Study 1

A New Statistical Model for Vrancea Earthquakes using Prior Information from before 1900

Pieter van Gelder, Delft University of Technology
Dan Lungu,
Technical University of Civil Engineering, Bucharest
Michael Thomas, University of Siegen.

Abstract. In this paper we present a new statistical model for the Gutenberg–Richter–magnitude (GR–magnitude) of earthquakes occurring from the Vrancea source in Romania. Mean return periods of Vrancea earthquakes are up to now calculated with statistical models of earthquake data of the last century. However, a very extensive historical research has been undertaken by Radu et al. to investigate the occurrences of earthquakes in Romania in the period 984–1900. This information has not been included in the statistical models for GR–magnitudes versus occurrence frequency so far.

Below new calculations will be presented that help to include this historical data in the current models by using a Bayesian modeling approach. This approach has proven to be successful in a similar problem regarding historical sea floods in the low lying country of the Netherlands (Van Gelder, 1996).

1. Introduction

Radu and Lungu (last version 1995) and Constantinescu and Marza (last version 1995) provided instrumental and historical catalogues of earthquake data occurring in the subcrustal (60–170km) Vrancea zone of the Carpathians in Romania over the 20th century and over the millennium 984-1995. The following subcatalogues have been filtered from the original Radu catalogue (Lungu et al., 1997).

Data set 1: GR–magnitudes above 4.1 in the period of 984–1900 (aftershocks are deleted) with 181 data points.

Data set 2: GR–magnitudes above 5.7 in the period of 1901–1995 (aftershocks are deleted) with 14 data points.

The use of the moment magnitude Mw is recommended as a systematic requirement for the seismic hazard assessment. However, since the available catalogues of Vrancea events were prepared by using the GR–magnitude M, the recurrence–magnitude relationships were determined by using that magnitude. The recurrence magnitude relationship for Vrancea source is determined by Radu's 20th century catalogue of intermediate depth magnitudes.

Lungu et al. (1995) investigated the exceedances above a threshold magnitude $v_0 = 6.0$ and found out that the average number per year of Vrancea earthquakes with magnitude equal to and greater than v is

$$\log m(v) = 3.49 - 0.72v.$$

If the magnitude is limited by an upper bound v_{max}, the recurrence relationship can be modified in order to satisfy the property of a probability distribution and, in the case of Vrancea source, Elnashai and Lungu (1995) used

$$m(v) = \exp(8.036 - 1.658v)\frac{1 - \exp(-1.658(v_{max} - v))}{1 - \exp(-1.658(v_{max} - v_0))}.$$

In the above equation the threshold magnitude is selected $v_0 = 6.0$, the maximum credible magnitude of the source is estimated $v_{max} = 7.8$. The estimated mean return GR–magnitudes are 7.65 (475–year) and 7.73 (1000-year)[1].

Another mathematical way to include a maximum credible magnitude is to use a generalized Pareto distribution rather than the above given exponential distributions. The two parameters of the generalized Pareto distribution (shape γ and scale σ) can be estimated by classical techniques. However, planning to incorporate historical data into the statistical models, we will also perform the estimation with Bayesian techniques.

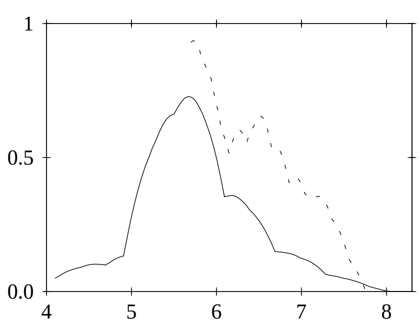

FIG. 1. Kernel density plots of data set 1 (solid) and data set 2 (dashed).

