Bulletin of Earthquake Engineering **1:** 321–347, 2003.

Uncertainty Analysis of Strong-Motion and Seismic Hazard

R. SIGBJÖRNSSON[1,2] and N. N. AMBRASEYS[1]
[1]*Department of Civil and Environmental Engineering, Imperial College of Science, Technology and Medicine, South Kensington Campus, London SW7 2AZ, UK;* [2]*Currently on leave from University of Iceland, Earthquake Engineering Research Centre, Austurvegur 2a, IS-800 Selfoss, Iceland (Tel: +354 525 4141; Fax: +354 525 4140; E-mail: Ragnar.Sigbjornsson@afl.hi.is; r.sigbjornsson@imperial.ac.uk)*

Received 24 June 2003; accepted 8 December 2003

Abstract. In this article we present the modelling of uncertainty in strong-motion studies for engineering applications, particularly for the assessment of earthquake hazard. We examine and quantify the sources of uncertainty in the basic variables involved in ground motion estimation equations, including those associated with the seismological parameters, which we derive from a considerable number of strong-motion records. Models derived from regression analysis result in ground motion equations with uncertain parameters, which are directly related to the selected basic variables thus providing an uncertainty measure for the derivative variable. These uncertainties are exemplified and quantified. An alternative approach is presented which is based on theoretical modelling defining a functional relationship on a set of independent basic variables. Uncertainty in the derivative variable is then readily obtained when the uncertainties of the basic variables have been defined. In order to simplify the presentation, only the case of shallow strike-slip earthquakes is presented. We conclude that the uncertainty is approximately the same as given by the residuals typical for regression modelling. This implies that uncertainty in ground motion modelling cannot be reduced below certain limits, which is in accordance with findings reported in the literature. Finally we discuss the implications of the presented methodology in hazard analyses, which is sensitive to the truncation of the internal error term, commonly given as an integral part of ground motion estimation equations. The presented methodology does not suffer from this shortcoming; it does not require truncation of the error term and yields realistic hazard estimates.

Key words: error analysis, ground motion estimation equation, hazard assessment, regression, strong-motion, uncertainty, uncertainty analysis

1. Introduction

A fundamental relation in seismic hazard and risk assessment is the attenuation scaling relationship, or ground motion estimation equation, which is needed to estimate the strong ground motion at a given site caused by an earthquake of given characteristics. With such a relationship it is possible to transfer the activity of a given seismogenic region into the seismic action required for the design of structures and for risk assessment. The uncertainties involved in this assessment can be divided into two main categories, referred to as aleatory and epistemic uncertainties. Epistemic uncertainties are due to lack of knowledge to describe fully the phenomenon. Obtaining new data and refining the modelling can reduce these uncertainties. Aleatory uncertainties, on the other hand, are related to the inherent unpredictability of earthquake processes. Such uncertainties cannot be reduced. It is important, however, to be able to quantify these uncertainties correctly in the design process to ensure adequate safety.

The objective of this paper is to discuss the nature of uncertainties connected with strong ground motion and quantify them. Furthermore, to give an example of uncertainty modelling and discuss the implications for engineering design and hazard assessment.

2. Uncertainty in strong-motion data

Uncertainties in the characterisation of strong ground motion are of two main types, i.e., uncertainties related to the functional form of the model (formal or internal) and uncertainties inherent in the input or basic variables used in the modelling. The basic variables commonly used can be divided into three main categories.

(i) In the first category we have the basic variables which describe the source, and include magnitude, or seismic moment, epicentral location, depth and source dimensions. Furthermore, we believe that this category should also include the type of focal mechanism, which is not always recognised, but is of importance, especially for distances to source less than four to five times the characteristic source dimensions.

(ii) The second category contains basic variables characterising the site. These include distance to the source, as well as variables describing the site conditions reflecting the local geology and topography.

(iii) The third category includes basic variables describing the wave propagation process and properties of the ray path from source to site. These variables include the mechanical properties of the material along the path, including its damping characteristics. It is common to treat some of these variables as derivative variables rather than basic variables, for instance, to relate anelastic attenuation to the source distance.

In the modelling process there is a general tendency towards simplifications since the best model is the model with the smallest number of basic variables that predict the derivatives with sufficient accuracy and reliability, conforming to available data. From the engineering point of view, such a model is preferable as it simplifies design decisions and makes the design process more robust than is the case for more complex models.

2.1. Data

Most of the strong-motion records in the European area, come from events already analysed by the International Seismological Centre (ISC, 2003), Harvard (2003) and from special studies. Many of the earthquakes are of moderate magnitude and are reported from a relatively large enough number of stations to ensure reasonable azimuthal coverage. The locations found by ISC are therefore not likely to be in serious formal error and can be used as initial values for refinement. The main uncertainty is in depth of focus. For certain parts of Europe focal depth is an important consideration, particularly for small magnitude events. Teleseismic locations are known to have larger uncertainties compared with those from local networks.

The use of a unified magnitude scale in attenuation studies is an important consideration. Our adoption of the surface-wave magnitude M_S rather than the local magnitude M_L stems from the fact that the former is not only the best estimator of the size of a crustal earthquake, but also from the fact that seismicity in Europe is generally evaluated in terms of M_S.

Equally important is the assessment of reliable source distance, particularly in the near-field, from the location of the recording sites. The distance or source-path one assigns to a strong-motion record has a significant influence on the close-in behaviour of attenuation curves, particularly for small events for which location errors can be many times the source dimension. These errors accrue owing to errors in source and station location.

For most of the larger earthquakes one may adopt the closest distance to the projection of the fault rupture. For small-magnitude crustal events the source distance is close to the epicentral distance. However, the locations of some of the smaller events are poorly known, and for this reason their position must be re-evaluated.

Local site conditions (soil, topography, instrument location, housing and characteristics) at many strong-motion stations are poorly known, particularly in the case of old sites that have been moved or abandoned, or for temporary stations used for aftershock studies. In terms of the soil conditions the majority of sites can only be described in very general terms at best, such as “soil” or “rock”. There are however, some stations for which there is no knowledge of the soil conditions.

The topographical details at most stations are even less well described. Where they do exist, they may be given only in terms of very broad descriptions, such as “at the top of a hill”, without any reference to the hill dimensions or the surrounding geomorphology.

Instrument data is usually more readily available, at least in general terms of the instrument type and the structure in which it is housed. However, it is not uncommon to have no knowledge of the specific characteristics of the instrument (sensitivity and damping). Furthermore, it is even less common to have detailed information regarding the structure in which the instrument is housed.

Some of the differences between ground motion estimation equations, particularly for near-field conditions, often arise from the size of the data sample used in their derivation, as well as from different distributions, biases and range of applicability of the variables.

The use of different magnitude scales also introduces a significant bias with respect to size and depth of the nucleation of the generating events, particularly for small ($M_S < 5.5$) and larger ($M_S > 6.7$) earthquakes. For instance, local magnitude M_L saturates much earlier than M_S, and for relatively small events, before the last decade, M_S cannot be compared directly with M_w. One of the reasons for this is that M_S is not a linear function of the logarithm of the seismic moment, $\log_{10}(M_o)$, for the whole range of magnitudes. For small events M_S and $\log_{10}(M_o)$, have a 1:1 scaling relation; for intermediate magnitudes this ratio is 1:1.5, reaching 1:2 for very large events (Ekstrom & Dziewonski, 1988). In other words, a model derived from M_w is a non-linear function in terms of M_S, for the same M_o increment, δM_S for smaller magnitudes being larger for small than for large shocks.

Additional differences between different equations arise from the modelling of ground motion estimation equations and fitting method used to regress the data. Results can be affected considerably by using a magnitude-dependent shape or a two-stage regression with weights.

There is no significant variation among different regions for shallow earthquakes, and there is remarkable agreement between Europe, western North America and New Zealand. One should not give much credence to differences in ground motion estimation equations between different countries; there is little physical basis for groupings within political boundaries. Some of these apparent differences arise from the limited subsets of data and their different distributions and biases. Ideally, individual tectonic regions should have their own relationships, but at present this is not feasible because of the limited available data. Comparison of results from the European dataset shows that regional differences are not very large, certainly for near-field predictions (Douglas, 2003b). They are all within the standard deviation of the residuals determinations, which are not better than by a factor of 1.7.

