11 WCEE
Paper No. 721. (quote when citing this article)
Eleventh World Conference on Earthquake Engineering
ISBN: 0 08 042822 3

BAYESIAN ESTIMATION OF SEISMIC HAZARD FOR SWITZERLAND

RÜTTENER Erik[1], MAYER-ROSA Dieter[1], EGOZCUE, Juan José[2]

[1]Swiss Seismological Service, Institute of Geophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland (erik@seismo.ifg.ethz.ch)

[2]Universidat Politécnica de Cataluña, ETSICCP, E-08034 Barcelona, Spain

ABSTRACT

The main primary inputs to a seismic hazard analysis are historical and instrumental earthquake catalogs and attenuation laws. Especially historic earthquake data are often imprecise, i.e. the earthquake size and location can only be estimated with uncertainties. Models based on probability distributions have been developed in order to quantify and represent these uncertainties. Traditional seismic hazard procedures do not take into account these uncertainties. Therefore, a procedure based on Bayesian statistics was developed to estimate return periods for different ground motion intensities (MSK scale). The method estimates the probability distribution of the number of occurrences in a Poisson process described by the parameter λ. The input are the historical occurrences of intensities for a particular site, represented by a discrete probability distribution for each earthquake. It can be shown, that the variance is smaller in regions with higher seismic activity. It can also be demonstrated that long return periods cannot be estimated with confidence, because the time period of observation is too short. This indicates, that the long return periods obtained by seismic source methods only reflects the delineated sources and the chosen earthquake size distribution law.

KEYWORDS

historical earthquakes; seismic hazard; bayesian estimation; data uncertainty; probability distribution; macroseismic intensity

INTRODUCTION

Earthquake hazard analysis is worldwide of considerable importance, either due to a significant seismic activity or the high economic values that are at risk. Seismic hazard analysis is usually based on historic earthquake records as well as instrumental data. In Switzerland the historic record goes back to the 13th century, but is only complete for earthquakes with intensity IX or greater for the complete time period. This historic earthquake data have to be used in seismic hazard analysis to get a realistic description of return periods, since the time period of instrumental recording is too short to reveal the complete pattern of seismic activity. However, historic earthquake data are often inaccurate and uncertain. In order to properly use these data, the uncertainty has to be quantified and taken into account in the hazard analysis.

Various procedures are available for evaluating seismic hazard (Cornell (1969), McGuire (1976)), which all use a three step procedure consisting of a source seismicity model, an attenuation model and an exposure evaluation model. These traditional methods of calculating seismic hazard, often called "deductive" or "seismic source" methods do not treat the observations uncertainties completely. Usually they take only the uncertainties in the ground motion attenuation model into account. Bender and Perkins (1987) developed a procedure which also treats earthquake locations uncertainties, but there exists yet no method which completely accounts for the errors of the location and size of each earthquake.

Therefore, a new approach based on Bayesian statistics was developed (cf. Rüttener (1995), Egozcue and Rüttener (1996)). In this method, the mean return period in a Poisson process and its variance is estimated taking into account the previously defined uncertainties of epicenter locations, earthquake size and attenuation relations. Before these uncertainties can be taken into account, they have to be quantified and appropriate models have to be defined. Therefore, probability distributions have been introduced to characterize the uncertainties of each earthquake individually.

HISTORICAL EARTHQUAKE DATA

The estimation of the location and size of historical earthquakes is based on historical observations. As historical reports are often ambiguous and incomplete, the derived earthquake parameters are always accompanied by some uncertainties. Usually, these uncertainties depend on the size of the earthquake, on the historic period in which it occurred and on the population density in the epicentral area. When a historic earthquake catalog is compiled, these information is rarely used to quantify the uncertainty of the size and location of the event. Therefore, the historic earthquake catalog was reevaluated and errors for size and location of historic earthquakes were introduced.

