1993

DEPARTMENT OF CIVIL ENGINEERING
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CALIFORNIA

B U M P
Bayesian Updating of Model Parameters

By

Philippe Geyskens
Armen Der Kiureghian
and
Paulo Monteiro

Report No. UCB/SEMM-93/06
Structural Engineering, Mechanics and Materials Program
Department of Civil Engineering
University of California
Berkeley, California, U.S.A.

June 1993

A lesser scientist than Einstein might have said, "But scientific knowledge comes from nature. Nature provides the hypothesis." But Einstein understood that nature does not. Nature provides only experimental data.

ZEN and the Art of Motorcycle Maintenance — R.M. PIRSIG

Abstract

This report introduces Bayesian updating as a convenient method of analyzing experimental data related to structural engineering applications. Bayesian model assessment is essential for a consistent evaluation of structural safety and the development of a reliability method that accounts for imperfect states of knowledge and recognizes all sources of uncertainty arising in structural problems.

After a short general introduction to the Bayesian theorem, the use of Bayesian methods in model assessment is highlighted. Bayesian model updating is discussed in the broader context of Bayesian structural reliability. An important subproblem of Bayesian updating is the efficient numerical integration of the updating formula. Different proposed methods are reviewed.

Current implementation in the program **BUMP** (`v1.07`) is described in detail and a full "User's Manual" is appended to the report. Problems associated with the Bayesian integration methods used in **BUMP** are documented.

Finally, the utility of Bayesian model assessment in general and of the **BUMP** program in particular, are demonstrated by applying the updating principles to asses a model relating the elastic modulus of concrete to its compressive strength and a model describing the evolution of the compressive strength of plain concrete with time. ACI models for the modulus of elasticity and for the compressive strength of aging concrete are calibrated using data obtained at the Department of Civil Engineering of the University of California at Berkeley.

Acknowledgments

This work was accomplished during the visit of the first author to the University of California in 1991. The first author would like to thank the Flemish Institute for the Encouragement of Scientific Research in Industry and Agronomy (IWONL) for making the visit possible. Additional support was received from a UC Berkeley Faculty Research Grant and from the National Science Foundation Grant No. MSM–8922077. These supports are gratefully acknowledged.

Contents

1. Introduction

In this section, an attempt is made to clarify the essentials of the Bayesian paradigm without going into the theoretical, often philosophical issues. For argumentation on the foundations of Bayesian thinking, the reader is referred to PRESS [37].

1.1 Bayes' theorem

Assume that $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ denote independent observations by random sampling from a density $f(\underline{x}|\underline{\theta})$. The density is conditioned on a set of *unobservable* random parameters $\underline{\theta}$ with density $p(\underline{\theta})$. The Bayes theorem states that

$$f(\underline{\theta}|\underline{x}) = \frac{f(\underline{x}_1|\underline{\theta}).f(\underline{x}_2|\underline{\theta}).\ldots.f(\underline{x}_n|\underline{\theta}).p(\underline{\theta})}{\int_\Theta f(\underline{x}_1|\underline{\theta}).\ldots.f(\underline{x}_n|\underline{\theta}).p(\underline{\theta})d\underline{\theta}} \tag{1.1}$$

with Θ the total definition space of the parameter set $\underline{\theta}$. The trivial proof of the theorem is based on the definition of conditional probability. The interpretation of the formula given by Bayes initiated a new approach to statistical reasoning.

The *prior density* $p(\underline{\theta})$ bundles all prior knowledge about the unknown parameters. Since it has no relation with the experiment under study, it has no objective basis whatsoever, but relies on subjective information about the parameter set. In order to be able to make this statement, one has to refute the *frequentist approach* to probability, saying that $P(A) = \lim_{N\to\infty} n/N$, with n the number of successful events on a total of N experiments, and accept probability to be a mathematical measure for the *"personal degree of disbelief or uncertainty"*.

Given the realization of the sample $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ and using Bayes' theorem for statistical inference, one makes an updated assessment of the behavior of the unknown parameters $\underline{\theta}$, in agreement with the acquired data. The density $f(\underline{\theta}|\underline{x})$ is therefore called the *posterior density* of the parameter set.

The *likelihood function* $L(\underline{x}_1, \ldots, \underline{x}_n|\underline{\theta})$ is proportional to the probability of making a specific observation, given the value $\underline{\theta}$ of the parameters. This is where the data enters the *updating* formula (1.1): the initial belief about the stochastic behavior of a set of parameters is updated using objective observations. The data is indirect: variables influenced by the unknown parameters $\underline{\theta}$ provide the necessary information.

Information can be supplied in many different ways. Observation of an independent sample $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ is easily integrated in the Bayesian formula: the likelihood is proportional to $L(\underline{x}_1, \ldots, \underline{x}_n|\underline{\theta}) = f(\underline{x}_1|\underline{\theta}).f(\underline{x}_2|\underline{\theta}).\ldots.f(\underline{x}_n|\underline{\theta})$, with $f(\underline{x}_i|\underline{\theta})$ proportional to the probability

of observing the values $\underline{x}_i$ when the parameters are equal to $\underline{\theta}$. Inference on less determinate data is equally important. Assume, for example, that the survival of a system is represented by the random variable x. A number $i = 1, \ldots, n$ of independent experiments consists of monitoring the first x_{l_i} life-cycles of the system under study. When the system survives, the experiment only provides a lower bound on the variable X: i.e. $X > x_l$. A more general formulation of the likelihood function, for continuous variables $\underline{x}$, is then

$$L(\underline{x}|\underline{\theta}) \propto \prod_{i=1}^{k} f(x_i|\underline{\theta}) \prod_{i=1}^{l} F(x_{u_i}|\underline{\theta}) \prod_{i=1}^{m} G(x_{l_i}|\underline{\theta}) \tag{1.2}$$

with x_l and x_u observed lower and upper bounds, respectively, and $G(x|\underline{\theta}) = [1 - F(x|\underline{\theta})]$ the complementary CDF.

The Bayesian theorem is often used in its simplified form: posterior density is proportional to likelihood multiplied by prior density

$$f_{\underline{\theta}}(\underline{\theta}) \propto L(\underline{\theta}).p_{\underline{\theta}}(\underline{\theta}) \tag{1.3}$$

with explicit reference to the data omitted in the notation above. It allows the integration of prior, subjective knowledge into a probabilistic assessment of the parameters. The proportionality factor is determined by applying the first probability axiom to the posterior density (total probability = 1). The simplified notation of (1.3) is convenient, since observations are fixed and variation with respect to them is of no further interest. One intuitively feels comfortable with the Bayesian formulation since engineering judgment or decision is almost invariably based on prior information and experience. It also allows for educated decisions based on small samples. A possible argument against Bayesian thinking is the cognitive aspect of the "personal degree of uncertainty", in contrast with the seemingly more "objective" mathematical foundation given by the frequentists.

1.2 Choice of priors

Since the *subjective uncertainty* on the behavior of the parameters enters the Bayesian formula through the prior $p_{\underline{\theta}}(\underline{\theta})$, significant research efforts have been spent on the educated choice of priors. A very complete description of the subject can be found in BOX AND TIAO [8].

The main requirement of any statistical inference is *robustness:* a statistical inference is considered robust, if it is not seriously affected by changes in the initial assumptions [26]. For a large data-set, sensitivity to the type of distribution and to the choice of prior is

minimal. The requirement becomes of major importance when the observed sample is small or moderate in size.

Extending Bayes' original use of the uniform prior, BOX AND TIAO [8] define a *locally uniform prior* to be constant in the significant range of the likelihood and small outside this range. A locally uniform prior can be improper (i.e. not abiding by the first probability axiom $\Rightarrow \int p_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} \neq 1$). When a locally uniform prior is chosen, the Bayesian formula simplifies to

$$f_{\underline{\theta}}(\underline{\theta}) = \frac{L(\underline{\theta}).p_{\underline{\theta}}(\underline{\theta})}{\int L(\underline{\theta}).p_{\underline{\theta}}(\underline{\theta}) d\underline{\theta}} = \frac{L(\underline{\theta})}{\int L(\underline{\theta}) d\underline{\theta}} \tag{1.4}$$

making the posterior result only a function of the likelihood ($\Rightarrow$ data).

This approach (uniform or locally uniform prior) strongly depends on the choice of parameterization: a diffuse prior, uniform in θ might provide information about a function of the parameter θ, such as θ^{-1} or $\log(\theta)$ (see [26, p.49ff]. In order to eliminate this unwanted effect, an attempt is made to find a metric $\phi(\underline{\theta})$ in which a locally uniform prior $\phi(\underline{\theta}) \equiv c$ can be regarded as *non-informative* for the specific transformation of parameters $\underline{\theta}$ used in the parameterized distribution $f(\underline{x}|\underline{\theta})$. An appropriate choice of metric ϕ makes the likelihood *data-translated* [8, p.26ff]:

$$L(\underline{x}|\underline{\theta}) = g(\phi(\underline{\theta}) - f(\underline{x})) \tag{1.5}$$

The likelihood is data-translated when the data $\underline{x}_i$ only influences the position of the likelihood in $\underline{\theta}$-space, not its spread. Identical considerations apply to the choice of prior densities on updatable parameters of mathematical models.

When analytical calculation of the updated density is attempted, one chooses *conjugate priors* [3] in order to simplify the calculations. The choice is one of convenience and considerations of robustness are discarded. In estimating model parameters, this choice is often not available.

2. Bayesian model assessment

2.1 Mathematical models

Studying the behavior of a physical system, one inevitably tries to find a mathematical model describing the system under study in order to make adequate predictions of system behavior. A mathematical model is a set of equations defining relations between *observable variables* $\underline{z} = (z_1, z_2, \ldots, z_k)$ [12]. The relation between aforementioned variables can be implicit when e.g. a numerical algorithm is used to solve a set of differential or integral equations. In order to formulate such expressions, one has to introduce a set of unobservable *parameters* $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_l)$ of which some might be known constants and others will be determined during the empirical calibration of the model.

In its most general form, the mathematical model will be an implicit function of variables and parameters

$$\underline{g}(\underline{z}, \underline{\theta}) = 0 \tag{2.1}$$

with $\underline{g}$ an s-dimensional vector of functions, $\underline{z}$ the m-dimensional vector of observable variables and $\underline{\theta}$ the n-dimensional vector of parameters. This formulation is known as the *structural set of equations* [5]. In most cases the variable set $\underline{z}$ can be partitioned in a subset of independent variables $\underline{x}$ and dependent variables $\underline{y}$. Solving the structural equations for the dependent variables, the *reduced model* is obtained

$$y_i = f_i(\underline{x}, \underline{\theta}) \quad i = 1, \ldots, r \tag{2.2}$$

In this study, we will deal with a one-equation ($r = 1$) reduced model. Extension to multi-equation problems is straightforward.

Because of unavoidable simplifications and uncertainties described in the next section, the model only has an approximate relation to the actual system behavior. Stochastic modeling of model error allows for a mathematical treatment of the model uncertainty.

2.2 Model uncertainty

The uncertainty in a mathematical model has three major sources [11]:

- *errors of ignorance or simplifications*: in most cases the underlying physical phenomena controlling the behavior of the system are insufficiently understood. Variable interaction

can remain undetected. Even the functional form of the model might be inappropriate. For reasons of convenience, one is also often compelled to simplify the model by eliminating supposedly insignificant terms from the equations or adopting simplified functional forms (e.g. linearized equations). One could argue about the modeling of ignorance and uncertainty errors as random phenomena. This modeling method is consistent with the Bayesian paradigm, where randomness can be interpreted as a degree of uncertainty.

- *measurement errors*: a mathematical model is calibrated, i.e. model parameters are determined by using experimental data. Data is acquired by measurement and measurements are prone to (random or systematic) errors. Using potentially inaccurate measured values to determine the model parameters, additional uncertainty is introduced into the model.

- *statistical errors*: when the model is exact and deterministic, a data-set of the same dimension as the set of unknown parameters is sufficient to determine aforementioned parameters. In dealing with an uncertain model, an infinite set of data is required to find "exact" point-estimates of the parameters. Model uncertainty will increase when only a limited data-set is available. This is particularly the case in structural engineering applications, where only few experiments can be performed, due to the high cost and the destructive nature of the tests.

- In some cases (see applications (§ 6.)), no direct relation exists between the different measured data-items. For example, in finding the relation between elastic modulus and compressive strength of concrete, the different measurements will be performed on different specimens of the same batch. As such, we will only be able to relate the measured elasticity to the population characteristics of the compressive strength. This uncertainty is also of statistical nature.

To integrate all these uncertainties into the model, model parameters are considered as random variables. Consistent with the Bayesian principle, a probability density function (PDF) is assigned to the set of parameters. Bayesian model assessment then means the updating of prior information on model parameters in light of observed data to construct a posterior distribution on the parameters. Characterization of marginal and joint statistics of the posterior distribution will be the prime focus of this study.

2.3 Theory

The uncertainty of the model is reflected in the uncertainty of its parameters. The determination of the model parameters can be done in two different ways

- calculation of *point estimates* is done by defining a procedural function between data and parameter estimate $\underline{\hat{\theta}} = \underline{h}(\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_n)$ which allows the determination of a "single" value of the parameter vector. The estimates $\underline{\hat{\theta}}$ are themselves random variables, characterized by a distribution which depends on both the functional relationship $\underline{h}$ and the data-distribution. The functional relationship can take different algorithmic forms: a point-estimator $\underline{\hat{\theta}}$ can maximize the likelihood $L(\underline{z}|\underline{\theta})$ (MLE: maximum likelihood estimator), can minimize the sum of squares of the residuals $e_i = y_i - f(\underline{x}_i, \underline{\theta}^*)$ (LSE: least square estimator), etc. A detailed description of point-estimation is given in [5].
- As mentioned before, the uncertainty on the parameters can be translated into a PDF $f_{\underline{\theta}}(\underline{\theta})$. Using Bayesian updating, prior information about the parameters can be integrated in the final parameter assessment. A posterior distribution carries more information than a point estimator. Small databases can efficiently be used. Point or interval estimates can still be obtained by statistical analysis of the posterior distribution. Using the PDF on the parameters, a predictive density of the dependent variable y can be obtained.

The posterior distribution of the parameters is obtained by applying the Bayesian updating formula (1.3)

$$f_{\underline{\theta}}(\underline{\theta}) \propto L(\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_n|\underline{\theta}) \cdot p_{\underline{\theta}}(\underline{\theta}) \tag{2.3}$$

An important issue in the use of the Bayesian updating formula for model assessment is the adequate selection of a likelihood function $L(\cdot)$.

2.3.1 Likelihood: exact model

Supposing that the model $y = g(\underline{x}, \underline{\theta})$ is exact, uncertainty originates from measurement error. In addition, the limited data set does not allow full quantification of aforementioned errors ($\Rightarrow$ statistical uncertainty). Assuming that only the dependent variable y is subject to inexact measurements, the error on observation k is equal to

$$\varepsilon_k = \hat{y}_k - g(\underline{x}_k, \underline{\theta}) \qquad k = 1, \ldots, n \tag{2.4}$$

where $\hat{y}_k$ stands for the measured value of the dependent variable y for observation x_k. Since the likelihood is proportional to the probability of observing errors $\underline{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ when $\underline{\theta}$ takes on a certain value, it has to be proportional to the joint PDF of the measurement errors:

$$L(\underline{\theta}, \underline{\eta}) \propto f_{\underline{\varepsilon}}(\hat{y}_1 - g(\underline{x}_1, \underline{\theta}), \hat{y}_2 - g(\underline{x}_2, \underline{\theta}), \ldots, \hat{y}_n - g(\underline{x}_n, \underline{\theta}) | \underline{\eta}) \tag{2.5}$$

with $\underline{\eta}$ the parameters of the joint distribution of $\underline{\varepsilon}$. These parameters are theoretically unknown, and should therefore be added to the updatable parameter set $\underline{\theta}$.

Taking measurement errors to be mutually independent with identical normal distribution is a valid engineering approximation [3]. If the measurement device was duly calibrated, the measurements are unbiased, meaning that the measurement error has zero mean and standard deviation σ_ε. Based on these assumptions, the likelihood is written as

$$L(\underline{\theta}, \sigma_\varepsilon) \propto \frac{1}{\sigma_\varepsilon^n} \exp\left[-\frac{1}{2}\sum_1^n \left(\frac{\hat{y}_k - g(\underline{x}_k, \underline{\theta})}{\sigma_\varepsilon}\right)^2\right] \tag{2.6}$$

If we further assume that σ_ε is a known characteristic of the measurement device, equation (2.6) reduces to

$$L(\underline{\theta}) \propto \exp\left[-\frac{1}{2}\sum_1^n \left(\frac{\hat{y}_k - g(\underline{x}_k, \underline{\theta})}{\sigma_\varepsilon}\right)^2\right] \tag{2.7}$$

2.3.2 Likelihood: inexact model

One way to include model inexactness in the probabilistic assessment, is to concentrate the uncertainty caused by both the inexactness of the model formulation and the influence of missing variables into a single correction factor γ

$$\hat{y}_k = g(\underline{x}_k, \underline{\theta}) + \gamma_k \tag{2.8}$$

If $f_{\underline{\gamma}}(\underline{\gamma}|\underline{\eta})$ is the PDF of the sample $\underline{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$, the formulation of the likelihood is similar to (2.5). It is convenient (but not necessarily exact) to assume that γ_k are mutually independent normals, with zero mean (unbiased model) and common standard deviation σ_γ. The likelihood function is then equal to (2.6), if measurement error ε is replaced by the model error parameter γ.

2.3.3 Likelihood: combination of uncertainties

When both sources of uncertainty are present, both correction factors ε and γ are introduced to obtain the observed value $\hat{y}_k$ of the dependent variable

$$\hat{y}_k = g(\underline{x}_k, \underline{\theta}) + \gamma_k + \varepsilon_k \tag{2.9}$$

The likelihood function is determined by the *joint* PDF of $\underline{\gamma} + \underline{\varepsilon}$. Under the aforementioned normal assumptions, the sum of these two variates is also normal with variance $\sigma^2 = \sigma_\varepsilon^2 + \sigma_\gamma^2$. Applying the same reasoning as in the previous sections, the likelihood becomes

$$L(\underline{\theta}, \sigma_\varepsilon^2, \sigma_\gamma^2) \propto (\sigma_\varepsilon^2 + \sigma_\gamma^2)^{-n/2} \exp\left[-\frac{1}{2}\sum_1^n \frac{(\hat{y}_k - g(\underline{x}_k, \underline{\theta})^2}{\sigma_\varepsilon^2 + \sigma_\gamma^2}\right] \tag{2.10}$$

Other derivations of the likelihood for more general formulations of the mathematical model and for other uncertainty assumptions are discussed by DER KIUREGHIAN [12].

2.3.4 Predictive use of statistical models

Once the posterior distribution of the variables is determined, different statistics can be obtained from by the general Bayesian integral

$$\int_\Theta w(\underline{\theta}) f_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} \tag{2.11}$$

with $f_{\underline{\theta}}(\underline{\theta})$ the posterior distribution of the parameters $\underline{\theta}$ and $w(\underline{\theta})$ a weight function, which is dependent on the statistic of interest: for $w(\underline{\theta}) \equiv 1$, the reciprocal of the proportionality constant of the Bayesian updating formula is obtained, for $w(\underline{\theta}) = \underline{\theta}$ the mean parameter-vector is calculated and for $w(\underline{\theta}) = \underline{\theta}.\underline{\theta}^t$ the mean-square matrix is evaluated. By restricting the integration to a subset of the parameter space, *nuisance parameters* are eliminated resulting in appropriate marginal densities and moments of the parameter subset.

Of main interest is the use of the calibrated probabilistic model for predictive purposes, e.g. for obtaining a probability density function of the dependent variable y for given deterministic or probabilistic values of the independent variables $\underline{x}$. According to the total probability rule, one can write

$$f_y(y) = \int_\Theta f_{y|\underline{\theta}}(y|\underline{\theta}) f_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} \tag{2.12}$$

The conditional distribution $f_{y|\underline{\theta}}(y|\underline{\theta})$ is implicitly defined by the functional dependence of y on $\underline{x}$. The theory of probability transformation [39] enables the determination of this distribution. The transformation procedure is best understood when applying it to a specific calibrated model (see § 6.).

