STATISTICAL METHODS TO ANALYZE SEISMICITY INTERACTION IN NEIGHBOUR ZONES; APPLICATION TO SIMULATED DATA

Dragomir Gospodinov[1], Boyko Ranguelov[1], Eleftheria Papadimitriou[2], Vassilis Karakostas [2]

[1] Geophysical Institute of the Bulgarian Academy of Sciences, Akad. G. Bonchev str., bl.3, 1113 Sofia, Bulgaria; e-mail: drago_pld@yahoo.com , boyko.ranguelov@geophys.bas.bg
[2]Geophysics Department, University of Thessaloniki, GR54124 Thessaloniki, Greece; e-mail: vkarak@geo.auth.gr , ritsa@geo.auth.gr

Key words: stochastic modeling, seismicity interaction, simulated data

Introduction

While the examination of seismicity patterns in separate zones has been the long aimed task of many studies, interaction between seismic activities in neighbor zones is a more difficult purpose as it needs a more profound physical hypothesis underlying the possible realization of such a process. The spatial aspect of this problem is being approached in recent years with the help of the Coulomb stress technique (CST), which depicts the areas of increased static stress after a strong earthquake. They could reveal the locations of the aftershock activity or the epicentral zone of a future strong event. The temporal features of a possible relation between seismic activities of two zones can be studied by modeling of these phenomena by stochastic models which incorporate eventual interaction (Ogata et al., 1982). This is most easy to be done for stationary processes and it is reasonable to exclude known deviations from stationarity (clusters, trends etc.) for any zone. The purpose of our study will be to verify a hypothesis of statistical independence between the seismic activities in two neighbor zones by considering them as stochastic point processes. The interpretation of such results can be supported by using available data on the seismotectonic features of the zones – tectonic structures, active and passive faulting, fragmentation, creep, etc. In this paper we shall analyze simulated data, which has been generated to follow the stationary Poisson process in time.

Methodology

The interaction between seismic activities of neighbor zones can be analyzed by examining the cross-correlation of some continuous variables like seismic energy release in the zones. In this paper, however, we apply a different approach using stochastic point processes to model the realization of earthquakes in the zones. Utsu & Ogata (IASPEI, 2000) have developed and offered a self- and externally excited point process model to consider possible dependence between the occurrences of events in the neighbor zones. The corresponding software program is called LINLIN and is included in the IASPEI Software Library Volume 6. This program computes the maximum likelihood estimates (MLEs) of the parameters of a point process model for seismic activities (Ogata et al., 1982) from a set of data on the occurrence times of earthquakes in a region. This region is called region A, and the process representing the earthquakes occurring in region A is called ‘master process’. The seismic activity in region A may be affected by the earthquakes occurring in another region (called region B and ‘external process’).

The data analyzed are the occurrence times t_i $(i = 1,2,...)$ of earthquakes in region A during a certain time interval. Optional data are the occurrence times u_j $(j = 1,2,...)$ of earthquakes in region B in the same time interval. The rate of earthquake occurrence in region (intensity function) is represented by

$$f(t) = P(t) + Q(t) + \sum g(t - t_i) + \sum h(t - u_j) \ . \quad (1)$$

In this equation, $P(t)$ is a polynomial representing a trend in activity in region, A, $Q(t)$ is a Fourier series $(K = 2,4,...)$ representing a periodic variation in activity in region A with a fundamental period P, $g(s)$ is a Laguerre type polynomial representing the effect of an earthquake at time t in region A on the seismic activity in the same region at time $t + s$ (self-exciting effect) and $h(s)$ has a similar form as $g(s)$ but it represents the effect of an earthquake at time t in region B on the activity at time $t + s$ in region

A (external effect). The two summations $\sum$ in equation (1) extend for all i satisfying $t_i < t$ and all j respectively $t_j < t$. The external process is not necessarily a series of earthquakes in region B. Other point events (e.g., volcanic eruptions in region A) may be considered as an external process.

Another statistical technique to verify the independent realization of earthquakes in neighbor zones was developed by Gospodinov (1998). By the probability theory this problem could be considered as follows:

Let's have two sequences of events $z_k^A, k = 1,2,...,N_A$ and $z_l^B, l = 1,2,...,N_B$ which occurred in two neighbor zones A and B in the same time period $T_1 \div T_2$. We form a sequence $z_m, m = 1,2,...,N$ with $N = N_A + N_B$ arranging the events chronologically. The aim is to verify whether a realization of an event in one of the zones is independent of the occurrences of events in the other zone. We can extract information about that by the distribution of the random variable $z_{m+1} \in i | z_m \in j; \quad i,j = A,B$ denoting by $z_{m+1} \in i$ an event with a sequential number $m+1$, which occurred in zone i while the previous one occurred in zone j. This is a discrete random variable with four possible states. Its distribution shall be determined following the assumption that the zones are independent that is an event realization in one of the zones does not change the probability for event occurring in the other zone. In the probability theory this problem is identical to the one of finding the distribution of all pairs of balls from a box with N_A white and N_B black balls. The probability of a certain combination is determined by the ratio of possible states for this combination $N_{ij} = N_i N_j$ versus the total number of possible states N^2 (Rozanov, 1989)

