2016-1

Joint ICTP/IAEA Advanced Workshop on Earthquake Engineering for Nuclear Facilities

30 November - 4 December, 2009

A Procedure for Probabilistic Scenario Based Seismic Risk Analysis How to develop meaningful scenarios?

Jens-Uwe Klügel

Kernkraftwerk Goesgen-Daeniken
Daeniken
Switzerland

Joint ICTP/IAEA Advanced Workshop on Earthquake Engineering for Nuclear Facilities

30 November- 4 December 2009

Kernkraftwerk Gösgen

A Procedure for Probabilistic Scenario-Based Seismic Risk Analysis – How to develop meaningful scenarios?

Dr. Jens-Uwe Klügel
Kernkraftwerk Goesgen-Daeniken, Kraftwerkstrasse, 4658 Daeniken, Switzerland

Member of the Seismological Society of America

A Procedure for Probabilistic Scenario-Based Seismic Risk Analysis – How to develop meaningful scenarios?

Dr. Jens-Uwe Klügel

Kernkraftwerk Goesgen-Daeniken, Kraftwerkstrasse, 4658 Daeniken, Switzerland

ABSTRACT: Traditional PSHA studies do not provide the input required for a probabilistic risk assessment, because their main result – the uniform hazard spectrum (UHS) is made up of earthquakes of different damaging effects. Therefore, internationally the development of scenario-based methods is observed. The lecture discusses different alternatives in use for the development of probabilistic earthquake scenarios. An advanced method based on direct development of earthquake scenarios from the seismo-tectonic, geological and geotechnical characteristics of the region of interest avoiding the need to construct uniform hazard spectra is described in detail. Its practical implementation is illustrated by two numerical examples. The proposed methodology is closely linked to the neodeterministic method of hazard analysis which is expanded for use in seismic risk analysis. The scenario-based methodology is strictly based on observable facts and data and complemented by physical modelling techniques, which can be submitted to a formalized validation process. By means of sensitivity analysis, knowledge gaps related to lack of data can be dealt with easily, due to the limited amount of scenarios to be investigated. The proposed seismic risk analysis can be used with confidence for planning, insurance and engineering applications.

1 Introduction

Earthquakes, as many other natural disasters, have both immediate and long-term economic and social effects. Seismic risk analysis traditionally has been based on the methodology of probabilistic seismic hazard analysis (PSHA) as developed by Cornell (1968), McGuire (1976,1995) and expanded for the treatment of uncertainties by using expert opinion (*S*enior *S*eismic *H*azard *A*nalysis *C*ommittee - SSHAC, 1997). The method has also been expanded to develop scenario-earthquakes by disaggregation of uniform hazard spectra (UHS) into distinct hazard bins – pairs of magnitude and distance - each of them reflecting a certain scenario-earthquake. Unfortunately the selection of scenario-earthquakes from disagregation of the UHS is not unique. It depends on the selection of the ground motion parameter as well as on the disaggregation method. Wang (2005) argued that a probabilistic seismic hazard analysis, as practiced today, leads to the loss of physical meaning in the results and provides the decision maker with an infinite choice for the selection of a design basis earthquake. Klügel (2005e) demonstrated that the results of a probabilistic seismic hazard analysis, presented as a uniform seismic hazard spectrum, do not provide the required input for a seismic probabilistic risk assessment (PRA), as required for risk informed regulation in nuclear technology. Therefore, the need for development of more adequate methods arises. Starting from a description of the limitations and mathematical drawbacks of traditional PSHA-methodology the lecture presents an improved methodology for seismic risk analysis which ensures a close and consistent link to traditional methods of deterministic seismic hazard analysis. Two numerical examples are given using the mathematical model of bivariate exponential distributions of Gumbel type II both for the temporal distribution of earthquake occurrence as well as for the spatial distribution of seismicity.

2 Limitations of traditional PSHA methods

The term traditional PSHA refers here to the Cornell-McGuire PSHA model as used as the baseline PSHA-model in the U.S.A. This model is based on the following assumptions:

- Seismicity is represented by a set of independent sources, S, each with spatially homogenous seismicity, sources can be represented by areal sources or linear soruces (faults)
- Ground motion prediction is characterised by a function $g(m,r)$ that yields the mean value of the ground motion parameter $\ln(a)$, given the magnitude, and the distance, r, of the event.
- The model assumes that earthquake events of “practical interest”, i.e. those above a predefined magnitude threshold, occur with a mean annual rate ν_i (source-specific, index i indicates the source)
- It is assumed, that these events occur at relative frequencies $f_M(m)$ with the complementary cumulative distribution function $G_M(m)$, which is the fraction of events with magnitude *m* or greater. The common assumption is that the form of $f_M(m)$ is exponential. As a further even stronger assumption it is introduced, that $f_M(m)$ is to be derived from the Gutenberg-Richter and therefore depends only on the b-values of this equation.

The methodology is based on the use of a stationary Poisson process for earthquake recurrence, stating that the exceedance probability for the reccurrence of an earthquake exceeding site ground motion level **z** can be computed as:

$$P[A>z,t]=1-e^{\lambda(z)t} \tag{1}$$

with

$$\lambda(z) = \sum_{i=1}^{S} \nu_i \iint \Phi_i^{'} \left(\ln(z) - g(m,r) \Big/ \sigma \right) f_{R,i}(r|m)\, f_{M,i}(m)\, dr dm \tag{2}$$

The equation (2) simply means that the contributions of different seismic sources to the frequency of occurrence of a site ground motion level exceeding **z** are treated as independent and additive.

Another mathematical formulation of equation (2) obtains the following format:

$$\lambda(S_a > z) = \sum_{i=1}^{n_{source}} N_i(M_{\min}) \int_m \int_R \int_\varepsilon f_{mi}(m) f_{Ri}(r,m)\, f_\varepsilon P(S_a > z | m, R, \varepsilon)\, d\varepsilon dR dm \tag{3}$$

The difference to equation (2) consists in the explicit use of spectral accelerations as the ground motion parameter (S_a) and in the generalised form of the distribution of the error term ε, which in (2) is assumed to be a normal distribution with a zero mean and the standard deviation σ, while in (3) another distribution can be selected. Additionally the truncation of the hazard at a lower boundary of $M_{\min}$ is expressed explicitly. The error term ε has its origin in the (empirical) ground motion attenuation model:

$$\log(S_a) = g(m, r, X_i) + \varepsilon \tag{4}$$

with the error term expressed as the multiple of the standard deviation $\varepsilon = a\sigma_{\log}$ (the use of decadal or natural logarithm is a matter of taste) describing the scatter of data used for the development of the attenuation model. X_i represent additional explanatory variables (or classification properties for the specific travel path from the seismic source to the site) of the attenuation model, which may or may not be considered in the model. Examples for these additional explanatory variables are:

- site conditions (e.g. shear wave velocity, depth of surface layer),
- topographical and directivity effects,
- hanging wall and footwall effects,
- fault style,
- aspect ratio of the seismic source,
- material properties of the travel path of seismic waves

Equations (2) and (3) are derived under the assumption

- that magnitude and distance are sufficient to describe wave propagation from source to site (this could be adjusted by using a more complicated functional form for the function

$g(m,r)$ expanding the integrand by additional functional terms reflecting the conditional probability of exceeding the ground motion level **z**)

- that the error term $\varepsilon = a\sigma_{\log}$ can be treated as independent from the functional form of the attenuation equation (separation of variables approach).

The use of the second assumption is apparent in the form of the hazard integral, where the probability of exceeding a certain error ε expressed by a certain number of standard deviations a is evaluated unconditioned on m or r. Unfortunately this second assumption is incorrect and all derivations based on this assumption are mathematically flawed. $\sigma_{\log}$ is the measurement error associated with the indirect measurement (in form of the regression of a large amount of data points derived under various measurement conditions) of the decadal logarithm of the spectral acceleration S_a in dependence of magnitude m, distance between earthquake location r and possible other explanatory variables X_i. The evaluation of m and r itself represents another, earlier stage of indirect measurement. Therefore, the standard deviation in equation (4) has to be calculated from standard error propagation techniques as:

$$\sigma_{\log} = \sqrt{\left({\partial g}/{\partial x_i}\, \sigma_{x_i} \right)^2 + \sum_{i,j} \rho_{i,j} \sigma_{x_i} \sigma_{x_j} + \sigma_{res}^2} \tag{5}$$

Here x_i represents the set of explanatory variables used in the attenuation model $g(m,r,X_i)$ explicitly and σ_{res} represents the residual error associated with other possible (and currently unknown or not estimable) explanatory variables. In the ideal case of a complete model (all explanatory variables are known and can be estimated) σ_{res} equals the regression error associated with the regression method used itself (close to zero). Among the explanatory variables explicitly used there are the magnitude m and the distance r. Therefore, the error term $\varepsilon = a\sigma_{\log}$ is correlated with the error associated with the evaluation of magnitude and distance (and the other explanatory variables) and ε cannot be separated from the regression function $g(m,r,X_i)$. Furthermore, some regression equations contain even a direct dependency of the standard deviation on magnitude, e.g. Ambraseys at all (2005). The conclusion is that the form of the Cornell-McGuire model for PSHA that includes integration over the error term (this is based on the uncertainty model of Ang and Tang (Ang, 1970, Ang and Tang 1975;1984) combining aleatory variability and epistemic uncertainty into a joint probability distribution (Ayyub and Klir, 2006) for hazard level exceedance is mathematically incorrect and should not be used. Furthermore, the combination of different types of uncertainty (aleatory variability and epistemic uncertainty) into a joint probability distribution is equivalent to the assumption that the knowledge of analysts (seismologists) about the frequency of earthquake occurrence is changing the frequency of earthquakes itself.

The development of equation (2) was based on a heuristic bias. Originally PSHA assumed that the uncertainty of the problem is completely concentrated in the error term ε of the attenuation equation, regarding all other modelling parameters as exactly known. This assumption was the result of the division of labour between different groups of geophysicists. One group was responsible for the evaluation of earthquake magnitude (or intensity) and earthquake epicentre location, while another group used this information to develop attenuation models assuming earthquake magnitude and location as exactly known. Later on, the uncertainty of other modelling parameters were included in the analysis (e. g. of a and b parameters of the Gutenberg-Richter equation, and of the epicentre location). Unfortunately, people forgot that the evaluation of magnitude and earthquake location is based on measurements. Therefore, the obtained values are not known exactly and (in a probabilistic approach) have to be treated as random parameters. This means that the error term ε in the attenuation equation does include the measurement uncertainties (from indirect measurements) associated with the evaluation of magnitude and epicentre location (and the effects of other explanatory variables not explicitly considered in the attenuation equation). Therefore, an attenuation equation as it is applied in a PSHA (as it is used in the hazard integral) represents a multivariate distribution of the considered ground motion parameter expressing its dependence on a set of random model parameters. For replacing this multivariate distribution by the simplified model of a lognormal distribution (or normal distribution in log-scale) for use in a PSHA logic tree (here magnitude and distance are "exactly known" for each single path through the tree) it would have been required to consider the dependency between the model parameters and the error term ε or to adjust the residual error. Additionally, it is rather questionable, whether the error term ε can at all be interpreted as aleatory variability, because by its origin it simply defines different confidence levels for the empirical ground motion prediction equation (4). Therefore, it cannot be interpreted as inherent variability if earthquake ground motion. Using another functional shape (e.g. in another format rather than in logspace) of the ground motion prediction equation will change the error term (in units of absolue accelerations).

Furthermore, the result of PSHA using equations (2) and (3) is clearly driven by the number of standard deviations considered as the boundary condition for the integral over ε (or the number of standard deviations $\sigma_{\log}$). The number of standard deviations considered is in principle unlimited, although physical boundaries (e. g. maximal ground motion) can be provided. Therefore, the conclusion is that the hazard integral (2) or (3) can include infinitely high accelerations (this means that the hazard does not converge at all) or it will converge to a maximum ground motion level set by the analyst in advance. From Chebyshev's inequality

$$\Pr\left(\left|X - E\left(X\right)\right| \geq a\sigma\right) \leq \frac{1}{a^2} \tag{6}$$

in conjunction with (3) and (4) it follows directly that the results of a PSHA are driven by the recordings of statistically rare time-histories, which (due to the second ergodic assumption (Klügel, 2005c)) frequently were recorded under measurement conditions completely different from the site of interest. Furthermore, the link between the observation of spike accelerations and the associated observed damage is lost because the damage is usually not recorded.

3 Methodology of probabilistic scenario-based seismic risk analysis

The methodology of probabilistic scenario-based seismic hazard analysis in this paper represents an extension of the methods, which have been used for deterministic seismic hazard analysis in high seismic areas like California for more than 30 years (Mualchin, 1996). The extension specializes in the treatment of problems specific to seismic hazard analysis for low to moderate seismic areas, incorporates physical modeling approaches and introduces a sound methodology for risk assessment. It is worth to mention that this method is principally different from approaches attempting to develop scenario earthquakes by disaggregation of unifom hazard spectra. It avoids the unnecessary step of developing uniform hazard spectra and the associated loss of information. To show this we discuss the principal alternatives that are based on traditional PSHA.

3.1 Elements of seismic risk analysis

For some applications, such as for safety analysis of critical infrastructures or for insurance companies, it is necessary to perform a detailed risk analysis. Such an analysis can be beneficial to assess the efficiency of design measures as well as to identify potential vulnerabilities especially for existing facilities. Risk analysis therefore provides a meaningful complementary tool to traditional safety analysis and deterministic design procedures. It is a common but erroneous belief that *only* a probabilistic seismic hazard analysis (traditional PSHA) is able to provide the required input for a probabilistic risk assessment (PRA) for critical infrastructures. People often prefer to believe in names (such as "probabilistic" seismic hazard analysis) instead of analysing the essential points of a topic. Even in official technical standards (Budnitz et al, 2003), this wrong belief is common. Unfortunately, the question is not that simple and is worth investigating in more detail. A deterministic scenario-based seismic hazard analysis result is appropriate to perform detailed risk analysis, as demonstrated below.

The key elements of a risk analysis (Kaplan & Garrick 1981) are:

1. Identification of events that can occur and have adverse consequences
2. Estimation of the likelihood of those events occurring
3. Estimation of the potential consequences.

Therefore, the results of a risk analysis can be presented as a set of triplets:

$$R = \langle H_i, P_i, C_i \rangle \tag{7}$$

H_i represents the set of *i* events with possible adverse consequences

P_i represents the associated probabilities of their occurrence

C_i represents the associated intolerable consequences.

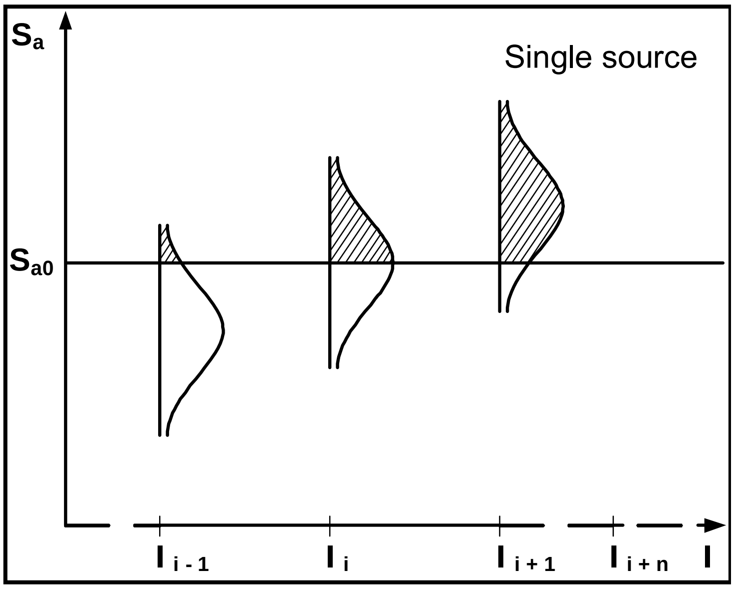
This means that a seismic hazard analysis shall provide the following information as an input for a probabilistic risk assessment (PRA):

- The events which may potentially endanger our infrastructure
- The frequency or probability of occurrence of these events.

The consequences of these events are evaluated by the risk model of the plant, which essentially represents a logic model mapping the hazards to be investigated to their consequences. What does a traditional PSHA provide? The standard output consists of a uniform hazard spectrum and a set of hazard curves, which represent the convoluted impact of a large amount or infinite (Wang, 2005) number of earthquakes with respect to the chances of causing certain level of ground accelerations at the site of interest. Therefore, traditional PSHA is not delivering the required frequency of events but exceedance probabilities of secondary properties. It is important to note that frequently damaging effects of an earthquake cannot be described by only a single secondary property (e. g. hazard curves expressed in terms of averaged spectral acceleration or even PGA). The impact of an earthquake event has to be described in the risk model of the plant, which can easily accommodate other impact effects besides the effects of acceleration (e. g. liquefaction, surface rupture below the basement).