It should be obvious from the previous discussion that an efficient determination of the Bayesian integral (2.11) is crucial for the success in Bayesian model assessment. Analytic solutions are only possible for few (academic) applications. Numerical procedures for the integration will be detailed in this report (§ 4.).

3. Bayesian structural reliability

3.1 Introduction

Bayesian model assessment, which is the main focus of this report, may be conducted in the broader context of Bayesian Structural Reliability (BSR). While strong foundations for the introduction of Bayesian decision theory were already laid by CORNELL in 1969 [10], reliability researchers seem to have altogether shunned the idea for many years. In recent publications by DER KIUREGHIAN [11][12], principles of Bayesian decision theory were introduced in the "classical" reliability formulation in order to distinguish between the different sources of uncertainty with which the reliability engineer is confronted. Distinction between reducible and inherent variability has prompted MAES [30] to use BSR for code calibration: users of the code are rewarded by smaller safety factors when design is based on more accurate modeling. The main features of BSR are summarized in the following sections.

3.2 Definitions

Calculating the reliability of a structural component boils down to the evaluation of the following probability integral

$$P_f = \int_{g(\underline{x})\leq 0} f_{\underline{x}}(\underline{x}) d\underline{x} \tag{3.1}$$

with $f_{\underline{x}}(\underline{x})$ the distribution of the basic variables $\underline{x} = (x_1, x_2, \ldots, x_n)$ and $g(\underline{x})$ the limit-state function. When $g(\underline{x}) \leq 0$, the component is considered to be in a failure state.

In order to distinguish between the sources of uncertainty, further parameterization is proposed:

$$P_f(\underline{\theta}) = \int_{g(\underline{x},\underline{\theta})\leq 0} f_{\underline{x}}(\underline{x},\underline{\theta}) d\underline{x} \tag{3.2}$$

The parameters $\underline{\theta}$ stand for the statistical estimation error and the model imperfection, the basic variables $\underline{x}$ describe the inherent variability of the structural system. Sources of model imperfection have been enumerated in the previous section. The imperfection of the probabilistic model $f_{\underline{x}}(\underline{x})$ is induced by lack of data for the statistical characterisation of the basic variables. Again, in accordance with the Bayesian paradigm, these errors or uncertainties are translated into PDFs of the parameters.

The most interesting feature of this parameterization is the distinction made between reducible and irreducible variability: inherent uncertainty on the basic variables is per definition irreducible, but model and estimation errors can be reduced by model refinement and/or further data-acquisition. The reducible uncertainty is bundled in the parameter set $\underline{\theta}$.

In "classical" reliability analysis, the lack of information has been tackled in a number of pragmatic ways, e.g. only the first- and second order moments of the basic variables were assumed to be known in the second-moment modeling approach, or a stochastic model is constructed from marginal densities of and correlation between the basic variables (Nataf model) [14].

In actual practice the amount of information taken from a sample is not restricted, as is assumed in the previous methods, but the quality of the information decreases with increasing order of the statistic. The parametric reliability method (3.2), combined with Bayesian updating, adequately handles the influence of "adding information" (= performing further experiments), thus allowing the engineer to make educated decisions about the extension of his experimental campaign. Finally, the Bayesian reliability format allows the integration of subjective information (engineering judgment) in the reliability assessment.

Since the failure probability is a function of the parameters $\underline{\theta}$, the (generalized) reliability index also becomes a *random function* of the parameters:

$$\beta(\underline{\theta}) = \Phi^{-1}(1 - P_f(\underline{\theta})) \tag{3.3}$$

in which $\Phi(.)$ denotes the standard normal cumulative probability density (CDF). One aim of Bayesian reliability analysis is to determine the PDF of $\beta(\underline{\theta})$ so that different consistent point-estimators of the reliability index can be calculated. DER KIUREGHIAN [11] defines the minimum penalty reliability index as the reliability index minimizing a predetermined penalty function. This weighted point estimation is a common technique in Bayesian decision theory [4]:

$$\beta_{mp} = \min_{\hat{\beta}} \mathrm{E}[h(\beta - \hat{\beta})] = \min_{\hat{\beta}} \int_{\beta_{min}}^{\beta_{max}} h(\beta - \hat{\beta}) dF_\beta(\beta) \tag{3.4}$$

where $h(\beta - \hat{\beta})$ is a penalty function. Predictive safety measures are also a necessary tool for reliability based decision making

$$\tilde{P}_f = \mathrm{E}[P_f(\underline{\theta})] = \int_\Theta P_f(\underline{\theta}) f_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} \quad \Rightarrow \quad \tilde{\beta} = \Phi^{-1}[1 - \tilde{P}_f] \tag{3.5}$$

$\tilde{\beta}$ is called the *predictor reliability index.*

3.3 Calculation

The integration of the probability integrals (3.1) and (3.2) has been a research topic for quite some time. First- and second-order reliability methods, simulation, etc. [29][31] give approximate solutions for the integral. Many of these methods also provide sensitivity measures of the failure probability with respect to any parameter of the problem.

Decision variables, such as the minimum penalty and the predictive reliability index require yet another integration level. For instance, the CDF of β is obtained by solving

$$F_\beta(\beta) = \int_{\beta(\underline{\theta})-\beta \leq 0} f_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} \tag{3.6}$$

This is similar to the probability integral (3.1), with integration domain $\beta(\underline{\theta}) - \beta \leq 0$. The boundary in itself requires the probability integral (3.2) to be solved. Computation of the above integral requires a nested application of the standard reliability methods cited above.

One way to perform the integration for the predictive failure probability (3.5) is to extend the set of basic variables $\underline{x}$ with the reducible parameters $\underline{\theta}$

$$\tilde{P}_f = \int_{g(\underline{v}) \leq 0} f_{\underline{v}}(\underline{v}) d\underline{v} \qquad \underline{v} = \begin{bmatrix} \underline{x} \\ \underline{\theta} \end{bmatrix} \tag{3.7}$$

allowing for the application of the aforementioned traditional techniques.

Alternatively, one can perform a *nested reliability calculation.* It can be proven [11] that the predictor reliability index $\tilde{\beta}$ is the solution of the reliability problem with

$$\tilde{g} = u + \beta(\underline{\theta}) \tag{3.8}$$

as limit-state function, where $[u, \underline{\theta}]^T$ is the set of basic variables and $u \sim N(0,1)$ is a standard-normal variate. Since the limit-state $\tilde{g}$ is a function of $\beta(\underline{\theta})$, a nested reliability solution is required for this approach.

Further approximations for the predictor β or alternative

point-estimators $\hat{\beta}$ can be found in [11].

4. Quadrature of the Bayesian updating formula

4.1 Introduction

When applying Bayesian methods for the determination of model parameters, using subjective information and experimental data, one comes across integrals of the type

$$I(w(\underline{\theta})) = \int_{\Theta} w(\underline{\theta}) L(\underline{x}|\underline{\theta}) p_{\underline{\theta}}(\underline{\theta}) d\underline{\theta} = \int_{\Theta} B(\underline{\theta}) d\underline{\theta} \tag{4.1}$$

with $L(\underline{x}|\underline{\theta})$ the likelihood function and $p_{\underline{\theta}}(\underline{\theta})$ the prior distribution of the parameters $\underline{\theta}$. The weight function $w(\underline{\theta})$ is determined by the statistic to be obtained

- $w(\underline{\theta}) \equiv 1 \Rightarrow$ inverse of the normalizing constant (proportionality constant) of the Bayesian updating formula.
- $w(\underline{\theta}) = \theta_i \Rightarrow$ expectation of parameter θ_i.
- $w(\underline{\theta}) = (\theta_i - \mathrm{E}(\theta_i)).(\theta_j - \mathrm{E}(\theta_j)) \Rightarrow$ covariance between parameters θ_i and θ_j; variance of θ_i for $i = j$.
- $w(\underline{\theta}) = f_{y|\underline{\theta}}(y|\underline{\theta}) \Rightarrow$ predictive distribution $\tilde{f}_y(y)$ of the dependent variable y.

By performing the integration over a subset of the parameter space, nuisance parameters [8] can be eliminated to obtain marginal statistics.

In the following paragraphs, the product of the weight $w(\underline{\theta})$, the likelihood $L(\underline{\theta})$ and the prior $p(\underline{\theta})$ will be referred to as the *Bayesian kernel* $B(\underline{\theta})$. A different formulation of (4.1), based on a logarithmic Bayesian kernel $\mathcal{B}(\underline{\theta})$, is frequently used for the derivation of approximation methods

$$I(w(\underline{\theta})) = \int_{\Theta} w(\underline{\theta}) \exp\{\mathcal{L}(\underline{\theta}) + \pi(\underline{\theta})\} d\underline{\theta} = \int_{\Theta} \mathcal{B}(\underline{\theta}) d\underline{\theta} \tag{4.2}$$

where $\mathcal{L}(\underline{\theta}) = \log L(\underline{\theta})$ is the log-likelihood function and $\pi(\underline{\theta}) = \log p_{\underline{\theta}}(\underline{\theta})$ is the logarithm of the prior distribution.

Only in very few cases is it possible to analytically solve the Bayesian integration (4.1). Analytical solving restricts the choice of priors to conjugate distributions, if they exist. In many practical cases one has to resort to numerical integration or approximate evaluation of $I(w(\underline{\theta}))$.

The high dimensionality of $\underline{\theta}$ and the often infinite integration domain constitute important problems. Moreover, given the diversity in weight functions $w(\underline{\theta})$ and priors $p_{\underline{\theta}}(\underline{\theta})$, a flexible method has to be selected. Many evaluations of integrals (4.1) have to be performed in order to obtain all necessary statistics of posterior and predictive variables. Efficiency of the selected method thus becomes an important issue. In this chapter, methods for the numerical integration or approximation of $I(w(\underline{\theta}))$ are critically reviewed. Press [37] gives a short review of the listed methods.

4.2 Expansion method

This approximation method, introduced by LINDLEY [27], is applicable for large databases (n samples) and low dimension ($p \leq 5$) of the parameter set. The large number of samples guarantees that the likelihood $L(\underline{\theta})$ is concentrated around a single maximum likelihood estimator $\hat{\underline{\theta}}$. The approximation method is based on a Taylor expansion of the logarithmic Bayesian kernel (4.2). The expansion method is more of historical interest. Indeed, it requires the third-order partial derivatives of the (log-)likelihood to be known, while the Laplace approximation, described in the following section, attains the same accuracy using only the Hessian of the (log-)likelihood.

4.3 Laplace approximation

An approximation of the integration (4.2) based on the asymptotic approximation of Laplace-type integrals is presented in a paper by TIERNEY AND KADANE [43]. The initially proposed method was restricted to the approximation of I with strictly positive weight functions $w(\underline{\theta})$. In [44] the method was extended to the calculation of expectation and variance by non-positive functions $w(\underline{\theta})$ through use of the moment generating function. A further application of the approximation is presented in [44].

The asymptotic ($n \rightarrow \infty$) approximation of a Laplace integral $\int \exp\{nh(\theta)\}d\theta$, with $h(\theta)$ an unimodal function, is equal to

$$\begin{aligned} \int_{-\infty}^{+\infty} \exp\{nh(\theta)\}d\theta &\approx \int_{-\infty}^{+\infty} \exp\{nh(\hat{\theta}) - n\frac{(\theta-\hat{\theta})^2}{(2\sigma)^2}\}d\theta \\ &= \sqrt{2\pi}\,\sigma\, n^{-1/2} \exp\{nh(\hat{\theta})\} \end{aligned} \tag{4.3}$$

with $\hat{\theta}$ equal to the mode of $h(\theta)$ and $\sigma^2 = -1/h''(\hat{\theta})$. Extension to multidimensional Laplace integrals is straightforward.

The integral $I(w(\theta))$ (4.2) is of Laplace-type and the aforementioned expansion (4.3) can be applied, n being equal to the size of the sampled data set

$$I(w(\underline{\theta})) \approx (2\pi)^{\frac{n}{2}} \det[\Sigma] \exp\{n\mathcal{H}(\hat{\underline{\theta}})\} \tag{4.4}$$

with $\mathcal{H} = \frac{1}{n}\{\mathcal{L}(\underline{\theta}) + \pi(\underline{\theta}) + \log(w(\underline{\theta}))\}$, $\hat{\underline{\theta}} \leftarrow \mathcal{H}(\hat{\underline{\theta}}) = \max_{\underline{t} \in \underline{\Theta}} \mathcal{H}(\underline{t})\}$ the mode of $\mathcal{H}(\underline{\theta})$ and Σ equal to the inverted Hessian of $\mathcal{H}(\underline{\theta})$, evaluated at the mode $\Sigma = [\nabla^2 \mathcal{H}(\hat{\underline{\theta}})]^{-1}$.

The approximation error of (4.3) is of the order $O(n^{-1})$. Solving a Bayesian updating problem, the fraction $I(w(\underline{\theta}))/I(1)$ needs to be taken to eliminate the proportionality constant. By taking the fraction, dominant error terms cancel and the Laplace approximation is of order $O(n^{-2})$, provided $w(\underline{\theta})$ is strictly positive. For non-positive functions $w(\underline{\theta})$, TIERNEY ET AL. [44] recommend the use of the moment-generating function $\exp\{sw(\underline{\theta})\}$ and the calculation of $M(s) = \mathrm{E}[\exp\{sw(\underline{\theta})\}]$. The expectation and variance of $w(\underline{\theta})$ can then be found by differentiation

$$\hat{\mathrm{E}}[w(\underline{\theta})] = \frac{d}{ds} \log \hat{M}(0) \tag{4.5}$$

$$\hat{\mathrm{E}}[(w(\underline{\theta}))^2] = \frac{d^2}{ds^2} \log \hat{M}(0) \tag{4.6}$$

where the hat stands for the fact that the integration was performed using the approximation (4.4).

The Laplace method is particularly suited to determine initial estimates of expected value and covariance of the parameter set.

4.4 The Gauss-Hermite quadrature

BAŽANT AND CHERN [6] use a GAUSS-HERMITE type integration to solve a particular Bayesian integral. A more general approach was initially proposed by NAYLOR AND SMITH in [34] and was further developed in a paper by SMITH ET AL. [42].

The integration method is based on the Gauss-Hermite quadrature for weighted integrals of the form

$$\int_{-\infty}^{+\infty} \exp\{-t^2\} h(t) dt \tag{4.7}$$

The integral is approximated using a k-point rule

$$\int_{-\infty}^{+\infty} \exp\{-t^2\} h(t) dt \approx \sum_{1}^{k} \omega_i h(t_i) \qquad \omega_i = \frac{2^{k-1}\, k!\, \sqrt{\pi}}{k^2 [H_{k-1}(t_i)]^2} \tag{4.8}$$

with H_{k-1} the Hermite polynomial of order $k-1$ and t_i the ith root of $H_k(t)$. The quadrature formula is exact when $h(t)$ is a polynomial of order $2k-1$.

Considering

$$B(t) = h(t)(2\pi\,\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\} \tag{4.9}$$

it is easy to prove [34] that, for $h(t)$ a well-behaved function

$$\int_{-\infty}^{+\infty} B(t) dt \approx \sum_{1}^{n} m_i B(z_i) \tag{4.10}$$

with

$$m_i = \omega_i \exp(t_i^2)\sqrt{2}\,\sigma \qquad z_i = \mu + \sqrt{2}\,\sigma t_i \tag{4.11}$$

The Gauss-Hermite quadrature is very efficient for $I(w(\theta))$ having a kernel that can be approximated by a function of the type "polynomial $\times$ normal". Since the likelihood generally behaves like a normal density for large sample sizes [8], most problems will satisfy this underlying requirement.

A multivariate extension of the quadrature formula is obtained by applying the Cartesian product-rule

$$\int \ldots \int B(t_1, \ldots, t_n) dt_1 \ldots dt_k \approx \sum m_{ik} \sum \ldots \sum m_{i1}\, B(z_{i1}, \ldots, z_{ik}) \tag{4.12}$$

A problem in the application of (4.10) and (4.12) for integration of the Bayesian integral (4.1) is an adequate choice of mean μ_k and variances σ_{kk}^2 of the Gaussian approximation of the likelihood. These quadrature parameters should be equal to the unknown mean and variance of the (unimodal) posterior density of each parameter θ_k. The maximum likelihood estimates (MLE)

$$\underline{\mu} = \hat{\underline{t}} \leftarrow \mathcal{L}(\hat{\underline{t}}) = \max_{\underline{t} \in E^n} \mathcal{L}(\underline{t}) \tag{4.13}$$

$$\sigma_{ii}^2 = -\{[\nabla^2 \mathcal{L}(\hat{\underline{t}})]^{-1}\}_{ii} \tag{4.14}$$

can be taken as starting values of an iteration in which the current approximations to the posterior mean and variance are used to construct grid z_{ik} and weights m_{ik} for the next iteration.

In order to be able to apply the Cartesian product-rule, one has to make assumptions about posterior independence of the parameters. When high posterior correlations between the parameters $\underline{t}$ exists, an appropriate linear (second-moment) transformation is performed to obtain a new orthogonal set of parameters (i.e. orthogonality in the statistical sense of the word: $\mathrm{E}[t_i.t_j] = 0$, for $i \neq j$).

The Gauss-Hermite method does not require the satisfaction of any asymptotic assumption, but the requirement of having an integrand which approximately behaves as a "polynomial $\times$ normal" might cause convergence problems when dealing with small data sets. Applications of the method have been restricted to low dimension cases ($p \leq 6$), due to the assumptions underlying the use of the Cartesian product rule.

4.5 Simulation

4.5.1 Introduction

Simulation has always been an appropriate but expensive method for numerical integration [40]. Application of the Monte Carlo method for solving (4.1) has been extensively publicized [37]. No restrictions are imposed on the integrand and the dimensionality of the integration. The estimate of the integral can be made with any desired accuracy, at least when one is prepared to pay the high computational cost. Monte Carlo simulation is rarely applied without some sort of variance reduction technique, such as importance sampling [40].

In two recent papers GELFAND AND SMITH [17] and GELFAND ET AL. [16] discuss the use of Gibbs-sampling for the reconstruction of marginal densities (and their moments) given a full set of conditionals. This method is described later in this section.

4.5.2 Importance sampling

An early account of the use of Monte-Carlo importance sampling method for the Bayesian estimation of the posterior moments of the parameters of a linear econometric model can be found in KLOEK AND VAN DIJCK [25]. The paper focused on the adequate choice of an importance sampling distribution. In [45] the authors extend some of the methods developed

in their original paper. GEWEKE [18] finally puts all heuristic considerations on the selection of a sampling distribution into a consistent mathematical framework.

Using importance sampling, the Bayesian integral (4.1) is modified in order to incorporate the *importance sampling density* $S(\underline{\theta})$

$$I(w(\underline{\theta})) = \int_{\Theta} \frac{B(\underline{\theta})}{S(\underline{\theta})}[S(\underline{\theta})d\underline{\theta}] \tag{4.15}$$

Moments of the posterior distributions obtained by Bayesian updating can then be estimated using

$$\hat{w}_N = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{B(\underline{\theta}^{(i)})}{S(\underline{\theta}^{(i)})}\right] \tag{4.16}$$

in which samples $\underline{\theta}^{(i)}$ are drawn from $S(\underline{\theta})$. It can be proven that, under weak assumptions, the estimator $\hat{w}_N$ converges to $I(w(\underline{\theta}))$ for $N \to \infty$.

GEWEKE [18] states that the importance sampling density should closely mimic the behavior of the posterior density, especially in the tail region: the tails of $S(\underline{\theta})$ should not decay more rapidly than the tails of the posterior $f_{\underline{\theta}}(\underline{\theta})$ and no subset of the posterior should be excluded. Empirical selection of an importance sampling density in many practical cases has proven to be cumbersome [46]. For seemingly sensible choices of $S(\underline{\theta})$, $\hat{w}_N$ can behave badly, resulting in unstable behavior of the estimator caused by extremely large values of $w(\underline{\theta}^{(i)})$. GEWEKE, therefore, approached the problem from an analytic perspective, using the local behavior of the posterior density at its mode. He introduced a new importance sampling density based on the Student-t distribution. Replication of these principles is beyond the scope of this report and the reader is therefore referred to the original paper [18].

