Istituto Universitario
di Studi Superiori di Pavia

Università degli Studi
di Pavia

**EUROPEAN SCHOOL OF ADVANCED STUDIES IN
REDUCTION OF SEISMIC RISK**

ROSE SCHOOL

FRAGILITY ANALYIS OF REINFORCED CONCRETE STRUCTURES USING A RESPONSE SURFACE APPROACH

**A Dissertation Submitted in Partial
Fulfilment of the Requirements for the Master Degree in**

EARTHQUAKE ENGINEERING

MARKO IJKE SCHOTANUS

Supervisor: Prof. PAOLO E. PINTO

July, 2002

The dissertation entitled “Fragility Analysis of Reinforced Concrete Structures Using a Response Surface Approach”, by Marko IJke Schotanus, has been approved in partial fulfilment of the requirements for the Master Degree in Earthquake Engineering.

Paolo E. Pinto _____

Gian Michele Calvi _____

Acknowledgements

This report has been written in partial fulfillment of the requirements for completing the Master programme offered by the European School for Advanced Studies in Reduction of Seismic Risk (ROSE) in Pavia, Italy .

The work that will be presented here was carried out under the supervision of Prof. P.E. Pinto at the Department of Structural and Geotechnical Engineering of the first University of Rome, "La Sapienza". I have benefited greatly from the people I have had the opportunity of working with in preparing this report, particularly Paolo Pinto, Paolo Franchin and Alessio Lupoi. Their assistance, advice and encouragement has taken many forms, and I am grateful to each of them.

Roma, luglio 2002, Marko Schotanus

Abstract

A general and sophisticated method for seismic fragility analysis of systems, originally proposed by Veneziano et.al. (1983, [15]), is further developed in this work and applied to a reinforced concrete frame. The method is not limited to a specific class of structures or systems, or to a particular representation of the random input. Sophistication lies in the incorporation of state-of-the-art mechanical models; use of realistic models for the seismic action; assignment of probabilistic models for mechanical parameters and capacity; and in accounting for multiple failure modes and their correlation.

The basic proposal is to use a response surface to represent the capacity part in an analytical limit state function (g-function) as input for SORM analysis. Response surface techniques are used to replace the algorithmic g-function, or the capacity part of it, with an explicit functional relationship, fitting a second order polynomial. Such an explicit format highly reduces the number of expensive numerical analyses needed compared to classical methods that determine the failure domain. In order to keep the method competitive, only a small number (with a maximum of about 6) of parameters are chosen to enter in the function as *explicit* variables, whose effect is denominated *fixed*. The response surface methodology is then used as an iterative and evolutive procedure, to help distinguishing the important variables from the less important and to support the choices made. Those random variables that are not explicitly incorporated are accumulated in an error term together with errors originating from the lack-of-fit of the model. This error term is itself random and transfers the uncertainty to the output quantity. More specifically, the effect of a large number of *implicit* variables can be grouped together in a few additive *random* variables (random effects), improving the model's descriptive power at a moderate additional cost. The effect of earthquake loading on the response and the spatial fluctuation of the mechanical parameters are treated in this way.

The unknown model parameters (fixed effect coefficients and distribution of random effects) are then statistically estimated from numerical analyses, carefully planned in the *experimental design*. After the model is fitted, a SORM procedure is used to obtain the fragility curve for the system. Further statistical analysis gives insight in the accuracy of the results obtained.

The developed theory is then applied to a three bay, six storey reinforced concrete frame, representing a system that is considered sufficiently complex to serve as a challenging test case. First only one random effect variable is introduced, describing dependence of

the response on the earthquake, then a second one, introducing spatial variability of the concrete ultimate strain. In depth investigations are carried out to test the stability and accuracy of the suggested procedure, consisting of: treating the model coefficients both as fixed and as random parameters in the limit state function; testing the contribution to the accuracy of the model of the basic input variables chosen; observing the sensitivity of the results to the location of the centre of the experimental design; determining the effect of the choice of different earthquake records; comparing some of the results with *exact* results found by Monte Carlo simulation.

The results of the computations, allow a number of general observations to be made that are summarised in a final discussion.

Contents

Chapter 1

Introduction

If both resistance R and load effect (demand) D of a structure are subject to variation with time, it is no longer straight-forward to anticipate whether a structure is safe or not. At a certain time instant t safety can be simply expressed as $S_t = R_t - D_t$, negative values implying an un-safe state, but all combinations occurring during the lifetime of a structure can hardly be expected to be known *a priori*. However, available information can indicate a range of realistically possible values for both R and D, ideally paired with a rate of their occurrence. This kind of information can be formally described in terms of *probability distributions* for R and D.

With the probability distributions of R and D, the probability of S can be computed for all possible combinations, but requiring $\Pr\{S > 0\}$ for all of them would lead to excessively expensive structures. It is therefore common to accept a limited number of unsafe combinations that might possibly occur during the design life of the structure. The sum of those failure states, weighted by their probability content usually goes under the name of *probability of failure* and can be seen as a measure of safety (its complement).

Treating the above outlined problem in a mathematically more rigourous way, regarding R and D as random quantities described by their probability density functions, denoted by $f_R()$ and $f_D()$ respectively, the probability of failure (p_f) of the system can be written in the following way (theorem of total probability)

$$p_f = \sum_x P[R = x \bigcap D \geq x] = \sum_x P[R = x]P[D \geq x] \tag{1.1}$$

if R and D are discrete variables with a finite set of values x. It has to be noted that the expression given in the right-hand side of Equation (1.1) only holds for the case when R and D are independent, which is an assumption generally made for loading and resistance terms. When x is continuous, having introduced the density functions, p_f can also be written as

$$p_f = \int_x f_R(x)[1 - F_D(x)]dx \tag{1.2}$$

This integral is also known as the *convolution integral*.

If the case of seismic loading, failure is traditionally expressed conditional on a ground-motion intensity measure Y. The total probability of failure, considering all possible intensities and their associated probabilities, can be obtained through a summation over all possible intensities Y:

$$p_f = \sum_Y P[D \geq R|Y]P[Y = y] \tag{1.3}$$

if Y takes on discrete values only or, when Y is continuous,

$$p_f = \int_Y P_f[Y = y] \cdot f_Y(y)\, dy = \int_Y \int_X [1 - F_D(x|y)] \cdot f_R(x|y) \cdot f_Y(y)\, dxdy \tag{1.4}$$

where $P_f[\dots]$ stands for *conditional probability of failure*. Still in an earthquake engineering (EE) context, the probability that the selected ground-motion intensity measure Y *exceeds* any given value y at a site is typically defined as the hazard function $H_Y(y) = 1 - F_Y(y)$, or simply *hazard*, and the density function f_Y is equal to the absolute value of its derivative. Taking this into account

$$P_f = \int_Y \int_X [1 - F_D(x|y)] \cdot f_R(x|y) \cdot \left| \frac{dH(y)}{dy} \right| dxdy \tag{1.5}$$

In Equation (1.4) the conditional probability of failure, the probability of the response exceeding a certain level of R as function of Y, is called the *fragility*. There are as many fragility functions as the response (damage) measures of interest.

1.1 Reliability problems

The purpose of the method presented in this report is to calculate fragility functions for systems, or, in other words, computing the conditional failure probability in (1.4), which can also be written as

$$P_f[Y = y] = \int_F f_{\mathbf{x}}(\mathbf{x})d\mathbf{x} \tag{1.6}$$

where $f_{\mathbf{x}}(\mathbf{x})$ is the joint probability density function (JPDF) of vector $\mathbf{x}$, and F the failure domain collecting all the values of $\mathbf{x}$ for which failure occurs. Vector $\mathbf{x}$ generally includes loading parameters, geometric data, material properties, etc.; all of uncertain nature. Using here the basic variables $\mathbf{x}$ and their joint distribution is simply a generalisation of the demand/capacity comparison used sofar, and both D and R are usually functions of $\mathbf{x}$.

Even though the formulation of Equation (1.6) makes the problem look harmless, computation of P_f results usually in considerable difficulties. Among the major ones is the determination of the joint probability model of the basic variables, which generally cannot be established reliably from available data but has to be chosen heuristically, based at least partially on qualitative physical arguments. The choice of this model is particularly critical when the value of P_f is very small in which case the result is sensitive to the distribution tails. Information on these latter is very difficult to collect, since they correspond to extremely rare events.

In the last twenty years this problem, evaluating (1.6), has attracted increasing attention in EE and much research has been devoted to the development of (more) efficient methods. Notwithstanding the progress, consolidated and affordable procedures are not yet available. The main operational complications depend on (1) the very large number of random variables, which becomes a practical problem in face of the complicated algorithmic definition of system response and (2) the fact that the ground-motion is modelled as a random process. The latter feature suggests using random vibration techniques, but the existing literature in this area is also of little help. The main obstacles in this case are the type of nonlinearity of the system, and the uncertain evolution in time of structural properties (e.g. stiffnesses) that depend on the deformation history of each member. A review of several recent and on-going studies which aim at setting up practical but approximate solutions is contained in Pinto (2001, [10]).

A trustworthy procedure should be comprehensive and affordable. Comprehensiveness implies that it should include and be able to deal with: realistic models for the seismic action (recorded accelerograms or random processes characterised by an evolutionary power spectral density); state-of-the-art mechanical models for the structure; randomness in the mechanical parameters of the structure; randomness in the capacity of the failure modes of the structure; multiple failure modes and their correlations. To be affordable implies: a formulation which involves mathematical notions available to an expert engineer but not necessarily a specialist; a number of non-linear analyses for computing the risk significantly smaller than those required by the classical Montecarlo simulation method.

The approach proposed in this study is an extension of response surface methods. Similar methodology is applied frequently for static cases, see for example Ellingwood (1993, [12]), but is not well explored for realistic dynamic problems. The basic idea is to use techniques from experimental design, analysis of variance, classical statistical inference and Level-2 reliability analysis in order to: (1) reduce the dimension of the random-variable space in which one needs to operate, (2) fit a simple regression relationship between the value of the spectral acceleration at failure $S_{a;f}$ (response) and the basic random variables in the

reduced space (regressors), and (3) replace the complicated finite-element algorithm with this simple relationship and use Level-2 reliability procedures for fragility calculation.

The main purpose of the present work is to investigate the potentiality and to develop further this methodology explored before only by Veneziano et al. (1983, [15]). The approach possesses most of the requisites listed above, and seems therefore possibly powerful. Two aspects of the method are given particular attention: (1) the choice of the more appropriate statistical solution tools, given their known delicate numerical stability, and (2) the implicit way multiple modes of failure are dealt with, a way that is difficult to justify in the formal framework of structural reliability theory, which makes the assessment of its validity more difficult.

Chapter 2

Response Surface Methodology

Availability of a limit state function, or simply g-function, in analytical form simplifies the application of FORM or SORM in approximating the probability of failure. The limit state function individuates with its values the safe domain and the failure domain for a system, and often assumes a *capacity minus demand* format. It is well known that in the area of structural analysis the response is expressed in algorithmic form, preventing a direct analytical expression for the g-function. However, using the observed response from a number of analyses, it is possible to construct an empirical relationship between a number of input variables $\mathbf{x}$ and the response quantity of interest Y through application of Response Surface techniques, thus providing the analytical input for the reliability analysis.

In this chapter the necessary framework for building an empirical limit state function using the Response Surface Methodology (RSM) will be developed. First the use of standard RSM for time-invariant reliability problems will be explained, then the theory will be expanded for the application in the domain of earthquake engineering, building a response surface for response under earthquake loading, to be used in a time-variant reliability problem.

2.1 Standard response surface methodology for time invariant problems

Let Y be a *measurable* random quantity. Let $\mathbf{x}$ be a vector of length k of measurable quantities upon which the *mean* of Y is believed to depend. The random quantity Y may be the value of the limit state function $g(\mathbf{x})$ or of a capacity $R(\mathbf{x})$ or demand $D(\mathbf{x})$ term appearing in g while the quantities $\mathbf{x}$ can be the basic variables describing random loading or material properties in the problem. Nonetheless it is important to point out at this stage that in the present context the variables $\mathbf{x}$ are regarded as measurable and *controllable*, meaning that they do not take on values at random, according to a joint PDF

$f(\mathbf{x})$, but can be assigned values to. Instead, Y is measurable but *not* controllable, in the sense that it can be predicted only to a certain degree based on the assigned values of $\mathbf{x}$, since it exhibits random fluctuations around this prediction. An *experiment* can now be defined as an observation (or measurement) of Y as *response* to some *input* values of $\mathbf{x}$. One then wishes to fit to the results of a set of n experiments an explicit functional relationship to relate the mean μ_Y of Y to $\mathbf{x}$. For instance, a simple complete quadratic model of the mean reads as follows

$$\mu_Y(\mathbf{x}) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{j \neq i}^{k} \beta_{ij} x_i x_j \tag{2.1}$$

Such a model for the mean of a random variable Y is a *statistical model*. A statistical model of this type, where dependence on the vector $\mathbf{x}$ is *explicit* and functional through a number of, yet unknown, regression coefficients $\boldsymbol{\beta}$, is called a regression model and in particular a *fixed effect* one. The model as defined in (2.1) is non linear in the *input* or *basic variables* $\mathbf{x}$, but it is still linear in the p *regression parameters* $\boldsymbol{\beta}$ and can thus be rewritten in the linear form

$$\mu_Y(\mathbf{x}) = \mathbf{z}(\mathbf{x})\boldsymbol{\beta} \tag{2.2}$$

where the row vector $\mathbf{z}(\mathbf{x})$ collects the so-called p *explanatory functions* (as many as the regression parameters), generally non-linear functions of $\mathbf{x}$ such as those appearing in expression (2.1).

Once the model for the mean is established the response variable takes the form

$$Y(\mathbf{x}) = \mu_Y(\mathbf{x}) + \varepsilon = \mathbf{z}(\mathbf{x})\boldsymbol{\beta} + \varepsilon \tag{2.3}$$

Where ε is a zero-mean random deviation about the mean, usually called *error term*. The distribution of Y clearly depends on the distribution of ε.

The results of the n experiments, starting from the response model (2.3), can be collectively written in the following matrix notation

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.4}$$

where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}(\mathbf{x}_1) \\ \vdots \\ \mathbf{z}(\mathbf{x}_n) \end{bmatrix} \tag{2.5}$$

is called the *design matrix* and collects the row vectors $\mathbf{z}$ from all n experiments. Determination of the regression coefficients $\boldsymbol{\beta}$ depends on the experimental conditions, i.e. the conditions under which the experiments are carried out. If the experiments are carried out in *homogeneous* conditions, i.e. all under the same identical settings (but for the controlled variables $\mathbf{x}$ that obviously change from experiment to experiment), then the experiments can be assumed to be *independent* and the vector of errors $\boldsymbol{\varepsilon}$ can be assumed to be distributed as $N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$. The latter assumption consists actually of three components, the first of which is the *normality* assumption, the second the *independence* assumption and the third the *constant variance* assumption.

The model described above represents the simplest possible linear regression model and its treatmnt is presented in the next section. In following sections, in extending the model to account for different experimental conditions, the normality assumptions will be retained, while the other two eventually will be removed (correlated errors with non-constant variance).

2.1.1 Ordinary least squares method

The well-known method of Ordinary Least Squares (OLS) allows estimation of the regression coefficients when $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$. Let $\mathbf{b}$ be the estimates of the parameters $\boldsymbol{\beta}$. The function *sum of squared errors* (S), often referred to as the *residual sum of squares* (RSS)

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}) \tag{2.6}$$

is a measure of the goodness of the model for the mean of Y. OLS consists in deriving the regression parameters as those that minimise $S(\boldsymbol{\beta})$. Differentiating (2.6) with respect to $\boldsymbol{\beta}$ and setting the derivatives to zero yields

$$\mathbf{Z}^T \mathbf{Z}\boldsymbol{\beta} - \mathbf{Z}^T \mathbf{Y} = \mathbf{0}$$

from which

$$\mathbf{b} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{Y} \tag{2.7}$$

It is observed that no prior knowledge of the error common variance σ_ε^2 is required to estimate the regression coefficients. The variance can be estimated through the expression (given here without the justification that could be given based on the ANalysis Of VAriance)

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\mathbf{b})}{n - (p+1)} \tag{2.8}$$

which shows that to decrease the error variance the number of experiments must be considerably larger than the number of parameters to be determined. The parameter estimates $\mathbf{b}$ are unbiased as shown by

$$E[\mathbf{b}] = E[(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y}] = E[(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{Z}\boldsymbol{\beta}+\boldsymbol{\varepsilon})] = \boldsymbol{\beta} \tag{2.9}$$

and have covariance matrix

$$\mathbf{C_{bb}} = \mathrm{Var}[\mathbf{b}\mathbf{b}^T] = E\Big[\Big((\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y}\Big)\Big((\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y}\Big)^T\Big] = \sigma_\varepsilon^2(\mathbf{Z}^T\mathbf{Z})^{-1} \tag{2.10}$$

where use is made of $\mathbf{C_{YY}} = \mathrm{Var}[\mathbf{Y}\mathbf{Y}^T] = \sigma_\varepsilon^2\mathbf{I}$ as shown in (2.11). An estimate of the covariance in (2.10) can be obtained replacing the unknown error variance σ_ε^2 with its estimate $\hat{\sigma}_\varepsilon^2$ in (2.8). Such an estimate is useful if statistical tests of significance are to be carried out on the model parameters to assess the relevance of the various terms included in the model for the mean of Y.

The method of maximum likelihood estimation (MLE) is known to produce estimates having desirable properties. Thus ordinary least squares is often justified by the argument that OLS estimates $\mathbf{b}$ of $\boldsymbol{\beta}$ are maximum likelihood estimators when $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2\mathbf{I})$. This is seen as follows. Taking into account (2.4), the covariance of $\mathbf{Y}$ is

$$\mathbf{C_{YY}} = E[(\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})(\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})^T] = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma_\varepsilon^2\mathbf{I} \tag{2.11}$$

from which, since

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}) \quad \rightarrow \quad \mathbf{Y} \sim N(\mathbf{Z}\boldsymbol{\beta}, \mathbf{C_{YY}})$$

the likelihood function, proportional to the probability of observing $\mathbf{Y}$ given $\boldsymbol{\beta}$ and $\mathbf{C_{YY}}$ (i.e. σ_ε^2), follows

$$L(\boldsymbol{\beta}, \sigma_\varepsilon^2|\mathbf{Y}) = \sigma_\varepsilon^{-n}\exp\left(\frac{-(\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})^T(\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right) = \sigma_\varepsilon^{-n}\exp\left(\frac{-S(\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right) \tag{2.12}$$

which shows that maximisation of L with respect to $\boldsymbol{\beta}$ is equivalent to minimisation of S, and therefore the estimates $\mathbf{b}$ of $\boldsymbol{\beta}$ can be determined independently of σ_ε^2. When deviation of the errors from the assumed distribution makes OLS inappropriate as a means to estimate regression parameters, MLE can still be resorted to.

2.1.2 Experimental design: preliminary remarks on planning the experiments

Regarding Equation (2.8) it was observed that the scatter of the model will decrease for increasing number of experiments n, and that the number of experiments should be considerably larger than p, the number of regression parameters. A similar observation can be made for the variance of the parameter estimates $\mathbf{b}$, as the covariance matrix given in (2.10) is directly dependent on the estimate of σ_ε. It is therefore clear that data for statistical parameter estimation should come from carefully planned numerical experiments, finding a balance between amount of experiments and accuracy in the parameter estimates. In the literature a great number of so called *experimental plans* is available, of which herein only the one more suitable for the chosen quadratic form of the response surface is discussed, the *Central Composite Design* (CCD).

The CCD consists of a *two-level factorial design*, augmented with a *star design*. A complete two-level factorial design in k variables consists in choosing 2 levels for each input variable x_i, and then performing the experiments for all their 2^k possible combinations. A complete factorial design can provide not only the direct effect on the response Y of each variable x_i, but also the effect on Y of the interactions among the variables, up to the k^{th} order.

Taking as an example a case with 3 variables, the complete factorial design is compatible with a model of the type:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3 + \beta_7 x_1 x_2 x_3 + \varepsilon \qquad (2.13)$$

For a complete 2^k design, the variance of the estimates of β_i is $\mathrm{Var}[b_i] = 4\sigma_\varepsilon^2/2^k$, which can be justified based on (2.10).

Assuming a complete number of experiments is carried out, the estimates $\mathbf{b}$ of the parameters $\boldsymbol{\beta}$ can of course be determined directly using Equation (2.7). In the classical regression analysis, however, an alternative estimation procedure is often encountered which is here recalled for the insight that can be gained from its application. The values of $\mathbf{b}$ can be evaluated on the basis of the following definitions:

- the *main* effect of a variable is half the difference in the average value of Y when the variable in question moves from the high to the low level. The average value of Y is meant over all remaining variables taking their upper and lower levels.
- *interaction* between two variables i and j (two-way interaction) exists if the main effect of i is different for the two levels of j, and is defined as half the difference of the main effects of i.
- similarly, the *interaction* between ij and l (three-way interaction) exists if the interaction ij is different for the two levels of l, and is defined as half the difference of the two values of the interaction ij.

It can be easily verified that the estimates of the β_i's evaluated with the Least Squares and according to the given definitions coincide. The evaluation of the coefficients according to the definitions can be made automatically through the use of a *table of signs*, which is now introduced because of its great usefulness in the understanding of the nature of the various factorial effects and because it will be used later on in section (2.4.1) in order to explain other features of experimental design.

For notational convenience we will first introduce the concept of *coded variables*. As experimental plans are usually described in terms of normalised *levels* of the input variables, the variables x_i are transformed into $\check{x}_i$ such that the $\check{x}_i$ take values of -1 (a minus sign in the table of signs) and 1 (a plus sign) when the x_i take the lower (x_l) and upper (x_u) value, respectively. This is achieved by

$$\check{x}_i = \frac{x_i - (x_u + x_l)/2}{(x_u - x_l)/2} \tag{2.14}$$

In those cases where the input variables are defined as random variables with a known distribution, the upper and lower levels are often chosen to be mean plus and minus one standard deviation.

	I	**1**	**2**	**3**	**12**	**13**	**23**	**123**
	+	-	-	-	+	+	+	-
	+	+	-	-	-	-	+	+
	+	-	+	-	-	+	-	+
	+	+	+	-	+	-	-	-
	+	-	-	+	+	-	-	+
	+	+	-	+	-	+	-	-
	+	-	+	+	-	-	+	-
	+	+	+	+	+	+	+	+
Divisor	8	4	4	4	4	4	4	4

Table 2.1: Columns of signs and the divisors for systematically obtaining the factorial effect in a 2^3 factorial design

With reference to the factorial model in Equation (2.13), the table of signs is as shown in Table 2.1. The rows in the table represent the eight possible combinations of the two levels of x_1, x_2 and x_3, defining the experiments needed for determining the eight values of the coefficients b_i [the estimates of the β_i in Equation (2.13)]. The numbers in the upper row denote the different effects (or explanatories). The arrangement of the signs in the first columns (after the one labelled **I**) has an easily detectable pattern, called *standard*, and extends in an obvious way to the case of more than three variables. It is easy to show that by combining the outcomes of the eight experiments with the signs of each of the first three columns and dividing the results by the factor at the bottom (2^{k-1}), the direct effects of the first three variables are obtained. The two-factor interactions 12, 13 and 23, as well as the three-factor interaction 123 are obtained similarly using the signs of their

respective columns. These latter are simply the results of taking the products of the signs of the two columns of interest, e.g. column 1 and 2 for the interaction 12, row by row. The initial column of plus signs, with the heading **I**, provides the average of the response values (b_0), and its divisor is obviously 2^k.

