Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations

Article *in* Journal of Engineering Mechanics · October 2002

DOI: 10.1061/(ASCE)0733-9399(2002)128:10(1024)

CITATIONS
477

READS
2,807

3 authors:

Paolo Gardoni
University of Illinois, Urbana-Champaign
298 PUBLICATIONS **7,017** CITATIONS

SEE PROFILE

Armen Der Kiureghian
University of California, Berkeley
102 PUBLICATIONS **9,885** CITATIONS

SEE PROFILE

Khalid Mosalam
University of California, Berkeley
183 PUBLICATIONS **4,776** CITATIONS

SEE PROFILE

Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations

Paolo Gardoni[1]; Armen Der Kiureghian, M.ASCE[2]; and Khalid M. Mosalam, M.ASCE[3]

Abstract: A methodology to construct probabilistic capacity models of structural components is developed. Bayesian updating is used to assess the unknown model parameters based on observational data. The approach properly accounts for both aleatory and epistemic uncertainties. The methodology is used to construct univariate and bivariate probabilistic models for deformation and shear capacities of circular reinforced concrete columns subjected to cyclic loads based on a large body of existing experimental observations. The probabilistic capacity models are used to estimate the fragility of structural components. Point and interval estimates of the fragility are formulated that implicitly or explicitly reflect the influence of epistemic uncertainties. As an example, the fragilities of a typical bridge column in terms of maximum deformation and shear demands are estimated.

CE Database keywords: Probability; Concrete columns; Cyclic loads; Deformation; Concrete, reinforced; Capacity; Models.

Introduction

Predictive capacity models in current structural engineering practice are typically deterministic and on the conservative side. These models typically 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 and explicitly account for all the prevailing uncertainties.

In this paper, we present a Bayesian framework for the development of multivariate probabilistic capacity models for structural components that properly account for all the prevailing 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 develop correction terms to existing deterministic capacity models in common use that properly account for the inherent bias and uncertainty in these models. Methods for assessing the model parameters on the basis of observed experimental data are described. 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 paper is aimed at developing probabilistic capacity models, the approach is general and can be applied to the assessment of models in many engineering mechanics problems. As an application, this methodology is used to develop probabilistic shear and deformation capacity models for reinforced concrete (RC) columns under cyclic loading.

Fragility is defined as the conditional probability of failure of a structural member or system for a given set of demand variables. The probabilistic capacity models are used in a formulation to estimate the fragility of structural components with special attention given to the treatment of aleatory and epistemic uncertainties. Univariate fragility curves and bivariate fragility contours are estimated for a representative bridge column under cyclic loading.

Capacity Models

In the context of this paper, 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, and 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 univariate when only one quantity is to be predicted and multivariate when several quantities are to be predicted. We begin our discussion with the univariate form of the model and then generalize to the multivariate case.

A univariate capacity model has the general form

$$C=C(\mathbf{x},\boldsymbol{\Theta}) \tag{1}$$

where $\boldsymbol{\Theta}$ denotes a set of parameters introduced to ''fit'' the model to observed data and C is the capacity quantity of interest. The function $C(\mathbf{x},\boldsymbol{\Theta})$ can have a general form involving algebraic expressions, integrals, or differentials. Ideally, it should be de-

[1]Graduate Student, Dept. of Civil & Environmental Engineering, Univ. of California, Berkeley, CA 94720.

[2]Taisei Professor of Civil Engineering, Dept. of Civil & Environmental Engineering, Univ. of California, Berkeley, CA 94720.

[3]Assistant Professor, Dept. of Civil & Environmental Engineering, Univ. of California, Berkeley, CA 94720.

Note. Associate Editor: George Deodatis. Discussion open until March 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on February 26, 2001; approved on January 4, 2002. This paper is part of the ***Journal of Engineering Mechanics***, Vol. 128, No. 10, October 1, 2002. ©ASCE, ISSN 0733-9399/2002/10-1024–1038/$8.00+$.50 per page.

rived from first principles, e.g., the rules of mechanics. Rather than developing new models, in this paper 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 univariate model form

$$C(\mathbf{x},\boldsymbol{\Theta}) = \hat{c}(\mathbf{x}) + \gamma(\mathbf{x},\boldsymbol{\theta}) + \sigma\varepsilon \tag{2}$$

where $\boldsymbol{\Theta}=(\boldsymbol{\theta},\sigma)$, $\boldsymbol{\theta}=(\theta_1,\theta_2,\ldots)$, denotes the set of unknown model parameters, $\hat{c}(\mathbf{x})$=selected deterministic model, $\gamma(\mathbf{x},\boldsymbol{\theta})$=correction term for the bias inherent in the deterministic model that is expressed as a function of the variables $\mathbf{x}$ and parameters $\boldsymbol{\theta}$, ε=random variable with zero mean and unit variance, and σ represents the standard deviation of the model error. Note that for given $\mathbf{x},\boldsymbol{\theta}$ and σ, we have $\mathrm{Var}[C(\mathbf{x},\boldsymbol{\Theta})]=\sigma^2$ as the variance of the model.

In formulating the model, we employ a suitable transformation of the quantity of interest to justify the following assumptions: (1) the model variance σ^2 is independent of $\mathbf{x}$ (homoskedasticity assumption), and (2) ε has the normal distribution (normality assumption). Box and Cox (1964) suggest a parameterized family of transformations for this purpose. However, in practice, 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

$$\gamma(\mathbf{x},\boldsymbol{\theta}) = \sum_{i=1}^{p} \theta_i h_i(\mathbf{x}) \tag{3}$$

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 reinforced concrete column has different capacities relative to failure in shear, bending, reinforcing bar slip or excessive deformation. For the analysis of such a component, we formulate a q-dimensional multivariate capacity model in the form

$$C_k(\mathbf{x},\boldsymbol{\theta}_k,\boldsymbol{\Sigma}) = \hat{c}_k(\mathbf{x}) + \gamma_k(\mathbf{x},\boldsymbol{\theta}_k) + \sigma_k\varepsilon_k, \quad k=1,\ldots,q \tag{4}$$

where

$$\gamma_k(\mathbf{x},\boldsymbol{\theta}_k) = \sum_{i=1}^{p_k} \theta_{ki} h_{ki}(\mathbf{x}), \quad k=1,\ldots,q \tag{5}$$

With the exception of the new term $\boldsymbol{\Sigma}$, all entries in the above expressions have definitions analogous to those of the univariate model. $\boldsymbol{\Sigma}$ denotes the covariance matrix of the variables $\sigma_k\varepsilon_k$, $k=1,\ldots,q$. The set of unknown parameters of the model in Eq. (4) 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}, $k=1,\ldots,q-1$, $l=k+1,\ldots,q$.

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 that are 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 matters, from errors 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 improved models, more accurate measurements and collection of additional samples. Below, we describe the specific types of uncertainties that arise in assessing capacity models. For simplicity in the notation, we use the formulation of a univariate model.

Model Inexactness

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 Eq. (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 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 between 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

As we will see shortly, the parameters of the model are assessed by use of a sample of observations C_i of the dependent variable 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, we 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 ith 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 are usually obtained through calibration of measurement devices and procedures. The mean values of these errors represent biases in the measurements (systematic error), whereas their variances represent the uncertainties inherent in the

measurements. In most cases 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

The accuracy of estimation of the model parameters $\boldsymbol{\Theta}$ depends on the observation sample size n. 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.

Bayesian Parameter Estimation

The philosophical framework for the approach developed in this paper is based on the Bayesian notion of probability. The eventual goal of developing probabilistic capacity 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, and subjective engineering judgment. It is equally important that the approach account for all the relevant uncertainties, including those that are aleatory in nature and those that are epistemic. The Bayesian framework employed in this paper is ideally suited for this purpose.