4.5.3 Directional sampling

Confronted with the aforementioned tail-problem, VAN DIJCK ET AL. [46] introduced a different approach to the simulation method. Their research was motivated by the observation of ill-behaved posterior densities for small data-bases. The directional sampling method still assumes the posterior to be unimodal, but it is able to capture the directional variance in tail-behavior. This method reduces the parameter space $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$ to an $n-1$ dimensional vector and a scalar. For the original mathematical construction of the method, the reader should consult [46]. We have opted for a different approach, based on the statement of the authors that their results can also be obtained by using a simple polar transformation

of the parameter space [23]. The presented alternate approach is also closely related to the well-known directional simulation principle in structural reliability [7][15][32].

The directional sampling method is based on a sequence of transformations. First, a *second-moment transformation* is performed, so that the parameter set becomes *standardized* (zero mean, unit covariance matrix) in the transformed space

$$\underline{s} = \underline{A}(\underline{\theta} - \mathrm{E}[\underline{\theta}]) \qquad \underline{\Sigma}_{\underline{\theta}\underline{\theta}} = (\underline{A}^{-1})(\underline{A}^{-1})^t \tag{4.17}$$

which transforms the Bayesian integral (4.1) into

$$I(w(\underline{s}(\underline{t}))) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} B(\underline{s}(\underline{t})).\det[\underline{A}^{-1}]d\underline{s} \tag{4.18}$$

The standardization is approximate since the values of expectation and covariance matrix of $\underline{\theta}$ are not exactly known. Next, the standardized variable $\underline{s}$ is transformed to hyper-polar coordinates [15][32][23]:

$$\underline{s} = r\underline{\omega} \qquad r = \|\underline{s}\| \qquad \underline{\omega} = \frac{\underline{s}}{\|\underline{s}\|} \qquad r \geq 0 \tag{4.19}$$

with $\underline{\omega}$ located on the unit hypersphere. The integral transforms into

$$I(w^*(r\,\underline{\omega})) = \int_{\Omega}\int_{0}^{+\infty} B^*(r\underline{\omega})\det[\underline{A}^{-1}]\frac{(2\pi)^{n/2}}{\Gamma(n/2)2^{(n/2-1)}}r^{(n-1)}dr\,d\underline{\omega} \tag{4.20}$$

with n the dimension of the parameter space. The Jacobian of the hyperpolar transformation is equal to the surface of the unit hypersphere. Simulation on the unit hypersphere produces the estimator

$$\hat{I} = \det[\underline{A}^{-1}]\frac{(2\pi)^{n/2}}{\Gamma(n/2)2^{(n/2-1)}}\hat{\mathrm{E}}_{\underline{\omega}}^{(N)}\left[\int_{0}^{+\infty} B^*(r\underline{\omega}^{(i)})r^{(n-1)}dr\right] \tag{4.21}$$

where the expectation $\hat{\mathrm{E}}$ stands for a summation over all samples $\underline{\omega}^{(i)}$, divided by the number of samples N. Counting on symmetry of the posterior density, the principle of *antithetic sampling* [40] can further reduce the variance on the sampling estimator $\hat{I}$

$$2\hat{I} \propto \hat{\mathrm{E}}_{\underline{\omega}}^{(N)}\left[\int_{-\infty}^{+\infty} B^*(r\underline{\omega}^{(i)})|r|^{(n-1)}dr\right] \tag{4.22}$$

Standard quadrature modules (mathematical software libraries) can be used to perform the infinite linear integration [e.g. [36]].

Directional sampling should not be too sensitive to the initial choices of mean vector $\hat{\underline{t}}$ and covariance array $\underline{\underline{C}}$, since the entire mass of the likelihood is included by the infinite integration in the radial direction. Convergence is therefore guaranteed. Initial values of first- and second-order moments of the parameters will most certainly influence the efficiency of the sampling. Indeed, the second-moment transform can be considered as an implicit variance reduction scheme by introduction of a normal importance sampling density, as is adequately demonstrated in [46]. Evolution of the sampling coefficient of variation (COV) should be a good indication for the quality of the initial estimates. MLEs provide decent first estimates of the moments of the parameter set.

4.5.4 Conditional sampling

GELFAND AND SMITH [17] developed the theoretical bases of the application of conditional sampling (or Gibbs sampling) for the determination of marginal densities by means of simulated samples of conditional distributions.

Consider the random variables $x_1, x_2, \ldots, x_n$ of which the complete set of *full conditional distributions* of $x_i|x_j(j \neq i)$ is available. In many practical situations, full conditioning will not be necessary: only a limited subset of $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ influences the behavior of variable x_i.

The Gibbs sampling updating scheme proceeds as follows: (i) select a set of starting values $x_1^{(0)}, \ldots, x_n^{(0)}$ and (ii) draw a sample $x_1^{(1)}$ from the density $f_{x_1|x_2,\ldots,x_n}(x_1|x_2^{(0)}, \ldots, x_n^{(0)})$, (iii) followed by a sample $x_2^{(1)}$ from the density $f_{x_2|x_1,x_3,\ldots,x_n}(x_2|x_1^{(1)}, x_3^{(0)}, \ldots, x_n^{(0)}) \ldots$, (iv) pursue in natural order, until a sample $x_n^{(1)}$ is obtained from $f_{x_n|x_1,\ldots,x_{n-1}}(x_n|x_1^{(1)}, \ldots, x_{n-1}^{(1)})$

The algorithm requires n random variate generators. After N iterations the sampled set $(x_1^{(N)}, x_2^{(N)}, \ldots, x_n^{(N)})$ is obtained. It can be proven that under mild conditions, the sampled set converges to a set $\underline{x}$ with joint distribution $f_{\underline{x}}(\underline{x})$ for $N \to \infty$. The convergence also holds for every component of the random vector $\underline{x}$ and the marginal density is estimated using

$$\hat{f}_{x_s}(x_s) = \frac{1}{N} \sum_{i=1}^{N} f_{x_s|x_r;r\neq s}(x_s|x_r = x_r^{(i)}; r \neq s) \tag{4.23}$$

Moments of a function $w(\underline{x})$ can easily be calculated since a sample of $\underline{x}$ is available.

In a following article GELFAND ET AL. [16] presented the application of Gibbs sampling for the calculation of marginal posterior densities. All distributions are considered to be conditioned on the data and the other components of the parameter set. The interested

reader can find many examples of Gibbs sampling in Bayesian inference in the previously mentioned paper.

The Gibbs sampling method is remarkably simple, requiring no intricate transformations of the parameter space. The computational expenses are claimed to be within reason. The only problem is the implementation of n random variate generators able to draw samples from problem dependent conditional distributions. The problem of an adequate convergence criterion for the iterative sampling procedure is only pragmatically addressed by GELFAND ET AL. [16].

5. BUMP: implementation

5.1 Introduction

Two of the numerical integration methods mentioned in chapter (§ 4.) have been implemented in the current version of **BUMP**: the directional sampling method (DS), and the Gauss-Hermite quadrature (GHQ) algorithm. In this section relevant specifics of the implementation are described, without going into actual programming details. Understanding of these implementation details will be a valuable help for the critical scrutiny of the **BUMP** output. Description of the input file is given in the appendices (§ A. and § B.) together with a summary of error messages (§ C.) and their probable causes.

It goes without saying that **BUMP** (v1.07) could gain from further improvements. A future version of **BUMP** should for example contain a simple regression module providing reasonably accurate initial estimates of first- and second-moments of the parameter set. Despite all these considerations, it is our opinion that this version already provides sufficient functionality to allow research in the applicability of Bayesian model assessment for many engineering applications.

5.2 Transformations

5.2.1 Variable transformation

It should be clear from the theoretical discussion in the previous chapter that the implemented methods for Bayesian integration (DS, GHQ) assume that the parameter space is unbounded. In many practical situations, this is *not* the case. Transformation from original to unbounded domain (and vice-versa) is therefore necessary.

In general, parameter θ_i is defined in a finite interval $\theta_i \in [a, b]$, with a the lower and b the upper bound of parameter θ_i. Both $|a|$ and $|b|$ can become infinite ($= \infty$), which basicly means that the variable is semi-bounded or unbounded. The transformation to the unbounded variables $\underline{t}$ depends on aforementioned type of bounding. Table 5.1 shows the different formulations of $\underline{t} = \underline{T}(\underline{\theta})$ and its associated inverse $\underline{\theta} = \underline{T}^{-1}(\underline{t})$.

Table 5.1: Variable transformation

	$\theta_i \in [-\infty, +\infty]$	$\theta_i \in [a, b]$
$t_i = T(\theta_i)$	θ_i	$\log\left(\frac{\theta_i - a}{b - \theta_i}\right)$
$[J_{\underline{T}}]_{ii} = \frac{\partial t_i}{\partial \theta_i}$	1	$\frac{b-a}{(b-\theta_i)(\theta_i - a)}$
$\theta_i = T^{-1}(t_i)$	t_i	$\frac{a + b\exp\{t_i\}}{1 + \exp\{t_i\}}$
$[J_{\underline{T}^{-1}}]_{ii} = \frac{\partial \theta_i}{\partial t_i}$	1	$\frac{b-a}{[1+\exp\{t\}]^2}$
$\theta_i = T_c^{-1}(t_i, \hat{\theta}_i)$	$\hat{\theta}_i + t_i$	$\frac{(\hat{\theta}_i - a)b\exp\{t_i\} + a(b - \hat{\theta}_i)}{(b - \hat{\theta}_i) + (\hat{\theta}_i - a)\exp\{t_i\}}$
$[J_{\underline{T}_c^{-1}}]_{ii} = \frac{\partial \theta_i}{\partial t_i}$	1	$\frac{(b-a)(\hat{\theta}_i - a)(b - \hat{\theta}_i)\exp\{t_i\})}{[(b - \hat{\theta}_i) + (\hat{\theta}_i - a)\exp\{t_i\}]^2}$

	$\theta_i \in [a, +\infty]$	$\theta_i \in [-\infty, b]$
$t_i = T(\theta_i)$	$\log(\theta_i - a)$	$-\log(b - \theta_i)$
$[J_{\underline{T}}]_{ii} = \frac{\partial t_i}{\partial \theta_i}$	$\frac{1}{\theta_i - a}$	$\frac{1}{b - \theta_i}$
$\theta_i = T^{-1}(t_i)$	$a + \exp\{t_i\}$	$b - \exp\{-t_i\}$
$[J_{\underline{T}^{-1}}]_{ii} = \frac{\partial \theta_i}{\partial t_i}$	$\exp\{t\}$	$\exp\{-t\}$
$\theta_i = T_c^{-1}(t_i, \hat{\theta}_i)$	$a + (\hat{\theta}_i - a)\exp\{t_i\}$	$b - (b - \hat{\theta}_i)\exp\{-t_i\}$
$[J_{\underline{T}_c^{-1}}]_{ii} = \frac{\partial \theta_i}{\partial t_i}$	$(\hat{\theta}_i - a)\exp\{t_i\}$	$(b - \hat{\theta}_i)\exp\{-t_i\}$

In both integration methods (DS, GHQ) the origin of the unbounded domain has to coincide with the current estimate of the expectation of the parameter set. With the simple (inverse) variable transform described above (§ 5.2.1), the origin of the unbounded domain is invariably transformed to the same location, independent of the expectation of the parameters (e.g. for

a parameter with lower bound $\theta_i \in [a, +\infty] : t_i = 0 \Rightarrow \underline{T}^{-1}(t_i) = \theta_i = 1 + a$). We therefore define a modified (inverse) transformation, labelled centered variable transformation $\underline{T}_c^{-1}$. This transformation makes the origin of the unbounded space coincide with a specific value $\hat{\theta}_i$ of the original parameter. The centered transformation is included in Table 5.1.

The transformations are introduced at integration level and, therefore, the presence of the Jacobian of the transformation is required. The Jacobian matrix of the transform defined above (or its inverse) is strictly diagonal. Derivatives of the transformation $\underline{T}$ and of the inverse transformation $\underline{T}^{-1}$ are also given in Table 5.1

The transformation of moments requires that the second-order derivative of the inverse transformation be known. We only print this derivative for the fully bounded case $\theta_i \in [a, b]$.

$$\frac{\partial^2 \theta_i}{\partial {t_i}^2} = \frac{(b-a)\exp\{t\}[1-\exp\{t\}]}{[1+\exp\{t\}]^3}$$

5.2.2 Transformation of moments

Initial estimates of expectation and covariance of the parameters are determined by optimization in the *unbounded domain.* These results have to be transformed to the original domain. A second-order approximation [4] is used to obtain the expectation in the original domain

$$\mathrm{E}[\theta_i] = E[\underline{T}^{-1}(t_i)] = \underline{T}^{-1}(\mathrm{E}[t_i]) + \mathrm{Var}[t_i]\frac{\partial^2 t_i}{\partial {\theta_i}^2} \tag{5.1}$$

A first-order approximation is used to determine the covariance in the original domain

$$\mathrm{Cov}(\theta_i, \theta_j) = \mathrm{Cov}(t_i, t_j)\frac{\partial \theta_i}{\partial t_i}\frac{\partial \theta_j}{\partial t_j} \tag{5.2}$$

5.3 Initial estimates

In the theoretical discussion, the MLE have been proposed as decent initial values for the iterative determination of expectation and covariance of the parameter set by the GHQ and the DS methods. This choice of initial values is based on the fact that for a sufficiently large sample of data (and under suitable regularity conditions) the likelihood $L(\underline{\theta}|\underline{x})$ closely approximates a normal distribution [8]

$$L(\underline{\theta}|\underline{x}) \approx \mathrm{N}(\underline{\hat{\theta}}, -[\nabla^2 \mathcal{L}(\underline{\hat{\theta}})]^{-1}) \tag{5.3}$$

with $\mathcal{L}(.) = \log L(.)$ the log-likelihood and $\underline{\hat{\theta}} \leftarrow \mathcal{L}(\underline{\hat{\theta}}) = \max_{\underline{t} \in \Theta} \mathcal{L}(\underline{t})$ the mode of the log-likelihood.

Determination of mode and inverted Hessian of the log-likelihood function is the next problem to tackle. The former demands an optimization scheme, the latter could be solved by using finite differences. **BUMP** uses a quasi-Newton optimization scheme to find the mode *and* to approximate the inverted Hessian. A similar approach was already employed in [20] solving the design point and the principal curvatures of the limit-state surface for reliability problems.

5.3.1 Quasi-Newton optimization

The mode is determined using an unconstrained quasi-Newton optimization scheme [28]. **BUMP** determines the minimum of the negative log-likelihood, after transformation to the unbounded domain

$$\min_{\underline{t} \in E^n} \{-\mathcal{L}(\underline{T}^{-1}(\underline{t})|\underline{x})\} \tag{5.4}$$

Selection of a minimization scheme (as opposed to maximization) is rather arbitrary, and has no other specific advantage than the immediate determination of the negative inverted Hessian.

At iteration k the search direction is $\underline{d}_k = \underline{H}_k \nabla \mathcal{L}(\underline{t}_k)$, with $\underline{t}_k$ the current approximation of the mode and $\underline{H}_k$ the current approximation of the inverted Hessian. The gradient of the log-likelihood is determined by central finite difference and Richardson extrapolation [9]. The next approximation of the optimum is equal to $\underline{t}_{k+1} = \underline{t}_k + \alpha_k \underline{d}_k$. The step length α_k is determined by an exact line-search, based on the golden section method [28]. The iteration is considered to have converged when two criteria are fulfilled: (i) $\|\nabla \mathcal{L}\| < \varepsilon_1$ and (ii) $\|\alpha_k \underline{d}_k\| < \varepsilon_2$

5.3.2 BFGS-update of the inverted Hessian

Quasi-Newton schemes are differentiated from (pure) Newton schemes, because second-order derivative information is not re-evaluated at every iteration step. An efficient class of methods uses an updating scheme that allows reconstructing the inverted Hessian, provided the updating is initiated with a positive definite matrix H_0. It is widely accepted that the BFGS-scheme is the most efficient updating algorithm available [21]. It is formulated as follows

$$\underline{H}_{k+1} = \underline{H}_k + \left(1 + \frac{\underline{g}_k^t \underline{H}_k \underline{g}_k}{\underline{g}_k^t \underline{s}_k}\right) \frac{\underline{s}_k \underline{s}_k^t}{\underline{g}_k^t \underline{s}_k} - \frac{\underline{s}_k \underline{g}_k^t \underline{H}_k + \underline{H}_k \underline{g}_k \underline{s}_k^t}{\underline{g}_k^t \underline{s}_k} \tag{5.5}$$

with $\underline{g}_k = \nabla\mathcal{L}(\underline{t}_k) - \nabla\mathcal{L}(\underline{t}_{k+1})$ and $\underline{s}_k = \underline{t}_{k+1} - \underline{t}_k = \alpha_k \underline{d}_k$

Convergence of the actual inverted Hessian is not checked. A specific construction of the program guarantees that the updated $\underline{H}_k$ is sufficiently close to $-[\nabla^2\mathcal{L}(\hat{\underline{\theta}})]^{-1}$:

- the implemented method is *partial*, meaning that every $p \times n$ steps the updating is re-initialised, in order to reduce influence from inadequate starting values; the programming parameter p is currently set to 3.
- the iteration is implemented as a double loop such that, even when convergence is achieved (according to aforementioned criteria), a sufficient number of additional iterations will guarantee a decent estimate of the inverted Hessian.

The motivation behind implementation of a partial method is one of convergence. The reason for the double looping lies in the fact the for a quadratic objective function, the inverted Hessian is *exactly* reconstructed using n BFGS updating steps [28].

5.3.3 Transformation

As stated earlier, optimization is performed in the unbounded domain. Initial estimates for expectation and covariance are also in reference to the unbounded variables $\underline{t}$. Moment transformation (§ 5.2.2) is used to find approximations in the original space.

5.4 Directional Sampling

The directional sampling method is implemented according to the theory described in section 4.5.3. Details regarding the sampling, line integration, variables transformation and convergence are documented below.

5.4.1 Sampling

A realization of the directional vector $\underline{\omega}$, uniform on the $n-1$ dimensional unit hypersphere Ω, is obtained by sampling a $n-1$ dimensional standard normal vector $\underline{u}$, followed by a normalization $\underline{\omega} = \underline{u}/\|\underline{u}\|$. A linear congruential generator ($m = 2^{31} - 1$, $a = 16807$) is used to generate uniform random variates [19]. The Box-Muller formula [38] transforms these uniform variates into variables drawn from a standard normal distribution.

5.4.2 Linear integration

The linear integration within the directional sampling estimator (4.22) is solved using an adaptive quadrature routine from **QUADPACK** [36]. Implementing the double infinite integration induces numerical problems (as was experienced in early development of the program). Knowing that the likelihood behaves approximately normal (more so with increasing size of the data set), integration is performed in a bounded interval $[-b, b]$. If the likelihood is assumed to be normal, the square of the polar coordinate r^2 will have a χ^2_n distribution

$$F_r(r) = \frac{1}{\Gamma(n/2)}.\Gamma\left(\frac{n}{2}, \frac{r^2}{2}\right) \tag{5.6}$$

with $\Gamma(\cdot,\cdot)$ the incomplete gamma function [1]. The bound b needs to be determined so that the bulk of the likelihood mass is included. Table 5.2 shows minimum values of b as a function of the dimension n and the tail probability α. In the current implementation of **BUMP**, b is set to 10.

Table 5.2: Minimum bound $b : F_r(b) \geq 1 - \alpha$

n	$\alpha = 0.01$	$\alpha = 0.001$
2	4.3	5.2
5	5.5	6.5
10	6.9	7.7
20	8.7	9.4

Note that convergence of the DS-iteration is no longer guaranteed by restricting the linear integration. Inadequate initial choices of expectation and covariance will lead to exclusion of likelihood mass. A good indication of this happening is when expectation and covariance from subsequent iterations tend towards zero values. A large and heavily fluctuating sampling COV is also a strong indicator for bad starting values.

5.4.3 Transform

The actual implementation of the sequence of transforms is slightly different from the theoretical sequence, described in (§ 4.5.3). Polar coordinates are randomly generated and inverse polar transform is used to obtain the standardized variables.

The second moment transformation (4.17) is not applicable as such. DS determines the first- and second-order moments in the original domain, hence, the second-moment characteristics of the unbounded parameter set are not available during the simulation. DS needs the latter statistics, since line integration along the sampled direction is performed in a standardized metric with the origin at the expectation of the likelihood mass.

