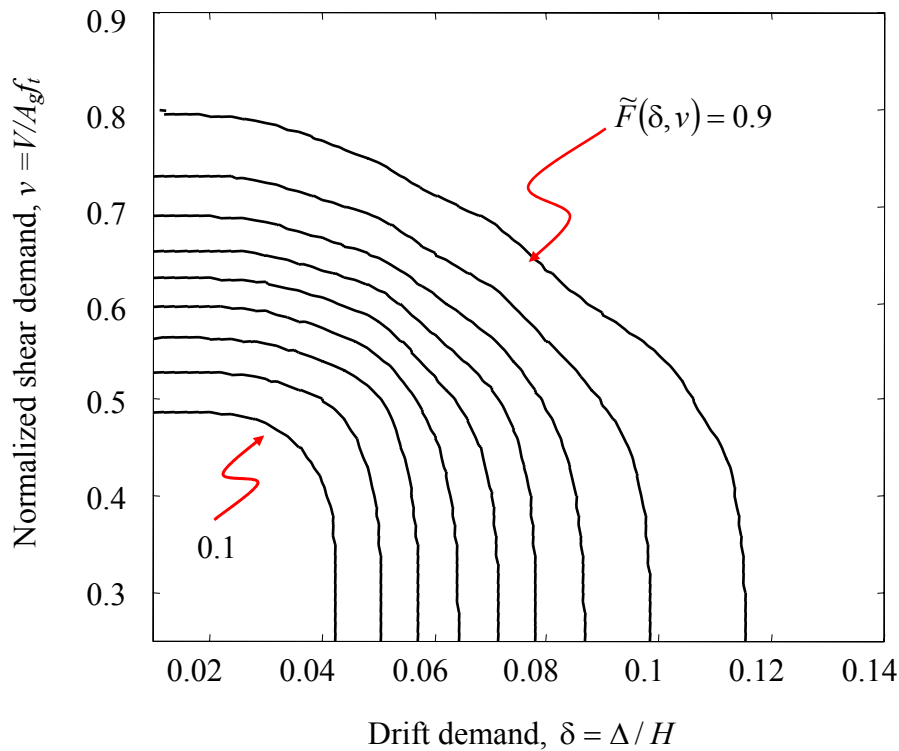
Figure 8.3. Contour plot of the predictive deformation-shear fragility surface of the example circular RC column.

8.3 FRAGILITY ESTIMATES FOR RC BRIDGE BENTS

The probabilistic demand models developed in Chapter 6 can be combined with the probabilistic capacity models (Chapter 4) to construct limit-state functions that properly account for all the relevant uncertainties. Such limit-state functions can be used to assess the structural fragilities of any single-column bent with circular cross section with specified geometry and material properties for a given set of ground motion parameters.

As an example, consider a column with the same geometry and material properties as defined in the previous section. The nonlinear pushover analysis of the bent that is now needed to compute the deterministic deformation and shear demands, $\hat{d}_k$, $k = \delta, v$, is performed using a nonlinear finite element model implemented in the PEER's OpenSees platform (McKenna, 2000). In this model, the column is represented using a two-dimensional nonlinear beam-column element with circular section having one layer of steel evenly distributed around the perimeter of

the confined core. The concrete is modeled using the Kent-Scott-Park stress-strain relation (Kent and Park, 1971) with degrading linear unloading/reloading and no strength in tension, and a uni-axial bilinear steel model with kinematic hardening with a post-yield stiffness equal to 5% of the pre-yield stiffness.

The probabilities of failure are computed with the Reliability module in OpenSees (Haukaas and Der Kiureghian, 2001) with an external limit-state function evaluator implemented in Matlab (1999). The reliability module in OpenSees provides a comprehensive tool for uncertainty analysis. Material, geometry, or load variables can be declared as random quantities with distributions assigned from a library or provided by the user. Dependence between random variables is accounted for. By specifying a set of "limit-state" functions that describe an event of interest, e.g., exceeding a deformation threshold, formation of a mechanism, the probability of the event or the mean rate of its occurrence in time can be computed. The statistical and model uncertainties, introduced in Chapter 3, are accounted for within a Bayesian reliability framework.

The solid line in Figure 8.4 indicates the predictive fragility estimate, $\widetilde{F}(\mathbf{s})$, for the example bent defined as the conditional probability of attaining or exceeding the deformation or shear limit states (series system) for a given value of S_a, where S_a, in units of g, is the spectral acceleration ordinate at the natural period T_n of the system vibrating within its linear elastic range. Note that T_n is not a constant but a function of the basic random variables previously defined. In this figure S_a ranges between 0.2 and 1.4, and these limits are within the range of values of S_a for the experimental data used to assess the probabilistic demand models. The dashed curve indicates the deformation fragility alone and the dotted curve indicates the shear fragility alone. We note that for a typical bridge bent in California, the deformation failure mode dominates the fragility estimate as intended in the design.
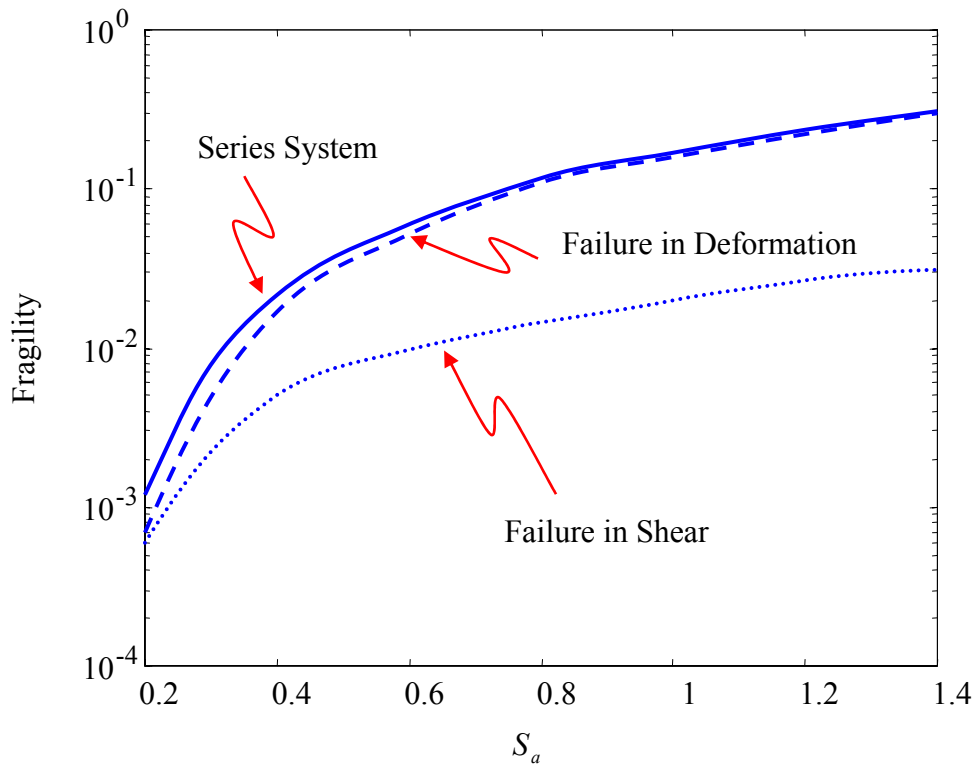
Figure 8.4.   Fragility estimates for the example single-column bridge bent.