3.2 Earthquake scenarios derived from UHS

Figure 1 illustrates the calculation process of an UHS (uniform hazard spectrum) in a



traditional PSHA for the case of a single source. The figure shows that traditional PSHA is computing the frequency of exceedence to a specific acceleration level independently from the damaging impact of the associated earthquakes. Therefore, a uniform hazard spectrum is not at all uniform with respect to the damaging effects of the earthquakes included into the hazard calculation.

Figure 1 Evaluation of an UHS in traditional PSHA, (single source, I – Intensity (damage index))

Meanwhile this problem is also understood by some engineering seismologists. Therefore different approaches have been developed to obtain scenarios from the results of a PSHA. A scenario-based seismic hazard analysis methodology is much better suited to provide the required and correct input for a seismic PRA because the physical impact associated with earthquake scenarios can easily be defined in subsequent engineering analysis.

The following approaches have been considered by different analysts.

1. Disaggregation of the uniform hazard spectrum into magnitude-distance bins.
2. Representing the uniform hazard spectrum by a set of equally weighted time-histories reflecting different parts of the spectrum (30 to 60 time-histories), therefore splitting of the energy content of the UHS into parts.

3. Subdividing the UHS into 2 or more conditional sub-spectra representing different frequency ranges of the spectrum (e.g. representing events with a dominant high-frequency response spectrum by one sub-spectrum (e.g. near-site events) and representing events with a dominant low frequency response spectrum (e.g. high magnitude distant events) by another sub-spectrum. The aggregation of the sub-spectra is assumed to match the total UHS spectrum. The sub-spectra are further decomposed to time-histories similarly as in approach #2.

The most meaningful approach is associated with the first approach. Essentially, it is an attempt to recover the lost information on the true damaging characteristics of earthquake scenarios aggregated into the uniform hazard spectrum. Figure 2 shows the process of developing scenario earthquakes from a traditional PSHA by disaggregation.

Nevertheless, the results of such a disaggregation represent merely mathematical artifacts and depend strongly on the mathematical assumptions used in the PSHA study. Therefore, it is always beneficial to compare and if necessary to correct the disaggregation results based on a comparison with the available geological and geomorphologic data available for the region. This approach was used for the update of the seismic PRA (Probabilistic Risk Assessment) of the nuclear power plant Gösgen, which is currently the largest seismic PSA study of the world. The approach was selected for legal reasons, because the development of an UHS in a PSHA was requested by officers of the Swiss nuclear safety inspectorate. Otherwise the approach presented in the second part of this paper would have been used.

The attempt to represent the uniform hazard spectrum by a set of different time-histories (approach # 2) contains a systematic mathematical bias. This is easy to understand. The uniform hazard spectrum (compare eqs (2) and (3)) represents the weighted result of summing up the contributions of many (essentially an infinite number limited only by the numerical discretization of the hazard code used) possible earthquake scenarios. Therefore the split-off of the uniform hazard spectrum has to follow the distribution of weights as used in the original PSHA study. Instead of this a uniform distribution is applied, therefore the time histories selected do not represent the damaging effects of the earthquakes included into the uniform hazard spectrum correctly.

The attempt to split-off the uniform hazard spectrum into conditional sub-spectra with a subsequent representation of the sub-spectra by different time histories smoothes the problems associated with approach #2, but it still remains. Another problem remaining here (this is true also for approach #2) is that the sub-spectra do represent groups of scenarios and not single or enveloping earthquake scenarios. Furthermore for soil sites it is questionable whether an UHS can be split-off into a high and a low frequency range, because the response spectra of earthquakes frequently show a double peak character. See for example recordings from the Beznau site in comparison with the PEGASOS UHS spectrum, scaled to the same level of peak ground accelerations, shown in figure 3. Although some attempts exist to treat this effect by developing correlation models (Baker & Cornell, 2006) for the spectral accelerations at different periods, this hardly will solve the problem. It simply increases the dependency of the results of the hazard analysis on the diverse mathematical assumptions made.

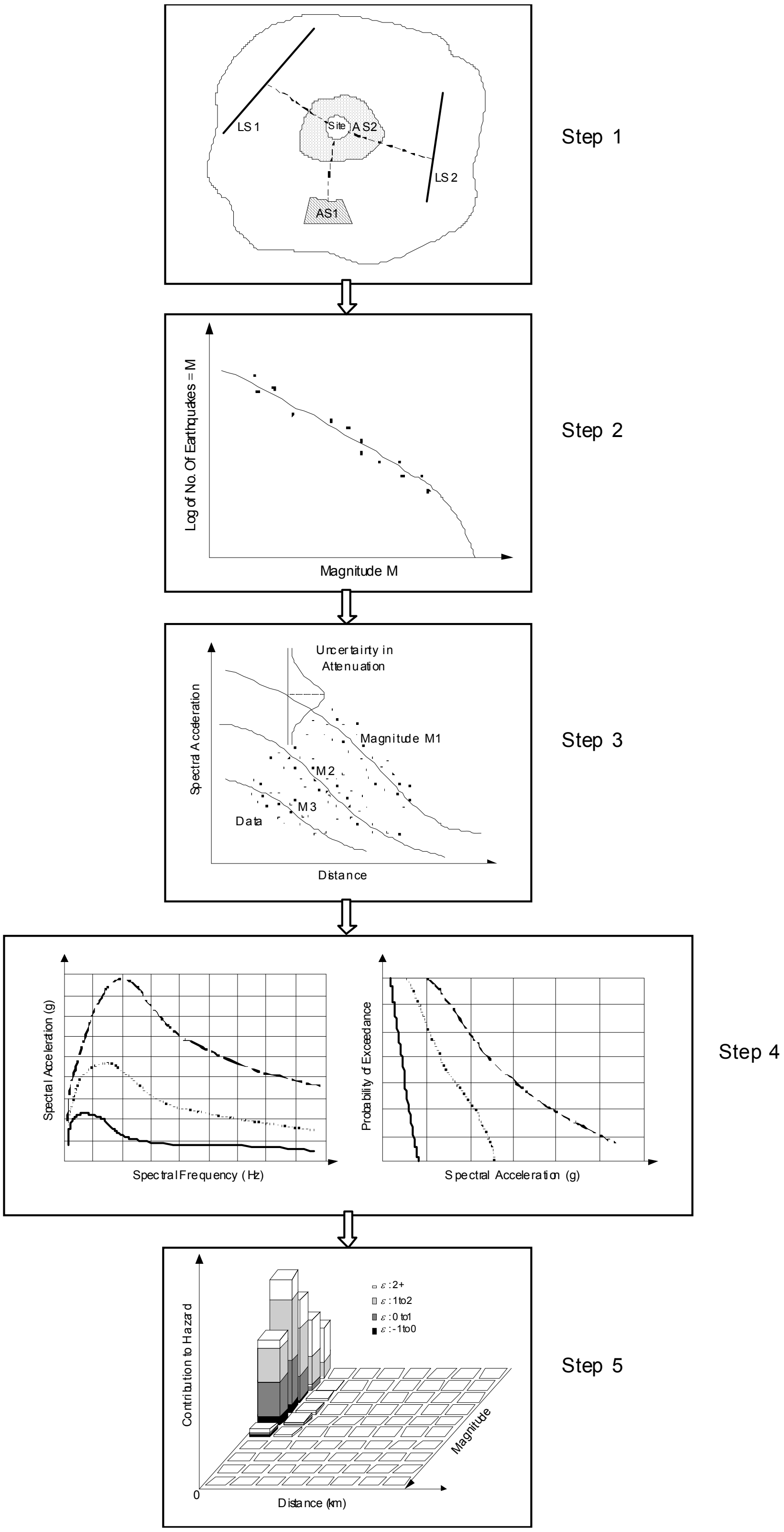


Figure 2 Workflow of a traditional PSHA including the development of scenario earthquakes from disaggregation

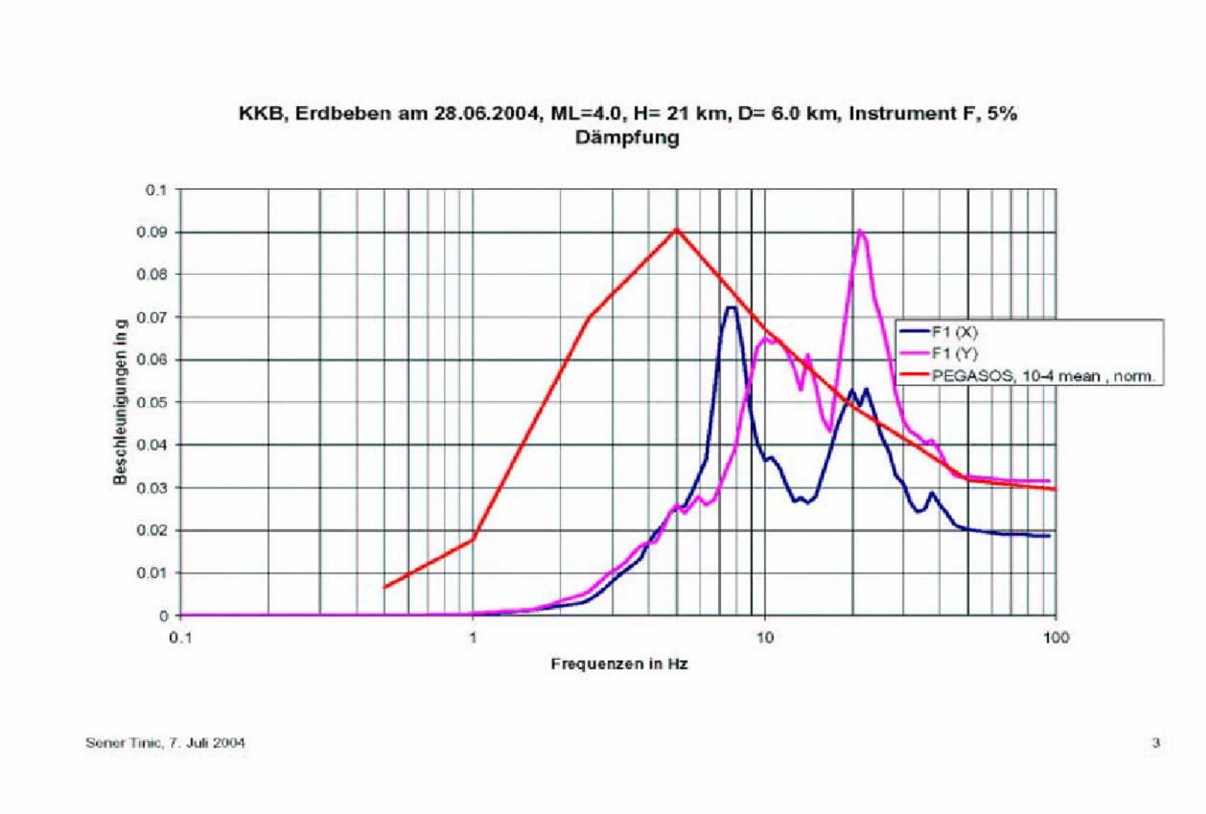


Figure 3 Earthquake recording at the site of the NPP Beznau scaled to the level of the PEGASOS pga level (10^{-4}/a)

3.3 Concept of direct scenario-based seismic risk assessment

A principal alternative to the development of earthquake scenarios from UHS consists in the development of earthquake scenarios directly from the seismo-tectonic model of the region (Klügel et al, 2006). This approach is usually characteristic for the neodeterministic seismic hazard analysis (NDSHA) which can be expanded for the purpose of probabilistic risk assessment.

Scenario earthquakes as developed by a neodeterministic analysis essentially represent the hazard events to be considered in the risk study (see section 3.1). The key point is that these scenarios have to be accompanied by an adequate probabilistic data model allowing estimating their frequency of occurrence. This allows expanding their use for probabilistic seismic risk assessment. If necessary some simplifications are possible to avoid the computation of too many scenarios. For example, the frequency of smaller earthquake events can be taken into account in the calculation of the frequency of occurrence of the stronger scenario earthquakes which envelope the impact of smaller events, by using a classification system. This corresponds exactly to how probabilistic risk assessments of nuclear power plants are performed for other initiating events (IAEA (1995), IAEA (2002), DOE (1996), Tregoning et al, (2005), Poloski et al, 1999). A nuclear power plant has, for example, a large amount of pipes in the reactor coolant circuit, which potentially could break causing a loss of coolant accident (LOCA) inside the reactor containment. The calculation of all possible scenarios associated with each single possible pipe break is not possible. Therefore, pipe breaks causing similar consequences are combined together and modeled by an enveloping, conservative, accident scenario. The frequency of the scenario is calculated as the sum of the frequencies of all underlying pipe breaks, assigned to the same class (e.g., small break LOCA, medium break LOCA, large break LOCA, etc.). The same approach is used in PRA for airplane crash analysis. Airplanes are classified by their impact characteristics and the risk contribution of airplane crashes is calculated as the sum of the contributions of each of the classes. The frequency assigned to each of the classes is developed from real data of airplane crashes and represents the total frequency of all crashes of airplanes belonging to the considered class.

The selection of one or a limited set of scenario earthquakes is the central concept of the methodology. The selection of scenario earthquake(s) includes the following steps:

- Characterisation of seismic sources for capacity/potential and location
- Selection of hazard parameter(s) to characterise the impact of an earthquake on the infrastructure
- Development of an attenuation model for the parameter to derive the values of the parameter(s) at the site
- Incorporation of site effects, and near-field and potential directivity/focusing factors
- Development of a probabilistic model for earthquake recurrence based on data analysis and definition of the scenario earthquake(s) modelled in the risk study

3.4 Characterisation of seismic sources

The selection of scenario earthquake(s) requires a detailed analysis of all regional seismogenic or active seismic sources surrounding the site of interest and assessment of their capability and potential to produce earthquakes of a significant size. For this step, all available information shall be explored. Figure 1 shows the concept in a schematic way.

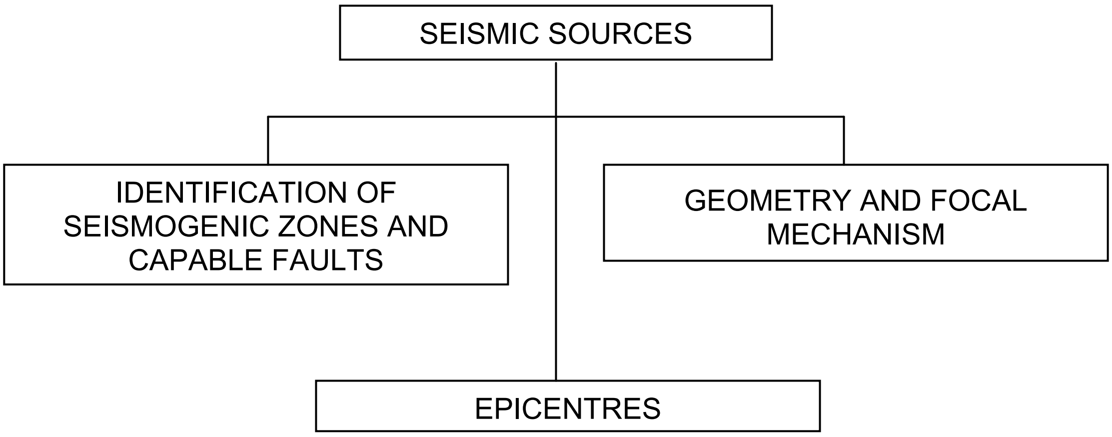


Figure 4 Information to characterize seismic sources

In the understanding of figure 4, a “capable fault” is a fault that has a significant potential for relative displacement at or near the ground surface.

In seismic highly active regions like California, the selection of seismic sources can be reduced to the identification and assessment of seismogenic faults, which can produce earthquakes of significant damaging potential. In less active regions and where instrumentally recorded earthquakes are not available, as is the case for several European areas, historical intensity data should be used to obtain an overall picture of the spatial distribution of the shaking intensity during written historical time. Although the epicentral locations and estimated magnitudes of historical earthquakes may not be as accurate as those of instrumentally recorded earthquakes, they can provide valuable, although incomplete, information on (1) the seismicity over long periods, (2) a rough delineation of seismic source zones and (3) reasonable estimates of future earthquake magnitudes, by assuming stable seismotectonic conditions for the region. It may even be possible to derive information on the frequency of large earthquakes, which are of interest for a scenario-based methodology, by time-series analysis. Different procedures for source modelling that elude source zones have been proposed. For example, one can make use of seismic parametric catalogues (historical and instrumental records) to define the possible locales of seismic events. This approach, called historical, has been widely applied in the past.