From experience with **BUMP**, it became evident that the sensitivity to the choice of the covariance matrix (scale parameter) is quite low, so that standardization in the unbounded domain using a first-order approximation (§ 5.2.2) of the covariance is sufficiently accurate. This is not the case with the expectation (location parameter). Convergence of the integration strongly depends on accurate estimation of the location of the origin of the standardized space, more so if the likelihood is strongly concentrated around its expectation (e.g. in the case of large data sets).

BUMP rescales the likelihood in the unbounded domain, using the first-order estimate of the covariance matrix. If the shift is also performed in the unbounded domain, only a first-order estimate of the expectation of the likelihood would be available. This inexact estimate can jeopardize convergence of the DS algorithm. The shift from origin of the standardized space to the expectation of the likelihood is therefore performed in the original domain using the centered transformation $\underline{T}_c^{-1}$ (§ 5.2.1). The centered transformation uses the expectation $\hat{\underline{\theta}}$ in the original domain, as it was obtained from the previous iteration.

5.4.4 Coefficient of variation

Bayesian statistics are obtained from the division of two integrals (4.1), since the unknown proportionality constant of the likelihood has to be eliminated

$$\mathrm{E}[w(\underline{\theta})] = \frac{I(w(\underline{\theta}))}{I(w(\underline{\theta}) \equiv 1)} \tag{5.7}$$

Both integrals are individually evaluated by directional sampling. The COV on each of the sampling results is calculated using the estimator [40]

$$\hat{\delta}_I^{(N)} = \frac{1}{\hat{\mathrm{E}}^{(N)}[I]} \sqrt{\frac{1}{N-1}\left\{\frac{1}{N}\left(\sum_{i=1}^{N}[I^{(i)}]^2\right) - \left(\hat{\mathrm{E}}^{(N)}[I]\right)^2\right\}} \tag{5.8}$$

with $\hat{\delta}_I^{(N)}$ the estimator of the COV on the Bayesian integral I after N simulations, $\hat{\mathrm{E}}^{(N)}[I]$ the estimate of the integral I and $I^{(i)}$ the result of a single sample.

The COV of the fraction (5.7) is obtained from a first-order approximation [4],

$$\hat{\delta}_{\mathrm{E}}^{(N)} = \sqrt{\left(\hat{\delta}_I^{(N)}\right)^2 + \left(\hat{\delta}_1^{(N)}\right)^2} \tag{5.9}$$

under the assumption that no correlation exists between estimators of the COV on the numerator and the denominator. The validity of this assumption is questionable when the approximate integration method contains systematic errors in evaluating the two integrals.

5.5 Gauss-Hermite quadrature

5.5.1 Transformation

The GHQ method is also implemented using a sequence of transforms. The grid of quadrature points is defined in standardized space $\underline{s}$. A second-moment transformation (4.17) rescales the parameter space, using the current estimate of the covariance matrix in the unbounded domain. The value of the parameters in the original domain is finally obtained after application of the centered transform (§ 5.2.1). The reasoning behind this transform sequence is similar to the one described for the DS-method (§ 5.4.3).

5.5.2 Cartesian product-rule

The Cartesian product-rule presents two interesting peculiarities: (i) it is an *approximate* approach to deal with multi-dimensional integration, (ii) the depth of summation is related to the actual dimensionality n of the problem, the dimension of the parameter set $\underline{\theta}$. Every term in the summation represents a single quadrature point, in which the value of the Bayesian kernel has to be evaluated. Summation is performed in a single loop, with the help of an index-vector of dimension n. The components of the vector are considered as digits of a number in base k representation, where k is the specified dimension of the grid. A particular gridpoint is associated with a single index combination. Incrementation modulo k of the index-number guarantees a summation over all gridpoints.

5.6 Joint statistics

Two types of statistics are always determined by BUMP, independent of the user-defined output: the expectation of the parameters $\underline{\theta}$

$$\hat{\underline{\theta}} = \hat{\mathrm{E}}[\underline{\theta}] = \frac{I(w(\underline{\theta}) \equiv \underline{\theta})}{I(w(\underline{\theta}) \equiv 1)} \tag{5.10}$$

and the covariance matrix

$$\hat{\text{Cov}}(\underline{\theta}) = \hat{\text{E}}[(\underline{\theta} - \hat{\underline{\theta}})(\underline{\theta} - \hat{\underline{\theta}})^t] = \frac{I(w(\underline{\theta}) \equiv (\underline{\theta} - \hat{\underline{\theta}})(\underline{\theta} - \hat{\underline{\theta}})^t)}{I(w(\underline{\theta}) \equiv 1)} \tag{5.11}$$

since both integration methods (DS, GHQ) rely on these estimates for standardization of the likelihood. Vector $\hat{\underline{\theta}}$ and matrix $\hat{\text{Cov}}$ are evaluated component per component. As was stressed before, the determination of these statistics is iterative: new estimates are calculated using the current estimates for standardization.

5.7 Marginal statistics

Marginal statistics are obtained by integrating out the *nuisance* parameters

$$M(\theta_i) = \frac{\int_{\Theta^*} w(\theta_i) L(\underline{\theta}|\underline{x}) p(\underline{\theta}) d\theta_1 \ldots d\theta_{i-1} d\theta_{i+1} \ldots d\theta_n}{I(w(\underline{\theta}) \equiv 1)} \tag{5.12}$$

BUMP takes a different approach. Only the marginal density is evaluated by integration. This is done by taking $w(\theta_i) \equiv 1$ in the numerator and by assigning a predefined set of values to the parameter θ_i. Evaluating (5.12) for each of these values generates a discrete number of points of the marginal density $f_{\theta_i}(\theta_i)$. A spline-curve is fitted to these points (using **FITPACK**). The fitted curve serves as the basis for calculation of the marginal distribution and moments.

5.8 Predictive statistics

A similar approach is taken for the determination of predictive statistics. In this case, the full set of updatable parameters $\underline{\theta}$ are considered to be "nuisance" parameters

$$P(\underline{y}) = \frac{I(w(\underline{y}|\underline{\theta}))}{I(w(\underline{\theta}) \equiv 1)} \tag{5.13}$$

Points of the predictive density are determined by evaluation of the integral (5.13) for a predetermined set of values $\underline{y}_i$ of the dependent variable. Curve fitting allows calculation of the predictive distribution and moments.

6. Examples

Bayesian model assessment is demonstrated using real engineering modeling problems. In the first example a mathematical model is used to relate the elastic modulus of concrete to its compressive strength. This problem is particularly interesting, since exact analytic solutions of the updating formula are available [13], making it possible to check the accuracy of BUMP. The second problem is of such complexity that only numerical analysis using BUMP will enable the Bayesian assessment of parameters and of the dependent variable. The model treated in the second example predicts the time-dependent behavior of the compressive strength of concrete. In this section, a full problem description is given, the likelihood function is constructed and weight-functions for predictive analysis are given. Selected BUMP input and output files are reproduced in appendix (§ D.)

6.1 Elastic modulus of concrete

6.1.1 Problem definition

Given the fact that testing for the elastic modulus is more expensive and elaborate than the simple compressive strength test of concrete, attempts to relate compressive strength to Young's modulus have regularly been made. Both the ACI Building Code 318R-83 [2] and the ACI Nuclear Safety Structures Code 349-27 [2] take PAUW's [35] empirical formula

$$E_c = a w^{3/2} \sqrt{f'_c} \tag{6.1}$$

as a density (w) dependent mathematical relation between elastic modulus E_c and compressive strength f'_c of the concrete batch under study. PAUW used linear regression on a large set of data to derive the constant $a = 33$. The ACI codes further suggest that, for normal weight concrete, equation (6.1) can be simplified to

$$E_c = A \sqrt{f'_c} \tag{6.2}$$

with $A = 57,000$ when both E_c and f'_c are expressed in psi. ACI 363R-21 [2] suggests using a modified expression for high-strength concrete ($f'_c > 6,000$ psi.) [33].

Bayesian model assessment is used to evaluate the simplified ACI relation (6.2). Variation on the parameters will be a strong indicator for the quality of the model. Predictive calculation of E_c will reflect all uncertainties involved.

6.1.2 Data

Tests were conducted at the University of California at Berkeley for the measurement of E_c and f'_c, 28 days after the concrete samples were cast. Experimental results are summarized in Table 6.1.

Table 6.1: Berkeley test series

$\{E_c\}_k$ ($\times 10^6$ psi)				
4.13	3.86	3.92	4.27	4.15
3.66	3.91	3.98	4.15	4.25

For the measurement of f'_c, 20 samples were taken from the same concrete batch. The compressive strength was found to be adequately represented by a log-normal distribution with mean $\mu_f = 5,430$ psi and coefficient of variation (COV) $\delta_f = 0.023$. The COV of the measured f'_c population is unusually small due to strict controls observed in the production and testing of the specimens.

A second series of measurements is extracted from SHIDELER's [41] and HANSON's [22] publications, which also provided the data for PAUW's study [35]. Both authors perform a vast comparative study of lightweight aggregates in different concrete mixes. Only one single aggregate produced a concrete that qualified as having normal weight (aggregate 8 [41][22]). Every mix with the latter aggregate is chosen so that a target compressive strength is approximately obtained. The data (for concrete tested at 28 days) is reproduced in Table 6.2.

Table 6.2: SHIDELER–HANSON test series

j	μ_{fj} (psi)	δ_{fj}	$\{E_c\}_k$ ($\times 10^6$ psi)	
1	3140	0.036	3.20	3.53
2	3885	0.049	3.22	3.82
3	4340	0.042	3.43	3.84
4	5270	0.030	3.82	4.01

SHIDELER and HANSON distinguish between wet and dry storage conditions for the aggregates. Since this distinction is irrelevant in equation (6.2), the associated measurements are considered as equivalent data. While 10 to 15 control cylinders were used to check the com-

pressive strength of each mix, the COV of the population was not reported. The latter COV was estimated by considering the reported strengths of wet and dry storage conditions as samples of the same population. The obtained coefficients agree with nominal values given by SHIDELER [41].

While for a field test the COV on the compressive strength might be larger than the one of the example data, the mathematical methods employed to predict the model would remain equally valid.

6.1.3 Bayesian Model Assessment

Single Mix of Concrete

For the convenience of subsequent analysis [13], the logarithm of model (6.2) is analyzed. This conversion in no way jeopardizes the accuracy and illustrative quality of the results. Statistical uncertainty caused by the inadequate empirical relation is considered by introducing a correction factor γ to account for the model error

$$D = \theta + 0.5F + \gamma \tag{6.3}$$

with $D = \log(E_c)$, $\theta = \log(A)$ and $F = \log(f'_c)$. The influence of the error in measuring E_c is incorporated in γ.

Testing procedures do not allow to measure the exact relation between E_c (or D) and f'_c (or F) for each specimen: one is only able to determine a relation between E_c and the *population* of f'_c, inducing additional statistical uncertainty in the model.

The k^{th} observation obeys $D_k - \theta = 0.5F_k + \gamma_k$, where F_k and γ_k are samples, respectively, of F and γ and D_k is the measured value of $\log(E_c)$. Assuming a lognormal density for f'_c, F is normally distributed with mean μ_F and standard deviation σ_F^2. Taking the normality assumption for γ (§ 2.3.2), the following likelihood function is obtained

$$L(\theta, \sigma_\gamma^2) \propto (0.25\sigma_F^2 + \sigma_\gamma^2)^{-n/2} \exp\left\{-\frac{1}{2}\left[\sum_{k=1}^{n} \frac{(D_k - \theta - 0.5\mu_F)^2}{(0.25\sigma_F^2 + \sigma_\gamma^2)}\right]\right\} \tag{6.4}$$

where the mean of γ is zero and σ_γ denotes its standard deviation. After expanding the sum in the argument of the exponential function, and taking into consideration $\frac{1}{n}\sum_{k=1}^{n} D_k^2 = s_D^2 + m_D^2$, the likelihood can be written as

$$L(\theta, \sigma_\gamma^2) \propto (0.25\sigma_F^2 + \sigma_\gamma^2)^{-n/2} \exp\left\{-\frac{1}{2}\left[\frac{(\theta - (m_D - 0.5\mu_F))^2 + s_D^2}{(0.25\sigma_F^2 + \sigma_\gamma^2)/n}\right]\right\} \tag{6.5}$$

with m_D the sample mean and s_D^2 the sample variance of the observations D_k.

In the case where no prior information is given on θ (locally uniform prior [8], $p(\theta) \equiv c$), it is easy to see from (6.5) that the conditional distribution of θ given σ_γ^2 is normal. Conditional and unconditional moments of θ are [13].

$$\mu_{\theta|\sigma_\gamma^2} = \mu_\theta = m_D - 0.5\mu_F \tag{6.6}$$

$$\sigma^2_{\theta|\sigma_\gamma^2} = \frac{0.25\sigma_F^2 + \sigma_\gamma^2}{n} \quad \Rightarrow \quad \sigma_\theta^2 = \frac{0.25\sigma_F^2 + \mathrm{E}[\sigma_\gamma^2]}{n} \tag{6.7}$$

The marginal density of σ_γ^2 is obtained by multiplying the likelihood with the prior density and dividing through the conditional PDF of $\theta|\sigma_\gamma^2$

$$f(\sigma_\gamma^2) = \frac{f(\theta, \sigma_\gamma^2)}{f(\theta|\sigma_\gamma^2)} \propto \frac{L(\theta, \sigma_\gamma^2)p(\theta)p(\sigma_\gamma^2)}{f(\theta|\sigma_\gamma^2)} \tag{6.8}$$

Box and Tiao [8] prove that a locally uniform prior ($p(\sigma_\gamma^2) \equiv c$) still carries information about the scale parameter σ_γ^2. The non-informative prior [8] is

$$p(\sigma_\gamma^2) \propto (0.25\sigma_F^2 + \sigma_\gamma^2)^{-1} \tag{6.9}$$

Hence, the marginal density of σ_γ^2 is proportional to

$$f(\sigma_\gamma^2) \propto (0.25\sigma_F^2 + \sigma_\gamma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2}\frac{n\, s_D^2}{(0.25\sigma_F^2 + \sigma_\gamma^2)}\right\} \tag{6.10}$$

with inverse proportionality factor

$$\begin{aligned}\int_0^\infty (0.25\sigma_F^2 + \sigma_\gamma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2}\frac{n\, s_D^2}{(0.25\sigma_F^2 + \sigma_\gamma^2)}\right\} d\sigma_\gamma^2 \\ = \Gamma\left(\frac{n-1}{2}, \frac{ns_D^2}{0.5\sigma_F^2}\right)\left(\frac{ns_D^2}{2}\right)^{-(n-1)/2}\end{aligned} \tag{6.11}$$

and $\Gamma(p, a) = \int_0^a t^{(p-1)} \exp\{-t\}dt$, the incomplete gamma function [1].

The expectation of the model variance σ_γ^2 is equal to

$$\mathrm{E}[\sigma_\gamma^2] = \frac{ns_D^2}{2}\frac{\Gamma\left(\frac{n-3}{2}, \frac{ns_D^2}{0.5\sigma_F^2}\right)}{\Gamma\left(\frac{n-1}{2}, \frac{ns_D^2}{0.5\sigma_F^2}\right)} - 0.25\sigma_F^2 \tag{6.12}$$

The data-based variance of θ is obtained by introducing (6.12) in (6.7)

$$\sigma_\theta^2 = \frac{s_D^2}{2} \frac{\Gamma\left(\frac{n-3}{2}, \frac{ns_D^2}{0.5\sigma_F^2}\right)}{\Gamma\left(\frac{n-1}{2}, \frac{ns_D^2}{0.5\sigma_F^2}\right)} \tag{6.13}$$

If the second argument of $\Gamma(p,a)$ is sufficiently large (e.g. $a > 50$ when $p = 10$ [1]), the incomplete gamma function is adequately approximated by the gamma function $\Gamma(p)$. Under this condition, equation (6.13) simplifies to

$$\sigma_\theta^2 = \frac{s_D^2}{n-3} \tag{6.14}$$

Different Concrete Mixes

Each individual mix is characterized by a likelihood (6.5). If independence between the data originating from the different concrete mixes is assumed, the likelihood becomes

$$L(\theta, \sigma_\gamma^2) \propto \prod_{j=1}^{k} (0.25\sigma_{Fj}^2 + \sigma_\gamma^2)^{-n_j/2} \exp\left\{-\frac{1}{2}\left[\frac{(\theta - (m_{Dj} + 0.5\mu_{Fj}))^2 + s_{Dj}^2}{(0.25\sigma_{Fj}^2 + \sigma_\gamma^2)/n_j}\right]\right\} \tag{6.15}$$

with μ_{Fj} and σ_{Fj} the mean and the standard deviation of the log-transformation of the compressive strength for each mix j of concrete and k the number of different mixes. The sample statistics m_{Dj} and s_{Dj} of D are determined for each different concrete mix j. A number n_j observations of D is performed on concrete mix j. The compound non-informative prior on σ_γ^2 is

$$p(\sigma_\gamma^2) \propto \prod_{j=1}^{k} (0.25\sigma_{Fj}^2 + \sigma_\gamma^2)^{-n_j/n_t} \tag{6.16}$$

with n_t the total number of observations ($= \sum_{j=1}^{k} n_i$).

The complexity of the likelihood and the prior make further analytical developments difficult. One has to resort to computational integration to determine posterior and predictive statistics.

Predictive Analysis using the Calibrated Model

Having calibrated the model (6.3), i.e. having determined the PDF of parameters θ and σ_γ^2, it should be put to profitable use. Predictive analysis on the calibrated model will give the statistical characteristics of Young's modulus E_c for a specific mix of concrete. The compressive strength f_c' of the mix is assumed to have a lognormal population with known mean μ_{fp} and standard deviation σ_{fp}. Hence, the statistics of the normal variable $F = \log(f_c')$ are easily determined.

Predictive analysis for model (6.3), calibrated with a single mix of concrete, is straightforward: D is the sum of normal variates. It is therefore also normal, with mean $\mu_D = \mu_\theta + 0.5\mu_{Fp}$ and variance $\sigma_D^2 = \sigma_\theta^2 + 0.25\sigma_{Fp}^2 + \mathrm{E}[\sigma_\gamma^2]$. E_c is then lognormal.

This analytical solution is not available when different mixes of concrete are used to calibrate the model. The predictive statistics have to be determined numerically.

Taking into account the initial (normality) assumptions on F and γ, it is convenient to replace $0.5F+\gamma$ by a single normal variable G, with mean $\mu_G = 0.5\mu_{Fp}$ and variance $\sigma^2_{G|\sigma_\gamma^2} = 0.25\sigma_{Fp}^2 + \sigma_\gamma^2$. The latter variance is conditional on the value of σ_γ^2. The joint, posterior density of θ and G is

$$f''(\theta, G) \propto \int_0^\infty (0.25\sigma_{Fp}^2 + \sigma_\gamma^2)^{-1/2} \exp\left\{-\frac{1}{2}\frac{(G-0.5\mu_{Fp})^2}{(0.25\sigma_{Fp}^2+\sigma_\gamma^2)}\right\} L(\theta, \sigma_\gamma^2) p(\sigma_\gamma^2) d\sigma_\gamma^2 \quad (6.17)$$

Prediction of the stochastic behavior of D is performed by considering the sequence of probability transformations

$$T_1 : \theta, G \mapsto D : D = \theta + G \qquad T_2 : D \mapsto E_c : E_c = \exp\{D\} \quad (6.18)$$

Applying transformation theory [39], the predictive distribution of E_c is obtained

$$f(E_c) \propto \int_{-\infty}^{+\infty} \int_0^\infty \frac{1}{E_c(0.25\sigma_{Fp}^2+\sigma_\gamma^2)^{1/2}} \exp\left\{-\frac{[\log(E_c)-\theta-0.5\mu_{Fp}]^2}{2(0.25\sigma_{Fp}^2+\sigma_\gamma^2)}\right\} L(\theta, \sigma_\gamma^2) p(\sigma_\gamma^2) d\sigma_\gamma^2 d\theta \quad (6.19)$$

Integral (6.19) is of Bayesian type (see equation (4.1) in § 4.). The weight function is equal to

$$w(E_c|\theta, \sigma_\gamma^2) = \frac{1}{E_c(0.25\sigma_{Fp}^2+\sigma_\gamma^2)^{1/2}} \exp\left\{-\frac{[\log(E_c)-\theta-0.5\mu_{Fp}]^2}{2(0.25\sigma_{Fp}^2+\sigma_\gamma^2)}\right\} \quad (6.20)$$

6.1.4 Results and Interpretation

Table 6.3 displays statistics of the model parameters obtained from single mix calibration using the Berkeley data. The sample characteristics for the corrected log-model (6.3) are $\mu_F = 8.6$ and $\sigma_F = 0.023$. The characteristics of the data entering the likelihood (6.5) are $m_D = 15.2, s_D = 0.049$ and $n = 10$. Exact results (§ 6.1.3) are compared with results obtained by the optimization for the MLE estimators, as well as results of numerical integration by directional sampling (DS) and by Gauss-Hermite quadrature (GHQ).