$$P(ij) = \frac{N_{ij}}{N^2} = \frac{N_i N_j}{N^2} \,. \tag{2}$$

Formula (2) gives the distribution function of the variable we are interested in. Applying it we may calculate the expected number for each possible combination (assuming independent zones)

$$Y_{ij}^o = P(ij)N = \frac{N_i N_j}{N} \qquad for \quad i,j = A,B. \tag{3}$$

Then we can form the statistics used by Paradyne and Rivette (1967)

$$\aleph_E^2 = \sum \sum \frac{(Y_{ij} - Y_{ij}^o)^2}{Y_{ij}^o} \qquad for \quad i,j = A,B \,, \tag{4}$$

where Y_{ij} denotes the theoretical number of cases $z_{m+1} \in i | z_m \in j; \quad i,j = A,B$. The statistics defined by formula (4) follows the $\aleph^2$-distribution and it is used to verify the hypothesis of independence of the seismic activities in neighbor zones (no interaction); the hypothesis is rejected for $\aleph_E^2 > \aleph_{\nu,\mu}^2$ where μ is the chosen significance level and ν is the degrees of freedom number ($\nu = 1$ for two zones).

Data and results

In this paper we tested the above methods on two simulated data samples. Both samples were generated following the stationary Poisson process with constant intensities; $\lambda_1 = .25$ for the master process with $N_A = 898$ events and $\lambda_2 = .08$ for the external process with $N_B = 300$ events. The simulation was done using the conditional intensity function (Ogata, 1999). Let $F(t \,|\, t_1,...,t_n)$ is the conditional distribution of an event occurring at time t, given a history of events which occurred at times $t_1, t_2, ..., t_n$. The time t_{n+1} of the next event can be simulated by solving the equation

$$F(t \,|\, t_1,...,t_n) = U_{n+1} \,, \tag{5}$$

where U_{n+1} is a uniform random variable. As in our case the conditional intensity is constant then the distribution function of the intervals is exponential and times are simulated by the formula

$$t_i = t_{i-1} - \log(U_i)/\lambda. \tag{6}$$

We first applied the LINLIN software to analyze possible interaction between both samples. We verified different cases including (or excluding) self-excitement, external effect (interaction), periodicity and trend. As the samples were generated to follow the stationary Poisson process it is expected that data would best support the model with no interaction (external effect), no self-excitement, periodicity and no trend. The variations of the model were compared using the Akaike information criterion (AIC), (Ogata, 1999)

$$AIC = -2 \max \log L + 2k \,, \tag{7}$$

Table I

Self-excitement	External effect	Periodicity	Trend	AIC
Yes	No	No	No	4241.738
No	No	No	No	4238.417
Yes	Yes	No	No	4243.630
No	Yes	No	No	4241.598
Yes	Yes	Yes	Yes	4242.265

Table II

Zones A-B	A		B		$\aleph_E^2$	$\aleph_{1;0.05}^2$
	Y_{ij}	Y_{ij}^o	Y_{ij}	Y_{ij}^o		
A	678	672.19	218	224.81	1.14	3.84
B	217	224.81	82	75.19		

where L is the maximum log-likelihood function and k is the number of parameters used in the model. The minimum AIC value identifies the best fit model. The results are presented in *Table I.* It can be seen that the LINLIN software correctly recognizes the best fit model for the simulated data. The model with smallest AIC is the one for which we have no self-excitement, no interaction, no periodicity and trend.

The two simulated samples were also analyzed with the help of the model offered by Gospodinov (1998). They cover one time period and the null hypothesis was that the two samples are independent that is a realization of an event in one of the zones does not change the probability for an event occurring in the other zone. Following this assumption and formula (3) we calculated the expected Y_{ij}^o values and compared them to the real Y_{ij} values. The results are shown in *Table II.* As one can see the calculated value $\aleph_E^2 < \aleph_{1;0.05}^2$ of the used statistic is smaller than the theoretical one which reveals that the data does not reject the null hypothesis at a significance level of $\mu = 0.05$. In fact both methods applied identify correctly the independence of the two simulated sequences and can be used to analyze possible interaction between real sequences of earthquakes in adjacent zones. One should use all existing on the geotectonic characteristics of the seismic zones (tectonic structure, faulting, faults interaction etc.) to make a most adequate interpretation of the results by such a study.

Acknowledgments: Supported by MON contract BG-9/05

References

Gospodinov, D. 1998. Statistical test for independence. Application to the analysis of data from adjacent seismic zones in the Upper-Thracian Valley. – *Geodynamical Investigations, Connected with 1928 Year Earthquakes in Chirpan-Plovdiv, Review Book of Papers from the Symposium,* 77-84 (in Bulgarian).

IASPEI 2000. *Software Library*, vol. 6.