Fig. 1 shows a kernel density plot of both data sets. One can recognize that the mode of data set 1 is at about $v = 5.6$. Although one would not expect such

[1]In terms of probability, 475 years correspond to 10% probability of exceedance in 50 years, and this probabilistic safety level is recommended by actual regulations from USA and Europe (ASCE 7-98 & 2000 and Eurocode 8, 1994).

a behaviour from GR–magnitudes of earthquakes, it can be explained with the fact that lighter earthquakes could usually not be reconstructed from historical documents. We therefore select a threshold of $v_0 = 5.7$ for the analysis of both data sets, leaving 80 observations for the years 984–1900.

2. Analysis Using Classical Techniques

A natural model for our data set are Poisson (λ, F) processes where λ is the intensity of a homogenous Poisson process controlling the arrival times of earthquakes, and F is a generalized Pareto distribution function describing the GR–magnitude of the earthquakes exceeding the threshold v_0. The reader is referred to section **??** for an introduction to these processes.

The intensity measure ν of a Poisson (λ, F) process is given by $\nu([0,t] \times [0,y]) = \lambda t F(y)$. The mean number of earthquakes exceeding a threshold $v > v_0$ within one year is therefore $m(v) := \nu([0,1] \times (v, \infty)) = \lambda(1 - F(v))$.

In the following sections, we employ different estimators for the generalized Pareto distribution $F = W_{\gamma, v_0, \sigma}$. The intensity λ is estimated by the number of earthquakes exceeding the GR–magnitude v_0 per year.

We start by analyzing data set 2. The accuracy of this data set is quite high as it consists partly (after 1934) of instrumental data. The number of data points is however quite low; 14 in a time span of 95 years, yielding the intensity $\lambda_2 = 0.147$.

In Fig. 2, an exponential and a generalized Pareto distribution are fitted to data set 2 employing the m.l. estimator for the exponential and the moment estimator for the GP distribution. The mean number $m(v)$ of exceedances above the threshold v is shown for both parametric distributions and the empirical distribution of the data.

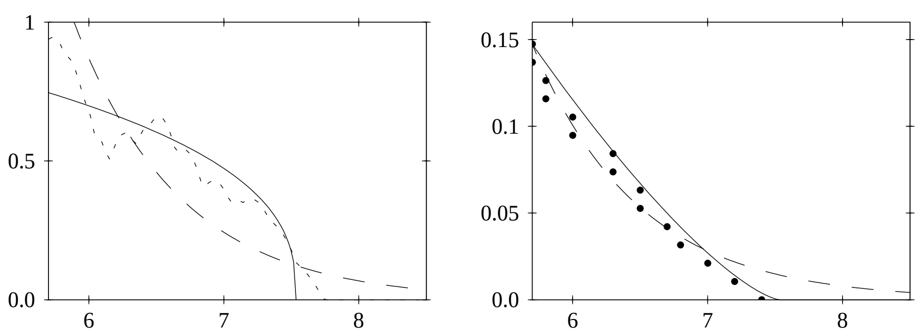

FIG. 2. (Left.) Kernel density of data set 2 (dotted), fitted GP distribution $W_{-0.73,5.7,1.34}$ (solid) and fitted exponential distribution $W_{0,5.7,0.79}$. (Right.) Mean number of exceedances $m(v)$ for the GP distribution (solid), exponential distribution (dashed) and empiric distribution of data (points).

We estimate a realistic 475–year return GR–magnitude of 7.45 when applying a GP distribution and we get an unrealistic GR–magnitude of 9.06 when applying an exponential distribution. The 1000–year return magnitudes are 7.49 and 9.64, respectively.

Moreover, the right endpoint of the GP distribution equals approximately 7.6 which might be understood as a maximum credible source magnitude. It is obvious that the exponential distribution overestimates the tail probabilities.