For the determination of a full quadratic form (i.e., including the effects of the squared terms) a two-level factorial design is not adequate, as each variable is applied at two levels only. A 3^k factorial plan, on the other hand, would rapidly require an excessive number of experiments with increasing number of variables, a number larger than strictly necessary for implementing a second order model. For this reason the CCD, illustrated for $k = 3$ in Figures 2.1 and 2.2 adds further points to the factorial design (the *cube* portion in Figure 2.1), located at the *centre* and at two points along each axis of the variables (*star* points).

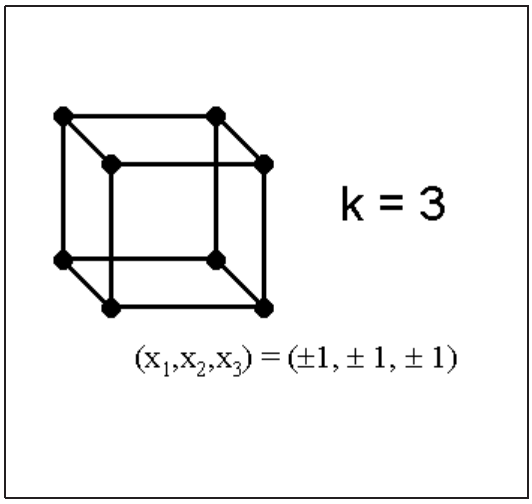

Figure 2.1: Factorial portion

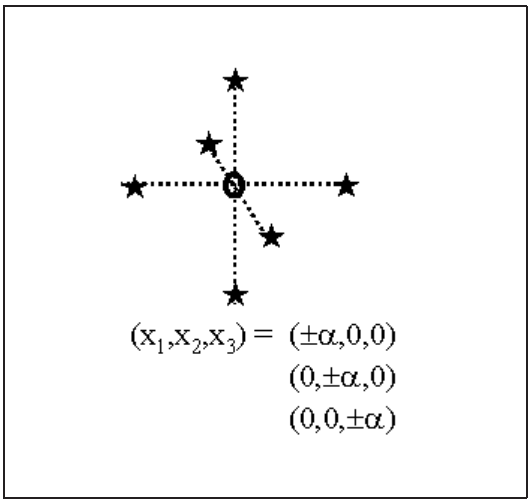

Figure 2.2: Star and centre points

To achieve *rotatability*, the coordinates of the star points must be located at distances $(2^k)^{1/4}$ from the centre. The number n_0 of central points is also related to rotatability (see e.g. Box and Draper (1987, [4]) for details). This property assures, according to its definition, a spherically uniform variance of the predicted response, and it is therefore intuitively appealing. However, approximate sphericity is all that is needed in practice, and, for example, a value of n_0 between 5 and 10 would be adequate for a problem with 6 variables.

2.2 Extension to time variant problems

It is useful to begin the discussion on how the tools presented sofar should be modified for the time-variant problem from the experimental design side. The selected CCD consists of $n = 2^k + 2k + n_0$ experimental runs. This means that each additional basic variable x_i introduced in the model almost doubles the number of necessary experiments. Practical considerations based either on economic or time-constraints set precise limits on the di-

mension of the vector $\mathbf{x}$. Rarely, in the applied statistics literature, one encounters designs for more than six variables, which already leads to close to a hundred experiments. Even if this limit can be reasonably raised in the context of EE, where experiments are numerical simulations, there is no interest in a procedure requiring hundreds of analyses, if the goal is to arrive at a method which is affordable when applied to realistic structures.

Seismic risk assessment, the specific time-variant problem dealt with here, is complicated by the very nature of the seismic action. The mathematical description of an earthquake is a challenge in itself. Choosing to use the language of random process theory, one could say that an earthquake is a non-Gaussian, fully non stationary process in the sense that both its amplitude and frequency content vary with time and that physical constraints limit the values of ground acceleration, velocity and displacement. Description of such a phenomenon, even in a simplified form, would require the introduction of a very large number of variables. For example, a common time-domain description in terms of a train of random acceleration pulses easily requires several hundreds, if not thousands, of variables to attain acceptable accuracy. Alternatively, a frequency-domain description requires discrete power spectral density and phase spectrum that, again, correspond to several hundreds of variables. These numbers are not compatible with the above mentioned limits.

These considerations answer a question that might have been asked from the beginning, i.e. whether the model in (2.3) is the most suited for dealing with the seismic fragility assessment problem.

Once the conclusion is reached that it is not possible to *explicitly* account for the effect of this large number of variables in the response model Y, the option of *implicitly* accounting for their effect in a somewhat global manner arises quite naturally. The basic variables vector $\mathbf{x}$ can then be imagined to be partitioned into two sub-vectors, $\mathbf{x_1}$ and $\mathbf{x_2}$, the first collecting the small number of most relevant variables whose effect is accounted for explicitly through the fixed effect type part of the model while the latter collects all the other (many) variables.

It remains to be chosen the way to account for the $\mathbf{x_2}$ variables in the model. This can be done introducing the concept of *random factors*. The earthquake, the complex aleatory phenomenon at hand, is treated as a factor that randomly affects the response Y in an additive manner. The model can thus be rewritten as

$$Y(\mathbf{x_1}, \mathbf{x_2}) = \mathbf{z}(\mathbf{x_1})\boldsymbol{\beta} + \delta + \varepsilon \tag{2.15}$$

where $\mathbf{z}(\mathbf{x_1})\boldsymbol{\beta}$ is the fixed effect that explicitly accounts for the effect of $\mathbf{x_1}$, while δ is the *random factor effect* (or simply *random effect*) and implicitly represents the influence on Y of the random factor $\mathbf{x_2}$. Thus, rather than estimating a prohibitive number of regression coefficients on $\mathbf{x_2}$, the problem is shifted to the estimation of the probability distribution of δ.

A common hypothesis is that δ is a zero-mean variable, which implies that the mean effect

of the random factor is accounted for in the constant term in the fixed effect portion of the model. If the other common hypothesis is made, that δ, as the error term ε, is Gaussian, the estimation of the distribution of δ reduces to that of its variance σ_δ^2.

Estimation of this variance component requires repeated observations of δ, each of which is the effect of a particular realisation of the earthquake. The latter realisation is called, in the statistical literature, a *level* of the random factor.

The necessary modifications in the statistical model, regarding the parameter estimation technique and the experimental design, will be presented in the next sections. For the former, the treatment of models such as that in (2.15), which are called *mixed* since both fixed and random effect are present, is discussed. As far as the latter is concerned, the concept of *blocking* of the experimental plan is introduced.

2.2.1 Blocks and random factor effects

Blocking of the experimental plan is a well-established technique often used when for some reason the experiments can not be carried out in homogeneous conditions. This may happen when the available experimental *units* do not come from the same batch, i.e. there is some uncontrollable and unknown source of dishomogeneity whose effect is to be determined. Alternatively the experiments can be carried out on purpose under non homogeneous conditions, in order to determine the effect of this variation, as it is done in the case under consideration where the factor that determines the dishomogeneity is the earthquake. When the experiments are divided in blocks the assumptions of independence and, possibly, constant variance of the errors $\boldsymbol{\varepsilon}$ have to be relaxed.

The case when the variance is still constant but the errors are correlated is considered first.

Blocking occurs when a factor is varying among groups of experiments (*blocks*). Here we make the common (and most often reasonable) assumption that this factor changes only additively the response Y (Fischer). If the n experimental runs are grouped into b blocks with $n_1, n_2, \ldots, n_b$ runs respectively such that

$$\sum_{i=1}^{b} n_i = n$$

the results of these experiments, taking into account the model form in (2.15), can be collectively written in matrix form as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \tag{2.16}$$

where $\boldsymbol{\delta}$ is a vector of block effects and the matrix $\mathbf{B}$ *collocates* these effects to the appro-

priate experiments. It is recalled that this model is a mixed (effect) model with $\mathbf{Z}\boldsymbol{\beta}$ being its fixed effect part and $\mathbf{B}\boldsymbol{\delta}$ its random effect one. The form of the collocation matrix $\mathbf{B}$ is

$$\mathbf{B} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_b} & \mathbf{0}_{n_b} & \cdots & \mathbf{1}_{n_b} \end{bmatrix} \tag{2.17}$$

where $\mathbf{1}_{n_i}$ and $\mathbf{0}_{n_i}$ are vectors of length n_i made up of ones and zeros respectively. When the model in (2.16) includes a constant term, i.e. when matrix $\mathbf{Z}$ has a column made up of ones, the model itself is no more full rank since the columns of $\mathbf{B}$ sum up to a column of ones too, thus giving two linearly dependent columns: $\mathbf{Z}\mathbf{1}_b = \mathbf{1}_N$. The block effects are confounded with the constant term of the model and unless one introduces an additional condition on the block effects they can not be determined separately from β_0. The model can be rewritten as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\tau} + \boldsymbol{\varepsilon} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}(\beta_0\mathbf{1}_b + \boldsymbol{\delta}) + \boldsymbol{\varepsilon} \tag{2.18}$$

The additional condition is commonly chosen as

$$\boldsymbol{\delta}^T\mathbf{1}_b = 0$$

which means that the block effects sum up to zero, so that only the mean block effect is confounded with the constant term. In this case one gets

$$\mathbf{1}_b^T\boldsymbol{\tau} = \mathbf{1}_b^T(\beta_0\mathbf{1}_b + \boldsymbol{\delta}) = b\beta_0$$

from which

$$\beta_0 = \frac{1}{b}\mathbf{1}_b^T\boldsymbol{\tau}$$

where the $\boldsymbol{\tau}$ are still unknown.

In cases other than the one considered here, one might be interested in determining explicitly the effect of the blocks on the response. In this case one would speak of the block effects as "fixed", and the determination of the regression parameters and of the block effects could be performed by OLS after having enlarged the regression problem. Equation (2.18) can be rewritten as

$$\mathbf{Y} = \widetilde{\mathbf{Z}}\tilde{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} = [\mathbf{Z}\ \mathbf{B}]\begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{\tau}\end{bmatrix} + \boldsymbol{\varepsilon} \tag{2.19}$$

which can be solved by OLS. In particular the regression parameters and fixed block effects estimates $\mathbf{b}$ and $\mathbf{t}$ are

$$\tilde{\boldsymbol{\beta}} = (\widetilde{\mathbf{Z}}^T\widetilde{\mathbf{Z}})^{-1}\widetilde{\mathbf{Z}}^T\mathbf{Y} \tag{2.20}$$

2.2.2 Single blocking

If the blocking of the experimental plan is determined by some random factor, as in the case at hand where a ground acceleration record (the realisation of a random ground-motion) determines the blocking, it is quite logical to assume that the block effect will be also random rather than fixed. In this case $\boldsymbol{\delta}$ in model (2.16) is a random vector. Data analysis for such a model is described in depth in Searle et.al. (1992, [13]) and a first brief introduction is given (for the case of single blocking) in Khuri and Cornell (1996, [8]). Assuming the random factor is distributed as $N(\mathbf{0}, \sigma_\delta^2\mathbf{I})$ independently of $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2\mathbf{I})$, the mean and covariance matrix of the response according to this model become

$$\boldsymbol{\mu}_\mathbf{Y}(\mathbf{x}) = \mathbf{Z}(\mathbf{x_1})\boldsymbol{\beta}$$

and

$$\begin{aligned}\mathbf{C_{YY}} &= E\Big[(\mathbf{Y}-\boldsymbol{\mu}_\mathbf{Y})(\mathbf{Y}-\boldsymbol{\mu}_\mathbf{Y})^T\Big] = E\Big[(\mathbf{B}\boldsymbol{\delta}+\boldsymbol{\varepsilon})(\mathbf{B}\boldsymbol{\delta}+\boldsymbol{\varepsilon})^T\Big] = \\ &= \sigma_\delta^2\mathbf{B}\mathbf{B}^T + \sigma_\varepsilon^2\mathbf{I}_n = \sigma_\varepsilon^2\mathbf{A}\end{aligned} \tag{2.21}$$

where $\mathbf{A} = \mathrm{diag}[\mathbf{A}_1 \ldots \mathbf{A}_b]$ is a block-diagonal matrix made up of matrices

$$\mathbf{A}_i = \gamma\mathbf{J}_{n_i} + \mathbf{I}_{n_i}$$

where $\mathbf{J}_{n_i}$ is $n_i \times n_i$ square matrix made up of ones coming from the matrix product $\mathbf{B}\mathbf{B}^T$, $\mathbf{I}_{n_i}$ is the $n_i \times n_i$ sub-matrix of the $n \times n$ identity $\mathbf{I}_n$ and $\gamma = \sigma_\delta^2/\sigma_\varepsilon^2$ is the ratio of the variances. The last equality in (2.21) introduces the so-called Hartley-Rao form of the covariance matrix for the mixed effects model [Searle et.al. (1992, [13])].

For example, the response covariance matrix form for a hypothetical case with $n = 6$ experiments divided in $b = 3$ blocks of two experiments each is given below.

$$\begin{aligned}
\mathbf{C_{YY}} &= \sigma_\varepsilon^2 \left[\begin{array}{cc:cc:cc} 1+\gamma & \gamma & 0 & 0 & 0 & 0 \\ \gamma & 1+\gamma & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 1+\gamma & \gamma & 0 & 0 \\ 0 & 0 & \gamma & 1+\gamma & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1+\gamma & \gamma \\ 0 & 0 & 0 & 0 & \gamma & 1+\gamma \end{array}\right] = \\
&= \breve{\sigma}_\varepsilon^2 \left[\begin{array}{cc:cc:cc} 1 & \rho & 0 & 0 & 0 & 0 \\ \rho & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 1 & \rho & 0 & 0 \\ 0 & 0 & \rho & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & 1 \end{array}\right]
\end{aligned} \tag{2.22}$$

where $\rho = \gamma/(1+\gamma)$ and $\breve{\sigma}_\varepsilon^2 = (1+\gamma)\sigma_\varepsilon^2$ which shows that the effect of blocking is to increase the error variance and to introduce *within-block* correlation among the errors. This within-block correlation ρ is equal under the distribution hypothesis for the block effects $\boldsymbol{\delta} \sim N(\mathbf{0}, \sigma_\delta^2 \mathbf{I})$ independent of $\boldsymbol{\varepsilon}$.

In the hypothetical case one knew the variance ratio γ the problem could be easily solved by MLE. Setting $\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma_\varepsilon^2 \mathbf{W}^{-1}$ the likelihood function in this case reads

$$\begin{aligned}
L(\boldsymbol{\beta}|\mathbf{Y}, \sigma_\varepsilon^2, \gamma) &= L(\boldsymbol{\beta}|\mathbf{Y}, \sigma_\varepsilon^2, \sigma_\delta^2) = \sigma_\varepsilon^{-n} |\mathbf{W}|^{-1/2} \exp\left[\frac{-(\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})^T \mathbf{W} (\mathbf{Y}-\mathbf{Z}\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right] = \\
&= \sigma_\varepsilon^{-n} |\mathbf{W}|^{-1/2} \exp\left[\frac{-Q(\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right]
\end{aligned} \tag{2.23}$$

which shows that maximisation of L with respect to $\boldsymbol{\beta}$ is equivalent to minimisation of $Q(\boldsymbol{\beta})$. The solution obtained

$$\mathbf{b} = (\mathbf{Z}^T \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \mathbf{Y} \tag{2.24}$$

is also known as *Weighted Least Squares* (WLS) and clearly coincides with OLS when $\mathbf{A} = \mathbf{I}$, i.e. the errors are uncorrelated. In real cases one does not know the variances

ratio *a priori* but has to estimate it by some method, for instance using maximum likelihood estimation. In this case the estimation of the regression parameters depends on the variances since the likelihood should be read as

$$L(\boldsymbol{\beta}|\sigma_\varepsilon^2,\gamma,\mathbf{Y}) \rightarrow L(\boldsymbol{\beta},\sigma_\varepsilon^2,\gamma|\mathbf{Y}) = \sigma_\varepsilon^{-n}|\mathbf{W}|^{-1/2}\exp\left[\frac{-Q(\boldsymbol{\beta},\gamma)}{2\sigma_\varepsilon^2}\right] \tag{2.25}$$

MLE of the variances requires a first estimate of the regression parameters and thus the problem must be solved iteratively. One possible way to proceed is to determine a first estimate of the parameters $\boldsymbol{\beta}$, $\mathbf{b}^{(1)}$, for example by OLS, i.e. assuming zero correlation, then using these estimates in (2.25) to estimate the variances only. With these variances one can then use (2.24) to obtain updated values of the parameters $\boldsymbol{\beta}$ repeating the two steps until convergence to a specified tolerance is attained. This is equivalent to splitting the solution of the $p+2$ dimensional constrained optimisation problem

$$\begin{aligned}
&\frac{\partial L(\boldsymbol{\beta},\sigma_\varepsilon^2,\gamma)}{\partial\boldsymbol{\beta}} = 0\\
&\frac{\partial L(\boldsymbol{\beta},\sigma_\varepsilon^2,\gamma)}{\partial\sigma_\varepsilon^2} = 0\\
&\frac{\partial L(\boldsymbol{\beta},\sigma_\varepsilon^2,\gamma)}{\partial\gamma} = 0\\
&\text{for } \sigma_\varepsilon^2 > 0, \gamma \geq 0
\end{aligned} \tag{2.26}$$

into two sequential optimisation problems whose dependence is accounted for by the iterative solution. An alternative way to proceed would be to solve directly (2.26) as a whole by one of the various optimisation techniques implemented in standard packages like Matlab. The solution of the above problem unfortunately turns out to be numerically awkward. The delicateness of the task of implementing a computer code for the determination of variance components for a mixed model is pointed out more than once in Searle et.al. (1992, [13]).

An alternative solution to the problem is attempted using a *conditional log-likelihood* often encountered in hierarchical models. The discussion of this alternative is postponed until the mixed effects model is extended to *multiple blocking* (multiple random factors) in the following section 2.3. The necessary modifications to the experimental plan, on the other hand, introducing the concept of blocking, will be presented in 2.4.

2.3 Introduction of additional random factors

The extension of the statistical model from a fixed effect model to a mixed effect model has been done in order to introduce dependence of the response on earthquake loading.

However, random effects can be useful in the characterisation of other sources of variability as well. Mechanical parameters, for example, characteristically exhibit a spatial variability over the whole structure. If this variability were to be described in terms of fixed effects, a large amount of input variables, such as those resulting from a random field discretisation, would be required. A practical solution can be obtained with the introduction of a random factor. In this way, locally, the variability will depend on the product of an uncertain mean value and a spatial fluctuation around the mean. The uncertain mean will be represented by a single basic variable in $\mathbf{x_1}$, and the fluctuation by a random factor, with mean equal to one.

Introduction of additional random factors gives rise to the need of a further extension of the theory developed so far.

2.3.1 Multiple blocking

Consider again the model in (2.15). This model can be easily extended to include additional random factor effects

$$Y = \mathbf{z}\boldsymbol{\beta} + \sum_i \delta_i + \varepsilon \tag{2.27}$$

Correspondingly, the results of n experiments are collectively written in matrix form, similarly to (2.16), as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \tag{2.28}$$

where $\mathbf{Z}\boldsymbol{\beta}$ are the fixed effects and $\boldsymbol{\delta}$ the random effects. Until now the vector $\boldsymbol{\delta}$ has collected the effects of all q levels of the single random factor present. The random factor was for instance the earthquake ground acceleration and its levels selected accelerograms. Now consider the extension of this model to the general case of r distinct random factors. In this model multiple blocking arises from multiple sources of dishomogeneity such as random ground acceleration, a random field representing the spatial variability of the concrete ultimate strain or steel yield strength, or, in general, of any significantly uncertain and fluctuating mechanical parameter in the structural model. The random effect portion of the model will then be partitioned as follows

$$\mathbf{B}\boldsymbol{\delta} = [\mathbf{B}_1 \cdots \mathbf{B}_r] \begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_r \end{bmatrix} = \sum_{i=1}^{r} \mathbf{B}_i \boldsymbol{\delta}_i \tag{2.29}$$

where $\boldsymbol{\delta}_i$ collects the effects of all the q_i levels of random factor i occurring in the data. The same assumptions as for the previous single-random-factor case are made here for the

r-random-factor model, i.e.

$$E[\boldsymbol{\delta}_i] = \mathbf{0} \quad \mathbf{C}_{\boldsymbol{\delta}_i\boldsymbol{\delta}_i} = E[\boldsymbol{\delta}_i\boldsymbol{\delta}_i^T] = \sigma_{\delta_i}^2\mathbf{I}_{q_i} \tag{2.30}$$

with the additional one

$$E[\boldsymbol{\delta}_i\boldsymbol{\delta}_i^T] = \mathbf{0} \qquad \forall i \neq j \tag{2.31}$$

from which it follows that the covariance matrix of $\boldsymbol{\delta}$ is block diagonal, made up of diagonal matrices of dimension q_i which can be written as

$$\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}} = \{_d\sigma_{\delta_i}^2\mathbf{I}_{q_i}\}_{i=1}^r \tag{2.32}$$

Further, as usual, for the error vector it is assumed that

$$E[\boldsymbol{\varepsilon}] = \mathbf{0} \quad \mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \sigma_\varepsilon^2\mathbf{I}_n \quad E[\boldsymbol{\varepsilon}\boldsymbol{\delta}_i^T] = \mathbf{0}\ \forall i \tag{2.33}$$

Using above assumptions leads, similarly to the case of a single random factor, to

$$E[\mathbf{Y}] = \mathbf{Z}\boldsymbol{\beta} \quad \mathbf{C}_{\mathbf{YY}} = \sum_{i=1}^{r}\mathbf{B}_i\mathbf{B}_i^T\sigma_{\delta_i}^2 + \sigma_\varepsilon^2\mathbf{I}_n \tag{2.34}$$

For notational convenience one can introduce the following relations

$$\delta_0 = \varepsilon \qquad q_0 = n \qquad \mathbf{B}_0 = \mathbf{I}_n \tag{2.35}$$

and write the model in (2.28) as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \sum_{i=0}^{r}\mathbf{B}_i\boldsymbol{\delta}_i + \boldsymbol{\varepsilon} \tag{2.36}$$

with the same mean and covariance matrix equal to

$$\mathbf{C}_{\mathbf{YY}} = \sum_{i=0}^{r}\mathbf{B}_i\mathbf{B}_i^T\sigma_{\delta_i}^2 \tag{2.37}$$

The Hartley-Rao form in this case is symbolically the same as for the single random factor case, i.e.

$$\mathbf{C_{YY}} = \mathbf{A}\sigma_\varepsilon^2 \tag{2.38}$$

where $\mathbf{A}$ has the same form as the covariance matrix except that where σ_ε^2 appears in $\mathbf{C_{YY}}$ there is a 1 in $\mathbf{A}$ and where there is a $\sigma_{\delta_i}^2$ in $\mathbf{C_{YY}}$ there is a $\gamma_i = \sigma_{\delta_i}^2/\sigma_\varepsilon^2$ in $\mathbf{A}$.