Other proposals based on the seismic catalogues are due to Veneziano et al. (1984), and Kijko and Graham (1998). In this context, Frankel (1995) also proposed a procedure using spatially-smoothed historical seismicity for the analysis of seismic hazard in Central and Eastern USA.

Woo (1996) suggested another procedure for areal sources statistically based on kernel estimation of the activity rate density inferred from regional seismic catalogue. Such approach considers that the form of kernel is governed by the concept of self-organised criticality and fractal geometry, with the bandwidth scaled according to magnitude. In general, the epicentre distribution of historical earthquakes gives a better indication of seismic zonation and generally leads to a non-uniform distribution of seismicity within the zone. Obviously, the most appropriate method suitable for the region of interest shall be selected based on the available data.

Another important item with respect to the characterization of the different seismic sources consists in the assessment of the maximum credible magnitudes to be considered in the analysis.

The size or magnitude of an earthquake can be estimated by several approaches. Fault length, area and displacement for known faults have been empirically correlated with moment magnitudes (Wells and Coppersmith, 1994). Improved correlations have been made possible by separating the data for different fault types. These relationships have been applied to seismogenic faults for estimating MCE magnitudes. An important assumption is the fault length used for MCE estimation (Mualchin, 1996). Empirical correlations for the assessment of earthquake magnitudes should not be applied outside the region they have been developed for. It should also be noted that fault mechanics (Scholz, 2002, p. 207) demonstrated different size regimes with respect to the scaling of moment and slip to the aspect ratio (length to width) of the source area, indicating different similarity regimes for earthquakes.

Correlations like Wells and Coppersmith (1994) are based on mixed data across these regimes and are compromise fits (Scholz, 2002). The use of mixed data can be a source of systematic error considered by some analysts as epistemic uncertainty. The different scaling regimes can be attributed, in part, to the way of propagation of the fault rupture. Seismic events with length less than the thickness of the brittle crust can propagate in all directions within a planar surface. Larger earthquakes, that rupture through the entire brittle crust (to the top of the ductile zone) can propagate farther only in the horizontal dimension. Thus, small and larger seismic events may be self-similar, but not to each other, and source scaling for interplate and intraplate tectonic regimes are different. Therefore, empirical correlations between magnitude and fault length should be based, as much as possible, on regional information. Figure 5 shows magnitude dependence on fault length from global earthquake data.

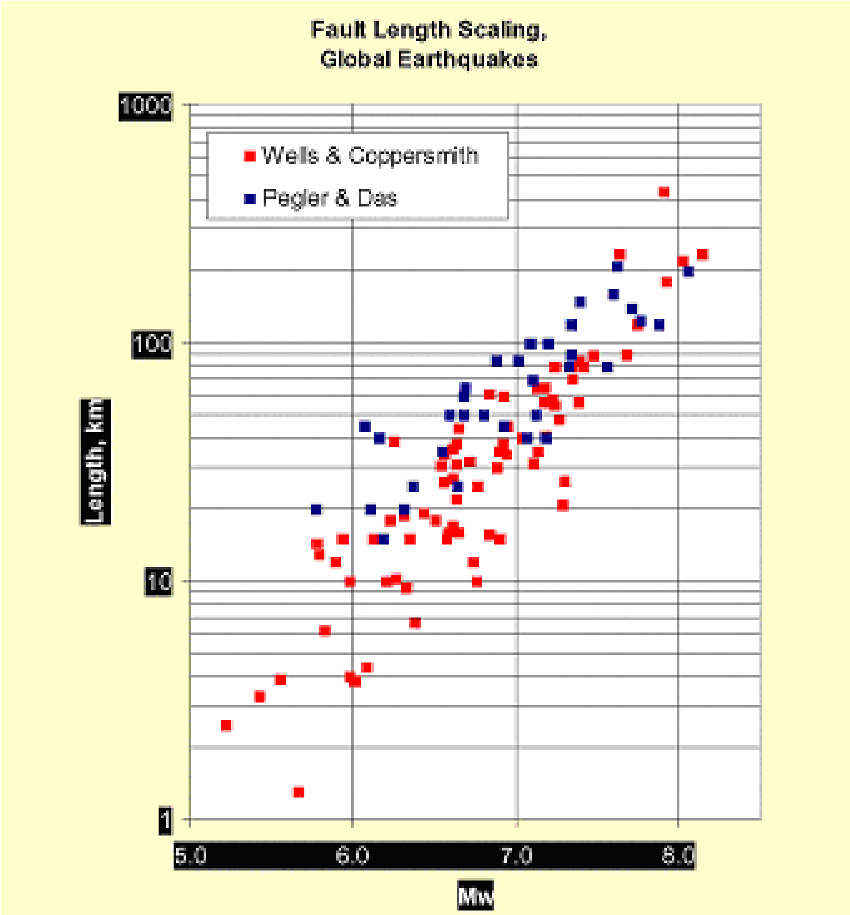


Figure 5 Fault Length Scaling to Magnitudes

In low seismic areas, the assessment of maximum credible earthquake magnitudes is more complicated. The solution to this problem is based on observed data. The data is based on seismic catalogues compiled from written records (historical approach). Fortunately, enough strong and damaging events in civilised areas are well recorded both in oral and written tradition. Statistical methods for the treatment of extreme values provide a meaningful means to assess maximum credible earthquake magnitudes in a region of interest. Possible methods are available, for example, by Noubary (2000):

- Bootstrap techniques (re-sampling of the distribution of observed maximum magnitude values)
- Threshold theory leading to the application of a Generalised Pareto Distribution (GPD)
- Traditional extreme value statistics like the Gumbel distribution.

It is worth to mention that these statistical methods have to apply in accordance with the mathematical prerequisites required for their application. For example in a threshold analysis the threshold to be analysed cannot be selected arbitrarily.

Additionally available information (e. g. from paleo-seismology) can be easily incorporated into the analysis. For example, paleo-seismological assessments of maximal magnitudes with the associated assessment of frequency (or recurrence period) can be incorporated into the empirical distribution of observed maximum values, which is re-sampled using a corresponding Monte-Carlo-Procedure. It is recommended to use the 95%-quantile of the re-sampled distribution as the maximum credible earthquake magnitude.

3.5 Selection of a parameter to characterise the impact of an earthquake

Different parameters are used by engineers to evaluate structural damage. For design purposes they often depend on national regulations and standards. Most standards are force-based. The design basis forces are derived typically from linear-elastic response spectra, taking into account some damping of the structure. These are adjusted by load correction factors for the required application. This is the reason why spectral accelerations (or even PGA) are traditional parameters for the representation of the results of seismic hazard analysis. In the past, following the original idea of Cancani (1904), PGA values were derived from intensity attenuation equations and therefore closely related to observed damage. At that time (before the mid 70's), measurements of ground motions were few, being limited by available instrumentation and seismic networks. The measured values were actually "peak-damped" without high frequency contents, because the latter were not measurable (high frequency peaks were filtered). Therefore, the physical meaning of the measured PGA values was quite close to the modern understanding of an effective ground acceleration (EGA) as used nowadays in engineering (with some minor difference in the values of the spectral amplification factors). This led to an implicit correlation of the observed intensities with the spectral acceleration reflecting the range of natural frequencies of civil structures. Indirectly, this correlation incorporates both the energy content of an earthquake, as well as the energy (defined by spectral shape and level) transfer into a structure. This picture has changed due to the development of modern seismic networks and instrumentations capable of recording high frequency contents of earthquake vibrations. Such high frequency vibrations, except for very brittle failure modes, generally do not cause damage to reasonably designed industrial structures and even to those not especially designed against earthquakes. Indeed, it is known that intensities (as a damage characteristic) correlate much better with peak ground velocity (PGV) or with the spectral acceleration corresponding to the first natural frequency of structures. Furthermore, ground motion measurements at a free surface (e. g. free standing soil column) are hardly representative for the interactions of seismic waves with massive buildings, which are considered by engineers. Therefore, the selection of appropriate physical parameter(s) to describe the damaging impact of an earthquake on structures more reliable is an important question for any methods of seismic hazard analysis. The selected parameter is important for later analysis steps, because the attenuation models to estimate the site hazard uses the same parameter. Generally, the parameter characteristics can be classified as structure-dependent or structure-independent.

3.5.1 Structure-independent parameters for impact characterisation

Due to the traditional division of labour between geophysicists and engineers, structure-independent impact parameters have some advantages due to their possible general-purpose applications. Spectral or peak ground accelerations are traditional structure-independent parameters for characterisation of the impact of earthquakes. As discussed in section 2.3, the sole use of spectral accelerations or spike peak ground accelerations may be misleading. Meaningful alternatives, which have found practical application, are the *Arias-Intensity* and the *Cumulative Absolute Velocity* (CAV).

The *Arias Intensity* (Arias, 1970) is defined as:

$$I_A = \frac{\pi}{2g} \int_0^{\tau} a^2(t)\, dt \tag{8}$$

where τ is the duration of the strong motion (eliminating the contribution of Coda waves) and a(t) is the acceleration time-history. Because *Arias-Intensity* represents a measure of the elastic energy content of an earthquake ground motion, it can be used to select design earthquakes in cases where inelastic behaviour of structures or components is not permitted (e.g., for brittle failure modes).

The *Cumulative Absolute Velocity* (EPRI, 1991) is calculated as:

$$CAV = \sum_{I=1}^{N} H\left(pga - 0.025\right) \int_{t_i}^{t_{i+1}} \left|a(t)\right| dt \tag{9}$$

where N is the number of 1-second time windows in the time series, pga is the peak ground acceleration in the i-th time window and $H(x)$ is the Heaviside function. CAV can be used to define the ductile (low cycle fatigue) failure mode condition of structures and components.

Another, more global structure-dependent parameter is intensity (e. g.according to EMS scale). Here the dependence on the design of the structure is given more implicitly by characterizing the damage observed for certain types of structures in terms of an intensity scale. Because the response of structures is maximised for resonance conditions intensities are implicitly correlated with the first natural frequency of the structures considered. For most structures these natural frequencies are in the range between 1 and 5 Hz.

3.5.2 Structure-dependent parameters for impact characterisation

Structure-dependent parameters can provide valuable information for the characterisation of the destructive potential of earthquakes. Most of them are based on an assessment of the energy transfer into a structure. The disadvantage is that the need to consider the physical characteristics of the structure may be too elaborate for general purpose applications. The analysis requires the use of appropriate time-histories, which can be synthetic seismograms and/or recorded data. More effort is needed here than for cases using structure-independent parameters. Therefore, the use of a structure-dependent parameter is recommended for specific structural analysis, for which deterministic scenario-based earthquakes approach is most appropriate.

Such an energy-based approach, more advanced than structures design by balancing energy demands and inputs, allows (1) proper characterisation of different types of time histories (impulsive, periodic with long-duration pulses, etc.) which may correspond to fairly realistic earthquake strong ground motions, and (2) simultaneous consideration of the dynamic response of a structure from elastic to ductile failure conditions.

The absolute energy input per unit of mass, can be expressed by:

$$E_I = \int_0^{\tau} \ddot{u}_t \, \dot{u}_g \, dt \qquad (10)$$

where $u_t = u + u_g$ is the absolute displacement of the mass, and u_g is the earthquake ground displacement. Another energy-based parameter, denoted as seismic hazard energy factor (AE_I), was introduced by Decanini et al., (1994), to take into account the global energy structural response amount. AE_I represents the area enclosed by the elastic energy input spectrum corresponding to different intervals of time, T (from T_1 to T_2) and is expressed by:

$$AE_I = \int_{T_1}^{T_2} E_I\left(\xi = 5\%, T\right) dT \qquad (11)$$

Other structure-dependent parameters for impact characterisation of earthquakes can be considered, too.

3.6 Ground motion prediction equations

For the selection of appropriate scenario earthquakes, as well as for the assessment of the impact on structures at a site, attenuation relationships are required. They should be developed for the selected parameter characterizing the impact of earthquakes on structures. In principle, the relationships show the parameter values as a function of distances for earthquake magnitudes. Still most popular are empirical attenuation equations with respect to spectral accelerations, spectral velocities or to intensities although direct simulation methods can be applied, too. Some equations for Arias-intensity or cumulative absolute velocity are also available (Travasarou et al, 2003, Kostov, 2005).

Empirical attenuation correlations for spectral accelerations (even developed specifically for a region) have limitations for near-fault conditions (e.g. Bolt and Abrahamson, 2002; Mollaioli, et al, 2003). These correlations are typically based on a model of simple amplitude decay with distance using a far-field approximation to a point (seismic) source characterisation (Aki and Richards, 2002). This approximation is not valid for near-fault conditions because it neglects multidimensional wave interference effects (Richwalski et al, 2004). Near-fault earthquake hazards can best be assessed by applying advanced dynamic source modelling. The use of broadband synthetic seismograms provides an effective means for performing this type of analysis (Romanelli et al,2003).

3.7 Incorporation of site effects

In general, site effects cannot be treated separately from the overall seismic waves propagation from causative seismic sources under consideration to the site through the propagating media (e.g., Panza et al, 2001; Field, 2000). In general there are two possibilities to deal with site effects. First of all site effects can be directly included into the attenuation model. This (to some extend) considers the dependency between the propagation of seismic waves from the source to the site. The second possibility consists in direct physical modelling. This approach allows incorporating the solution of the attenuation problem with site effects in a physically correct manner. This represents the most advanced approach for treating site effects. The disadvantage of this approach is the increased effort for the analysis. It is important to note that the models should conform to the principle of empirical control. Accordingly they have to be checked against earthquake recordings from the region when available. On the other hand a scenario-based approach reduces the number of scenarios to be analysed substantially in comparison to traditional PSHA. Large logic trees represent multi-millions of branches. For a correct representation of at least the important branches a large amount of scenarios has to be analysed.

3.8 Probabilistic model of earthquake recurrence and selection of risk-relevant scenario-earthquakes

Let us have a look how the frequency of scenario earthquakes can be calculated starting from the most general case for an area source, A, which completely encloses our site of interest (e.g., area with radius/distance of 300 km or less from the site). Because the occurrence of earthquakes is not invariant in time and space, the calculation of an average frequency of occurrence for a certain earthquake (magnitude) class requires the solution of the following equation:

$$F\left(M_i\right)=\frac{1}{T_{life}}\int_{A}\int_{0}^{T_{Life}}\int_{M_{low}}^{M_{upper}} f_1\left(r,m,t\right)dmdrdt \tag{12}$$

Here, $M_i \in \left(M_{low}, M_{upper}\right)$ is the magnitude value associated to the considered earthquake class, M_{low} is the lower interval limit for the considered class, M_{upper} is the upper interval limit for the considered earthquake class, F is the average frequency of the earthquake class, r is the distance from a point seismic source located inside the seismic area source A to the site, f_1 is the multivariate frequency density distribution of earthquakes within the considered area source, T_{Life} is the expected (residual) life time of the infrastructure analysed in the study, m is the magnitude, t is time. It is easy to understand that only a few earthquake classes have to be considered in a risk analysis (not more than 3 or 4). The impact assigned to each of the earthquake classes can be defined by the solution of the optimisation problem:

$$find\left(r_{opt}\right)\rightarrow \max\int_{A}\int_{0}^{T_{life}}\int_{M_{low}}^{M_{upper}} f_1\left(m,r,t\right)g\left(r\middle|m\right)dmdrdt \tag{13}$$

where $g\left(r\middle|m\right)$ calculates the value of the selected impact parameter (energy-based measure, spectral acceleration, etc.) as a function of the distance from the location of the earthquake with magnitude m to the site. The calculated r_{opt} defines the location of the deterministic scenario earthquake considered for this class. In many practical cases, a simplification of the problem is possible by separating the spatial distribution of seismicity from the frequency distribution of earthquakes depending on magnitude size and time. This means that the frequency density distribution f_1 can be represented as:

$$f_1\left(r,m,t\right)= f_2\left(m,t\right) f_3\left(r\middle|m\right) \tag{14}$$

Equation (14) reflects the assumption that the spatial distribution of seismic activity is invariant with time. This is of course a rather strong assumption, which for a short-lived structure can be justified by the assumption of stable seismotectonic conditions in the area of interest. The required density distribution f_2 can be obtained much more easily, for example, using bivariate extreme value distributions (Noubary, 2000) or Markoff or Semi-Markoff models.

In cases where the seismic activity can be allocated to specific faults, the problem is simplified to a very large extent. The frequency of an earthquake belonging to the class *i* can be calculated as:

$$F\left(M_i\right) = \frac{1}{T_{Life}} \sum_{j=1}^{N} \int_{0}^{T_{Life}} \int_{M_{lower}}^{M_{Upper}} f_j\left(m,t\right) dm dt \tag{15}$$

Here, j is the summation index for the relevant faults and N is the total number of faults potentially contributing to the magnitude class i. The optimisation problem of equation (13) can also be simplified under these conditions by making the bounding assumption that the shortest distance between fault and site will be selected.