2.2. Regression modelling

Modelling based on regression analysis provides far the most common approach to establish ground motion estimation equations, often called attenuation relationship. Douglas (2003a) has given a comprehensive overview of these models encountered in the literature.

Both linear and non-linear regression methods are applied. Most ground motion estimation equations found in the literature have the following basic form:

$$\log_{10}(a) = f(M, R, source, soil) + P\sigma \tag{1}$$

where, a denotes the absolute value of the strong-motion variable, for instance, the larger component of the horizontal peak ground acceleration; M is the earthquake magnitude; R is the distance to the causative fault, while source and soil refer to the source parameters and soil conditions at the site, respectively; σ reflects the standard error obtained fitting the model to the data, and P is standard normal distribution with zero mean and unit standard deviation, as it is common to assume the error term normally distributed (Douglas, 2003a). Hence, P denotes the number of standard deviations used when estimating the strong-motion variable, a, from this expression.

The usual definition of the error residuals is the logarithm base 10 of the difference between the observed and predicted peak ground acceleration, or:

$$\varepsilon = \log_{10}\left(\frac{a_{observed}}{a_{predicted}}\right) \tag{2}$$

where, $a_{observed}$ and $a_{predicted}$ refer, respectively, to a data point and the corresponding value predicted by the regression curve. The standard deviation of ε is the above-mentioned σ, which typically has values in the range 0.2 to 0.3 (Douglas, 2003a).

We find that the standard error is a function of the inherent uncertainty in the variables M and R as well as in other variables included in the ground motion estimation equations. This is evident in the case of site conditions as the standard error derived with uniform soil conditions is smaller than the standard error for a sample with mixed soil properties. The same applies to source mechanics and depth. Furthermore, the standard error obtained for a single event is smaller than the error derived from a sample containing many earthquakes, even in the case where influences from variables other than M and R are kept as small as possible. This was first pointed out by Brillinger and Preisler (1985). Applying a ground motion estimation equation with two variables, M and R, they find that the standard error could be split into two parts: (a) a contribution related to the variability between earthquakes, $\sigma_M = 0.2284$, and (b) a contribution related to variability between records from the same earthquake, $\sigma_R = 0.1223$. This gave a total error $\sigma = \sqrt{0.2284^2 + 0.1223^2} = 0.259$.

This indicates that the total standard error can be considered as being composed of contributions related to uncertainties inherent in the quantities governing the physical process and hence the variables of the mathematical model fitted to the data. Therefore it is not obvious that increasing the number of model variables will lead to a reduction of the total standard error. On the other hand, a refined model may better explain the sources of uncertainties than can be done using a simplified model.

To exemplify the uncertainties involved, let us consider the following simplified model:

$$\log_{10}(PGA) = b_0 + b_1 M - \log_{10}(R) + b_2 R \tag{3}$$

Here *PGA* denotes peak ground acceleration (the derivative variable); M is earthquake magnitude; R is a distance parameter defined as $R = \sqrt{D^2 + h^2}$, where D is the fault distance and h is a depth parameter. The quantities M, D and h are the basic variables of this model, while *PGA* is the derivative variable. The basic variables can be assumed (for most purposes) as independent, while the derivative variable clearly is not.

Within the framework of the linear regression analysis this model can be expressed as (see for instance Draper and Smith, 1998):

$$\mathbf{y} = \mathbf{Ab} \tag{4a}$$

where:

$$\mathbf{b} = [b_0\ b_1\ b_2]^T \tag{4b}$$

and:

$$\mathbf{A} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \tag{4c}$$

Here we introduce $x_{1k} = M$ and $x_{2k} = R$, where the index k refers to the record number; furthermore:

$$\mathbf{y} = \begin{bmatrix} \log_{10}(PGA_1) + \log_{10}(R_1) \\ \log_{10}(PGA_2) + \log_{10}(R_2) \\ \ldots \\ \log_{10}(PGA_n) + \log_{10}(R_n) \end{bmatrix} \tag{4d}$$

The least square estimates of the model parameters, Eq. (4b), are obtained as:

$$\mathbf{b} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{y} \tag{4e}$$

The residuals can then be estimated using the following expression:

$$s_\varepsilon^2 = var(\varepsilon) = \frac{1}{n-4}(\mathbf{y} - \mathbf{Ab})^T(\mathbf{y} - \mathbf{Ab}) \tag{4f}$$

and, furthermore, the covariance matrix for the **b**-coefficients is:

$$\mathbf{C_b} = cov(\mathbf{b}) = s_\varepsilon^2\left(\mathbf{A}^T\mathbf{A}\right)^{-1} \tag{4g}$$

It is worth noting that this leads to **b**-parameters, which are uncertain, and that this uncertainty is related directly to the basic variables and the functional form of the selected model. As outlined above there are clearly well established uncertainties in M as well as in the epicentre location, which transfers to the source distance measure applied. These uncertainties are imbedded in the data and will be visualised in the regression model through the distributions of the **b**-parameters and the residuals.

To give an example of the uncertainties involved in this procedure, an analysis was carried out using data from the Imperial College Strong-Motion Databank (see also Ambraseys *et al.*, 2002). The following selection criteria were used: epicentral

Table I. Results of the regression analysis

	Mean value	Standard deviation	Coefficient of variation	Confidence interval (95% confidence level)	
b_0	-1.2780	0.1909	0.1494	-1.6528	-0.9033
b_1	0.2853	0.0316	0.1107	0.2233	0.3472
b_2	$-1.730\ 10^{-3}$	$2.132\ 10^{-4}$	0.1232	-0.0021	-0.0013

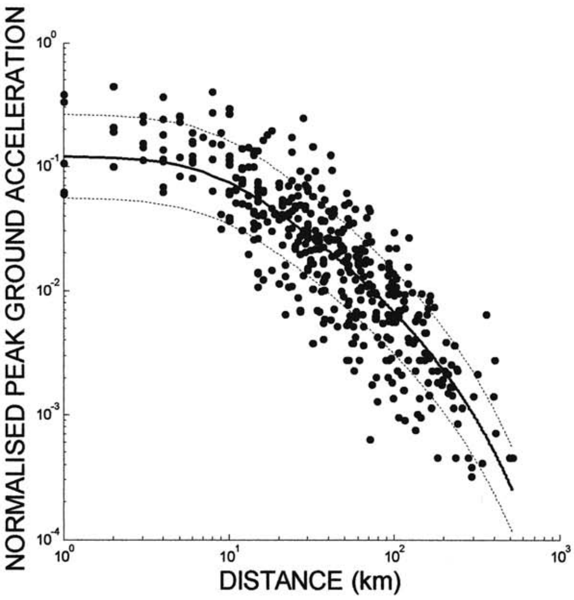
Table II. Correlation coefficients for the regression

	b_0	b_1	b_2
b_0	1	-0.9938	-0.0679
b_1		1	-0.0076
b_2	*symmetric*		1

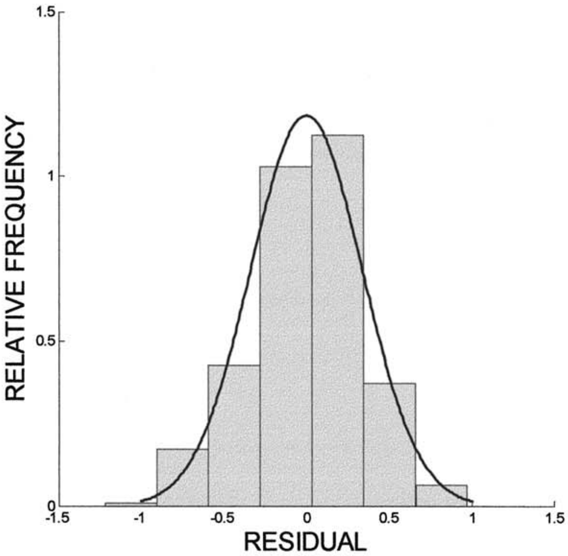
distance < 1000 km; magnitude (M_w or M_S) in the range 5 to 7; depth < 20 km and strike slip mechanism. This resulted in 465 records with mixed site conditions. In the analysis the shortest distance to fault was used as a source distance whenever available; otherwise the epicentral distance was used. In the analysis only the larger horizontal component of peak ground acceleration was selected. Furthermore, the depth parameter was arbitrarily fixed to 8 km. The results of the analysis are displayed in Figure 1, including the distribution of the residuals. The regression coefficients are given in Tables I and II along with basic statistics.