For example, if for a hypothetical case with $n = 4$ experiments and $r = 2$ random factors, sub-divided in $b = 2$ blocks of two experiments each ($q_i = 2$) the matrices $\mathbf{B}_i$ are defined as

$$\mathbf{B}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{B}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.39}$$

the response covariance matrix would take the form

$$\begin{aligned} \mathbf{C_{YY}} &= \left[\begin{array}{cc:cc} \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & \sigma_{\delta_1}^2 & \sigma_{\delta_2}^2 & 0 \\ \sigma_{\delta_1}^2 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & 0 & \sigma_{\delta_2}^2 \\ \hdashline \sigma_{\delta_2}^2 & 0 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & \sigma_{\delta_1}^2 \\ 0 & \sigma_{\delta_2}^2 & \sigma_{\delta_1}^2 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 \end{array}\right] = \\ &= \breve{\sigma}_\varepsilon^2 \left[\begin{array}{cc:cc} 1 & \rho_1 & \rho_2 & 0 \\ \rho_1 & 1 & 0 & \rho_2 \\ \hdashline \rho_2 & 0 & 1 & \rho_1 \\ 0 & \rho_2 & \rho_1 & 1 \end{array}\right] \end{aligned} \tag{2.40}$$

where $\rho_i = \sigma_{\delta_i}/(\sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2)$ and $\breve{\sigma}_\varepsilon^2 = (\sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2)$, which shows that the effect of multiple blocking no longer leads to a block diagonal correlation matrix, but that next to the within block correlation also *inter-block* correlation is introduced among the errors. It can be easily understood that extension to a situation with more blocks will eventually lead to a full correlation matrix.

2.3.2 Hierarchical model formulation and conditional likelihood

Hierarchical techniques can help both in understanding the statistical models and in estimating and interpreting the model parameters. A hierarchical model is one that is specified

in stages, with each stage built upon the previous one. The advantage of building a model in stages is that each stage can be relatively simple and easy to understand, while the entire model may be rather complicated.

It is recalled that the general mixed model equation for a response Y and n experiments has been written as (2.28)

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \tag{2.41}$$

where $\boldsymbol{\beta}$ is an unknown vector of fixed parameters, $\boldsymbol{\delta}$ is an unknown vector of random (factor) effects and $\boldsymbol{\varepsilon}$ is an unknown vector of random error terms. The matrices $\mathbf{Z}$ and $\mathbf{B}$ are fixed and known (they are determined by the experimental plan and the way it is blocked).

The ordinary likelihood function for the model above (i.e. according to a non-hierarchical approach) has the form

$$L(\boldsymbol{\beta}, \mathbf{C_{YY}}|\mathbf{Y}) = |\mathbf{C_{YY}}|^{-\frac{1}{2}} \exp\Big(-(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta})^T \mathbf{C_{YY}^{-1}} (\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta})\Big) \tag{2.42}$$

which closely resembles (2.25) with the difference that now the unknown terms of the symmetric covariance matrix $\mathbf{C_{YY}}$ of (2.37) are $r + 1$, with a maximum of $([n \times (n - 1)]/2) + 1$. Techniques for the estimation of the parameters (both regression on the mean $\boldsymbol{\beta}$ and variance components $\mathbf{C_{YY}}$) of the model, starting from the likelihood in (2.42) are presented in Searle et.al. (1992, [13]). It has already been observed that this is not a straightforward task, and that considerable numerical difficulties arise. The *hierarchical formulation* of the model offers an alternative estimation technique.

In the classical approach for analysis of data using a mixed model, the distinction of fixed versus random, known versus unknown, is important. This classification dictates the type of estimation and inference that is possible. When analysing the mixed model (or any other model) with a hierarchical approach, however, it only matters whether a specified quantity is *observable* or *unobservable*.

In model (2.41) $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{B}$ are observable (because their values are known), while $\boldsymbol{\beta}$, $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are unobservable (for the opposite reason). Both fixed and random effects $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are treated in the same way, because they are unobservable.

A hierarchy is established when the distribution of the observable response Y is specified conditional on the unobservable vectors $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ as a first stage of the model, and then the marginal distributions of the latter are specified as its second stage. This can be written as

$$\mathbf{Y} \sim f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}})$$

$$\boldsymbol{\beta} \sim f(\boldsymbol{\beta}|\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta\beta}}) \qquad \boldsymbol{\delta} \sim f(\boldsymbol{\delta}|\mathbf{C}_{\boldsymbol{\delta\delta}})$$

The joint distribution of $\mathbf{Y}$, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ is then obviously given by

$$f(\mathbf{Y}, \boldsymbol{\beta}, \boldsymbol{\delta}|\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta\beta}}, \mathbf{C}_{\boldsymbol{\delta\delta}}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}) = f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}) f(\boldsymbol{\beta}|\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta\beta}}) f(\boldsymbol{\delta}|\mathbf{C}_{\boldsymbol{\delta\delta}})$$

and the likelihood function, proportional to the probability of observing $\mathbf{Y}$, is given by

$$L(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta\beta}}, \mathbf{C}_{\boldsymbol{\delta\delta}}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}|\mathbf{Y}) = \int\int f(\mathbf{Y}, \boldsymbol{\beta}, \boldsymbol{\delta}|\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta\beta}}, \mathbf{C}_{\boldsymbol{\delta\delta}}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}) \mathrm{d}\boldsymbol{\beta} \mathrm{d}\boldsymbol{\delta} \tag{2.43}$$

The latter is also called *full likelihood* and coincides with (2.42) in the case of the *normal model* with a *point-mass* density for $\boldsymbol{\beta}$, i.e. a density that concentrates all the probability mass at a point, which is equivalent to considering $\boldsymbol{\beta}$ as deterministic.

If the unobservable $\boldsymbol{\delta}$ are not integrated out as in (2.43) one obtains, for the normal model with point-mass density for $\boldsymbol{\beta}$, the following *conditional likelihood*

$$L(\boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{C}_{\boldsymbol{\delta\delta}}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}|\mathbf{Y}) = f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}) f(\boldsymbol{\delta}|\mathbf{C}_{\boldsymbol{\delta\delta}}) \tag{2.44}$$

which is equal to

$$\begin{aligned}
&L(\boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{C}_{\boldsymbol{\delta\delta}}, \mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}|\mathbf{Y}) = \\
&\quad = |\mathbf{C}_{\mathbf{YY}|\boldsymbol{\delta}}|^{-1} \exp\Big[-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}|\boldsymbol{\delta}})^T \mathbf{C}^{-\mathbf{1}}_{\mathbf{YY}|\boldsymbol{\delta}} (\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}|\boldsymbol{\delta}}) \Big] \times \\
&\quad \times |\mathbf{C}_{\boldsymbol{\delta\delta}}|^{-1} \exp\Big[-\frac{1}{2}(\boldsymbol{\delta} - \boldsymbol{\mu}_{\boldsymbol{\delta}})^T \mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\delta\delta}} (\boldsymbol{\delta} - \boldsymbol{\mu}_{\boldsymbol{\delta}}) \Big] = \\
&\quad = |\mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}|^{-1} \exp\Big[-\frac{1}{2}(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta})^T \mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\varepsilon\varepsilon}} (\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}) \Big] |\mathbf{C}_{\boldsymbol{\delta\delta}}|^{-1} \exp\Big[-\frac{1}{2}\boldsymbol{\delta}^T \mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\delta\delta}} \boldsymbol{\delta} \Big] = \\
&\quad = \sigma_{\varepsilon}^{-n} \prod_{i=1}^{r} \sigma_i^{-q_i} \exp\left[-\frac{1}{2\sigma_{\varepsilon}^2} \Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big)^T \Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big)^T - \sum_{i=1}^{r} \frac{1}{2\sigma_i^2} \boldsymbol{\delta}_i^T \boldsymbol{\delta}_i \right]
\end{aligned} \tag{2.45}$$

The corresponding log-likelihood is

$$\begin{aligned}
l(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_{\varepsilon}^2, \sigma_1^2, \ldots, \sigma_r^2) &= \log L(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_{\varepsilon}^2, \sigma_1^2, \ldots, \sigma_r^2) = \\
&= -n \log \sigma_{\varepsilon} - \sum_{i=1}^{r} q_i \log \sigma_i + \\
&+ \frac{1}{2\sigma_{\varepsilon}^2} \Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big)^T \Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big) + \sum_{i=1}^{r} \frac{1}{2\sigma_i^2} \boldsymbol{\delta}_i^T \boldsymbol{\delta}_i
\end{aligned} \tag{2.46}$$

Maximisation of l requires setting to zero its derivatives with respect to $\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_\varepsilon^2, \sigma_i^2, \ldots, \sigma_r^2$. These are

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\beta}} &= \frac{1}{\sigma_\varepsilon^2}\mathbf{Z}^T\Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big) \\
\frac{\partial l}{\partial \boldsymbol{\delta}} &= (\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-\mathbf{1}} + \mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{B})\boldsymbol{\delta} - \mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-1}\Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}\Big) \\
\frac{\partial l}{\partial \sigma_\varepsilon^2} &= -\frac{n}{\sigma_\varepsilon} + \frac{1}{\sigma_\varepsilon^3}\Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big)^T\Big(\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{B}\boldsymbol{\delta}\Big) \\
\frac{\partial l}{\partial \sigma_i^2} &= -\frac{q_i}{\sigma_i} + \frac{1}{\sigma_i^3}\boldsymbol{\delta}_i^T\boldsymbol{\delta}_i \quad i = 1, \ldots, r
\end{aligned} \tag{2.47}$$

These, set to zero, give the non linear equations for the general case

$$\text{sub-set 1:} \left\{ \begin{array}{l} \mathbf{b} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{Y} - \mathbf{B}\boldsymbol{\delta}) \\ \boldsymbol{\delta} = (\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-\mathbf{1}} + \mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{Y} - \mathbf{Z}\mathbf{b} \end{array} \right.$$

$$\begin{aligned}
\left\{ \begin{array}{c} \mathbf{b} \\ \boldsymbol{\delta} \end{array} \right\} = \left[\begin{array}{cc} \mathbf{I}_p & (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{B} \\ (\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-\mathbf{1}} + \mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{Z} & \mathbf{I}_q \end{array} \right]^{-1} \times \\
\times \left\{ \begin{array}{c} (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ (\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-\mathbf{1}} + \mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}}\mathbf{Y} \end{array} \right\}
\end{aligned} \tag{2.48}$$

$$\text{sub-set 2:} \left\{ \begin{array}{l} \sigma_\varepsilon^2 = n^{-1}(\mathbf{Y} - \mathbf{Z}\mathbf{b} - \mathbf{B}\boldsymbol{\delta})^T(\mathbf{Y} - \mathbf{Z}\mathbf{b} - \mathbf{B}\boldsymbol{\delta}) \\ \sigma_i^2 = q^{-1}\boldsymbol{\delta}_i^T\boldsymbol{\delta}_i \qquad i = 1, \ldots, r \end{array} \right.$$

These are solved iteratively, first sub-set 1, then sub-set 2 and so on, starting the process with initial assumed values for the covariance matrices equal to $\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-\mathbf{1}} = \mathbf{0}$ and $\mathbf{C}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}^{-\mathbf{1}} = \mathbf{I}$. The first estimates of the regression coefficients $\mathbf{b}^{(1)}$ and of the random factors levels $\boldsymbol{\delta}^{(1)}$ are given by

$$\left\{ \begin{array}{c} \mathbf{b}^{(1)} \\ \boldsymbol{\delta}^{(2)} \end{array} \right\} = \left[\begin{array}{cc} \mathbf{I}_p & (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{B} \\ (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Z} & \mathbf{I}_q \end{array} \right]^{-1} \left\{ \begin{array}{c} (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y} \end{array} \right\} \tag{2.49}$$

It can be observed finally that using this conditional form of the likelihood, for the same number of experiments, increases convergence at the expense of increased uncertainty in the determined parameters, since it augments the number of parameters to be considered moving the effects $\boldsymbol{\delta}$ of the random factor levels from the covariance of $\mathbf{Y}$ to its mean.

2.3.3 Covariance matrix of the parameters estimates

The covariance matrix of the parameter estimates is needed if, later, they are modelled as random variables in reliability analysis, or when significance tests are performed on them. The covariance matrix of the parameters can be easily obtained once sub-set 1 in (2.48) is rewritten as follows

$$\left\{ \begin{array}{c} \mathbf{b} \\ \boldsymbol{\delta} \end{array} \right\} = \mathbf{AY} = \left[\begin{array}{c} \mathbf{A}_1 \\ \mathbf{A}_2 \end{array} \right] \mathbf{Y}$$

It immediately follows that the covariance of the parameters $\mathbf{b}$ is given by

$$\mathbf{C_{bb}} = \mathbf{A_1 C_{YY} A_1^T} \tag{2.50}$$

where, of course, the matrices $\mathbf{C}_{\boldsymbol{\delta\delta}}$ and $\mathbf{C}_{\boldsymbol{\varepsilon\varepsilon}}$ are those obtained at convergence of the iterative process in (2.48). The existence of a covariance matrix for the estimates $\mathbf{b}$ of $\boldsymbol{\beta}$ may seem inconsistent with the choice for point-mass density of the parameters $\boldsymbol{\beta}$. However, considering $\boldsymbol{\beta}$ deterministic does not necessarily imply that $\mathbf{b}$ are deterministic as well, they are instead random and their uncertainty, which is epistemic, stems from the limited statistical basis on which their estimation relies.

2.4 Experimental design for the mixed model

In the foregoing discussion of the mixed model in (2.28) it has always been tacitly assumed that experimental data, in the form of n results Y_i, were available. How these experiments had been carried out, i.e. with which combinations of the input variables $\mathbf{x}$, was supposedly prescribed by an experimental plan. Design of an experimental plan plays a central role in establishing any statistical model. The way the experiments are carried out determines properties of the data-set that influence their statistical analysis in terms of which effects can be reliably estimated and which others cannot.

In section 2.1.2 a brief introduction to experimental design has been given, limited, however, to the central composite design (CCD) that has been adopted in this work. This design allows estimation of fixed effects only and is thus still *unfit* to be used in conjunction with the mixed model. It is therefore necessary to modify it, in order to allow estimation of the random factor effects. This extension consists of subdividing the planned experiments into groups, called blocks, to each of which a level of the random factor is associated. This operation is referred to as *blocking of the experimental plan*. The groups that result from blocking are mutually exclusive and collectively exhaustive, i.e. each experiment belongs to one block only. This holds as long as a single random factor is present in the model. With more than one random factor, however, multiple blocking of the plan is necessary

Run	1	2	3	4	5	6	123456	Block
1	-	-	-	-	-	-	+	II
2	+	-	-	-	-	-	-	I
3	-	+	-	-	-	-	-	I
4	+	+	-	-	-	-	+	II
5	-	-	+	-	-	-	-	I
6	+	-	+	-	-	-	+	II
7	-	+	+	-	-	-	+	II
8	+	+	+	-	-	-	-	I
9	-	-	-	+	-	-	-	I
10	+	-	-	+	-	-	+	II
11	-	+	-	+	-	-	+	II
12	+	+	-	+	-	-	-	I
13	-	-	+	+	-	-	+	II
14	+	-	+	+	-	-	-	I
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
63	-	+	+	+	+	+	-	I
64	+	+	+	+	+	+	+	II

Table 2.2: Blocking the 2^6 design into two blocks using B = 123456

and, while each blocking is made up of mutually exclusive and collectively exhaustive blocks, each experiment falls into more than one block, one from each blocking.

In the next two sections single and multiple blocking are presented. In particular the problem of confounded effects is discussed in order to arrive at a criterion to assign experiments to blocks. Further, the notion of orthogonal blocking is introduced and the feasibility of a unique plan for the estimation of both fixed and random effects is discussed.

2.4.1 Single blocking of the experimental plan

Various issues arise in subdividing the factorial part of the CCD in *suitably defined blocks* are described, so as to prevent possible bias in the estimates of the parameters that might be expected from the fact that not all the experimental units belong to a single statistical population for all the experiments. Following Box and Draper (1987, [4]), the problem is discussed making reference to an example, assuming that a complete two-level factorial design is to be carried out for $k = 6$. Such a design consists of $2^6 = 64$ experiments and the discussion is started assuming that these are to be divided into two equal blocks of 32 experiments each.

Consider the partial table of signs, Table 2.2, which shows the 2^6 design arranged in the so-called *standard* order (whose structure was discussed before). As a *criterion for blocking*, the sign resulting from the product of the signs of the six variables in each row is taken:

whenever this product, indicated by 123456 in the table, is negative, the run is assigned to block I, when it is positive to block II. The above subdivision in blocks has an important property, which consists in the fact that the columns with the signs of the variables for evaluating the effects are orthogonal with each other. This is a property of factorial combinations, valid whatever is the sequence of the experiments. The consequence of this property is that if, due to differences in the population of some variables, the response values in each block are changed by the same fixed quantity, the estimates of all effects do not change, except for the interaction 123456. This latter, however, is in most cases of too high-order to be significant. In the terminology of the discipline, it is said that the block effect is *confounded* with the 123456 interaction. In the same framework, $\mathcal{B} = 123456$ is called a *block generator*.

Before proceeding with subdivisions in more numerous, smaller, blocks of the same 2^6 design, it is helpful to go deeper into the evaluation of interactions between higher-order effects, an analysis started with the two-factor and three-factor interactions based on the table of signs, Table 2.1.

The main point is to note that the product of any column (sequence of signs for the calculation of any direct or interaction effect) by itself, gives a column of plus signs, which is called *identity* and denoted by $\mathbf{I}$ (check with the columns of Table 2.1), e.g.

$$1 \times 1 = 1^2 = \mathbf{I}^2 = \mathbf{I}$$

or

$$12 \times 12 = 1^2 2^2 = \mathbf{I}^2 = \mathbf{I}$$

If the above results are now applied to the product of two different multi-factor interactions, as for example 123456 and 23456, the following is obtained

$$123456 \times 23456 = 12^2 3^2 4^2 5^2 6^2 = 1 \cdot \mathbf{I}^5 = 1$$

It can be seen then that the interaction between the two multiple interactions (or *contrasts*, as they are frequently called in the statistical literature) is simply the direct effect 1. It would therefore be a really bad choice to use the two previous interactions as block generators, since the interaction between the two blocks would be confounded with the direct effect of the variable 1, which is certainly a main objective of the analysis.

The above considerations are immediately helpful for the purpose of planning a blocking scheme for a specific problem. The first decision to be taken by the experimenter is the highest order of the interactions among variables that he considers to be significant and hence not to be confounded. In the simple model adopted here, Equation (2.1), for example all of the three-factor interactions have already been sacrificed. In some cases, even two-factor interactions may be discarded from the beginning, based on physical reasoning.

For example, if one would choose the generators

$$\mathcal{B}_1 = 1234 \qquad \mathcal{B}_2 = 3456 \qquad \mathcal{B}_3 = 136$$

the following interactions will be confounded

$$\mathcal{B}_1 = 1234 \quad \mathcal{B}_2 = 3456 \quad \mathcal{B}_1\mathcal{B}_2 = 1256$$
$$\mathcal{B}_3 = 136 \quad \mathcal{B}_1\mathcal{B}_3 = 246 \quad \mathcal{B}_2\mathcal{B}3 = 145 \quad \mathcal{B}_1\mathcal{B}_2\mathcal{B}_3 = 235$$

For the current example, the three generators give rise to eight blocks, each containing eight runs, with a number of three and higher-factor interactions lost, but all main effects and two-factor effects estimable. The experiments contained in the eight blocks will be those giving the following combinations of signs for the three generators: (+ + +), (+ + -), (+ - -), (- - -), (- - +), (- + +), (- + -), (+ - +). If, for some reason, one or more of the lost three-factor interactions is of interest one can easily modify the generators to account for this.

Tables giving possible arrangements and their respective confounded contrasts are available, e.g. in Box and Draper (1987, [4]).

An important question arising when dealing with blocked experiments is whether the estimates of the fixed effects can be made independent from the estimates of the random effects. An experimental plan leading to such a desirable property is said to *block orthogonally*, and the condition for it to be achieved is (see for ex. Khuri and Cornell 1996 [8])

$$\mathbf{Z}^T\mathbf{B} = \mathbf{0} \tag{2.51}$$

CCD designs can be made to block orthogonally in the case of single blocking for specific arrangements of the experiments as shown for ex. in [8], but full orthogonality is not achievable in the general case.

2.4.2 Multiple blocking of the experimental plan

It has been indicated that *orthogonal blocking* of the CCD for a single random factor can be achieved in specific particular cases. The question now arises whether it is possible to design a unique experimental plan that allows estimation of the parameters of the mixed model with *multiple blocking* represented by (2.28) and (2.29) retaining the orthogonality property. In this case each experiment would fall into more than one block, e.g. a block from the *earthquake*-blocking, one for the *concrete ultimate strain*-blocking, and similarly for every other diffused random variation modelled with a random field. Under

these conditions the satisfaction of (2.51) is not guaranteed. There are some particular combinations of number of variables $\mathbf{x}$ and of random factors for which an arrangement of the CCD that blocks orthogonally is known, but there is no general rule. The point is therefore whether loosing the orthogonality property constitutes such an impairing flaw that it compromises the whole procedure

As a partial answer to this question, one might note that loosing the orthogonality property would be more of a problem if a method to estimate the interaction between fixed and random effects were not available. Such a method exists and it consists of a rather straightforward extension of model (2.28).

In Khuri and Cornell (1996, [8]) an extension of the mixed model to include *interaction terms* between the fixed and the random effects is presented. The model can be written as

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\delta} + \sum_{i=1}^{p} \mathbf{U}_i\boldsymbol{\Delta}_i + \boldsymbol{\varepsilon} \tag{2.52}$$

The summation in (2.52) accounts explicitly for the interaction. The matrix product $\mathbf{U}_i\boldsymbol{\Delta}_i$ accounts for the interaction between the i-th fixed effect and all the q random effects. The j-th column of the matrix $\mathbf{U}_i$ is given by the element-wise product of the i-th column of $\mathbf{Z}$ times the j-th column of $\mathbf{B}$. The random vector $\boldsymbol{\Delta}_i^T = [\boldsymbol{\Delta}_{i1}^T \dots \boldsymbol{\Delta}_{ib}^T]$ collects the interaction between all the $q = q_1 + \cdots + q_r$ levels of the r random factors and the i-th fixed effect. This vector is assumed to be normally distributed as $N(\mathbf{0}, \mathbf{I}_q\sigma^2_{\Delta i})$ independently of the other $\boldsymbol{\Delta}$'s as well as of the block effects and the error term.

As this enhanced model gives the possibility to quantify the interaction terms, one may adopt an experimental plan with multiple blocking without worrying about making it orthogonal. The suggested format divides the factorial part into a desired number of blocks, then adding two star points and a centre point to each of the first k blocks in a random way.

The experimental results can then be analysed based on model (2.52), using the complete model and two reduced versions, one without the *fixed-random interaction* effects and one without any random effect at all, i.e. $\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. From the estimates of the regression parameters and of the variance components the effect of lacking the block-orthogonality in the problem can be assessed.