Table 6.3: Comparison of Bayesian integration methods

	Exact	MLE	DS	GHQ
μ_θ	10.90	10.91	10.90	10.91
σ_θ	0.0185	0.0153	0.0188	0.0189
$E[\sigma_\gamma^2]$	0.0033	0.0023	0.0034	0.0034

Directional sampling results were obtained after a single iteration of 20,000 samples. Convergence to the expectation of the parameters is almost immediate, but further iterations will result in a fluctuating estimate of the variance on parameter θ.

The Gauss-Hermite quadrature was stopped after 3 iterations, using a grid of 31^2 points. Additional quadrature iterations only reduced the cross-covariance between θ and σ_γ^2 to zero, variance and expectation of the parameters did not change. Decreasing the gridsize underestimates the expectation of the parameters and jeopardizes convergence, since part of the likelihood mass is left out.

GHQ is much more efficient than DS, and generally provides accurate results. This conclusion has been confirmed in other applications. In cases where GHQ is likely to produce erroneous results, i.e. cases where the likelihood does not comply to the "normal × polynomial" assumption (§ 4.4)), DS constitutes a necessary validation method. Note that in this case the mode and expectation of θ are equal, indicating that the likelihood closely approximates a normal distribution. Sampling becomes prohibitively expensive when it is used for accurate evaluation of marginal or predictive PDF's.

Table 6.4 compares statistics of the model parameters obtained by single and multiple mix calibrations. The statistics were calculated using the Gauss-Hermite quadrature method.

The change of the standard deviation on the parameter θ and of the modeling variance σ_γ^2 is to be expected. The single mix calibration looks at the accuracy of the model for a limited

Table 6.4: Comparison of the calibration methods

	single mix	multi-mix
μ_θ	10.91	10.94
σ_θ	0.0189	0.0254
$E[\sigma_\gamma^2]$	0.0034	0.0047

interval of f_c'. Multi-mix calibration uses information of a number of such intervals, each describing a different "area" of the curve. Multi-mix calibration thus gives a global appreciation of the lack of fit of the model. Hence, the increase in uncertainty is comprehensible. Both calibration methods have their specific use. If a general model, e.g. for use in concrete codes, is to be derived, *global calibration* is recommended. Single mix calibration (or *local calibration*) is particularly suited for a situation in which a known concrete mix is going to be used throughout the construction process. It tailors the model to the specific mix of the concrete.

In Figure 6.1 the marginal PDF of σ_θ^2 for both single mix and multi-mix calibration is displayed. Figure 6.2 does the same for the variance on the modeling error. The marginals were obtained using Gauss-Hermite quadrature.

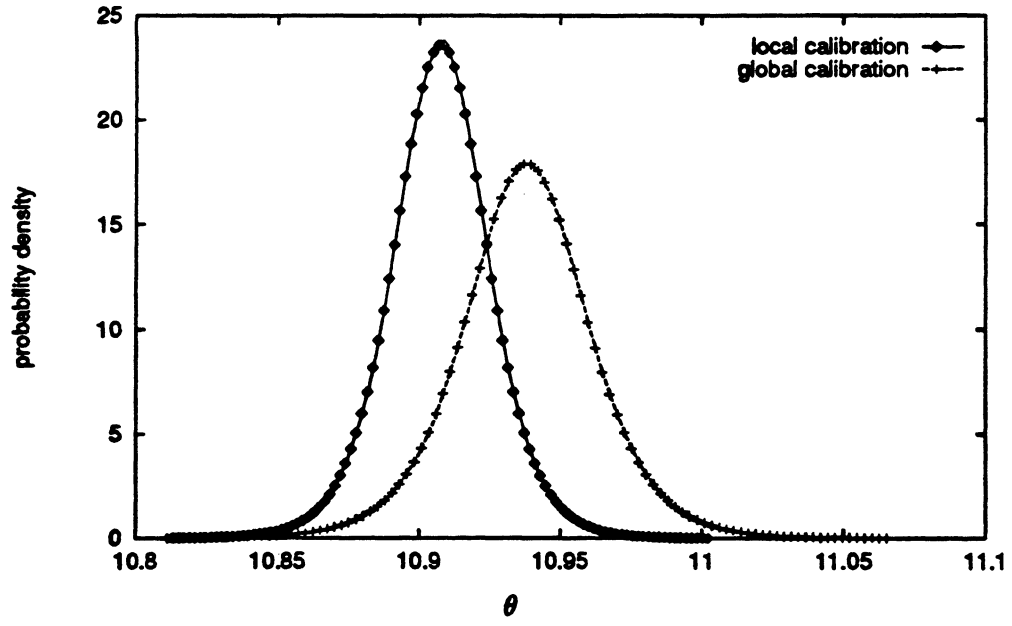

Figure 6.1: PDF of the model parameter θ

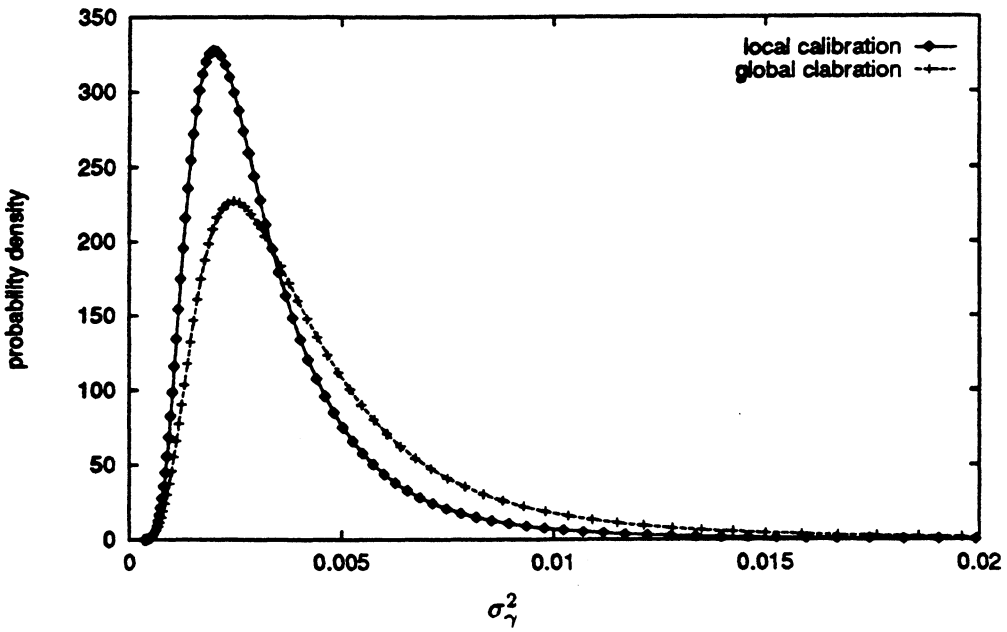

Figure 6.2: PDF of the variance on the model error σ_γ^2

The calibrated model is used to determine the statistical characteristics of E_c for a mix of concrete with μ_{fp} = 5430 psi and σ_{fp} = 125 psi. These are the characteristics of the mix used for the single-mix calibration. Figure 6.3 shows the predictive density (PDF) of E_c for both calibration types while the cumulative density (CDF) is shown in Figure 6.4.

Modelling and measurement errors are represented by the the random correction parameter γ. The uncertainty on parameter θ is caused by all sources of uncertainty: lack of data, randomness in f'_c, measurement error and modeling error. The predictive analysis showed (§ 6.1.3) that the variance of $D = \log(E_c)$ has three distinct contributions: $\sigma_D^2 = \sigma_\theta^2 + 0.25\sigma_F^2 + \mathrm{E}[\sigma_\gamma^2]$. Hence, The lack of fit of the model is evaluated by focussing on the variances of the model parameter σ_θ^2 and the correction factor $\mathrm{E}[\sigma_\gamma^2]$. The model COV $\hat{\delta}_{E_c}$, based on $\hat{\sigma}_D = \sqrt{\sigma_\theta^2 + \mathrm{E}[\sigma_\gamma^2]}$ and determined for a specific value of f'_c (or F), can be an indicator for the modeling accuracy.

Table 6.5 compares the model COV $\hat{\delta}_{E_c}$ with the predictive COV δ_{E_c}. It is clear that the bulk of the uncertainty is due to model imperfection (as opposed to the variability arising from the randomness of f'_c). This clearly indicates the importance of assessing model uncertainty.

Finally, it is worth while to compare the value of E_c predicted by the ACI relation with the results obtained by Bayesian predictive analysis. The ACI relation (6.2) needs *nominal values* of the dependent and independent variables. If (6.2) is used in reliability-based design, the

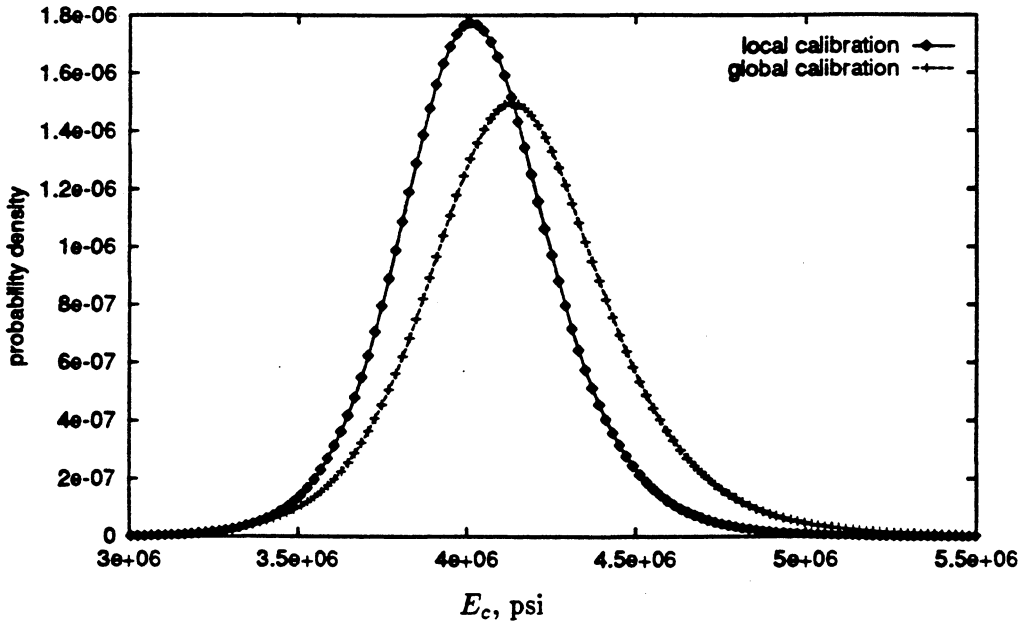

Figure 6.3: Predictive PDF of Young's Modulus E_c ($\mu_{fp} = 5340$, $\delta_{fp} = 0.023$)

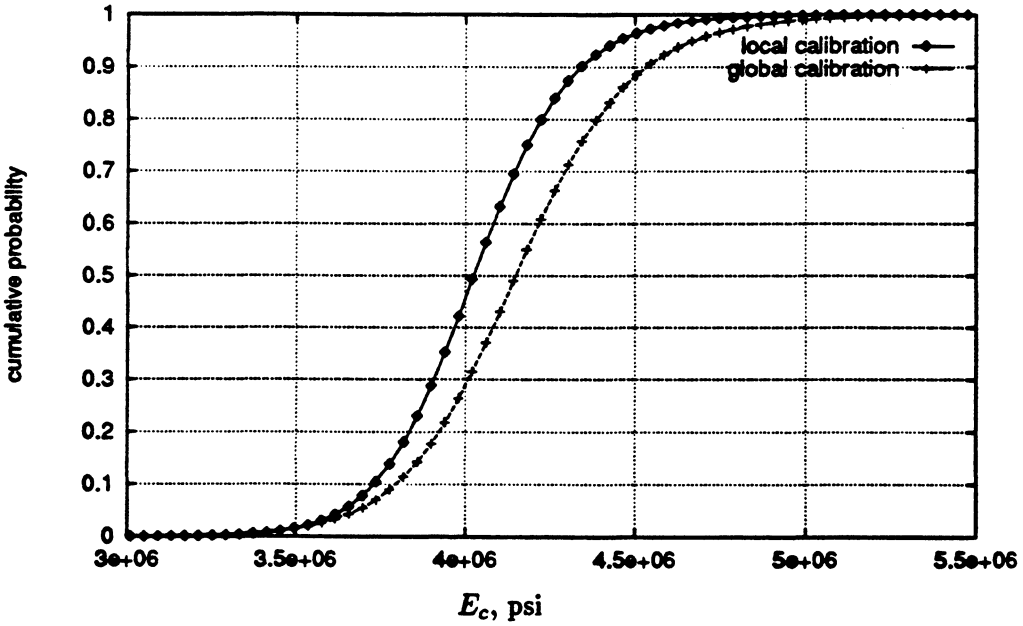

Figure 6.4: Predictive CDF of Young's Modulus E_c ($\mu_{fp} = 5340$, $\delta_{fp} = 0.023$)

lower α-quantile is used as *characteristic value* E_{ck} ($\Pr(E_c \leq E_{ck}) = \alpha$). Indeed, both f'_c and E_c generally provide resistance in a structural assembly. This characteristic value is obtained

Table 6.5: Predictive and model COVs (μ_{fp} = 5430 psi, δ_{fp} = 0.023)

	μ_D	$(\sigma_\theta^2 + \mathrm{E}[\sigma_\gamma^2])^{1/2}$	μ_{E_c}	δ_{E_c}	$\hat{\delta}_{E_c}$
single mix	15.21	0.0613	4.03×10^6 psi	6.3 %	6.1 %
multi-mix	15.24	0.0731	4.16×10^6 psi	7.4 %	7.3 %

by either plugging in the associated characteristic value of the compressive strength f'_{ck} in the ACI formula, or by finding the appropriate quantile of the predictive distribution of E_c obtained from Bayesian analysis. Table 6.6 shows a comparison of both values.

Table 6.6: Comparison of characteristic values of E_c (μ_{fp} = 5430 psi, δ_{fp} = 0.023)

	ACI-318		BUMP	
			single batch	multi-batch
α	f'_{ck} (psi)	E_{ck} ($\times10^6$ psi)	E_{ck} ($\times10^6$ psi)	E_{ck} ($\times10^6$ psi)
50 %	5430	4.20	4.02	4.15
10 %	5270	4.14	3.74	3.80
5 %	5225	4.12	3.64	3.68
1 %	5135	4.08	3.45	3.44

The characteric values obtained from the ACI-318 relation (6.2) seem to be grossly unconservative when used within a reliability context. The large difference between the ACI formula and the Bayesian predictive distribution is to be ascribed to the neglect of model uncertainty in the former. The characteristic value according to ACI-318 only incoporates the stochastic nature of the concrete mix under study. This large difference most certainly stresses the supremacy of a Bayesian approach for mathematical modeling of physical phenomena when accounts for all different sources of uncertainty are present.

6.2 Strength of Aging Concrete

6.2.1 Problem definition

Prediction of concrete strength as a function of time has always been of interest to construction engineers, since it largely determines the progress on the construction site. ACI-

Commitee 209 [2] recommends use of the expression

$$f'_c(t) = f'_{c28}\left(\frac{t}{a+bt}\right) \tag{6.21}$$

for modeling the compressive strength gain of concrete with time [33]. The compressive strength f'_c at time t is compared with the reference value f'_{c28}, obtained after 28 days. Values of the coefficients a and b can vary within a wide range, depending on the type of cement and curing used. Values of interest for the subsequent analysis are associated with moist-cured concrete containing Type I cement: $a = 4.0$ and $b = 0.85$ [2]. The experimental data used in this section stems from such concrete.

The ACI-209 model (6.21) is calibrated by a Bayesian procedure. Because of the destructive testing procedure, we are again faced with a problem where no direct correspondence exists between measured $f'_c(t)$ and f'_{c28} values. Simple regression analysis cannot assimilate this type of relation between model variables.

6.2.2 Data

Bayesian model assessment is applied for optimal inference on data acquired at the University of California, Berkeley (Table 6.7). Strict controls observed during casting and during the experimentation produced data of exceptional quality (small COV, for a given time).

The compressive strength at 28 days after casting has a lognormal population with mean $\mu_{f_{c28}} = 5431$ psi and standard deviation $\sigma_{f_{c28}} = 125$ psi ($\delta = 2.3\%$).

A second data-set is taken from literature: KLIEGER [24] reports on an extensive campaign to assess the time-variant behavior of cement as principal component of the concrete mix. KLIEGER's data was part of the original data-set for the regression analysis on formula (6.21).

The 28-days compressive strength of KLIEGER's concrete sample has a lognormal population with mean $\mu_{f_{c28}} = 5281$ psi and standard deviation $\sigma_{f_{c28}} = 542$ psi ($\delta = 10.3\%$). It should be mentioned that the data extracted from [24] for Type I cements. Since formula ACI-209 (6.21) does not distinguish between members of the latter family, amalgamation of the data was deemed valid.

6.2.3 Bayesian analysis

Considering the logarithmic model for sample k, we can write

$$F_{tk} = F_{28k} + \log[t_k/(a+bt_k)] + \gamma_k \tag{6.22}$$

Table 6.7: Berkeley data for $f'_c(t)$

$\{f'_c(t)\}_k$ psi.						
3 days	7 days	21 days	28 days			
2680	3960	5120	5530	5460	5360	5490
2680	3970	4960	5250	5070	5500	5410
2680	3800	5030	5410	5390	5460	5620
	3910		5430	5470	5450	5650
	3870		5500	5350	5480	5350
60 days	90 days	120 days	180 days	210 days		
5730	6030	6135	5930	6400		
5580	6110	6300	6100	6250		
5170	5760	6045	6250	6310		
6095		6215				
6490		6055				

Table 6.8: KLIEGER'S data for $f'_c(t)$

$\{f'_c(t)\}_k$ (psi)					
1 day	7 days	28 days	90 days	360 days	1080 days
780	4210	5870	6390	7020	6910
580	3800	5710	6650	7010	7380
890	2740	4730	5760	6400	6750
800	3480	5190	5700	6540	7030
1260	4730	5830	6280	6440	6360
1020	4000	5820	6520	7060	7640
740	3480	5770	6410	6560	6890
670	3600	5600	6230	6540	6810

with $F_{tk} = \log f'_{ck}(t)$, $F_{28} = \log f'_{c28}$ and γ the Gaussian, random correction accounting for model uncertainty and measurement error. The population of F_{28} is normally distributed

with μ_F and σ_F. The likelihood function is proportional to

$$L(a,b,\sigma_\gamma^2|t) \propto (\sigma_F^2+\sigma_\gamma^2)^{-n/2} \exp\left\{-\frac{1}{2(\sigma_F^2+\sigma_\gamma^2)}\sum_1^n\left[F_{tk}-\log\left(\frac{t_k}{a+bt_k}\right)-\mu_F\right]^2\right\} \tag{6.23}$$

The complexity of the likelihood does not allow an analytical approach.