It is worth noting that the coefficient of variation for the estimated **b**-coefficients is in the range 11 to 15% and also that b_0 and b_1 are very strongly correlated with a negative correlation coefficient, which indicates that increase in b_0 implies reduction in b_1. Furthermore, b_0 and b_1 are almost uncorrelated with b_2. The strong negative correlation between b_0 and b_1 appears logical as the peak ground acceleration is governed by $b_0 + b_1 M$ as the distance D approaches zero.

The lack of correlation between b_2, on the one hand, and b_0 or b_1, on the other, can be interpreted as a result of zero correlation between M and D which is obvious. Hence, we may conclude that the statistical properties of the **b**-coefficients seem in accordance with the statistical properties of the basic variables. It should be pointed out that for the data used there is a positive correlation between the peak ground acceleration and magnitude, but negative correlation between the peak ground acceleration and distance. Also, this is in accordance with the physics of the process, and it is important to note that peak ground acceleration is correlated with the basic variables.



(a)



(b)

Figure 1. Attenuation of peak ground acceleration in shallow strike slip earthquakes. The larger horizontal component is displayed. The magnitude and distance measure used are, respectively, surface-wave magnitude and shortest distance to fault. (a) Normalised horizontal peak ground acceleration as a function of distance. The solid line is the regression line, and the dotted lines represent the mean $\pm$ one standard deviation. Dots are the data points. (b) Distribution of residuals compared to normal probability density. The standard deviation is equal to 0.3368.

The following model was also studied, using regression analysis (Ambraseys, 1995):

$$\log_{10}(PGA) = b_0 + b_1 M + b_2 R + b_3 \log_{10}(R) \tag{5}$$

This model gives slightly smaller residuals than the model given in Eq. (3). Furthermore, the b_3-parameter is smaller than -1, which is interesting in view of the results published in the literature (see for instance Douglas, 2003a). The behaviour of the **b**-parameters in this case showed the same main trend as in the previous case using Eq. (3).

It is also worth noting that this equation can be rewritten as follows, introducing the magnitude as a derivative response variable:

$$M = a_0 + a_1 \log_{10}(PGA) + a_2 R + a_3 \log_{10}(R) \tag{6}$$

The regression analysis based on this equation, using the above-mentioned data set, shows that the residuals were approximately normally distributed and with a standard deviation of the same magnitude as for the peak ground acceleration. We will apply this result in Section 4.

2.3. Error Analysis

In the foregoing we pointed out the uncertainty of the regression parameters and their functional relationship to the basic variables. This type of uncertainty is normally not included in the assessment of strong ground motion, and the **b**-parameters are treated as constants. On the other hand, the error term in the regression equation, Eq. (1), is usually accounted for, at least partly, providing a measure of the total uncertainty in the strong-motion variable. In design decisions it may be useful to be able to assess the sensitivity of the strong-motion variable to changes in basic variables. This can be achieved by error analysis.

To give an example of such analysis, let us take a ground motion model:

$$\log_{10}(a) = A + B\, M_S + C\, r + D\, \log_{10}(r) \tag{7a}$$

where

$$r = \sqrt{d^2 + h_0^2} \tag{7b}$$

Here, a denotes the larger horizontal component of peak ground acceleration (PGA) in g; d is source distance in km; h_0 is a depth parameter in km, and M_S is surface-wave magnitude. The data set applied contained earthquakes in the range $4.0 \leq M_S \leq 7.4$. The regression parameters are given as follows (Ambraseys, 1995): (1) for horizontal PGA not including focal depth, $A = -1.09$, $B = 0.238$, $C = -0.00050$, $D = -1$, $h_0 = 6.0$ and $\sigma = 0.28$; (2) for vertical PGA not including focal depth, $A = -1.34$, $B = 0.230$, $C = 0$, $D = -1$, $h_0 = 6.0$ and $\sigma = 0.27$; (3) for horizontal PGA including focal depth, $A = -0.87$, $B = 0.217$, $C = -0.00117$,

$D = -1$, $h_0 = h$ and $\sigma = 0.26$ and (4) for vertical PGA including focal depth, $A = -1.10$, $B = 0.200$, $C = -0.00015$, $D = -1$, $h_0 = h$ and $\sigma = 0.26$. This particular model is selected as its fits data from shallow strike slip earthquakes fairly well and is hence in accordance with other models considered in this study. The error in $\log_{10}(a)$ is readily derived as follows, taking the depth, h, as a variable:

$$\delta\left(\log_{10}(a)\right) = B\,\delta M_S + \left(C\frac{d}{\sqrt{d^2+h^2}} + D\frac{\log_{10}(e)\ d}{d^2+h^2}\right)\delta d + \left(C\frac{h}{\sqrt{d^2+h^2}} + D\frac{\log_{10}(e)\ h}{d^2+h^2}\right)\delta h \tag{7c}$$

where δd and δh are uncertainties in source distance and depth, respectively, and $e \approx 2.71828$. If the depth is selected as a fixed value the last term vanishes.

The results are shown in Figure 2, taking the depth as a constant, and in Figure 3 assuming the depth is a variable. It can be seen that the error induced by distance is greatest close to the epicentre and for shallow focus. By examining Figure 2, it is seen that the error in peak ground acceleration can be quite high even for realistic error in magnitude and distance. For instance, $M_S = 7 \pm 0.1$ and $d = 30 \pm 2$ km yield $\delta(\log_{10}(a)) = 0.055$, which results, approximately, in 13% error in peak ground acceleration. This indicates that quite large errors can be expected. Similar results are obtained by examining Figure 3.

3. Modelling of uncertainties

Up to now we dealt with uncertainties in strong-motion data and associated seismological parameters qualitatively, by describing the main pathological features encountered, and quantitatively, through regression analysis as well as error analysis. However, to quantify the uncertainties better, and reveal their statistical interrelations, a more thorough analysis is needed. A suitable tool for this type of analysis is furnished in uncertainty modelling (Ditlevsen, 1981), using the so-called reliability or performance index concept.

To be able to apply this methodology, we need a well-defined set of basic variables. Furthermore, we need a functional relationship relating the ground motion parameters to these basic variables. This is desirable only if this functional relationship is derived from the basic principles of mechanics and reflects all the main aspects and core ingredients needed for a theoretical description of the problem. The theoretical ground motion model adopted for our discussion here is described in some detail by Ólafsson (1999). It is based on the widely applied Brune spectra (Brune, 1970, 1971) and is found valid for shallow strike slip earthquakes with approximately circular faults, i.e., the thickness of the seismogenic zone is not a significant constraint. From a practical point of view one of the shortcomings of this model is that it contains many variables, some of which may be difficult to obtain. A way out of this is to use constraint optimisation to define the parameters, provided we have reliable data. As pointed out earlier, the lack of data prevents us