The analysis of model (2.52) requires that the conditional likelihood be rewritten, conditioning on the *fixed-random interaction* terms too. The derivatives of the conditional log-likelihood set to zero give the equations in the unknown parameters

$$
\begin{aligned}
\mathbf{b} &= (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\Big(\mathbf{Y} - \mathbf{Zb} - \mathbf{B}\boldsymbol{\delta} - \sum_{i=1}^{p}\mathbf{U_i}\boldsymbol{\Delta_i}\Big) \\
\boldsymbol{\delta} &= (\mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\delta}\boldsymbol{\delta}} + \mathbf{B}^T\mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}^{-1}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}\Big(\mathbf{Y} - \mathbf{Zb} - \sum_{i=1}^{p}\mathbf{U_i}\boldsymbol{\Delta_i}\Big) \\
\boldsymbol{\Delta_i} &= (\mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\Delta_i}\boldsymbol{\Delta_i}} + \mathbf{U}_i^T\mathbf{C}^{-\mathbf{1}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}\mathbf{U}_i)^{-1}\mathbf{U}_i^T\mathbf{C}^{-1}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}\Big(\mathbf{Y} - \mathbf{Zb} - \mathbf{B}\boldsymbol{\delta} - \sum_{i=1}^{p}\mathbf{U_i}\boldsymbol{\Delta_i}\Big) \quad i = 1, \ldots, p \\
\sigma^2_{\varepsilon} &= \frac{1}{n}\Big(\mathbf{Y} - \mathbf{Zb} - \mathbf{B}\boldsymbol{\delta} - \sum_{i=1}^{p}\mathbf{U_i}\boldsymbol{\Delta_i}\Big)^T\Big(\mathbf{Y} - \mathbf{Zb} - \mathbf{B}\boldsymbol{\delta} - \sum_{i=1}^{p}\mathbf{U_i}\boldsymbol{\Delta_i}\Big) \\
\sigma^2_i &= \frac{1}{q_i}\boldsymbol{\delta}_i^T\boldsymbol{\delta}_i \quad i = 1, \ldots, r \\
\sigma^2_{\Delta_i} &= \frac{1}{q}\boldsymbol{\Delta}_i^T\boldsymbol{\Delta}_i \quad i = 1, \ldots, p
\end{aligned}
\tag{2.53}
$$

The set of equations (2.53) is more complicated to solve than the previous set given by Equations (2.48). The total number of unknown has increased to $p + q + pq + 1 + r + p$.

It is easily recognised that estimating the interaction between parameters, even if we would be only interested in determining the interaction between mean and random factor for one or two mechanical parameters, increases the number of parameters to be estimated from the experiments rapidly. Therefore, to maintain the same level of accuracy, it will be necessary to increase the amount of data, i.e. to perform more experiments. This might be done by performing replicates of the CCD, which means performing the same set of experiments more than once, obviously using different levels of the random factors.

As a final remark, one could observe that in principle the problem of determining a *fixed-random* interaction exists even in the case of a block-orthogonal design. This happens when the randomness of spatially distributed uncertain quantities is modelled, as suggested in Section 2.3, by separating in into a *global* effect, such as for instance their mean or median, described by a random variable and included in $\mathbf{x_1}$, from their *local* random fluctuation described by a random field, discretised and included in $\mathbf{x_2}$. It is intuitively clear that an interaction between fixed and random effects naturally arises when the same physical source of variability is split into the random vectors $\mathbf{x_1}$ and $\mathbf{x_2}$.

Model (2.52) offers a general framework in which interaction from both sources, lack of orthogonality and the above mentioned split in randomness description, can be estimated.

2.5 Reliability analysis

The response surface discussed in this work is used to model the capacity of a structure as a function of a number of variables representing mechanical parameters as well as loading. The capacity is expressed in terms of the *intensity* of the seismic action producing a state of *failure*. The chosen intensity measure is the spectral acceleration $S_{a,f}$ close to the (median) fundamental period of the structure. Failure is assumed to be attained as soon as one of all the possible failure mechanisms occurs. This implicitly accounts for the interaction between failure modes.

In the original formulation the response Y is Gaussian, as δ and ε are Gaussian variables, thus its domain of definition is the interval $[-\infty, \infty]$. Since the spectral acceleration is a positive definite quantity, it is necessary to introduce a simple transformation $Y = \ln(S_{a,f})$ to be consistent with the assumptions on which the response surface model is based.

The limit state function $[G(\mathbf{x})]$ for the structure can now be defined as

$$G(\mathbf{x}) = \ln(S_{a,failure}(\mathbf{x})) - \ln(S_a).$$

where $\ln(S_{a,failure}(\mathbf{x}))$ is the structure's capacity represented by the response surface and S_a is the demand spectral acceleration. FORM or SORM can be employed for parametrically varying values of $\ln(S_a)$ so as to obtain numerically the fragility function for the structure. Beside the basic variables $\mathbf{x}$ describing the randomness in the structure, the other random variables entering in the limit-state function above are: the parameters estimates $\mathbf{b}$, assumed as Gaussian with mean the estimated values at convergence and covariance matrix as described in Section 2.3.3, the random factors effects δ's, Gaussian with zero mean and covariance $\mathbf{C}_{\boldsymbol{\delta\delta}}$ and the error term ε, zero-mean Gaussian variable with variance σ^2_ε.

Chapter 3

Application of methodology to an RC-frame

3.1 Introduction

In this chapter the methodology described in Chapter 2 is applied to a reinforced concrete plane frame, designed solely for gravity loads. Such a structure has been selected to serve as a test case for the method, because it is regarded as more interesting than a frame designed for high lateral loads. The latter, if designed to resist seismic actions according to a modern seismic design philosophy, is supposed to fail in a prescribed failure mechanism, a feature that would not allow to assess the capability of the method of taking into account multiple failure modes and their interaction. The selected frame is shown in Figure 3.1, a 6 storey, 3 bay structure. Storey height is equal for all floors except the ground floor, a typical configuration for buildings with a public function at street level. Spacing between parallel frames in the direction orthogonal to the plane of the frame is 5 m.

The vertical load is taken to be equal to 5 kN per m^2 floor surface. This includes structure self weight, dead loads and part of the live load. The section of the columns is reduced proportional to the acting axial load, resulting in decreasing column dimensions with height. The different sections are indicated in Figure 3.1.

The structure has been designed according to an *allowable stress design* procedure. The concrete used is a class B25 and for the steel a mild steel S400 was chosen.

The columns have been dimensioned for an admissible compression stress of $\tilde{\sigma}_c^- = 4\,\text{MPa}$, and are reinforced with a minimum reinforcement percentage of $A_{s;\text{min}} = 0.8\%$.

The beam height has been estimated according to $h_b = {}^1\!/_{10} \times l_b$. The beam width was chosen in accordance with the column depth. In a following step, beam reinforcement has been designed for an admissible stress of $\tilde{\sigma}_s^+ = 255\,\text{MPa}$. Sectional properties are

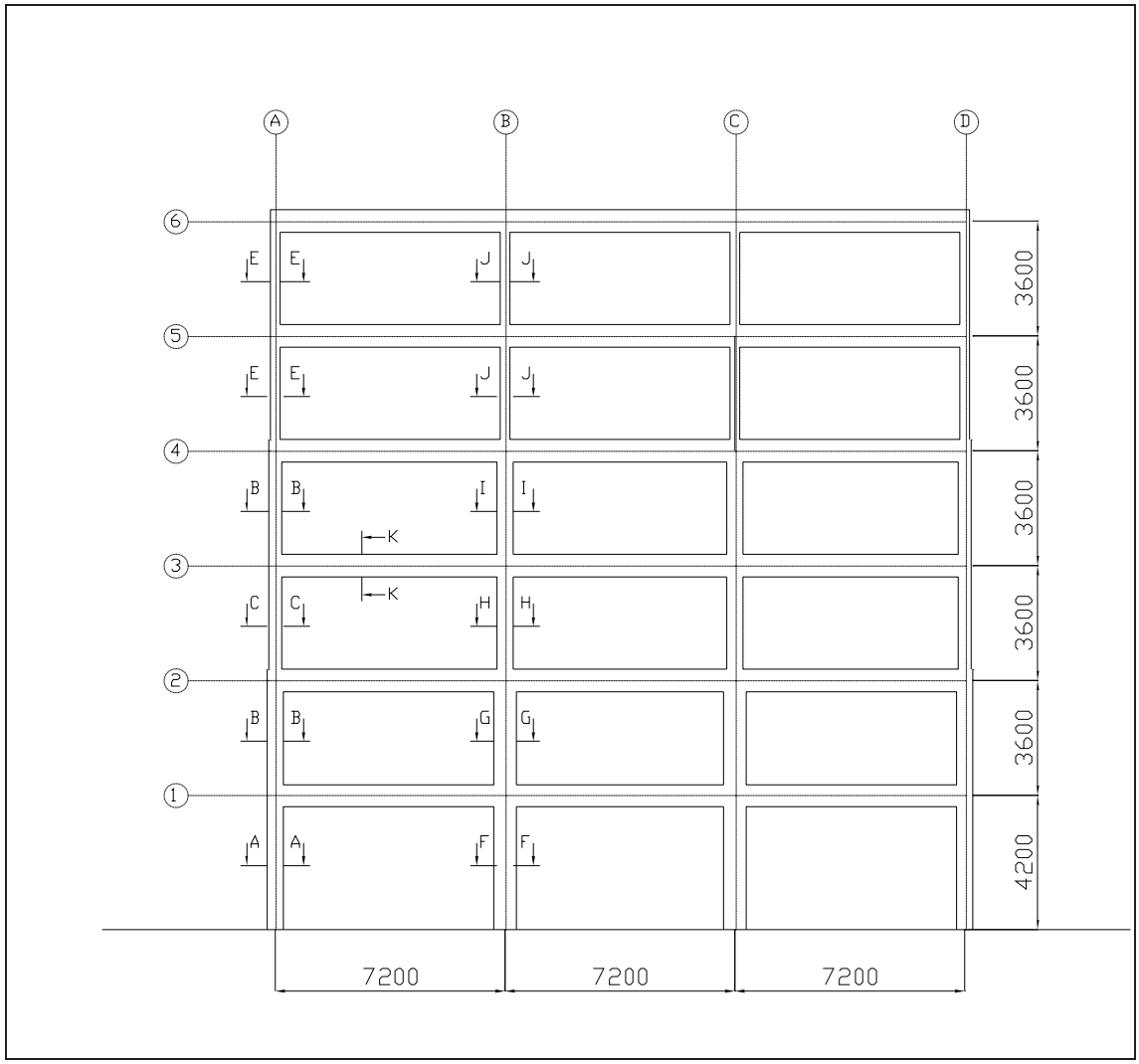

Figure 3.1: 6 storey, 3 bay concrete frame

summarised in Table 3.1 for the different beam and column sections. Concrete cover depth is uniform and equal to 30 *mm*.

3.1.1 Structural model

In order to estimate the parameters of the quadratic response surface a number of *experiments* have to be carried out. In this case, the experiments consist of non-linear numerical time-history analyses, using different earthquake ground-motion records and varying the mechanical parameters. The inelastic dynamic analysis program Ruaumoko (Carr, 1998, [5]) is used for this purpose.

The structural model adopted is of the lumped mass type, with mass concentrated in the

nodes at beam/column joints. These masses will only contribute to the diagonal terms in the mass matrix, as terms associated with rotational degrees of freedom are taken to be zero. The mass, associated with the nodes of Figure 3.1 varies only between the nodes on axes A and D (exterior) and axes B and C (interior) respectively:

	Mass	**Unit**
Exterior	90	[tons]
Interior	180	[tons]

A traditional Rayleigh damping model is used, where damping is proportional to mass and stiffness. The coefficients α and β, multiplying Mass and Stiffness matrices $\mathbf{M}$ and $\mathbf{K}$ respectively, are computed to give a 5% level of viscous damping at the first and second mode of free vibration.

In Ruaumoko columns and beams are both modelled as frame elements. The inelastic behaviour of the frame member is described with the Giberson one-component model (Sharpe, 1974), which is of the concentrated plasticity type and has a possible plastic hinge at both ends of the elastic central length of the member. The hinge stiffness is controlled by the tangent stiffness the appropriate moment-curvature hysteretic rule. The stiffness of the hinge is such that the sum of the rotation of the hinge and the rotation associated with the elastic curvature of the beam over the hinge length is the same as the rotation associated with the curvature over the hinge length with the inelastic properties in the hinge zone.

Several (empirical) expressions are available for the definition of the plastic hinge length. Here the well known expression given in Paulay and Priestley (1992, [9]) is used,

$$l_p = 0.08h_e + 0.022f_y d_{bl} \tag{3.1}$$

where h_e is the distance to the point of contra flexure, f_y the effective steel yield strength and d_{bl} the longitudinal bar diameter. The distance to the point of contra-flexure varies in general as a function of the moment distribution, but is approximately constant and equal to half the member length.

The Modified Takeda hysteresis model (Otani, 1974), widely used for reinforced concrete sections, is used to represent the moment-curvature behaviour in the hinge region of the member. The model parameters, besides elastic stiffness k_0 and bilinear factor r, are ALFA, BETA and NF, defining the unloading and reloading behaviour. For the purpose of these analyses the values of the parameters have been chosen to be equal to: ALFA = 0 (unloading stiffness k_u is equal to initial loading stiffness k_0), BETA = 0 (reloading points towards maximum moment-curvature point reached), NF = 1.0 (recommended value for reloading stiffness power factor).

The elastic stiffness k_0 of the elements is computed according to the cracked section ap-

Section	**b × h** [m×m]	$\mathbf{A_c}$ [m^2]	$\mathbf{I_{g;y}}$ [m^4]	$\mathbf{I_{cr;y}}$ [m^4]	$\mathbf{A_{st,design}}$ [m^2]	$\mathbf{l_p}$ [m]
Exterior						
A	0.35x0.5	0.175	0.003646	0.002188	0.0011	0.33
B	0.35x0.5	0.175	0.003646	0.002188	0.0009	0.30
C	0.35x0.4	0.140	0.001867	0.001120	0.0008	0.30
D	0.35x0.4	0.140	0.001867	0.001120	0.0007	0.30
E	0.35x0.3	0.105	0.0007875	0.0004725	0.0007	0.30
Interior						
F	0.35x0.7	0.245	0.01000	0.006000	0.0022	0.33
G	0.35x0.7	0.245	0.01000	0.006000	0.0018	0.30
H	0.35x0.6	0.210	0.006300	0.003780	0.0015	0.30
I	0.35x0.6	0.210	0.006300	0.003780	0.0011	0.30
J	0.35x0.5	0.175	0.003646	0.002188	0.0007	0.30
Beams						
K	0.3x0.7	0.335	0.01889	0.006613	0.000308/0.000769 top/bottom	0.41

Table 3.1: Section properties

proach. Again following [9], the second moment of inertia is reduced, imposing $I_{cracked} = 0.35I_{gross}$ for T-beams and $I_{cr} = 0.6I_g$ for moderately loaded columns. The bilinear factor r, expressing the post-yield stiffness as a percentage of initial stiffness, is taken to be 0.01. Table 3.1 collects the data evaluated as above.

A beam-column member differs from a beam member in the definition of the yield moment. For beam elements yield is simply defined by a positive ($\mathbf{M}_y^+$) and negative yield moment ($\mathbf{M}_y^-$). Columns, on the other hand, are characterised by a yield surface which takes into account the interaction between the axial force and the moment capacity. This yield surface can be defined by four points, approximating the surface with a tri-linear curve. In Figure 3.2 the characteristic points of the yield surface are shown, of which PYC, (PB,MB), M_0 and PYT are used. Table 3.2 gives values for the different column and beam sections for the *mean* structure (the structure obtained assigning the mean values for all those parameters that are considered random in the subsequent analysis).

The static loads are applied on the beam members, in the form of fixed end forces, a shear (V) and a moment (M), uniform on all beams

Force		**Unit**
V	90	[kN]
M	108	[kNm]

The static analysis is the starting point of the dynamic analysis as this ensures that all

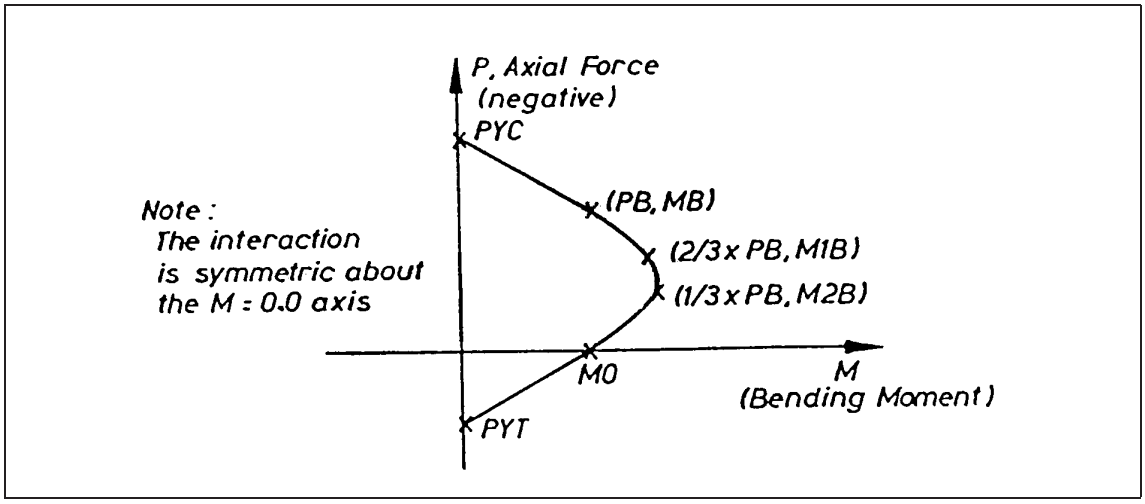

Figure 3.2: Concrete Beam - Column yield surface, ([5])

members start with the correct force distribution before commencing the time-history analysis. In the time history analysis an integration time-step of $dt = 0.002$ (s) has been used.

In Figure B.1 (Appendix B) the input model for Ruaumoko is shown. All beams and columns are subdivided in two elements, mainly for the purpose of the random field mesh.

The dynamic excitation is applied in terms of a ground acceleration time-history applied to the fixed horizontal degrees of freedom (column base).

3.1.2 Choice of basic variables

As discussed before, only a limited number of variables is used as fixed effect variables $\mathbf{x}_1$ for the construction of the response surface. These variables can be subdivided in two categories: variables related to strength/stiffness of the structure and variables related to its geometry. It is convenient to work in the space of basic material properties. For what concerns concrete, if the strength f'_c and the ultimate deformation ε_{cu} are used as basic variables, other properties such as the modulus E_c and the tensile strength f_t should be expressed as functions of these basic properties. For what concerns steel the only property considered as random is the yield strength f_y. Given the geometry, the three chosen material variables fully determine the elements' stiffness and strength. Their mean values are shown in Table 3.3, together with the distribution type and values of the coefficient of variation commonly associated with these mechanical parameters. The values have been taken from [7]. Corresponding sectional properties, evaluated for the mean values of the mechanical variables, are shown in Table 3.2.

The assumed mean values of f'_c and f_y correspond to values frequently encountered in

Section	**PYC** [kN]	**PB** [kN]	**MB** [kNm]	**M0** [kNm]	**PYT** [kN]	**My+** [kNm]	**My-** [kNm]
External							
A	-4815.0	-1645.8	353.5	96.8	440.0		
B	-4735.0	-1645.8	335.9	79.6	360.0		
C	-3820.0	-1295.6	217.7	55.4	320.0		
D	-3780.0	-1295.6	210.9	48.6	280.0		
E	-2905.0	-945.5	124.4	35.0	280.0		
Internal							
F	-7005.0	-2346.2	788.4	275.4	880.0		
G	-6845.0	-2346.2	737.2	226.2	720.0		
H	-5850.0	-1996.0	533.3	160.1	600.0		
I	-5690.0	-1996.0	490.1	118.2	440.0		
J	-4655.0	-1645.8	318.3	62.2	280.0		
Beams							
K						73.5	-183.5

Table 3.2: Section yield properties

	Mean	**Units**	**Type**	**COV**
f'_c	25	[MPa]	LN	0.20
E_c	25000	[MPa]		
ε_{cu}	0.003		LN	0.30
f_y	400	[MPa]	LN	0.10
E_s	200000	[MPa]		

Table 3.3: Material properties

reality. For what concerns the mean of ε_{cu} since the example frame is not designed for seismic forces, and hence the percentage of transverse reinforcement in the members is usually low, confinement in the plastic hinge region is expected to be also low. These considerations have led to the choice of values of $\varepsilon_{cu} = 0.003$.

The vector of the *fixed effect* variables $\mathbf{x}_1$ is completed by the geometric random variables which could have included both the gross concrete section dimensions and the area of the reinforcement. The uncertainty on the former variables has been considered as negligible compared with that of the latter which are also more influential on the behaviour of the structure. To obtain the maximum saving in the number of the variables recourse has been made to a widely known index which is also proposed in regulatory documents related to assessment. This index, called *storey index*, is considered as informative of the type of mechanism, beam-sway or column-sway, likely to form under seismic excitation. It is here used in the following slightly modified form

	Mean	**Type**	**COV**
S_1	0.5875	LN	0.15
S_2	0.8852	LN	0.15
S_3	1.5386	LN	0.15

Table 3.4: Geometric properties

$$S_i = \frac{\sum (A_{s;beam,left} + A_{s;beam,right})}{\sum (A_{s,column,above} + A_{s,column,below})} \tag{3.2}$$

where the summation extends to all beam-column joints of a given floor: the steel areas obviously refer to the steel in tension according to a sway mechanism, e.g., for beams, to the top steel on one side of the joint and the bottom steel on the other. To stay with the maximum practicable number of variables it has been deiced to group the six storey indexes two by two. Table 3.4 gives the properties of the three indices chosen. In their treatment as random variables the three storey indices are considered as uncorrelated variables.

3.1.3 Definition of capacity

In the current investigation only flexural failure in the columns is considered. Different failure modes are therefore not so much related to a different type of failure but rather to the location of the failed section. Failure is thus defined as any of the 48 column sections reaching their curvature capacity. The curvature capacity of a concrete column section with two symmetric layers of steel can be defined in an approximative way using an equivalent stress block, and assuming tension steel to be at yield (see Figure 3.3). For this special case, section translation equilibrium gives:

$$0.85 f_c' b * 0.8 x_u + \frac{\varepsilon_{cu}\,(x_u - c)}{x_u} E_s A_{s;top} - f_y A_{s;bottom} + N = 0 \tag{3.3}$$

from which the neutral axis depth x_u can be obtained. Defining $\alpha = {}^2\!/_3 f_c' b$, $\beta = \varepsilon_{cu} E_s A_{s;top} - f_y A_{s;bottom} + N$ and $\gamma = -\varepsilon_{cu} E_s A_{s;top} c$ and given that ultimate curvature is defined as $\phi_u = \varepsilon_{cu}/x_u$, the ultimate curvature capacity can be written as

$$\phi_u = \frac{2\alpha\varepsilon_{cu}}{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}} \tag{3.4}$$

Equation (3.3) holds as long as the compression steel has not yet yielded. When this is not the case the ultimate curvature in the column becomes simply

$$\phi_u = \frac{\alpha \varepsilon_{cu}}{-N} \tag{3.5}$$

where the curvature capacity is a function of concrete parameters only, since yielded steel areas contributions cancel out.