Four different error classes were used to describe the error of the epicentral intensity estimates. These four classes are: intensity error of (1) 0 degrees, (2) ± 0.5 degrees, (3) ± 1.0 degrees and (4) ± 2.0 degrees. These error classes were converted into probability distributions, whereby the following conditions have been taken into account: a general tendency in historical sources to report maximum observed damages, and a probable underestimation of epicentral intensities due to location errors. Figure 1 shows the used probability distribution for the 4 error classes. It is clear, that the probability distributions are based on subjective interpretations of historical earthquake data. Therefore, they have to be tested in combination with intensity attenuation laws. The tests which have been carried out (Rüttener, 1995) have shown that these probability distributions in combination with the attenuation laws for Switzerland model the observed intensity distribution pattern accurately. It is worth to mention that the description of errors by probability distributions is flexible enough to represent a great variety of historical information. If for example historical observations can be summarized as "the earthquake had an epicentral intensity of at least VII (MSK) but was not greater than intensity IX", a probability distribution with equal probabilities for the intensity range from VII to IX can be used to model this earthquake.

Also the epicenter location errors are modeled by probability distributions. Instrumental as well as macroseismic locations can be seen as maximum likelihood estimations. In this context a maximum likelihood estimation of the earthquake location can best be modelled by a two-dimensional normal distribution, whose density for circular symmetry with covariance equal 0 is given by (1). More explicitly, the probability P that the epicenter lies in an area $(\Delta x \Delta y)$, assuming that the most probable epicenter is at (0,0) and that the standard deviation is σ, is equal to the integral over this area of the probability density function (2):

$$f(x_1, y_1) = \left(\frac{1}{2\pi\sigma^2}\right)\exp\left(-\frac{1}{2\sigma^2}\left(x_1{}^2 + y_1{}^2\right)\right) \qquad (1)$$

and

$$P(\Delta x, \Delta y) = \int_{\Delta x \Delta y} f(x, y)\, dx dy \qquad (2)$$

where x_1, y_1 are the point of interest, σ the standard deviation (location error) and $\Delta x, \Delta y$ the area for which the probability is calculated.

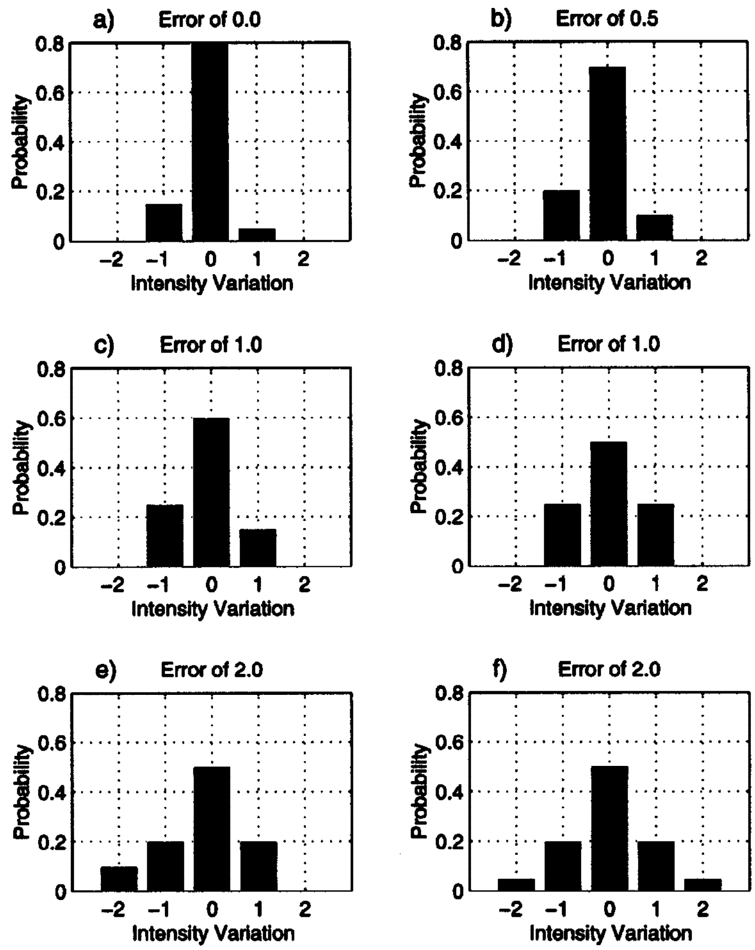

Figure 1: Conversion scheme for intensity errors into probability distributions. Probability distribution for: a) error of 0, b) error of ± 0.5, c) error of ± 1 and location error ≤ 10 km, d) error of ± 1 and location error >10 km, e) error of ± 2 and location error ≤ 10 km, and f) error of ± 2 degrees and location error >10 km.