8.4   FRAGILITY ESTIMATES FOR RC BRIDGES

Two configurations of typical new California highway overpass bridges are considered. These have been designed by Mackie and Stojadinović (2001) according to Caltrans's Bridge Design Specification and Seismic Design Criteria (Caltrans, 1999), which incorporate recommendations from ATC-32 (1996). The first configuration (Figure 8.5) is a single-bent overpass and the second (Figure 8.6) is a two-bent overpass. Both configurations have single-column bents, Type I integral pile-shaft foundations extending the columns with the same cross sections into the soil, as designed by Caltrans (Yashinsky and Ostrom, 2000). The bridge decks have a typical Caltrans box section for a three-lane wide roadway. The circular column cross sections have perimeter longitudinal bars and spiral confining reinforcement. The design parameters for the example overpass bridges are defined in Figures 8.5 and 8.6, and the numerical values are listed in Table 8.1.
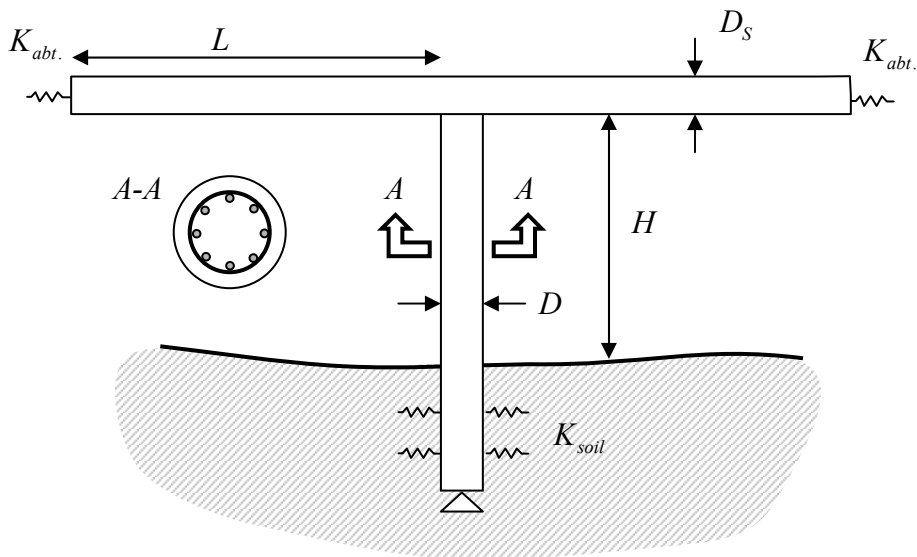
Figure 8.5. Design parameters for the example single-bent overpass bridge (not to scale).
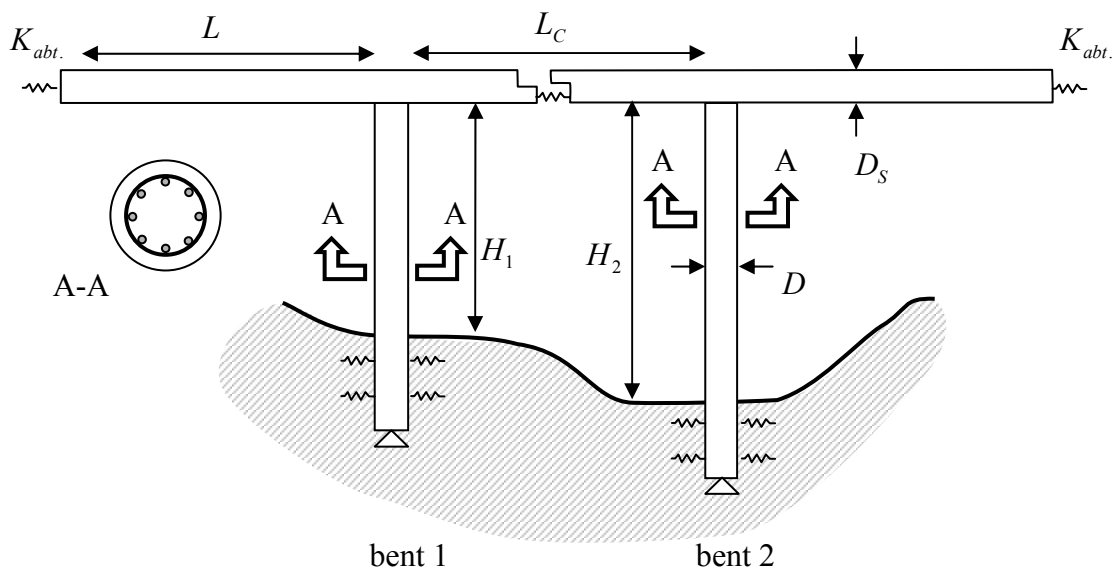
Figure 8.6. Design parameters for the example two-bent overpass bridge (not to scale).

To account for material variability, we assume the compressive strength of concrete, f'_c, is lognormally distributed with mean 27.6MPa and 10% coefficient of variation, and the yield stress of longitudinal reinforcement, f_y, is lognormally distributed with mean 448.2MPa and 5% coefficient of variation. To account for variability in the axial load for the single-bent overpass,

138

we assume the additional bridge dead load (e.g., weight of vehicles) $r = W_{\text{additional dead load}} / W_{\text{deadweight}}$ is normally distributed with mean equal to 10% of the dead weight and a 25% coefficient of variation. For the two-bent overpass, in order to account for the variability in the axial load of the two bents, we assume the additional dead loads for bent 1, r_1, and for bent 2, r_2, to be statistically independent and identically distributed random variables with the same distribution as r.

Table 8.1.   List of the important variables for the systems considered.

Description	Parameter	Value/Distribution
Span length (right and left)	L [mm]	18,300
Center span length	L_C [mm]	44,200
Span-to-column height ratio	L/H , L/H_1	2.4
	L/H_2	2.0
Column-to-superstructure dimension ratio	D/D_s	0.75
Concrete nominal strength	f'_c [MPa]	$LN(27.6, 2.76)$
Reinforcement nominal yield strength	f_y [MPa]	$LN(448.2, 22.4)$
Column longitudinal reinforcement ratio	ρ_l	2.0%
Column transverse reinforcement ratio	ρ_s	0. 7%
Soil stiffness based on NEHRP groups (FEMA-273, 1996)	K_{soil}	B
Additional bridge dead load (as a ratio of the deadweight)	r, r_1, r_2	$IN(0.1, 0.025)$

The nonlinear pushover analyses in the transverse direction of the two structures needed to compute their deterministic deformation and shear demands, $\hat{d}_k$, $k = \delta, v$, are performed with PEER's OpenSees platform (McKenna, 2000) by using a nonlinear model developed by Mackie and Stojadinović (2001). In these models, the columns and pile shafts are modeled using a three-dimensional flexibility-based nonlinear beam-column element with fiber cross section. A simple

elastic-plastic material with a post-yield stiffness equal to 1.5% of the pre-yield stiffness is used to model all reinforcing steel. The confined concrete is modeled using the Kent-Scott-Park stress-strain relation (Kent and Park, 1971). Similarly to the analysis by Fenves and Ellery (1998), used in Chapter 6 as virtual experiments, the soil-structure interaction is modeled using bilinear springs located along the pile shaft length. The deck is modeled as a linear elastic beam with cracked section stiffness. In order to account for the opening and closing of the gap between the deck and the abutments, the abutments are modeled using nonlinear elastic-perfectly plastic spring-gap elements.

The probabilities of failure are computed by important sampling centered around the origin with the Reliability toolbox of OpenSees (Haukaas and Der Kiureghian, 2001) with an external limit-state function evaluator implemented in Matlab (1999).

The solid line in Figure 8.7 indicates the predictive fragility estimate, $\widetilde{F}(\mathbf{s})$, for the single-bent overpass defined as the conditional probability of attaining or exceeding the deformation or shear limit states (series system) for a given value of the spectral acceleration S_a. The dashed curve indicates the deformation fragility alone and the dotted curve indicates the shear fragility alone. Note that the deformation failure is dominant over the shear failure. This is consistent with the displacement-based capacity design approach used by Caltrans. Increasing S_a, we observe a more gradual increase of the failure probabilities than observed for the single-bent bridge (Figure 8.4). This difference is possibly due to the higher redundancy of the bridge structure due to the presence of the two abutments, which allows redistribution of the loads near failure.
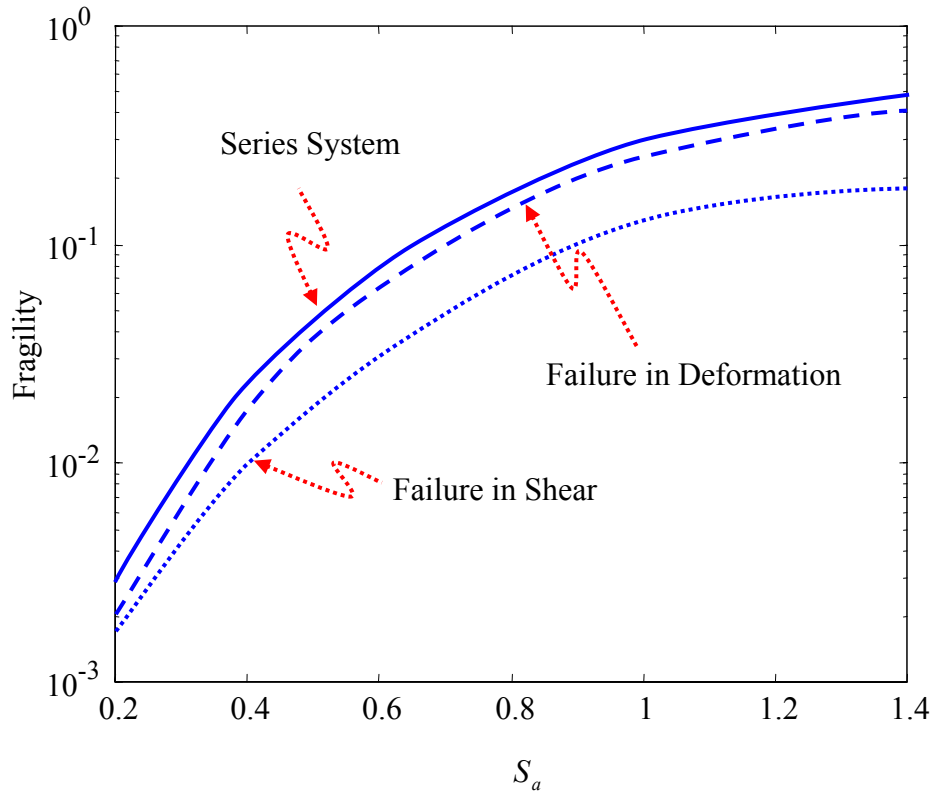
Figure 8.7.   Fragility estimates for the example single-bent overpass bridge.

