

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Change Point Analysis in Earthquake Data

Ayten Yiğiter

*Hacettepe University, Faculty of Science, Department of Statistics Beytepe-Ankara
Turkey*

1. Introduction

Earthquake forecasts are very important in human life in terms of estimating hazard and managing emergency systems. Defining of earthquake characteristics plays an important role in these forecasts. Of these characteristics, one is the frequency distribution of earthquakes and the other is the magnitude distribution of the earthquakes. Each statistical distribution has many parameters describing the actual distribution.

There are various statistical distributions used to model the earthquake numbers. As is well known, these are binomial, Poisson, geometric and negative binomial distributions. It is generally assumed that earthquake occurrences are well described by the Poisson distribution because of its certain characteristics (for some details, see Kagan, 2010; Leonard & Papasouliotis, 2001). In their study, Rydelek & Sacks (1989) used the Poisson distribution of earthquakes at any magnitude level. The Poisson distribution is generally used for earthquakes of a large magnitude, and the earthquake occurrences with time/space can be modeled with the Poisson process in which, as is known, the Poisson distribution is one that counts the events that have occurred over a certain period of time.

There is a significant amount of research on the change point as applied to earthquake data. Amorese (2007) used a nonparametric method for the detection of change points in frequency magnitude distributions. Yigiter & İnal (2010) used earthquake data for their method developed for the estimator of the change point in Poisson process. Aktaş et al. (2009) investigated a change point in Turkish earthquake data. Rotondi & Garavaglia (2002) applied the hierarchical Bayesian method for the change point in data, taken from the Italian NT4.1.1 catalogue.

Recently, much research in the literature has focused on whether there is an increase in the frequency of earthquake occurrences. It is further suggested that any increase in the frequency of earthquakes, in some aspects, is due to climate change in the world. There is considerable debate on whether climate change really does increase the frequency of natural disasters such as earthquakes and volcano eruptions. In many studies, it is emphasized that there is serious concern about impact of climate change on the frequencies of hazardous events (Peduzzi, 2005; Lindsey, 2007; Mandewille, 2007 etc.). In Peduzzi's study (2005), there are some indicators about increasing number of the earthquakes especially affecting human settlements, and it is also reported that there is an increase in the percentage of earthquakes affecting human settlements from 1980 onwards. The change point analysis can be used to study the increase or decrease in the frequency of the earthquake occurrences.

The change point analysis is a very useful statistical tool to detect an abrupt or a structural change in the characteristic of the data observed sequentially or chronologically. Many different statistical methods are available to detect the change point in the distribution of a sequence of random variables (Smith, 1975; Hinkley, 1970; Boudjellaba et al., 2001 in many others). Multiple change points have also been investigated for many years (Hendersen & Matthews, 1993; Yao, 1993; Chen & Gupta, 1997, among others.).

In this study, we aimed to find an evidence of increased or decreased in world seismic activities after 1901. Earthquake frequencies are modeled by the Poisson distribution as the most commonly used discrete distribution. We modeled the earthquakes in the world with the Poisson process since the number of earthquakes is counted with Poisson distribution. We investigated any abrupt change point(s) in parameters of the model using the frequentist (maximum likelihood) and Bayesian method. When the magnitude of the earthquakes is taken into account in the process, it is then called the compound Poisson process. We also investigated a change in the magnitudes of the world earthquakes. For this purpose, Poisson process and compound Poisson process are introduced in Section 2. The frequentist and Bayesian methods used for change point estimates are explained in Section 3. Worldwide earthquake data and change point analysis of this data are given in Section 4.

2. Poisson process

The Poisson process has an important place in stochastic processes. It is a Markov chain with a continuous parameter.

A stochastic process $\{N_t, t \geq 0\}$ is said to be a homogeneous Poisson process if

- i. $\{N_t, t \geq 0\}$ has stationary independent increments
- ii. for any times s and t such that $s < t$ the number of events occurred in time interval (s, t) has Poisson distribution with parameter $\lambda(t-s)$.

In the homogeneous Poisson Process, events occur independently throughout time. The number of events occurred time interval $(0, t]$ has Poisson distribution with parameter λt ,

$$P(N_t = i) = e^{-\lambda t} \frac{(\lambda t)^i}{i!}, \quad i = 0, 1, \dots \quad (1)$$

$$= 0, \quad \text{otherwise}$$

where λ is called the occurrence rate of the event in unit time. The occurrence rate is assumed to be a constant throughout the process.

Let S_0, S_1, S_2, \dots be occurrence times of the events where $S_0 = 0$ and $T_1 = S_1 - S_0$, $T_2 = S_2 - S_1, \dots$ be time intervals between the events. T_1, T_2, \dots are independent identically exponential random variables with parameter λ . The distribution of the random variables T_1, T_2, \dots is:

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0 \quad (2)$$

$$= 0, \quad t < 0$$

2.1 Compound poisson process

A stochastic process $\{X_t, t \geq 0\}$ is said to be a compound Poisson process if it can be presented, for $t \geq 0$, by

$$X_t = \sum_{i=1}^{N_t} Y_i \quad (3)$$

in which $\{N_t, t \geq 0\}$ is a Poisson process, and $\{Y_i, i = 1, 2, \dots\}$ are independent identically distributed (i.i.d.) random variables. The compound Poisson distribution has stationary independent increments, and the mean and variance of the X_t ,

$$\begin{aligned} E(X_t) &= E(N_t)E(Y) \\ V(X_t) &= E(N_t)E(Y^2). \end{aligned} \quad (4)$$

See Parzen (1962) for some details about Poisson processes.

3. Methods for Estimating Change Point

The Maximum likelihood method and the Bayesian method are basic methods in statistical change point analyses for point estimation and hypothesis testing.