From the instrumental data catalog accurate information about the depth distribution of earthquakes in Switzerland is gained. The focal depth distribution for well-constrained earthquakes reveals that in the Alpine Belt, focal depths are restricted to the upper most 15 km of the crust. In contrary, the earthquakes are distributed throughout the entire crust (up to 30 km) in the Alpine foreland (Deichmann and Baer (1990)). This depth distribution is considered in the final model for hazard analysis also by probability distributions.

Once the models for earthquake location uncertainty, earthquake size uncertainty, focal depth distribution and intensity attenuation are established, the probable ground motion for a specific earthquake can be modelled for any site. In this study, a probabilistic intensity attenuation law was used, which takes into account the discrete character of macroseismic intensities and their uncertainties (cf. figure 1). If the probable ground motion that is produced at a particular site is calculated for each event in the earthquake catalog, the historic occurrences of ground motions for this particular intensity is obtained. These ground motion occurrences will be called "Earthquake Site Catalog" (cf. figure 2). The earthquake site catalog contains for each earthquake a discrete probability distribution of the ground motion that occurred at this particular site.

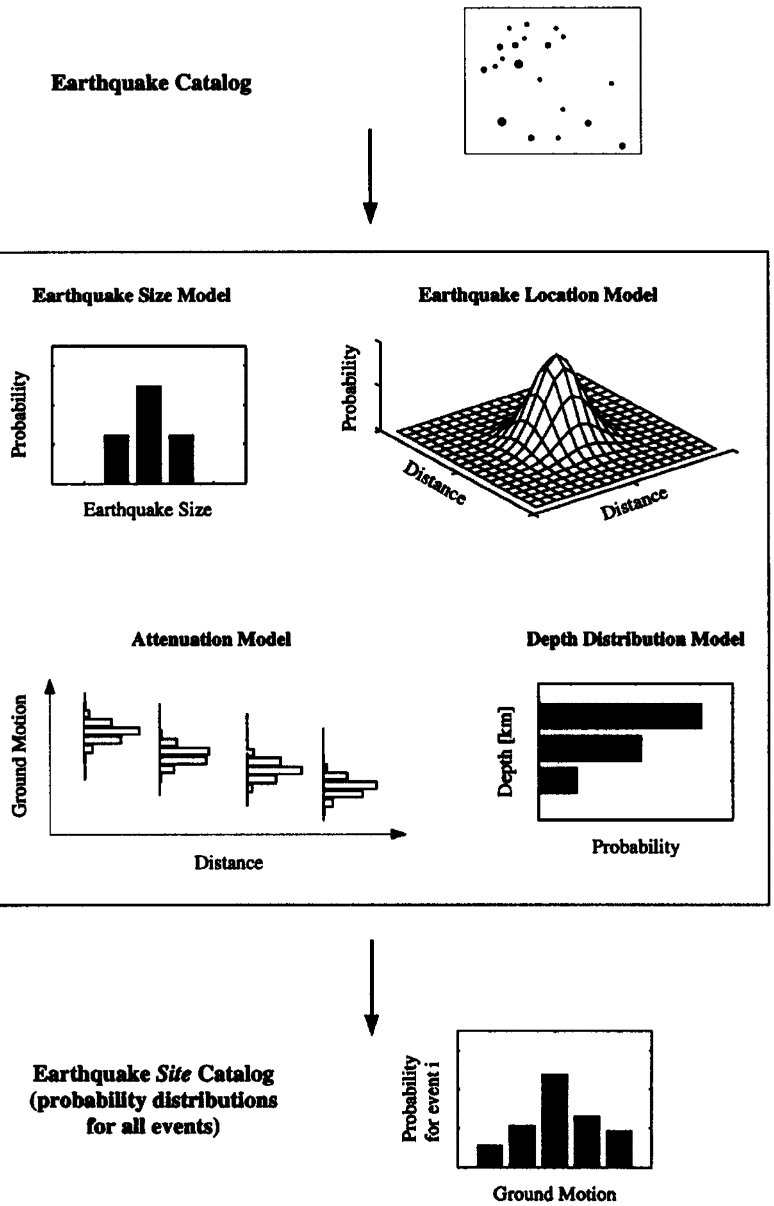

Figure 2: Steps involved in the calculation of "Earthquake Site Catalogs". The ground motion probability distribution is calculated for each earthquake taking into account uncertainties in size, location, ground motion attenuation and depth distribution model.

BAYESIAN ESTIMATION OF SEISMIC HAZARD

The method developed is based on Bayesian statistic. Bayesian statistical theory provides a mathematical model for incorporating statistical and model uncertainties as well as individual and more subjective elements (Benjamin and Cornell (1970), Press (1989)). Bayesian estimation techniques have been applied in seismic hazard analysis either in order to estimate single input parameters for standard seismic hazard approaches (Mortgat and Shah, 1979) or to update the results of the seismic hazard analysis by the observed data (Egozcue et al., 1991).