Considering the second bridge configuration, the double-bent overpass, the solid line in Figure 8.8 (top chart) indicates the predictive fragility estimate, $\widetilde{F}(\mathbf{s})$, for the single-column bent 1, the dashed curve indicates the deformation fragility alone, and the dotted curve indicates the shear fragility alone. Similarly, Figure 8.8 (bottom chart) indicates the fragility estimates for the single-column bent 2. In a stand-alone configuration, the deformation demand of the taller bent would be larger than that of the shorter bent. However, as components of a structural system, the two bents have a dynamic interaction due to the connecting deck. Bent 2 drags bent 1 and forces it to displace more than it would if it were in the stand-alone configuration and, likewise, bent 1 holds bent 2 back so that its deformation demand is smaller than that in the stand-alone configuration. On the other hand, the column capacities are not modified by the dynamic interaction of the bents. As a consequence, for a given S_a, the probability of failure in

deformation of bent 2 is smaller then the corresponding probability of failure for bent 1. Finally, the solid line in Figure 8.9 indicates the predictive fragility estimate, $\widetilde{F}(\mathbf{s})$, for the bridge system defined as a series system of the two bents. The dashed curve indicates the deformation fragility alone of the bridge and the dotted curve indicates its shear fragility alone. Note that both at the component level and at the system level, the deformation failure mode dominates the fragility estimate, as intended by the design.
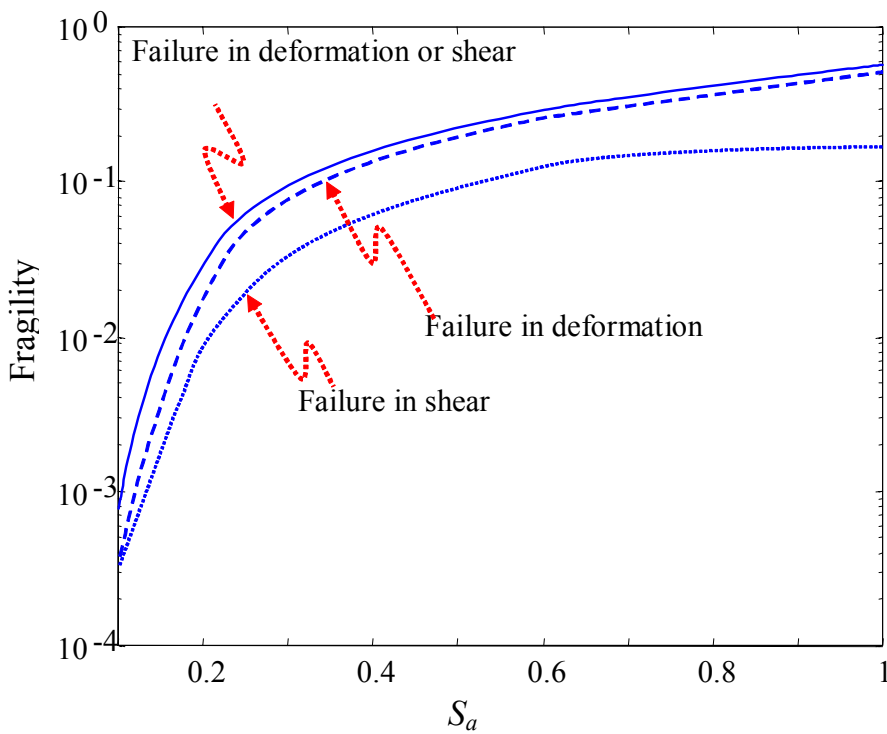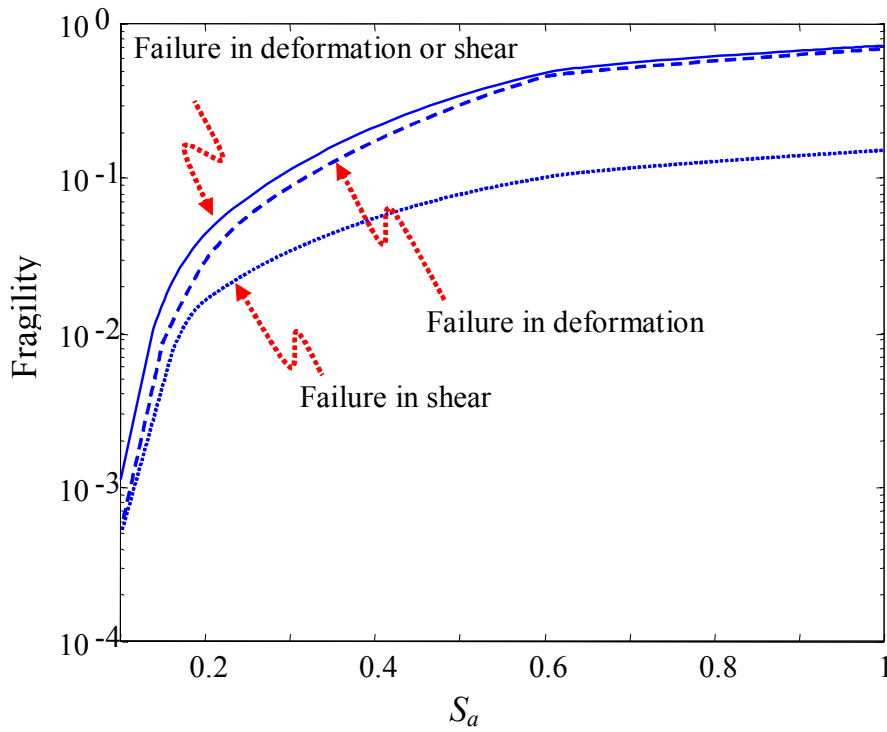
Figure 8.8. Fragility estimates for bent 1 (top) and bent 2 (bottom) for the example two-bent overpass bridge.
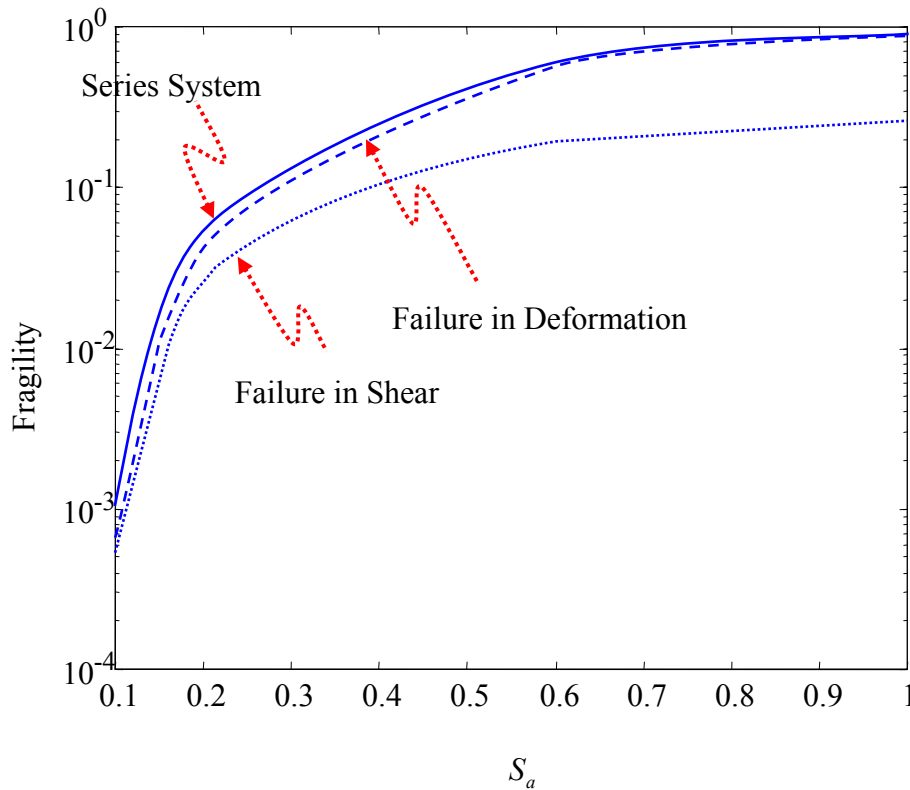
Figure 8.9.   Fragility estimates for the example two-bent overpass bridge.