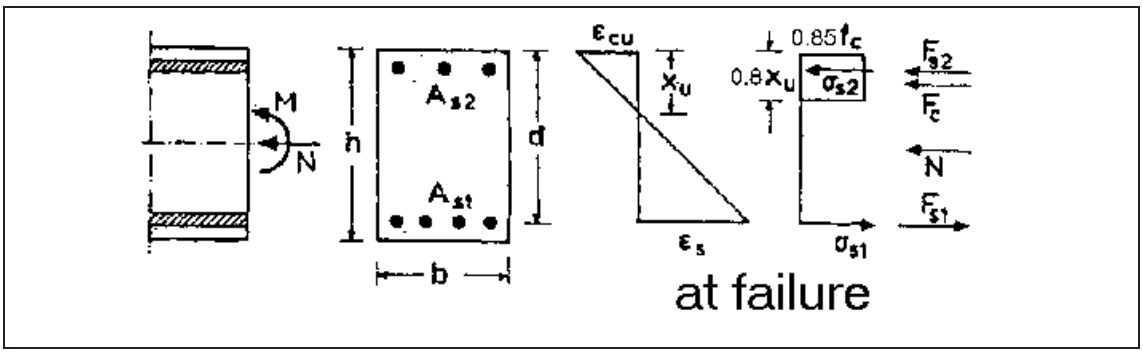

Figure 3.3: Capacity of concrete column

From Equations (3.4) and (3.5) it is apparent that the curvature capacity depends on the axial force. This latter varies in the course of the dynamic response. The critical situation, however, occurs when the axial force reaches its (compressive) maximum since the maximum curvature demand usually occurs at the same time instant. It is therefore sufficient to monitor the maximum value of curvature demand and then compare it with the corresponding capacity for the associated value of axial force[1].

Additional failure mechanisms such as joint tension or compression failure, shear failure and weak storey failure have not been considered in this application, though their inclusion would not add further significant computational demand.

3.1.4 Statistical model

A second order polynomial is now developed on the six chosen variables. The model is described in (3.6): based on judgement, all linear and quadratic terms have been included, as well as the two-way interaction terms with the exception of those between material parameters and storey indices; all the higher order terms have been neglected. Two random factors have been included: δ_{eq} is the effect of the random factor representing the earthquake, while $\delta_{\varepsilon_{cu}}$ is the effect of the random factor that accounts for the spatial variability of the ultimate concrete strain (ε_{cu}).

[1]This is fortunate, as most of the structural analysis programs do not routinely print curvature or strain at every time-step in the standard output file and, even if they would, the time-consumption associated with post-processing operations would become too onerous.

$$\ln(S_{a;failure}) = \beta_0 + \beta_1 \cdot f'_c + \beta_2 \cdot \varepsilon_{cu} + \beta_3 \cdot f_y + \beta_4 \cdot S_1 + \beta_5 \cdot S_2 + \beta_6 \cdot S_3 + \beta_7 \cdot f'^2_c + \beta_8 \cdot \varepsilon^2_{cu} + \beta_9 \cdot f^2_y + \beta_{10} \cdot S^2_1 + \beta_{11} \cdot S^2_2 + \beta_{12} \cdot S^2_3 + \beta_{13} \cdot f'_c \cdot \varepsilon_{cu} + \beta_{14} \cdot f'_c \cdot f_y + \beta_{15} \cdot \varepsilon_{cu} \cdot f_y + \beta_{16} \cdot S_1 \cdot S_2 + \beta_{17} \cdot S_1 \cdot S_3 + \beta_{18} \cdot S_2 \cdot S_3 + \delta_{eq} + \delta_{\varepsilon_{cu}} + \varepsilon \quad (3.6)$$

In the following section the seismic fragility of the frame is first computed with reference to an incomplete model with the first random factor only, i.e. taking into account the randomness due to earthquake loading only. For this case a number of sensitivity analyses have been performed that will be discussed in section **??**. Finally, in section 3.3, the experiments are performed for the full model as described in Equation (3.6).

3.2 Single blocking

A first set of experiments has been run, including only the earthquake random factor in the model. As anticipated in Chapter 2, the experimental plan is the central composite design, suitably subdivided in blocks. To estimate the distribution of the random factor effect, with which the influence on the response to ground-motion variability is represented, one needs a number of random factor *levels* to be associated with the blocks in which the CCD is subdivided. In the next sections the following issues are discussed: the choice of the accelerograms representing the earthquake action; the effect on the fragility of considering the regression parameters **b** either as random or deterministic; the significance of the terms in model (3.6) in view of its possible simplification; the sensitivity of the fragility to the location of the centre of the experimental plan, to alternative choices of sets of accelerograms and, finally, validation of the obtained fragility through Monte Carlo simulation.

3.2.1 Selection of earthquake accelerograms

Among possible alternatives the decision has been taken to use in this work recorded accelerograms. A well recognised problem connected with the choice of natural accelerograms is the admissibility of modifying their intensity if required by the analysis. This is actually the case in the present study where the dynamic analyses need to be run for varying intensities until a state of failure is attained. This scaling issue, whose importance in risk studies is unquestionable, is still subject to debate and by its very nature does easily lend itself to rigorous treatment. Reference is made to a recent comprehensive research whose results are summarised, for instance, in [14], and of which, for the sake of the present study, it is sufficient to retain the following concluding remarks. Conditional on the selection as intensity parameter of the spectral ordinate close to the fundamental period of the structure and if the response of interest is a peak quantity, scaling up or down this parameter by factors as large as 30 does not introduce bias in the results. This

is shown by selecting sets of recorded accelerograms belonging to widely different ranges of both magnitude and distance. The accelerograms of each set are then scaled to a common S_a value and the statistics of the response of multi-degree of freedom structures are determined from each set separately and from the larger set that collects all the accelerograms: the results are approximately the same. This amounts to say that, within the limits of the considered data set, peak responses are practically independent of the magnitude and distance characterising the record used for the analysis.

Based on the preceding considerations, in the current application records for rock and stiff soil-types, with a source-to-site distance between 25 and 75 kilometres and a magnitude ranging from 5.5 to 7.5 are used. Table A.1 in Appendix A shows the earthquake strong-motion records used[2], with their specific properties. The 16 records selected are split up in two groups of 8 records each, Set I and Set II, to be later used in the sensitivity analysis. Their acceleration response spectra are shown in Figures 3.4 and 3.5.

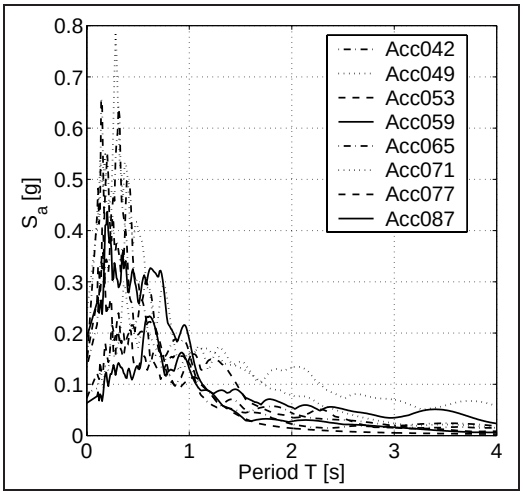

Figure 3.4: Earthquake spectra, Set I

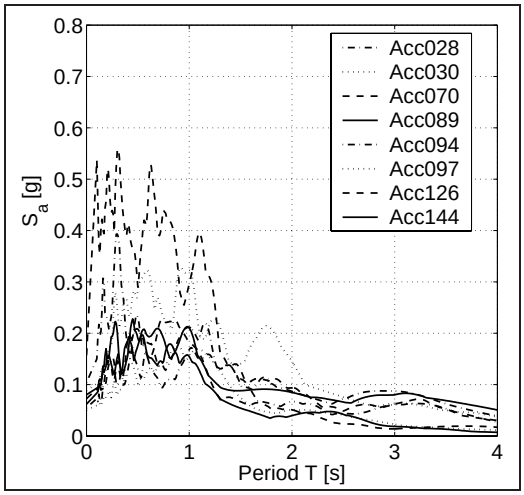

Figure 3.5: Earthquake spectra, Set II

3.2.2 Statistical analysis of the numerical experiments

For the current case study, with 6 input variables, the CCD requires performing $n = 2^6 + 2 \cdot 6 + 6 = 82$ experiments. These 82 experiments are subdivided in 8 blocks consistently with the choice of records. The factorial part of the design is grouped using block generators $\mathcal{B}_1 = 135$, $\mathcal{B}_2 = 1256$ and $\mathcal{B}_3 = 1234$, that confound the effects 135, 1256, 1234, 236, 245, 3456 and 146, all three and four-way interaction terms that are not used in the model. For the star part of the design it has been decided to form six groups of three points, each group collecting the two star levels of a basic variable and a centre point, and to assign them to the first 6 blocks of the factorial design. This results in 6 blocks with 11 experiments and 2 blocks with 8 experiments.

[2]All ground-motion records used in this study are taken from the PEER database, website: http:\\peer.berkeley.edu\smcat

b	EQ Set I	EQ Set II	Shifted	Reduced
intercept	-1.539	-6.43	-6.47	-3.37
f_c'	0.0678	0.200	0.1159	0.0932
ε_{cu}	0.743	0.806	1.017	0.928
f_y	$-2.27\cdot10^{-3}$	$-1.960\cdot10^{-4}$	0.01575	
S_1	-0.917	2.94	-3.93	
S_2	-2.12	2.44	7.70	
S_3	0.773	0.1828	-3.14	
$f_c'^2$	$-8.03\cdot10^{-4}$	$-3.29\cdot10^{-3}$	$-2.30\cdot10^{-3}$	$-7.84\cdot10^{-4}$
ε_{cu}^2	-0.0846	-0.1006	-0.1081	-0.085
f_y^2	$-1.392\cdot10^{-6}$	$-5.49\cdot10^{-6}$	$-2.50\cdot10^{-5}$	
S_1^2	0.921	0.652	3.48	
S_2^2	0.438	-0.820	-4.45	
S_3^2	-0.214	0.220	1.110	
$f_c' \cdot \varepsilon_{cu}$	$-6.47\cdot10^{-3}$	-0.01050	$-6.99\cdot10^{-3}$	$-6.64\cdot10^{-3}$
$f_c' \cdot f_y$	$6.49\cdot10^{-5}$	$8.76\cdot10^{-5}$	$1.375\cdot10^{-4}$	
$\varepsilon_{cu} \cdot f_y$	$4.48\cdot10^{-4}$	$7.74\cdot10^{-4}$	$1.645\cdot10^{-4}$	
$S_1 \cdot S_2$	1.229	-1.757	0.444	
$S_1 \cdot S_3$	-0.828	-1.314	-0.503	
$S_2 \cdot S_3$	0.344	-0.1052	-0.131	
σ_ε	0.1087	0.1088	0.1575	0.1161
$\sigma_{\delta_{eq}}$	0.3105	0.2041	0.2569	0.3151

Table 3.5: Regression coefficients

The results of the numerical analyses are summarised in Table B.1 in Appendix B. From the regression analysis the estimates of the parameters $\boldsymbol{\beta}$ and the variances of the random factor effect δ_{eq} and of the error term ε are calculated. Results are given in the second column of Table 3.5. The 8 block effects $\delta_{eq,i}$ are collected in Table 3.6.

A partial plot of the response surface in terms of the two (dominant) variables f_c' and ε_{cu} is shown in Figure 3.6. All other basic variables are kept at their respective mean values.

3.2.3 Fragility analysis

The determined response surface is used as the capacity in the limit state function and repeated FORM and SORM analyses are performed to construct the fragility curve. The SORM procedure used to (quadratically) approximate the limit state function is of the curvature-fitting type, i.e. it uses the main curvatures at the *design point*. Fragility curves from FORM and SORM analysis are shown in Figures 3.7 and 3.8. The fragility will be illustrated by two figures routinely. The second using a logarithmic scale for the vertical axis, to show differences for the low probabilities of failure more clearly. The figures

Block	**EQ 1**	**EQ 2**	**Shift**	**Reduced**
B_1	-0.1773	-0.2526	-0.1956	-0.1832
B_2	-0.3604	-0.2288	-0.3015	-0.3601
B_3	0.2301	0.0433	0.1723	0.2264
B_4	-0.1638	0.2035	-0.0939	-0.1538
B_5	0.3529	-0.0574	0.3331	0.3631
B_6	-0.4755	0.1242	-0.3770	-0.4919
B_7	0.3362	0.3496	0.2401	0.3355
B_8	0.2578	-0.1817	0.2225	0.2640

Table 3.6: Block effects

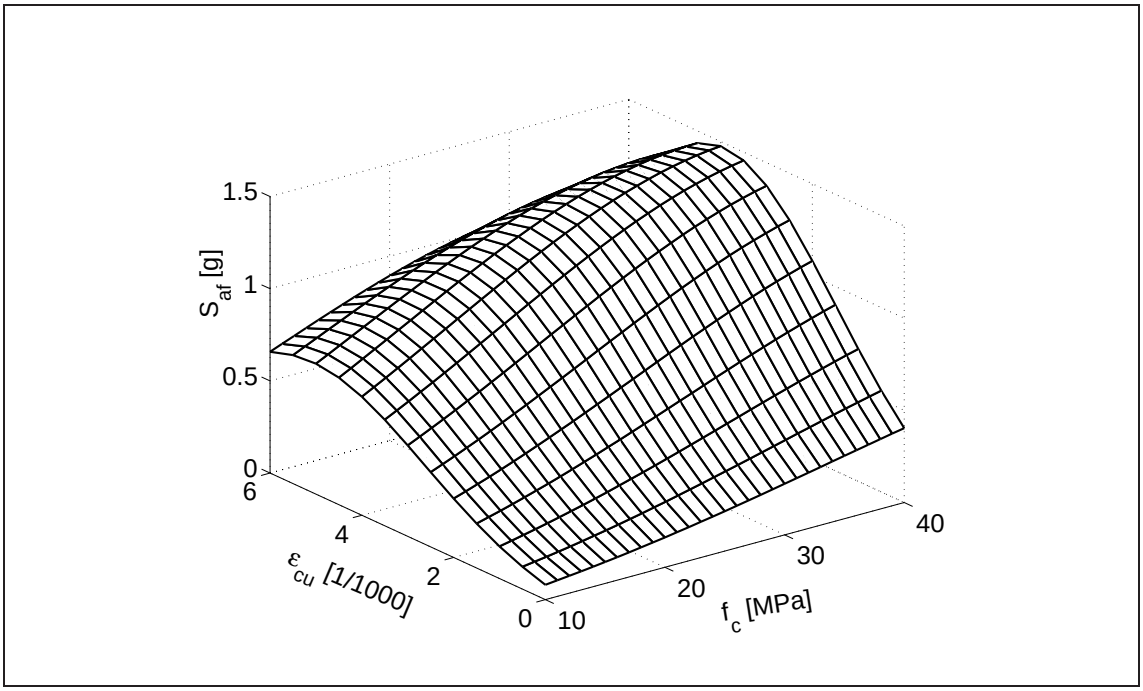

Figure 3.6: Response surface in terms of ε_{cu} and f'_c (all other basic variables kept at their mean)

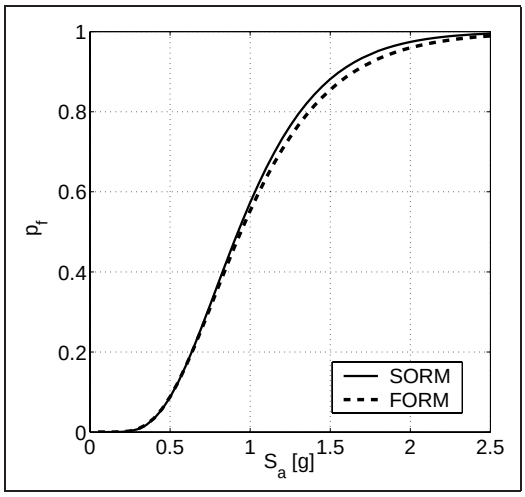

Figure 3.7: Fragility functions: FORM vs. SORM

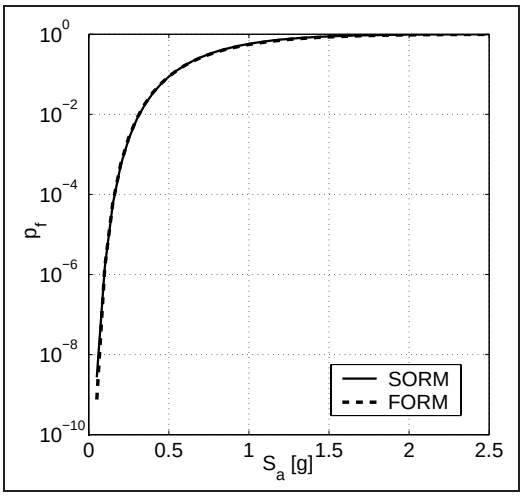

Figure 3.8: Fragility functions: FORM vs. SORM, logarithmic scale

show that, in this case, there is little difference between FORM and SORM results, which indicates that the curvature of the failure boundary is not very pronounced, at least close to the design point.

3.2.4 Sensitivity to the randomness of the regression coefficients

FORM and SORM analyses have been run considering the maximum likelihood estimates $\mathbf{b}$ of the model parameters both as random, with covariance given in (2.50), and as deterministic quantities. The resulting fragilities are shown in Figures 3.9 and 3.10 from which it is apparent that, in this particular application, the model uncertainty has a negligible effect.

3.2.5 Significance of the terms in the model

Economy in the number of the terms in the model is obviously desirable since it allows a more reliable estimate of the truly significant ones. A preliminary selection of explanatory functions assumed as meaningful is usually made based on judgment. Statistical tests of significance can then be used to assess the relative importance of the model terms. Terms whose inclusion in the model does not lead to a significant increase in explanatory power can be removed from the model. The parameters of the reduced model can then be estimated more accurately on the same statistical basis.

A possible way of proceeding is to conduct tests of hypothesis on each of the b_i or some linear combination of them. It might be of interest to construct a joint test involving $m \leq p$ of the model coefficients, for which the hypotheses read as given in Equation (3.7).

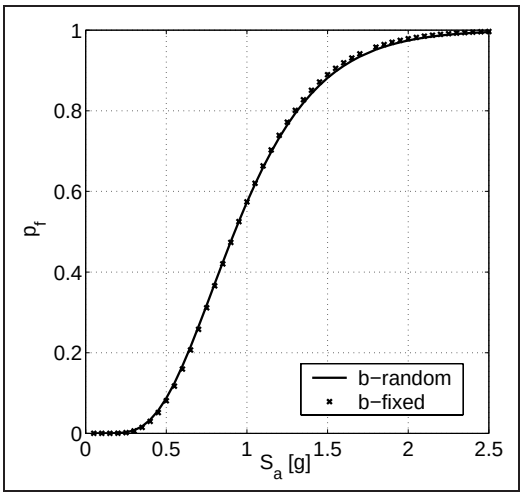

Figure 3.9: Fragility functions: **b** random vs. **b** deterministic

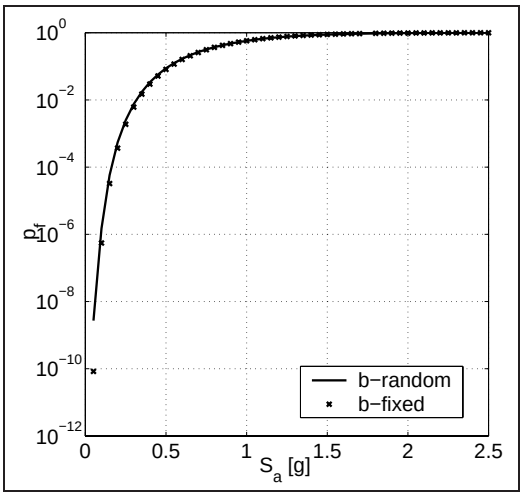

Figure 3.10: Fragility functions: **b** random vs. **b** deterministic, logarithmic scale

$$\begin{aligned} H_0 &: \beta_i = \beta_i^* \quad i = 1, \ldots, m \leq p \\ H_1 &: \beta_i \neq \beta_i^* \quad \text{for at least one } i \end{aligned} \tag{3.7}$$

where β_i^* has a selected value, not equal to the one given by the regression.

Now, let Ω denote the complete model and ω a simplified model, in which the parameters of the model satisfy the constraints as specified in the null hypothesis, H_0. For example, as complete and simplified models might be chosen

$$\begin{aligned} \Omega &\rightarrow Y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + \varepsilon \\ \omega &\rightarrow Y = b_0 + \varepsilon \end{aligned} \tag{3.8}$$

The hypothesis tested comparing these two models is that only the intercept, b_0, matters. The residual sum of squares (RSS) is computed from the regression results on the complete and the reduced models.

$$\begin{aligned} RSS_\Omega &= \sum (Y_i - b_0 + b_1 x_{1i} + b_2 x_{2i} + b_{12} x_{1i} x_{2i})^2 \\ RSS_\omega &= \sum (Y_i - b_0)^2 \end{aligned} \tag{3.9}$$

The ANOVA table for these two models can be written as

	Sum of Squares	**dof**
Intercept	TSS-RSS$_\omega$	$n-(n-1)=1$
Additional variables	RSS$_\omega$- RSS$_\Omega$	$(n-1)-(n-p)=p-1$
Residual	RSS$_\Omega$	$n-p$
Total	TSS	n

b	**EQ 1**
$intercept$	0.318
f'_c	2.31
ε_{cu}	10.06
f_y	0.0593
S_1	0.1084
S_2	1.192
S_3	0.473
f'^2_c	1.918
ε^2_{cu}	22.51
f^2_y	0.0337

b	**EQ 1**
S^2_1	0.251
S^2_2	0.226
S^2_3	0.566
$f'_c \cdot \varepsilon_{cu}$	2.99
$f'_c \cdot f_y$	0.537
$\varepsilon_{cu} \cdot f_y$	0.947
$S_1 \cdot S_2$	0.754
$S_1 \cdot S_3$	1.033
$S_2 \cdot S_3$	0.404

Table 3.7: F-test results

where TSS= $\sum(Y_i - E[Y])^2$ is the total sum of squares, TSS-RSS$_\omega$ is the variation in Y explained by the intercept, RSS$_\omega$-RSS$_\Omega$ is the variation in Y explained by the m additional variables in the complete model, with respect to the reduced one and RSS$_\Omega$ is the residual sum of squares for the full model.

To compare the two models one can use the following test statistic

$$F_{p-1,n-p} = \frac{\frac{RSS_\omega - RSS_\Omega}{p-1}}{\frac{RSS_\Omega}{n-p}} \tag{3.10}$$

Notice that the degrees of freedom in the numerator are the number of parameters set to zero to obtain the reduced model. If the observed test statistic value is larger than an appropriate critical one, it must be concluded that one, or some, or all of the b_i for $i = 2, \ldots, p$ are not zero, i.e. they should not be eliminated from the model.

For the present application 19 tests are performed with test statistic distributed as $F_{1,(n-p-q-r-1)}$: each test aims at assessing the importance of one of the 19 fixed-effect terms in the model. The degrees of freedom in the denominator correspond to the number of experiments n minus the number of parameters estimated from the data: p coefficients β, q coefficients δ, r variances σ_δ and the error variance σ_ε. The results for these tests are summarised in Table 3.7.