3.1 Change point estimation with the maximum likelihood method (frequentist method)

Let X_1, X_2, \dots, X_n be a sequence of the random variables such that X_i ($i = 1, \dots, v$) has probability density function $f(x, \theta_1)$ and X_i ($i = v + 1, \dots, n$) has probability density function $f(x, \theta_2)$ where the change point v is unknown discrete parameter and the parameters θ_1, θ_2 can be assumed to be either known or unknown. The single change point model in the sequence of the random variables is written:

$$\begin{aligned} X_1, X_2, \dots, X_v &\sim f(x, \theta_1) \\ X_{v+1}, X_{v+2}, \dots, X_n &\sim f(x, \theta_2). \end{aligned} \quad (5)$$

Under model Eq.(5), for the observed values of the sequence of the random variables, the likelihood function is

$$L(\theta_1, \theta_2, v | x_1, \dots, x_n) = L(\theta_1, \theta_2, v) = \prod_{i=1}^v f(x_i, \theta_1) \prod_{i=v+1}^n f(x_i, \theta_2) \quad (6)$$

and the logarithm of the likelihood function is:

$$\ell n L(\theta_1, \theta_2, v) = \sum_{i=1}^v \ell n f(x_i, \theta_1) + \sum_{i=v+1}^n \ell n f(x_i, \theta_2) \quad (7)$$

After adding the expression $\pm \sum_{i=1}^v \ell n f(x_i, \theta_2)$ in Eq.(7) then Eq.(7) can be re-written as follows:

$$\ell n L(\theta_1, \theta_2, v) \propto \sum_{i=1}^v [\ell n f(x_i, \theta_1) - \ell n f(x_i, \theta_2)] \quad (8)$$

When the parameters, θ_1, θ_2 are unknown, for any fixed values of change point v , the maximum likelihood estimator of the parameters θ_1, θ_2 are found to be the derivative of Eq.(8) θ_1 and θ_2 respectively:

$$\frac{\partial \ln L(\theta_1, \theta_2; v)}{\partial \theta_1} = 0, \quad (9)$$

$$\frac{\partial \ln L(\theta_1, \theta_2; v)}{\partial \theta_2} = 0. \quad (10)$$

After solving the equations system given in Eq.(9) and (10), the maximum likelihood estimator of the parameters θ_1 , θ_2 , $\hat{\theta}_1$, $\hat{\theta}_2$ are obtained. The maximum likelihood estimate of the change point v is:

$$\hat{v} = \arg \max_{k=1, \dots, n-1} \sum_{i=1}^k [\ln f(x_i, \hat{\theta}_1) - \ln f(x_i, \hat{\theta}_2)] \quad (11)$$

3.2 Change point estimation with the Bayesian method

The Bayesian method differs from the frequentist method in that each parameter is assumed to be a random variable and each one has a probability function called prior distribution. The estimate of the unknown parameter is obtained by deriving a posterior distribution on the basis of the prior distributions and the likelihood function. The posterior distribution is obtained:

$$\text{Posterior} \propto \text{likelihood} \times \text{prior}.$$

Under change point model given in Eq.(5), let $p_0(\theta_1, \theta_2 | v)$ be the joint prior distribution of parameters θ_1 , θ_2 and let $p_0(v)$ be the prior distribution of change point v . The likelihood function is given by Eq.(6) and so the joint posterior distribution is written as follows:

$$p_1(\theta_1, \theta_2, v) \propto L(\theta_1, \theta_2, v) p_0(\theta_1, \theta_2 | v) p_0(v). \quad (12)$$

Integrate Eq.(12) with respect to the parameters, the marginal posterior distribution of change point is proportional to

$$p_1(v) \propto \int \int_{\theta_1, \theta_2} L(\theta_1, \theta_2, v) p_0(\theta_1, \theta_2 | v) p_0(v) d\theta_1 d\theta_2 \quad (13)$$

and the Bayesian estimate \hat{v} of the change point v is found by maximizing the marginal posterior distribution given by Eq.(13). Assuming uniform priors, the joint posterior mode, which gives the maximum likelihood estimates, is at \hat{v} , $\hat{\theta}_1$, $\hat{\theta}_2$ (Smith, 1975).

4. Detecting change point in worldwide earthquake data and findings

This study investigates whether there is a change point in worldwide earthquake activities such as the number of earthquake occurrences and their magnitude. At first, assuming that the occurrences of the earthquakes follow the homogeneous Poisson process, a change point is investigated in the occurrence rate of the earthquake in unit time; secondly, the magnitudes of the earthquake are assumed to be normally distributed random variables; a change point is investigated in the mean of the normally distributed random variables describing the earthquake magnitudes.

To detect a change point in earthquakes, the relevant data is taken from the website of the U.S. Geological Survey (2011), consisting of 819 earthquakes worldwide of magnitude 4.0 or above covering the period from 3-March-1901 until 11-March-2011. It is assumed that the earthquake occurrences follow the homogeneous Poisson process. The earthquakes are observed at times S_1, S_2, \dots, S_{818} in the continuous time interval $(0, T]$. For the earthquakes data, the starting point S_0 is in 3-March-1901 and the end point $S_{818}=T$ is in 11-March-2011. T_1, T_2, \dots, T_{818} are exponentially distributed random variables as given in Eq.(2). Under the assumption that τ is equal to the occurrence time of an event, the likelihood function can be written under the change point model in the sequence of exponentially distributed random variables:

$$L(\lambda_1, \lambda_2, \tau) = \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n-N_\tau} e^{-\lambda_2 (T-\tau)}. \quad (14)$$

Where τ shows change point and is a continuous parameter defined time interval $(0, T]$, λ_2 is the occurrence rate in the time interval $(\tau, T]$ and N_τ is the number of events that occurred in the time interval $(0, \tau]$ and T is the sum of the time interval between the earthquake occurrences, that is $T = \sum_{i=1}^{818} t_i$ (Raftery & Akman, 1986; Akman & Raftery, 1986).

The logarithm of the likelihood function is

$$\ell n L(\lambda_1, \lambda_2, \tau) = N_\tau \ell n(\lambda_1 / \lambda_2) - \tau(\lambda_1 - \lambda_2) + n \ell n \lambda_2 - \lambda_2 T \quad (15)$$

After the derivation of the log likelihood function given by Eq.(15) for λ_1 and λ_2 , the maximum likelihood estimations of λ_1 and λ_2 are obtained respectively:

$$\hat{\lambda}_1 = \frac{N_\tau}{\tau}, \quad \hat{\lambda}_2 = \frac{n - N_\tau}{T - \tau} \quad (16)$$

For the possible values of τ , s_1, s_2, \dots, s_n , and the variable N_τ takes the values $1, 2, \dots, n$, and so the estimations $(\hat{\lambda}_{11}, \hat{\lambda}_{21}), (\hat{\lambda}_{12}, \hat{\lambda}_{22}), \dots, (\hat{\lambda}_{1n}, \hat{\lambda}_{2n})$ are calculated from Eq.(16) and then these estimations are substituted into Eq.(15). One of $(\hat{\lambda}_{1i}, \hat{\lambda}_{2i}, s_i)$, $i = 1, 2, \dots, n$, gives the maximum of this function, say $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\tau}$. Hence $\hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\tau}$ are the maximum likelihood estimation for λ_1, λ_2 and τ .