In the Bayesian approach, the parameters $\boldsymbol{\Theta}$ are estimated by use of the well-known updating rule (Box and Tiao 1992)

$$f(\boldsymbol{\Theta}) = \kappa L(\boldsymbol{\Theta}) p(\boldsymbol{\Theta}) \tag{6}$$

where $f(\boldsymbol{\Theta})$=posterior distribution representing our updated state of knowledge about $\boldsymbol{\Theta}$; $L(\boldsymbol{\Theta})$=likelihood function representing the objective information on $\boldsymbol{\Theta}$ contained in a set of observations; $p(\boldsymbol{\Theta})$=prior distribution reflecting our state of knowledge about $\boldsymbol{\Theta}$ prior to obtaining the observations; and $\kappa = [\int L(\boldsymbol{\Theta}) p(\boldsymbol{\Theta}) d\boldsymbol{\Theta}]^{-1}$=normalizing factor. The likelihood is a function that is proportional to the conditional probability of making the observations for a given value of $\boldsymbol{\Theta}$. Specific formulations of this function are described in the next section. The prior distribution may incorporate any information about $\boldsymbol{\Theta}$ that is based on our past experience or engineering judgment. When no such information is available, one should use a prior distribution that has minimal influence on the posterior distribution, so that inferences are unaffected by information external to the observations. For the set of parameters $\boldsymbol{\Theta}=(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, it is generally assumed that $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are approximately independent so that $p(\boldsymbol{\Theta}) \approx p(\boldsymbol{\theta}) p(\boldsymbol{\Sigma})$. Using Jeffrey's rule (1961), Box and Tiao (1992) have shown that the noninformative prior for the parameters $\boldsymbol{\theta}$ is locally uniform such that $p(\boldsymbol{\Theta}) \cong p(\boldsymbol{\Sigma})$. Furthermore, they have shown that the noninformative prior for the elements Σ_{ij} of $\boldsymbol{\Sigma}$ has the form

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(q+1)/2} \tag{7}$$

where $|\cdot|$ denotes the determinant. When the unknown parameters are the standard deviations σ_i and the correlation coefficients ρ_{ij}, the noninformative prior takes the form (Gardoni 2002)

$$p(\boldsymbol{\Sigma}) \propto |\mathbf{R}|^{-(q+1)/2} \prod_{i=1}^{q} \frac{1}{\sigma_i} \tag{8}$$

where $\mathbf{R}=[\rho_{ij}]$=correlation matrix. For a univariate model, the above simplifies to

$$p(\sigma) \propto \frac{1}{\sigma} \tag{9}$$

We note that the posterior distribution is centered at a point that represents a compromise between the prior information and the data, and the compromise is increasingly controlled by the data as the sample size increases. For large or even moderate-sized samples, a fairly drastic modification of the prior distribution may only lead to a minor modification of the posterior distribution (Box and Tiao 1992).

Once the posterior distribution of $\boldsymbol{\Theta}$ is determined, one can compute its mean vector and covariance matrix. We denote these as $\mathbf{M}_{\boldsymbol{\Theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}}$, respectively. Computation of these statistics, as well as the normalizing constant κ is not a simple matter, as it requires multifold integration over the Bayesian kernel $L(\boldsymbol{\Theta}) p(\boldsymbol{\Theta})$. An algorithm for computing these statistics is described in Gardoni (2002).

Likelihood Functions

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 univariate model with exact measurements. The effect of measurement error is then incorporated in an approximate manner. The formulation is next extended to multivariate 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: (1) the demand is measured at the instant of failure, in which case the measured demand represents the component capacity; (2) the component does not fail, in which case the measured demand represents a lower bound to the component capacity; and (3) 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. Using Eq. (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 $r_i(\boldsymbol{\theta}) = C_i - \hat{c}(\mathbf{x}_i) - \gamma(\mathbf{x}_i, \boldsymbol{\theta})$ and ε_i denotes the outcome of the model error term at the ith observation.

Lower Bound Datum

Observed 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 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 bound data are often referred to as censored data.

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

$$L(\boldsymbol{\theta},\sigma)\propto \prod_{\text{failure data}} P[\sigma\varepsilon_i = r_i(\boldsymbol{\theta})] \times \prod_{\text{lower bound data}} P[\sigma\varepsilon_i > r_i(\boldsymbol{\theta})] \times \prod_{\text{upper bound data}} P[\sigma\varepsilon_i < r_i(\boldsymbol{\theta})] \tag{10}$$

Since ε has the standard normal distribution, we can write

$$L(\boldsymbol{\theta},\sigma)\propto \prod_{\text{failure data}} \left\{\frac{1}{\sigma}\varphi\left[\frac{r_i(\boldsymbol{\theta})}{\sigma}\right]\right\} \times \prod_{\text{lower bound data}} \Phi\left[-\frac{r_i(\boldsymbol{\theta})}{\sigma}\right] \times \prod_{\text{upper bound data}} \Phi\left[\frac{r_i(\boldsymbol{\theta})}{\sigma}\right] \tag{11}$$

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

Now consider the case where measurement errors are present. Denote $\hat{C}_i$ and $\hat{\mathbf{x}}_i$ as the measured values in the ith 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 zero. Let s_i^2 and $\boldsymbol{\Sigma}_i$ denote the variance of e_{Ci} and the covariance matrix of $\mathbf{e}_{\mathbf{x}i}$, respectively. As it 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. We also assume that the error terms are normally distributed. For the failure data we have $\hat{C}_i + e_{Ci} = \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \boldsymbol{\theta}) + \sigma\varepsilon_i$, for the lower bound data we have $\hat{C}_i + e_{Ci} < \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \boldsymbol{\theta}) + \sigma\varepsilon_i$ and for the upper bound data we have $\hat{C}_i + e_{Ci} > \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) + \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \boldsymbol{\theta}) + \sigma\varepsilon_i$. Defining

$$r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i}) = \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}) - \gamma(\hat{\mathbf{x}}_i + \mathbf{e}_{\mathbf{x}i}, \boldsymbol{\theta}) \tag{12}$$

the conditions for the three types of data can be written as $\sigma\varepsilon_i - e_{Ci} = r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i})$, $\sigma\varepsilon_i - e_{Ci} > r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i})$, and $\sigma\varepsilon_i - e_{Ci} < r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i})$, respectively. Unfortunately $r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i})$ 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 $\hat{\mathbf{x}}_i$, we use a first-order approximation to express $r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i})$ 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

$$r_i(\boldsymbol{\theta}, \mathbf{e}_{\mathbf{x}i}) \cong \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i) - \gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) - [\nabla_{\hat{\mathbf{x}}_i}\hat{c}(\hat{\mathbf{x}}_i) + \nabla_{\hat{\mathbf{x}}_i}\gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta})]\mathbf{e}_{\mathbf{x}i} = \hat{r}_i(\boldsymbol{\theta}) + \nabla_{\hat{\mathbf{x}}_i}\hat{r}_i(\boldsymbol{\theta})\mathbf{e}_{\mathbf{x}i} \tag{13}$$

where $\nabla_{\mathbf{x}}$ denotes the gradient row vector with respect to $\mathbf{x}$ and $\hat{r}_i(\boldsymbol{\theta}) = \hat{C}_i - \hat{c}(\hat{\mathbf{x}}_i) - \gamma(\hat{\mathbf{x}}_i, \boldsymbol{\theta})$. 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 $\hat{\sigma}^2(\boldsymbol{\theta},\sigma) = \sigma^2 + s_i^2 + \nabla_{\hat{\mathbf{x}}_i}\hat{r}_i(\boldsymbol{\theta})\boldsymbol{\Sigma}_i\nabla_{\hat{\mathbf{x}}_i}\hat{r}_i(\boldsymbol{\theta})^T$. Hence, in presence of measurement errors, the likelihood function takes the form

$$L(\boldsymbol{\theta},\sigma)\propto \prod_{\text{failure data}} \left\{\frac{1}{\hat{\sigma}(\boldsymbol{\theta},\sigma)}\varphi\left[\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta},\sigma)}\right]\right\} \times \prod_{\text{lower bound data}} \Phi\left[-\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta},\sigma)}\right] \times \prod_{\text{upper bound data}} \Phi\left[\frac{\hat{r}_i(\boldsymbol{\theta})}{\hat{\sigma}(\boldsymbol{\theta},\sigma)}\right] \tag{14}$$

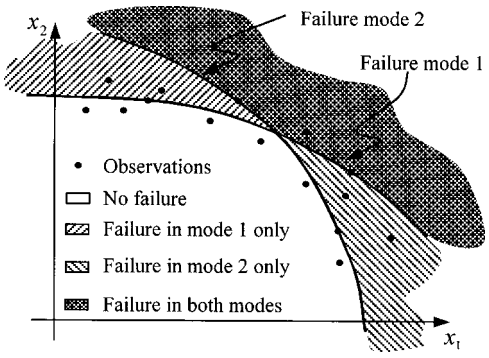

Fig. 1. Representation of data types

We now consider the multivariate model in Eq. (4) under exact measurements. For the ith 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. Fig. 1 illustrates a conceptual representation of the various ways that the data for a bivariate 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 nonfailure 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.