8.5   SUMMARY

First, the probabilistic capacity models developed in Chapter 4 are used to estimate uni-variate fragility curves and bi-variate fragility contours for an example RC bridge column for given shear and deformation demands. Point estimates of the fragility curves based on predictive analysis are derived, together with confidence intervals on the fragilities that reflect the influence of epistemic uncertainties. Significant interaction between the two failure modes, particularly at high demand levels, is observed.

Then, the probabilistic demand models developed in Chapter 6 are combined with the probabilistic capacity models to construct limit-state functions that properly account for all the relevant uncertainties and are used to assess the structural fragilities of an example RC single-column bent for a given set of ground motion parameters. For the example bent the deformation

failure is dominant over the shear failure. This is consistent with the displacement-based capacity design approach used by Caltrans.

Finally, the analysis is extended to the fragility assessment of two configurations of typical new California highway overpass bridges. Fragilities are estimated both at the component level and at the system level. Consistent with the displacement-based capacity design approach used to design the bridge systems, the deformation failure mode dominates the fragility estimate.

9 Conclusions

9.1 SUMMARY OF MAJOR FINDINGS

A methodology is developed for constructing component and system fragility estimates that is based on solving reliability problems that involve the structural capacities at the component level and the corresponding demands due to an earthquake ground motion. The probability of failure of a component is computed as the probability that at least one demand measure is larger or equal to the corresponding capacity measure considering different modes of failure. Similarly, by means of the concepts of structural reliability, the probability of failure of a structural system is defined as the failure of one or more components.

A comprehensive Bayesian framework for constructing probabilistic capacity and demand models is formulated. The models correct the conservatism inherent in the deterministic models and explicitly account for the most relevant uncertainties: model errors arising from an inaccurate model form or missing variables, measurement errors, and statistical uncertainty. With the aim of facilitating their use in practice, the models are constructed by developing correction terms for existing deterministic models derived from first principles, e.g., rules of mechanics. The deterministic deformation capacity model used in this study is based on the notion of decomposing the total displacement of the RC column into its basic components. Specifically, the column displacement is considered to be composed of elastic and inelastic components, with the elastic component itself consisting of contributions from the flexural and shear deformations and from the slip of the longitudinal reinforcing bars. For the shear capacity model, owing to the complex nature of the underlying load transfer mechanisms, a unique consensus model does not exist. Here, two alternative deterministic models used in practice are considered and objective measures of their relative qualities are assessed. The deterministic procedure used to develop the demand models was proposed by Chopra and Goel (1999). The procedure is an improvement of

the ATC-40 (ATC, 1996) and FEMA-273 (FEMA, 1997) capacity-demand diagram methods, which use the well-known constant-ductility spectrum for the demand diagram. In this procedure, first, a nonlinear static analysis of the structure subjected to a monotonically increasing lateral load is performed (pushover analysis). The distribution of the lateral forces corresponds to an assumed displacement shape weighted by tributary masses. Then, an equivalent single-degree-of-freedom (SDOF) system with a bilinear force-displacement relationship is derived from the pushover curve of the structure. The deformation demand of the equivalent SDOF system is estimated by response spectrum analysis using inelastic spectra. Finally, the local seismic demands are determined by pushing the original structure to the maximum displacement determined in the previous step.

Through a model selection process that makes use of a set of "explanatory" functions, the terms that effectively correct the inherent conservatism in the existing model forms are identified. Moreover, the explanatory functions provide means to gain insight into the underlying behavioral phenomena and to select ground motion parameters that are most relevant to the seismic demands. Methods for assessing the unknown model parameters on the basis of observational data are described.

Although the methodology presented is aimed at developing probabilistic capacity and demand models, the approach is quite general and can be applied to the assessment of models (i.e., model selection and parameter estimation) in many engineering problems.

As the principal application of the methodology, probabilistic models for the deformation and shear capacities of RC circular columns subjected to cyclic loading are developed using existing experimental data for cyclically tested columns. Probabilistic models for the deformation and shear demands of RC bridges subjected to earthquake ground motions are developed using experimental observations on single-column bridge bents and virtual experiments on a bridge system. The conceptual idea of virtual experiments is introduced and the effectiveness of ground motion parameters in improving the predictive models is explored.

The probabilistic capacity and demand models are used in a formulation to assess the fragility of structural components and systems, with due consideration given to the different natures of aleatory and epistemic uncertainties. Point estimates of the fragility based on posterior estimates and predictive analyses, as well as confidence intervals on fragility that reflect the influence of epistemic uncertainties are presented.

First, the probabilistic capacity models developed are used to estimate uni-variate deformation and shear fragility curves and bi-variate deformation-shear fragility contours for an example RC bridge column. Significant interaction between the two failure modes, particularly at high demand levels, is observed. Point estimates of the fragility curves based on predictive analysis are derived together with confidence intervals on the fragilities that reflect the influence of epistemic uncertainties.

Then, the probabilistic demand models developed are combined with the probabilistic capacity models to construct limit-state functions that properly account for all the relevant uncertainties and are used to assess the structural fragilities of an example RC single-column bent for a given set of ground motion parameters. The deformation failure is found to be dominant over the shear failure. This is consistent with the displacement-based capacity design approach followed by Caltrans and used to design the example bent.

Finally, the analysis is extended to the fragility assessment of two configurations of typical new California highway overpass bridges. Fragilities are estimated both at the component level and at the system level. Consistent with the displacement-based capacity design approach used to design the bridge systems, the deformation failure modes dominate the fragility estimates both at the component and at the system level.

The developed probabilistic models are properly applicable only to components and structural systems that have geometry and material properties within the range of the observations used to assess the models. The normalized form that is used in the model formulation allows the extension of the applicability to cases in which at least the values of the normalized quantities that enter in the models are within the range of the observations. Application of these models to components and systems significantly different from the ones used for the model assessment (e.g., multi-column bents) is not appropriate without further investigation.

9.2   FURTHER STUDY

In order to increase the range of applicability of the models and to improve their quality, reducing the statistical uncertainty, more data should be collected and used in the model assessment. In particular tests on the system level and instrumentation of actual structural systems would provide additional data that would allow estimating the component demands accounting for the

system interaction without having to use virtual experiments that have an intrinsic model error. Furthermore, data for different structural configurations (e.g., multi-column bents) would broaden the applicability of the models. The Bayesian approach used in this work is perfectly suited to incorporate new information, as they become available. This can simply be done by updating the latest posterior distribution of the parameters.

The uncertain or random nature of phenomena that are of interest in the engineering practice make probabilistic analysis methods and statistics essential in decision-making and risk assessment. Furthermore, structural reliability and reliability/probabilistic-based design require experience both in engineering mechanics to properly model the structural behavior and limiting states and in statistics and probability in order to estimate and predict under conditions of uncertainty. The methodology presented here is general and can be applied to the assessment of models and limit states in many engineering problems.

In particular, in civil engineering statistics is a fundamental tool for constructing a probabilistic foundation for the development of performance-based guidelines, and for building a structure for coordinating, combining, and assessing the many considerations implicit in performance-based seismic assessment and design. Performance-based earthquake engineering (PBEE) implies design, evaluation, and construction of engineered facilities whose performance under common and extreme loads responds to the diverse needs and objectives of owner, users, and the society. PBEE is based on the premise that performance can be predicted and evaluated with sufficient confidence for the engineer and client jointly to make intelligent and informed decisions based on building life-cycle considerations rather than on construction costs alone. Implementation of such a design decision process necessitates a shift away from prescriptive codes and toward a design and assessment process more firmly rooted in the realistic prediction of structural behavior and load environment, with full account of the underlying uncertainties. This process implies a shift toward a probabilistic-oriented design approach, with an emphasis on accurate characterization and prediction that employs statistics more than it did in the past. In particular, a probabilistic approach is essential for a consistent treatment of risk and uncertainties, for the identification and quantification of performance parameters, for the assessment of uncertainties in demands and capacities, for addressing life-cycle cost issues, for the formulation of engineering limit states, and for the fragility assessment of structural components and systems.

Together with the increasing need of reliability/probabilistic-based design and analysis methods in engineering, the advent of new technologies that enable engineers to collect, manipu-

late, and display extensive data with minimal human intervention will drastically increase the diffusion and implementation of probabilistic methods and statistics in the future.