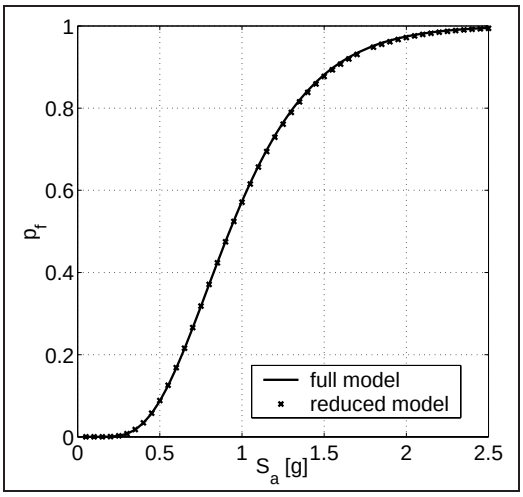

Figure 3.11: Fragility function: Complete model vs. reduced model

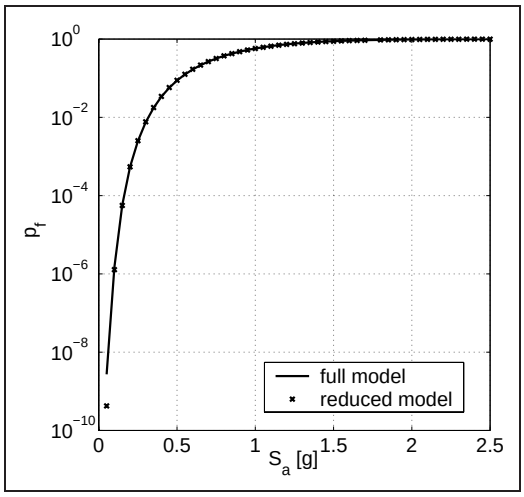

Figure 3.12: Fragility function: Complete model vs. reduced model: logarithmic scale

These results suggest that variables f_c' and ε_{cu} contribute much more significantly in explaining the response than the other four basic variables. If it is decided that only those parameters should be kept in the model that are significant at the 85% confidence level, the corresponding critical value for the test statistic (the decision parameter) is equal to $F(0.85, 1, 53) = 1.68$. This leads to use a reduced model that includes only the two above mentioned parameters, in their linear, quadratic and interaction terms.

The greatly simplified model is given in Equation (3.11), for which the estimated parameters can be found in the fifth column of Tables 3.5 and 3.6.

$$\begin{aligned}\ln(S_{a;fail}) = \beta_0 + \beta_1 \cdot f_c' + \beta_2 \cdot \varepsilon_{cu} + \\ \beta_3 \cdot f_c'^2 + \beta_4 \cdot \varepsilon_{cu}^2 + \beta_5 \cdot f_c' \cdot \varepsilon_{cu} + \delta_{eq} + \varepsilon\end{aligned} \tag{3.11}$$

It is clear from the results that reducing the number of parameters in the model changes the estimates of the coefficients. The error term, and thus the lack-of-fit of the model, has increased as well, even though only slightly. Introducing this simplified model into the limit state function gives the fragility function shown in Figures 3.11 and 3.12. The new fragility almost coincides with that from the complete model, from which it can be concluded that, in this specific case, the random variables f_y, S_1, S_2, and S_3 are not influential in explaining the response variation.

3.2.6 Sensitivity to the location of the experimental design centre

In time-invariant problems, response surface methodologies are commonly used to find an empirical relationship for a structural response quantity entering a limit-state. As the

response surface is most accurate close to the centre of design, the value of the input variables that represents the zero level, it is apparent that, in order to increase accuracy on the failure probability evaluation, the response surface should be centred around the design point. This requires constructing the surface iteratively, for instance centring the experimental design to establish a new surface at the design point of the previous surface/iteration. A review of the techniques used is contained in [12].

In the time-variant problem, where the dynamic action is characterised by such a high variability, the problem of optimising the location of the centre point is fortunately less relevant. This is because in the limit-state function the variable δ_{eq} that accounts for the earthquake effect dominates over the "structural" ones, whose design point values are not far from their respective means.

As a first insight on the approximation involved in using a fixed centre of design (a "first iteration" response surface), an alternative design plan has been used in which the position of the design centre along the concrete ultimate strain axis is shifted half a standard deviation downwards.

The design-points obtained from FORM analysis with the basic experimental design are given in Table B.2 in Appendix B as a function of the spectral acceleration S_a. It can be observed that for the variables f_y, S_1, S_2 and S_3 the design-point values are practically constant and remain close to their respective mean. The other four variables show a larger variation, although their average lies close to their mean. The results of the analyses for the "shifted" experimental plan are given in Table B.3 in Appendix B. Statistical analysis of these results leads to the parameters reported in the fourth column of Tables 3.5 and 3.6. Introducing the latter in the limit state function yields the fragility function shown in Figures 3.13 and 3.14, where the original fragility is also plotted for ease of comparison. It can be seen that the results are not very different in the two cases. The above conclusion is encouraging, although much more work would be needed in order to extend the validity of this provisional statement, which could only be done by comparing the fragility obtained by use of a fixed centre of design with that obtained through the mentioned iterative fitting performed at each value of S_a.

3.2.7 Sensitivity to choice of earthquake records

As discussed before, in Section 3.2.1, two sets of accelerograms have been selected in order to study the stability of the response when using a reduced number of randomly selected records. Results of the dynamic analyses performed using Set II records are given in Table B.4 in Appendix B. The parameters found through the statistical analysis are given in the third column of Tables 3.5 and 3.6. It is noted that even though the average value of the capacity $S_{a,f}$, as well as its maximum and minimum values, are very close for both sets of accelerograms (compare the bottom lines in Table B.4 with B.1) the estimated model parameters exhibit a larger variation. The model accuracy, expressed by σ_ε, remains the same.

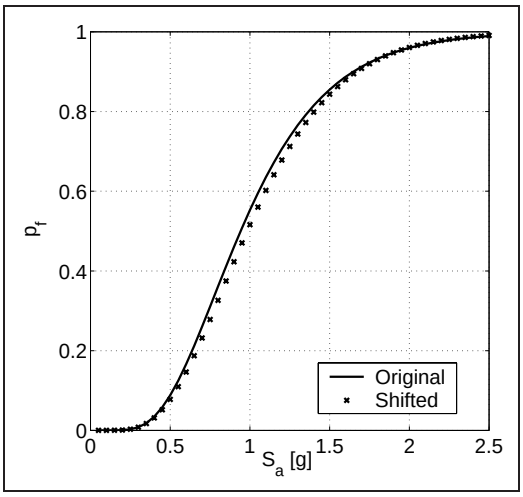

Figure 3.13: Fragility functions for different centres of design

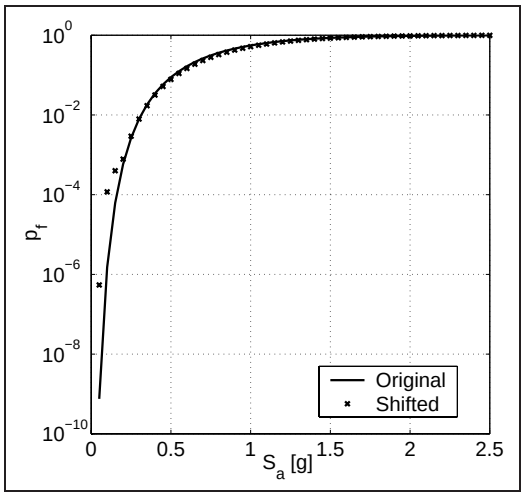

Figure 3.14: Fragility functions for different centres of design: logarithmic scale

	Risk
Set I	0.0053
Set II	0.0041
Shifted centre of design	0.0050

Table 3.8: Unconditional risk computed using the hazard function in Figure 3.17

The fragility functions for the two sets are shown in Figures 3.15 and 3.16. It can be observed that the two functions differ to some extent. To judge the significance of this difference the comparison, however, has to be pursued further up to the level of the unconditional seismic risk, which represents the ultimate measure of safety of the structure. Computation of the risk requires convolution of the fragility with the hazard. This has been done using the hazard function shown in Figure 3.17: the computed risk for the two sets is reported in Table 3.8. It is remarkable to note how the two values are, for all practical purposes, essentially coincident.

3.2.8 Validation by Monte Carlo simulation

A sample check of the accuracy of the results obtained for this single-blocking case has been done using Monte Carlo (MC) simulation. The simulations have been performed for two values of the spectral acceleration. Values of the mechanical properties have been sampled from their respective distributions and have been associated with a record randomly selected from the same set of eight earthquakes (Set I) used to establish the surface and scaled to the appropriate S_a. Any difference emerging from the comparison between MC and RS results is necessarily due to either the response surface approximation

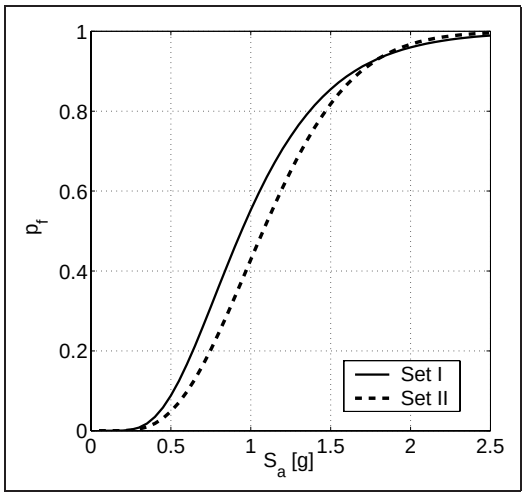

Figure 3.15: Fragility functions for different choice of earthquake records

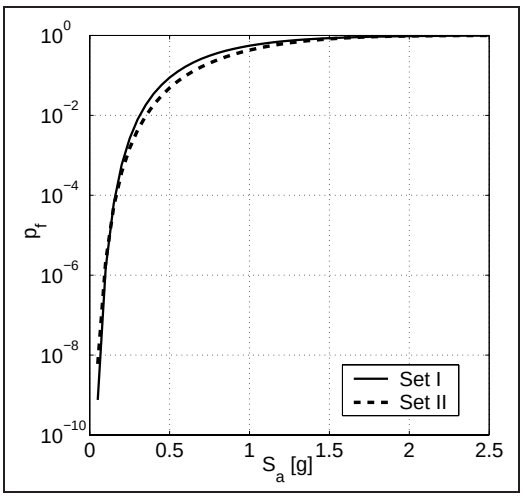

Figure 3.16: Fragility functions for different choice of earthquake records: logarithmic scale

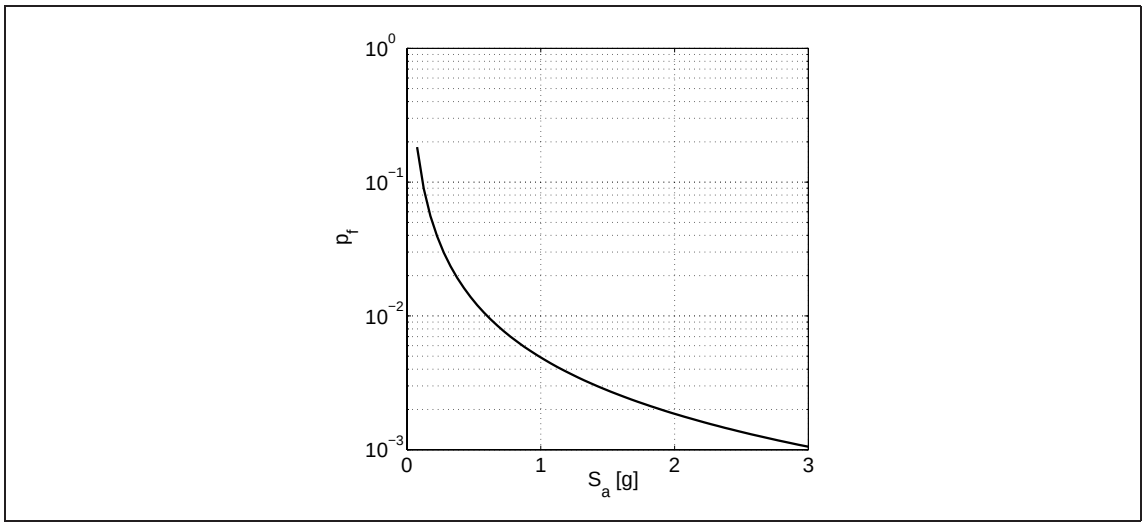

Figure 3.17: Hazard function

or the SORM one.

Figure 3.18 shows that for the two points examined the agreement is excellent.

The MC simulation provides also a measure of the differences in computational effort. The two points obtained by MC required 1297 dynamic analyses for a COV of 0.05, as opposed to the 406 performed in constructing the response surface from which the 50-points fragility function has been computed.

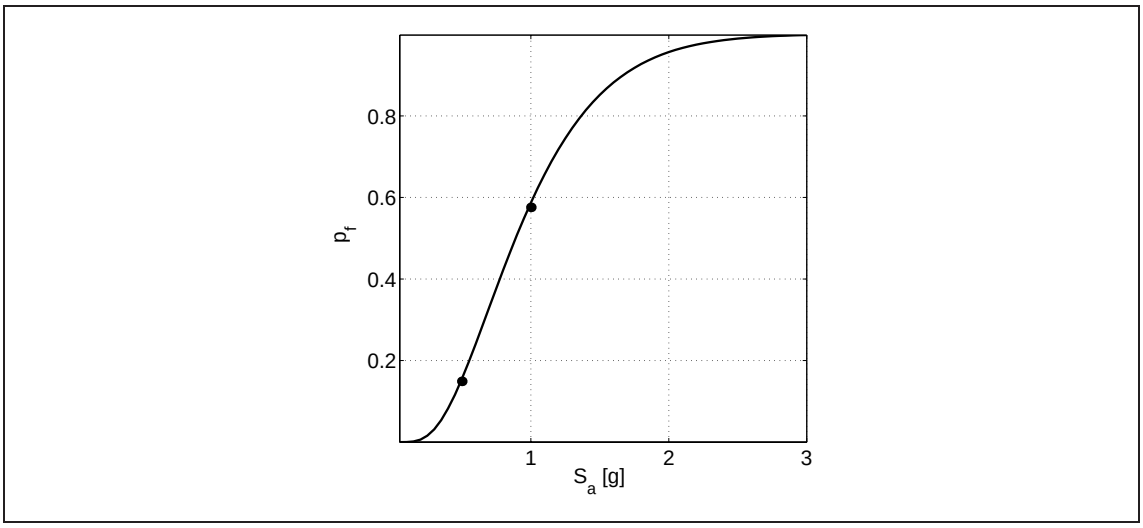

Figure 3.18: Validation of results using MC simulation

3.3 Multiple blocking

3.3.1 Introduction of spatial variability

In many cases the variability of mechanical parameters throughout the structure can be described as the product of two terms, the first one representing the spatial average value of the parameter and the second representing the local random fluctuation about the average. Uncertainty usually exists on the average due to incomplete information which leads to describe this quantity as a random variable. The fluctuation, on the other hand, can be effectively described by means of a unit mean random field.

It is obvious that introduction of these local variations is only worthwhile for those variables that have an important contribution to the response at failure. In this application, it has been shown that ε_{cu} is one of the parameters governing failure, hence the random field model has been used for this quantity. At any given location in the structure the model reads as

$$\varepsilon_{cu} = \bar{\varepsilon}_{cu} \times \tilde{\varepsilon}_{cu} \tag{3.12}$$

where the mean $\bar{\varepsilon}_{cu}$ is represented, as before, by a random variable in $\mathbf{x}_1$, while the fluctuation $\tilde{\varepsilon}_{cu}$ is represented by a unit median random field discretised into a large set of random variables included in $\mathbf{x}_2$.

The levels of the random factor to be used for the experiments are realisations of the random field discretised at as many locations as required by the finite element mesh. The product of the general mean and of a level of the random factor then gives the input value for ε_{cu} to be entered in the finite element model.

For the single-blocking case the ultimate concrete strain has been modelled as a single Lognormal random variable. This choice is maintained for the product in (3.12) which implies that both factors are Lognormal. It remains to decide how to distribute the total variability among the two factors. The coefficient of variation to be used for the random field can be computed once values for the coefficients of variation of the product and of the mean in (3.12) are chosen. For the product of Lognormal variables, see e.g. Benjamin and Cornell (1970, [3]), the following relations hold for the logarithmic mean and standard deviation, respectively

$$\lambda = \bar{\lambda} + \tilde{\lambda} \quad \text{and} \quad \zeta = \bar{\zeta} + \tilde{\zeta} \tag{3.13}$$

where $\lambda = \ln \mu - 0.5\zeta$ and $\zeta = \ln(1 + \delta^2)$. If we set the conditions that $\mu = \bar{\mu} = 0.003$, and recall that $\delta = 0.30$ and assume that $\bar{\delta} = 0.20$ the mean and standard deviation for the random field can be calculated using Equations (3.13) and the expressions for λ and ζ. The expression on the left-hand-side in Equations (3.13) gives

$$\ln \tilde{\mu} - 0.5\tilde{\zeta} = (\ln \mu - 0.5\zeta) - (\ln \bar{\mu} - 0.5\bar{\zeta}) \tag{3.14}$$

Now substituting the right-hand expression of (3.13)

$$\ln \tilde{\mu} = 0.5\bar{\zeta} - 0.5\zeta + 0.5\tilde{\zeta} = 0.5\bar{\zeta} - 0.5\zeta + 0.5(\zeta - \bar{\zeta}) = 0 \tag{3.15}$$

Similarly, starting from the expression on the right-hand side one can find that

$$\ln(1 + \tilde{\delta}^2) = \ln(1 + \delta^2) - \ln(1 + \bar{\delta}^2) \tag{3.16}$$

$$\tilde{\delta}^2 = \frac{\tilde{\sigma}^2}{\tilde{\mu}^2} = \tilde{\sigma}^2_{(\tilde{\mu}=1)} = \left(\frac{1 + \delta^2}{1 + \bar{\delta}^2}\right) - 1 = \frac{\delta^2 - \bar{\delta}^2}{1 + \bar{\delta}^2} \cong \delta^2 - \bar{\delta}^2 \tag{3.17}$$

b	**SB**	**DB**
intercept	-1.539	-4.71
f'_c	0.0678	0.0435
ε_{cu}	0.743	1.112
f_y	-2.27·10^{-3}	4.62·10^{-3}
S_1	-0.917	3.07
S_2	-2.12	-1.3982
S_3	0.773	0.873
f'^2_c	-8.03 ·10^{-4}	-4.49·10^{-4}
ε^2_{cu}	-0.0846	-0.0821
f^2_y	-1.302·10^{-6}	-9.26·10^{-6}
S^2_1	0.921	-2.79
S^2_2	0.438	0.1268
S^2_3	-0.214	-0.339
$f'_c \cdot \varepsilon_{cu}$	-6.47·10^{-3}	-8.41·10^{-3}
$f'_c \cdot f_y$	6.49·10^{-5}	1.078·10^{-4}
$\varepsilon_{cu} \cdot f_y$	4.48·10^{-4}	-2.48·10^{-4}
$S_1 \cdot S_2$	1.229	0.789
$S_1 \cdot S_3$	-0.828	-0.540
$S_2 \cdot S_3$	0.344	0.407
σ_ε	0.1087	0.1027
$\sigma_{\delta_{eq}}$	0.3105	0.2979
$\sigma_{\delta_{\varepsilon_{cu}}}$		0.0698

Table 3.9: Regression coefficients for single (SB)and double blocking (DB)

where the approximation in Equation (3.17) holds for $\bar{\delta} \leq 0.20$. For $\delta = 0.30$ and $\bar{\delta} = 0.20$ it can be found that the standard deviation of the fluctuation field must be $\tilde{\sigma} = 0.22$.

3.3.2 Statistical analysis of the numerical experiments

For the case of multiple blocking the CCD requires performing the same 82 experiments as for the case of single blocking. The design was already broken down in 8 blocks for the random factor representing the earthquake, and for simplicity the second random factor, representing the spatial variability of ε_{cu}, will be treated similarly. The factorial part of the design is thus again grouped in 8 blocks, now using block generators $\mathcal{B}_1 = 346$, $\mathcal{B}_2 = 1345$ and $\mathcal{B}_3 = 1236$, confounding 346, 1345, 1236, 156, 124, 2456 and 235. The star part is now assigned randomly to the first 6 blocks of the factorial design, still adding two star points and a centre point to each block.

Performing the experiments, subdivided in 8 blocks for the earthquake loading, and in 8 different blocks for the spatial variability gives the results summarised in Table B.5 in

Block	SB	DB	Block	DB
B_{11}	-0.1773	-0.2773	B_{21}	-0.1704
B_{12}	-0.3604	-0.2874	B_{22}	-0.0131
B_{13}	0.2301	0.1828	B_{23}	-0.0059
B_{14}	-0.1638	-0.1433	B_{24}	0.0364
B_{15}	0.3529	0.3605	B_{25}	0.0574
B_{16}	-0.4755	-0.4309	B_{26}	0.0496
B_{17}	0.3362	0.3398	B_{27}	-0.0056
B_{18}	0.2578	0.2559	B_{28}	0.0516

Table 3.10: Block effects for single and double blocking

Appendix B. Statistical analysis of these results yields the estimates for the parameters $\boldsymbol{\beta}$ and for the distribution of the random factors effects δ_{eq} and $\delta_{\varepsilon_{cu}}$, and of the error ε, given in Tables 3.9 and 3.10. In these tables previous results for the case of single blocking are reported for the ease of comparison. The difference in the parameters estimates does not lead to appreciable change in the response surface, shown in Figure 3.19, and consequently in the fragility function shown in Figures 3.20 and 3.21. The small increase in the fragility values is essentially due to the presence of the variance of the additional random factor which is however small. The variances of δ_{eq} and ε are practically unchanged.

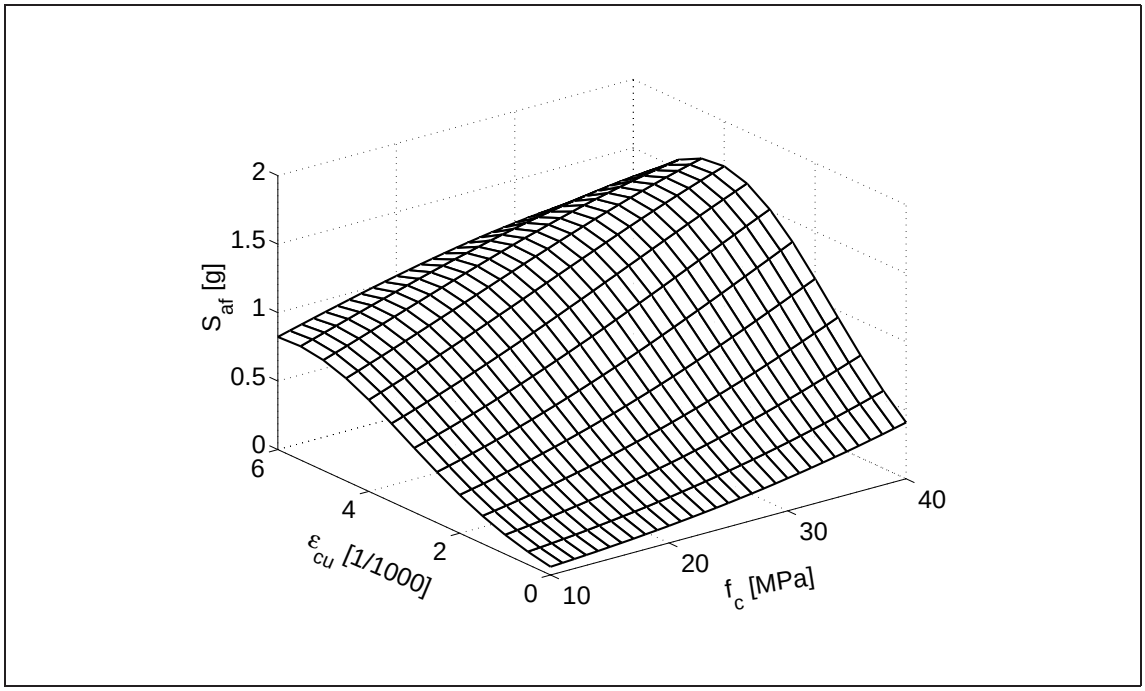

Figure 3.19: Response surface for the case of double blocking, using EQ Set I

Fragility curves are shown in Figure 3.20 and 3.21, where the results for the case of double blocking is shown together with the previous results for the case of single blocking. With

respect to the case of single blocking we observe an increase in the variability.