For the ith observation of the kth capacity model, define $r_{ki}(\boldsymbol{\theta}_k) = C_{ki} - \hat{c}_k(\mathbf{x}_i) - \gamma_k(\mathbf{x}_i, \boldsymbol{\theta}_k)$, where C_{ki} is the measured value of the kth capacity or its lower or upper bound. Also, let ε_{ki} be the outcome of the error term for the kth capacity model in the ith 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

$$L(\boldsymbol{\Theta})\propto \prod_{\text{observation } i} P\{[\sigma_k\varepsilon_{ki} = r_{ki}(\boldsymbol{\theta}_k)] \underset{\text{data } k}{\cap_{\text{failure}}} [\sigma_k\varepsilon_{ki} > r_{ki}(\boldsymbol{\theta}_k)] \underset{\text{data } k}{\cap_{\text{lower bound}}} [\sigma_k\varepsilon_{ki} < r_{ki}(\boldsymbol{\theta}_k)] \underset{\text{data } k}{\cap_{\text{upper bound}}} \} \tag{15}$$

Table 1. Probability Terms for Bivariate Capacity Model with Lower Bound and Failure Data

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

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 $\mu_{1|\zeta}=\rho_{12}(\sigma_1/\sigma_2)\zeta$.

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 using the multinormal probability density and cumulative distribution functions. As an example, Table 1 lists the expressions for these terms for a bivariate model with only lower bound and failure data. 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 multivariate model remains similar to that in Eq. (15) with the term $\sigma_k\varepsilon_{ki}$ for the ith observation of the kth model replaced by

$$\sigma_k\varepsilon_{ki}-e_{Cki}-\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\mathbf{e}_{\mathbf{x}i} \tag{16}$$

and $r_{ki}(\boldsymbol{\theta}_k)$ replaced by $\hat{r}_{ki}(\boldsymbol{\theta}_k)$, which is equivalent to $\hat{r}_i(\boldsymbol{\theta})$ for the kth model. The term Eq. (16) is a zero mean normal random variable with variance $\hat{\sigma}_k^2(\boldsymbol{\theta}_k,\sigma_k)=\sigma_k^2+s_{ki}^2+\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\boldsymbol{\Sigma}_i\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)^T$. Furthermore, the terms for the kth and lth models have the covariance $\hat{\rho}_{kl}\hat{\sigma}_k\hat{\sigma}_l=\rho_{kl}\sigma_k\sigma_l+\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{ki}(\boldsymbol{\theta}_k)\boldsymbol{\Sigma}_i\nabla_{\hat{\mathbf{x}}_i}\hat{r}_{li}(\boldsymbol{\theta}_l)^T$, where $\hat{\rho}_{kl}$ denotes the correlation coefficient. For the bivariate model described in Table 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.

Model Selection

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

The probabilistic model in Eqs. (2) and (3) 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 prediction purposes, the selection process should aim at a model that is unbiased, accurate, and easily implementable in practice. Furthermore, from a statistical standpoint, it is desirable that the correction term $\gamma(\mathbf{x},\boldsymbol{\theta})$ have a parsimonious parameterization (i.e., have 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 overfit of the data.

The model form in Eq. (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, the one that has the smallest σ can be considered to be the most accurate. Therefore, an estimate of the parameter σ, e.g., its posterior mean, can be used to select the most accurate model among several viable candidates.

The explanatory functions $h_i(\mathbf{x})$ should be selected so as to enhance the predictive capacity 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 intuition 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 stepwise 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 $|\rho_{\theta_i\theta_j}|=\max_{k\neq i}|\rho_{\theta_i\theta_k}|$ between θ_i and the remaining parameters θ_k, $k\neq i$. A value of $|\rho_{\theta_i\theta_j}|$ close to 1, say $0.7\leqslant|\rho_{\theta_i\theta_j}|$, 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 $|\rho_{\theta_i\theta_j}|$ small in relation to 1, say $|\rho_{\theta_i\theta_j}|\leqslant 0.5$, is an indication that the information content in $h_i(\mathbf{x})$ is not closely related to that in the remaining terms. If $0.7\leqslant|\rho_{\theta_i\theta_j}|$, one can choose to replace θ_i by
$$\hat{\theta}_i=\mu_{\theta_i}+\rho_{\theta_i\theta_j}\frac{\sigma_{\theta_i}}{\sigma_{\theta_j}}(\theta_j-\mu_{\theta_j}) \tag{17}$$
where μ_{θ_i} and σ_{θ_i}=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.

Assessment of Structural Component Fragility

For a structural component, fragility is defined as the conditional probability of attaining or exceeding prescribed limit states for a given set of boundary variables. Following the conventional notation in structural reliability theory (Ditlevsen and Madsen 1996), let $g_k(\mathbf{x},\boldsymbol{\Theta})$ be a mathematical model describing the kth limit state of interest for the structural component, where as in the previous sections $\mathbf{x}$ denotes a vector of measurable variables and $\boldsymbol{\Theta}$ denotes a vector of model parameters. This function is defined such that the event $\{g_k(\mathbf{x},\boldsymbol{\Theta})\leq 0\}$ denotes the attainment or exceedance of the kth limit state by the structural component. 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.

Using the capacity models described earlier, the limit state functions for a structural component can be formulated as

$$g_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})=C_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})-D_k(\mathbf{r},\mathbf{s}) \quad k=1,\ldots,q \tag{18}$$

where $D_k(\mathbf{r},\mathbf{s})$ denotes the demand for the kth failure mode. For example, for failure in shear of a reinforced concrete column, $C_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})=$ maximum shear force that the column can sustain, whereas $D_k(\mathbf{r},\mathbf{s})=$maximum applied shear force. Note that both quantities can be functions of other demand variables, e.g., applied bending moment or axial force. Therefore, the functions $C_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})$ and $D_k(\mathbf{r},\mathbf{s})$ generally could include both $\mathbf{r}$ and $\mathbf{s}$ as arguments.

The fragility of the structural component is stated as

$$F(\mathbf{s},\boldsymbol{\Theta})=P\left[\bigcup_k\{g_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})\leq 0\}\middle|\mathbf{s},\boldsymbol{\Theta}\right] \tag{19}$$

where $P[A|\mathbf{s}]$ denotes the conditional probability of event A for the given values of variables $\mathbf{s}$. The uncertainty in the event for given $\mathbf{s}$ arises from the inherent randomness in the capacity variables $\mathbf{r}$, the inexact nature of the limit state model $g_k(\mathbf{r},\mathbf{s},\boldsymbol{\Theta})$ (or its submodels), and the uncertainty inherent in the model parameters $\boldsymbol{\Theta}$. We have expressed the fragility as a function of the parameters so as to highlight the fact that an estimate of the fragility depends on how we treat the model parameters. Various estimates of fragility can be developed depending on how we treat the parameter uncertainties. These are described in the following subsections.

Point Estimates of Fragility

A point estimate of the fragility is obtained by ignoring the uncertainty in the model parameters and using a point estimate $\hat{\boldsymbol{\Theta}}$ in place of $\boldsymbol{\Theta}$. Most commonly the posterior mean $\mathbf{M}_{\boldsymbol{\Theta}}$ or the maximum likelihood estimate (MLE) $\boldsymbol{\Theta}_{\mathrm{MLE}}$ is used. We denote the corresponding point estimate of the fragility as

$$\hat{F}(\mathbf{s})=F(\mathbf{s},\hat{\boldsymbol{\Theta}}) \tag{20}$$

The uncertainty in this estimation arises from the intrinsic variability in $\mathbf{r}$ and from the random model correction terms ε_k, 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 multinormal probability distribution of ε_k. More generally, a multifold integral involving the joint distribution of $\mathbf{r}$ and ε_k over the failure domain must be computed. Methods for numerical computation of such probability terms are well developed in the field of structural reliability (Ditlevsen and Madsen 1996).

Predictive Estimate of Fragility

The point estimate of fragility in Eq. (20) does not incorporate the epistemic uncertainties inherent in the model parameters $\boldsymbol{\Theta}$. To incorporate these uncertainties, we must consider $\boldsymbol{\Theta}$ as random variables. The predictive estimate of fragility $\tilde{F}(\mathbf{s})$ is the expected value of $F(\mathbf{s},\boldsymbol{\Theta})$ over the posterior distribution of $\boldsymbol{\Theta}$, i.e.,

$$\tilde{F}(\mathbf{s})=\int F(\mathbf{s},\boldsymbol{\Theta})f(\boldsymbol{\Theta})d\boldsymbol{\Theta} \tag{21}$$

This estimate of fragility incorporates the epistemic uncertainties in an average sense. However, it does not distinguish between the different natures of the aleatory and epistemic uncertainties.