Once the probabilistic scenario-earthquakes are selected and their frequency is calculated (this is the required frequency of an initiating seismic event), it is easy to calculate scenario-specific hazard spectra, which will provide the input for subsequent analysis within the framework of a seismic PRA. Within this probabilistic framework it is possible to calculate uncertainty bounds on the average frequencies obtained from equation (12) or (15) by performing sensitivity analysis. It is also possible to calculate uncertainty bounds for the hazard spectra associated with each magnitude class, taking into account the total empirically observed uncertainty associated with the attenuation of seismic waves in the region of interest. Such estimates can easily be performed by propagating the uncertainties associated with the lack of knowledge of the values of the model parameters used through the model. Direct Monte Carlo analysis or response-surface analysis techniques can be used in dependence of the complexity of the model.

The proposed probabilistic extension of the neodeterministic scenario-based seismic hazard analysis method avoids the problems associated with the development of uniform hazard spectra (UHS) in traditional PSHA. Furthermore, it is focused on the output as typically requested by risk analysts (frequency of critical events instead of exceedance probabilities of secondary hazard parameters).

4 Numerical Example 1

A numerical example will illustrate the suggested scenario-based procedure. For simplicity, the solution of the optimisation problem will be performed, using the simplifying assumption of equation (14).

4.1 Task specification

Figure 6 illustrates the task. A critical infrastructure shall be designed against earthquakes. Additionally it is required to develop the input for a seismic PRA. The critical infrastructure is located in the centre of the circle shown in figure 6. From the responsible project engineers it is known that modern design rules ensuring a ductile design of structures will be applied. It is also known that the characteristic first natural frequencies of the new structures are expected to be in the range of 3Hz. The design lifetime of the critical infrastructure is 40 years. The very detailed site investigation performed allows the definition of an exclusion zone with respect to the existence of active, capable faults within a radius of 5 km around the site. This means that inside this area only small and deep earthquake events are feasible (M_w < 5.0). From the available geological and seismological database, it was concluded that in the surroundings of the site two significant linear sources (faults) have to be considered. The shortest distances to site are D1=30 km and D2=25 km. The length of surface projection of the first fault (line source LS1) is 21 km and of the second fault (line source LS2) 15 km. For simplicity, it is assumed that the perpendicular from the site to the faults subdivides the fault surface projections of both line sources into two parts at a ratio of 2:1. Available data does not indicate any preferred location of epicentres along both faults, therefore a non-informative distribution of epicentre location has to be assumed. Based on historical data two areal sources (AS1 and AS2) with some past seismic activity have been discovered, which have to be considered in the analyses. The shortest distance of both areal sources to the site is 5 km (joining the exclusion zone). Areal source AS1 is extended up to a distance of 65 km, while areal source AS1 is extended up to a distance of 98 km from the site. Detailed statistical analyses have been performed to develop temporal and spatial frequency distributions of earthquake occurrences at the different sources including spatial distributions of epicentres for the areal sources. For simplification, the model of bivariate exponential distributions (Gumbel Type 2, see Noubary, 2000) is used both for the temporal distributions as well as for the spatial distributions. Detailed statistical analysis showed that the 95%-quantil of the magnitude distribution corresponds to a magnitude of 5.9 for source AS1 and 6.3 for source AS2.

The joint distribution function of the bivariate exponential distribution (X, T being the variates) is given as:

$$F(x,t) = \left(1 - e^{-\lambda_1 x}\right)\left(1 - e^{-\lambda_2 t}\right)\left[1 + \alpha e^{-(\lambda_1 x + \lambda_2 t)}\right] \tag{16}$$

The joint density is given as:

$$f\left(x,t\right) = \lambda_1 \lambda_2 e^{\left(-\lambda_1 x - \lambda_2 t\right)}\left[1 + \alpha\left(2e^{\lambda_1 x} - 1\right)\left(2e^{\lambda_2 t} - 1\right)\right] \tag{17}$$

Maximum likelihood estimators for the parameters λ_1 and λ_2 are based on the empirical mean of X and T and calculated simply as:

$$\hat{\lambda}_1 = \frac{1}{\overline{X}} \tag{18}$$

and

$$\hat{\lambda}_2 = \frac{1}{\overline{T}} \tag{19}$$

α is calculated based on the empirical correlation coefficient ρ:

$$\hat{\alpha} = 4\hat{\rho} \tag{20}$$

In our application the random parameter X has the meaning of magnitude, while the random parameter T corresponds either to the elapsed time between two earthquakes (temporal distribution) or to the distance between epicentre location and site. Other statistical distributions, continuous as well as discrete ones, can be used in dependence of the results of data analysis.

Upper limit estimates for the statistical models can be provided by accounting for the error in the magnitude and location estimates. The simplest procedure consists of an estimate of the upper limit for the maximum magnitude (e. g. mean + 2σ, or 95%-quantile) and the lower limit for the distance (in case of a spatial distribution for an areal source, e. g. mean -2σ). The procedure is similar with respect to the elapsed time between events.

Statistical analysis was performed in units of moment magnitudes, years (time) and km (distance).

Table 1 shows the information available for the line sources. Table 2 shows the available information for the areal sources for the considered case.

Table 1: Data for line sources

Source	Fault length, km	Fault length error, standard deviation in km	Shortest distance to site, km	Applicable statistical model for f_2 (see equation (11))	Parameters of the model			Parameters of the model for f_2, upper limit		
					λ_1	λ_2	α	λ_1	λ_2	α
LS1	25	5	30	Bivariate exponential (Gumbel)	0.17	0.009	0.79	0.15	0.013	0.79
LS2	17	3	25	Bivariate exponential (Gumbel)	0.23	0.0051	0.69	0.20	0.0082	0.68

Table 2a: Data for areal sources, distributions for f_2 (Eq. (11))

Source	Shortest distance to site, km	Applicable statistical model for f_2 (see equation (11))	Parameters of the model for f_2			Parameters of the model for f_2, upper limit		
			λ_1	λ_2	α	λ_1	λ_2	α
AS1	5	Bivariate exponential (Gumbel)	0.35	0.132	0.84	0.32	0.143	0.81
AS2	5	Bivariate exponential (Gumbel	0.31	0.124	0.89	0.28	0.17	0.88

Table 2b: Data for areal sources, distributions for f_3 (Eq. (11))

Source	Shortest distance to site, km	Applicable statistical model for f_3 (see equation (11))	Parameters of the model for f_3			Parameters of the model for f_3, upper limit		
			λ_1	λ_2	α	λ_1	λ_2	α
AS1	5	Conditional probability, based on a bivariate exponential model (Gumbel)	0.35	0.031	0.8	0.32	0.035	0.65
AS2	5	Conditional probability, based on a bivariate exponential model (Gumbel)	0.31	0.022	0.72	0.28	0.03	0.64

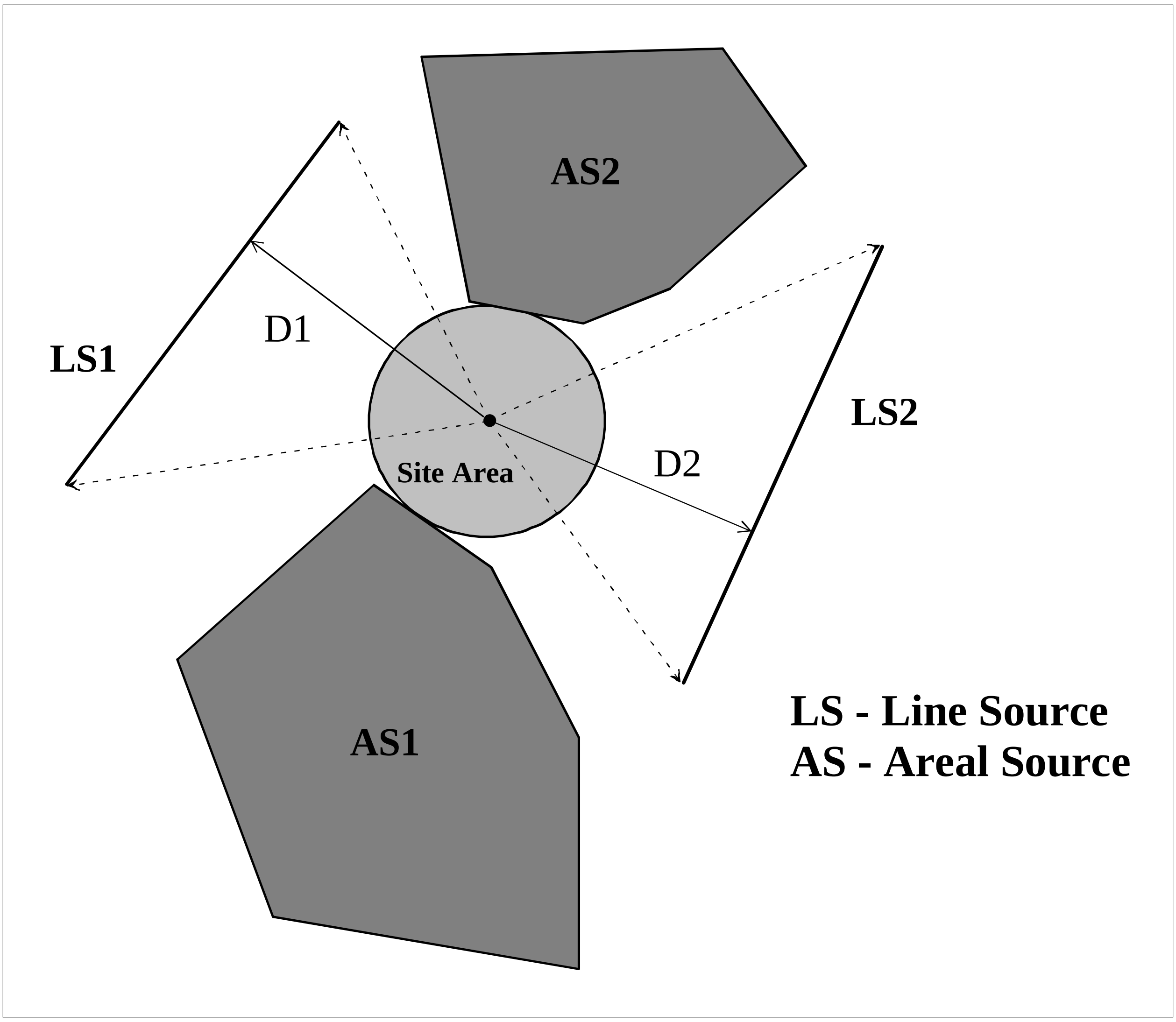


Figure 6 Illustration of the numerical example

In addition, for our example it is assumed that

- detailed physical modelling has confirmed that for the relevant sources a simple amplitude-decay model for ground motion attenuation is acceptable,
- validated attenuation models for each of the sources have been established in terms of ground motion (spectral accelerations).

With respect to attenuation equations a set of four source-specific equations is available, reflecting the different topographical and directivity conditions with respect to seismic wave propagation from the different sources to the site. The general format for these equations is:

$$\log\left(S_a\right) = a + bM_w + c\log\left(R\right) + dR + \sigma P \qquad (21)$$

with $R = \left(D^2{}_{JB} + h^2\right)^{0.5}$ and D_{JB} representing the Joyner-Boore distance.

Table 3 shows the coefficients of the equation for the line source LS1, table 4 for LS2 , table 3 for the areal source AS1 and table 4 for the areal source AS2. These equations have been developed especially for this analysis by modifying the baseline equation of table 3. They are not to be used for any other purpose. For simplicity a constant standard deviation for all equations of 0.28 (in log-scale) is assumed.

Table 3: Coefficients of attenuation model for LS1

Spectral frequency, Hz	a	b	c	d	h
PGA (50)	-1.5537	0.2396	-0.62494	-0.0081622	5.4294
35	-1.5558	0.2648	-0.66713	-0.0085626	5.658
25	-1.6455	0.27332	-0.63828	-0.0087011	5.0448
20	-1.3713	0.23727	-0.63121	-0.0086357	4.9516
13.33	-1.3756	0.24517	-0.63336	-0.0086132	5.268
10	-1.2412	0.23763	-0.63708	-0.0086018	5.607
6.67	-0.96632	0.21371	-0.62504	-0.0083936	6.1966
5	-1.0168	0.21242	-0.57166	-0.0082279	5.8137
4	-1.103	0.22025	-0.56614	-0.0081654	6.765
2.5	-2.053	0.31787	-0.50505	-0.0079937	4.8624
2	-2.5039	0.35523	-0.46556	-0.0079405	4.6353
1.34	-2.6029	0.357	-0.45591	-0.0078623	4.617
1	-3.0338	0.38841	-0.42746	-0.0078021	4.0694
0.667	-3.521	0.42579	-0.41148	-0.0077495	4.5939
0.5	-3.9299	0.46231	-0.41078	-0.0077495	4.7113

Table 4: Coefficients of attenuation model for LS2

Spectral frequency, Hz	a	b	c	d	h
PGA (50)	-1.320645	0.27554	-0.687434	-0.0081622	6.6294
35	-1.32243	0.30452	-0.733843	-0.0085626	6.858
25	-1.398675	0.314318	-0.702108	-0.0087011	6.2448
20	-1.165605	0.2728605	-0.694331	-0.0086357	6.1516
13.33	-1.16926	0.2819455	-0.696696	-0.0086132	6.468
10	-1.05502	0.2732745	-0.700788	-0.0086018	6.807
6.67	-0.821372	0.2457665	-0.687544	-0.0083936	7.3966
5	-0.86428	0.244283	-0.628826	-0.0082279	7.0137
4	-0.93755	0.2532875	-0.622754	-0.0081654	7.965

Spectral frequency, Hz	a	b	c	d	h
2.5	-1.74505	0.3655505	-0.555555	-0.0079937	6.0624
2	-2.128315	0.4085145	-0.512116	-0.0079405	5.8353
1.34	-2.212465	0.41055	-0.501501	-0.0078623	5.817
1	-2.57873	0.4466715	-0.470206	-0.0078021	5.2694
0.667	-2.99285	0.4896585	-0.452628	-0.0077495	5.7939
0.5	-3.340415	0.5316565	-0.451858	-0.0077495	5.9113

Table 5: Coefficients of attenuation model for AS1

Spectral frequency, Hz	a	b	c	d	h
PGA (50)	-1.70907	0.27554	-0.812422	-0.0081622	5.4294
35	-1.71138	0.30452	-0.867269	-0.0085626	5.658
25	-1.81005	0.314318	-0.829764	-0.0087011	5.0448
20	-1.50843	0.2728605	-0.820573	-0.0086357	4.9516
13.33	-1.51316	0.2819455	-0.823368	-0.0086132	5.268
10	-1.36532	0.2732745	-0.828204	-0.0086018	5.607
6.67	-1.062952	0.2457665	-0.812552	-0.0083936	6.1966
5	-1.11848	0.244283	-0.743158	-0.0082279	5.8137
4	-1.2133	0.2532875	-0.735982	-0.0081654	6.765
2.5	-2.2583	0.3655505	-0.656565	-0.0079937	4.8624
2	-2.75429	0.4085145	-0.605228	-0.0079405	4.6353
1.34	-2.86319	0.41055	-0.592683	-0.0078623	4.617
1	-3.33718	0.4466715	-0.555698	-0.0078021	4.0694
0.667	-3.8731	0.4896585	-0.534924	-0.0077495	4.5939
0.5	-4.32289	0.5316565	-0.534014	-0.0077495	4.7113

Table 6: Coefficients of attenuation model for AS2

Spectral frequency, Hz	a	b	c	d	h
PGA (50)	-1.39833	0.2396	-0.718681	-0.0081622	5.4294
35	-1.40022	0.2648	-0.7671995	-0.0085626	5.658
25	-1.48095	0.27332	-0.734022	-0.0087011	5.0448
20	-1.23417	0.23727	-0.7258915	-0.0086357	4.9516
13.33	-1.23804	0.24517	-0.728364	-0.0086132	5.268
10	-1.11708	0.23763	-0.732642	-0.0086018	5.607
6.67	-0.869688	0.21371	-0.718796	-0.0083936	6.1966
5	-0.91512	0.21242	-0.657409	-0.0082279	5.8137
4	-0.9927	0.22025	-0.651061	-0.0081654	6.765
2.5	-1.8477	0.31787	-0.5808075	-0.0079937	4.8624
2	-2.25351	0.35523	-0.535394	-0.0079405	4.6353
1.34	-2.34261	0.357	-0.5242965	-0.0078623	4.617
1	-2.73042	0.38841	-0.491579	-0.0078021	4.0694
0.667	-3.1689	0.42579	-0.473202	-0.0077495	4.5939
0.5	-3.53691	0.46231	-0.472397	-0.0077495	4.7113

For our example, it is assumed that from the information in the geological and seismological database the following correlations for the relationship between fault rupture length and moment magnitude as well as between fault length and moment magnitude have been established.

$$\log\left(L_R\right) = -3.6 + 0.75M + P\sigma \tag{22}$$

with $\sigma = 0.1$ and

$$\log\left(L_{fault}\right) = -3.25 + 0.72M + P\sigma \tag{23}$$

with $\sigma = 0.1$.