The scatter plot of the time intervals between the earthquakes is shown in Figure 1. From Figure 1, it is clearly seen that there is a change (or at least one change) in the occurrence rate of the earthquakes.

When we assume a single change point in the data and the method explained above is applied this data and the results are given in Table 1.

As is shown in Table 1, the occurrence rate of the earthquake in unit time (in day or in year) increased considerably after February 3, 2002, approximately ten times of the occurrence rate of the earthquakes in unit time before February 3, 2002.

From the log likelihood function of the change point in Figure 2, it can be seen that there are at least two change points in the data. There many methods are used to detect multiple change points in the data. One of the basic methods used for this purpose is binary

segmentation procedure (Chen & Gupta, 1997; Yang & Kuo, 2001). In the binary segmentation procedure, the data is divided into two homogeneous groups according to the estimated change point, and a change point is searched in each subdivided data until there is no change in the subdivided data. The results are given in Figure 3.

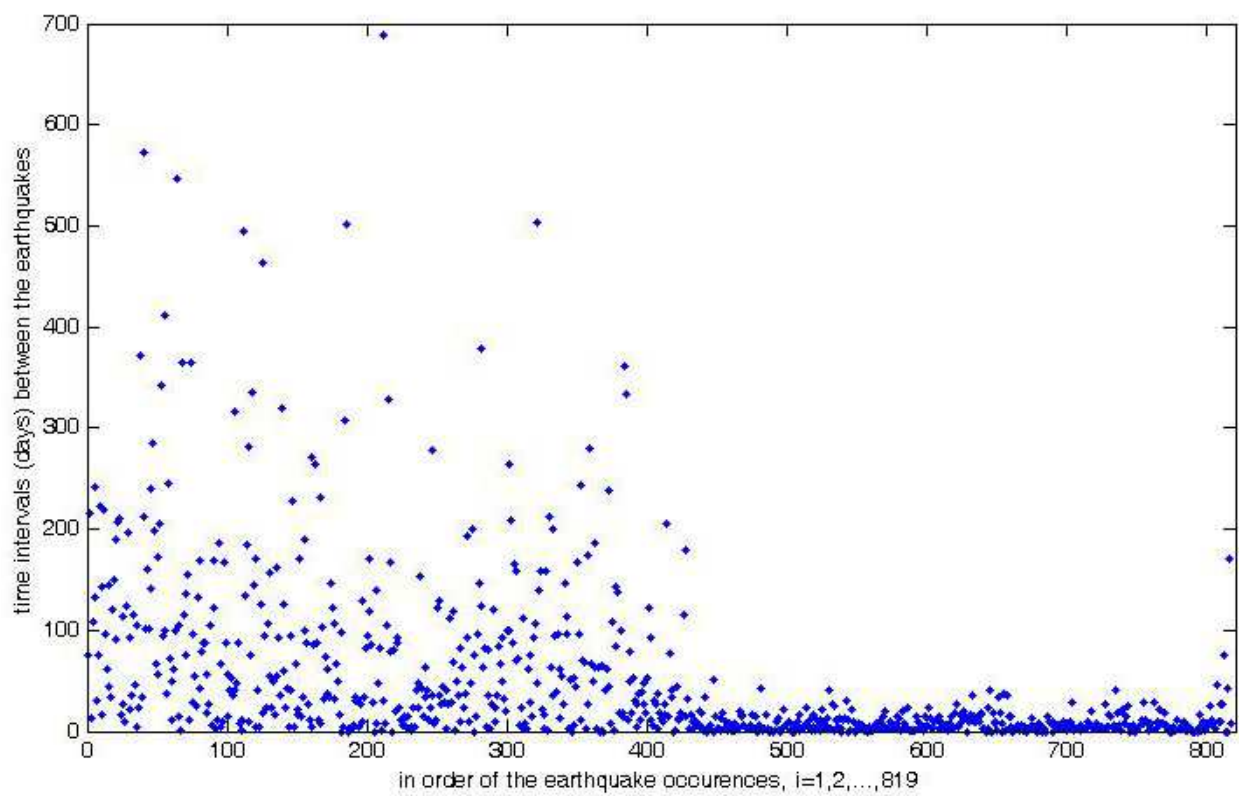


Fig. 1. The time intervals (days) between the earthquakes according to the earthquake occurrences.

n = 818, T= 40185 days		
Before the change point $\hat{\lambda}_1$	Estimated change point $\hat{\tau}$ $N_{\hat{\tau}}=429$	After the change point $\hat{\lambda}_2$
0.0116380 (in days) 4.1896804 (in years)	03-Feb-2002	0.1170628 (in days) 42.1426080 (in years)

Table 1. Estimated parameters for the earthquake data.