We assume to have local uniform priors on parameters a and b. The prior density of the variance of the model error σ_γ^2 is chosen to be non informative [8]

$$p(\sigma_\gamma^2) = (\sigma_F^2+\sigma_\gamma^2)^{-1} \tag{6.24}$$

Predictive analysis using the calibrated model requires a double probability transform: (i) $T_1 : (a,b,F_{28},\gamma) \mapsto F_t$, (ii) $T_2 : F_t \mapsto f_c'(t)$. Applying probability transformation theory [3], the predictive distribution of $f_c'(t)$ at time t, is found by evaluating

$$f(f_c'|t) \propto \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{(\sigma_\gamma^2+\sigma_F^2)f_c'} \exp\left\{-\frac{[\log(f_c')-\log(t/(a+bt))-\mu_F]^2}{2(\sigma_\gamma^2+\sigma_F^2)}\right\} L(a,b,\sigma_\gamma^2|t)\,da\,db\,d\sigma_\gamma^2 \tag{6.25}$$

6.2.4 Results and Interpretation

Both directional sampling and quadrature produce the same results. The moments obtained by GHQ and DS are remarkably close to the initial MLE estimates especially for the Berkeley data. Sampling was performed using a single iteration of 1000 samples, but reasonable results were already obtained using only 100 simulations. Three iterations of 11^3 points were sufficient to produce accurate quadrature results. These facts indicate that the likelihood not only approximates a normal density, but is also strongly concentrated around its mean/mode. This is caused by the fairly large data-set available for this problem. The large dispersion of the KLIEGER data produces a less pronounced concentration effect.

The joint statistics of the parameter set obtained from Bayesian integration are reproduced in Table 6.9. Entries in the table represent the expectation or variance / covariances of the parameters shown in rows and columns.

Close scrutiny of Table 6.9 leads to a number of important observations. First, the parameter b is nearly degenerate (very small variance). This means that there exists an almost deterministic relation between the 28-days compressive strength f_{c28}' and the compressive strength $f_c'(t \to \infty)$ the concrete will attain after a long time.

Table 6.9: Moments of parameter set

	ACI	Berkeley data				KLIEGER DATA			
		E	1	a	b	E	1	a	b
a	4.0	a	3.59	0.0107	−0.000362	a	5.42	0.0662	−0.00150
b	0.85	b	0.87	−0.000362	3.55×10^{-5}	b	0.76	−0.00150	0.000380
		σ_γ^2	0.00619	≈ 0	≈ 0	σ_γ^2	0.00720	≈ 0	≈ 0

Secondly, there exists significant (negative) correlation between parameters a and b ($\rho \approx 45\%$). Parameter b is the reciprocal of the asymptotic ($t \to \infty$) compressive strength; a is the reciprocal of the gradient of the compressive strength at time $t = 0$. Negative correlation implies that if a specific sample of b is in the lower tail of its distribution, parameter a will most likely take a value larger than its mean. Note that this is a probabilistic relation ("most likely"), not a functional one. Therefore, higher asymptotic values are likely to be reached after shorter periods of time.

The fairly large difference between the moments obtained from the Berkeley data and those obtained from the KLIEGER data may disturbing at first glance. The larger variances of the model parameters for the KLIEGER data are clearly due to the larger dispersion inherent in this data set. The large difference between the mean values however, is an indication that the parameter values strongly depend on the concrete mix. ACI 209R-9 [2] indeed reports wide ranges in which the parameters can be found: $0.05 < a < 9.25$ and $0.67 < b < 0.98$. The values given in Table 6.9 are typical (nominal) values recommended by ACI for moist cured concrete based on a Type I cement. The wide ranges of parameters, confirmed by the Bayesian analysis here,suggests that perhaps formula (6.21) should be re-assessed for different types of concrete mixes or be refined to include variables that characterize the mix. The Bayesian assessment method described in this report is particularly suited for such re-evaluation.

Integration of (6.25) produces the predictive PDF of the compressive strength f'_c as a function of time. The evolution of the mean strength is shown in Figure 6.5 for both data-sets. The upper diagram displays the evolution of the compressive strength during the first three years. The lower plot is restricted to 60 days after casting in order to highlight the early evolution of the concrete strength. Both plots show the recommended ACI value of $f'_c(t)$ based on the corresponding mean of f'_{c28}.

The comparison between the nominal ACI relation and the mean predictive values of $f'_c(t)$ reveals that the Berkeley data produces a mean relation which is in close agreement with the

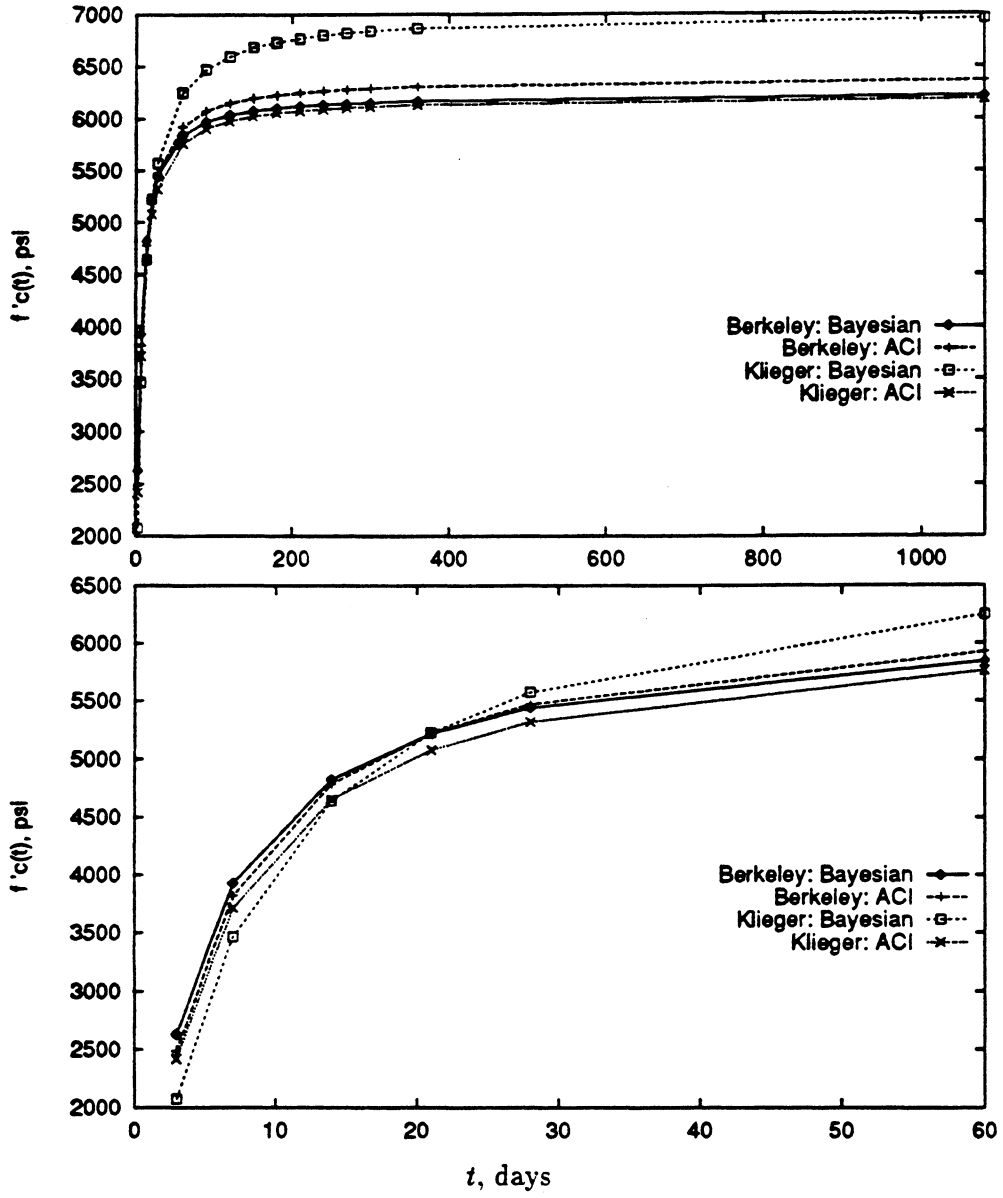

Figure 6.5: Expectation of compressive strength as a function of time

ACI formula for the mean value of $f'_c(t)$. This is not true for the KLIEGER data-set. This observation is further accentuated in Table 6.10, where the lower 2σ bounds of $f'_c(t)$ obtained by Bayesian updating are compared with the ACI-209 formula using the lower 2σ bounds on f'_{c28}. These bounds may be considered as characteristic values of the strength for us in a partial factor design approach.

The comparison reveals that the ACI formula provides inconsistent predictions, both on

Table 6.10: Lower 2σ bound on the compressive strength $f'_c(t)$

	$f'_c(t)$ (psi)			
	Berkeley data			
t	7 days	28 days	90 days	360 days
ACI	3647	5184	5795	6020
BUMP	3652	5068	5559	5745
	Klieger data			
t	7 days	28 days	90 days	360 days
ACI	2950	4223	4657	4869
BUMP	2525	4071	4729	4991

the conservative and unconservative sides. This inconsistency has obvious ramifications on reliability-based design of structures.

Figure 6.6 shows the PDFs of $f'_c(t)$ for different times based bon the KLIEGER data, clearly revealing the increase of the predictive variance with time. The predictive COV remains practically constant with time. The model distributes uncertainty, originating from lack of data or dispersion of data at a certain moment in time, over the whole time scale. The predictive COV of the model calibrated with the KLIEGER data is close to 13.5%, while the Berkeley data induces a COV of 3.4% in the predictive distribution.

The large COV of the KLIEGER data is only partially caused by the large dispersion of the data. The previous discussion has emphasized the lack of fit of the recommended ACI-209 model as far as the KLIEGER data is concerned. A major part of the predictive COV is attributed to the modeling uncertainty.

Predictive Bayesian analysis enables the analyst to produce graphs that are particularly useful to the field engineer. For example, Figure 6.7 shows mean and mean$\pm\sigma$ bounds on compressive strength as a function of time. The curves are based on the predictive analysis performed on the model calibrated with the KLIEGER data.

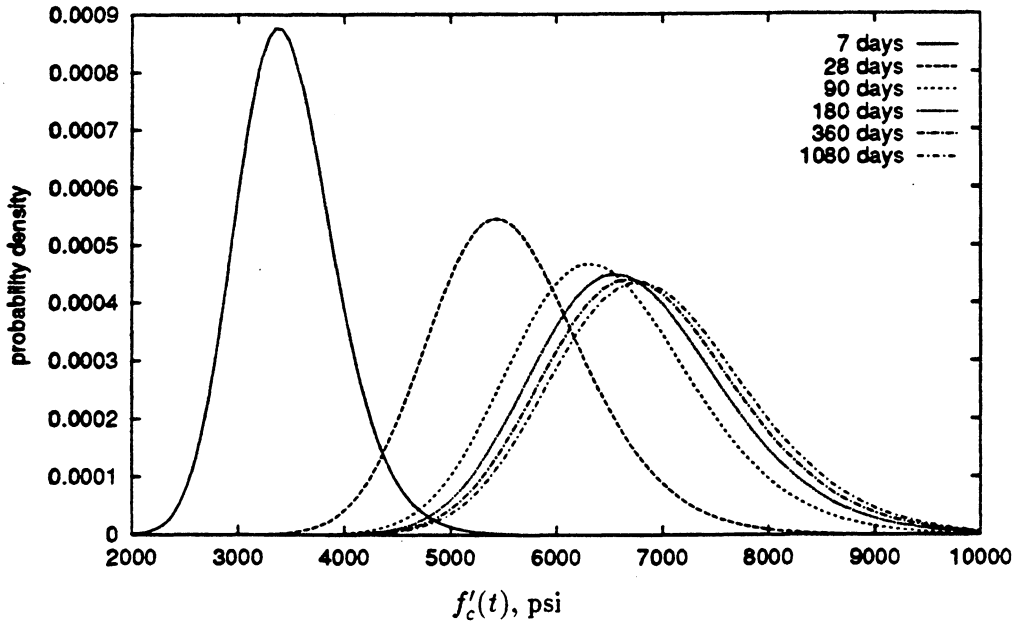

Figure 6.6: PDF of the compressive strength at different times (KLIEGER data)

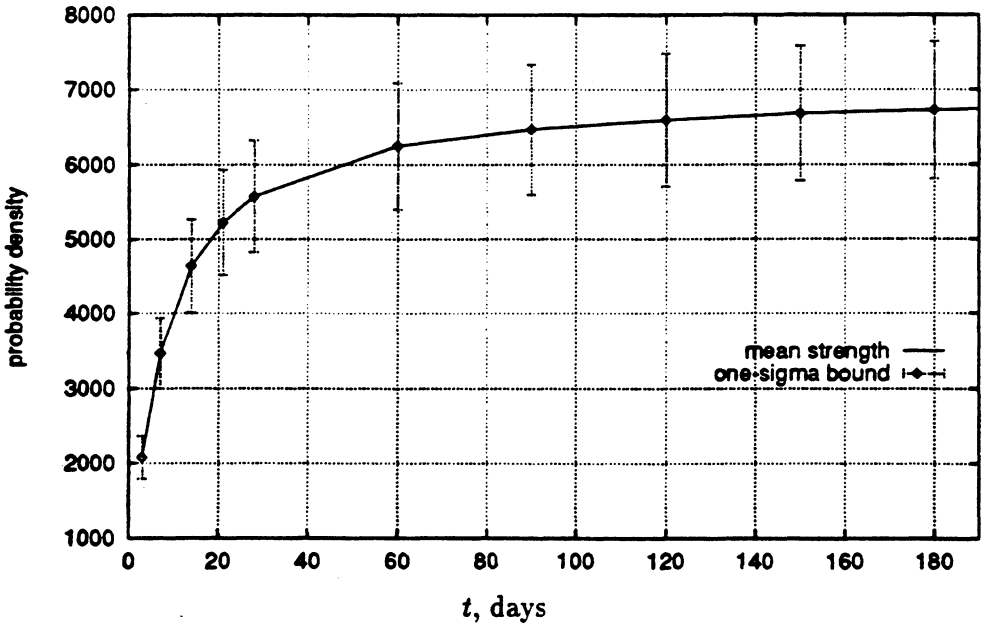

Figure 6.7: Mean and mean$\pm\sigma$ bounds on $f'_c(t)$ (KLIEGER data)

7. Summary and Conclusion

The theory behind Bayesian model updating (BMU) has been detailed in this report. It has been shown that practical application of BMU requires efficient numerical procedures for the integration of the Bayesian kernel. Different methods have been reviewed, two of them (directional sampling and Gauss-Hermite quadrature) are implemented in the program BUMP.

Examples have shown that Bayesian model updating in general and BUMP in particular are powerful tools for engineering modeling and inference. From the different integration methods used, Gauss-Hermite quadrature is the most efficient: it provides accurate results with minimal effort, provided the likelihood resembles a "polynomial × normal" form. Directional sampling is found to be unsuitable for marginal or predictive analysis. Nevertheless, it is still an important tool for Bayesian analysis, as it provides a control method for the determination of joint, second-moment statistics of a parameter set."Exact" results will be obtained, provided computer resources and time are invested in the calculation. This characteristic is important when dealing with ill-behaved (not approximately normal) likelihood functions in order to verify results obtained from the GHQ method.

The examples have further demonstrated the importance of the Bayesian model assessment. The approach presented herein has proven to be an effective and systematic framework for the inclusion of all possible types of uncertainty in model calibration. It was shown that the incorporation of model uncertainty, measurement error, etc. has a strong influence on statistics to be used in a reliability context.

The Bayesian model assessment may also take into account subjective information available on the model and its parameters. This is introduce into the analysis through the prior distribution. The adequate definition of the likelihood allows the inclusion of observed data, independent of the size of the data set.

The Bayesian format is particularly suited for flexible coding purposes. The deterministic mathematical model frequently fails to cover the broad domain it is supposed to describe. Its parameters are often chosen by compromise. This was amply demonstrated by the analysis of the ACI-209 model for the compressive strength of aging concrete.

A coded Bayesian format should prescribe distributional assumptions on parameters, dependent and independent variables. The accumulated experience of prior calibration campaigns can easily be introduced at the level of the prior distribution. The field engineer would then update the model with the field data that is directly related to the application of interest. The engineer then has a model that is an optimal blend of the problem-specific data and of past experience.

At its current stage, the research into the Bayesian model assessment has proven that in many civil engineering applications deterministic modeling can easily be replaced by the proposed stochastic modeling technique. It adds important flexibility and reliability to the modeling process for a negligible computational price.

A. BUMP: user's manual

A.1 Input

A.1.1 General remarks

BUMP reads input from the standard input device. Its commands are best grouped in an input file, which can then be submitted the program using input redirection (for both DOS and UNIX operating systems)

```
% bump < {infile}
```

with {infile} the name of the input file.

The input file consists of a number of independent sections, each having a body of data (Figure 6.8). The data structure is specific for each section. Sections are separated by a slash "/". In this chapter, sections are described in order of appearance in the file. The user will be flagged if a section is missing. Errors will occur if the prescribed order is not maintained.

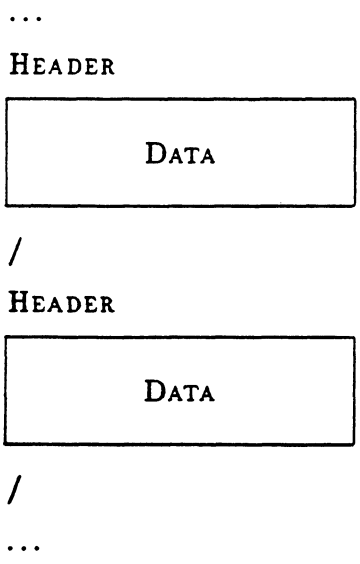

Figure 6.8: General structure of the input file

The data of a section is built up from input lines. An input line consists of different items, separated by a comma or by a space. Input following a percentage sign % is discarded, allowing

the inclusion of comments in the file. The **BUMP** line parser identifies if an item is numeric or alphabetic. Alphabetic data is truncated to 4 significant characters. **BUMP** is insensitive to case. The parser does not make any difference between real and integer numbers.

Each section activates a reading routine, which will then call upon the parser to analyze the input. Consistency of the input is not checked at parser level, but is controlled by the section-related read routines. Clear error messages (§ C.) are displayed when inconsistent data is encountered. In the following paragraphs, each individual data section will be fully described. Example input files can be found in appendix (§ D.).

A.1.2 Identifier

The file identifier is the only section not having a body. It simply consists of the header `bump`.

A.1.3 Parameter definition

Header: `parameter`

Body: `{name},{type},{range/value},{init}`

{name}: parameter identifier <char>

{type}: `updatable` or `known` <char>

{range/value}:

if type = `updatable` then
  range = {`lowerbound`},{`upperbound`}

if type = `known` then
  range = {`value`},`0`.

{init}: initial value for optimization (⇒ MLE)

Remarks:

- if lowerbound $\leq -10^{35}$ then lowerbound $= -\infty$
  if upperbound $\geq +10^{35}$ then upperbound $= +\infty$
- a **known** parameter has a fixed value throughout the analysis

A.1.4 Prior density

Header: `prior`

Body: `{name},{type},{p1},{p2},{p3},{p4}`

{name}: name of considered parameter <char>

{type}: type of prior density <char>

- **none**
- **normal**
- **lognormal**
- **gamma**
- **beta**
- **uniform**
- **user**

{p*}: distribution parameters <numr> (table 6.11)

Remarks:

- all updatable parameters have to be included in the prior list

Table 6.11: Prior density library

name	PDF	p1	p2	p3	p4
normal	$\exp\{-(x-\mu)^2/2\sigma^2)\}/\sqrt{2\pi}\sigma$	μ	$0<\sigma$	-	-
lognormal	$\exp\{-(\log(x)-\lambda)^2/2\eta^2)\}/\sqrt{2\pi}\eta x$	λ	$0<\eta$	-	-
gamma	$\lambda(\lambda x)^{k-1}\exp\{-\lambda x\}/\Gamma(k) \quad x \geq 0$	$0<\lambda$	$0<k$	-	-
beta	$(x-a)^{q-1}(x-b)^{r-1}/\{B(q,r)(b-a)^{q+r-1}\}$ $a \leq x \leq b$	$0<q$	$0<r$	a	b
uniform	$(b-a)^{-1} \quad a \leq x \leq b$	a	b	-	-

A.1.5 Observations

Header: **observations**

Body: first line: **nobs** = {nn},**nvar** = {mm}

{nn}: number of observations <numr>

{mm}: number of observed variables <numr>

followed by **nn** rows of **mm** data-items

Remarks:

- the data is *not* parsed, but read using **FORTRAN77** list directed input.