Bounds on Fragility

Uncertainty inherent in the fragility estimate due to the epistemic uncertainties is reflected in the probability distribution of $F(\mathbf{s},\boldsymbol{\Theta})$ relative to the parameters $\boldsymbol{\Theta}$. 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.

The reliability index corresponding to the conditional fragility in Eq. (19) is defined as

$$\beta(\mathbf{s},\boldsymbol{\Theta})=\Phi^{-1}[1-F(\mathbf{s},\boldsymbol{\Theta})] \tag{22}$$

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 $F(\mathbf{s},\boldsymbol{\Theta})$ is. Using a first-order Taylor series expansion around the mean point $\mathbf{M}_{\boldsymbol{\Theta}}$, the variance of $\beta(\mathbf{s},\boldsymbol{\Theta})$ is approximately given by

$$\sigma_\beta^2(\mathbf{s})\approx\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}}\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})^T \tag{23}$$

where $\nabla_{\boldsymbol{\Theta}}\beta(\mathbf{s})=$gradient row vector of $\beta(\mathbf{s},\boldsymbol{\Theta})$ at the mean point and $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}}$ denotes the posterior covariance matrix. 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, $\tilde{\beta}(\mathbf{s})\pm\sigma_\beta(\mathbf{s})$, where $\tilde{\beta}(\mathbf{s})=\Phi^{-1}[1-\tilde{F}(\mathbf{s})]$ denotes the mean ± 1 SD bounds of the reliability index. Transforming these back into the probability space, one obtains

$$\{\Phi[-\tilde{\beta}(\mathbf{s})-\sigma_\beta(\mathbf{s})],\quad \Phi[-\tilde{\beta}(\mathbf{s})+\sigma_\beta(\mathbf{s})]\} \tag{24}$$

as ''1 SD'' bounds of the fragility estimate. These bounds approximately correspond to 15% and 85% probability levels.

Applications

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 sections to develop deformation and shear capacity models for RC circular columns under cyclic loading. This specific class of structural component is selected because of their predominant use for bridge structures in many seismically active regions of the world.

As described in ''Capacity Models'', the probabilistic models can be built upon existing deterministic models. The deformation capacity model used in this study is based on the notion of de-

Table 2. Ranges of Variables from 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

composing 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 longitudinal reinforcing bar slip. 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 bivariate deformation–shear capacity model. Finally, the probabilistic capacity models are used to compute univariate fragility curves and bivariate fragility contours for a representative bridge column.

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. In this database, columns 38, 69, 75, and 82 did not include spiral reinforcement, column 106 had some 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 2. 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.

Deformation Capacity Model

Following common practice (Park and Paulay 1975; Lynn et al. 1996), we define the deformation capacity of a column as the displacement Δ corresponding to a drop in the lateral force resistance equal to 20% of its peak value. Fig. 2(a) 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%. Fig. 2(b) represents 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. Fig. 2(c) illustrates such a case, where stiffness deterioration occurs without reaching the lateral force resistance. The measured displacement in this case (i.e., that cor-

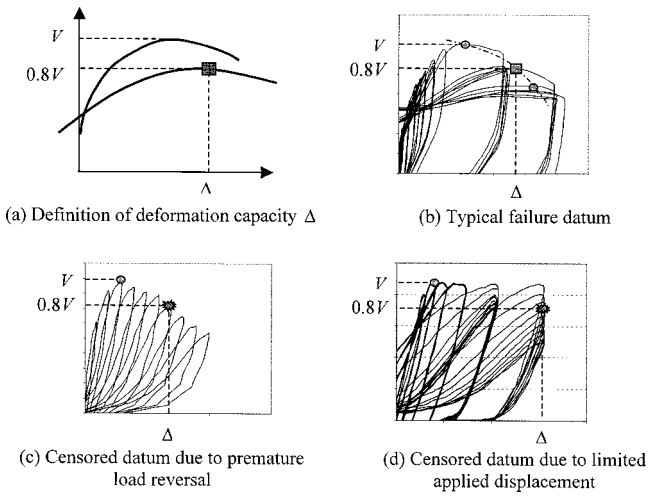

Fig. 2. Deformation capacity definition and data types

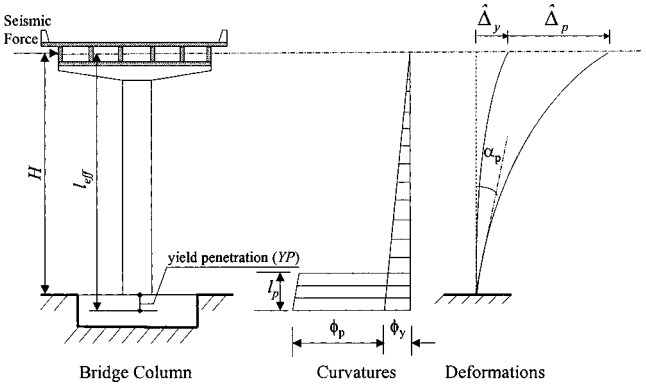

Fig. 3. Decomposition of lateral displacement of single-column bridge bent

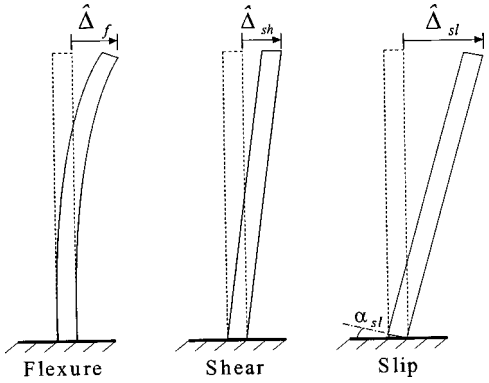

Fig. 4. Components of yield displacement $\hat{\Delta}_y$ for reinforced concrete column

responding to 80% of the peak lateral load) is obviously a lower bound to the deformation capacity Δ. This type of observation is identified as a ''lower bound'' datum. Third are columns whose deformation capacity is not reached because of limitation in the applied maximum displacement. Fig. 2(d) 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.

To develop the deformation capacity model, we employ the drift ratio capacity $d=\Delta/H$. This is a dimensionless quantity, convenient for model formulation. Let $\hat{d}(\mathbf{x})$ be an existing deterministic model for predicting d, 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

$$\ln[d(\mathbf{x},\boldsymbol{\Theta})]=\ln[\hat{d}(\mathbf{x})]+\gamma_d(\mathbf{x},\boldsymbol{\theta})+\sigma\varepsilon \tag{25}$$

where $\boldsymbol{\Theta}=(\boldsymbol{\theta},\sigma)$=set of unknown model parameters and $\gamma_d(\mathbf{x},\boldsymbol{\theta})$, σ, and ε, are as described earlier. Because of the employed logarithmic transformation, one can show that σ is approximately equal to the coefficient of variation (COV) of the drift ratio. The following two subsections describe the formulation used for $\hat{d}(\mathbf{x})$ and $\gamma_d(\mathbf{x},\boldsymbol{\theta})$.

Deterministic Model

It is 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 Fig. 3 for a single RC column bridge bent. Accordingly,

$$\hat{d}(\mathbf{x})=\frac{1}{H}(\hat{\Delta}_y+\hat{\Delta}_p) \tag{26}$$

For a RC column responding as a cantilever (i.e., with a single curvature) with fixed base, the yield displacement is comprised of a flexural component $\hat{\Delta}_f$ based on a linear curvature distribution along the full column height (Fig. 3), a shear component $\hat{\Delta}_{\text{sh}}$ due to shear distortion, and a slip component $\hat{\Delta}_{\text{sl}}$ due to the local rotation at the base caused by slipping of the longitudinal bar reinforcement. These three components are illustrated in Fig. 4. Thus, we have

$$\hat{\Delta}_y=\hat{\Delta}_f+\hat{\Delta}_{\text{sh}}+\hat{\Delta}_{\text{sl}} \tag{27}$$

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

$$\hat{\Delta}_f=\tfrac{1}{3}\phi_y l_{\text{eff}}^2 \tag{28}$$

where $l_{\text{eff}}=H+\text{YP}$=effective length of the column, in which YP denotes the depth of the yield penetration into the column base (Fig. 3). The latter term accounts for the additional rotation of the critical section resulting from strain penetration of the longitudinal reinforcement into the column footing. According to Priestley et al. (1996), YP is estimated as $\text{YP}=0.022f_y d_b$, where d_b=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

$$\hat{\Delta}_{\text{sh}}=\frac{V_y H}{GA_{\text{ve}}} \tag{29}$$

where V_y=shear force at yield, G=shear modulus of concrete, and A_{ve}=effective shear area. The latter is computed as $A_{\text{ve}}=k_I k_s A_g$, where A_g=gross cross sectional area, $k_s=0.9$=shape factor for a circular cross section, and the factor k_I reflects the increased shear deformation in a flexurally cracked RC column. Due 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=effective moment of inertia determined from the moment–curvature relationship of the column cross section, as demonstrated in Fig. 5, and I_g=gross moment of inertia of the cross section.