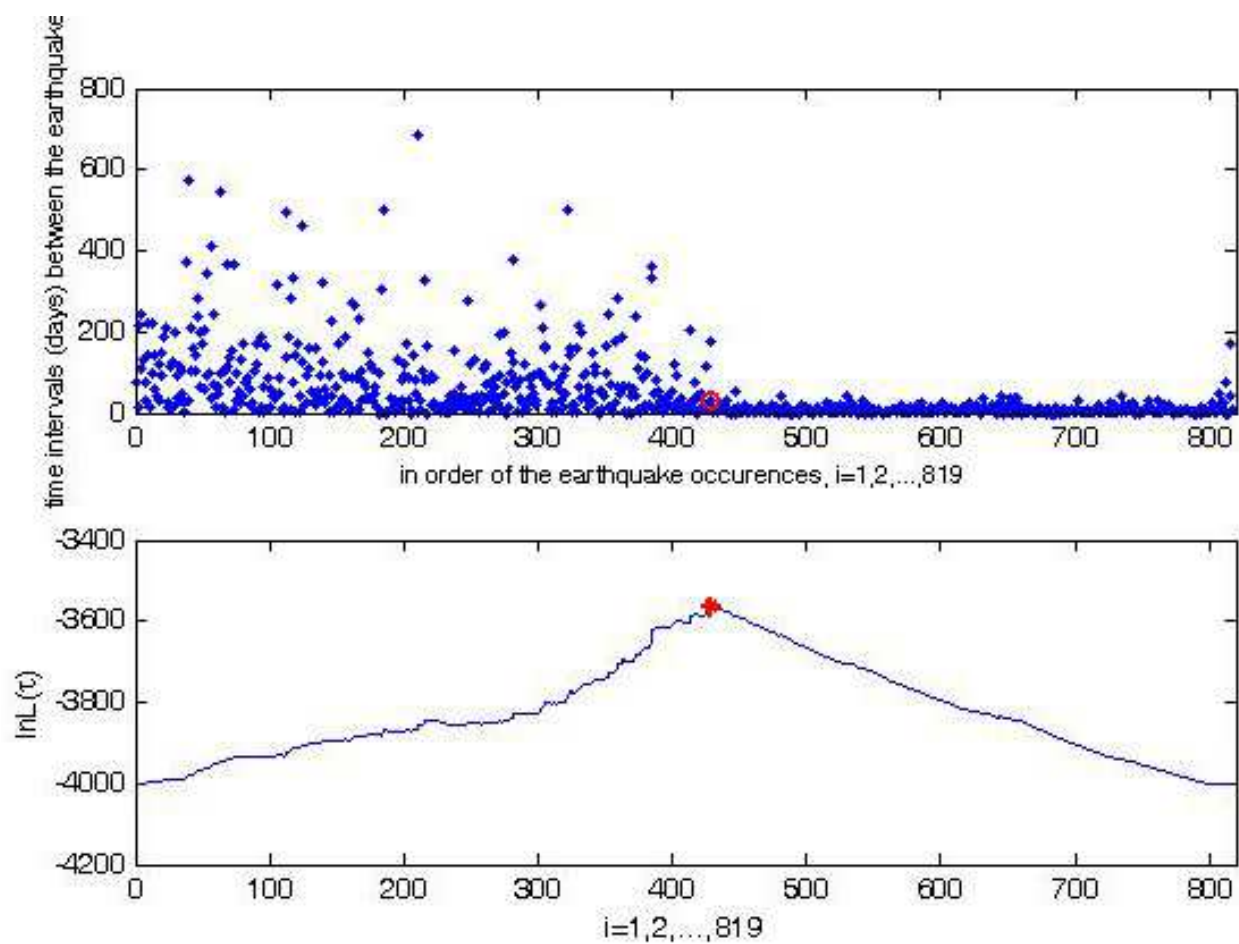


Fig. 2. The maximum point of log likelihood function given in Eq.(16) is shown as a red star and the corresponding change point estimate is shown with a red circle.

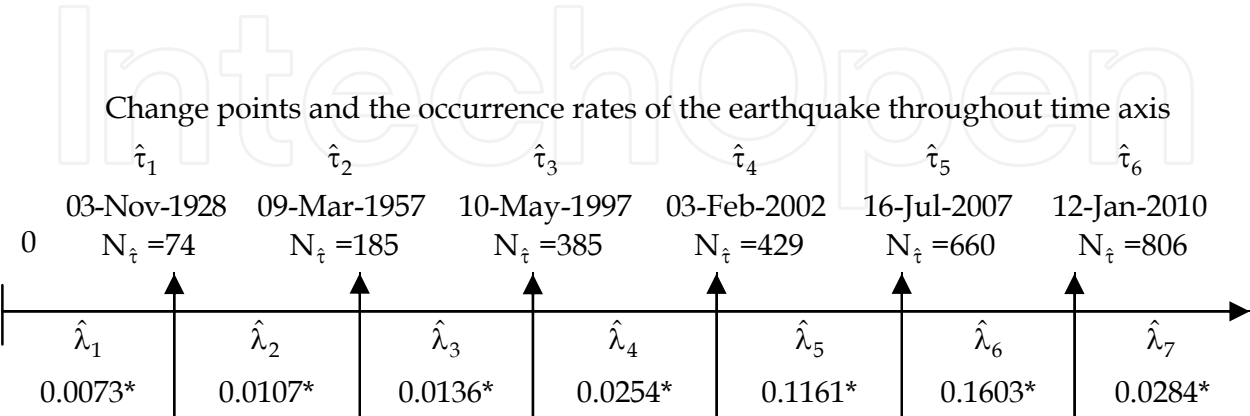


Fig. 3. The maximum likelihood estimations of parameters for multiple change points in the earthquake data (*in days).

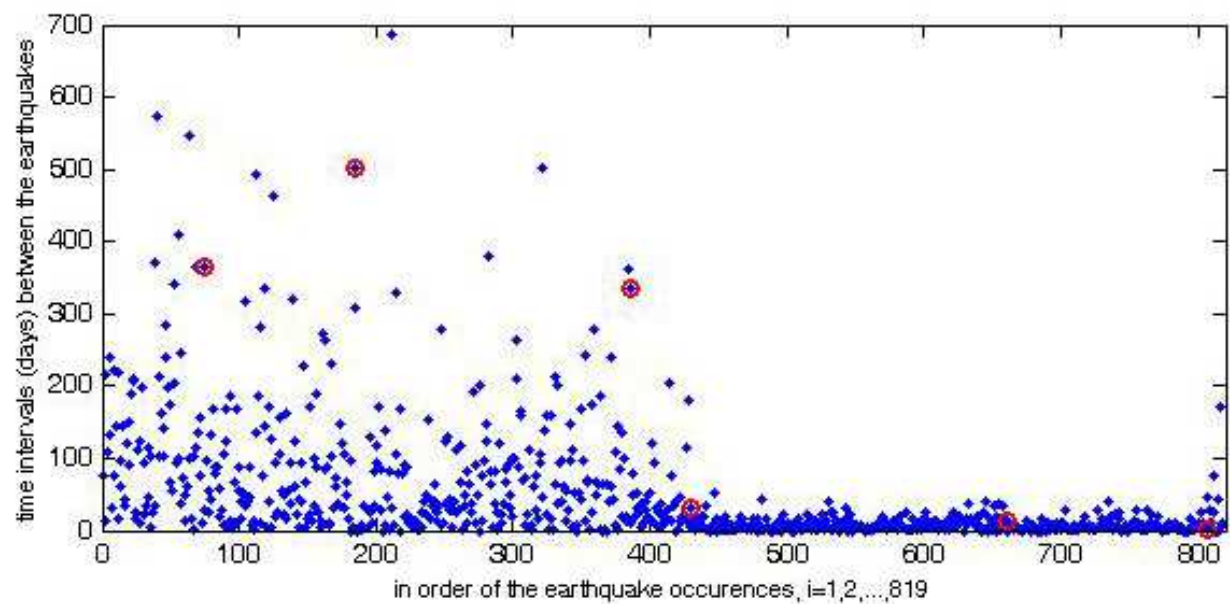


Fig. 4. Multiple change points in the earthquake data.