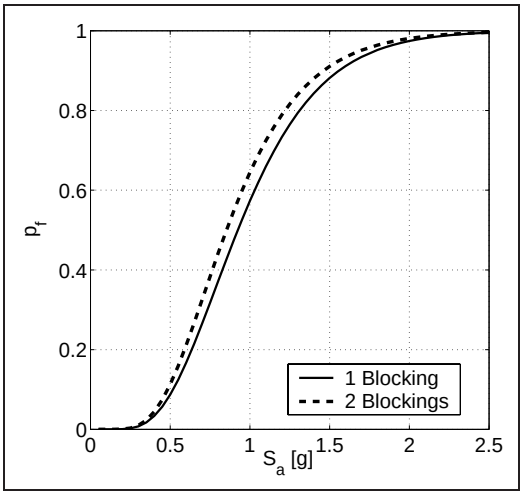

Figure 3.20: Single and double blocking

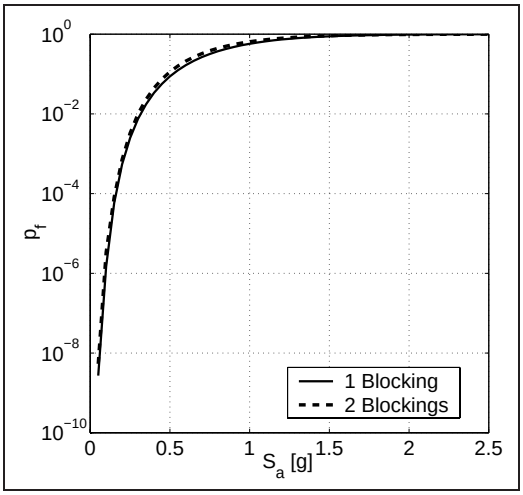

Figure 3.21: LOG scale

3.3.3 Validation by Monte Carlo simulation

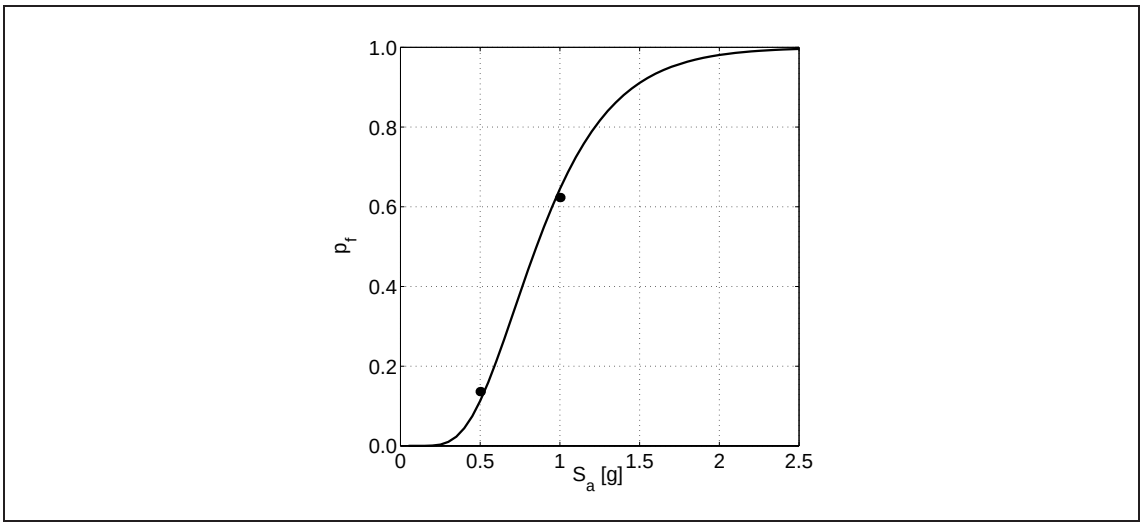

Figure 3.22: Validation of results using MC simulation

A sample check of the accuracy of the results has been done using Monte Carlo simulation also for this double-blocking case. The simulations have been again performed for two values of the spectral acceleration only. Values of the mechanical properties have been sampled from their respective distributions and have been associated with a earthquake record and a random field realisation randomly selected from the same set of eight earthquakes (Set I) and the same suite of field realisations used to establish the surface. Any difference emerging from the comparison between MC and RS results is necessarily due to either the response surface approximation or the SORM one.

For this case the two points obtained by MC have required 2782 dynamic analyses for a COV of 0.05, as opposed to the 387 performed in constructing the response surface used in calculating the 50 points that make up the fragility curve.

Comparing the two points found by MC with the estimates it can be seen that there is still good agreement even if the difference is slightly larger than for the previous case of single blocking.

The reason for more limited accordance between results should be sought mainly in the way the spatial variability is introduced in the statistical model. It was already observed in Chapter 2 that splitting the effect in two terms, one capturing the mean, the other the fluctuation around the mean, introduces interaction between these parameters that is neglected in the current model.

Chapter 4

4.1 Summary and final remarks

A statistical approach for time variant reliability problems has been developed and investigated in this study. The basic proposal is to use a response surface to represent the capacity part in an analytical limit state function as input for SORM analysis.

In order to include a maximum number of input variables, without taking account of them all explicitly, the response surface is expressed by a mixed model in terms of mechanical parameters (random variables and random factors) and seismic action (random factor). Using a random factor to characterise the variability due to the ground motion and constructing the response surface for the maximum response in time, the mixed model eliminates time-variance and gives a time-invariant limit state function for the time-variant problem. As the response surface represents the capacity part in the limit state function, the maximum response of interest is the response for which the structure just reaches failure. Defining failure as the union of all significant mechanisms, or in other words, defining failure as the collapse of any section due to any prescribed failure mechanism, the response takes the intersection between different failure mechanisms implicitly into account. The fragility function for the system can now be obtained through the performance of repeated SORM analyses, with the constructed empirical limit state function as input.

The developed theory has been applied to a three bay, six storey reinforced concrete frame, representing a system that is considered sufficiently complex to serve as a challenging test case. First only one random factor has been introduced, describing dependence of the response on the earthquake, then a second one, introducing spatial variability of the concrete ultimate strain.

Some investigations have been carried out to check the stability and accuracy of the suggested procedure, consisting of: treating the model coefficients both as deterministic and as random parameters in the limit state function; checking the significance of the chosen explanatory functions; observing the sensitivity of the results to the location of the centre of the design; determining the effect of the choice of different earthquake records; comparing some of the results with *exact* results found by Monte Carlo simulation.

From the results presented in the previous chapters one can make the following general remarks.

The proposed procedure is *general* in the sense that it can be used in conjunction with state-of-the-art mechanical models, and that the variability in the response due to uncertainty in the input ground-motion is realistically represented. Further, variability of the mechanical parameters can be included and is accurately represented in its spatial fluctuation. Finally, the computation of the failure probability accounts implicity for any type and number of different failure modes.

The method can be regarded *affordable* as the number of computations can still be considered as acceptably low. The exact number of analyses needed is a function of the system under consideration and of the iterative procedure used to find failure in the structure. In the case-study considered, the number that has been observed for 82 experiments varies around 400 analyses.

The method involves only *two sources of approximation*: the representation of the collapse condition through a response surface and the estimation of failure probability using FORM/SORM analysis. A limited validation of the fragility function through Monte Carlo simulation, for a small number of points, has shown that the approximations do not seem to introduce significant errors.

The example application has also indicated *difficulties* with the parameter estimation procedure, which are, however, not unfamiliar to users of advanced estimation techniques. Some *sensitivity to the experimental plan* was observed. The plan determines the form of the design matrix, and the product of the latter times its transpose has often turned out in a quasi-singular matrix, thus leading to inaccurate numerical solutions and accumulation of error, during the iterative parameter estimation procedure. Further, as already indicated in [8] and [15], the maximisation of the conditional likelihood has been found to be unstable in certain cases. All cases in which instability was observed concerned the second random factor in the situation of multiple blocking.

Finally, it is admitted that the procedure requires knowledge of statistical tools that are usually not part of the background of engineers.

4.2 Further research

The idea of the existence of a smooth response surface representing a failure condition that might be due to a multiplicity of mechanisms, introduced for the first time in [15] and adopted in the present work, can not be rigorously supported and deserves more thorough confirmation through extensive different numerical applications.

The experienced sensitivity to the experimental plan seems to be mainly due to the values attributed to the basic variables. When these latter are characterised by small variability the experimental points are not enough wide apart and this leads to rows of the design

matrix that are almost proportional (collinearity problem). This is related to the common procedure in factorial design of assigning upper and lower values located one standard deviation above and below the mean respectively. On the other hand, increasing the distance from the mean can lead to star points that are outside the admissible random variable range of values (for instance for Lognormal variables). A possible solution could therefore be to work in the space of coded variables, where all variables have the same order of magnitude, transforming the input variables directly to the standard normal space, and perform the regression there.

Beside the stability problem, one more reason prompts to search for different types of experimental plans in order to understand better possible sources of bias in the parameters introduced due to the use of a non-orthogonal design.

The used parameter estimation method and hierarchical model formulation present some convergence problems especially as it regards the variance components: it was already observed by Veneziano [15] that the variance estimators are biased to small values, and that the iterative nature of the procedure introduces bias itself. A natural advance from the so called hierarchical model, might be recourse to Bayesian parameter estimation.

Appendices

Appendix A

Selected earthquake records

No.	**Name**	**Date** (m/d/y)	**Station name**	**Rec.**	**M**	**R** [km]	**PGA** [g]	**Td** [s]	**S_a(1.046)** [g]
Set I									
Acc042	Trinidad	11/08/80	Rio Dell Overpass E	270	7.2	71.9	0.134	22.00	0.126
Acc049	Northridge	01/17/94	LA - Baldwin Hills	090	6.7	31.3	0.239	40.00	0.158
Acc053	Palm Springs	07/08/86	Cranston Forest	225	6.0	35.3	0.153	11.00	0.114
Acc059	Loma Prieta	10/18/89	Fremont Emerson Court	090	7.1	43.4	0.192	40.00	0.144
Acc065	Northridge	01/17/94	San Gabriel - E Grand	180	6.7	41.7	0.141	35.00	0.115
Acc071	Coalinga	05/02/83	Parkfield - Gold Hill 2W	000	6.5	36.6	0.083	40.00	0.162
Acc077	Coalinga	05/02/83	Parkfield - Gold Hill 5W	000	6.5	43.7	0.073	40.00	0.159
Acc087	Coalinga	05/02/83	Parkfield - Vineyard Cany 4W	000	6.5	37.7	0.064	30.00	0.108
Set II									
Acc028	Loma Prieta	10/18/89	Apeel 9 Crystal Springs Res.	227	7.1	46.9	0.104	39.00	0.203
Acc030	Loma Prieta	10/18/89	Apeel 10 Skyline	090	7.1	47.8	0.088	40.00	0.287
Acc070	Imperial Valley	10/15/79	Cerro Prieto	237	6.9	26.5	0.157	64.00	0.340
Acc089	Coalinga	05/02/83	Vineyard Cany 5W	090	6.5	37.1	0.062	40.00	0.142
Acc094	Coalinga	05/02/83	Parkfield - Fault Zone 6	000	6.5	32.8	0.055	40.00	0.165
Acc097	Coalinga	05/02/83	Parkfield - Fault Zone 9	090	6.5	31.9	0.050	32.00	0.167
Acc126	Loma Prieta	10/18/89	Sago South Surface	261	7.1	34.7	0.073	40.00	0.175
Acc144	Loma Prieta	10/18/89	Woodside	000	7.1	39.9	0.080	40.00	0.186

Table A.1: Accelerograms used in numerical experiments

Appendix B

Results from analysis

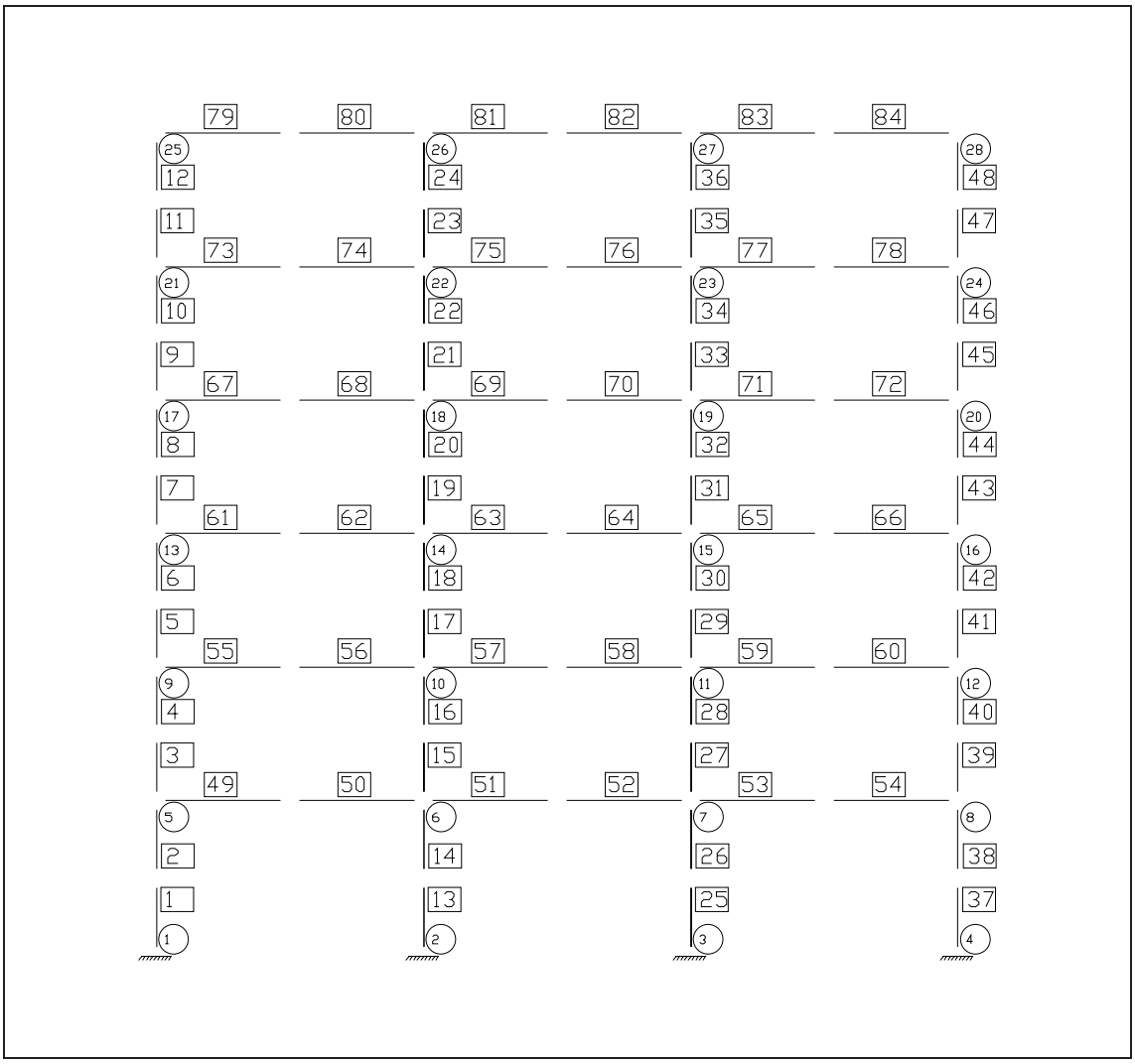

Figure B.1: Ruaumoko model, with node and element numbers

Experiment	Period [s]	$S_a(T_{mean})$ [g]	Scale	$S_{a;f}(T_{mean})$ [g]	Convergence	Element	Section	No. runs
1	1.170	0.126	6.54	0.822	1.07	25	1	8
2	0.955	0.126	6.13	0.772	0.98	34	2	3
3	1.170	0.126	4.37	0.549	1.05	34	2	4
4	0.955	0.126	8.06	1.015	1.07	25	1	4
5	1.170	0.126	6.67	0.839	1.01	34	2	5
6	0.955	0.126	6.10	0.767	1.06	34	2	5
7	0.955	0.126	6.17	0.777	0.94	34	2	6
8	1.170	0.126	5.41	0.680	1.02	34	2	3
9	0.955	0.158	5.78	0.916	0.93	25	1	4
10	1.170	0.158	3.22	0.509	1.01	8	2	6
11	0.955	0.158	3.52	0.558	1.05	25	1	4
12	1.170	0.158	4.52	0.717	0.98	6	2	6
13	1.170	0.158	3.72	0.589	1.03	25	1	4
14	0.955	0.158	3.73	0.591	1.03	25	1	4
15	1.170	0.158	2.99	0.473	0.97	34	2	15
16	0.955	0.158	5.56	0.880	1.06	25	1	4
17	0.955	0.114	9.90	1.129	0.93	25	1	4
18	1.170	0.114	12.20	1.390	0.96	25	1	4
19	0.955	0.114	15.87	1.810	1.01	25	1	3
20	1.170	0.114	5.49	0.626	1.02	22	2	4
21	1.170	0.114	6.62	0.755	0.99	22	2	3
22	0.955	0.114	14.49	1.652	0.99	25	1	4
23	1.170	0.114	11.63	1.326	1.05	22	2	3
24	0.955	0.114	9.71	1.107	0.98	25	1	4
25	1.170	0.144	4.33	0.622	1.02	10	2	2
26	0.955	0.144	8.20	1.177	0.96	25	1	4
27	1.170	0.144	5.75	0.825	1.03	25	1	7
28	0.955	0.144	5.99	0.860	0.98	25	1	7
29	0.955	0.144	5.88	0.845	1.06	25	1	9
30	1.170	0.144	5.29	0.760	1.05	25	1	6
31	0.955	0.144	7.09	1.018	0.98	25	1	4
32	1.170	0.144	2.66	0.382	0.96	34	2	3
33	0.955	0.115	10.75	1.234	1.05	25	1	11
34	1.170	0.115	13.33	1.531	0.98	25	1	7
35	0.955	0.115	13.33	1.531	1.03	25	1	3
36	1.170	0.115	9.52	1.093	0.95	42	2	4
37	1.170	0.115	8.20	0.941	0.93	22	2	4
38	0.955	0.115	14.29	1.640	1.00	25	1	4
39	1.170	0.115	11.63	1.335	1.06	25	1	2
40	0.955	0.115	11.63	1.335	0.93	25	1	5
41	1.170	0.162	2.74	0.444	1.00	8	2	6
42	0.955	0.162	5.56	0.901	0.98	25	1	3
43	1.170	0.162	3.15	0.512	1.00	25	1	5
44	0.955	0.162	3.34	0.542	0.99	25	1	4
45	0.955	0.162	3.48	0.565	1.00	25	1	5
46	1.170	0.162	3.37	0.546	0.99	25	1	5
47	0.955	0.162	4.90	0.795	0.95	25	1	4
48	1.170	0.162	2.65	0.430	1.02	6	2	7
49	1.170	0.159	10.31	1.639	1.04	25	1	3
50	0.955	0.159	7.81	1.242	1.02	25	1	8
51	1.170	0.159	5.46	0.869	0.95	8	2	6
52	0.955	0.159	10.87	1.728	0.99	25	1	3
53	0.955	0.159	12.20	1.939	0.94	25	1	15
54	1.170	0.159	4.63	0.736	1.05	8	2	3
55	0.955	0.159	6.99	1.112	1.04	25	1	4
56	1.170	0.159	8.93	1.420	1.06	22	2	8
57	0.955	0.108	15.38	1.655	1.01	25	1	4
58	1.170	0.108	5.88	0.633	1.03	44	2	3
59	0.955	0.108	13.16	1.416	0.94	22	2	5
60	1.170	0.108	13.70	1.474	1.06	25	1	5
61	0.955	0.108	16.13	1.735	0.94	25	1	3
62	1.170	0.108	6.90	0.742	1.01	34	2	4
63	1.170	0.108	6.71	0.722	1.03	22	2	4
64	0.955	0.108	15.63	1.681	1.06	25	1	2
65	1.046	0.126	6.76	0.850	0.96	25	1	6
66	1.046	0.158	4.20	0.667	0.95	25	1	8
67	1.046	0.114	10.64	1.211	1.04	25	1	3
68	1.046	0.144	5.81	0.836	0.94	25	1	8
69	1.046	0.115	12.05	1.389	1.07	25	1	5
70	1.046	0.162	3.32	0.539	0.99	25	1	8
71	0.838	0.126	9.17	1.154	1.01	25	1	4
72	1.577	0.126	3.13	0.394	1.02	34	2	5
73	1.046	0.158	5.99	0.948	1.07	25	1	5
74	1.046	0.158	0.97	0.153	1.01	25	1	1
75	1.046	0.114	9.52	1.090	1.06	22	2	3
76	1.052	0.114	10.99	1.257	1.01	25	1	4
77	1.046	0.144	6.58	0.942	0.98	25	1	5
78	1.046	0.144	5.35	0.767	1.01	25	1	6
79	1.046	0.115	12.05	1.380	0.95	25	1	7
80	1.046	0.115	12.99	1.485	0.94	25	1	5
81	1.046	0.162	3.65	0.592	1.01	25	1	3
82	1.046	0.162	2.99	0.486	1.05	25	1	5
Average			7.59	0.98		55×25		4.95
Min	0.84		0.97	0.15				1
Max	1.58		16.13	1.94				15
Sum								406