The contribution to the yield deformation due to slippage of the longitudinal reinforcing bars is related to the local rotation at

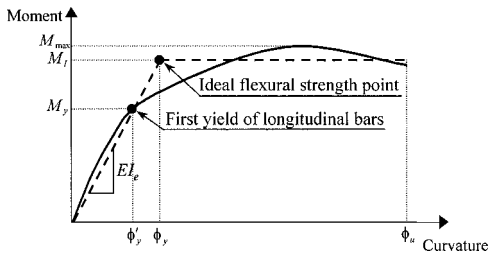

Fig. 5. Generic moment–curvature diagram

the base of the column α_{sl} (Fig. 4). We adopt the assumptions by Pujol et al. (1999), whereby the bond stress at yield 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,

$$\hat{\Delta}_{sl} = \frac{\phi_y f_y d_b H}{8.64\sqrt{f'_c}} \tag{30}$$

The contributions from postelastic flexural behavior, diagonal tension cracking, and yield penetration are manifested in the so-called plastic hinge rotation α_p, as shown in Fig. 3. An equivalent rectangular plastic curvature is commonly assumed (Priestley et al. 1996). Accordingly, the plastic deformation $\hat{\Delta}_p$ in Eq. (26) is obtained from

$$\hat{\Delta}_p = \alpha_p H = \phi_p l_p H \tag{31}$$

where $l_p = 0.08H + \text{YP} \geq 0.044 f_y d_b$ = 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 $\phi_p = \phi_u - \phi_y$ = plastic curvature, where ϕ_u denotes the ultimate curvature.

A generic moment–curvature relationship with elastic–perfectly plastic idealization is illustrated in Fig. 5 showing the definitions of ϕ_y and ϕ_u. In this figure, M_y = moment that induces the first yielding in the column longitudinal reinforcement, ϕ'_y = corresponding curvature, and M_I = 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 setting $\varepsilon_c = 0.005$. For each tested RC column, we compute all the moments and the corresponding curvatures defined above using fiber–element section analysis (Thewalt and Stojadinovic 1994).

From the elastic–perfectly plastic idealization in Fig. 5, ϕ_y is determined from the linear extrapolation $EI_e = M_y/\phi'_y = M_I/\phi_y$. 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

$$\varepsilon_{cu} = 0.004 + \frac{1.4\rho_s f_{yh}\varepsilon_{su}}{f'_{cc}} \tag{32}$$

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

$$f'_{cc} = f'_c\left(2.254\sqrt{1 + \frac{7.94f'_l}{f'_c}} - 2\frac{f'_l}{f'_c} - 1.254\right) \tag{33}$$

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

Model Correction

The term $\gamma_d(x,\theta)$ on the right-hand side of Eq. (25) is intended to correct the bias inherent in the deterministic model $\ln[\hat{d}(\mathbf{x})]$. We select the linear form in Eq. (3) for this function, where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ is a vector of unknown 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 variables $\mathbf{x}$, we select $h_1(\mathbf{x}) = 1$. To detect any possible under- or overestimation of the individual contributions defined in Eqs. (26) and (27) 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 $\ln[\hat{d}(\mathbf{x})]$ 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, we select $h_8(\mathbf{x}) = \rho_s(f_{yh}/f'_c)(D_c/D_g)$. 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 we employ $h_{10}(\mathbf{x}) = f_y/f'_c$ and $h_{11}(\mathbf{x}) = \varepsilon_{cu}$. Note that these explanatory functions are all dimensionless. As a result, the parameters $\boldsymbol{\theta}$ 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 deformation capacity of RC columns.

Parameter Estimation

Having defined the deterministic model $\hat{d}(\mathbf{x})$ and the correction term $\gamma_d(x,\theta)$, we are now ready to assess the probabilistic model in Eq. (25), i.e., estimate its parameters $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_{11}, \sigma)$ by the Bayesian updating formula described in "Bayesian Parameter Estimation". Having no prior information on these parameters, we select the noninformative prior in Eq. (9). 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 "Model Selection" section, a stepwise deletion procedure was described for reducing the number of terms in $\gamma_d(x,\theta)$ to achieve a compromise between model simplicity (few correction terms) and model accuracy (small σ). Since σ is approximately equal to the COV of the predicted drift ratio, the accuracy of the model is not expected to improve by including a term that has a COV much greater than σ.

Fig. 6 summarizes the stepwise term deletion process for the deformation model. For each step, the figure shows the posterior COV of the model parameters θ_i (solid dots) and the posterior mean of the model standard deviation σ (open square). At step 1 with the complete 12-parameter model, the posterior mean of σ is 0.306 and the parameter with the largest COV (=0.71) 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 Fig. 6. 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 COV now is θ_2. We remove the term $\theta_2 h_2(\mathbf{x})$ and continue the same procedure. After eight steps, we

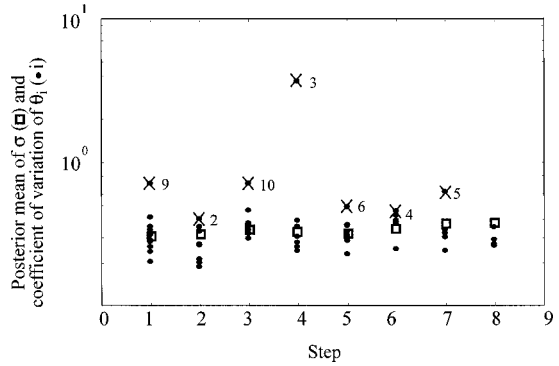

Fig. 6. Stepwise deletion process for deformation capacity model where a superposed cross (×) indicates term to be removed

find the largest COV (for parameter θ_7) to be close in magnitude 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 ($\rho = -0.85$). Considering the definitions of $h_8(\mathbf{x})$ and $h_{11}(\mathbf{x})$ and Eq. (32), a strong correlation between these parameters is not surprising. As a further simplification, using the posterior estimates of the four-parameter model, θ_8, is expressed by its linear regression on θ_{11} [see Eq. (17)] as $\hat{\theta}_8 = -6.035 - 1.034\theta_{11}$. Thus, the reduced correction term takes the form

$$\gamma_d(\mathbf{x},\boldsymbol{\theta}) = \theta_1 + \theta_7 \frac{4V_I}{\pi D_g^2 f_t'} + (-6.035 - 1.034\theta_{11}) \frac{\rho_s f_{yh} D_c}{f_c' D_g} + \theta_{11}\varepsilon_{cu} \quad (34)$$

with only three unknown parameters.

Table 3 lists the posterior statistics of the parameters $\boldsymbol{\Theta} = (\theta_1, \theta_7, \theta_{11}, \sigma)$ of the reduced model. The following observations are noteworthy: (1) The positive mean of θ_1 indicates that, independent of the variables $\mathbf{x}$, the deterministic model $\hat{d}(\mathbf{x})$ tends to underestimate the deformation capacity of the column. (2) The positive 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 "Deterministic Model". The negative estimate of θ_{11} indicates that the deterministic model tends to underestimate the contribution of transverse reinforcement and overestimate the contribution of the ultimate concrete strain.

Table 3. Posterior Statistics of Parameters in Deformation Model

Parameter	Mean	Standard deviation	Correlation Coefficient			
			θ_1	θ_7	θ_{11}	σ
θ_1	0.531	0.119	1			
θ_7	0.701	0.204	−0.37	1		
θ_{11}	−48.4	13.6	−0.59	−0.38	1	
σ	0.383	0.050	−0.04	0.14	0.20	1

Fig. 7 shows a comparison between the measured and predicted values of the drift ratio capacities for the test columns based on the deterministic (left chart) and the probabilistic (right chart) models. For the probabilistic model, median predictions ($\varepsilon = 0$) are shown. The failure data are shown as solid dots and the censored data are shown as open triangles. For a perfect model, the failure data should line up along the 1:1 dashed line and the censored data should lie above it. The deterministic model on the left is strongly biased on the conservative side since most of the failure and many of the censored data lie below the 1:1 line. The probabilistic model on the right clearly corrects this bias. The dotted lines in the right figure delimit the region within 1 SD of the median. We note that a majority of the failure data points fall within 1 SD 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 prevailing uncertainties.