Estimated change points are shown on the scatter plot of the time intervals between the earthquakes.

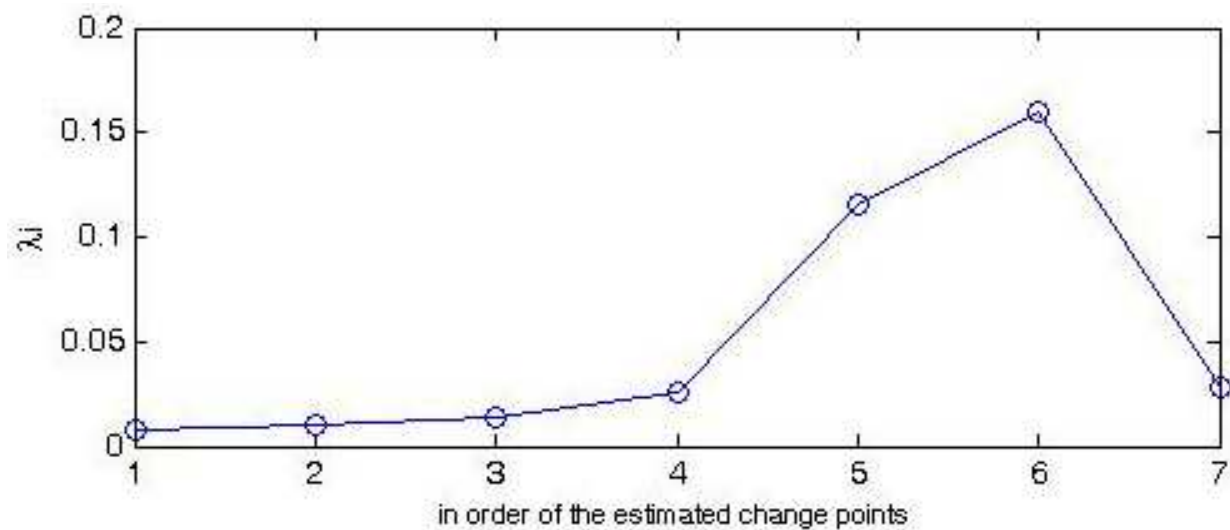


Fig. 5. The estimated occurrence rate of earthquakes in order of change points.

As can be seen in Figure 4 and Figure 5, the occurrence rate of the earthquakes increases slowly up until 03-February-2002, goes up sharply between 03- February-2002 and 12- January-2010, and tends to fall thereafter.

When using Bayesian method, the prior distributions of the parameters λ_1, λ_2 and τ , are taken to be respectively,

$$\begin{aligned} p_0(\lambda_1, \lambda_2 | \tau) &= \frac{1}{\lambda_1 \lambda_2}, \quad \lambda_1, \lambda_2 > 0 \\ p_0(\tau) &= \frac{1}{T}, \quad 0 < \tau < T \end{aligned} \tag{17}$$

where the prior distributions are called uninformative priors. The joint posterior distribution of the parameters is can be written as:

$$p_1(\lambda_1, \lambda_2, \tau) \propto L(\lambda_1, \lambda_2, \tau) p_0(\lambda_1, \lambda_2 | \tau) p_0(\tau) \\ \propto \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n-N_\tau} e^{-\lambda_2 (T-\tau)} \frac{1}{\lambda_1 \lambda_2 T} \quad (18)$$

With integrate Eq. (18) with respect to the parameters λ_1, λ_2 , the marginal posterior distribution of change point is proportional to

$$p_1(\tau) \propto \frac{\Gamma(N_\tau) \Gamma(n - N_\tau)}{\tau^{N_\tau} (T - \tau)^{n - N_\tau}} \left(\frac{1}{T} \right) \quad (19)$$

where $\Gamma(\cdot)$ is gamma function. The marginal posterior distributions of λ_1 and λ_2 are not obtained analytically because of discontinuity in τ and the close form of the marginal posterior distributions of λ_1 and λ_2 are respectively proportional to

$$p_1(\lambda_1) \propto \frac{1}{T} \int_0^T \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \frac{\Gamma(n - N_\tau)}{(T - \tau)^{n - N_\tau}} d\tau, \\ p_1(\lambda_2) \propto \frac{1}{T} \int_0^T \lambda_2^{n - N_\tau} e^{-\lambda_2 (T - \tau)} \frac{\Gamma(N_\tau)}{\tau^{N_\tau}} d\tau$$

where, for the possible values of τ , s_1, s_2, \dots, s_n , and the variable N_τ takes the values $1, 2, \dots, n$.

Form Eq.(19), the Bayesian estimate $\hat{\tau}$ of the change point in the earthquake data is found to be the date 03-February-2002 2002 which corresponds to the posterior mode. The posterior distribution of the change point is given in Figure 6.

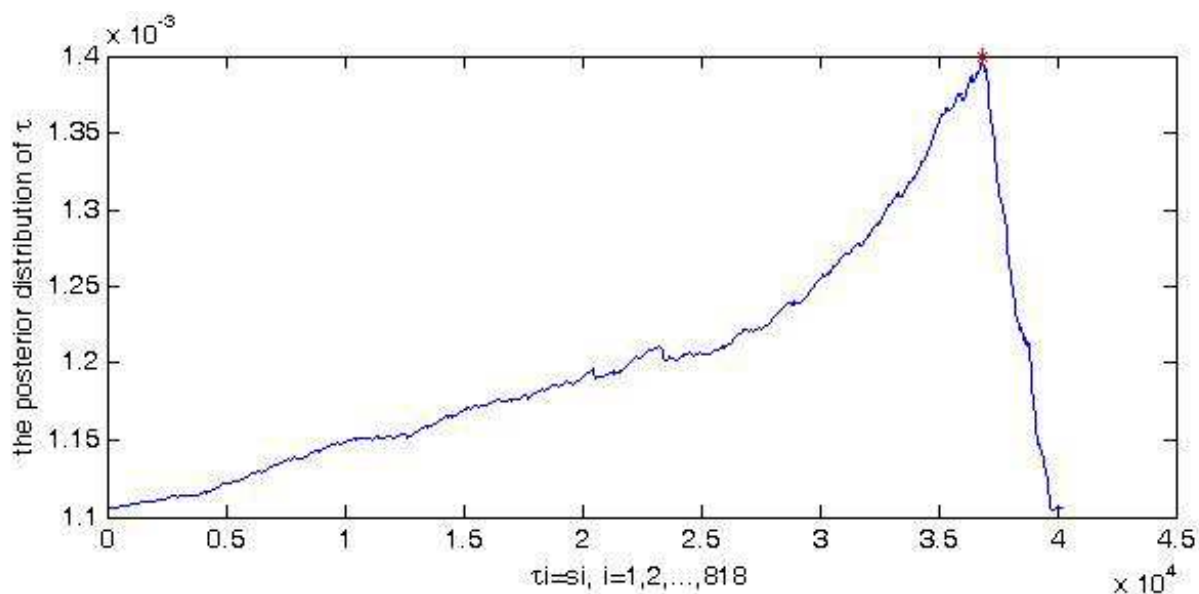


Fig. 6. The posterior distribution of change point and the posterior mode.