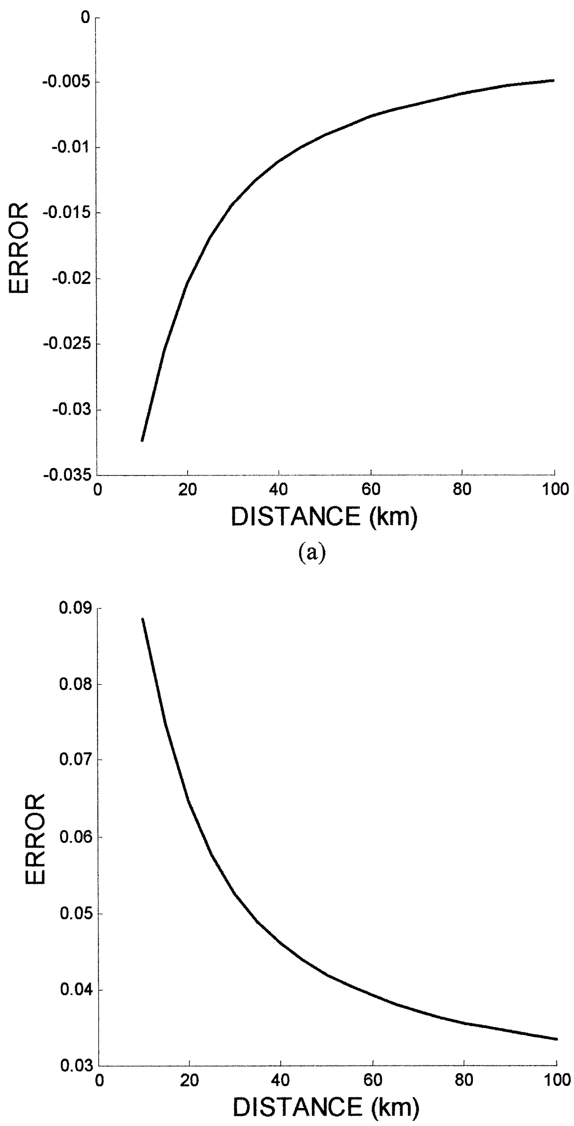
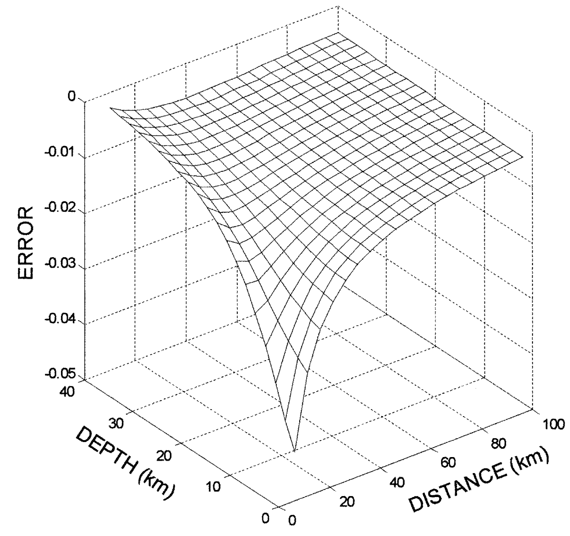
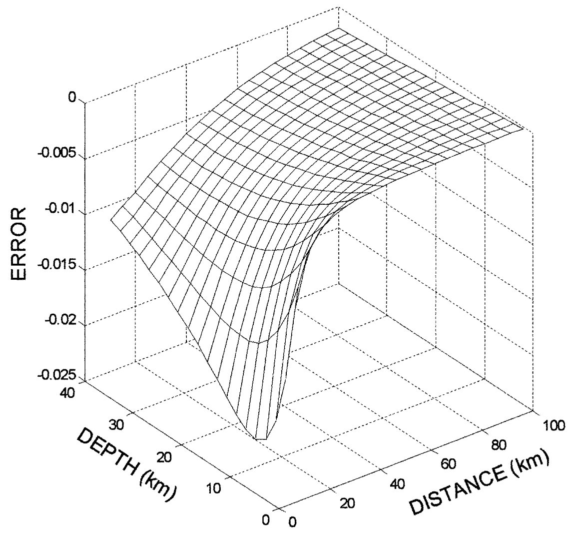


Figure 2. Error in $\log_{10}(a)$ according to Eq. (7c) as a function of epicentral distance for unit uncertainty in basic variables. Based on Eq. (7a) for horizontal peak ground acceleration not including focal depth as a variable but taking parameters $A = -1.09$, $B = 0.238$, $C = -0.00050$, $D = -1$, $h_0 = 6.0$ and $\sigma = 0.28$. (a) Error term proportional to δd, and (b) total error given by Eq. (7c) by taking the absolute value of individual terms and substituting $\delta d = 2$ km $\delta M = 0.1$ (and $\delta h = 0$).

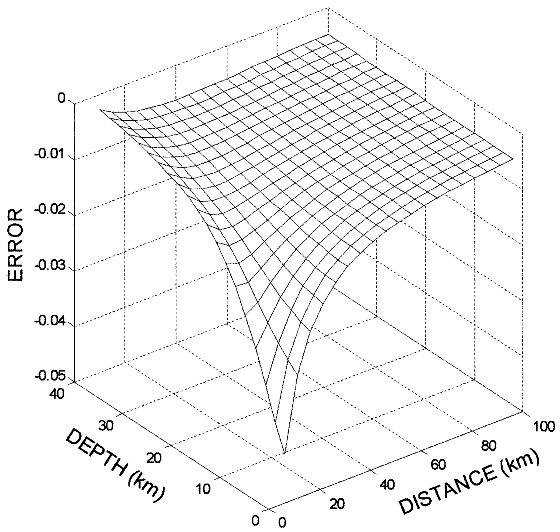


(a)

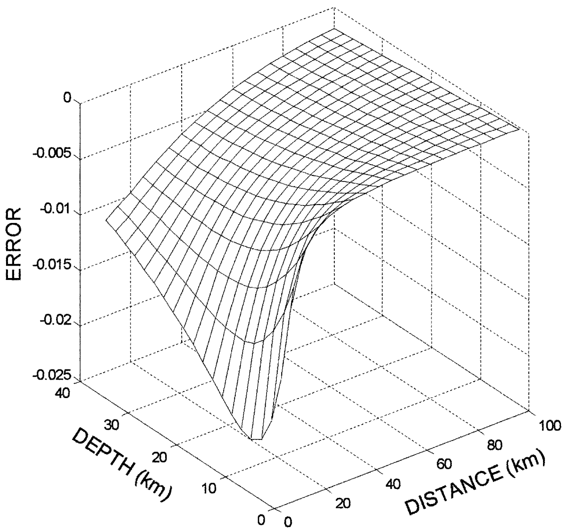


(b)

Figure 3. Error in $\log_{10}(a)$ according to Eq. (7c) as a function of epicentral distance and depth for unit uncertainty in basic variables. Based on Eq. (7a) for horizontal peak ground acceleration including focal depth as a variable and taking parameters $A = -0.87$, $B = 0.217$, $C = -0.00117$, $D = -1$, $h_0 = h$ and $\sigma = 0.26$ (a) error term proportional to δd, (b) error term proportional to δh, (c) error terms proportional to δd and δh, respectively, and (d) total error given by Eq. (7c) by taking the absolute value of individual terms and substituting $\delta d = \delta h = 2$ km and $\delta M = 0.1$.



(a)



(b)

Figure 3. (Continued).

from carrying this out, except for a limited number of special cases. Therefore, a less rigorous approach will be adopted, using statistical information when available and heuristic a priori assumptions otherwise. To facilitate the discussion in the following, we start with some preliminary definitions. This is done to illustrate the methodology rather than provide accurate statistical information.

3.1. Definitions and basic assumptions

It is assumed that the performance of a system subjected to (external) environmental disturbances, for instance, ground response or structural response due to earthquakes, can be described in general by a finite number of uncertain, measurable variables, termed basic variables:

$$\mathbf{X} = \left\{ X_1, \ X_2, \ X_3, ..., X_n \right\} = \left\{ M, \ R, \ ... \right\} \tag{8}$$

In the following, the basic variables are modelled as independent, stochastic variables. This implies that the variables are uncorrelated. Furthermore, it is assumed that the response or performance of the system can be described by a mathematical expression or function, called herein the response function of the system, exemplified in the following by the ground motion estimation equation That is:

$$f(\mathbf{X}) = 0 \tag{9}$$

We may consider this function as a hyper-surface in an n-dimensional space of the basic variables. The response hyper-surface divides the space into two regions, that is, a region where $f(\mathbf{X}) > 0$ and a region where $f(\mathbf{X}) < 0$, which we could call, respectively, the safe and the unsafe region if $f(\mathbf{X}) = 0$ describes some sort of a limit state behaviour. Furthermore, as the basic variables are assumed to be modelled as stochastic variables, the safe performance of the system can only be expressed in probabilistic settings, for example, as follows:

$$\Pr\left[f(\mathbf{X}) \geq 0 \right] = C_p \tag{10}$$

Here, $Pr[\cdot]$ denotes the probability (of performance), and C_p is a number quantifying this probability.