Table B.1: Results from dynamic analyses: Single Blocking, EQ Set I

$\mathbf{S_a}$	**Design point**							
	f'_c	ε_{cu}	f_y	S_1	S_2	S_3	δ_{eq}	ε
0.05	15	1.24	497	0.5843	0.8916	1.5670	-1.0421	-0.1276
0.10	17	1.41	430	0.5837	0.8903	1.5582	-0.9178	-0.1124
0.15	18	1.58	419	0.5831	0.8884	1.5503	-0.7606	-0.0932
0.20	19	1.71	414	0.5827	0.8868	1.5451	-0.6449	-0.0790
0.25	20	1.84	410	0.5824	0.8855	1.5413	-0.5543	-0.0679
0.30	20	1.95	407	0.5822	0.8843	1.5384	-0.4804	-0.0588
0.35	21	2.05	406	0.5820	0.8833	1.5360	-0.4178	-0.0512
0.40	21	2.14	404	0.5819	0.8824	1.5340	-0.3635	-0.0445
0.45	22	2.23	403	0.5817	0.8815	1.5323	-0.3157	-0.0387
0.50	22	2.31	402	0.5816	0.8807	1.5307	-0.2725	-0.0334
0.55	22	2.39	401	0.5815	0.8800	1.5293	-0.2333	-0.0286
0.60	23	2.46	400	0.5814	0.8793	1.5281	-0.1972	-0.0241
0.65	23	2.53	400	0.5814	0.8787	1.5269	-0.1636	-0.0200
0.70	23	2.60	399	0.5813	0.8781	1.5259	-0.1320	-0.0162
0.75	24	2.66	399	0.5812	0.8775	1.5249	-0.1023	-0.0125
0.80	24	2.72	399	0.5811	0.8769	1.5239	-0.0740	-0.0091
0.85	24	2.78	398	0.5811	0.8764	1.5231	-0.0469	-0.0057
0.90	24	2.83	398	0.5810	0.8758	1.5222	-0.0209	-0.0026
0.95	25	2.88	398	0.5809	0.8753	1.5214	0.0041	0.0005
1.00	25	2.93	398	0.5809	0.8748	1.5207	0.0284	0.0035
1.05	25	2.98	398	0.5808	0.8743	1.5199	0.0519	0.0064
1.10	25	3.02	398	0.5808	0.8738	1.5192	0.0748	0.0092
1.15	25	3.06	398	0.5807	0.8733	1.5185	0.0971	0.0119
1.20	26	3.10	398	0.5807	0.8729	1.5179	0.1187	0.0145
1.25	26	3.14	397	0.5806	0.8724	1.5172	0.1400	0.0171
1.30	26	3.17	397	0.5806	0.8719	1.5166	0.1607	0.0197
1.35	26	3.20	397	0.5805	0.8715	1.5160	0.1810	0.0222
1.40	26	3.23	397	0.5805	0.8710	1.5154	0.2010	0.0246
1.45	26	3.26	398	0.5804	0.8706	1.5148	0.2206	0.0270
1.50	26	3.29	398	0.5804	0.8701	1.5143	0.2397	0.0294
1.55	27	3.32	398	0.5803	0.8697	1.5137	0.2584	0.0316
1.60	27	3.34	398	0.5803	0.8693	1.5132	0.2767	0.0339
1.65	27	3.36	398	0.5802	0.8689	1.5127	0.2949	0.0361
1.70	27	3.38	398	0.5802	0.8685	1.5122	0.3126	0.0383
1.75	27	3.40	398	0.5801	0.8680	1.5117	0.3299	0.0404
1.80	27	3.42	398	0.5801	0.8676	1.5112	0.3470	0.0425
1.85	27	3.44	398	0.5800	0.8672	1.5108	0.3638	0.0446
1.90	27	3.46	398	0.5800	0.8668	1.5103	0.3802	0.0466
1.95	28	3.47	398	0.5799	0.8664	1.5099	0.3963	0.0485
2.00	28	3.49	398	0.5799	0.8661	1.5094	0.4124	0.0505
2.05	28	3.50	398	0.5798	0.8657	1.5090	0.4278	0.0524
2.10	28	3.52	399	0.5798	0.8653	1.5086	0.4433	0.0543
2.15	28	3.53	399	0.5798	0.8649	1.5082	0.4583	0.0561
2.20	28	3.54	399	0.5797	0.8646	1.5078	0.4730	0.0579
2.25	28	3.55	399	0.5797	0.8642	1.5074	0.4875	0.0597
2.30	28	3.56	399	0.5796	0.8638	1.5070	0.5018	0.0615
2.35	28	3.57	399	0.5796	0.8635	1.5066	0.5160	0.0632
2.40	29	3.58	399	0.5795	0.8631	1.5063	0.5297	0.0649
2.45	29	3.59	399	0.5795	0.8628	1.5059	0.5433	0.0665
2.50	29	3.60	399	0.5795	0.8625	1.5056	0.5567	0.0682
Mean	25	3	400	0.588	0.885	1.539	0	0

Table B.2: Design point coordinates for complete model: Single Blocking, EQ Set I

Experiment	Period [s]	$S_a(T_{mean})$ [g]	Scale	$S_{a;f}(T_{mean})$ [g]	Convergence	Element	Section	No. runs
1	1.170	0.126	6.41	0.806	0.97	25	1	6
2	0.955	0.126	5.13	0.645	0.99	34	2	3
3	1.170	0.126	4.12	0.518	0.99	34	2	4
4	0.955	0.126	7.58	0.953	1.02	25	1	4
5	1.170	0.126	5.43	0.684	0.97	34	2	3
6	0.955	0.126	5.29	0.666	0.98	34	2	3
7	0.955	0.126	5.46	0.687	0.94	34	2	4
8	1.170	0.126	4.41	0.554	1.01	34	2	4
9	0.955	0.158	4.95	0.784	1.05	25	1	5
10	1.170	0.158	3.19	0.506	0.94	8	2	5
11	0.955	0.158	3.31	0.525	0.97	25	1	4
12	1.170	0.158	4.20	0.666	1.01	6	2	5
13	1.170	0.158	3.88	0.614	1.02	25	1	37
14	0.955	0.158	3.77	0.598	0.99	25	1	7
15	1.170	0.158	2.62	0.416	0.95	6	2	5
16	0.955	0.158	4.81	0.762	1.05	25	1	4
17	0.955	0.114	8.85	1.009	0.99	25	1	4
18	1.170	0.114	10.75	1.226	1.05	25	1	2
19	0.955	0.114	14.08	1.606	0.99	25	1	4
20	1.170	0.114	5.00	0.570	1.02	22	2	2
21	1.170	0.114	5.49	0.626	1.05	22	2	3
22	0.955	0.114	12.99	1.481	0.99	25	1	4
23	1.170	0.114	9.09	1.036	1.06	22	2	3
24	0.955	0.114	7.81	0.891	1.00	25	1	4
25	1.170	0.144	3.58	0.515	0.98	10	2	4
26	0.955	0.144	7.52	1.080	0.94	25	1	6
27	1.170	0.144	5.52	0.793	1.02	25	1	6
28	0.955	0.144	5.41	0.776	1.02	25	1	2
29	0.955	0.144	4.88	0.700	1.05	10	2	2
30	1.170	0.144	5.59	0.802	0.93	25	1	7
31	0.955	0.144	6.58	0.945	0.94	25	1	4
32	1.170	0.144	2.48	0.355	0.99	22	2	3
33	0.955	0.115	10.64	1.221	0.96	25	1	5
34	1.170	0.115	12.99	1.491	0.95	25	1	5
35	0.955	0.115	14.08	1.617	0.94	25	1	4
36	1.170	0.115	7.69	0.883	0.95	46	2	3
37	1.170	0.115	7.19	0.826	1.02	22	2	5
38	0.955	0.115	12.82	1.472	1.02	25	1	4
39	1.170	0.115	11.24	1.290	1.00	25	1	4
40	0.955	0.115	10.20	1.171	0.98	22	2	3
41	1.170	0.162	2.49	0.404	1.05	8	2	2
42	0.955	0.162	4.76	0.772	1.01	25	1	4
43	1.170	0.162	2.88	0.467	0.99	25	1	5
44	0.955	0.162	2.94	0.477	1.02	25	1	5
45	0.955	0.162	3.11	0.504	1.02	25	1	21
46	1.170	0.162	3.21	0.520	1.00	25	1	5
47	0.955	0.162	4.41	0.715	1.03	25	1	5
48	1.170	0.162	2.32	0.376	1.06	6	2	36
49	1.170	0.159	8.40	1.336	1.04	25	1	4
50	0.955	0.159	6.90	1.097	1.01	22	2	4
51	1.170	0.159	5.15	0.820	0.98	8	2	4
52	0.955	0.159	9.35	1.486	1.02	25	1	4
53	0.955	0.159	10.53	1.674	1.03	25	1	5
54	1.170	0.159	4.03	0.641	1.00	8	2	4
55	0.955	0.159	6.25	0.994	1.07	25	1	4
56	1.170	0.159	7.52	1.195	1.01	22	2	7
57	0.955	0.108	14.93	1.606	1.06	25	1	4
58	1.170	0.108	4.98	0.535	1.06	44	2	3
59	0.955	0.108	12.35	1.328	0.95	8	2	3
60	1.170	0.108	13.33	1.435	1.01	25	1	3
61	0.955	0.108	16.95	1.824	0.94	25	1	5
62	1.170	0.108	5.56	0.598	0.99	22	2	4
63	1.170	0.108	5.99	0.644	0.94	44	2	3
64	0.955	0.108	15.63	1.681	1.02	25	1	5
65	1.046	0.126	5.92	0.818	0.94	34	2	4
66	1.046	0.158	3.73	0.749	1.05	25	1	8
67	1.046	0.114	10.31	1.164	0.99	22	2	3
68	1.046	0.144	6.17	1.443	0.93	25	1	6
69	1.046	0.115	11.76	1.067	0.95	25	1	5
70	1.046	0.162	2.95	1.336	0.99	25	1	5
71	0.838	0.126	8.70	1.094	0.94	25	1	4
72	1.577	0.126	2.89	0.364	1.02	10	2	6
73	1.046	0.158	5.59	0.997	0.96	25	1	6
74	1.046	0.158	0.77	0.200	1.06	25	1	2
75	1.046	0.114	7.81	1.018	1.03	22	2	3
76	1.052	0.114	10.42	1.081	0.99	25	1	4
77	1.046	0.144	5.18	1.298	0.94	34	2	4
78	1.046	0.144	5.29	1.458	0.93	25	1	6
79	1.046	0.115	10.87	0.862	1.06	25	1	3
80	1.046	0.115	11.76	0.953	0.98	25	1	7
81	1.046	0.162	2.92	1.315	1.01	25	1	5
82	1.046	0.162	2.95	1.366	1.02	25	1	5
Average			6.91	0.93		49×25		5.26
Min	0.84		0.77	0.20				2
Max	1.58		16.95	1.82				37
Sum								431

Table B.3: Results from dynamic analyses: Single blocking, EQ Set I, shifted centre of design

Experiment	Period [s]	$S_a(T_{mean})$ [g]	Scale	$S_{a;f}(T_{mean})$ [g]	Convergence	Element	Section	No. runs
1	1.170	0.203	3.75	0.759	1.01	25	1	5
2	0.955	0.203	3.97	0.804	1.01	25	1	5
3	1.170	0.203	2.05	0.416	0.99	42	2	3
4	0.955	0.203	5.49	1.113	1.01	25	1	5
5	1.170	0.203	4.12	0.834	1.01	25	1	5
6	0.955	0.203	3.92	0.795	1.07	25	1	4
7	0.955	0.203	3.53	0.716	1.05	25	1	4
8	1.170	0.203	4.29	0.870	0.95	42	2	4
9	0.955	0.287	3.22	0.923	1.02	25	1	8
10	1.170	0.287	2.48	0.710	0.99	8	2	2
11	0.955	0.287	2.27	0.652	1.03	25	1	4
12	1.170	0.287	3.29	0.944	1.02	25	1	4
13	1.170	0.287	3.32	0.953	0.93	25	1	4
14	0.955	0.287	2.99	0.857	0.94	25	1	19
15	1.170	0.287	1.71	0.490	0.94	42	2	5
16	0.955	0.287	3.05	0.875	1.05	25	1	4
17	0.955	0.340	3.29	1.118	1.04	34	2	8
18	1.170	0.340	3.19	1.086	0.95	34	2	2
19	0.955	0.340	3.92	1.333	0.95	25	1	5
20	1.170	0.340	1.64	0.557	1.06	44	2	4
21	1.170	0.340	2.12	0.720	1.06	44	2	2
22	0.955	0.340	4.48	1.525	0.97	25	1	6
23	1.170	0.340	3.08	1.046	0.98	22	2	3
24	0.955	0.340	3.23	1.097	0.97	22	2	6
25	1.170	0.142	5.78	0.823	1.04	8	2	3
26	0.955	0.142	13.89	1.978	1.00	25	1	3
27	1.170	0.142	8.77	1.249	1.07	6	2	4
28	0.955	0.142	8.47	1.207	0.93	25	1	3
29	0.955	0.142	8.06	1.148	0.98	34	2	3
30	1.170	0.142	8.55	1.217	1.01	34	2	3
31	0.955	0.142	12.20	1.737	0.95	25	1	5
32	1.170	0.142	4.78	0.681	1.06	6	2	2
33	0.955	0.165	5.75	0.949	0.94	25	1	6
34	1.170	0.165	5.99	0.989	0.94	25	1	4
35	0.955	0.165	7.63	1.261	1.01	25	1	8
36	1.170	0.165	3.83	0.633	1.03	44	2	2
37	1.170	0.165	3.98	0.658	1.02	44	2	4
38	0.955	0.165	7.69	1.271	1.01	25	1	4
39	1.170	0.165	5.95	0.983	0.96	25	1	6
40	0.955	0.165	4.78	0.790	1.06	25	1	3
41	1.170	0.167	3.76	0.626	1.03	44	2	3
42	0.955	0.167	9.80	1.633	0.99	25	1	4
43	1.170	0.167	7.58	1.262	0.98	25	1	4
44	0.955	0.167	6.41	1.068	0.95	25	1	3
45	0.955	0.167	6.80	1.133	0.98	34	2	3
46	1.170	0.167	7.94	1.322	0.98	34	2	3
47	0.955	0.167	10.10	1.683	0.99	25	1	4
48	1.170	0.167	3.62	0.604	0.97	44	2	3
49	1.170	0.175	8.85	1.547	0.98	25	1	6
50	0.955	0.175	7.94	1.387	0.99	25	1	6
51	1.170	0.175	5.03	0.878	1.02	8	2	3
52	0.955	0.175	12.05	2.106	0.93	25	1	4
53	0.955	0.175	9.80	1.714	1.00	25	1	4
54	1.170	0.175	5.38	0.940	1.02	44	2	5
55	0.955	0.175	7.63	1.334	1.05	25	1	6
56	1.170	0.175	8.55	1.494	0.96	25	1	4
57	0.955	0.186	5.38	1.001	1.06	25	1	6
58	1.170	0.186	3.09	0.575	1.04	8	2	3
59	0.955	0.186	4.20	0.782	1.06	25	1	5
60	1.170	0.186	4.27	0.796	0.98	25	1	4
61	0.955	0.186	5.65	1.052	1.05	25	1	4
62	1.170	0.186	3.16	0.589	1.00	8	2	5
63	1.170	0.186	2.98	0.554	0.95	25	1	6
64	0.955	0.186	6.06	1.128	0.96	25	1	5
65	1.046	0.203	4.03	0.818	1.04	25	1	4
66	1.046	0.287	2.61	0.749	1.01	25	1	5
67	1.046	0.340	3.42	1.164	0.94	34	2	4
68	1.046	0.142	10.10	1.443	1.02	25	1	3
69	1.046	0.165	6.45	1.067	1.04	25	1	5
70	1.046	0.203	8.00	1.336	0.94	25	1	4
71	0.838	0.203	5.59	1.132	0.95	25	1	4
72	1.577	0.094	1.78	0.166	1.00	44	2	3
73	1.046	0.287	3.47	0.997	1.04	25	1	4
74	1.046	0.287	0.70	0.200	0.97	25	1	2
75	1.046	0.340	2.99	1.018	1.06	22	2	3
76	1.052	0.347	3.12	1.081	1.03	25	1	6
77	1.046	0.142	9.09	1.298	1.06	22	2	4
78	1.046	0.142	10.20	1.458	0.98	25	1	4
79	1.046	0.165	5.21	0.862	0.96	6	2	3
80	1.046	0.165	5.78	0.953	1.03	34	2	4
81	1.046	0.167	7.87	1.315	1.02	25	1	3
82	1.046	0.167	8.20	1.366	0.98	25	1	6
Average			5.48	1.03		51×25		4.37
Min	0.84		0.70	0.17				2
Max	1.58		13.89	2.11				19
Sum								358

Table B.4: Results from dynamic analyses: Single blocking, EQ Set II

Experiment	Period [s]	$S_a(T_{mean})$ [g]	Scale	$S_{a;f}(T_{mean})$ [g]	Convergence	Element	Section	No. runs
1	1.170	0.126	6.58	0.828	0.97	34	2	6
2	0.955	0.126	4.81	0.605	1.04	9	1	3
3	1.170	0.126	3.91	0.491	1.02	34	2	4
4	0.955	0.126	7.04	0.886	1.00	25	1	5
5	1.170	0.126	4.00	0.503	1.05	34	2	3
6	0.955	0.126	5.29	0.666	1.05	34	2	3
7	0.955	0.126	5.46	0.687	1.04	22	2	3
8	1.170	0.126	4.44	0.559	0.96	33	1	5
9	0.955	0.158	6.45	1.022	0.99	34	2	4
10	1.170	0.158	3.13	0.497	1.03	25	1	4
11	0.955	0.158	3.73	0.591	1.06	25	1	4
12	1.170	0.158	4.18	0.663	0.96	9	1	8
13	1.170	0.158	4.22	0.668	0.94	25	1	16
14	0.955	0.158	4.39	0.695	0.96	25	1	6
15	1.170	0.158	2.75	0.435	0.95	34	2	4
16	0.955	0.158	4.31	0.683	1.03	25	1	5
17	0.955	0.114	10.10	1.152	1.05	45	1	2
18	1.170	0.114	11.76	1.341	1.05	25	1	3
19	0.955	0.114	16.67	1.900	0.99	25	1	3
20	1.170	0.114	4.88	0.556	0.96	34	2	4
21	1.170	0.114	5.99	0.683	1.02	34	2	3
22	0.955	0.114	11.49	1.310	1.01	25	1	4
23	1.170	0.114	7.41	0.844	1.01	45	1	5
24	0.955	0.114	9.62	1.096	0.96	22	2	3
25	1.170	0.144	3.62	0.520	1.00	34	2	4
26	0.955	0.144	6.90	0.990	0.96	25	1	6
27	1.170	0.144	5.88	0.845	0.96	25	1	5
28	0.955	0.144	6.06	0.870	1.04	25	1	6
29	0.955	0.144	6.21	0.892	0.98	34	2	8
30	1.170	0.144	4.59	0.659	1.04	9	1	3
31	0.955	0.144	8.47	1.217	0.96	1	1	6
32	1.170	0.144	2.83	0.407	0.94	33	1	4
33	0.955	0.115	10.10	1.160	0.96	25	1	4
34	1.170	0.115	12.99	1.491	1.02	34	2	3
35	0.955	0.115	12.99	1.491	1.03	25	1	4
36	1.170	0.115	7.94	0.911	1.00	46	2	3
37	1.170	0.115	7.41	0.850	1.02	22	2	5
38	0.955	0.115	13.89	1.594	0.97	25	1	4
39	1.170	0.115	10.99	1.262	0.97	25	1	4
40	0.955	0.115	11.24	1.290	1.05	34	2	4
41	1.170	0.162	2.78	0.451	1.04	8	2	7
42	0.955	0.162	5.13	0.832	1.00	25	1	4
43	1.170	0.162	2.58	0.419	1.06	25	1	5
44	0.955	0.162	4.00	0.649	1.06	25	1	4
45	0.955	0.162	3.27	0.530	1.04	25	1	4
46	1.170	0.162	3.50	0.567	0.98	1	1	5
47	0.955	0.162	4.46	0.724	0.98	25	1	5
48	1.170	0.162	2.63	0.427	1.03	25	1	6
49	1.170	0.159	8.26	1.314	0.96	25	1	3
50	0.955	0.159	7.58	1.205	1.06	34	2	13
51	1.170	0.159	5.56	0.883	0.98	25	1	7
52	0.955	0.159	9.80	1.559	1.00	25	1	3
53	0.955	0.159	11.63	1.849	0.95	25	1	4
54	1.170	0.159	5.10	0.811	1.07	22	2	3
55	0.955	0.159	6.94	1.104	0.93	25	1	2
56	1.170	0.159	7.52	1.195	0.94	7	1	2
57	0.955	0.108	14.29	1.537	0.96	25	1	5
58	1.170	0.108	5.18	0.558	0.98	45	1	6
59	0.955	0.108	11.63	1.251	1.01	13	1	7
60	1.170	0.108	11.90	1.281	0.98	9	1	4
61	0.955	0.108	15.63	1.681	0.96	34	2	4
62	1.170	0.108	4.98	0.535	1.03	34	2	3
63	1.170	0.108	9.62	1.035	0.93	44	2	5
64	0.955	0.108	15.63	1.681	0.97	25	1	4
65	1.046	0.126	6.62	0.832	1.05	25	1	6
66	1.046	0.158	3.83	0.608	0.93	25	1	7
67	1.046	0.114	9.26	1.060	1.02	22	2	4
68	1.046	0.144	4.88	0.702	1.06	34	2	4
69	1.046	0.115	12.05	1.387	0.99	25	1	5
70	1.046	0.162	3.61	0.585	0.93	25	1	6
71	0.838	0.126	8.93	1.123	0.99	25	1	5
72	1.577	0.126	2.84	0.357	0.97	34	2	3
73	1.046	0.158	6.80	1.080	0.94	34	2	4
74	1.046	0.158	1.86	0.295	1.07	8	2	6
75	1.046	0.114	6.85	0.778	1.06	34	2	3
76	1.052	0.114	10.75	1.230	0.99	25	1	4
77	1.046	0.144	3.34	0.480	1.04	34	2	6
78	1.046	0.144	5.24	0.751	0.99	25	1	8
79	1.046	0.115	10.75	1.239	1.02	42	2	3
80	1.046	0.115	13.70	1.563	0.94	34	2	4
81	1.046	0.162	2.92	0.475	1.03	25	1	4
82	1.046	0.162	2.81	0.456	1.00	25	1	7
Average			7.09	0.913		38×25		4.72
Min	0.84		1.86	0.30				2
Max	1.58		16.67	1.90				16
Sum								387

Table B.5: Results from dynamic analyses: Double blocking, EQ Set I

Bibliography

[1] Paolo Bazzurro and Allin Cornell. *Seismic Hazard Analysis of Nonlinear Structures. I: Methodology*. Journal of Structural Engineering, ASCE, Vol. 120, No. 11, November 1994, p.3320-3344.

[2] Paolo Bazzurro and Allin Cornell. *Seismic Hazard Analysis of Nonlinear Structures. II: Applications*. Journal of Structural Engineering, ASCE, Vol. 120, No. 11, November 1994, p.3345-3365.

[3] Jack R. Benjamin and C. Allin Cornell. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill, 1970.

[4] George E.P. Box and Norman R. Draper. *Empirical Model-Building and Response Surfaces*. John Wiley & Sons, 1987.

[5] Athol J. Carr. *Ruaumoko, FORTRAN Program Code for Inelastic Dynamic Analysis*. University of Canterbury, 1998.

[6] C. Allin Cornell. *Calculating Building Seismic Performance Reliability: a basis for multi-level design norms*. 11th WCEE, Acapulco, Mexico, STS 13, 1996.

[7] CUR. *Kansen in de civiele techniek, Deel 1: Probabilistisch ontwerpen in theorie*. CUR, Directoraat-Generaal Rijkswaterstaat, 1997.

[8] A. Khuri and J. Cornell. *Response Surfaces: Designs and Analyses*. Marcel Dekker, 1996.

[9] T. Paulay and M.J.Nigel Priestley. *Seismic Design of Reinforced Concrete and Masonry Buildings*. John Wiley & Sons, 1992.

[10] Paolo E. Pinto. *Reliability methods in earthquake engineering*. Progr.Struct.Eng.Mater., **3**(1), 2001, p.76-85.

[11] M.J. Nigel Priestley, F. Seible, G.M. Calvi. *Seimic Design an Retrofit of Bridges*. John Wiley & Sons, 1996.

[12] Malur R. Rajashekhar and Bruce R. Ellingwood. *A new look at the response surface approach for reliability analysis*. Structural Safety, 12, 1993, p.205-220.

[13] Shayle R. Searle, George Casella and Charles E. McCulloch. *Variance Components*. Prentice-Hall, 1992.

[14] N. Shome, A.C. Cornell, P. Bazzurro and I.E. Carbalho. *Earthquakes, records and non-linear responses*. Earthquake Spectra, **14**(3), 1998, p.469-500.

[15] D. Veneziano, F. Casciati and L. Faravelli. *Method of seismic fragility for complicated systems*. Proceedings of the 2nd Committee of Safety of Nuclear Installation (CNSI) Specialist Meeting on Probabilistic Methods in Seismic Risk Assessment for NPP. Lawrence Livermore Laboratory, CA.