Ogata, Y., H. Akaike, K. Katsura 1982. The application of linear intensity models to the investigation of causal relations between a point process and another stochastic process. – *Ann. Inst. Statist. Math.*, *34*, 373-387.

Ogata, Y. (1999). Seismicity analyses through point-process modelling: A review. – In: Wyss, M., K. Shimazaki, A. Ito (eds.), *Seismicity Patterns, Their Statistical Significance and Physical Meaning,* (*Pure and Applied Geophysics, 155.* Basel, Birkhauser Verlag, 471-507

Paradyne, K., B. Rivette 1967. *Statistical Methods in Tchniques*. Sofia, Tehnika Publ. House.

Rozanov, Y. (1989). *Probability Theory, Stochastic Processes and Mathematical Statistics*. Nauka, Moscow, 312 p.

СТАТИСТИЧЕСКИ АНАЛИЗ НА ВЗАИМОДЕЙСТВИЕТО МЕЖДУ СЕИЗМИЧНИТЕ АКТИВНОСТИ В СЪСЕДНИ ЗОНИ: ПРИЛОЖЕНИЕ ЗА СИМУЛИРАНИ КАТАЛОЗИ

Драгомир Господинов[1], Бойко Рангелов[1], Елефтерия Пападимитриу[2], Василис Каракостас[2]

Много изследвания показват, че при изследване на сеизмичността в определена сеизмична зона трябва да се има предвид и възможното взаимодействие между процесите в съседни зони. В пространствен аспект това взаимодействие се изучава с помощта на така наречената техника на Кулон за определяне на промяната в полето на напрежение след реализацията на силно земетресение. Тази техника очертава зоната на афтершокова активност или епицентралната област на бъдещо силно събитие. Характеристиките на евентуално взаимодействие във времето могат да се анализират чрез моделирането му с помощта на случайни процеси. От математическа гледна точка това е най-лесно осъществимо за стационарни процеси и затова е целесъобразно предварително да се отстранят известните отклонения от стационарност (групиране, наличие на тренд и др.). Целта на настоящата работа е да се провери хипотезата за статистическа независимост между сеизмичните активности в съседни зони чрез разглеждането им като случайни процеси. Използвани са два подхода, които са приложени върху две случайно генерирани последователности от събития, следващи прост Поасонов процес във времето. Интерпретацията на резултатите за реални зони трябва да се прави с помоща на наличната геотектонска информация за зоните – структура, разломяване, крип и др.

При единия подход Utsu & Ogata (IASPEI, 2000) предлагат самовъзбуждащ се процес с опции за включване на външен ефект (взаимодействие с друга зона). Разработена е и компютърна програма LINLIN за оценка на моделните параметри по метода на максималното правдоподобие. Според модела интензивността на сеизмичния процес в дадена зона A се дава с

$$f(t) = P(t) + Q(t) + \sum g(t - t_i) + \sum h(t - u_j) \quad (1)$$

като t_i $(i = 1,2,...)$ са времената на събитията в зона A, u_j $(j = 1,2,...)$ са времената на събитията в зона B, $P(t)$ е полином, отразяващ тренда в активността на зона A, $Q(t)$ представя периодичността, а $g(s)$ - въздействието на предишни земетресения върху активността. Подобна форма има и $h(s)$, като той отразява влиянието на предишни земетресения от зона B (взаимодействие) върху сеизмичността на зона A.

Друга статистическа техника за анализ на независимата реализация на събития в съседни зони е разработена от Господинов (1998). В теорията на вероятностите този проблем може да се третира по следния начин. Нека имаме две последователности от земетресения $Z_k^A, k = 1,2,...,N_A$ и $Z_l^B, k = 1,2,...,N_B$, които са възникнали в зони A и B в един времеви период $T_1 \div T_2$. Образуваме обща извадка $Z_m, m = 1,2,...,N$ с $N = N_A + N_B$, подреждайки събитията хронологично. За да се провери дали има взаимодействие между зоните се използва случайната величина $Z_{m+1} \in i | Z_m \in j; \quad i, j = A, B$. Обозначаваме с $Z_{m+1} \in i$ събитие с номер $m+1$, станало в зона i, докато предишното е станало в зона j. Разпределението на тази дискретна случайна величина при хипотеза за независимост се дава с

$$P(ij) = \frac{N_{ij}}{N^2} = \frac{N_i N_j}{N^2} \quad i, j = A, B \qquad (2)$$

като $P(ij)$ е вероятността за реализация на съответната комбинация. Разгледаните по-горе два метода бяха приложени за анализ на възможното взаимодействие между две случайно генерирани последователности следващи прост Поасонов процес с различна интензивност. И двата метода коректно идентифицираха липса на взаимодействие между последователностите, както може да се очаква от метода на генерирането им. Горната методика може да се приложи и за анализ на реални сеизмични зони, като при интерпретацията трябва да се имат предвид сеизмотектонските свойства на зоните – наличие на активни/пасивни разломи, крип, вертикални и хоризонтални движения и др.