Within the framework of the reliability index approach (Ditlevsen, 1981), the basic variables, $\mathbf{X}$, are transformed to a normalised, Gaussian space, where the transformed variables, $\mathbf{u}$, are normally distributed with zero mean and unit standard deviation. Hence, the reliability index can be obtained as:

$$\beta = \min \ \mathbf{u}^T\mathbf{u} \quad \text{for} \quad \mathbf{u} \in \left\{ \mathbf{u} : f\left(\mathbf{u}\right) \right\} \tag{11}$$

where, $\mathbf{u}$ denotes the vector of basic variables in the normalised, Gaussian space and $f(\mathbf{u})$ is the corresponding response function. The point on the response surface with the highest probability density describes the 'most likely' performance of the system. For linear hyper-surfaces, it can be shown that (Ditlevsen, 1981):

$$\beta = \frac{E\left[f\left(\mathbf{X}\right)\right]}{D\left[f\left(\mathbf{X}\right)\right]} \tag{12}$$

where $E[]$ denotes the expectation operator and $D[]$ the standard deviation operator. Hence, the index can be interpreted as the number of standard deviations that we have between the mean value and the 'critical' boarder. Furthermore, we have:

$$\Phi\left(-\beta\right) = C_P \tag{13}$$

where, Φ denotes the standardised normal distributions. If the response surface is well behaved, we assume that the above model applies with a reasonable degree of accuracy, which is the case when the performance surface can be described with a hyper-plane in the vicinity of the performance point.

3.2. Probabilistic Modelling of Basic Variables

The basic variables used to describe the adopted strong-motion model can be selected in different ways. Without going into details, we have selected the following basic variables, which we judge applicable for the adopted ground motion model (see Ólafsson and Sigbjörnsson, 1999; Sigbjörnsson and Ólafsson, 2004): magnitude, distance to source, depth, fault radius, shear wave velocity, density, spectral decay and peak factor. We assume, for the time being, that these variables can be treated as independent stochastic variables, and that other variables are either treated as stochastic derivative variables or approximated as deterministic.

The first variable is magnitude, which was discussed in Section 2.1, including the problems that may arise due to the mixing up of different magnitude scales and temporal inhomogeneous data. In general, we adopt the surface-wave magnitude scale for reasons explained earlier.

We find that the uncertainty in $M_S$magnitude estimates tends towards normal distribution. However, the standard deviations obtained depend strongly on the number of stations as well as on their azimuthal distribution.

For Iceland and the Iceland Region we find that the mean value of the standard deviations for individual events is 0.24, which is close to the values obtained for events during the last decade for which we have recordings from many stations (Ambraseys, and Sigbjörnsson, 2000). Hence we can reasonably assume that the standard deviation of magnitude estimates is quite high.

The ground motion model used does not include magnitude as an explicit variable but instead the seismic moment, M_o. We therefore treat the seismic moment as a derivative stochastic variable.

For magnitudes of about 6, however, there is a small difference between the moment magnitude scale, M_w, and the surface-wave magnitude scale, M_S. In that case it is possible to use the Hanks-Kanamori relation (Hanks and Kanamori, 1979) to relate seismic moment and magnitude, i.e., $M_w = \frac{2}{3}\log_{10}(M_o) - 10.7$, if $M_w > 6$ (see also Section 2.1 for large events). These assumptions lead to the following distribution for the seismic moment:

$$p_{M_o}(M_o) = \frac{\log_{10}(e)}{\sqrt{2\pi}\sigma_{M_w} c}\frac{1}{M_o}\exp\left(-\frac{1}{2}\left(\frac{\frac{3}{2}\log_{10}\left(M_o/c\right) - \mu_{M_w}}{\sigma_{M_w}}\right)^2\right) \qquad (14)$$

where $c = 10^{1.5\times 10.7}$, $e \approx 2.71828$, μ_{M_w} and σ_{M_w} denote respectively the mean value and the standard deviation quantifying the normal distribution of the magnitude. It should be noted that the seismic moment is a positive quantity.

To quantify the uncertainty in the seismic moment, let us assume that the magnitude has a mean value of 7.0 and standard deviation 0.24, which gives a coefficient of variation equal to 0.0343. Substituting these values into Eq. (8) and integrating gives a mean value of the seismic moment equal to 5.00×10^{26} dyn·cm and standard deviation 4.97×10^{26} dyn·cm. This corresponds to a coefficient of variation of 0.99, reflecting the great uncertainty inherent in the seismic moment. In this context it is also worth pointing out the skewness of the distribution given in Eq. (8) emerging in a modal value (i.e., the most probable value) equal to 1.78×10^{26} dyn·cm and a median value of 3.55×10^{26} dyn·cm. This great variability is in accordance with our experience in computation of seismic moments from individual records obtained from the same event.

The epicentral distance is Rayleigh distributed (Ditlevsen and Madsen, 1996) if we assume the coordinates of the epicentre and site location to be normal distributed. For distances in the far-field, if the coefficient of variation is small, the distribution can be approximated as normal. In the intermediate field, however, the Rayleigh distribution is to be preferred. As pointed out earlier, the uncertainty in distance depends on the source of information, ranging from a few hundred meters up to several kilometres. For source distances of intermediate range, it appears that the uncertainty is commonly in the range 1 to 5 km.

Depth is not a well-defined quantity. From a theoretical viewpoint it is seen that the distribution of depth must be defined on a closed interval, i.e., ranging from close to the surface, to the thickness of the seismogenic crust. A distribution easily adopted to these constraints is the beta-distribution. In addition, it can take the form of a bell-shaped curve in cases where constraints from the boundaries are not significant. In such cases we believe that a normal approximation can be applied as well. In such cases we find that a coefficient of variation is in the range 0.1 to 0.2.

The fault radius is a quantity with great inherent uncertainty, which is difficult to quantify. It is also questionable whether the fault radius should be regarded as a basic variable, as it is functionally related to the seismic moment, shear modulus and the amount of slip. We will comment further on this problem later on. Here we treat the slip as a derivative variable, rather than the fault radius. By definition the fault radius must be greater than zero, furthermore, it seems natural to have some upper bounds on its length, for instance, half of the thickness of the seismogenic zone. Experience shows that the model adopted (see Sigbjörnsson and Ólafsson, 2004) can be applied with reasonable accuracy even for events with surface faulting as long as the fault length is not much greater than the thickness of the seismogenic zone. In these cases we can define the radius for an equivalent circular area equal to the size of the fault. This may lead to a radius exceeding half of the thickness of the seismogenic zone by 10% to 20%. In view of this we suggest that the fault radius can be modelled by beta-distribution that may be approximated by a normal distribution if there are no significant constraints from the boundaries.

The mechanical properties like shear wave velocity and density are, in most cases, well-defined quantities, which we assume can be modelled by log-normal distribution. When doing so, we must keep in mind that these quantities are by definition positive quantities. The other mechanical properties needed in the adopted model for this study, like shear modulus, are defined as a derivative variable, using the shear wave velocity and density.

The spectral decay parameter is applied to shape the high frequency tail of the acceleration spectrum. By definition this variable must be positive, to secure a bounded integral defining the rms-acceleration. As this variable shows clear normal tendency, it can be modelled by log-normal distribution. However, a normal approximation can also be adopted in cases where the values are far enough from zero.

The peak factor is last on the list of our suggested basic variables. It is defined in probabilistic settings as the ratio between the peak ground acceleration and the corresponding rms-value. The peak factor must hence be greater than 1. The distribution of this variable can be obtained using the theory of extremes (Vanmarcke and Lai, 1980). However, for simplification it appears acceptable to approximate its distribution as log-normal.

Other variables included in the adopted strong-motion model are treated as derivative variables or simply approximated as deterministic constants. It is important, also for the derivative variables, to investigate any physical boundary or constraints that may be connected with the variables. Furthermore, it is necessary, when the uncertainty analysis described above is exercised, to check whether the critical events can be regarded as physically realisable. This means that all variables must be within reasonable physical limits.

4. Applications to ground motion estimation models

The uncertainty analysis outlined above is readily applicable to peak ground acceleration and can be used to get estimates for the standard deviation that can be compared with the standard deviation of the residuals of the regression analysis. Therefore we can assess whether it is probable that complex models can reduce the uncertainty of the derivative variables.

To illustrate our point, we use a simplified model for strike-slip earthquakes, but exclude soil conditions commonly influencing site response. Hence we assume

Table III. Basic variables for the near-field model assumed normal distributed

Name	Symbol	Unit	Mean value	Coefficient of variation
Magnitude	M_w	–	6.6	0.02
Fault radius	r	km	7.5	0.07
Shear wave velocity	υ	km/s	3.5	0.10
Density	ρ	g/cm^3	2.8	0.07
Spectral decay	κ_o	–	0.02	0.12
Peak factor	p	–	2.94	0.15

that the peak ground acceleration induced by shear waves in the near-field can be approximated by the following expression (Sigbjörnsson and Ólafsson, 2004):

$$a = \frac{1}{\sqrt{\pi}} \frac{7}{8} \frac{p\, M_o\, C_p}{\rho\, \upsilon\, r^3 \sqrt{\kappa_o}} \left(\frac{\Psi_o}{T_o} \right)^{1/2} \tag{15}$$

where M_o is the seismic moment; υ is the shear wave velocity; ρ is density of rock; r is characteristic fault dimension (radius); κ_o is the spectral decay factor; p is the peak factor; Ψ_o is the dispersion function (for further details see Sigbjörnsson and Ólafsson, 2004); T_o is the duration, which can be related to source dimension and shear wave velocity, and C_p is a partitioning factor.