Shear Capacity Models

In this section we construct and compare two probabilistic shear capacity models for RC circular columns. For this purpose, we classify the maximum lateral load measured in each experiment as a "failure" datum if the tested column failed in shear, and as a lower-bound "censored" datum if the tested column failed in

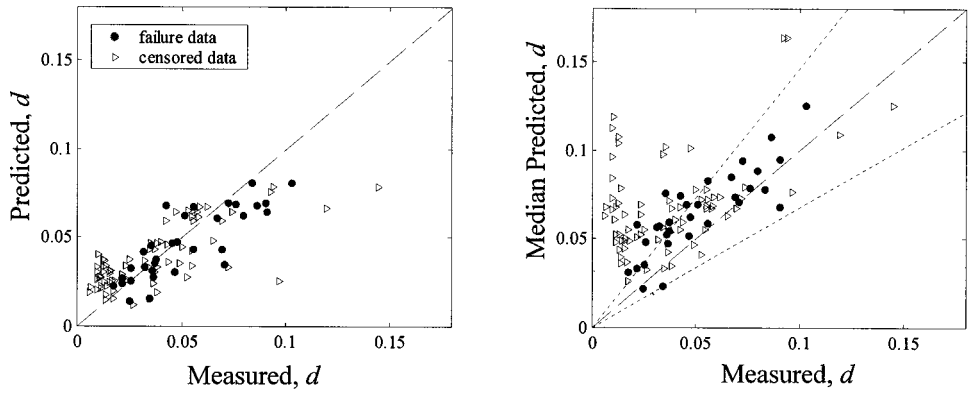

Fig. 7. Comparison between measured and predicted drift ratio capacities based on deterministic (left) and probabilistic (right) models

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 data category.

To develop a dimensionless model for the shear force capacity V, we consider the normalized quantity $\nu = V/(A_g f_t')$, 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

$$\ln[\nu(\mathbf{x},\boldsymbol{\Theta})] = \ln[\hat{\nu}(\mathbf{x})] + \gamma_\nu(\mathbf{x},\boldsymbol{\theta}) + \sigma\varepsilon \tag{35}$$

where $\boldsymbol{\Theta} = (\boldsymbol{\theta},\sigma)$ = set of unknown model parameters, $\hat{\nu}(\mathbf{x})$ = existing deterministic model for predicting ν, and $\gamma_\nu(\mathbf{x},\boldsymbol{\theta})$, σ, and ε are as defined earlier. Due to the logarithmic transformation used, σ is approximately equal to the coefficient of variation of ν. The following subsections describe the formulations of $\hat{\nu}(\mathbf{x})$ and $\gamma_\nu(\mathbf{x},\boldsymbol{\theta})$.

Deterministic Models

Two predictive models for the shear capacity of RC columns are considered. 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 Federal Emergency Management Agency 273 (FEMA 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 steel $\hat{V}_s$, i.e.,

$$\hat{V} = \hat{V}_c + \hat{V}_s \tag{36}$$

where $\hat{V}$ denotes the deterministic prediction of the shear capacity (often denoted nominal 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

$$\hat{V}_c = \nu_b A_e + \frac{M_d}{a} \tag{37}$$

where $\nu_b = (0.067 + 10\rho_t)\sqrt{f_c'} \leq 0.2\sqrt{f_c'}$ with units of MPa is the "basic" shear strength of concrete, in which $\rho_t = 0.5\rho_l$ = longitudinal tension reinforcement ratio, A_e = effective shear area taken as $0.8A_g$ for circular sections (Priestly et al. 1996), $M_d = PI_g/(A_g y_t) = PD_g/8$ = decompression moment with the axial load P and $y_t = D_g/2$, and $a = M/V$ = shear span expressed as the ratio of the moment to shear at the critical section (for a cantilever column $a = H$). The contribution from the transverse steel in Eq. (36) is based on the well-known truss model and is given by

$$\hat{V}_s = \frac{A_v f_{yh} D_e}{S} \tag{38}$$

where $A_v = 2A_h$ = total area in a layer of the transverse reinforcement in the direction of the shear force, D_e = effective depth commonly taken as $0.8D_g$ for circular cross sections, and S = spacing of transverse reinforcement. Substituting Eqs. (37) and (38) in Eq. (36), we have

$$\hat{V} = 0.8\nu_b A_g + 0.125\frac{PD_g}{H} + 1.6\frac{A_h f_{yh} D_g}{S} \tag{39}$$

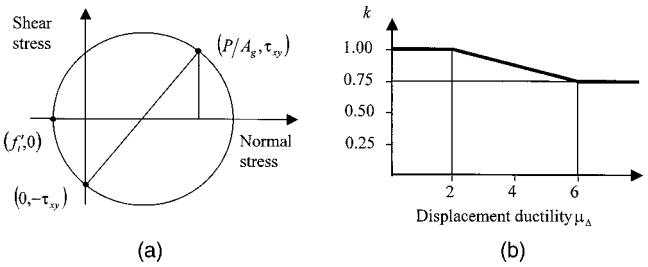

Fig. 8. Shear failure model by Moehle et al. (1999, 2000)

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 Fig. 8(a), interaction between flexural and shear stresses, and strength degradation within the plastic hinge, the final form of this model is

$$\hat{V}_c = k\left(\frac{f_t'}{a/D_e}\sqrt{1 + \frac{P}{f_t' A_g}}\right)A_e \tag{40}$$

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 $\mu_\Delta = \Delta/\hat{\Delta}_y = d/\hat{d}_y$, as defined in Fig. 8(b).

We consider two alternatives for the drift ratio to be used in the expression for μ_Δ. The first is simply $d = \hat{d}(\mathbf{x})$, the deterministic deformation model described in "Deterministic Model". The second is the median of the probabilistic deformation capacity model, $d = \hat{d}(\mathbf{x})\exp[\gamma_d(\mathbf{x},\mathbf{M}_{\boldsymbol{\Theta}})]$, where the parameters are fixed at the posterior mean estimates, $\mathbf{M}_{\boldsymbol{\Theta}}$. In the two papers by Moehle et al. (1999, 2000), different approaches are proposed to introduce 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 paper, we do not introduce any reduction factor for $\hat{V}_s$ in the deterministic model. However, we assess 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 subsection.

Model Correction

The model correction term $\gamma_\nu(\mathbf{x},\boldsymbol{\theta})$ in Eq. (35) is intended to capture the inherent bias in $\ln[\hat{\nu}(\mathbf{x})]$. We select the linear form in Eq. (3). To capture a potential constant bias in $\ln[\hat{\nu}(\mathbf{x})]$, we select $h_1(\mathbf{x}) = 1$. To account for a possible correction in the contribution of the longitudinal steel 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 modification in the contribution from the transverse steel, 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.

Table 4. Reduced Model Correction Terms and Posterior Means and Standard Deviations of σ Selected for Shear Models

Deterministic shear capacity model	$\gamma_\nu(\mathbf{x};\boldsymbol{\theta})$	Mean of σ	Standard deviation of σ
ASCE–ACI 426 model	$\theta_1+\theta_2 h_2$	0.189	0.019
Deformation-dependent model with $d=\hat{d}(\mathbf{x})$ used in computing k	$\theta_2 h_2+\theta_4 h_4$	0.179	0.019
Deformation-dependent model with $d=\hat{d}(\mathbf{x})\exp[\gamma_d(\mathbf{x};\mathbf{M}_{\boldsymbol{\Theta}})]$ used in computing k	$\theta_2 h_2+\theta_4 h_4$	0.153	0.013

Parameter Estimation

We now estimate the parameters $\boldsymbol{\Theta}=(\theta_1,\ldots,\theta_4,\sigma)$ for the shear capacity models using Bayesian inference. Owing to the lack of prior information, the noninformative prior $p(\boldsymbol{\Theta})\propto\sigma^{-1}$ is selected. The stepwise deletion procedure described in the ''Model Selection'' section is used to detect superfluous explanatory functions, which are then dropped to simplify each model. The reduced form of $\gamma_\nu(\mathbf{x},\boldsymbol{\theta})$ is different for the two deterministic shear capacity models. Furthermore, the parameter estimates depend on whether $d=\hat{d}(\mathbf{x})$ or $d=\hat{d}(\mathbf{x})\exp[\gamma_d(\mathbf{x},\mathbf{M}_{\boldsymbol{\Theta}})]$ is used to compute the factor k. Table 4 lists the expression of $\gamma_\nu(\mathbf{x},\boldsymbol{\theta})$ and the posterior mean and standard deviation of σ for each model.