REFERENCES

Applied Technology Council (1985). *Earthquake damage evaluation data for California*. Applied Technology Council Report ATC-13, Redwood City, California.

Applied Technology Council (1996). *Improved seismic design criteria for California bridges*. Applied Technology Council Report ATC-32, Palo Alto, California.

Applied Technology Council (1996). *Seismic evaluation and retrofit of concrete buildings*. Applied Technology Council Report ATC-40, Redwood City, California.

Arici, Y., and Mosalam, K.M. (2000). System identification and modeling of bridge systems for assessing current design procedures. *Proceedings of SMIP2000 Seminar*. Sacramento, California: 77–95.

ASCE-ACI Joint Task Committee 426 (1973). Shear strength of reinforced concrete members. *Journal of Structural Engineering*, ASCE, **99** (6): 1091–1187.

Bradley, P.C., and Thomas, A.L. (1998). *Bayes and empirical Bayes methods for data analysis*. Chapman and Hall, Boca Raton, Florida.

Basoz, N., and Kiremidjian, A.S. (1997). Risk assessment of bridges and highway systems from the Northridge earthquake. *Proceedings of the National Seismic Conference on Bridges and Highways: "Progress in Research and Practice."* Sacramento, California: 65–79.

Box, G.E.P., and Cox, D.R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society*, Series B, **26**: 211–52.

Box, G.E.P., and Tiao, G.C. (1992). *Bayesian inference in statistical analysis*. Wiley, New York.

Caltrans (1999). *Seismic design criteria*. California Department of Transportation. 1.1 Edition.

Cano, R. (1992). *On the efficiency of the Bayesian bootstrap*. Department of Statistics, Stockholm University. Stockholm, Sweden.

Casciati, F., and Faravelli, L. (1991) *Fragility analysis of complex structural systems*. Research Studies Press Ltd., United Kingdom.

Chernick, M.R. (1999). *Bootstrap methods, a practitioner's guide*. Wiley, New York, New York.

Chopra, A.K., and Goel, R.K. (1999). *Capacity-demand-diagram methods for estimating seismic deformation of inelastic structures: SDF systems*. Report Number PEER-1999/02, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Chopra, A.K., and Goel, R.K. (2001). *A modal pushover analysis procedure to estimate seismic demands for buildings: theory and preliminary evaluation*. Report Number PEER-2001/03, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Davison, A. C., and Hinkley, D.V. (1997). *Bootstrap methods and their application.* Cambridge University Press, New York.

Der Kiureghian, A. (1989). Measures of structural safety under imperfect states of knowledge. *Journal of Structural Engineering*, ASCE, **115**(5): 1119–40.

Der Kiureghian, A. (1999). A Bayesian framework for fragility assessment. *Proceedings of the ICASP8 Conference*. Sidney, Australia: 1003–9.

DIANA user's manual, version 7.1 (1999). TNO Building and Construction Research. Delft, The Netherlands.

Ditlevsen, O., and Madsen, H. O. (1996). *Structural reliability methods*. J. Wiley, New York.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statistics,* **7**: 231–46.

Efron, B. (1982). *The Jackknife, the Bootstrap, and other Resampling Plans*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Efron, B., and Tibshirani, R.J. (1993). *An introduction to the bootstrap*. Chapman & Hall, New York.

Fajfar, P. (2000). A nonlinear analysis method for performance-based seismic design. *Earthquake Spectra,* **16**(3): 573–92.

Fajfar, P., and Fischinger, M. (1987). Non-linear seismic analysis of RC buildings: implications of a case study. *European Earthquake Engineering*, **1**: 31-43.

Fajfar, P., Fischinger, M. (1988). N2--a method for non-linear seismic analysis of regular buildings. *Proceedings of the 9th World Conference on Earthquake Engineering*. Tokyo, Japan, **V**: 111–16.

Fajfar, P., Gašperšič, P. and Drobnič, D., (1997). A simplified nonlinear method for seismic damage analysis of structures. *Seismic design methodologies for the next generation of codes*. P. Fajfar and H. Krawinkler, Eds., Balkema, Rotterdam, The Netherlands: 183–94.

Fenves, G.L., and Ellery, M. (1998). *Behavior and failure analysis of a multiple-frame highway bridge in the 1994 Northridge earthquake*. Report Number PEER-1998/08, Berkeley: Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

FEMA 273/FEMA274, NEHRP (1997). *Guidelines for the seismic rehabilitation of buildings*, FEMA 273; and *Commentary on the guidelines for the seismic rehabilitation of buildings*, FEMA 274, *Federal Emergency Management Agency*. Washington, D.C.

Fisher, R. A. (1922). On the mathematical formulation of theoretical statistics. *Phil. Trans. Roy. Soc., Series A* **222**: 309.

Fisher, R. A. (1925). Theory of statistical estimation. *Proc. Camb. Phil. Soc.,* **22**: 700.

Freeman, S. A., Nicoletti, J. P., and Tyrell, J. V. (1975). Evaluations of existing buildings for seismic risk: A case study of Puget Sound Naval Shipyard, Bremerton, Washington. *Proceedings of 1st U.S. National Conference on Earthquake Engineering, EERI*. Berkeley, California: 113-122.

Fukushima, S., Kai, Y., and Yashiro, K. (1996). Study on the fragility of system — Part 1: structure with brittle elements in its stories. *Proceedings of the 11th World Conference on*

Earthquake Engineering. Pergamon, Elsevier Science Ltd., Oxford, United Kingdom, **1**, Paper No. 333.

Gardoni P., Der Kiureghian A., Mosalam K. M. (2002). Probabilistic capacity models and fragility estimates for RC columns based on experimental observations. *Journal of Engineering Mechanics,* ASCE, accepted for publication.

Geyskens, P., Der Kiureghian, A., and Monteiro, P. (June 1993). BUMP: *Bayesian updating of model parameters.* Report UCB/SEMM-93/06, Structural Engineering, Mechanics and Materials, Department of Civil Engineering. University of California, Berkeley, California.

Gelman, A., Carlin, J.B., Stern, H. S., and Rubin, D.B. (1998). *Bayesian data analysis*. Chapman & Hall, London, United Kingdom.

Goel, R. K., and Chopra, A. K. (1997). Evaluation of bridge abutment capacity and stiffness during earthquakes. *Earthquake Spectra*, **13**(1): 1–23.

Haukaas, T., and Der Kiureghian, A. (2001). A computer program for nonlinear finite element reliability analysis. *Proceeding of the 8th International Conference on Structural Safety and Reliability*. Newport Beach, California.

Horton, S. P., and Barstow, N. (1995). *Simulation of ground motion at the I5/Route 14 Interchange due to the 1994 Northridge, California earthquake*. Report Number NCEER-95-XXXX, National Center of Earthquake Engineering Research, Buffalo, New York. (Draft Report.)

Hutchings, L., and Jarpe, S. (1996). Ground-motion variability at the Highway 14 and I-5 interchange in the northern San Fernando Valley. *Bulletin of the Seismological Society of America*, **86**(1): S289-S299.

Hwang, H.H.M., and Huo, J-R. (1994). *Generation of hazard-consistent fragility curves for seismic loss estimation studies.* Technical Report NCEER-94-0015, National Center for Earthquake Engineering Research, State University of New York at Buffalo, Buffalo, New York.

Johnson, R.A. (1967). An asymptotic expansion for posterior distributions. *Ann. Math. Statist*. **38**: 1899.

Johnson, R.A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist*. **41**: 851.

Kai, Y., and Fukushima, S. (1996). Study on the fragility of system — Part 2: system with ductile elements in its stories. *Proceedings of the 11th World Conference on Earthquake Engineering*. Pergamon, Elsevier Science Ltd., Oxford, UK, **1**, Paper No. 334.

Kent, D.C., and Park, R. (1971). Flexural members with confined concrete. *Journal of Structural Engineering*, **97**(7): 1969–90.

Kramer, S. L. (1996). *Geotechnical earthquake engineering*. Prentice Hall. Upper Saddle River, New Jersey.

Karim, K.R., and Yamazaki, F. (2001). Effect of earthquake ground motions on fragility curves of highway bridge piers based on numerical simulation. *Earthquake Engineering & Structural Dynamics*, **30**(12): 1839–56.

Krawinkler, H., and Nassar, A.A. (1992). Seismic design based on ductility and cumulative damage demands and capacities. *Nonlinear Seismic Analysis and Design of Reinforced Concrete Buildings*. Eds P. Fajfar and H. Krawinkler, Elsevier Applied Science, New York.

Liu, P.-L., and Der Kiureghian, A. (1986). Multi-variate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, **1**(2): 105–12.