A.1.6 Execution control

Header: **execute**

Body: {**method**},{p1},{p2},{p3}

{method}: identification of integration method <char>

- method = **samp** ⇒ directional sampling
- method = **quad** ⇒ Gauss-Hermite quadrature

{p*}: integration parameters <numr>

if method = **samp** then

{p1} = generator seed

{p2} = number of iterations

{p3} = number of samples per iteration

if method = **quad** then

{p1} = number of iterations

{p2} = number of quadrature points per dimension

Remarks:

- **BUMP** uses the integration methods with the specified parameters in the order given; the second-moment results of the previous execution are used as initial values of the next integration method.
- the last integration in the sequence is used to calculate marginal and predictive statistics.

A.1.7 Output control

Header: **output**

Body: {**type**},{**name/number**},{p1},{p2},{p3},{w*}

{type}: identification of the requested output <char>

- **marginal**
- **predictive**

if type = **marginal** then

{name}: name of marginal parameter <char>

{p1} = number of increments in interval

{p2},{p3} = interval definition

if type = **predictive** then

{number}: number of user defined weightfunction ⇒ identification of dependent variable

{p1} = number of increments in interval

{p2},{p3} = interval definition

{w*}: parameters passed on to the weight function <numr>

Interval definition

default interval

{p2} = 0.

{p3} = 0.

the interval is defined as the 5σ boundary around the expectation

relative interval definition

{p2} = {k} <numr>

{p3} = 0.

the interval is defined as the $k\sigma$ boundary around the expectation

absolute interval definition

{p2} = {lowerbound} <numr>

{p3} = {upperbound} <numr>

the interval is defined as [{lowerbound},{upperbound}]

Remarks:

- σ is approximated by the square root of the appropriate diagonal element of the covariance matrix.
- predictive analysis requires the definition of the lowerbound and upperbound of the interval; relative interval definition will prompt an error message
- when a marginal density is determined by sampling, explicit definition of the interval bounds is strongly recommended; an inexact estimate of σ might produce an interval that contains no likelihood mass.

A.2 User defined subroutines

BUMP needs 3 subroutines to be supplied by the user.

- **ulhood**: the user-defined likelihood function
- **uprior**: the user-defined prior density
- **uwghtf**: the weight function (for predictive analysis)

Implementation of **ulhood** is imperative since the whole purpose of **BUMP** is the integration of a Bayesian kernel, containing a problem-related likelihood function. Definition of a prior density is only required when the **user** option is activated in the **prior** section (§ A.1.4).

The user-defined weight function is currently only used when predictive statistics need to be calculated.

The source listing of file `user.for` is reproduced in appendix (§ B.). Description of the parameters of the subroutines can be found in the comments to the source.

A.3 Precision

BUMP, and the library routines it uses, captures as many over- and underflows as possible and tries to take full advantage of the precision of the computer on which it is running. Precision parameters therefore need to be defined. The subroutine `dlmach.for` contains machine parameters for most of the common computer platforms. The user must change this routine to comply to the double precision implementation of the particular computer in use. It is currently set for the precision definition according to the IEEE-standard, standard by which most workstations and personal computers abide.

A.4 Output

BUMP writes its results to the standard output device. Device redirection can be used to capture the output in a file. The typical call to BUMP would look like (DOS/UNIX)

```
% bump < {infile} > {outfile}
```

with {infile} the name of the input file containing the BUMP commands and {outfile} the name of the file in which the results will be collected.

The BUMP output consists of following major parts

parser trace the input is echoed to the output, allowing the user to trace back the command being executed the moment an error interrupts the program.

data printout observations are printed to the output in order to check if all data has correctly been read.

MLE optimization trace optimization parameters are printed at regular intervals so that divergence of the iteration can be spotted and its cause can be identified. Note that trace values are referring to the unbounded domain.

MLE optimization output the final result of the optimization, the MLE of the parameter set, is printed after transformation to the original variables.

execution trace joint second-moment results are printed after every iteration, for every line in the **exec** input section (§ A.1.6). When integration is performed by the directional sampling method, COV values will be included in the trace.

execution output second-moment statistics are printed at the end of all iterations requested in single execution definition. Repeated for all executions.

marginal/predictive output Marginal and predictive analysis uses the second-moment results obtained after the last iteration. The integration method is determined by the last line in the **exec** section. Expectation and variance of the marginal / predictive variable is followed by a printout of density and distribution: values of both functions are printed in the the user-defined abscissae.

Example printouts of the output file are included in appendix (§ D.).

B. Source listing of USER.FOR

```
c
      double precision function ulhood(mparam,mobsrv,mvarbl,nparam,
     &                                 nobsrv,nvarbl,observ,parval)
c
c=======================================================================

c     USER-DEFINED LIKELIHOOD FUNCTION
c
c     parameters
c       mparam : maximum number of parameters (= 20) <dim>
c       mobsrv : maximum number of observations (= 500) <dim>
c       mvarbl : maximum number of observed variables (= 10) <dim>
c       nparam : number of parameters <in>
c       parval : current value parameter <in>
c       nobsrv : number of observations <in>
c       nvarbl : number of variables per observation <in>
c       observ : observations <in>
c       ulhood : value of likelihood function <out>
c     remark
c       parameters are numbered according to the definition
c       sequence in the input file
c
c-----------------------------------------------------------------------
c
      integer            mparam,mobsrv,mvarbl,nparam,nobsrv,nvarbl
      real               observ(mobsrv,mvarbl)
      double precision   parval(mparam)
c
      ulhood = 1d0
c
      return
      end
c
c
      double precision function uprior(mparam,mpripr,nparam,pripar,
     &                                 parval)
c
c=======================================================================
c
c     USER-DEFINED PRIOR DENSITY
c
c     parameters
c       mparam : maximum number of parameters (= 20) <dim>
c       mpripr : maximum number of parameters for prior (= 4) <dim>
c       nparam : actual number of parameters <in>
c       pripar : parameters of the prior distribution <in>
c       parval : current value of the parameters (original) <in>
c       uprior : value of the prior for parameters = parval <out>
c     remark
c       parameters are numbered according to the definition
c       sequence in the input file
c
```

```
c-----------------------------------------------------------------------
c
      integer           mparam,mpripr,nparam
      real              pripar(mparam,mpripr)
      double precision  parval(mparam)
c
      uprior = 1d0
      return
      end
c
c
      double precision function uwghtf(mparam,mwgtpr,nparam,parval,
     &                                 wgtnum,prdvar,wgtpar)
c
c=======================================================================
c
c     USER-DEFINED WEIGTH FUNCTION (eg. for predictive distribution)
c
c     parameters :
c       mparam : maximum number of parameters to analyse  <dim>
c       mwgtpr : maximum number of parameters for weight function <dim>
c       nparam : number of parameters <in>
c       parval : parameter values in original domain <in>
c       wgtnum : identification of the weightfunction <in>
c       prdvar : value of the predictive parameter <in>
c       wgtpar : parameters of the weight function <in>
c       uwghtf : value of the user defined weight function
c
c-----------------------------------------------------------------------
c
      integer            mparam,mwgtpr,nparam,wgtnum
      double precision   parval(mparam),prdvar
      real               wgtpar(mwgtpr)
c
      if (wgtnum.eq.1) then
        uwghtf = 1d0
      else if (wgtnum.eq.2) then
        uwghtf = 1d0
      else
        uwghtf = 1d0
      endif
      return
      end
```

C. Error messages

Although the program is not "fool-proof", it will capture most important errors occurring. Execution of the program will be stopped on occurrence of an error.

This appendix contains a review of error messages and their probable causes. Errors caused by invalid input are easily traced. When the implementation of the likelihood function is faulty, messages may become confusing. Tracking the error down will be quite difficult. Extreme care should therefore be taken when implementing the user-defined subroutines. If you really cannot find the problem, send an EMail to "jhdaa04@cc4.kuleuven.ac.be" containing problem description, input file and user-defined subroutines. This will help the first author find remaining bugs.

The most common error messages are related to incorrect input. The are detected before the actual calculation is started.

***** Error: invalid input-file [rinput]**

the **bump** file identifier (§ A.1.2) was not found and the program assumes you have submitted an invalid input file

***** Error: invalid data [lparse]**

the line parser captured a **FORTRAN** error that occurred while reading your command line; it probably contains an invalid character, you might also have typed an integer larger than $2^{31} - 1, \ldots$.

***** Error: incomplete input file [rinput]**

check if you have all required sections in your input file.

***** Error: no parameter definition [rparam]**

BUMP thinks it found an empty **parameter** section, check if you complied to the structure of this section (§ A.1.3).

***** Error: incomplete definition [rprior]**

you have forgotten to specify a prior density for an updatable parameter.

***** Error: parameter unknown [rprior]**

the **prior** definition is referring to a parameter that was not previously defined; check the sequence of sections in the input file or the spelling of the parameter name

***** Error: prior of known variable [rprior]**

the **prior** definition for a **known** parameter does not really make sense.

***** Error: impossible value for parameter [rprior]**

you have given an invalid parameter value to a prior density. Check possible parameter values in table 6.11 (§ A.1.4).

***** Error: parameters not defined [robsrv]**

the first line of the data-section does not contain the compulsory dimensioning variables (§ A.1.5).

***** Error: invalid data [robsrv]**

the subroutine reading the experimental data captured an error that occurred during the unformatted, list-directed **FORTRAN** read. An invalid input character is the probable cause.

***** Error: method not specified [rexctl]**

BUMP has not found the name of the integration method you want to use. Check the format of the **exec** section (§ A.1.6)

***** Error: method not implemented [rexctl]**

this might be wishful thinking on your part: **BUMP** only uses sampling or quadrature to solve the Bayesian integration problem! Or perhaps you mistyped the identifier string for the integrations method (§ A.1.6)?

***** Error: incomplete specification [rexctl]**

you have omitted integration parameters. Check the definition of the **exec** section in the user's manual (§ A.1.6)

***** Error: grid too large [rexctl]**

you have selected too may gridpoints for the Gauss-Hermite quadrature. The maximum is currently set at ngrid = 300.

***** Error: parameter unknown [routct]**

you are requesting marginal information from a parameter that was not previously defined. Check to see if you have not misspelled the name of the parameter. You might have requested output before defining the actual parameters, which is a bit confusing to **BUMP**

***** Error: output for known variable [routct]**

asking for marginal statistics from a known variable is rather awkward to **BUMP**; you must probably have confused parameter names.

***** Error: incomplete interval specification [routct]**

there is something wrong with the specification of the interval in which the marginal /predictive density is going to be evaluated. Read through (§ A.1.7) to see what you have done wrong.

***** Error: predictive requires bounds [routct]**

you cannot use the default or relative definition of the interval for a predictive density, use absolute bounds **[{lowerbound},{upperbound}]**.

***** Error: too many steps for density [routct]**

the number of points in which the density is going to be evaluated exceeds the capacity of the program; the maximum is currently set at 500 points.

The following enumeration of messages concerns run-time errors captured by **BUMP**

***** Error: zero likelihood [derivt]**

the finite difference routine is confronted with a zero value of the likelihood ⇒ triple-check your implementation of the function **ulhood**. This error can also occur when the optimization routine has reached the maximum before both convergence criteria are true. A remedy to this unlikely event is the selection of other starting values for the iteration (§ A.1.3)

***** Error: overflow in exponential [qdcore]**

the quadrature formula contains an exponential with positive argument; when the number of gridpoints is to high, this exponent will overflow. The current maximum number of gridpoints is chosen so that overflow will not occur on an IEEE machine. If your computer uses a different double precision implementation (eg. IBM370) this error is likely to occur for large grids. Reduce the number of gridpoints until the problem has disappeared.

***** Error: initial values in likelihood tail [qnewtn]**

your initial values are so far in the likelihood tail that the gradient of the (log-)likelihood is zero; change your initial values. A simple regression procedure might help to make educated guesses about adequate initial values for the optimization.

***** Error: no convergence [gosect]**

the golden section line search has not found an optimum along the direction given by the quadratic approximation. A pragmatic approach suggests to use different initial values for the iteration. If that does not help, than there is probably something wrong with your likelihood function.

***** Error: inaccurate quadrature [smp*]**

the **QUADPACK** adaptive linear integration routine was unable to perform the integration to the preset accuracy. The likelihood is again to blame: either it is wrongly implemented, or it is really ill-behaved.

This final list of messages are originating from errors within some of the library routines **BUMP** uses. These errors should, in theory, never occur. You have most probably stumbled upon a bug. Send me an EMail, and I will try to fix it as soon as I can.

***** Error: internal [dgamma]**

***** Error: illegal parameter [dgemv]**

***** Error: invalid parameter [dgerup]**

***** Error: unable to perform curve fit on density [cumdis]**

D. BUMP output

D.1 Example 1: quadrature results — Berkeley data

```
B U M P (v1.05)
Bayesian Updating of Model Parameters
================================================================================


INPUT

> bump
> parameters
>   thta,update,-1e36,1e36,11.
>   muf,known,8.6
>   sgf,known,0.023
>   errv,update,0e0,1e36,0.005
>
> prior
>   thta=none
>   errv=user
>
> data
>   nobs=10 nvar=1
>
> method
>   quad,3,31
>
> output
>   marginal,thta,100,5
>   marginal,errv,100,3
>   predictive,1,100,14e0,16e0
>
> run


OBSERVATIONS

Data :   10 observations of    1 variables

                1
  1    0.4130E+07
  2    0.3860E+07
  3    0.3920E+07
  4    0.4270E+07
  5    0.4150E+07
  6    0.3660E+07
  7    0.3910E+07
  8    0.3980E+07
  9    0.4150E+07
 10    0.4250E+07


OPTIMIZATION TRACE (unbounded domain)
```

```
iteration 12
optimum
   0.1091D+02    0.8600D+01    0.2300D-01   -0.6209D+01
gradient
  -0.5331D-03    0.0000D+00    0.0000D+00   -0.2805D-04
inverted Hessian
   0.2350D-03    0.0000D+00    0.0000D+00   -0.3795D-03
   0.0000D+00    0.1000D+01    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
  -0.3795D-03    0.0000D+00    0.0000D+00    0.2340D+00
```

INITIAL ESTIMATES

Expectation (SO-approx)

	thta	muf	sgf	errv
E	0.1091E+02	0.8600E+01	0.2300E-01	0.2247E-02

Covariance (FO-approx)

	thta	muf	sgf	errv
thta	0.2350E-03	0.0000E+00	0.0000E+00	-0.7634E-06
muf	0.0000E+00	0.1000E+01	0.0000E+00	0.0000E+00
sgf	0.0000E+00	0.0000E+00	0.1000E+01	0.0000E+00
errv	-0.7634E-06	0.0000E+00	0.0000E+00	0.9469E-06

QUADRATURE: EXECUTION TRACE

```
iteration  1
proportionality factor
   0.3603D+09
expectation
   0.1091D+02    0.8600D+01    0.2300D-01    0.3440D-02
covariance
   0.3570D-03    0.0000D+00    0.0000D+00   -0.4258D-09
   0.0000D+00    0.1000D+01    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
  -0.4258D-09    0.0000D+00    0.0000D+00    0.7764D-05

iteration  2
proportionality factor
   0.3603D+09
expectation
   0.1091D+02    0.8600D+01    0.2300D-01    0.3440D-02
covariance
   0.3572D-03    0.0000D+00    0.0000D+00    0.9060D-16
   0.0000D+00    0.1000D+01    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.9060D-16    0.0000D+00    0.0000D+00    0.6365D-05

iteration  3
proportionality factor
```

```
   0.3603D+09
expectation
   0.1091D+02    0.8600D+01    0.2300D-01    0.3440D-02
covariance
   0.3572D-03    0.0000D+00    0.0000D+00   -0.2138D-18
   0.0000D+00    0.1000D+01    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
  -0.2138D-18    0.0000D+00    0.0000D+00    0.6365D-05


JOINT MOMENTS by GAUSS-HERMITE QUADRATURE
niter =  3, ngrid =  31

Expectation

              thta          muf          sgf          errv
   E    0.1091E+02   0.8600E+01   0.2300E-01   0.3440E-02

Covariance

              thta          muf          sgf          errv
thta    0.3572E-03   0.0000E+00   0.0000E+00  -0.2138E-18
 muf    0.0000E+00   0.1000E+01   0.0000E+00   0.0000E+00
 sgf    0.0000E+00   0.0000E+00   0.1000E+01   0.0000E+00
errv   -0.2138E-18   0.0000E+00   0.0000E+00   0.6365E-05


MARGINAL STATISTICS by GAUSS-HERMITE QUADRATURE
parameter = thta

Expectation =     0.1090E+02
Variance    =     0.3733E-03

          abscissa        density        distrib
  1     0.1081E+02     0.1460E-01     0.0000E+00
  2     0.1082E+02     0.1692E-01     0.3002E-04
  3     0.1082E+02     0.1964E-01     0.6486E-04
  4     0.1082E+02     0.2285E-01     0.1053E-03
  5     0.1082E+02     0.2663E-01     0.1525E-03

...

 95     0.1099E+02     0.3110E-01     0.9994E+00
 96     0.1099E+02     0.2664E-01     0.9994E+00
 97     0.1100E+02     0.2286E-01     0.9995E+00
 98     0.1100E+02     0.1965E-01     0.9995E+00
 99     0.1100E+02     0.1692E-01     0.9995E+00
100     0.1100E+02     0.1460E-01     0.9996E+00


MARGINAL STATISTICS by GAUSS-HERMITE QUADRATURE
parameter = errv

Expectation =     0.3421E-02
```

Variance = 0.5594E-05

	abscissa	density	distrib
1	0.3826E-03	0.5524E-01	0.0000E+00
2	0.4000E-03	0.9237E-01	0.1256E-05
3	0.4182E-03	0.1519E+00	0.3439E-05
4	0.4372E-03	0.2455E+00	0.7155E-05
5	0.4570E-03	0.3901E+00	0.1338E-04
...			
95	0.2495E-01	0.1421E+00	0.9990E+00
96	0.2608E-01	0.1158E+00	0.9992E+00
97	0.2726E-01	0.9428E-01	0.9993E+00
98	0.2850E-01	0.7667E-01	0.9994E+00
99	0.2980E-01	0.6229E-01	0.9995E+00
100	0.3115E-01	0.5056E-01	0.9996E+00

PREDICTIVE STATISTICS by GAUSS-HERMITE QUADRATURE

Expectation = 0.1521E+02
Variance = 0.3926E-02

	abscissa	density	distrib
1	0.1400E+02	0.2751E-07	0.0000E+00
...			
50	0.1499E+02	0.4967E-01	0.1918E-02
51	0.1501E+02	0.8760E-01	0.3267E-02
52	0.1503E+02	0.1566E+00	0.5665E-02
53	0.1505E+02	0.2826E+00	0.9977E-02
54	0.1507E+02	0.5105E+00	0.1777E-01
55	0.1509E+02	0.9127E+00	0.3178E-01
56	0.1511E+02	0.1588E+01	0.5650E-01
57	0.1513E+02	0.2632E+01	0.9848E-01
58	0.1515E+02	0.4046E+01	0.1654E+00
59	0.1517E+02	0.5599E+01	0.2630E+00
60	0.1519E+02	0.6793E+01	0.3892E+00
61	0.1521E+02	0.7097E+01	0.5313E+00
62	0.1523E+02	0.6355E+01	0.6687E+00
63	0.1525E+02	0.4932E+01	0.7834E+00
64	0.1527E+02	0.3392E+01	0.8672E+00
65	0.1529E+02	0.2128E+01	0.9224E+00
66	0.1531E+02	0.1253E+01	0.9559E+00
67	0.1533E+02	0.7102E+00	0.9753E+00
68	0.1535E+02	0.3949E+00	0.9862E+00
69	0.1537E+02	0.2186E+00	0.9922E+00
70	0.1539E+02	0.1216E+00	0.9955E+00
71	0.1541E+02	0.6846E-01	0.9974E+00
72	0.1543E+02	0.3916E-01	0.9984E+00
73	0.1545E+02	0.2280E-01	0.9991E+00
74	0.1547E+02	0.1353E-01	0.9994E+00

```
 75      0.1549E+02      0.8191E-02      0.9996E+00
 76      0.1552E+02      0.5053E-02      0.9998E+00
 77      0.1554E+02      0.3175E-02      0.9998E+00
 78      0.1556E+02      0.2032E-02      0.9999E+00
 79      0.1558E+02      0.1322E-02      0.9999E+00
 80      0.1560E+02      0.8746E-03      0.1000E+01
```