The selected basic variables are given in Table III, where mean values and coefficient of variance are also listed. As a first crude approximation all variables are taken to be normally distributed. Other variables are assumed to be derivatives or constants. This applies to the duration parameter $\mathrm{T}_o = 3{\cdot}r/2{\cdot}\upsilon$, the dispersion Ψ_o, and partitioning factor ($C_p = 1/\sqrt{2}$). Furthermore, the seismic moment is derived from the moment magnitude applying the Hanks-Kanamori relation.

The estimation of the standard deviation of the peak ground acceleration, using the approach outlined in Section 3.1, gives $\sigma = 0.239$ (corresponding to $\beta = 1$). This value is comparable with the standard deviations reported in the literature (Douglas, 2003a; Ambraseys and Douglas, 2003).

The sensitivity factors are shown in Figure 4. They represent the sensitivity of the standardised response surface at the performance point to changes in the basic variables (Ditlevsen and Madsen, 1996). A low sensitivity factor for a particular basic variable indicates that there is not a great need to increase statistical information about this variable. This may even suggest that this variable can be treated as deterministic rather than as a stochastic. For the data presented in Figure 4, the sensitivity factors indicate that the shear wave velocity, density and spectral decay could perhaps be treated as deterministic. The figure also shows that the improved statistical information is especially beneficial for the magnitude and the source radius. On the other hand, reduction of the uncertainties of these variables will increase the sensitivity factors of the other variables and thereby their importance.

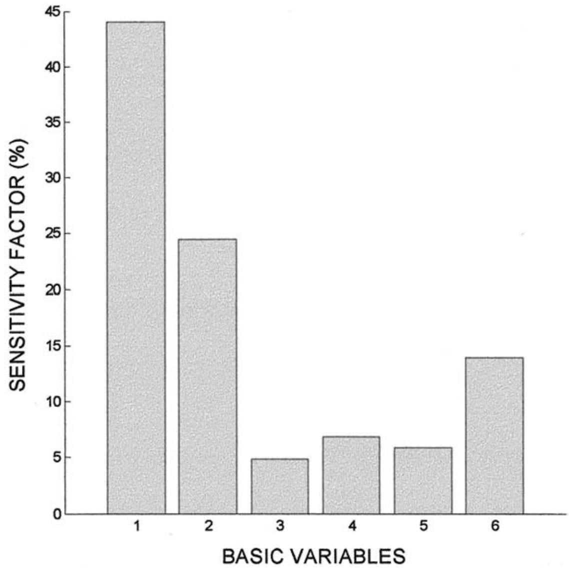


Figure 4. The relative contribution of the basic variables to the standard deviation. Denotation of the basic variables: 1 - magnitude, 2 - fault radius, 3 - shear wave velocity, 4 - density, 5 - spectral decay, and 6 - peak factor.

As a further illustration we may assume that the peak ground acceleration, a, induced by shear waves in the far-field, which we assume can be approximated as follows (Ólafsson and Sigbjörnsson, 1999):

$$a = \frac{\left(2\sqrt{7}\right)^{2/3}}{2\sqrt{\pi}} \frac{p\, C_P\, R_{\theta\phi}\, \Delta\sigma^{2/3}}{v\, \rho\, \sqrt{\kappa}} \sqrt{\frac{\Psi}{T_d}} \frac{M_o^{1/3}}{R} \tag{16}$$

Here the following notation is used: M_o is the seismic moment; R is the source distance; v is the shear wave velocity; ρ is density of rock; $\Delta\sigma$ is the stress drop related to characteristic fault dimension (radius), r (see Ólafsson, 1999); κ is the spectral decay factor; p is the peak factor; Ψ is the dispersion function (Ólafsson and Sigbjörnsson, 1999), T_d is the duration, which can be related to source dimension, shear wave velocity and distance (Vanmarcke and Lai, 1980), and C_p is a partitioning factor ($C_p = 1/\sqrt{2}$)

The selected basic variables are listed in Table IV. We assume at this stage that all the basic variables can be approximated by normal distributions with the parameters given in Table IV. Other variables are evaluated as derivatives using the formulae given in Ólafsson and Sigbjörnsson (1999); Ólafsson (1999), and Sigbjörnsson and Ólafsson (2004).

The obtained standard deviations of the peak ground acceleration for distance to fault equal to 10 and 50 km are 0.227 and 0.237, respectively ($\beta = 1$). The sensitivity factors are shown in Figure 5. The uncertainty assigned to magnitude is the greatest single contribution, as was the case for the near-field data shown in

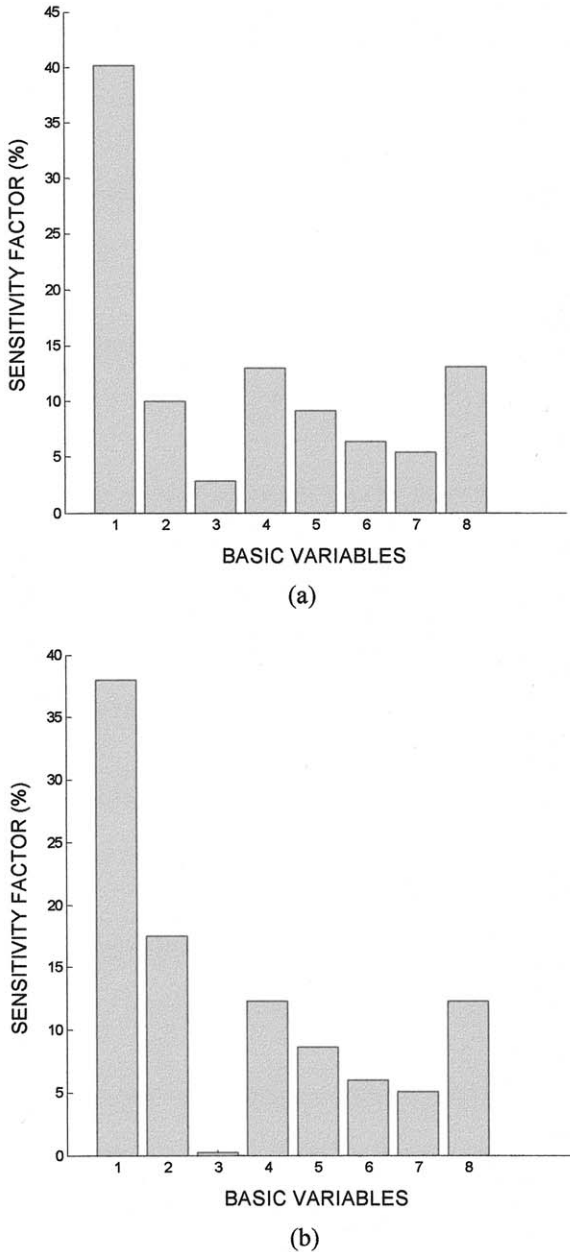


Figure 5. The relative contribution of the basic variables to the standard deviation. Denotation of the basic variables: 1 - magnitude, 2 - distance to fault, 3 - depth, 4 - fault radius, 5 - shear wave velocity, 6 - density, 7 - spectral decay, and 8 - peak factor. (a) Distance to fault 10 km, (b) distance to fault 50 km.

Table IV. Basic variables for the far-field model assumed normal distributed

Name	Symbol	Unit	Mean value	Coefficient of variation
Magnitude	M_w	–	6.6	0.02
Distance to fault	D	km	variable	0.15
Depth	h	km	9.0	0.10
Fault radius	r	km	7.5	0.07
Shear wave velocity	v	km/s	3.5	0.10
Density	ρ	g/cm^3	2.8	0.07
Spectral decay	κ	–	0.02	0.12
Peak factor	p	–	2.94	0.15

Figure 4. It is worth pointing out that the sensitivity factor for depth decreases from 5% for a 10-km source distance to almost zero for a 50-km distance. This seems in accordance with previous experience and expectance. Furthermore, this implies that for greater (epicentral) distances the effect of depth is negligible.