As mentioned earlier, a measure of the predictive accuracy of each model is the posterior mean of σ. The last two columns of Table 4 show that the deformation-dependent shear capacity model with $d=\hat{d}(\mathbf{x})\exp[\gamma_d(\mathbf{x};\mathbf{M}_{\boldsymbol{\Theta}})]$ used in computing the factor k has the smallest error standard deviation and, therefore, is the most accurate model. In the remainder of this paper, we present results only for this model.

Fig. 9 summarizes the stepwise deletion process for the selected shear capacity model. For each step, the figure shows the COV 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 five-parameter model, the COV of θ_3 is 0.340 and the mean of σ is 0.136. To simplify the model, we drop the term $\theta_3 h_3(\mathbf{x})$. In step 2 for the reduced four-parameter model, the mean of σ is 0.144, which indicates no appreciable deterioration of the model. The parameter with the highest COV (=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 COV (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

$$\gamma_\nu(\mathbf{x},\boldsymbol{\theta})=\theta_2\rho_l+\theta_4 A_\nu f_{\mathrm{yh}} D_g/A_g f_t' S \tag{41}$$

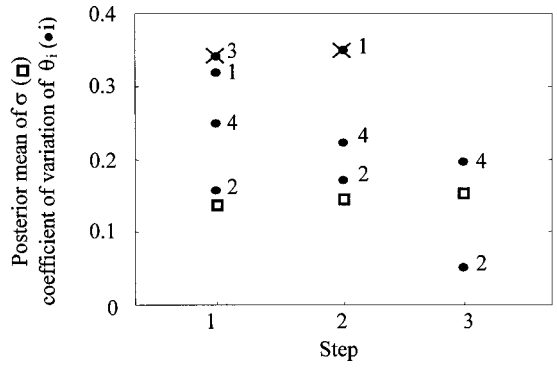

Fig. 9. Stepwise deletion process for shear capacity model, where superposed cross (×) indicates term to be removed

Table 5 lists the posterior statistics of the parameters θ_2, θ_4, and σ for the reduced model.

The following noteworthy observations can be made from the preceding results: (1) the fact that the explanatory function $h_1=1$ is not informative indicates that there is no constant bias in the deterministic model. (2) The presence of $h_2(\mathbf{x})=\rho_l$ with a positive coefficient in Eq. (41) is an indication that the contribution of the longitudinal reinforcement to the shear capacity is underestimated in the deterministic model. This could be due to the fact that 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. 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 the estimate 23.1 for the selected model. (3) The fact that the explanatory function $h_3(\mathbf{x})=PD_g/(A_g f_t' H)$ is 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 [Fig. 8(a)]. (4) The presence of the explanatory function $h_4(\mathbf{x})=A_\nu f_{\mathrm{yh}} D_g/(A_g f_t' S)$ with a negative coefficient in Eq. (41) represents the effect of strength degradation in the contribution from the transverse steel that is needed for a more accurate prediction.

Fig. 10 shows a comparison between the measured and predicted values of the normalized shear capacities for the test columns based on the deterministic (left chart) and probabilistic (right chart) shear capacity models. Same definitions as in Fig. 7 apply. It is seen that the deterministic model on the left is strongly biased on the conservative side. The probabilistic model on the right clearly corrects this bias. The dotted lines in this figure delimit the region within 1 SD of the model.

Bivariate Deformation–Shear Capacity Model

In this section we construct a bivariate deformation–shear capacity model that accounts for the correlation between the two model errors. The subscripts d and ν are used to indicate quantities related to the deformation and shear capacities, respectively. Using Eq. (25) and (35), the bivariate capacity model is written as

$$\ln[d(\mathbf{x},\boldsymbol{\theta}_d,\sigma_d,\rho)]=\ln[\hat{d}(\mathbf{x})]+\gamma_d(\mathbf{x},\boldsymbol{\theta}_d)+\sigma_d\varepsilon_d \tag{42a}$$

Table 5. Posterior Statistics of Parameters in Selected Shear Model

			Correlation Coefficient		
Parameter	Mean	Standard deviation	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

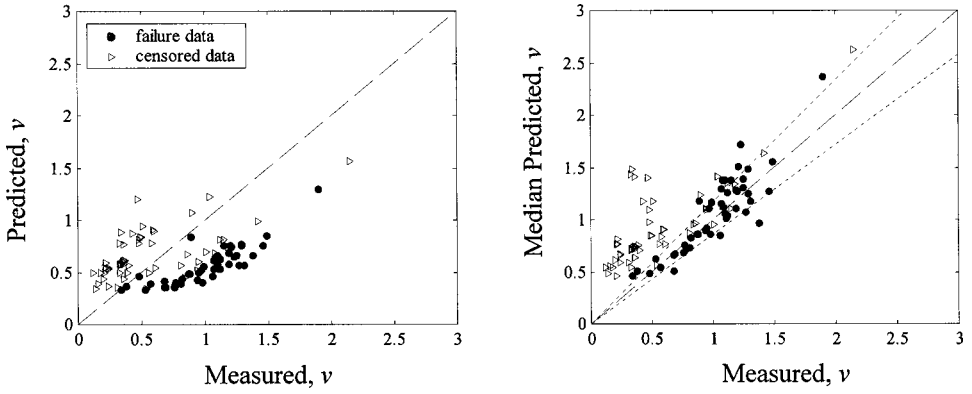

Fig. 10. Comparison between measured and predicted shear capacities based on deterministic (left) and probabilistic (right) models

$$\ln[\nu(\mathbf{x},\boldsymbol{\theta}_\nu,\sigma_\nu,\rho)] = \ln[\hat{\nu}(\mathbf{x})] + \gamma_\nu(\mathbf{x},\boldsymbol{\theta}_\nu) + \sigma_\nu\varepsilon_\nu \quad (42b)$$

where all terms are as defined earlier and ρ=correlation coefficient between the model errors ε_d and ε_ν. The unknown model parameters $\boldsymbol{\Theta}=(\boldsymbol{\theta}_d,\sigma_d,\boldsymbol{\theta}_\nu,\sigma_\nu,\rho)$ are estimated by Bayesian inference. Having already determined the informative explanatory functions for each model, the reduced model correction terms in Eqs. (34) and (41) are used. Owing to the lack of prior information, we select the noninformative prior $p(\boldsymbol{\Theta})\propto(1-\rho^2)^{-3/2}/(\sigma_d\,\sigma_\nu)$.

Table 6 lists the posterior statistics of the parameters $\boldsymbol{\Theta}$. As expected, the estimates of $\boldsymbol{\theta}_d$, $\boldsymbol{\theta}_\nu$, σ_d, and σ_ν are nearly the same as the corresponding estimates for the individual 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.

Univariate and Bivariate Fragility Estimates

The probabilistic deformation and shear capacity models developed in the preceding sections can be used to assess the fragility function for any circular column with specified geometry, material properties, and applied compression force. ''Assessment of Structural Component Fragility'' describes the computational framework for this purpose. An example column with geometry and material properties that are representative of currently constructed RC highway bridge columns in California is considered (Naito 2000). The column has the longitudinal reinforcement ratio $\rho_l=1.99\%$, gross diameter $D_g=1{,}520$ mm, and 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 $f_{\text{yh}}=493$ MPa. To account for material variability, we assume the compressive strength of concrete f_c' is lognormally distributed with mean 35.8 MPa and 10% COV, and the yield stress of longitudinal reinforcement f_y is lognormal with mean 475 MPa and 5% COV. To account for variability in the axial load, we assume P is normally distributed with mean 4,450 kN (corresponding to 7% of the axial capacity based on the gross cross-sectional area) and 25% COV. Finally, to account for variability in construction, we assume the effective moment of inertia I_e is lognormal with mean 2.126×10^{11} mm^4 and 10% COV.

As defined in ''Assessment of Structural Component Fragility'', 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 $\tilde{F}(\mathbf{s})$ is the expected fragility estimate with respect to the distribution of the model parameters $\boldsymbol{\Theta}$. This estimate accounts for the effect of epistemic uncertainties (uncertainty in the model parameters) in an average sense. An explicit account of the variability in the fragility estimate due to epistemic uncertainties is provided by confidence bounds at specified probability levels. In the following, we present these estimates for the example column.