The Bayesian estimates of the parameters λ_1 and λ_2 would be the same as the maximum likelihood estimates because of uninformative prior distributions.

For multiple change points, the binary Bayesian segmentation procedure can be easily used. The procedure is employed as the mode of the posterior distribution of change point decreases considerably until the one before the posterior mode is found. The results are given in Figure 7.

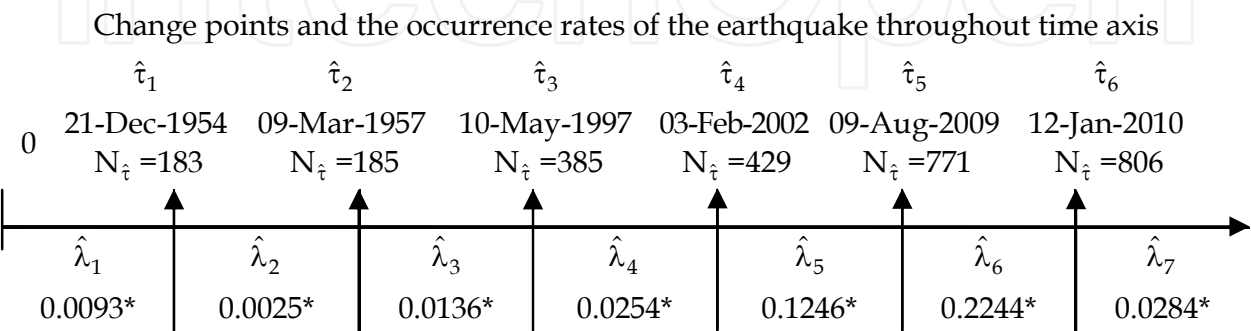


Fig. 7. The Bayesian estimations of parameters for multiple change points in the earthquake data (*in days).

When comparing the maximum likelihood estimates and the Bayesian estimates of the change points, we can see that the change points corresponding to dates such as 9-March-1957, 10-May-1997, 03-February-2002 and 12-January-2010 overlap. These dates are investigated in depth from many aspects, including overall world temperature, other disasters, and astronomical events.

When we look at the magnitudes of the earthquakes in the data, those of magnitude 6 to 6.9 number 291 (35.57%) and those of magnitude 7 to 7.9 number 258 (31.54%) (Table 2). The histogram of the magnitudes is given in Figure 6. Furthermore, using the Kolmogorov-Smirnov test, the distribution of the magnitudes is found to be almost normal ($p=0.000$).

Magnitude	4-4.9	5-5.9	6-6.9	7-7.9	8-8.9	≥ 9
The number of earthquakes	54	93	291	258	40	82
%	0.0660	0.1137	0.3557	0.3154	0.0489	0.1002

Table 2. The number of earthquakes with respect to magnitudes.

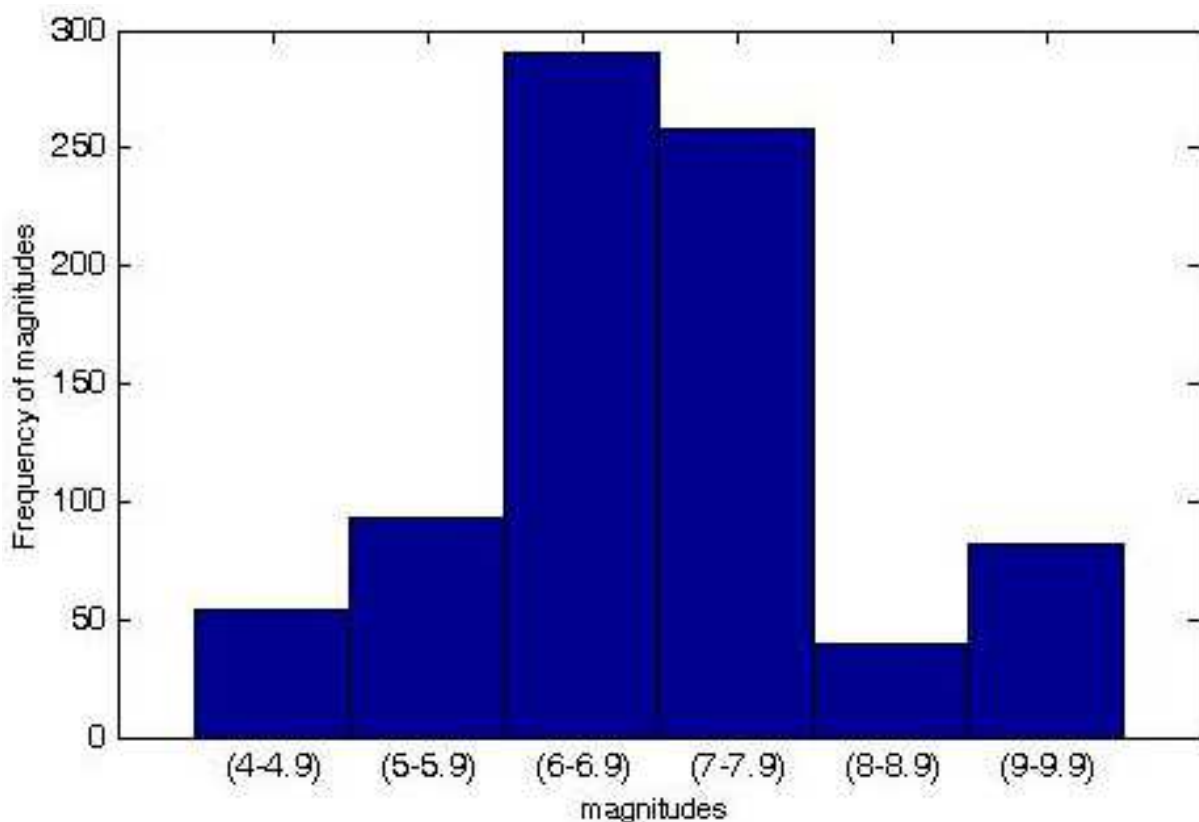


Fig. 8. Histogram of the earthquake magnitudes.