Liu, P.-L., Lin, H.-Z., and Der Kiureghian, A. (1989). *CalREL user's manual*. Report Number UCB/SEMM-9/8, Department of Civil and Environmental Engineering, University of California, Berkeley, California.

Lynn, A.C., Moehle, J.P., and Mahin, S.A. (1996). Seismic evaluation of existing reinforced concrete building columns. *Earthquake Spectra*, **12**(4): 715–39.

Mackie, K., and Stojadinović, B. (2001). Probabilistic seismic demand model for California Highway Bridges. *Journal of Bridge Engineering*, ASCE, **6**(6): 468–81.

Mander, J.B., and Basoz, N. (1999). Seismic fragility curve theory for highway bridges. *Proceedings of the 5th U.S. Conference on Lifeline Earthquake Engineering*. Reston, Virginia: 31–40.

Mander, J.B., Priestley, M.J.N., and Park, R. (1988). Theoretical stress-strain model for confined concrete. *Journal of Structural Division,* ASCE, **114**(8): 1804–26.

Matlab reference manual; version 5.3 (1999). MathWorks, Inc., Natick, Massachusetts.

McKenna, F., and Fenves, G.L. (2000). An object-oriented software design for parallel structural analysis. *Proceedings of the 2000 Structures Congress & Exposition Advanced Technology in Structural Engineering*. Philadelphia, Pennsylvania.

Moehle, J.P., Elwood, K., and Sezen, H. (2000). Shear failure and axial load collapse of existing reinforced concrete columns. *Proceedings of the 2nd U.S.-Japan Workshop on Performance-Based Design Methodology for Reinforced Concrete Building Structures*. Sapporo, Japan: 241-255.

Moehle, J.P., Lynn, A.C., Elwood, K., and Sezen, H. (1999). Gravity load collapse of reinforced concrete frames during earthquakes. *Proceedings of the 1st U.S.-Japan Workshop on Performance-Based Design Methodology for Reinforced Concrete Building Structures*. Maui, Hawaii: 175–89.

Mosteller, F., and Wallance, D.L. (1964). *Inference and disputed authorship: The federalist*. Addison-Wesley, Reading, Massachusetts.

Nelder, J.A. (1977). A reformulation of linear models (with discussion). *Journal of the Royal Statistical Society*. A **140**: 48–76.

Naito, C.J. (2000). *Experimental and computational evaluation of reinforced concrete beam-column connections for seismic performance*. Ph.D. Dissertation, University of California, Berkeley.

Paret, T. E., Sasaki, K. K., Eilbekc, D. H., and Freeman, S. A. (1996). Approximate inelastic procedures to identify failure mechanisms form higher mode effects. Paper No 966, *11th Word Conference on Earthquake Engineering*. Acapulco, Mexico.

Park, Y.-J., and Ang, A. (1985). Mechanistic seismic damage model for reinforced concrete. *Journal of Structural Engineering*, ASCE, **111**(4): 722-739.

Park, R., and Paulay, T. (1975). *Reinforced concrete structures*. John Wiley & Sons, Inc., New York.

Priestley, M.J.N., Seible, F., and Calvi, G.M. (1996). *Seismic design and retrofit of bridges*. John Wiley & Sons, Inc., New York.

Priestley, M.J.N., and Park, R. (1987). Strength and ductility of concrete bridge columns under seismic loading. *ACI Structural Journal*, **84**(1): 61–76.

Pujol, S., Ramfrez, J.A., and Sozen, M. A. (1999). Drift capacity of reinforced concrete columns subjected to cyclic shear reversals. *Seismic response of concrete bridges,* ACI International. Farmington Hills, Michigan: 255-274.

Rao, C. R., and Toutenburg, H. (1997). *Linear models, least squares and alternatives*. Springer, New York.

Richards, F.S.G. (1961). A method of maximum likelihood estimation. *Journal of the Royal Statistical Society,* **23**: 469-475.

Singhal, A., and Kiremidjian, A. (1998). Bayesian updating of fragilities with application to RC frames. *Journal of Structural Engineering*, ASCE, **124**(8): 922-929.

Shinozuka, M., Feng, M.Q., Kim, H.-K., and Kim, S.-H. (2000a). Nonlinear static procedure for fragility curve development. *Journal of Engineering Mechanics,* ASCE, **126**(12): 1287-1295.

Shinozuka, M., Feng, M.Q., Lee, J., and Naganuma, T. (2000b). Statistical analysis of fragility curves. *Journal of Engineering Mechanics*, ASCE, **126**(12): 1224-1231.

Stone, J. C. (1996). *A course in probability and statistics*. Duxbury Press, Belmont, California.

Tanaka, S., Kameda, H., Nojima, N., and Ohnishi, S. (2000). Evaluation of seismic fragility for highway transportation systems. *Proceedings of the 12th World Conference on Earthquake Engineering*. Upper Hutt, New Zealand: Paper No. 0546.

Thewalt, C. R., and Stojadinović, B. (1994). Stable reinforced concrete section analysis procedure. *Journal of Structural Engineering,* ASCE, **120**(10): 3012-3024.

Trifunac, M.D., and Brady, A.G. (1975). On the correlation of seismic intensity with peaks of recorded strong ground motion. *Bulletin of Seismological Society of America,* **65**: 139-162.

Watson, S., and Park, R. (1994). Simulated seismic load tests on reinforced concrete columns. *Journal of Structural Engineering,* ASCE, **120**(6): 1825-1849.

Williams, M.S., and Sexsmith, R.G. (1995). Seismic damage indices for concrete structures: a state-of-the-art review. *Earthquake Spectra*, **11**(2): 319-349.

Yamazaki, F., Motomura, H., and Hamada, T. (2000). Damage assessment of expressway networks in Japan based on seismic monitoring. *Proceeding of the 12th World Conference on Earthquake Engineering*. Upper Hutt, New Zealand, Paper No. 0551.

Yashinsky, M., and Ostrom, T. (2000). Caltrans' new seismic design criteria for bridges. *Earthquake Spectra*, **16**(1): 285–307.

Appendix Invertible Transformations

After a brief description on invertible transformations, the probability density function in (2.42) is derived from (2.41).

Following Stone (1996), we start by considering a vector of random variables $\mathbf{X} = [X_1,\ldots,X_q]^T$, distributed with probability density function $p_{\mathbf{X}}$ on $\mathbb{R}^q$, and an open set $\mathbb{X}$ in $\mathbb{R}^q$ such that $p_{\mathbf{X}} = 0$ on its complement $\mathbb{X}^c$. Let $\mathbf{g} = [g_1,\ldots,g_q]^T$ where g_k, $k = 1,\ldots,q$, are real-valued functions on $\mathbb{X}$ and let $\mathbf{Y} = [Y_1,\ldots,Y_q]^T$ be the transform of $\mathbf{X}$ defined by $\mathbf{Y} = \mathbf{g}(\mathbf{X})$, so that $Y_k = g_k(\mathbf{X})$ for $k = 1,\ldots,q$. Then, under appropriate conditions on $\mathbf{g}$, $\mathbf{Y}$ has a probability density function that can be expressed explicitly in terms of $\mathbf{g}$ and $p_{\mathbf{X}}$.

As a first assumption, let $\mathbf{g}$ be a one-to-one mapping from $\mathbb{X}$ onto an open set $\mathbb{Y}$ in $\mathbb{R}^q$ (i.e., for each outcome $\mathbf{x}$ of $\mathbf{X}$ there is one and only one outcome $\mathbf{y}$ of $\mathbf{Y}$), so $\mathbf{g}$ has an inverse mapping $\mathbf{f} = \mathbf{g}^{-1} = [f_1,\ldots,f_q]$ from $\mathbb{Y}$ to $\mathbb{X}$.