Figs. 11 and 12, respectively, show the univariate fragility curves with respect to drift ratio demand d and normalized shear demand ν for the example column. The solid lines represent the

Table 6. Posterior Statistics of Parameters in Bivariate Deformation–Shear Model

	$\theta_{d,1}$	$\theta_{d,7}$	$\theta_{d,11}$	σ_d	$\theta_{\nu,2}$	$\theta_{\nu,4}$	σ_ν	ρ
Mean	0.512	0.828	-50.8	0.383	21.6	-0.584	0.189	−0.535
Standard deviation	0.054	0.198	11.5	0.048	1.4	0.180	0.019	0.166
				Correlation coefficients				
$\theta_{d,7}$	−0.38							
$\theta_{d,11}$	−0.60	−0.51						
σ_d	−0.41	0.36	0.07					
$\theta_{\nu,2}$	0.02	−0.10	0.07	−0.06				
$\theta_{\nu,4}$	0.12	0.00	−0.11	−0.01	−0.84			
σ_ν	0.06	0.05	−0.10	0.06	0.02	0.03		
ρ	0.14	−0.29	0.12	−0.26	0.12	0.01	−0.15	

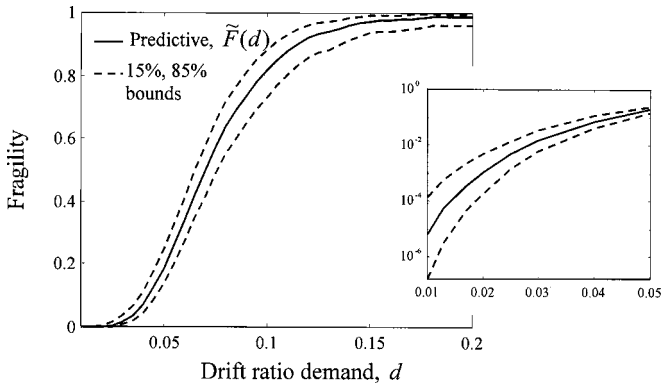

Fig. 11. Fragility estimate for deformation failure of example circular reinforced concrete column

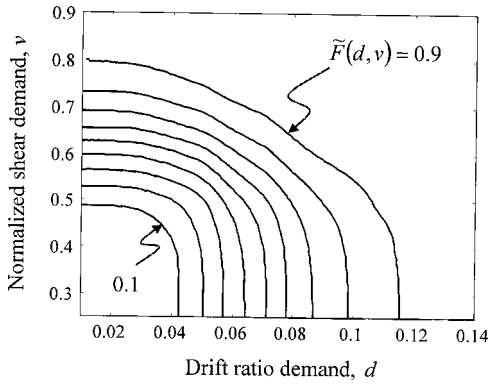

Fig. 13. Contour plot of predictive deformation-shear fragility surface of example circular reinforced concrete column

predictive estimates $\widetilde{F}(d)$ and $\widetilde{F}(v)$ and the dashed lines indicate the 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 $\boldsymbol{\Theta}$).

The bivariate deformation–shear fragility estimates for the example column are obtained using the bivariate capacity model developed in the ''Bivariate Deformation–Shear Capacity Model'' section. 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. Fig. 13 shows the contour plot of the predictive fragility surface $\widetilde{F}(d,v)$ in terms of the drift ratio demand d and normalized shear demand v. Each contour in this plot connects pairs of values of the demands d 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.

Conclusions

A comprehensive Bayesian framework for constructing univariate and multivariate predictive capacity models based on experimental observations is formulated. The models are unbiased and explicitly account for all the prevailing uncertainties. With the aim of facilitating their use in practice, the probabilistic models are constructed by developing correction terms to existing deterministic models in common use. Through a model selection process that makes use of a set of ''explanatory'' functions, terms that effectively correct the bias in the existing model forms are identified and insight into the underling behavioral phenomena is gained.

Although the methodology presented in this paper is aimed at developing probabilistic capacity models for structural components, the approach is quite general and can be applied to the assessment of models in many engineering fields. As an application, univariate and bivariate probabilistic models for the deformation and shear capacities of RC circular columns subjected to cyclic loading are developed using existing experimental data.

The probabilistic capacity models are used in a formulation to assess the fragility of structural components, 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. This methodology is used to estimate univariate fragility curves and bivariate fragility contours for an example RC bridge column. The effect of the cyclic nature of the loading is implicitly accounted for through the experimental data used.

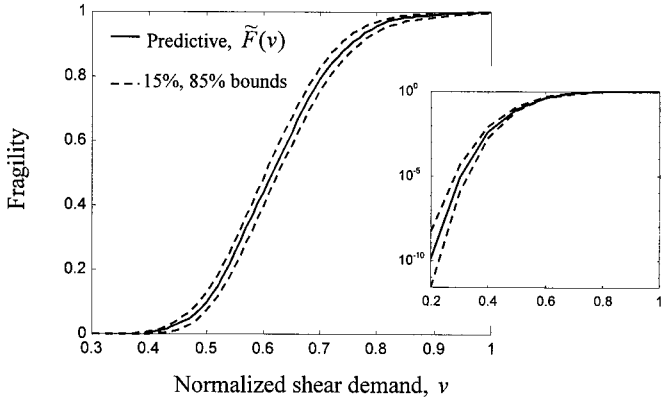

Fig. 12. Fragility estimate for shear failure of example circular reinforced concrete column

Acknowledgments

This work was supported primarily by the Earthquake Engineering Research Centers Program of the National Science Foundation under Award No. EEC-9701568 to the Pacific Earthquake Engineering Research (PEER) Center at the University of California, Berkeley. Opinions and findings presented are those of the writers and do not necessarily reflect the views of the sponsor or PEER.

References

ASCE-ACI Joint Task Committee 426. (1973). ''Shear strength of reinforced concrete members.'' *J. Struct. Eng.*, 99(6), 1091–1187.

Box, G. E. P., and Cox, D. R. (1964). ''An analysis of transformations (with discussion).'' *J. R. Stat. Soc., Series B*, 26, 211–252.

Box, G. E. P., and Tiao, G. C. (1992). *Bayesian inference in statistical analysis*, Addison–Wesley, Reading, Mass.

Der Kiureghian, A. (1989). ''Measures of structural safety under imperfect states of knowledge.'' *J. Struct. Eng.*, 115(5), 1119–1140.
Ditlevsen, O., and Madsen, H. O. (1996). *Structural reliability methods*, Wiley, New York.
FEMA. (1997). ''Guidelines for the seismic rehabilitation of buildings.'' *FEMA 273*, National Earthquake Hazard Reduction Program, Washington D.C.
Gardoni, P. (2002). ''Probabilistic models and fragility estimates for structural components and systems.'' PhD dissertation, Univ. of California, Berkeley, Berkeley, Calif.
Jeffreys, H. (1961). *Theory of probability*, 3rd Ed., Clarendon, Oxford, U.K.
Lynn, A. C., Moehle, J. P., and Mahin, S. A. (1996). ''Seismic evaluation of existing reinforced concrete building columns.'' *Earthquake Spectra*, 12(4), 715–739.
Mander, J. B., Priestley, M. J. N., and Park, R. (1988). ''Theoretical stress-strain model for confined concrete.'' *J. Struct. Eng.*, 114(8), 1804–1826.
Moehle, J. P., Elwood, K., and Sezen, H. (2000). ''Shear failure and axial load collapse of existing reinforced concrete columns.'' *Proc., 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.'' *Proc., 1st U.S.–Japan Workshop on Performance-Based Design Methodology for Reinforced Concrete Building Structures*, Maui, Hawaii, 175–189.
Naito, C. J. (2000). ''Experimental and computational evaluation of reinforced concrete beam-column connections for seismic performance.'' PhD dissertation, Univ. of California, Berkeley, Berkeley, Calif.
Park, Y.-J., and Ang, A. (1985). ''Mechanistic seismic damage model for reinforced concrete.'' *J. Struct. Eng.*, 111(4), 722–739.
Park, R. and Paulay, T. (1975). *Reinforced concrete structures*, Wiley, New York.
Priestley, M. J. N., Seible, F., and Calvi, G. M. (1996). *Seismic design and retrofit of bridges*, Wiley, New York.
Priestley, M. J. N., and Park, R. (1987). ''Strength and ductility of concrete bridge columns under seismic loading.'' *ACI Struct. J.*, 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, Mich., 255–274.
Rao, C. R., and Toutenburg, H. (1997). *Linear models, least squares and alternatives*, Springer, New York.
Stone, J. C. (1996). *A course in probability and statistics*, Duxbury, Belmont, Calif.
Thewalt, C. R., and Stojadinovic, B. (1994). ''Stable reinforced concrete section analysis procedure.'' *J. Struct. Eng.*, 120(10), 3012–3024.
Watson, S., and Park, R. (1994). ''Simulated seismic load tests on reinforced concrete columns.'' *J. Struct. Eng.*, 120(6), 1825–1849.