In this study a Bayesian method is presented that estimates the probability distribution of the mean number of occurrences in a Poisson process described by the parameter λ, taking into account the previously defined uncertainties of the input parameters. The parameter λ corresponds to the mean number of occurrences of a given intensity, but the method can be used similarly if λ corresponds to an intensity range or an intensity threshold. The inherent uncertainty of the parameter λ requires a treatment of λ as a random variable. We further assume that the probability distribution of λ is proportional to a Gamma distribution with parameters ν and κ. This somehow arbitrary choice is justified, because the Gamma function is able to fit a large variety of shapes and, therefore, does not introduce substantial limitations in the model. The prior estimate of λ can thus be written as (Benjamin and Cornell, 1970):

$$f'_{\lambda}(\lambda) = \frac{\nu^{\kappa}}{\Gamma(\kappa)}\lambda^{\kappa-1}\exp(-\nu\lambda) \tag{3}$$

where Γ is Euler's Gamma function and ν, κ are the parameters of the Gamma distribution. The parameters ν and κ are directly related to the expectation and variance of λ:

$$E[\lambda] = \frac{\kappa}{\nu} \text{ and } Var[\lambda] = \frac{\kappa}{\nu^2} \tag{4}$$

We have assumed that the number of occurrences of a ground motion level follows a Poisson process with parameter λ. The sample likelihood function on λ, when n occurrences have been observed in the time period T, is:

$$L(\lambda|T, N=n) = \frac{(\lambda T)^n}{n!}\exp(-\lambda T) \tag{5}$$

where T is the time period of observation and n the number of occurrences in this time period. By applying Bayes' rule (Press, 1989), the posterior distribution f''_{λ} of λ is:

$$f''_{\lambda}(\lambda) = f''_{\lambda|N=n}(\lambda) = NL(\lambda|T, N=n)f'_{\lambda}(\lambda) \tag{6}$$

$$= \frac{(\nu+T)^{\kappa+n}}{\Gamma(\kappa+n)}\lambda^{\kappa+n-1}\exp(-(\nu+T)\lambda) \tag{7}$$

From formulas (3) and (7) follows, that:

$$E''[\lambda|(\nu,\kappa)] = \frac{\kappa+n}{\nu+T} \text{ and } Var''[\lambda|(\nu,\kappa)] = \frac{\kappa+n}{(\nu+T)^2} \tag{8}$$

From formula (8) it can easily be seen that the larger T is, the smaller will be the variance in λ, i.e. that the longer the time period of observation is, the smaller will be the uncertainty in λ.