Second, assuming that $\mathbf{g}$ is continuously differentiable on $\mathbb{X}$, that is, for $1 \leq k,l \leq q$, the partial derivative of $y_k = g_k(\mathbf{x})$ with respect to x_l exists for $x \in \mathbb{X}$ and is a continuous function of $\mathbf{x}$. These partial derivatives define the *derivative matrix* of $\mathbf{g}$ at $\mathbf{x}$ as

$$
\begin{bmatrix}
\dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_q} \\
\vdots & & \vdots \\
\dfrac{\partial y_q}{\partial x_1} & \cdots & \dfrac{\partial y_q}{\partial x_q}
\end{bmatrix}
\qquad\qquad\qquad (A.1)
$$

The *Jacobian* of $\mathbf{g}$ at $\mathbf{x}$ is then the determinant of the derivative matrix of $\mathbf{g}$ at $\mathbf{x}$

$$
\frac{\partial(y_1,\ldots,y_q)}{\partial(x_1,\ldots,x_q)} =
\begin{vmatrix}
\dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_q} \\
\vdots & & \vdots \\
\dfrac{\partial y_q}{\partial x_1} & \cdots & \dfrac{\partial y_q}{\partial x_q}
\end{vmatrix}
\qquad\qquad (A.2)
$$

Thirdly, we assume that the derivative matrix of $\mathbf{g}$ is invertible everywhere on X, that is, its Jacobian is nonzero everywhere on X. It follows that $\mathbf{f}$ is continuously differentiable on Y, the derivative matrix of $\mathbf{f}$ at $\mathbf{y}$

$$
\begin{bmatrix}
\dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_q} \\
\vdots & & \vdots \\
\dfrac{\partial x_q}{\partial y_1} & \cdots & \dfrac{\partial x_q}{\partial y_q}
\end{bmatrix}
\qquad\qquad\qquad (A.3)
$$

is the inverse of the derivative matrix of $\mathbf{g}$ at $\mathbf{x} = \mathbf{f}(\mathbf{y})$ (by using the chain rule) and the Jacobian of $\mathbf{f}$ at $\mathbf{y}$

$$
\frac{\partial(x_1,\ldots,x_q)}{\partial(y_1,\ldots,y_q)}
\qquad\qquad\qquad (A.4)
$$

is the reciprocal of the Jacobian of $\mathbf{g}$ at $\mathbf{x} = \mathbf{f}(\mathbf{y})$.

Under these three assumptions on $\mathbf{g}$ (or the equivalent assumptions on the inverse mapping $\mathbf{f}$), the transformation of $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ has the density function on $\mathbb{R}^q$ given by

$$p_{\mathbf{Y}}(\mathbf{y}) = \left| \frac{\partial (x_1,\ldots,x_q)}{\partial (y_1,\ldots,y_q)} \right| p_X(\mathbf{f}(\mathbf{y})) \tag{A.5}$$

and $p_{\mathbf{Y}}(\mathbf{y}) = 0$ for $\mathbf{y} \notin \mathbb{Y}$, where $|\ |$ indicates the absolute value. A rigorous proof of this result can be found in Stone (1996).

Let $\mathbf{x} = [\sigma_1^2,\ldots,\sigma_q^2,\rho_{12}\sigma_1\sigma_2,\ldots,\rho_{1q}\sigma_1\sigma_q,\ldots,\rho_{q-1q}\sigma_{q-1}\sigma_q]^T$ have probability density function $p_X \propto |\mathbf{\Sigma}|^{-(q+1)/2}$ (see equation (2.41)). Introducing the diagonal matrix of standard deviations, $\mathbf{D}$, and the correlation coefficient matrix, $\mathbf{R}$, as

$$\mathbf{D} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \sigma_q \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} 1 & 0 & & \text{sym.} \\ \rho_{21} & 1 & & 0 \\ \vdots & \vdots & \ddots & \\ \rho_{q1} & \rho_{q2} & & 1 \end{bmatrix} \tag{A.6}$$

$|\mathbf{\Sigma}|$ can be written as

$$|\mathbf{\Sigma}| = |\mathbf{DRD}| = |\mathbf{D}||\mathbf{R}||\mathbf{D}| \tag{A.7}$$

Therefore, $\mathbf{X}$ can be obtained as a transform of $\mathbf{Y} = [\sigma_1,\ldots,\sigma_q,\rho_{12},\ldots,\rho_{1q},\ldots,\rho_{q-1q}]^T$ that satisfies the three assumptions listed above. Hence, (A.5) and (A.7) can be used to obtain the density function $p_{\mathbf{Y}}(\mathbf{y})$ on $\mathbb{R}^q$ as

$$p_{\mathbf{Y}}(\mathbf{y}) = \left| \frac{\partial \left(\sigma_1^2,\ldots,\sigma_q^2,\rho_{12}\sigma_1\sigma_2,\ldots,\rho_{1q}\sigma_1\sigma_q,\ldots,\rho_{q-1q}\sigma_{q-1}\sigma_q \right)}{\partial \left(\sigma_1,\ldots,\sigma_q,\rho_{12},\ldots,\rho_{1q},\ldots,\rho_{q-1q} \right)} \right| \left(|\mathbf{D}||\mathbf{R}||\mathbf{D}| \right)^{-(q+1)/2} \tag{A.8}$$

and $p_{\mathbf{Y}}(\mathbf{y}) = 0$ for $\mathbf{y} \notin \mathbb{Y}$. Since the derivative matrix of the transformation at $\mathbf{x}$ is a lower triangular matrix, it follows that the Jacobian in (A.8) is simply the product of its diagonal elements $(2\sigma_1,\ldots,2\sigma_q,\sigma_1\sigma_2,\ldots,\sigma_1\sigma_q,\ldots,\sigma_{q-1}\sigma_q)$ that is

$$\left| \frac{\partial \left(\sigma_1^2,\ldots,\sigma_q^2,\rho_{12}\sigma_1\sigma_2,\ldots,\rho_{1q}\sigma_1\sigma_q,\ldots,\rho_{q-1q}\sigma_{q-1}\sigma_q \right)}{\partial \left(\sigma_1,\ldots,\sigma_q,\rho_{12},\ldots,\rho_{1q},\ldots,\rho_{q-1q} \right)} \right| = 2^q \sigma_1^q \cdots \sigma_q^q \tag{A.9}$$

For the right-hand side of (A.9), the following proportionality holds

$$2^q \sigma_1^q \cdots \sigma_q^q \propto |\mathbf{D}|^q \qquad (A.10)$$

Substituting (A.10) in (A.8), the probability density of $p_\mathbf{Y}(\mathbf{y})$ is given by

$$p_\mathbf{Y}(\mathbf{y}) \propto |\mathbf{D}|^q \left(|\mathbf{D}||\mathbf{R}||\mathbf{D}|\right)^{-(q+1)/2} = |\mathbf{R}|^{-(q+1)/2}|\mathbf{D}|^{-1} = |\mathbf{R}|^{-(q+1)/2} \prod_{i=1}^{q} \frac{1}{\sigma_i} \qquad (A.11)$$

According to the above result, we conclude that the non-informative prior density in (2.42) expressed in terms of variances and covariances is equivalent to the density function in (2.41) when working with standard deviations and correlation coefficients.

PEER REPORTS

PEER reports are available from the National Information Service for Earthquake Engineering (NISEE). To order PEER reports, please contact the Pacific Earthquake Engineering Research Center, 1301 South 46th Street, Richmond, California 94804-4698. Tel.: (510) 231-9468; Fax: (510) 231-9461.

PEER 2002/13 *Probabilistic Models and Fragility Estimates for Bridge Components and Systems.* Paolo Gardoni, Armen Der Kiureghian, and Khalid M. Mosalam. June 2002.

PEER 2002/11 Analytical and Experimental Study of Fiber-Reinforced Strip Isolators. *James M. Kelly and Shakzhod M. Takhirov.* September 2002.

PEER 2002/08 *Component Testing, Stability Analysis and Characterization of Buckling-Restrained Unbonded Braces^{TM}.* Cameron Black, Nicos Makris, and Ian Aiken. September 2002.

PEER 2002/07 *Seismic Performance of Pile-Wharf Connections.* Charles W. Roeder, Robert Graff, Jennifer Soderstrom, and Jun Han Yoo. December 2001.

PEER 2002/06 *The Use of Benefit-Cost Analysis for Evaluation of Performance-Based Earthquake Engineering Decisions.* Richard O. Zerbe and Anthony Falit-Baiamonte. September 2001.

PEER 2002/05 *Guidelines, Specifications, and Seismic Performance Characterization of Nonstructural Building Components and Equipment.* André Filiatrault, Constantin Christopoulos, and Christopher Stearns. September 2001.

PEER 2002/03 *Investigation of Sensitivity of Building Loss Estimates to Major Uncertain Variables for the Van Nuys Testbed.* Keith A. Porter, James L. Beck, and Rustem V. Shaikhutdinov. August 2002.

PEER 2002/02 *The Third U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures.* July 2002.

PEER 2002/01 *Nonstructural Loss Estimation: The UC Berkeley Case Study. Mary C. Comerio and John C. Stallmeyer.* December 2001.

PEER 2001/16 *Statistics of SDF-System Estimate of Roof Displacement for Pushover Analysis of Buildings.* Anil K. Chopra, Rakesh K. Goel, and Chatpan Chintanapakdee. December 2001.

PEER 2001/15 *Damage to Bridges during the 2001 Nisqually Earthquake.* R. Tyler Ranf, Marc O. Eberhard, and Michael P. Berry. November 2001.

PEER 2001/14 *Rocking Response of Equipment Anchored to a Base Foundation.* Nicos Makris and Cameron J. Black. September 2001.