It is also assumed that the regression technique used for the development of these equations possesses the property of orthogonality.

4.2 Neodeterministic scenario-based analysis

According to the procedure of the neodeterministic scenario-based seismic hazard analysis, the first step consists in the evaluation of the maximum credible earthquake.

A conservative way of performing this task consists in the assumption that the whole fault length established by measurement could rupture. Additionally, the uncertainty of the measurement should be considered. The analysis is performed for a critical infrastructure. Therefore, we base our analysis on the mean +1 σ value of the estimated fault length length as well as on the mean - 1σ value (inverse problem) obtained from equations (19) and (20).

So, we obtain for the linear source LS1 a MCE-value of M_w=6.9. In case we want to base our analysis on the more realistic correlation between fault length and magnitude, removing the assumption of a complete rupture of the fault, the result would be M_w=6.7. Therefore, the difference is not very large. Neglecting the uncertainty but keeping the assumption that the fault can rupture completely results in a magnitude value of 6.8. So the discussion confirms that MCE-magnitudes behave robustly with respect to a modification of data on fault or rupture lengths.

Therefore, we accept the following value

$$MCE_{lS1} = 6.9 .$$

Repeating the same procedure for line source LS2, we obtain a magnitude value of

$$MCE_{LS2} = 6.7 .$$

The next task consists in the evaluation of the MCE-magnitudes for the areal sources. According to our procedure we use the 95%-quantile of the historical magnitude distribution as the magnitude value for the MCE. The resulting MCE for the areal source AS1 is

$$MCE_{AS1} = 5.9 .$$

Accordingly, we obtain for the second areal source:

$$MCE_{AS2} = 6.3 .$$

Once we have established the maximum credible earthquakes (in practical applications the values may be rounded off to the next larger quarter of a magnitude unit, therefore the final values would be 7.0, 6.75 for the line sources and 6.5 and 6.0 for the areal sources), we are able to calculate the corresponding hazard spectra and the associated CAV-values assuming the shortest distance between source and site.

Because in our case we know that the new construction will correspond to modern requirements of a ductile design, we may use the CAV-value as the criterion which scenario to select for the design of the new infrastructure. The alternative could simply consist in the use of an envelope of the hazard spectra for all four MCE-scenarios using the source-specific attenuation models. The later approach is more conservative. It is close to the approach frequently used by practitioners, where the most conservative of all attenuation equations for a region is used if source specific models are not available.

Figure 7 shows the resulting hazard spectra for the 4 scenario-earthquakes based on the mean regression models. Figure 8 shows the same comparison for the mean +1 σ value. It is observed that at lower frequencies the hazard is dominated by the line source 2, while in the higher frequency range the areal sources contribute to the hazard envelope. Nevertheless, a design based on scenario 2 only, is sufficient, because it results in the highest spectral acceleration values in the range of the first natural frequency of the considered construction. Furthermore, the differences to the other scenarios at higher frequencies are low. Additionally, we may prefer to consider the information available with respect to the epicentre distribution in the areal sources. They indicate that the expected value for the distance to the site is much higher than 5 km (28.6 to 33.3 km, according to the statistical analysis). This consideration would allow the exclusion of the areal sources from further consideration.

Figure 9 shows the enveloping hazard spectra for the mean regression and the mean +1 σ model.

It is interesting to observe that the most critical scenario results from line source LS2 with a smaller maximum credible magnitude than line source LS1. This is the result of the shorter minimal distance and the large differences between the source-specific attenuation equations of the two sources. This emphasises the importance of the development of a source-specific attenuation model or the use of detailed wave propagation models (e.g. the use of synthetic seismograms).

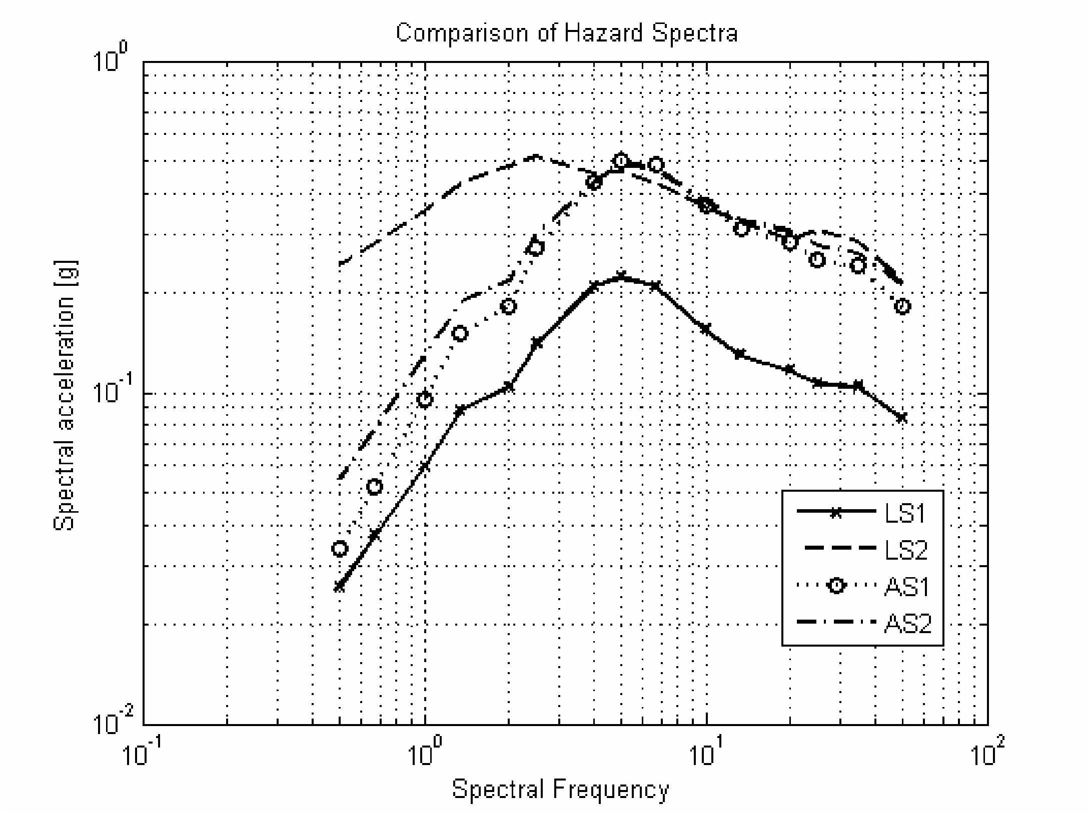


Figure 7 Comparison of scenario hazard spectra – mean regression model

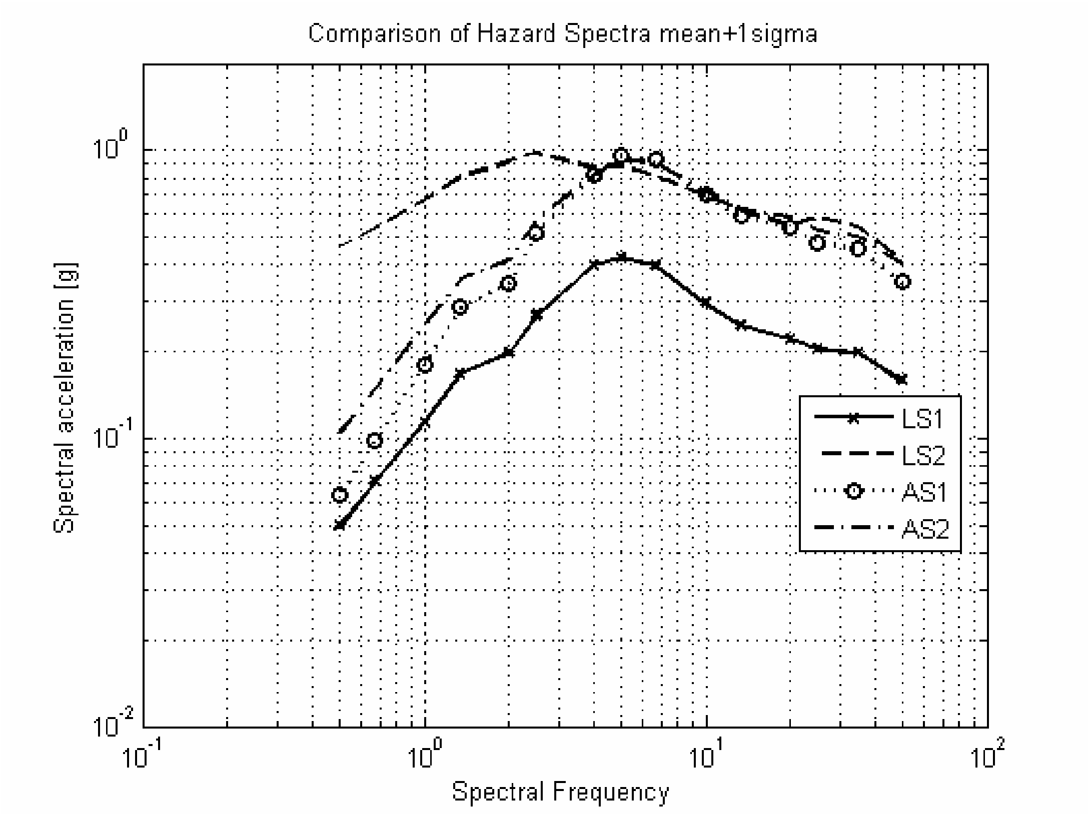


Figure 8 Comparison of scenario hazard spectra – mean + 1 sigma model

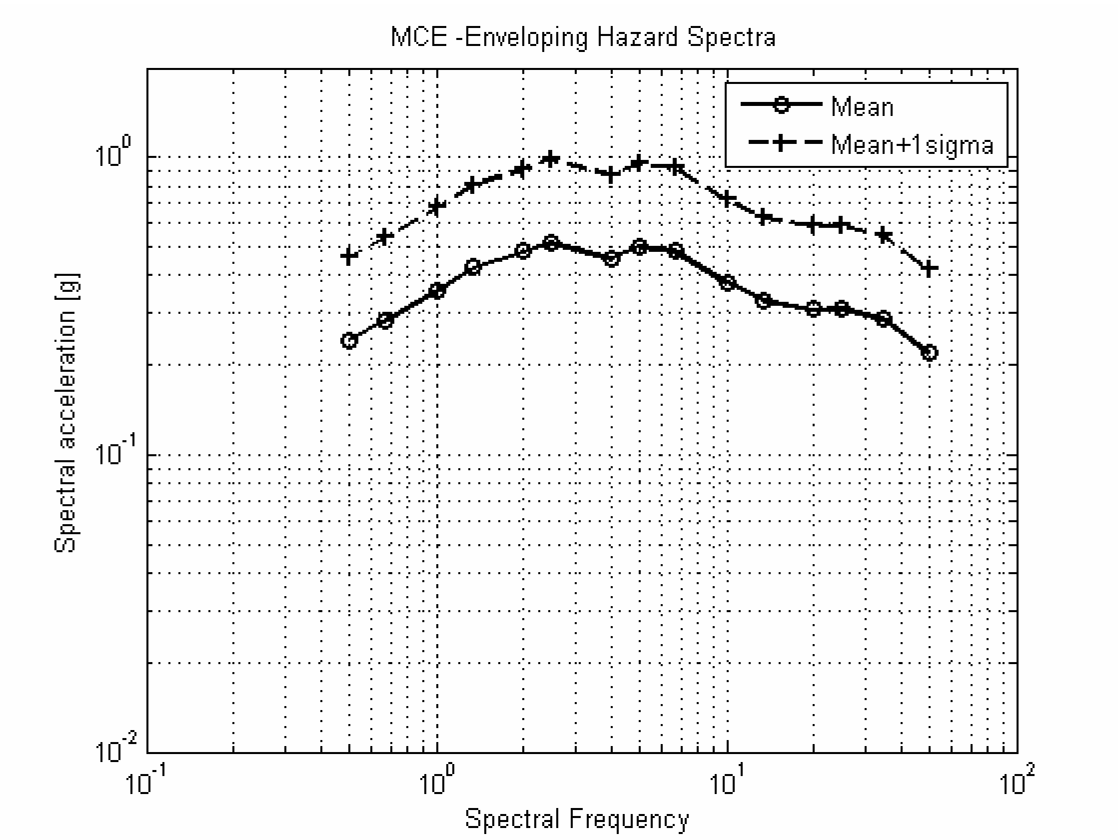


Figure 9 Comparison of enveloping hazard spectra

Note that any smaller seismic event at any of the four sources will not exceed the enveloping hazard developed from the scenarios.

4.3 Probabilistic scenario-based hazard analysis

4.3.1 Introductory discussion

Note that an upper limit for the probability of the critical scenario 2 (this is not to be set equal to the total frequency of scenarios in the same magnitude class as described in section 3.7 can be assessed with the help of the recommended bivariate exponential model (neglecting for simplicity the truncation at the "physical limit of m=6.7" in our introductory discussion). The conditional probability of occurrence of an earthquake with magnitude X exceeding a specified value given a certain length of time (the lifetime of our structure) for the bivariate exponential model is calculated as:

$$P\left(X > m \middle| T > T_{Life}\right) = \frac{P\left(X > m, T > T_{Life}\right)}{P\left(T > T_{Life}\right)} \tag{24}$$

By using $F(x,t)$ and the marginal distribution of T, yields

$$P\left(X > m \middle| T > T_{Life}\right) = e^{-\lambda_1 m}\left[1+\alpha\left(1-e^{-\lambda_1 m}\right)\left(1-e^{-\lambda_2 T_{Life}}\right)\right] \quad (25)$$

The conditional probability of earthquake occurrence within an interval (m_1, m_2) is given as:

$$P\left(X > m_1 \middle| T > T_{Life}\right) - P\left(X > m_2 \middle| T > T_{life}\right) \quad (26)$$

The averaged annual frequency is obtained dividing the result by the lifetime of the structure. This approach can also be used for line source 1 and, in a similar way, for the areal sources neglecting the spatial distribution of seismic activity and for other magnitude values. Combining the obtained frequencies with the worst case scenario (deterministic scenario earthquake at LS2 with magnitude 6.7) and summing up over all frequencies of the corresponding magnitude class leads to a conservative risk model for the infrastructure for the considered seismic initiating event, because the impact is maximized under our assumptions

- the scenario earthquake it located at the shortest distance to the site,
- analysis of seismic wave attenuation indicated the applicability of a simple amplitude-decay model.

Risk analysts are interested in a more realistic assessment. Therefore, a more detailed probabilistic analysis following the procedure of section 3.7 is required.

4.3.2 Detailed probabilistic analysis

In a first step it is necessary to scale the suggested probabilistic models (the bivariate exponential distribution), which in principle allow infinite values of X (meaning magnitude or distance to site), for application in an interval. Due to the correlation between X and time, the calibration factor K is time dependent. The factor can be calculated from the joint distribution function (equation (17)):

$$K(t) = \left[\frac{F(\infty,t)-F(0,t)}{F(x_u,t)-F(x_l,t)}\right] \quad (27)$$

For $t \to \infty$ the calibration coefficient obtains its usual univariate format:

$$K = \frac{e^{-\lambda_1 x_l}}{1-e^{-\lambda_1(x_u - x_l)}} \quad (28)$$

Here, x_u and x_l are the upper and lower limits of the random variable X (here meaning magnitude or distance).

The first step in our risk analysis consists in the calculation of the frequency of initiating events for each class of events. After the initial seismological analysis, we decided to consider the following event classes:

- magnitude between 6.5 and 6.9 – magnitude class 1
- magnitude between 6.0 and 6.5 – magnitude class 2
- magnitude between 5.5 and 6.0 – magnitude class 3

Because the design of the considered infrastructure will be very robust (designed against a conservative MCE-scenario), it is not necessary to consider more events in the analysis.

First, we calculate the frequency of class 1 events. Only the linear sources LS1 and LS2 contribute to this class. Additionally, the magnitude truncation at magnitude 6.7 has to be considered for LS2. Therefore, the frequency of the events in class 1 can be calculated as the integral over time of a sum of two integrals :

$$F\left(6.5 \le m \le 6.9\right) = {}^{1}\!/_{T_{life}} \int_{0}^{T_{life}} \left[\int_{6.5}^{6.9} f_{LS1}\left(m,t\right) dm + \int_{6.5}^{6.7} f_{lS2}\left(m,t\right) dm \right] dt \tag{29}$$

Here, f_{LS1} and f_{LS2} are the calibrated joint density functions of the bivariate exponential model for the line sources LS1 and LS2 correspondingly.