The sequence of the earthquake magnitudes Y_1, Y_2, \dots, Y_n can be assumed to be normally distributed random variables. The change point model in the mean of the normally distributed random sequences is written,

$$\begin{aligned} Y_1, Y_2, \dots, Y_v &\sim N(\mu_1, \sigma^2) \\ Y_{v+1}, Y_{v+2}, \dots, Y_n &\sim N(\mu_2, \sigma^2) \end{aligned} \quad (20)$$

where v is an unknown change point in the mean of the sequence of the normally distributed random variables. To investigate whether there is a change point in the earthquake magnitudes, under the change point model, the likelihood function of the observed values, y_1, y_2, \dots, y_n ,

$$\begin{aligned} L(\theta_1, \theta_2, v) &= \prod_{i=1}^v f(y_i, \mu_1, \sigma^2) \prod_{i=v+1}^n f(y_i, \mu_2, \sigma^2) \\ &\propto \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum_{i=1}^v (y_i - \mu_1)^2 - \sum_{i=v+1}^n (y_i - \mu_2)^2 \right) \right\} \end{aligned} \quad (21)$$

and the logarithm of the likelihood function in Eq. (21) is:

$$\ell n L(\theta_1, \theta_2, v) \propto -\frac{1}{2\sigma^2} \left(\sum_{i=1}^v (y_i - \mu_1)^2 - \sum_{i=v+1}^n (y_i - \mu_2)^2 \right) \quad (22)$$

For any given value of v , the maximum estimates of unknown parameters μ_1, μ_2 , are $\bar{y}_1 = \frac{\sum_{i=1}^v y_i}{v}$ and $\bar{y}_2 = \frac{\sum_{i=v+1}^n y_i}{n-v}$ respectively. The estimates are substituted into Eq.(22), and with the maximized the log likelihood function given by Eq.(22), we can obtain the maximum likelihood estimate of the change point,

$$\hat{v} = \arg \max_{k=1, \dots, n-1} \left[\frac{1}{2\sigma^2} \left(v \bar{y}_1^{-2} - (n-v) \bar{y}_2^{-2} \right) \right] \tag{23}$$

From Eq. (23), the change point is estimated to be the date 28-February-2011 corresponding to $v=817$. The estimated change point is close to end of the sequence such that it refers to no change point in the sequence. We can model both the number of the earthquake occurrences and the magnitudes of the earthquakes with compound Poisson process. Let $\{X_t, t \geq 0\}$ be a compound Poisson process as defined in Subsection 2.1

$$X_t = \sum_{i=1}^{N_t} Y_i$$

where N_t is the number of earthquakes in time interval $(0, t]$, which is a Poisson distributed random variable with parameter λt , and $\{Y_i, i = 1, 2, \dots\}$ are the magnitudes of the earthquakes, which are normally distributed random variables with parameters μ_i and σ^2 . The mean and variance of the magnitudes of the earthquakes according to the estimated change points dates, (9-March-1957, 10-May-1997, 03-February-2002 and 12-January-2010) are given in Table 3.

Estimated Change points	Up to 9-Mar-1957	Between 9-Mar-1957 and 10-May-1997	Between 10-May-1997 and 03-Feb-2002	Between 03-Feb-2002 and 12-Jan-2010	After 12-Jan-2010
Occurrence rate of the earthquakes	0.0090*	0.0136*	0.0254*	0.1300*	0.0284*
Mean of the magnitudes	6.9622	6.7250	7.0795	6.4562	7.1167
Variance of the magnitudes	0.9305	0.8821	0.5579	0.9845	1.2706

Table 3. Means and variances of the magnitudes with respect to estimated change point (*in days).

Using Table 3, we compute the recurrence periods for certain earthquakes of magnitude y , corresponding to the change point, 12-January-2010. The recurrence period of an earthquake with magnitude y is computed by

$$\text{Reccurence period} = 1 / (\text{Expected number of earthquakes}).$$

The expected number of earthquakes of magnitude y is computed by the multiplying the probability of occurrence class of magnitude with the expected number of earthquakes over a certain period of time.

Class Magnitude (y)	$F_Y(L)$ Expected	$F_Y(U)$ Expected	$P(L \leq y \leq U)$ Expected	Expected number of earthquakes (year)	Average recurrence period (year)
4-4.9	0.0028	0.0246	0.0218	0.2257	4.4306
5-5.9	0.0302	0.1402	0.1100	1.1404	0.8769
6-6.9	0.1609	0.4238	0.2629	2.7248	0.3670
7-7.9	0.4588	0.7565	0.2977	3.0856	0.3241
8-8.9	0.7834	0.9432	0.1598	1.6566	0.6037
≥ 9	0.9526	1.0000	0.0474	0.4912	2.0360

Table 4. The estimation of recurrence periods for certain earthquakes after the change point 12-January-2010.

Time	Expected number of earthquakes	Expected total magnitudes of earthquakes
30 days	0.8520	6.0634
3 months	2.5560	18.1902
6 months	5.1120	36.3805
1 years	10.3660	73.7717
2 years	20.7320	147.5434
3 years	31.0980	221.3151
4 years	41.4640	295.0868
5 years	51.8300	368.8585
10 years	103.6600	737.7171
20 years	207.3200	1475.4342

Table 5. Expected number of earthquakes and the expected total magnitudes of earthquakes with respect to the change point 12-January-2010.