These results do not point towards reduction in uncertainty in the peak ground acceleration, even if the adopted model accounts for more parameters than are usually used in regression models. Therefore, it seems that complex models are not likely to decrease uncertainty. On the other hand, the complex model can explain the sources of uncertainties, and how they contribute to the uncertainty of the derivative response variable, better than can be done using models with few parameters. This study indicates in particular that the main source of uncertainty is attached to magnitude and its inherent uncertainty. The most obvious remedy to reduce uncertainty appears to be enhancement of magnitude determination or perhaps seismic moment.

5. Applications to earthquake hazard assessment

In the foregoing it was assumed that all basic variables followed a bell-shaped distribution, which conforms to the results of the regression analysis. These distributions could be approximated by normal distribution, at least in cases where ‘tail sensitivity’ was not of importance. When applying the ground motion estimation equation in hazard and risk assessment, the distribution of magnitude and distance have to be redefined to reflect the statistical properties of the seismogenic area to be the studied. In this context the normal distribution and related bell-shaped distributions are clearly not applicable.

The distribution of epicentral distances for a particular site is derived from the distribution of epicentres in terms of their geographical coordinates. For a line source with uniform seismic activity, it is common to treat the epicentres as uniformly distributed along the line. For an area source, on the other hand, it is usual to assume the epicentres uniformly distributed within the area. When assessing

Table V. Basic variables for the near-field model describing a predefined source with assumed distributions

Name	Symbol	Unit	Distribution	Mean value	Coefficient of variation
Magnitude	M_w	–	seismicity dependent[1]	–	–
Fault radius	r	km	normal	conditional[1]	0.07
Shear wave velocity	v	km/s	normal	3.5	0.10
Density	ρ	g/cm^3	normal	2.8	0.07
Spectral decay	κ_o	–	normal	0.02	0.12
Peak factor	p	–	normal	2.94	0.15

[1] See text

the distribution for fault distance, information on fault size and fault orientation is required in addition to the distribution of epicentres. In both cases this leads to distribution for distance that deviates significantly from the bell-shaped normal type distributions.

The magnitude distribution is commonly assumed to be of the exponential type, often mapped on a closed interval ranging from the magnitude of the smallest earthquakes, judged to have significant effect on structures, to the magnitude of the largest earthquake that can credibly originate within the seismic zone in question. In addition, we need the number of earthquakes originating within each seismic zone that belong to the predefined magnitude interval. This leads to a magnitude distribution that is not of the bell type assumed in the previous sections.

The remaining basic variables can in principle be assumed to follow the distributions discussed earlier. Some of the parameters of these distributions may on the other hand be dependent on the earthquake magnitude. This applies especially to the size of the fault, which cannot be treated as independent of magnitude. In view of this it can be argued that fault size should not be regarded as a basic variable. However, it is possible to overcome this difficulty by using the principles of conditional distribution. Other basic variables, discussed in Section 3.2, appear to fulfil the requirements of independence and will therefore be retained as basic variables in the hazards assessment.

The methodology described above has been used to obtain hazard curves. Two cases are considered. The first case refers to site in the near-field and the second one to a far-field location. The variables adopted and the corresponding distribution parameters conform to those used earlier as discussed above. The data are summarised in Tables V and VI. In both cases we use a single seismic source area defined as follows: Line source of length 50 km with uniformly distributed epicentres, parameters of the magnitude distribution, $M_{min} = 4$, $M_{max} = 6.3$, a = 10 and b = 1.57. The site is at the middle of the fault with the shortest distance to fault equal to 10 km. In the case of the near-field model only events with epicentres within

Table VI. Basic variables for the far-field model describing a predefined source with assumed distributions

Name	Symbol	Unit	Distribution	Mean value	Coefficient of variation
Magnitude	M_w	–	seismicity dependent[1]	–	–
Distance to fault	D	km	source zone dependent[1]	–	–
Depth	h	km	normal	9.0	0.05
Fault radius	r	km	normal	conditional[1]	0.07
Shear wave velocity	v	km/s	normal	3.5	0.10
Density	ρ	g/cm^3	normal	2.8	0.07
Spectral decay	κ_o	–	normal	0.02	0.12
Peak factor	p	–	normal	2.94	0.15

[1] See text

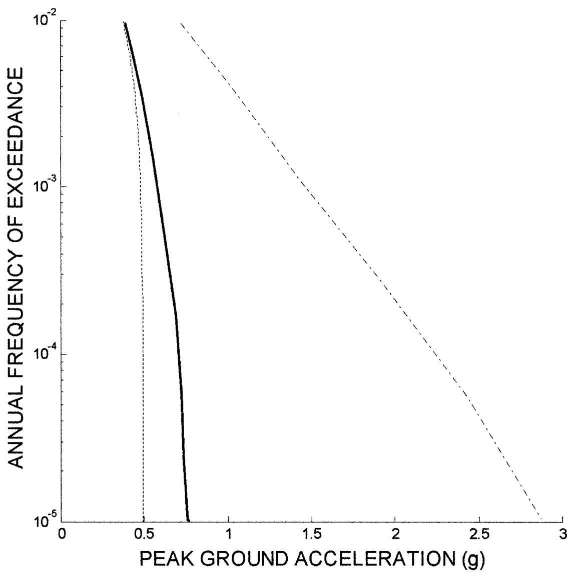


Figure 6. Hazard curve for peak ground acceleration derived using the near-field model. Properties of seismic source area are described in the text. The following notation is used: dotted curve based on mean values neglecting uncertainty, solid curve based on suggested uncertainty modelling and data in Table V, dash-dot curve based on mean values and the introduction of residual distribution for the peak ground acceleration with $\sigma = 0.25$.

radius of 12 km are considered. In the case of the far-field model all earthquakes described by the source are included.

The results are given in Figures 6 and 7. It is seen that there is a great difference between the hazard curves obtained, using mean values neglecting uncertainties, on the one hand (dotted curve), and the hazard curve we get by using traditional methods based on mean values but additionally introducing uncertainties in the

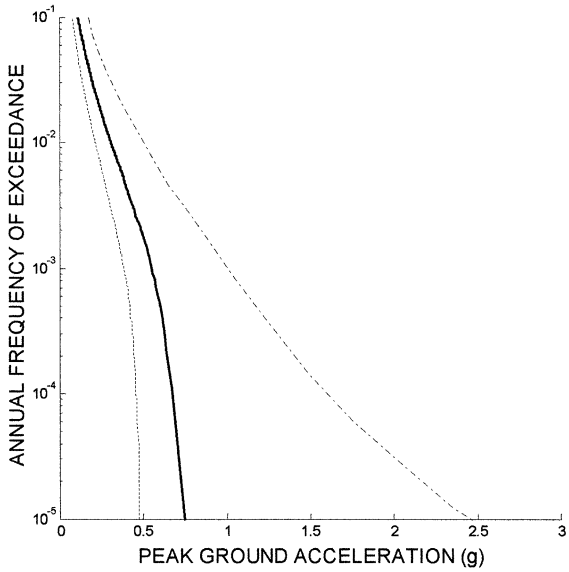


Figure 7. Hazard curve for peak ground acceleration derived using the far-field model. Properties of seismic source area are described in the text. The following notation is used: dotted curve based on mean values neglecting uncertainty, solid curve based on suggested uncertainty modelling and data in Table VI, dash-dot curve based on mean values and the introduction of residual distribution for the peak ground acceleration with $\sigma = 0.25$.

form of residuals of the peak ground acceleration (dash-dot curve), especially for small exceedance probabilities. The results obtained using the suggested uncertainty model are shown by the solid curve. It is seen that this curve shows roughly the same behaviour as the 'mean' curve, the dotted one. Furthermore, it is seen that the 'mean' curve has an upper bound, while the hazard curve, derived using the un-truncated residual distribution is not bounded as the probability of exceedance approaches zero. (dash-dot curve), On the other hand, this appears to be the case for the presented uncertainty model (solid curve). The main advantage of this model is that it produces values that seem to be sensible and in accordance with experience as far as they can be inferred.