Formula (7) is valid only if the number of occurrences N=n in time T is known precisely. If the sample is imprecise itself, i.e. if the number n of occurrences can only be estimated with uncertainty, the number n has also to be treated as a random variable. This means that the number of occurrences N also follows a probability distribution P[N=n] with n=0,1,,,. The Bayesian estimate (7) has now to be rewritten for the value of N=n, which yields the weighted Bayesian estimate:

$$\tilde{f}_{\lambda}(\lambda) = \sum_{n=0}^{\infty} f''_{\lambda|N=n} \cdot P[N=n] \tag{9}$$

Formula (7) can now be expressed as:

$$\tilde{f}_{\lambda}(\lambda) = \sum_{n=0}^{\infty} \frac{(\nu+T)^{\kappa+n}}{\Gamma(\kappa+n)} \lambda^{\kappa+n-1} \exp(-(\nu+T)\lambda) P[N=n] \quad (10)$$

Combining (8) with (10) yields:

$$\tilde{f}_{\lambda}(\lambda) = \sum_{n=0}^{\infty} \frac{\nu''^{\kappa''}}{\Gamma(\kappa'')} \lambda^{\kappa''-1} \exp(-\nu''\lambda) P[N=n] \quad (11)$$

Formula (11) will be used for the calculation of posterior distributions of the parameter λ, if the sample data are imprecise. The probabilities P[N=n] are calculated from the "Earthquake Site Catalog" (ESC). The given probabilities p_i in the ESC can be interpreted as probabilities of a Bernoulli trial. Then P[N=n] is obtained by a standard Bernoulli trial with varying probabilities p_i for each earthquake (Egozcue and Rüttener, 1996). The expectation and variance of λ is now:

$$\tilde{E}[\lambda|(\nu,\kappa)] = \frac{\kappa+E(N)}{\nu+T} \text{ and } \tilde{Var}[\lambda|(\nu,\kappa)] = \frac{\kappa+E(N)+Var(N)}{(\nu+T)^2} \quad (12)$$

This result shows that the mean of the weighted posterior distribution is that of the standard posterior distribution (8), if the number of observations were E(N). But the uncertainty of n influences the variance of the posterior distributions, i.e. the larger the variance in the number of observations is, the larger the variance of λ will be.

The posterior distribution of the Poisson parameter λ allows to calculate both, point estimates of λ (e.g. mean values, modes, median) and interval estimates of λ. Point estimates from probability densities, e.g. the mean of the posterior density of λ, give a representative point of the value of λ, but do not show the variance of the obtained results. Therefore, whenever possible, probability intervals should be given to indicate the accuracy of the recurrence rate λ. This is valid for the standard Bayesian estimation (with exact data) as well as for the weighted Bayesian estimation (with imprecise data).

APPLICATION

The developed method is used to estimate seismic hazard for two points in Switzerland: for the city of Zurich, which has a moderate seismicity and where no damaging earthquakes in the historical earthquake catalog are known, and for the city of Brig, which suffered several damaging earthquakes. In the first step, probability distributions of the ground motion of each individual earthquake are calculated for these two sites, i.e. the "Earthquake Site Catalogs" are calculated taking into account the observed uncertainties (cf. figure 2). Since each earthquake can be interpreted as a trial in a Bernoulli process, the number of occurrences of a given ground motion level at a site is obtained in the form of a discrete probability distribution. Figure 3 shows the calculated probabilities versus number of occurrences of intensity VI from the year 1750 for the city of Brig. The peak in the number of occurrences is recognized to be 7, i.e. most probably Brig experienced 7 times intensity VI in the last 243 years. However, there is a great uncertainty about the number of occurrences, i.e. all occurrences between 5 and 10 times have a significant probability.