D.2 Example 1: sampling results — Berkeley data

```
B U M P (v1.05)
Bayesian Updating of Model Parameters
================================================================================


INPUT

> bump
> parameters
>   thta,update,-1e36,1e36,11.
>   muf,known,8.6
>   sgf,known,0.023
>   errv,update,0e0,1e36,0.005
>
> prior
>   thta=none
>   errv=user
>
> data
>   nobs=10 nvar=1
>
> method
>   samp,12076492,1,20000
>
> output
>
> run

...

DIRECTIONAL SAMPLING: EXECUTION TRACE

iteration  1
proportionality factor
 0.36E+09 (.00)
expectation
 0.11E+02 (.00)  0.86E+01 (.00)  0.23E-01 (.00)  0.34E-02 (.00)
covariance
 0.35E-03 (.00)  0.00E+00 (.00)  0.00E+00 (.00)  0.63E-07 (***)
 0.00E+00 (.00)  0.10E+01 (.00)  0.00E+00 (.00)  0.00E+00 (.00)
 0.00E+00 (.00)  0.00E+00 (.00)  0.10E+01 (.00)  0.00E+00 (.00)
 0.63E-07 (***)  0.00E+00 (.00)  0.00E+00 (.00)  0.77E-05 (.01)
```

```
JOINT MOMENTS by DIRECTIONAL SAMPLING
seed =   12076492, niter =  1, nsamp =  20000

Expectation

              thta          muf          sgf          errv
   E    0.1090E+02   0.8600E+01   0.2300E-01   0.3437E-02

Covariance

              thta          muf          sgf          errv
thta    0.3549E-03   0.0000E+00   0.0000E+00   0.6265E-07
 muf    0.0000E+00   0.1000E+01   0.0000E+00   0.0000E+00
 sgf    0.0000E+00   0.0000E+00   0.1000E+01   0.0000E+00
errv    0.6265E-07   0.0000E+00   0.0000E+00   0.7710E-05
```

D.3 Example 2: quadrature results — Berkeley data

```
B U M P (v1.05)
Bayesian Updating of Model Parameters
================================================================================


INPUT

> bump
> parameters
>   aa,update,0e0,1e36,4.0
>   bb,update,0e0,1e36,0.9
>   errv,update,0e0,1e36,0.01
>   muf,known,8.6
>   sgf,known,0.023
>
> prior
>   aa=none
>   bb=none
>   errv=none
>
> data
>   nobs=50  nvar=2
>
> method
>   quad,3,11
>
> output
>   predictive,1,50,2200e0,3200e0,3
>   predictive,1,50,3000e0,4500e0,7
>   predictive,1,50,4000e0,5500e0,14
>   predictive,1,50,4500e0,6000e0,21

...

>   predictive,1,50,5500e0,7000e0,240
```

```
>   predictive,1,50,5500e0,7000e0,270
>   predictive,1,50,5500e0,7000e0,300
>   predictive,1,50,5500e0,7000e0,360
>
> run
```

OBSERVATIONS

Data : 50 observations of 2 variables

	1	2
1	0.3000E+01	0.2680E+04
2	0.3000E+01	0.2680E+04
3	0.3000E+01	0.2680E+04
4	0.7000E+01	0.3960E+04
5	0.7000E+01	0.3970E+04
6	0.7000E+01	0.3800E+04
7	0.7000E+01	0.3910E+04
8	0.7000E+01	0.3870E+04
9	0.2100E+02	0.5120E+04
10	0.2100E+02	0.4960E+04
11	0.2100E+02	0.5030E+04
12	0.2800E+02	0.5530E+04
13	0.2800E+02	0.5250E+04
14	0.2800E+02	0.5410E+04
15	0.2800E+02	0.5430E+04
16	0.2800E+02	0.5500E+04
17	0.2800E+02	0.5460E+04
18	0.2800E+02	0.5070E+04
19	0.2800E+02	0.5390E+04
20	0.2800E+02	0.5470E+04
21	0.2800E+02	0.5350E+04
22	0.2800E+02	0.5360E+04
23	0.2800E+02	0.5500E+04
24	0.2800E+02	0.5460E+04
25	0.2800E+02	0.5450E+04
26	0.2800E+02	0.5480E+04
27	0.2800E+02	0.5490E+04
28	0.2800E+02	0.5410E+04
29	0.2800E+02	0.5620E+04
30	0.2800E+02	0.5650E+04
31	0.2800E+02	0.5350E+04
32	0.6000E+02	0.5730E+04
33	0.6000E+02	0.5580E+04
34	0.6000E+02	0.5170E+04
35	0.9000E+02	0.6030E+04
36	0.9000E+02	0.6110E+04
37	0.9000E+02	0.5760E+04
38	0.9000E+02	0.6095E+04
39	0.9000E+02	0.6490E+04
40	0.1200E+03	0.6135E+04
41	0.1200E+03	0.6300E+04
42	0.1200E+03	0.6045E+04
43	0.1200E+03	0.6215E+04

```
 44     0.1200E+03     0.6055E+04
 45     0.1800E+03     0.5930E+04
 46     0.1800E+03     0.6100E+04
 47     0.1800E+03     0.6250E+04
 48     0.2100E+03     0.6400E+04
 49     0.2100E+03     0.6250E+04
 50     0.2100E+03     0.6310E+04

OPTIMIZATION TRACE (unbounded domain)

iteration 15
optimum
   0.1279D+01   -0.1388D+00   -0.7592D+01    0.8600D+01    0.2300D-01
gradient
  -0.2104D-01   -0.4508D-01    0.1050D-02    0.0000D+00    0.0000D+00
inverted Hessian
   0.7541D-03   -0.1109D-03   -0.7554D-04    0.0000D+00    0.0000D+00
  -0.1109D-03    0.4286D-04   -0.1084D-03    0.0000D+00    0.0000D+00
  -0.7554D-04   -0.1084D-03    0.1617D+00    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01

iteration 26
optimum
   0.1279D+01   -0.1388D+00   -0.7592D+01    0.8600D+01    0.2300D-01
gradient
   0.3956D-06    0.2129D-05   -0.3650D-07    0.0000D+00    0.0000D+00
inverted Hessian
   0.7431D-03   -0.1040D-03    0.7018D-06    0.0000D+00    0.0000D+00
  -0.1040D-03    0.4218D-04    0.3885D-06    0.0000D+00    0.0000D+00
   0.7018D-06    0.3885D-06    0.1681D+00    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01

INITIAL ESTIMATES

Expectation (SO-approx)

                aa             bb            errv            muf            sgf
   E    0.3594E+01     0.8704E+00    0.5467E-03     0.8600E+01     0.2300E-01

Covariance (FO-approx)

                aa             bb            errv            muf            sgf
  aa    0.9592E-02    -0.3252E-03    0.1272E-08     0.0000E+00     0.0000E+00
  bb   -0.3252E-03     0.3196E-04    0.1705E-09     0.0000E+00     0.0000E+00
errv    0.1272E-08     0.1705E-09    0.4275E-07     0.0000E+00     0.0000E+00
 muf    0.0000E+00     0.0000E+00    0.0000E+00     0.1000E+01     0.0000E+00
 sgf    0.0000E+00     0.0000E+00    0.0000E+00     0.0000E+00     0.1000E+01

QUADRATURE: EXECUTION TRACE
```

```
iteration  1
proportionality factor
   0.9998D-02
expectation
   0.3595D+01    0.8704D+00    0.6453D-03    0.8600D+01    0.2300D-01
covariance
   0.1091D-01   -0.3697D-03    0.1144D-06    0.0000D+00    0.0000D+00
  -0.3697D-03    0.3632D-04   -0.9615D-09    0.0000D+00    0.0000D+00
   0.1144D-06   -0.9615D-09    0.7534D-07    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01

iteration  2
proportionality factor
   0.9998D-02
expectation
   0.3595D+01    0.8704D+00    0.6453D-03    0.8600D+01    0.2300D-01
covariance
   0.1091D-01   -0.3697D-03    0.8899D-07    0.0000D+00    0.0000D+00
  -0.3697D-03    0.3632D-04    0.3072D-09    0.0000D+00    0.0000D+00
   0.8899D-07    0.3072D-09    0.6562D-07    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01

iteration  3
proportionality factor
   0.9998D-02
expectation
   0.3595D+01    0.8704D+00    0.6454D-03    0.8600D+01    0.2300D-01
covariance
   0.1091D-01   -0.3697D-03    0.8897D-07    0.0000D+00    0.0000D+00
  -0.3697D-03    0.3632D-04    0.3073D-09    0.0000D+00    0.0000D+00
   0.8897D-07    0.3073D-09    0.6560D-07    0.0000D+00    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01    0.0000D+00
   0.0000D+00    0.0000D+00    0.0000D+00    0.0000D+00    0.1000D+01


JOINT MOMENTS by GAUSS-HERMITE QUADRATURE
niter =  3, ngrid =  11

Expectation

                aa            bb            errv          muf           sgf
   E    0.3595E+01    0.8704E+00    0.6454E-03    0.8600E+01    0.2300E-01

Covariance

                aa            bb            errv          muf           sgf
  aa    0.1091E-01   -0.3697E-03    0.8897E-07    0.0000E+00    0.0000E+00
  bb   -0.3697E-03    0.3632E-04    0.3073E-09    0.0000E+00    0.0000E+00
errv    0.8897E-07    0.3073E-09    0.6560E-07    0.0000E+00    0.0000E+00
 muf    0.0000E+00    0.0000E+00    0.0000E+00    0.1000E+01    0.0000E+00
 sgf    0.0000E+00    0.0000E+00    0.0000E+00    0.0000E+00    0.1000E+01
```

```
PREDICTIVE STATISTICS by GAUSS-HERMITE QUADRATURE

...
```

D.4 Example 2: sampling results — Berkeley data

```
B U M P (v1.05)
Bayesian Updating of Model Parameters
=====================================================================


INPUT

> bump
> parameters
>   aa,update,0e0,1e36,4.0
>   bb,update,0e0,1e36,0.85
>   errv,update,0e0,1e36,0.01
>   muf,known,8.6
>   sgf,known,0.023
>
> prior
>   aa=none
>   bb=none
>   errv=none
>
> data
>   nobs=50  nvar=2
>
> method
>   samp,1695678,1,1000
>
> output
>
> run

...

DIRECTIONAL SAMPLING: EXECUTION TRACE

iteration  1
proportionality factor
 0.10E-01 (.00)
expectation
 0.36E+01 (.00)  0.87E+00 (.00)  0.65E-03 (.00)  0.86E+01 (.00)  0.23E-01 (.00)
covariance
 0.11E-01 (.03) -0.36E-03 (.05) -0.30E-06 (***)  0.00E+00 (.00)  0.00E+00 (.00)
-0.36E-03 (.05)  0.36E-04 (.03)  0.15E-07 (***)  0.00E+00 (.00)  0.00E+00 (.00)
-0.30E-06 (***)  0.15E-07 (***)  0.77E-07 (.02)  0.00E+00 (.00)  0.00E+00 (.00)
 0.00E+00 (.00)  0.00E+00 (.00)  0.00E+00 (.00)  0.10E+01 (.00)  0.00E+00 (.00)
 0.00E+00 (.00)  0.00E+00 (.00)  0.00E+00 (.00)  0.00E+00 (.00)  0.10E+01 (.00)
```

JOINT MOMENTS by DIRECTIONAL SAMPLING
seed = 1695678, niter = 1, nsamp = 1000

Expectation

	aa	bb	errv	muf	sgf
E	0.3591E+01	0.8714E+00	0.6463E-03	0.8600E+01	0.2300E-01

Covariance

	aa	bb	errv	muf	sgf
aa	0.1066E-01	-0.3600E-03	-0.3045E-06	0.0000E+00	0.0000E+00
bb	-0.3600E-03	0.3581E-04	0.1535E-07	0.0000E+00	0.0000E+00
errv	-0.3045E-06	0.1535E-07	0.7746E-07	0.0000E+00	0.0000E+00
muf	0.0000E+00	0.0000E+00	0.0000E+00	0.1000E+01	0.0000E+00
sgf	0.0000E+00	0.0000E+00	0.0000E+00	0.0000E+00	0.1000E+01

Bibliography

[1] M. Abramowitz and I.A. Stegun (1972). *Handbook of Mathematical Functions.* Dover Publications, New York, USA.

[2] American Concrete Institute (1992). *ACI Manual of Concrete Practice: Part 1 - Part 5.* ACI Publications, Detroit, Michigan, USA.

[3] A-S. Ang and W.H. Tang (1975). *Probability Concepts in Engineering Planning and Design. Volume I: Basic Principles.* John Wiley & Sons, New York, USA.

[4] A-S. Ang and W.H. Tang (1984). *Probability Concepts in Engineering Planning and Design. Volume II: Decision, Risk and Reliability.* John Wiley & Sons, New York, USA.

[5] Y. Bard (1974). *Nonlinear Parameter Estimation.* Academic Press, Orlando, Florida.

[6] Z.P. Bažant and J-C. Chern (1984). "Bayesian Statistical Prediction of Concrete Creep and Shrinkage". *ACI Journal*, 81(4):319–330.

[7] P. Bjerager (1988). "Probability Integration by Directional Simulation". *Journal of Engineering Mechanics, ASCE*, 114(8):1285–1302.

[8] G.E.P. Box and G.C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Addison–Wesley, Reading, Massachusetts, USA.

[9] W. Cheney and D. Kincaid (1980). *Numerical Mathematics and Computing.* Brooks and Cole, Monterey, California, USA.

[10] C.A. Cornell (1972). "Bayesian Statistical Decision Theory and Reliability-Based Design". In A.M. Freudenthal, editor, *Proceedings of the International Conference on Structural Safety and Reliability*, pages 47–68, Oxford, UK. Pergamon Press. Conference held at the Smithsonian Institution, Washington DC., 1969.

[11] A. Der Kiureghian (1989). "Measures of Structural Safety under Imperfect States of Knowledge". *Journal of Structural Engineering, ASCE*, 115(5):1119–1140.

[12] A. Der Kiureghian (1990). "Bayesian Analysis of Model Uncertainty in Structural Reliability". In *Proceedings of the 3rd IFIP WG7.5 Working Conference on Reliability and Optimization of Structural Systems, Berkeley, California*, pages 211–221, Berlin, Germany. Springer–Verlag.

[13] A. Der Kiureghian (1991). "Prediction Model for Pull-Out Capacity of Steel Plate Embedded in Concrete Slab". Draft report, Department of Civil Engineering, University of California at Berkeley, Berkeley, California, USA.

[14] A. Der Kiureghian and P-L. Liu (1986). "Structural Reliability under Incomplete Probability Information". *Journal of the Engineering Mechanics Division, ASCE*, 112(1):85–104.

[15] O. Ditlevsen, P. Bjerager, R. Oleson, and A.M. Hasofer (1988). "Directional Simulation in Gaussian Processes". *Probabilistic Engineering Mechanics*, 3(4):207–217.

[16] A.E. Gelfand, S.E. Hills, A. Racine-Poon, and A.F.M. Smith (1990). "Illustration of Bayesian Inference in Normal Data Models using Gibbs Sampling". *Journal of the American Statistical Association*, 85(412):972–985.

[17] A.E. Gelfand and A.F.M. Smith (1990). "Sampling-Based Approaches to Calculating Marginal Densities". *Journal of the American Statistical Association*, 85(410):398–409.

[18] J. Geweke (1989). "Bayesian Inference in Econometric Models using Monte Carlo Integration". *Econometrica*, 57(6):1317–1339.

[19] Ph. Geyskens and A. Der Kiureghian (1992). "TURNS – Testing of Pseudo-Random Number Sequences". Internal Report.

[20] Ph. Geyskens, A. Der Kiureghian, and G. De Roeck (1991). "SORM Analysis using Quasi-Newton Optimization". In P.D. Spanos and C.A. Brebbia, editors, *Computational Stochastic Mechanics*, pages 89–100, Barking, Essex, UK. Elsevier Science Publishers.

[21] P.E. Gill, W. Murray, and M.H. Wright (1981). *Practical Optimization.* Academic Press, New York, USA.

[22] J.A. Hanson (1958). "Shear Strength of Lightweight Reinforced Concrete Beams". *ACI Journal*, 30(3):387–403.

[23] M. KENDALL AND A. STUART (1977). *The Advanced Theory of Statistics. Volume 1: Distribution Theory.* Charles Griffin & Company Ltd., London, UK.

[24] P. KLIEGER (1957). "Long-Time Study of Cement Performance in Concrete. Chapter 10 – Progress Report on Strength and Elastic Properties of Concrete". *ACI Journal*, 29(6):481–504.

[25] T. KLOEK AND H.K. VAN DIJCK (1978). "Bayesian Estimates of Equation System Parameters: an Application of Integration by Monte Carlo". *Econometrica*, 46(1):1–19.

[26] P.M. LEE (1989). *Bayesian Statistics: an Introduction.* Oxford University Press, New York, USA.

[27] D.V. LINDLEY (1980). "Approximate Bayesian Methods". In J.M. Bernardo, M.H. De Groot, D.V. Lindley, and A.F.M. Smith, editors, *Bayesian Statistics*, pages 223–245, Valencia, Spain. Valencia Press.

[28] D.G. LUENBERGER (1984). *Linear and Nonlinear Programming.* Addison-Wesley, Reading, Massachusetts, USA, second edition.

[29] H.O. MADSEN, S. KRENK, AND N.C. LIND (1986). *Methods of Structural Safety.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

[30] M.A. MAES (1991). "Codifications of Design Load Criteria subject to Modeling Uncertainty". *Journal of Structural Engineering, ASCE*, 117(10):2988–3007.

[31] R.E. MELCHERS (1987). *Structural Reliability: Analysis and Prediction.* Ellis Horwood, Chichester, UK.

[32] R.E. MELCHERS (1990). "Radial Importance Sampling for Structural Reliability". *Journal of Engineering Mechanics, ASCE*, 116(1):189–203.

[33] P.J.M. MONTEIRO, J.P. MOEHLE, AND H.T. TANG (1991). "Aging Effects on Mechanical Properties of Power Plant Concretes". Technical report, Department of Civil Engineering, University of California at Berkeley, Berkeley, California, USA.

[34] J.C. NAYLOR AND A.F.M. SMITH (1982). "Applications of a Method for the Efficient Computation of Posterior Distributions". *Applied Statistics*, 31(3):214–225.

[35] A. PAUW (1960). "Static Modulus of Elasticity of Concrete as Affected by Density". *ACI Journal*, 32(6):679–687.

[36] R. PIESSENS, E. DE DONCKER, C.W. ÜBERHUBER, AND D.K. KAHANER (1983). *Quadpack - a Subroutine Package for Automatic Integration.* Springer–Verlag, Berlin, Germany.

[37] S.J. PRESS (1989). *Bayesian Statistics, Principles, Models, and Applications.* John Wiley & Sons, New York.

[38] W.H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY, AND W.T. VETTERLING (1986). *Numerical Recipes.* Cambridge University Press, Cambridge, Massachusetts, USA.

[39] G.G. ROUSSAS (1973). *A First Course in Mathematical Statistics.* Addison-Wesley, Reading, Massachusetts, USA.

[40] R.Y. RUBINSTEIN (1981). *Simulation and the Monte-Carlo method.* John Wiley & Sons, New York, USA.

[41] J.J. SHIDELER (1957). "Lightweight-Aggregate Concrete for Structural Use". *ACI Journal*, 29(4):299–328.

[42] A.F.M. SMITH, A.M. SKENE, J.E.H. SHAW, AND J.C. NAYLOR (1987). "Progress with Numerical and Graphical Methods for Practical Bayesian Statistics". *The Statistician*, 36:75–82.

[43] L. TIERNEY AND J.B. KADANE (1986). "Accurate Approximations for Posterior Moments and Marginal Densities". *Journal of the American Statistical Association*, 81(393):82–86.

[44] L. TIERNEY, R.E. KASS, AND J.B. KADANE (1989). "Fully Exponential Laplace Approximations to Expectations and Variances of Nonpositive Functions". *Journal of the American Statistical Association*, 84(407):710–716.

[45] H.K. VAN DIJCK AND T. KLOEK (1980). "Further Experience in Bayesian Analysis using Monte Carlo Integration". *Journal of Econometrics*, 14:307–328.

[46] H.K. VAN DIJCK, T. KLOEK, AND C.G. BOENDER (1985). "Posterior Moments Computed by Mixed Integration". *Journal of Econometrics*, 29:3–18.