PEER 2001/13 *Modeling Soil Liquefaction Hazards for Performance-Based Earthquake Engineering.* Steven L. Kramer and Ahmed-W. Elgamal. February 2001.

PEER 2001/12 *Development of Geotechnical Capabilities in OpenSees.* Boris Jeremic. September 2001.

PEER 2001/11   *Analytical and Experimental Study of Fiber-Reinforced Elastomeric Isolators*. James M. Kelly and Shakhzod M. Takhirov. September 2001.

PEER 2001/10   *Amplification Factors for Spectral Acceleration in Active Regions.* Jonathan P. Stewart, Andrew H. Liu, Yoojoong Choi, and Mehmet B. Baturay. December 2001.

PEER 2001/09   *Ground Motion Evaluation Procedures for Performance-Based Design.* Jonathan P. Stewart, Shyh-Jeng Chiou, Jonathan D. Bray, Robert W. Graves, Paul G. Somerville, and Norman A. Abrahamson. September 2001.

PEER 2001/08   *Experimental and Computational Evaluation of Reinforced Concrete Bridge Beam-Column Connections for Seismic Performance*. Clay J. Naito, Jack P. Moehle, and Khalid M. Mosalam. November 2001.

PEER 2001/07   *The Rocking Spectrum and the Shortcomings of Design Guidelines.* Nicos Makris and Dimitrios Konstantinidis. August 2001.

PEER 2001/06   *Development of an Electrical Substation Equipment Performance Database for Evaluation of Equipment Fragilities*. Thalia Agnanos. April 1999.

PEER 2001/05   *Stiffness Analysis of Fiber-Reinforced Elastomeric Isolators.* Hsiang-Chuan Tsai and James M. Kelly. May 2001.

PEER 2001/04   *Organizational and Societal Considerations for Performance-Based Earthquake Engineering.* Peter J. May. April 2001.

PEER 2001/03   *A Modal Pushover Analysis Procedure to Estimate Seismic Demands for Buildings: Theory and Preliminary Evaluation.* Anil K. Chopra and Rakesh K. Goel. January 2001.

PEER 2001/02   *Seismic Response Analysis of Highway Overcrossings Including Soil-Structure Interaction.* Jian Zhang and Nicos Makris. March 2001.

PEER 2001/01   *Experimental Study of Large Seismic Steel Beam-to-Column Connections.* Egor P. Popov and Shakhzod M. Takhirov. November 2000.

PEER 2000/10   *The Second U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures.* March 2000.

PEER 2000/09   *Structural Engineering Reconnaissance of the August 17, 1999 Earthquake: Kocaeli (Izmit), Turkey.* Halil Sezen, Kenneth J. Elwood, Andrew S. Whittaker, Khalid Mosalam, John J. Wallace, and John F. Stanton. December 2000.

PEER 2000/08   *Behavior of Reinforced Concrete Bridge Columns Having Varying Aspect Ratios and Varying Lengths of Confinement.* Anthony J. Calderone, Dawn E. Lehman, and Jack P. Moehle. January 2001.

PEER 2000/07   *Cover-Plate and Flange-Plate Reinforced Steel Moment-Resisting Connections.* Taejin Kim, Andrew S. Whittaker, Amir S. Gilani, Vitelmo V. Bertero, and Shakhzod M. Takhirov. September 2000.

PEER 2000/06   *Seismic Evaluation and Analysis of 230-kV Disconnect Switches.* Amir S. J. Gilani, Andrew S. Whittaker, Gregory L. Fenves, Chun-Hao Chen, Henry Ho, and Eric Fujisaki. July 2000.

PEER 2000/05   *Performance-Based Evaluation of Exterior Reinforced Concrete Building Joints for Seismic Excitation.* Chandra Clyde, Chris P. Pantelides, and Lawrence D. Reaveley. July 2000.

PEER 2000/04   *An Evaluation of Seismic Energy Demand: An Attenuation Approach.* Chung-Che Chou and Chia-Ming Uang. July 1999.

PEER 2000/03   *Framing Earthquake Retrofitting Decisions: The Case of Hillside Homes in Los Angeles.* Detlof von Winterfeldt, Nels Roselund, and Alicia Kitsuse. March 2000.

PEER 2000/02   *U.S.-Japan Workshop on the Effects of Near-Field Earthquake Shaking.* Andrew Whittaker, ed. July 2000.

PEER 2000/01   *Further Studies on Seismic Interaction in Interconnected Electrical Substation Equipment.* Armen Der Kiureghian, Kee-Jeung Hong, and Jerome L. Sackman. November 1999.

PEER 1999/14   *Seismic Evaluation and Retrofit of 230-kV Porcelain Transformer Bushings*. Amir S. Gilani, Andrew S. Whittaker, Gregory L. Fenves, and Eric Fujisaki. December 1999.

PEER 1999/13   *Building Vulnerability Studies: Modeling and Evaluation of Tilt-up and Steel Reinforced Concrete Buildings.* John W. Wallace, Jonathan P. Stewart, and Andrew S. Whittaker, editors. December 1999.

PEER 1999/12   *Rehabilitation of Nonductile RC Frame Building Using Encasement Plates and Energy-Dissipating Devices.* Mehrdad Sasani, Vitelmo V. Bertero, James C. Anderson. December 1999.

PEER 1999/11   *Performance Evaluation Database for Concrete Bridge Components and Systems under Simulated Seismic Loads.* Yael D. Hose and Frieder Seible. November 1999.

PEER 1999/10   *U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures.* December 1999.

PEER 1999/09   *Performance Improvement of Long Period Building Structures Subjected to Severe Pulse-Type Ground Motions*. James C. Anderson, Vitelmo V. Bertero, and Raul Bertero. October 1999.

PEER 1999/08   *Envelopes for Seismic Response Vectors.* Charles Menun and Armen Der Kiureghian. July 1999.

PEER 1999/07   *Documentation of Strengths and Weaknesses of Current Computer Analysis Methods for Seismic Performance of Reinforced Concrete Members.* William F. Cofer. November 1999.

PEER 1999/06   *Rocking Response and Overturning of Anchored Equipment under Seismic Excitations.* Nicos Makris and Jian Zhang. November 1999.

PEER 1999/05   *Seismic Evaluation of 550 kV Porcelain Transformer Bushings.* Amir S. Gilani, Andrew S. Whittaker, Gregory L. Fenves, and Eric Fujisaki. October 1999.

PEER 1999/04   *Adoption and Enforcement of Earthquake Risk-Reduction Measures.* Peter J. May, Raymond J. Burby, T. Jens Feeley, and Robert Wood.

PEER 1999/03   *Task 3 Characterization of Site Response General Site Categories*. Adrian Rodriguez-Marek, Jonathan D. Bray, and Norman Abrahamson. February 1999.

PEER 1999/02   *Capacity-Demand-Diagram Methods for Estimating Seismic Deformation of Inelastic Structures: SDF Systems*. Anil K. Chopra and Rakesh Goel. April 1999.

PEER 1999/01   *Interaction in Interconnected Electrical Substation Equipment Subjected to Earthquake Ground Motions.* Armen Der Kiureghian, Jerome L. Sackman, and Kee-Jeung Hong. February 1999.

PEER 1998/08   *Behavior and Failure Analysis of a Multiple-Frame Highway Bridge in the 1994 Northridge Earthquake.* Gregory L. Fenves and Michael Ellery. December 1998.

PEER 1998/07   *Empirical Evaluation of Inertial Soil-Structure Interaction Effects.* Jonathan P. Stewart, Raymond B. Seed, and Gregory L. Fenves. November 1998.

PEER 1998/06   *Effect of Damping Mechanisms on the Response of Seismic Isolated Structures.* Nicos Makris and Shih-Po Chang. November 1998.

PEER 1998/05   *Rocking Response and Overturning of Equipment under Horizontal Pulse-Type Motions.* Nicos Makris and Yiannis Roussos. October 1998.

PEER 1998/04   *Pacific Earthquake Engineering Research Invitational Workshop Proceedings, May 14–15, 1998: Defining the Links between Planning, Policy Analysis, Economics and Earthquake Engineering.* Mary Comerio and Peter Gordon. September 1998.

PEER 1998/03   *Repair/Upgrade Procedures for Welded Beam to Column Connections.* James C. Anderson and Xiaojing Duan. May 1998.

PEER 1998/02   *Seismic Evaluation of 196 kV Porcelain Transformer Bushings*. Amir S. Gilani, Juan W. Chavez, Gregory L. Fenves, and Andrew S. Whittaker. May 1998.

PEER 1998/01   *Seismic Performance of Well-Confined Concrete Bridge Columns.* Dawn E. Lehman and Jack P. Moehle. December 2000.