For event class 2 areal source AS2 has to be considered additionally besides the line sources. Because our analysis is based on equation (14), an integration over the area is not required for the evaluation of the total frequency of earthquake events in this class. Therefore, the resulting equation is again an integral over time of a sum of integrals:

$$F\left(6.0 \le m \le 6.5\right) = {}^{1}\!/_{T_{life}} \int_{0}^{T_{Life}} \left[\int_{6.0}^{6.5} \left(f_{LS1}\left(m,t\right) + f_{LS2}\left(m,t\right) \right) dm + \int_{6.0}^{6.3} f_{AS2}\left(m,t\right) dm \right] dt \tag{30}$$

Similarly we obtain the frequency for the event class 3:

$$F\left(6.0 \le m \le 6.5\right) = {}^{1}\!/_{T_{life}} \int_{0}^{T_{Life}} \left[\int_{6.0}^{6.5} \left(f_{LS1}\left(m,t\right) + f_{LS2}\left(m,t\right) + f_{AS2}\left(m,t\right) \right) dm + \int_{5.5}^{5.9} f_{AS1}\left(m,t\right) dm \right] dt \tag{31}$$

The calculation's results of the frequency of events are shown in table 7. A lower magnitude level of $m_l = 2.0$ was used in the analysis.

Table 7: Initiating event frequencies of the scenario earthquakes

Scenario earthquake (magnitude class)	Magnitude range	Frequency (best estimate)	Frequency, upper limit
1	6.5-6.9	0.00107	0.00113
2	6.0-6.5	0.0284	0.14
3	5.5-6.0	0.0742	0.316

4.3.3 Solution of the optimisation problem

For a more realistic derivation of the scenarios the optimisation problem according to equation (13) has to be solved. The optimisation problem can be simplified under certain conditions. For example, if

- the selected ground motion characteristic follows a simple amplitude-decay model,
- and the spatial distribution over the source is non-informative (uniform distribution or beta-distribution with shape parameters smaller than 1 within the distance interval),

then the scenario earthquake can be assumed to occur at the shortest distance between source and site. Under these conditions a simple comparison between the resulting hazard spectra (as performed for the deterministic case in section 2) is sufficient to identify the critical scenario for each magnitude class.

In our example, these conditions are fulfilled for the line sources but not for the areal sources. Because the first natural frequency of the considered infrastructure is in the range of 3 Hz, we solve the optimisation problem with respect to the spectral acceleration at 3Hz.

4.3.3.1 Magnitude class 1

Only the two line sources actually contribute to this magnitude class. According to the task description we don't have any relevant information on the spatial distribution of seismicity along the faults. Therefore, the earthquake scenario to be considered in the risk study corresponds to the deterministic scenario-earthquake occurring at the closest distance between line source LS2 and the site. Figure 7 shows the corresponding hazard spectrum (regression mean).

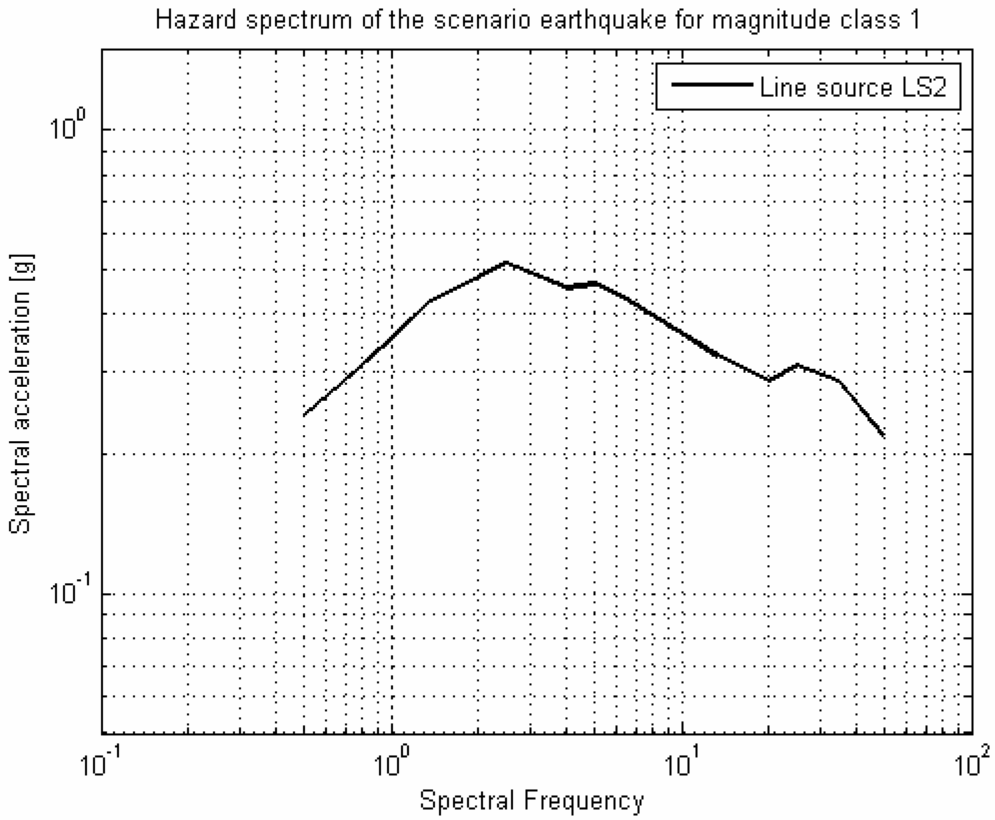


Figure 7. Hazard spectrum of the scenario earthquake of magnitude class 1.

4.3.3.2 Magnitude class 2

Contributors to this class are both line sources and the areal source AS2. For the areal source a probabilistic model for the spatial distribution of seismicity is given. For the line sources once again a simplified analysis is sufficient assuming the occurrence of the candidate scenario earthquakes at the shortest distance to the site. Therefore, the optimisation problem converts into the task of finding the location of the candidate scenario earthquake for area source 2 and a comparison of the hazard spectra of all candidate scenarios. Because for the areal source we also apply an amplitude-decay model it is sufficient to solve the reduced optimisation problem

$$find\left(r_{opt}\right) \rightarrow \max \int_{0}^{T_{Life}} \int_{5.5}^{5.9} \int_{r_{\min}}^{r_{\max}} f_{AS2}\left(m,t\right) f_{AS2}\left(r\left|m\right.\right) dm dr dt \tag{32}$$

to find the candidate scenario earthquake for the areal source 2. The integration variable can be separated. Therefore, it is possible to perform a further reduction of the optimisation problem:

$$find\left(r_{opt}\right) \rightarrow \max \int_{r_{\min}}^{r_{\max}} f_{AS2}\left(r\left|m\right.\right) \tag{33}$$

The conditional probability can be calculated in analogy to equation (22). For the solution it is sufficient to find the location r_{opt} maximising the conditional probability for the lower magnitude value of the considered interval (5.5). From equation (22) it can be concluded that the candidate scenario earthquake for the areal source AS2 is also located at the boundary of the source (the modal value is located at the shortest distance). Therefore, for the final selection of the scenario-earthquake for magnitude class 2 we have to compare the hazard spectra from the 3 contributing sources LS1, LS2 and AS2. For the line sources, the magnitude values to be considered are m=6.5, while for the areal source the magnitude value is 6.3 (maximal value). Figure 10 shows the comparison of the hazard spectra for the 3 candidate scenarios. The hazard spectrum of candidate scenario from line source 2 shows the highest value for the spectral acceleration at 3Hz although the corresponding value for areal source 2 is close. Because the candidate scenario earthquake from line source LS2 is associated with a larger magnitude value (with a larger energy content), the candidate scenario from line source LS2 has to be selected as the final scenario earthquake for magnitude class 2.

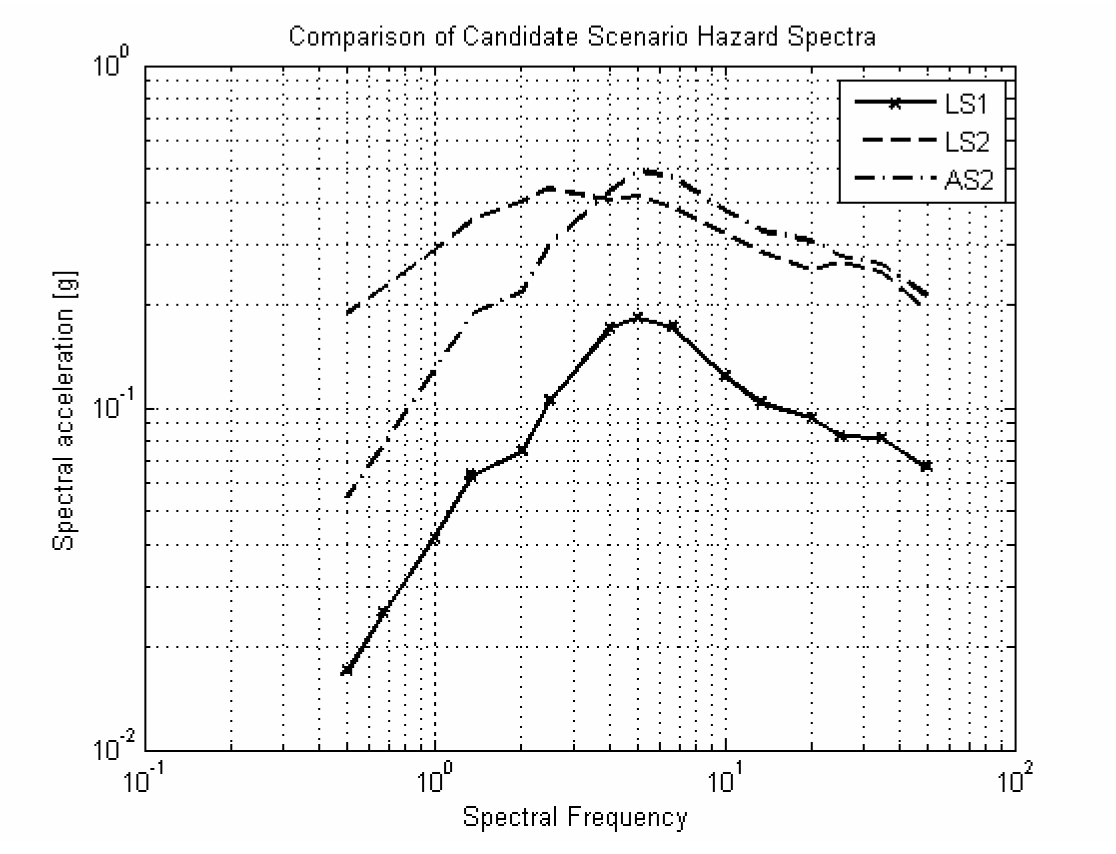


Figure 10 Hazard spectra of the candidate scenario earthquakes of magnitude class 2.

4.3.3.3 Magnitude class 3

The solution of the optimisation problem for magnitude class 3 follows the discussion in section 3.3.2. All sources do contribute to this magnitude class. Once again the candidate scenario earthquakes are located at the boundary of the areal sources and at the shortest distance between the line sources and the site. Therefore, the final scenario earthquake is to be selected by a comparison of the hazard spectra of the candidate scenarios from each source. Figure 11 shows the comparison. Again the candidate scenario earthquake of line source LS2 leads to the largest spectral acceleration at 3 Hz. Therefore, it has to be selected as the final scenario earthquake for magnitude class 3.

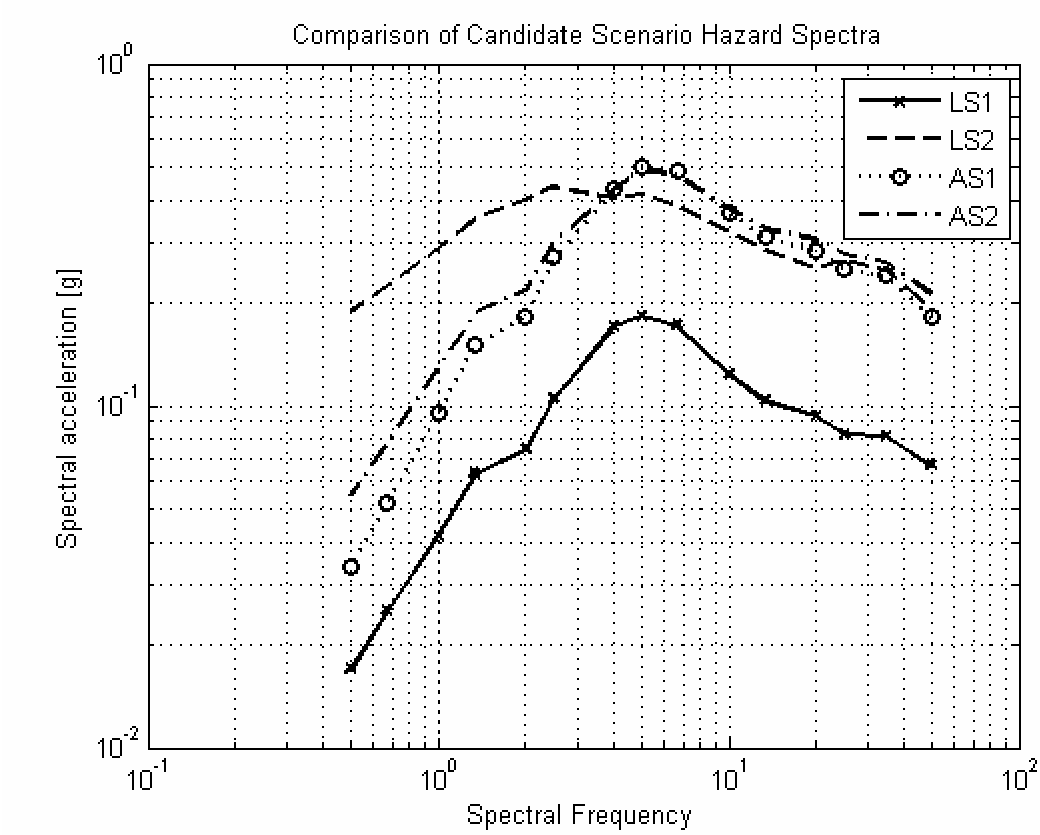


Figure 11 Hazard spectra of the candidate scenario earthquakes of magnitude class 3.

4.4 Seismic risk evaluation

Most seismic risk studies (e.g. for nuclear power plants) as well as the corresponding software are based on hazard curves. The new methodology does not require the development of hazard curves because the frequency of seismic initiating events is calculated directly. Instead of hazard curves it is required to calculate the conditional probability of exceedance of the scenario earthquakes' hazard spectra including the corresponding uncertainty distribution. Together with the calculated frequencies of initiating events this allows to use the e-xisting risk software to perform a seismic PRA (Probabilistic Risk Assessment).

For the calculation of the conditional hazard spectra exceedance probability it is possible to use the model of a lognormal distribution of spectral accelerations for a given scenario earthquake. To use this model correctly, we have to adjust the uncertainty values of our attenuation equations. Attenuation equations represent multivariate distributions of spectral accelerations in dependence of magnitude, distance and additional parameters not used explicitly as explanatory variables in the equation. The uncertainty caused by these additional explanatory variables is frequently confused with inherent randomness of earthquakes and named aleatory uncertainty (Abrahamson, 2006, SSHAC, 1997). Because this uncertainty is epistemic by nature, it is more appropriate to call this uncertainty "(temporarily) irreducible epistemic uncertainty". This irreducible part has to be treated as random in our model. The contribution of uncertainty of magnitude and distance can be eliminated from our probabilistic model because the selected scenarios are characterised by a fixed (upper estimate) and known magnitude value and a fixed and known distance between the earthquake location and the site. Furthermore, the selected scenarios are conservative with respect to all scenarios within the same magnitude class. Considering that the error term in our attenuation equation (18) can be represented as

$$\sigma = \sqrt{\left(\frac{\partial g(m,r)}{\partial m}\sigma_m\right)^2 + 2\rho\sigma_m\sigma_r + \left(\frac{\partial g(m,r)}{\partial r}\sigma_r\right)^2 + \sigma_{ired}^2} \tag{34}$$

we can calculate the irreducible, residual part of uncertainty σ_{ired} to be considered in the probabilistic model. *g* is the attenuation equation functional form (equation (18)). The correlation coefficient ρ can be set to 1, because a strong physical correlation exists between magnitude and epicentre location at the fault rupture plane. Furthermore, these two parameters are correlated in our case because the scenarios in terms of magnitude and distance pairs represent the solution of an optimisation problem. The errors of magnitude and distance can be evaluated. As an example we perform the calculation for magnitude class 1. For σ_m we have to consider a value of 0.4 magnitude units, because the selected scenario earthquake completely envelopes all scenarios within this magnitude class with respect to the used impact parameter (S_a). The value for σ_r should be evaluated from the spatial distribution of seismicity in the area surrounding the site. For magnitude class 1 we have to consider the two line sources as contributors. For our analysis we use the minimal value for σ_r of both faults (conservative assessment). Based on the data in our example and the theorem of Pythagoras we get for each of the line sources the following relation for the error

$$\sigma_r = \sqrt{a^2 + D^2} - D \tag{35}$$

where *a* is the larger of the two fault sections formed by the perpendicular between the line source and the site; D is the length of the perpendicular (shortest distance in our example). In our example we obtain the value of $\sigma_r = 1.9$ km for line source LS2.