In the second step, the probability density of return periods is calculated using the formula developed (11). A "non-informative" prior distribution of the parameter λ with the condition that return periods increase with increasing intensities are used. The "non-informative" prior distribution is represented by $\kappa = 1, \nu = 0$ as parameters of the Gamma distribution. This assures that the results are only determined by the data sample and not by the form of the prior distribution. Representative points of the estimated return periods, e.g. median values or intervals estimates, are calculated from the probability distributions in order to draw seismic hazard curves. Figure 4 shows the result for the two sites in Switzerland. Dotted lines with crosses give median return periods, dotted lines with black points represent the 0.25 and 0.75 bounds of the probability intervals, and dotted lines with open circles represent the 0.05 and 0.95 bounds of the probability intervals. Also the results obtained by Mueller and Mayer-Rosa (1980) following a modified Cornell approach are indicated (by dashed lines). It can be seen that the variance increases with higher intensities, which reflects the short data sample for the bigger earthquakes. It can also be seen that the variance in the estimated return periods is smaller in Brig than in Zurich. A result of the longer return periods which are estimated for Zurich. In Zurich, i.e. in a seismic low active region, the time interval of the data is very short compared to the return periods that are estimated. The variance is generally larger in regions with a lower seismic activity than in regions with an increased seismic activity, if the data catalog is of a similar length in time.

Results of the earlier study (Mueller and Mayer-Rosa (1980)) lie within the 90% probability interval of the Bayesian estimates. But it has to be stressed that for the Bayesian estimation of the seismic hazard neither seismic sources have been delineated nor earthquake size distribution laws (e.g. Gutenberg-Richter relation) have been applied. The obtained results purely reflects the past seismicity taking all the uncertainties into account. If long return periods in comparison to the time interval of the data catalog are estimated, the variance gets very big and the results are almost meaningless. However, it demonstrates that the results obtained by seismic source methods purely reflects the delineated sources and the defined earthquake size distribution and cannot be confirmed by the data.

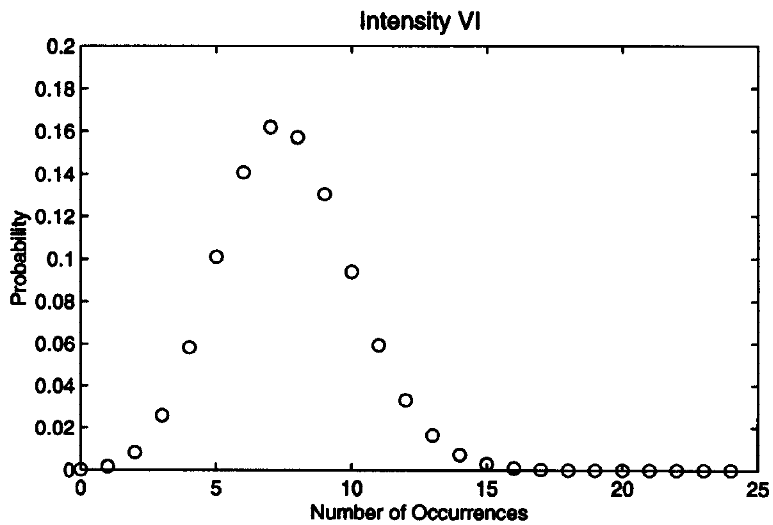

Figure 3: Probability versus number of occurrences of intensity VI for the city of Brig.

CONCLUSIONS

Earthquake data catalogs and ground motion attenuation laws are the main geophysical input into seismic hazard analysis. However, most of these input data are affected by large uncertainties. These uncertainties can be properly represented by probability distributions. Bayesian Statistics is an appropriate technique to estimate return periods when uncertainties are relevant. A Bayesian model for estimating seismic hazard analysis has been developed. It takes the uncertainties in the input data into account. The resulting seismic hazard reflects these uncertainties in probability distributions of the return period. It can be observed that in regions with low seismic activity the estimated return periods have a greater variance than in more active regions. Large return periods cannot be determined with confidence, because the historic record is too short, neither can the results obtained by seismic source methods be confirmed. Additional data (e.g. paleoseismological data) have to be used in order to reduce the spread of the probability interval. However, when such data are available, they can readily be integrated into the method developed.