3. A Bayesian Approach

A Bayesian approach (see, e.g., Carlin and Louis , 1996) can be applied to data set 2 nicely by obtaining a prior distribution of the GR–magnitudes from data set 1. We employ the likelihood function

$$L(\gamma, \sigma | \boldsymbol{x}) = \prod_{i=1}^{n} w_{\gamma, v_o, \sigma}(x_i)$$

and a prior $p(\gamma, \sigma)$ that is specified later. The Bayes estimators of γ and σ are then given by $\hat{\gamma}(\boldsymbol{x}) = \int \gamma p(\gamma, \sigma | \boldsymbol{x}) \, d\gamma \, d\sigma$ and $\hat{\sigma}(\boldsymbol{x}) = \int \sigma p(\gamma, \sigma | \boldsymbol{x}) \, d\gamma \, d\sigma$ where

$$p(\gamma, \sigma | \boldsymbol{x}) = \frac{L(\gamma, \sigma | \boldsymbol{x}) p(\gamma, \sigma)}{\int L(\gamma, \sigma | \boldsymbol{x}) p(\gamma, \sigma) \, d\gamma \, d\sigma}$$

is the posterior distribution.

The following steps are performed to include the historical data into our analysis.

1. Apply the m.l. estimator for γ and σ in the GP model to data set 1 and determine a normal approximation $p(\gamma, \sigma)$ to its distribution. Because the m.l. estimator maximizes the posterior distribution when a non–informative prior is applied, this step can also be interpreted as a Bayesian analysis.

2. Use the bivariate normal distribution p obtained in the first step as prior distribution for the Bayesian estimation of γ and σ.

The m.l. estimator fits the generalized Pareto distribution $W_{-0.442,5.7,0.903}$ to data set 1. The normal approximation to the distribution of the estimator is performed by means of a simulation under the estimated parameters.

One obtains a bivariate normal with mean vector $\boldsymbol{\mu}^t = (-0.442, 0.903)$ and covariances

$$\Sigma = \begin{pmatrix} 0.0118 & -0.0135 \\ -0.0135 & 0.0187 \end{pmatrix}.$$

Using that normal as a prior in the second step, the Bayes estimator yields $\gamma = -0.472$ and $\sigma = 0.955$. The result is shown by the dotted lines in Fig. 3.

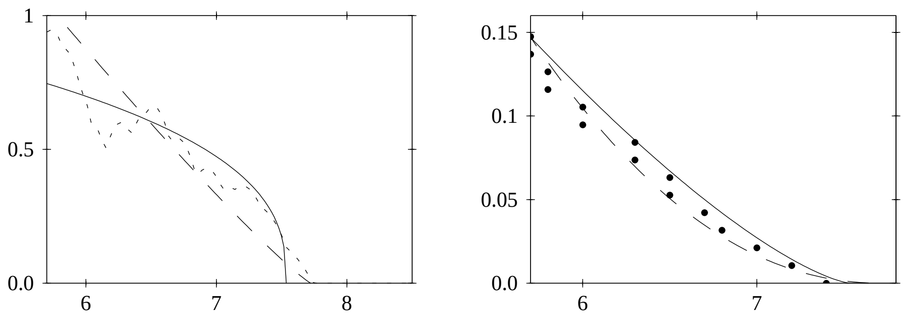

FIG. 3. (Left.) Kernel density of data set 2 (dotted), moment estimate $W_{-0.73,5.7,1.34}$ (solid) and Bayes estimate $W_{-0.472,5.7,0.955}$ (dashed). (Right.) Mean number of exceedances $m(v)$ using moment estimate (solid), Bayes estimate (dashed) and empiric distribution of data (points).

The estimated 475– and 1000–year return GR–magnitudes are 7.45 and 7.53. One can see that a heavier tail is estimated by including the historical information into the model.

To take account of the inaccuracy of data set 1, a different weighting procedure was applied. This can be done within the Bayesian framework by making the prior distributions more vague. We therefore employ a normal prior with covariance matrices

$$\Sigma_1 = \begin{pmatrix} 0.04 & -0.05 \\ -0.05 & 0.07 \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} 0.16 & -0.2 \\ -0.2 & 0.28 \end{pmatrix},$$

yielding estimates $\gamma = -0.521, \sigma = 1.04$ for Σ_1 and $\gamma = -0.561, \sigma = 1.145$ for Σ_2. The pertaining return levels are shown in Table 1.