It is important to mention that the location uncertainty associated with an areal source is much higher than for a line source.

The partial derivatives can be calculated from the source-specific attenuation equations. In our example, all considered scenarios originate from line source LS2. Therefore, we have to use the attenuation equation for the line source LS2 for the calculation of the partial derivatives. The partial derivative for m is just the coefficient *b* in our equation (18). For simplicity, we evaluate the uncertainty (as an example) for the spectral frequency of 2.5 Hz. Therefore, b= 0.31787. Then the resulting contribution of magnitude uncertainty to the uncertainty of the attenuation equation is 0.127. The partial derivative with respect to r is:

$$\frac{\partial g(m,r)}{\partial r} = c\frac{1}{r\ln(10)} + d \tag{36}$$

We evaluate the derivative for r at the shortest distance between fault and site neglecting the contribution of depth:

$$r \approx D_{JB} = 25$$

Because the coefficient d is very small we can neglect its contribution. For c=0.556 (line source LS2, 2.5 Hz) we obtain for the resulting contribution of location uncertainty to the uncertainty of the attenuation equation a value of 0.02. Based on equation (18) we can calculate the irreducible part of uncertainty. This irreducible uncertainty is $\sigma_{ired} = 0.191$ (instead of 0.28 obtained from regression). Using the model of a lognormal distribution of spectral accelerations for a given scenario, we can calculate a “mean” hazard spectrum and the required quantile spectra.

We can also calculate the conditional probability of exceedance of our design spectrum. This delivers the required information for a subsequent probabilistic risk assessment.

Figure 12 shows a comparison between the probabilistic “mean” hazard spectrum with the “median” (from our attenuation equations (18) representing the regression mean) and with the deterministic design spectra (the “mean” spectrum and the “mean +1σ” spectrum).The comparison shows that the probabilistic “mean” spectrum always lies below the deterministic “mean +1σ” spectrum and is very close to the deterministic “mean” spectrum. Considering the uncertainty reduction in the example it can be concluded that the deterministic “mean+

1σ"-spectrum effectively corresponds to a "median $+1.5\sigma$" spectrum in the probabilistic analysis. The likelihood that a deterministic design spectrum, which is based on the "mean+1sigma" approach, will be exceeded is very low.

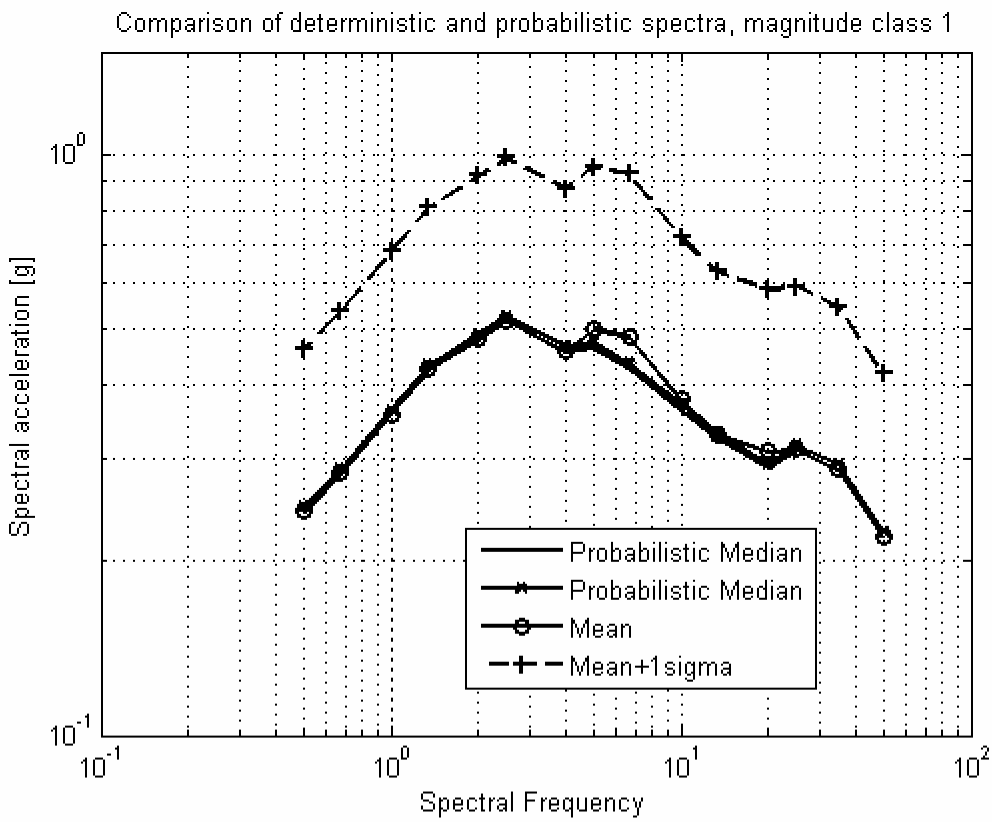


Figure 12 Comparison of probabilistic (magnitude class 1) and deterministic hazard spectra

5 Numerical example 2 – data analysis of an earthquake catalogue

Because the analysis of data from an earthquake catalogue is an important step in the methodology of probabilistic scenario-based seismic risk analysis a second numerical example is given which is based on the Goesgen specific earthquake catalogue. This catalogue covers an area of about 250 km around the site of the Nuclear Power Plant Goesgen. Once again the model of two bivariate exponential distributions of the Gumbel type 2 is used.

5.1 Data from the earthquake catalogue

Table 8 shows all earthquakes with an intensity larger (or equal) 7 registered in the Goesgen earthquake catalogue.

Table 8 Earthquakes of intensity VII (or larger) registered in the surroundings of NPP Goesgen

Location	Day	Month	Year	Time	Latitude	Longitude	Distance to Goesgen [km]	Depth [km]	Moment magnitude M_W	Intensity I_0
Kaiseraugst (Augusta Raurica)	1	1	250	0:00:00	47.5	7.7	25.03212706		6.9	9.0
Strasbourg/F	2	9	1279	0:00:00	48.58	7.75	135.8921768		6.2	8.0
Chur/GR	4	9	1295	0:00:00	46.83	9.53	132.85399	12	6.5	8.0
Basel/BS	18	10	1356	17:00:00	47.55	7.6	34.3975157		6.2	7.5
Basel/BS	18	10	1356	21:00:00	47.46	7.6	29.62901624	12	6.9	9.0
Strassburg/F	14	5	1357	0:00:00	48.16	7.5	94.90919617		5.5	7.0
Thann, Haut-Rhin/F	24	6	1363	0:00:00	47.8	7.1	81.13853185		5.5	7.0
Carpignano,Novara /I	1	2	1369	0:00:00	45.58	8.22	199.7224626		5.1	7.5
Mühlhausen/F	1	6	1372	0:00:00	47.82	7.14	80.10533582		5.5	7.0
Monza	26	11	1396	0:00:00	45.58	9.27	222.5324126		4.8	7.5
Basel/BS	13	12	1428	0:00:00	47.53	7.6	33.13633713		5	7.0
NOERDLIN-GEN/RIES	1	5	1471	0:00:00	48.83	10.5	249.1623822		5.4	7.0
Ardez/GR	1	3	1504	0:00:00	46.78	10.19	181.0528366		5	7.0
Ardon/VS	1	4	1524	0:00:00	46.27	7.27	133.1750937	12	6.4	8.0
Basel/BS	1	6	1572	0:00:00	47.56	7.59	35.6641006		5	7.0
OFFENBURG	1	1	1574	0:00:00	48.51	7.9	127.2263063		5.4	7.0
Geneve/GE	3	5	1574	0:00:00	46.2	6.2	187.3445375		5.5	7.0
Aigle/VD	11	3	1584	11:00:00	46.33	6.96	138.6616285	12	6.4	7.0
NOERDLIN-GEN/RIES	6	2	1593	0:00:00	48.83	10.5	249.1623822		5.4	7.0
Unterwalden	18	9	1601	1:00:00	46.92	8.36	57.97921448	12	6.2	7.0
Basel/BS	29	11	1610	0:00:00	47.56	7.59	35.6641006		5	7.0
Fetan/GR	3	8	1622	0:00:00	46.82	10.23	182.2974659		5	7.0
TUEBINGEN	29	3	1655	0:00:00	48.51	9.07	151.4860812		5.8	7.5
BERGAMASCO	12	3	1661	0:00:00	45.7	9.85	235.0500796		5	7.5
HAUTES-VOSGES (REMIREMONT)	12	5	1682	2:30:00	47.97	6.51	128.4358125		6	8.0
RASTATT	3	8	1728	16:30:00	48.83	8.22	163.8133202	16	4.9	7.5
KARLSRUHE; RASTATT	18	5	1737	21:45:00	48.51	8.13	127.7013878	8	4.4	7.0
Brig, Naters/VS	9	12	1755	13:30:00	46.32	7.98	116.4950578	12	6.1	8.0
Altdorf/UR	10	9	1774	15:30:00	46.85	8.67	78.45040587	12	5.9	7.0
Wisserlen, Kerns OW	7	2	1777	1:00:00	46.9	8.29	57.44723829		5.1	7.0

Location	Day	Month	Year	Time	Latitude	Longitude	Distance to Goesgen [km]	Depth [km]	Moment magnitude M_W	Intensity I_0
Wildhaus/SG	6	12	1795	0:00:00	47.2	9.41	110.7340291		5.3	7.0
Grabs/SG	20	4	1796	6:12:00	47.2	9.41	110.7340291	5	5.3	7.0
MASSIF DU MONT-BLANC (CHAMONIX)	11	3	1817	21:25:00	45.91	6.83	184.0338171		4.8	7.0
BUGEY (BELLEY)	19	2	1822	8:45:00	45.81	5.81	239.5165373		5.6	7.5
AVANT-PAYS JURASSIEN (BESANCON)	30	10	1828	7:20:00	47.28	6.09	142.3664036		5.4	7.0
Birgisch VS	24	1	1837	1:00:00	46.31	7.96	117.605044	12	5.7	7.0
UNTERRIEXIN-GEN	7	2	1839	21:00:00	48.9	9.02	187.6591591	3	4.2	7.0
AVANT-PAYS SAVOYARD (ANNECY)	11	8	1839	20:00:00	45.91	6.13	214.6839965		4.8	7.0
Törbel VS	25	7	1855	11:50:00	46.23	7.85	126.8203435	12	6.4	8.0
Stalden,Visp/VS	26	7	1855	9:15:00	46.23	7.88	126.6782984	12	5.6	7.0
Stalden,Visp/VS	28	7	1855	10:00:00	46.25	7.82	124.78816	12	5.2	7.0
FAUCIGNY (LA ROCHE/FORON)	8	10	1877	5:12:00	46.08	6.32	190.7455028		4.8	7.0
CHABLAIS (ST-JEAN-D'AULPS)	30	12	1879	12:27:00	46.21	6.65	163.415726		5.5	7.0
Bern/BE	27	1	1881	13:20:00	46.9	7.5	62.95772048	12	5	7.0
SCHUTTERWALD	9	10	1886	18:20:00	48.45	7.92	120.5022213	2	4.1	7.0
Nassereith	28	11	1886	22:30:00	47.32	10.84	217.2090567	8	5.1	7.5
PONT S. MARTIN	5	3	1892	0:00:00	45.61	7.8	195.8701352		4.8	7.0
KANDEL	22	3	1903	5:08:00	49.08	8.17	191.1438922	2	4.1	7.0
MASSIF DU MONT-BLANC (LAC D'EMOSSON)	29	4	1905	1:59:00	46.09	6.9	163.8677324		5.7	7.5
MASSIF DU MONT-BLANC (CHAMONIX)	13	8	1905	10:22:00	45.98	6.98	171.8360247		5.2	7.0
Nassereith	13	7	1910	8:32:00	47.32	10.84	217.2090567	8	4.8	7.0
EBINGEN	16	11	1911	21:25:48	48.22	9	122.4097982	10	5.8	8.0
EBINGEN	20	7	1913	12:06:22	48.23	9.01	123.7410777	9	5.2	7.0
KAISERSTUHL	28	6	1926	22:00:40	48.13	7.68	87.55091801	8	4.4	7.0
Bioley-Magnoux VD	1	3	1929	10:32:00	46.73	6.72	118.431959	5	5.3	7.0
Namlos	7	10	1930	23:27:00	47.36	10.66	203.4627812	9	5.3	7.5
RASTATT	8	2	1933	7:07:12	48.85	8.2	165.8619856	6	4.9	7.0
Moudon VD	12	8	1933	9:58:58	46.66	6.8	118.6752414		5	7.0

Location	**Day**	**Month**	**Year**	**Time**	**Latitude**	**Longitude**	**Distance to Goesgen [km]**	**Depth [km]**	**Moment magnitude M_W**	**Intensity I_0**
SAULGAU	27	6	1935	17:19:30	48.04	9.47	135.3618804	10	5.7	7.5
ONSTMETTINGEN	2	5	1943	1:08:02	48.27	8.98	125.8380354	13	5.1	7.0
ONSTMETTINGEN	28	5	1943	1:24:08	48.27	8.98	125.8380354	9	5.6	8.0
Ayent VS	25	1	1946	17:32:00	46.35	7.4	121.1655748	12	6.1	8.0
Ayent VS	30	5	1946	3:41:00	46.3	7.41	126.121243	12	6	7.0
FORCH-HEIM/RHEIN	7	6	1948	7:15:19	48.97	8.33	180.3361366	6	4.3	7.0
LUDWIGSHA-FEN/RHEIN, WORMS	24	2	1952	21:25:30	49.5	8.32	238.7211049	8	4.3	7.0
OUTRE-FORET (WISSEMBOURG)	8	10	1952	5:17:00	48.95	7.98	176.0881761		4.7	7.0
Brig/VS	23	3	1960	23:10:00	46.37	8.02	111.0019708	12	5.3	8.0
Flüeli OW	17	2	1964	12:20:00	46.88	8.26	58.56983263	5	5	7.0
Alpnach/OW	14	3	1964	2:39:00	46.86	8.31	62.12555249		5.7	7.0
Chablais (Abon-dance)	19	8	1968	0:36:41	46.29	6.55	161.4701804	10	5.2	7.0
Balingen/Swabian Jura	26	2	1969	1:28:01	48.29	9.01	128.960037	8	4.7	7.0
EBINGEN	22	1	1970	15:25:17	48.28	9.03	128.9996274	8	4.8	7.0
JURA	21	6	1971	7:25:00	46.35	5.7	206.7325786		4.8	7.0
Ebingen/Swabian Jura	18	5	1972	8:11:01	48.28	9.03	128.9996274	8	4.4	7.0
Ebingen/Swabian Jura	3	9	1978	5:08:32	48.28	9.03	128.9996274	6	5.2	7.5
AVANT-PAYS SAVOYARD (EPAGNY-ANNECY)	15	7	1996	0:13:31	45.93	6.09	215.0295066	2	4.59	7.0
Rambervillers, Saint Dié, F	22	2	2003	20:41:03	48.37	6.64	149.4	10	4.8	7.1[1]
Lago di Garda	24	11	2004	22:59:00	45.6	10.6	282	9	5	7.5

Based on equations (18) to (20) we can develop the model parameters for our bivariate distributions describing the dependence of earthquake size (magnitude) on the elapsed time between two sequential events as well as between earthquake size (magnitude) and distance between earthquake location and site.

---

[1] Calculated value

The parameters of the bivariate distribution models (with respect to the dependency between size and elapsed time and between size and distance) are as follows:

- Dependence between Moment magnitude M_W – elapsed time t (between two sequential earthquakes)
    - o λ_{Mw} = 0.189049662
    - o λ_t = 0.044447897

with a sample correlation coefficient

- o ρ_{Mwt} = 0.184420097 and a corresponding value of
- o α = 0.737680388.