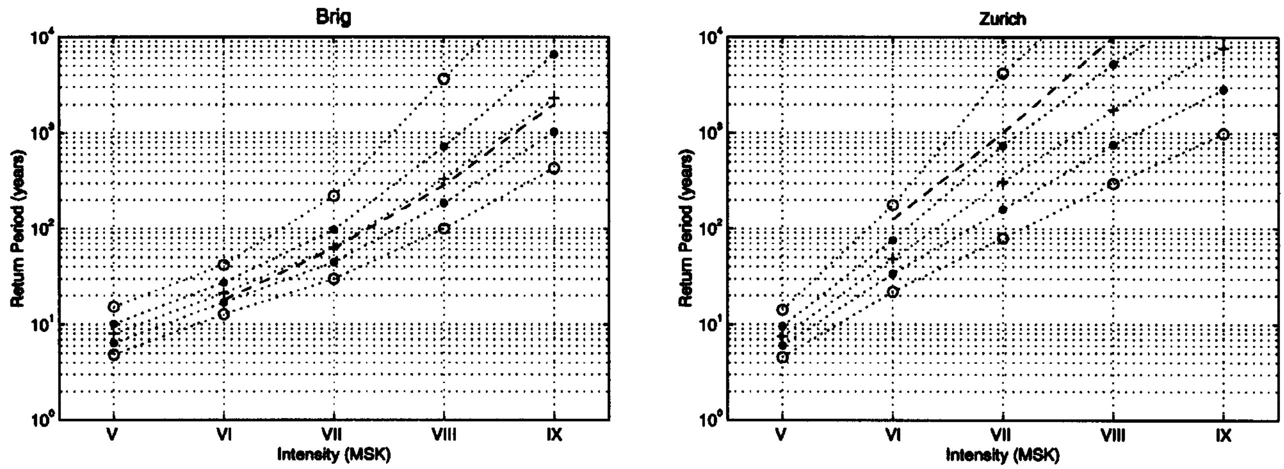

Figure 4: Return period versus intensity for two locations in Switzerland.

REFERENCES

Benjamin, J.R. and C.A. Cornell (1970): Probability, Statistics and Decision for Civil Engineers. McGraw-Hill, New York, 684 pp.

Bender, B. and D.M. Perkins (1987): Seisrisk III: A Computer Program for Seismic Hazard Estimation. *U.S. Geological Survey Bulletin* 1772, Denver, Colorado.

Cornell, C.A. (1968): Engineering Seismic Risk Analysis. *Bull. Seism. Soc. Am.*, 58, 1583-1606.

Deichmann, N. and M. Baer (1990): Earthquake Focal Depths Below the Alps and Northern Alpine Foreland of Switzerland. In: Freeman, R. P. Giese and St. Mueller (Eds.): The European Geotraverse: Integrative Studies. European Science Foundation, Strasbourg, France, 277-288.

Egozcue, J.J., A. Barbat, J.A. Canas, J. Miquel and E. Banda (1991): A Method to Estimate Intensity Occurrence Probabilities in Low Seismic Activity Regions. *Earthquake Engineering and Structural Dynamics*, 20, 43-60.

Egozcue, J.J. and E. Rüttener (1996): Bayesian Techniques for Seismic Hazard Assessment Based on Imprecise Data. in prep.

Mayer-Rosa, D. and H. Merz (1976): Seismic Risk Maps of Switzerland. Description of the Probabilistic Method and Discussion of the Input Parameters. In: A.R. Ritsema (Ed.): Seismic Risk Assessment for Nuclear Power Plants, Proc. of ESC Symposium, Luxembourg. KNMI De Bilt, The Netherlands, 45-52.

McGuire, R. K. (1976): FORTRAN Computer Program for Seismic Risk Analysis. *U.S. Geological Survey, Open-File Report* 76-67, 90 pp.

Mortgat, C.P. and H.C. Shah (1979): A Bayesian Model for Seismic Hazard Mapping. *Bull. Seism. Soc. Am.*, 69, 1237-1251.

Mueller, St. and D. Mayer-Rosa (1980): The New Seismic Hazard Maps for Switzerland. *Rev. Geofisica*, 13, 7-19.

Press, S.J. (1989): Bayesian Statistics: Principles, Models, and Applications. John Wiley & Sons, New York, 237 pp.

Rüttener E.(1995): Earthquake Hazard Evaluation for Switzerland. Ph. D. Thesis, ETH Zurich.