4. Conclusion

In this paper new models have been presented in the analysis of a frequency distribution for the occurrence of earthquakes in the Romanian subcrustal (60–170 km) Vrancea source.

Instead of a truncated exponential distribution, a generalized Pareto distribution has been suggested, which enables us to model a maximum credible magnitude of the source more easily. A second new element in the paper is the combination of the quasi–instrumental catalogue of the twentieth century with the large historical catalogue of earthquakes in the Vrancea over a millenium. To combine these two very different catalogues, a Bayesian framework was developed.

Table 1 summarizes our results.

TABLE 1. Estimated return GR–magnitudes.

Estimation method	475–year	1000–year
Lungu et al. (1995)	7.65	7.73
Moment(GP)	7.45	7.49
Bayesian (Σ)	7.45	7.53
Bayesian (Σ_1)	7.48	7.55
Bayesian (Σ_2)	7.55	7.62

Combining the historical data with the instrumental data causes a slight increase of the return magnitudes. It is still recommended to analyze the sensitivity of the threshold parameters (now taken as 5.7) in both datasets and to include a maximum credible magnitude v_1 as a boundary condition in the Bayesian optimization procedure. We also did not investigate the intensity parameter λ further, where we obtain an intensity $\lambda_1 = 0.087$ for data set 1 and $\lambda_2 = 0.147$ for data set 2. It would be interesting to extend the Bayesian approach to it.

REFERENCES

Carlin, B.P. and Louis, T.A. (1996). Bayes and Empirical Bayes Methods for Data Analysis, Chapman & Hall.

Constantinescu, L. and Marza, V.I. (1980). A computer–compiled and computer–oriented catalogue of Romanias earthquakes during a millennium. Revue Roumaine de Geologie, Geophysique et Geographie. Tome 24, No. 2, Editura Academiei R.S.Romania, Bucharest, 193–206.

Elnashai A. and Lungu D. (1995). Zonation as a tool for retrofit and design of new facilities. Report of the Session A1.2 – 5th International Conference on Seismic Zonation, Nice, France, Oct. 16–19. Proceedings Vol. 3. Quest Editions, Presses Academiques, Nantes, p. 2057–2082.

Marza, V.I., Kijko, A., and Mäntyniemi, P. (1991). Estimate of earthquake hazard in the Vrancea (Romania) region. Pageoph 136, Birkhäuser, Basel, 143–154.

Lungu, D., Cornea, T., and Coman, O. (1995). Probabilistic hazard analysis to the Vrancea earthquakes in Romania. Part I in “Experience database of Romanian facilities subjected to the last 3 Vrancea earthquakes”. Research report to the International Atomic Energy Agency, Vienna. Contract No. 8223/EN. Stevenson & Assoc. Office in Bucharest.

Lungu D., Cornea T., Aldea A. and Zaicenco A. (1997). Basic representation of seismic action. In: Design of structures in seismic zones: Eurocode 8—Worked examples. TEMPUS PHARE CM Project 01198: Implementing of structural Eurocodes in Romanian civil engineering standards. Edited by D. Lungu, F. Mazzolani and S. Savidis, Bridgeman Ltd., Timisoara, p. 1–60.

Radu, C. (1970, 1986, 1995). Catalogue of strong earthquakes ($I_0 \geq 6.0$) occurred in Romania in the period 984–1995, (from Radu's manuscripts by Lungu, D., 1997).

Van Gelder, P.H.A.J.M. (1996). A new statistical model for extreme water levels along the Dutch coast. IAHR Proceedings on Stochastic Hydraulics, Vol. 7, pp. 243–251, Mackay, Australia.