- Dependence between Moment magnitude M_W – distance d (between the location of the earthquake and Goesgen site)
    - o λ_{Mw} = 0.189049662
    - o λ_d = 0.007079397

with a sample correlation coefficient ρ_d = -0.391348109 and a corresponding α = -1.565392435.

The lower magnitude truncation limit is $m_l =$ 4.1, while the upper bound magnitude based on historical data was assessed to be $m_u =$7.5. This value was found from a bootstrap procedure after rounding to the first decimal.

Once correlation is observed it shall not be neglected in the later analysis. Nevertheless it is meaningful to provide a statistical test procedure.

5.2 Procedure to test the statistical significance of the model

For practical application with respect to risk analysis for the NPP Goesgen it is important to test the statistical significance of the model against the earthquake data from the Goesgen catalogue (Müller, 2005). In our case it is sufficient to test whether the assumption of correlation between earthquake size and elapsed time as well as between earthquake size and distance between earthquake location and site cannot be rejected. It is well known that to test the null hypothesis $H_0 : \rho = 0$ against $H_1 : \rho \neq 0$, one can use the statistic, which has a Students' t distribution with n-2 degrees of freedom.

$$T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \tag{37}$$

Thus, for a sufficiently large value of n, in order to obtain significance of the test, at say a significance level of 0.05, we need to have

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} > 1.96 \tag{38}$$

For a fixed number of datasets n, in order to capture a significant test, we need to have

$$|r| > \frac{1.96}{\sqrt{n+1.96}} \tag{39}$$

This criterion can easily be checked for the data from the Goesgen specific earthquake catalogue calculating the empirical correlation coefficient $\hat{\rho}$.

In our case a statistical correlation between magnitude and elapsed time cannot be excluded (5% confidence level, 78 records) if the absolute value of the correlation coefficient is larger than 0.219. The observed correlation coefficient is smaller, but not by much. A correlation between magnitude and distance between earthquake location and site cannot be excluded. This correlation is very pronounced and negative. This means that with increasing size earthquakes tend to occur closer to the site of the plant.

5.3 Application of the model

The developed probabilistic model can be used to calculated the magnitude of the largest earthquake to be expected during the residual life time of the Goesgen plant (33 years, total 60 years of operation) as well as the expected value of the distance between earthquake location and site.

Figures 13 and 14 show the dependence of earthquake size (magnitude) on the elapsed time between two subsequent events and o the distance between earthquake location and site.

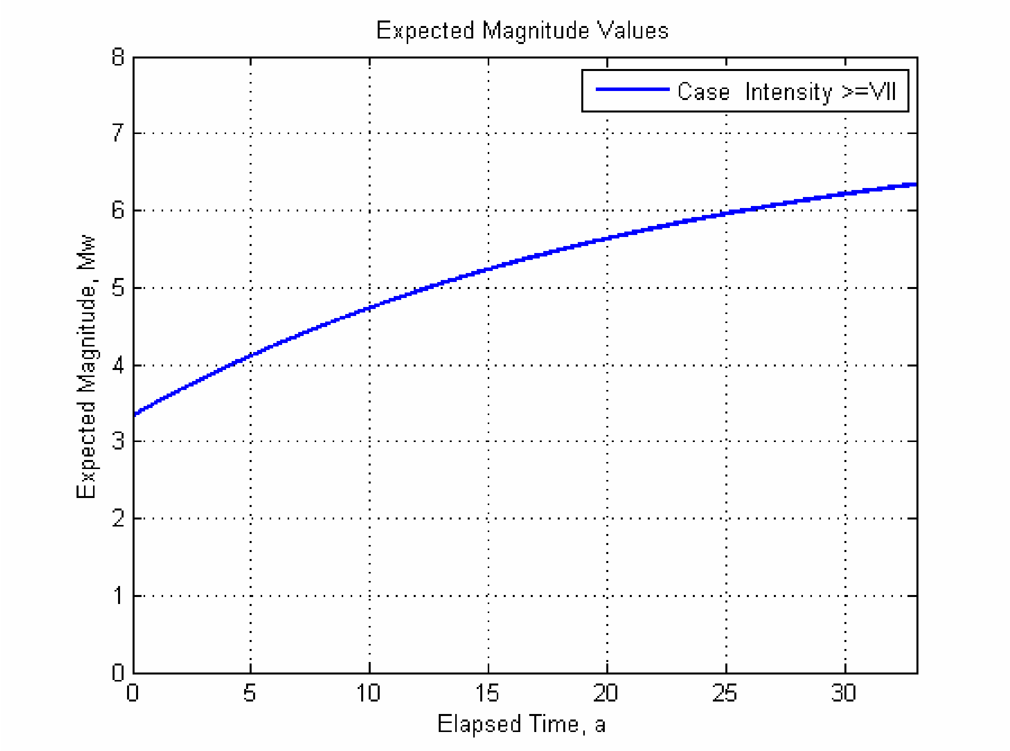


Figure 13 Expected maximum earthquake during the residual lifetime of NPP Goesgen

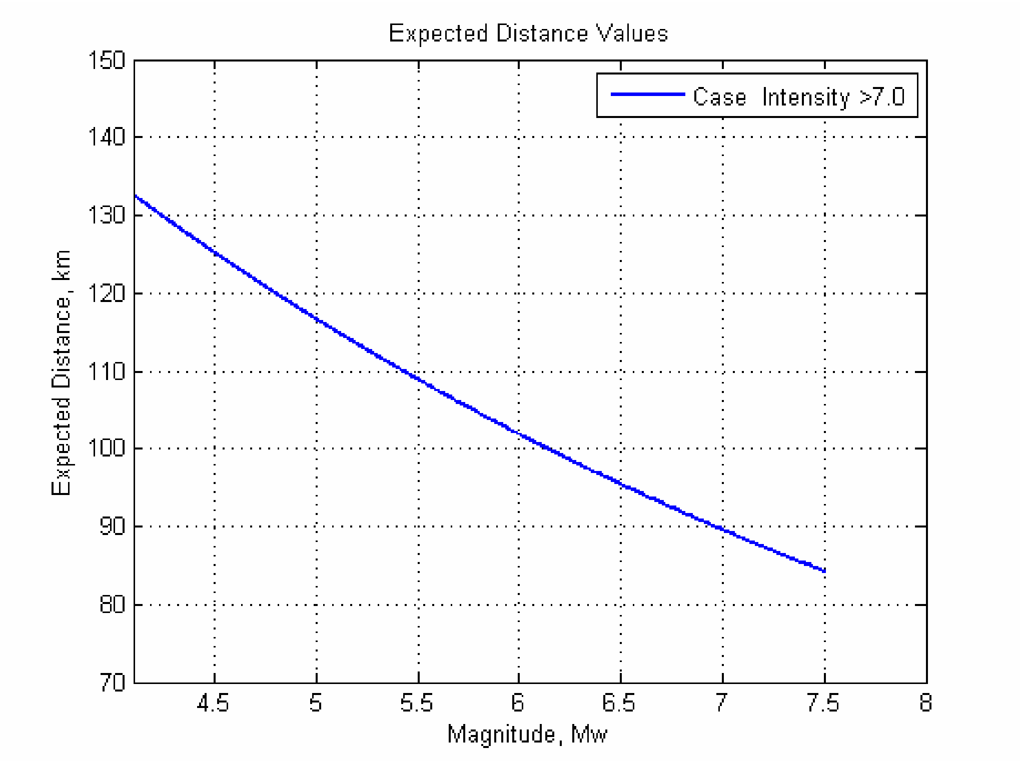


Figure 14 Expected location of earthquakes (distance to Goesgen site) in dependence of earthquake magnitude.

According to the results an earthquake of magnitude 6.3 at a distance of about 97 km from the Goesgen site represents the expected maximum scenario earthquake which cannot be excluded during the residual lifetime of NPP Goesgen. This earthquake can be regarded as a functional design basis (operational design basis) earthquake. The design of the plant shall ensure a continuation of plant operation after the occurrence of such an event.

6 Summary and Conclusions

A procedure for a probabilistic scenario-based seismic risk analysis was developed that allows to incorporate all available information (in a geological, seismo-tectonic and geotechnical database of the site of interest), and advanced physical modelling techniques to provide a reliable and robust basis for seismic risk analysis. The procedure is in full compliance with the likelihood prionciple and contemporary methods of risk analysis. Probabilistic scenario-based seismic hazard analysis can produce the necessary input for probabilistic risk assessment (PRA), as required by safety analysts and insurance companies. The scenario-based approach removes the ambiguity of the results of traditional probabilistic seismic hazard analysis (PSHA), which relies on the projections of Gutenberg-Richter (GR) equation, as practiced in some countries. Two numerical examples illustrate the application of the method based on the model of bivariate exponential distributions of the Gumbel type 2 for the dependency between earthquake size (magnitude) and elapsed time and the corresponding conditional probability distribution between earthquake size (magnitude) and distance between earthquake location.

7 References

Abrahamson, N.A., 2006. Deterministic and Probabilistic Assessments, Northwest Dam Safety Regional Forum, Portland, Oregon. (http://www.ferc.gov/industries/hydropower/safety/wkshps/nw-forum.asp)

Aki, K. and Richards P.G., 2002. Quantitative Seismology, Second Edition. University Science Books, 700 p.

Ang, A. H-S. (1970). Extended Reliability Basis of Structural Design Under Uncertainties, Annals of Reliability and Maintainability Vol. 9, AIAA, 642-649.

Ang, A., Tang, W., (1975), Probability Concepts in Engineering Planning and Design, Vol. I. John Wiley, New York.

Ang, A., and Tang, W., (1984), Probability Concepts in Engineering Planning and Design, Vol. II. John Wiley, New York.

Arias, A., 1970. A measure of earthquake intensity. In Seismic Design for Nuclear Power Plants, Ed. R.J. Hansen, MIT Press, Cambridge, Massachusetts.

Ayyub, B.M., and Klir, G.J., Uncertainty Modeling and Analysis in Engineering and the Sciences, Chapman &Hall/CRC, Boca Raton, London, New York, 2006, 378 p,.

Baker, J.W. and Cornell, C.A., 2006. Correlation of Response Spectral Values for Multicomponent Ground Motions, BSSA, 96, pp. 215-227.

Bolt, B.A. and Abrahamson, N.A., 2002. Estimation of Strong Ground Motion, International Handbook of Earthquake Engineering Seismology, IASPEI, 2002.

Budnitz, R.J., Chokshi, N.C., Henries, W., Ravindra, M.K., Stevenson, J. and Yee, T., 2003. External Events PRA Methodology Standard, American Nuclear Society Standard BSR/ANS 58.21.

Cancani, A. 1904. Sur l'emploi d'une double echelle seismique des intensites, empirique et absolue, G. Beitr., 2, 281-283.

Cornell, C. A., 1968. Engineering seismic risk analysis, Bull. Seis. Soc. Am., 58, 1583-1606.

Decanini, L., Mollaioli, F. and Oliveto, G., 1994. Observations and lessons learned from the earthquake of 13th December 1990 in South-East Sicily, Proc. of the 10th European Conference on Earthquake Engineering, Duma (Editor), Balkema, Wien, 1935-1943.

DOE, 1996. Accident analysis for aircraft crash into hazardous facilities, Department of Energy, DOE-STD-3014-96.

Edwards, A.W.F., 1972. Likelihood. Cambridge: University Press.

EPRI, 1991. Standardization of the Cumulative Absolute Velocity. TR-10082.

EPRI, 2005. Program on Technology Innovation: Use of CAV in determining effects of small magnitude earthquakes on seismic hazard analyses, TI-1012965.

Field, E.H., 2000. The SCEC Phase III Working Group. Accounting for site effects in probabilistic seismic hazard analyses of Southern California: overview of the SCEC Phase III report," Bull. Seism. Soc. Am., 90, 6B, S1-S31.

Frankel, A. 1995. "Mapping seismic hazard in the Central and Eastern United States," Seism. Res. Lett., 66(4), 8-21.

IAEA, 1995. Treatment of external hazards in probabilistic safety assessment for nuclear power plants, Safety Series No. 50-P-7. Vienna.

IAEA, 2002a. Evaluation of seismic hazards for nuclear power plants, Safety Guide No. NS-G-3.3, Vienna.IAEA, 2002b, Review of probabilistic safety assessments by regulatory bodies, Safety Report Series No. 25, Vienna.

Kaplan, S., Garrick, B.J. 1981. On the Quantitative Definition of Risk, Risk Analysis, 1: 11-27,

Kijko, A., Graham, G. (1998). Parametric-historical procedure for probabilistic seismic hazard analysis, Proceedings Eleventh European Conference on Earthquake Engineering, Paris.

Klügel, J.-U., 2005c. Reply to the comment on J.U. Klügel's: "Problems in the Application of the SSHAC Probability Method for Assessing Earthquake Hazards at Swiss Nuclear Power Plants", Engineering Geology, Vol. 78, pp. 285-307, by Musson et al. Engineering Geology, Vol. 82, pp. 56-65.

Klügel, J-U., 2005e. Paper KM02_2. On the Use of Probabilistic Seismic Hazard Analysis as an Input for Seismic PSA. 18th International Conference on Structural Mechanics in Reactor Technology (SMiRT 18) Beijing, China, August 7-12, 2005.

Klügel, J.-U., Mualchin, L., Panza, G.F., 2006, A scenario-based procedure for seismic risk analysis, Engineering geology 88 (2006), pp. 1-22.

Kostov, M.K., 2005. Site specific estimation of cumulative absolute velocity. 18th International Conference on Structural Mechanics in Reactor Technology (SMiRT 18), Beijing, SMiRT18-K03-4.

McGuire, R. K., 1976. FORTRAN computer program for seismic risk analysis: U.S. Geological Survey Open-File Report 76-67.

Mc Guire, R. K., 1995. Probabilistic seismic hazard analysis and design earthquakes: Closing the loop: Bulletin of the Seismological Society of America, v. 85 p. 1,275-1,284.

Mollaioli, F., Decanini, L., Bruno, S. and Panza, G.F., 2003. Analysis of the response behaviour of structures subjected to damaging pulse-type ground motions. OECD Workshop on the Relations between Seismological DATA and Seismic Engineering, Istanbul, 16-18 October 2002, NEA/CSNI/R (2003) 18, 109-119.

Noubary, R., 2000. Statistical Methods for Earthquake Hazard Assessment and Risk Analysis. New York: Nova Science Publishers, Inc.

Panza, G.F., Romanelli, F., Vaccari, F., Decanini, L.and Mollaioli, F. 2000b. Contribution of the deterministic approach to the characterisation of seismic input, OECD-NEA Workshop on Engineering characterisation of Seismic Input, BNL, Upton, New York, 15-17 November, 1999, NEA/CSNI/R(2000) 2, 655-715.

Poloski, J.P., Marksberry, D.G., Atwood, C.L. and Galyean, W.J., 1999. Rates of initiating events at U.S. Nuclear Power Plants: 1987 – 1995, NUREG/CR-5750.

Richwalski,S.M., Roth, F. and Wang, R., 2004. Erdbebensimulation von Bodenerschütterungen im Bereich kritischer Bauwerke. In: Merz, B., Apel, H. (Eds.), Risiken durch Naturgefahren in Deutschland. Abschlussbericht des BMBF-Verbundprojektes Deutsches Forschungsnetz Naturkatastrophen (DFNK), pp. 161-170.

Romanelli, F., Vaccari, F., Panza, G.F., 2003. Realistic Modelling of the Seismic Input: Site Effects and Parametric Studies, JSEE, Fall 2003, Vol. 5 (3) pp. 27-39.

Scholz, C.H., 2002. The Mechanics of Earthquakes and Faulting. Cambridge: Cambridge University Press (2nd edition).

Senior Seismic Hazard Analysis Committee (SSHAC). 1997. Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts. NUREG/CR-6372.

Travasarou, T., Bray, J.D. and Abrahamson, N.A. 2003. Empirical attenuation relationship for arias intensity. Earthquake Engineering and Structural Dynamics, 32,1133-1155.

Tregoning, R., Abramson, L. and Scott, P., 2005. Estimating Loss of Coolant Accident (LOCA) frequencies through elicitation process, NUREG-1829, Draft report for comment.

Veneziano, D., Cornell, C.A. and O'Hara, T. (1984). Historical method of seismic hazard analysis, Electrical Power Research Institute Report NP-3438, Palo Alto, California.

Wang, Z. (2005). Discussion, Klügel, J.-U. 2005. Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants, Engineering geology, Vol. 78, pp. 285-307., Engineering Geology, in Press.

Wells, D. L. and Coppersmith, K. J., 1994. New Empirical Relationships Among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement: Bulletin of Seismological Society of America, v. 84, No. 4, pp. 974ff.

Woo, G. (1996). Kernel estimation methods for seismic hazard area source modelling, Bull. Seism. Soc. Am., 86, 353-362.