In this context we note that extending the hazard curves to probabilities of the order 10^{-6} or 10^{-7} may have some formal meaning in statistics. Such low probabilities may reflect also the level of formal risk that the designer is willing to accept. On the other hand they do not say much when we address the real physical problem of regional continental seismicity. We know many regions, which have been active during the last few hundreds of years, and which border faults ceased to be active 10^3 to 10^4 years ago; the reverse is also true. In the time scale of more than about 10^5 years, regional seismicity is predominantly itinerant, and return periods of the order 10^6 to 10^7 are extremely judgemental in nature. Statistical extrapolations from 20th century data have little validity for periods of this great length.

6. Discussion and conclusions

Uncertainties in strong-motion recordings and associated seismological parameters have been discussed both qualitatively and quantitatively. Emphasis has been put on shallow strike-slip earthquakes and peak ground acceleration. It is found that the uncertainties in the peak ground acceleration can be explicitly related to uncertainties in a few basic variables, making it possible to quantify how much each basic variable contributes to peak ground acceleration or response spectra. It is found important to select basic variables that are statistically independent. If that is not possible, a transformation of dependent variables into independent is recommended, for instance, by using Rosenblatt transformation (Rosenblatt, 1952; Ditlevsen and Madsen, 1996). In some cases conditional distributions can be applied to simplify this process.

It is found that increasing the number of variables in the ground motion model for peak ground acceleration does not apparently decrease the standard deviation of the residuals. This is due to the intrinsic uncertainty of the basic variables. It is also found that the magnitude (surface- or moment-) contributes most to the uncertainty in peak ground acceleration. It seems, therefore, that a method to reduce the uncertainty in magnitude or seismic moment is a remedy that would be very beneficial. The advantages of applying seismic moment are widely appreciated. However, we find that the seismic moment has inherently great uncertainty, furnished in a skewed statistical distribution with a standard deviation of the same order of magnitude as the mean value. Reduction of this uncertainty would have positive effects on the residuals. Therefore, sophistication of the ground motion estimation model by the inclusion of additional independent parameters, such as the source dimension, stress drop, and seismogenic thickness, to mention a few, might be desirable if the dataset were adequate to determine reliably their influence on strong-motion. However, even if this were the case, this then places the onus on the engineer to assess a priori parameters, which are difficult enough to assess even after an earthquake.

Even though multi-parameter analytical ground motion models as put forward in this study do not reduce the inherent uncertainty in ground motion, they are found useful in the analysis of uncertainties as they make it possible to quantify, to a certain extent, the contribution of individual parameters to the overall uncertainty in strong-motion variables, like peak ground acceleration and response spectra.

We find that it is of importance to account for source mechanism when deriving ground motion estimation equations. We suggest that the source mechanics should be added as a ‘parameter’ to the parameters commonly in use like magnitude, source distance and soil conditions. Available data already make this feasible.

The assessment of seismic hazard involves M_S in both constituent functions, i.e. in the ground motion estimation equation as well as in the magnitude-frequency distribution. The former function is based on observations derived from data covering a long period of time. From the preceding it is obvious that the uncertainty in M_S is significant not only for the assessment of strong-motion estimates for modern earthquakes, say of the last three decades, but more so for earlier events for which the standard error in event magnitude M_S rises to 0.35. For historical events, whose M_S is estimated from semi-empirical scaling laws, σ values may reach 0.5 or more. Thus the uncertainty in M_S generally increases as we go back in time, particularly for the more rare, but important large early events, which plot near where the magnitude-frequency distribution curve steepens.

The application of the presented models in seismic hazard analysis makes the treatment of uncertainties more realistic than in the traditional approach. Results obtained in traditional hazard assessment are sensitive to the truncation of the error term commonly given as an integral part of ground motion estimation equations. The presented approach does not suffer from this shortcoming and yields apparently reasonable hazard curves without introduction of artificial constraints. The benefit of being able to assess hazard without having to invoke arbitrary truncation limits, is obvious.

The derivation of ground motion estimation equations that follows political or national boundaries are found not to be desirable and, in principle, without scientific foundation. We recognise, however, that there may be differences between seismic regions and possibly also seismogenic zones, even though this difference may not be very significant for the near-source areas, which are often of highest importance for engineering design. Limitations in data, however, are the main obstacle to practical use of such information.

Acknowledgements

This work was carried out by the writers with the support of EPSRC, Grant/52114, entitled Earthquake spectra and real time ground motions for design purposes. The authors are indebted to Dr J. Douglas for his valuable comments. Furthermore, the suggestions and corrections from two anonymous reviewers enhanced the quality of the paper.

References

Ambraseys, N.N. (1995) The prediction of earthquake peak ground acceleration in Europe, *Earthquake Engineering and Structural Dynamics*, **24** (4), 467–490.

Ambraseys, N.N. *et al.* (2002) *ISESD – Internet Site for European Strong-Motion Data*. http://www.isesd.cv.ic.ac.uk, European Commission, Research-Directorate General, Environment and Climate Programme.

Ambraseys, N.N. *et al.* (in press) Internet Site for European Strong-Motion Data, *Bollettino di Geofisica teorica ed applicata.*

Ambraseys, N.N. and Sigbjörnsson, R. (2000) *Re-appraisal of the seismicity of Iceland.* In: Polytechnica – Engineering seismology, Earthquake Engineering Research Centre, University of Iceland.

Ambraseys, N.N. and Douglas, J. (2003) Near-field horizontal and vertical earthquake ground motions, *Soil Dynamics and Earthquake Engineering*, **23** (1), 1–18.

Boore, D.M. (1983) Stochastic simulation of high frequency ground motion based on seismological modelsof the radiated spectra, *Bulletin of the Seismological Society of America*, **73**, 1865–1894.

Brillinger, D.R. and Preisler, H.K. (1985) Further analysis of the Joyner-Boore attenuation data, *Bulletin of the Seismological Society of America*, **75**(2), 611–614.

Brune, J. (1970) Tectonic stress and the spectra of shear waves from earthquakes, *Journal of Geophysical Research*, **75**, 4997–5009.

Brune, J. (1971) Correction. *Journal of Geophysical Research*, **76**, 5002.

Ditlevsen, O. (1981) *Uncertainty Modeling with Applications to Multidimensional Civil Engineering Systems.* McGraw-Hill International, 1981.

Ditlevsen, O. and Madsen, H.O. (1996) *Structural Reliability Methods*, John Wiley & Sons, Chichester, UK (Internet edition, May 2003, http://www.mek.dtu.dk/staff/od/books.htm).

Douglas, J. (2003a) Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates, *Earth-Science Reviews*, **61**(1–2), 43–104.

Douglas, J. (2003b) Use of analysis of variance for the investigation of regional differences in strong ground motions (submitted for publication).

Draper, N.R. and Smith, H. (1998) *Applied Regression Analysis*, Wiley Series in Probability & Statistics.

Ekstrom, G. and Dziewonski, A. (1988) Evidence of bias in estimation of earthquake size, *Nature*, **332**, 319–323.

Hanks, T.C. and Kanamori, H. (1979) A moment magnitude scale, *Journal of Geophysical Research*, **84**, 2348–2350.

Harward (2003) http://www.seismology.harvard.edu/

ISC (2003) http://www.isc.ac.uk/

Ólafsson, S. 1999: *Estimation of earthquake-induced response.* Department of Structural Engineering, Norwegian University of Science and Technology, Trondheim.

Ólafsson, S. and Sigbjörnsson, R. (1999) A theoretical attenuation model for earthquake-induced ground motion, *Journal of earthquake engineering*, **3**, 287–315.

Ólafsson, S. and Sigbjörnsson, R. (2002) Attenuation of strong-motion in the South Iceland earthquakes 2000. *In: Proceedings of the 12th European Conference on Earthquake Engineering*, London, UK. Paper no. 123. Elsevier Science.

Rosenblatt, M. (1952) Remarks on a multivariate transformation, *Ann. Mathematical Statistics*, **23**, 470–472.

Sigbjörnsson, R. and Ólafsson, S. (2004) On the South Iceland earthquakes in June 2000: Strong-motion effects and damage, *Bollettino di Geofisica teorica ed applicata* (in press).

Vanmarcke, E.H. and Lai, S.P. (1980) Strong motion duration and rms amplitude of earthquake records, *Bulletin of the Seismological Society of America*, **70**, 1293–1307.