Number of earth.	30 days 11-Feb-10	3 months 12-Apr-10	6 months 11-Jul-10	1 year 12-Jan-11	2 years 11-Jan-12	3 years 11-Jan-13	4 years 11-Jan-14	5 years 11-Jan-15	10 years 11-Jan-20	20 years 11-Jan-30
0	0.426560956	0.077614579	0.006024023	3.1485E-05	9.91304E-10	3.12112E-14	0	0	0	0
1	0.363429935	0.198382863	0.030794805	0.000326373	2.05517E-08	9.70605E-13	0	0	0	0
2	0.154821152	0.253533299	0.078711521	0.001691593	2.13039E-07	1.50919E-11	0	0	0	0
3	0.043969207	0.216010371	0.134124431	0.005845017	1.47224E-06	1.56443E-10	1.16755E-14	0	0	0
4	0.009365441	0.138030627	0.171411023	0.015147361	7.63063E-06	1.21627E-09	1.21028E-13	0	0	0
5	0.001595871	0.070561256	0.17525063	0.031403509	3.16396E-05	7.56469E-09	1.00366E-12	0	0	0
6	0.000226614	0.030059095	0.149313537	0.054254796	0.000109326	3.92078E-08	6.93599E-12	0	0	0
7	2.75821E-05	0.010975864	0.109041543	0.080343602	0.000323791	1.74183E-07	4.10848E-11	0	0	0
8	2.9375E-06	0.003506789	0.069677546	0.104105222	0.000839104	6.77095E-07	2.12943E-10	3.99619E-14	0	0
9	2.78083E-07	0.000995928	0.039576846	0.119906081	0.001932923	2.33959E-06	9.8105E-10	2.30136E-13	0	0
10	2.36927E-08	0.000254559	0.020231684	0.124294643	0.004007336	7.27565E-06	4.06783E-09	1.1928E-12	0	0
11	1.83511E-09	5.91503E-05	0.009402215	0.117130752	0.007552736	2.05689E-05	1.53335E-08	5.62024E-12	0	0
12	1.30292E-10	1.2599E-05	0.004005344	0.101181448	0.01304861	5.33044E-05	5.29823E-08	2.42747E-11	0	0
13	8.53917E-12	2.47716E-06	0.001575024	0.08068053	0.020809521	0.000127512	1.68989E-07	9.67815E-11	3.26026E-14	0
14	5.19669E-13	4.52259E-07	0.000575109	0.05973817	0.030815928	0.000283241	5.00498E-07	3.58299E-10	1.44839E-13	0
15	2.95172E-14	7.70649E-08	0.000195997	0.041283058	0.042591721	0.000587215	1.38351E-06	1.23804E-09	6.00562E-13	0
16	0	1.23111E-08	6.26211E-05	0.026746261	0.055188223	0.001141327	3.58536E-06	4.01048E-09	2.33453E-12	0
17	0	1.85101E-09	1.88305E-05	0.016308926	0.067303661	0.002087822	8.74491E-06	1.22273E-08	8.5411E-12	0
18	0	2.62844E-10	5.34787E-06	0.009392129	0.077518861	0.003607061	2.01444E-05	3.52077E-08	2.95123E-11	0
19	0	3.53594E-11	1.43886E-06	0.005124148	0.084585317	0.005903809	4.39614E-05	9.60429E-08	9.66079E-11	0
20	0	4.51893E-12	3.67772E-07	0.002655846	0.08768114	0.009179833	9.11408E-05	2.48895E-07	3.00431E-10	0
25	0	0	2.01378E-10	4.98583E-05	0.052673373	0.041877043	0.001752055	1.46016E-05	4.38567E-08	0
30	0	0	4.11104E-14	3.48963E-07	0.011797303	0.071223597	0.012557101	0.000319369	2.3869E-06	0
35	0	0	0	1.07217E-09	0.001159888	0.05317575	0.039506887	0.003066385	5.70261E-05	0
40	0	0	0	1.6252E-12	5.62613E-05	0.019586801	0.061322008	0.014525139	0.000672161	4.94125E-13
150	0	0	0	0	0	0	0	0	0	3.67978E-06

Table 6. Probability of earthquake occurrences after the change point 12-January-2010.

5. Conclusion

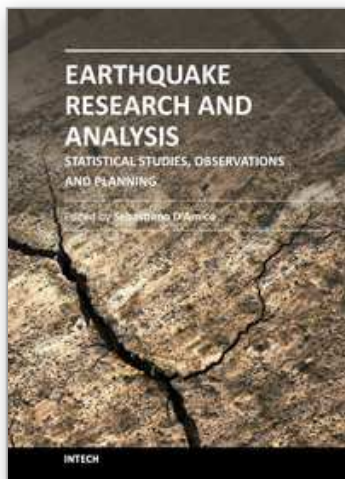
Predicting earthquakes is of crucial importance these days. Many researchers have studied earlier attempts at earthquake detection. The change point analysis is used in both backward (off-line) and forward (on-line) statistical research. In this study, it is used with the backward approach in the worldwide earthquake data. The change points found in the worldwide earthquake data are useful in making reliable inferences and interpreting the results for further research. Each date found as the change point in the earthquake data should be carefully investigated with respect to other geographical, ecological, and geological events or structures in the world.

6. References

- Akman V. E. & Raftery A. E. (1986). Asymptotic inference for a change-point Poisson process, *The Annals of Statistics*, Vol.14, No.4, pp. 1583–1590.
- Aktaş, S.; Konşuk, H. & Yiğiter, A. (2009). Estimation of change point and compound Poisson process parameters for the earthquake data in Turkey, *Environmetrics*, Vol. 20, No.4, pp. 416-427
- Amorese, D. (2007). Applying a Change-Point Detection Method on Frequency-magnitude Distributions, *Bulletin of the Seismological of America*, Vol.97, No.5, pp.1742-1749
- Boudjellaba, H.; MacGibbon, B. & Sawyer, P. (2001). On exact inference for change in a Poisson Sequence. *Communications in Statistics A: Theory and Methods*, Vol.30, No.3, pp. 407–434
- Chen J. & Gupta A. K. (1997). Testing and locating variance change points with application to stock prices, *Journal of the American Statistical Association: JASA*, Vol.92, No. 438, pp. 739-747
- Fotopoulos S. B. & Jandhyala V. K. (2001). Maximum likelihood estimation of a change-point for exponentially distributed random variables, *Statistics & Probability Letters*, 51: 423–429
- Kagan, Y. Y. (2010). Statistical Distributions of Earthquake Numbers: Consequence of Branching Process, *Geophysical Journal International*, Vol.180, No.3, pp.1313-1328
- Leonard T. & Papasouliotis, O. (2001). A Poisson Model for Identifying Characteristic Size Effects in Frequency Data: Application to Frequency-size Distributions for Global Earthquakes, “Starquakes”, and Fault Lengths, *Journal of Geophysical Research*, Vol.106, No.B7, pp.13,413-13,484
- Lindsey H. (June 2007). 'Natural' catastrophes increasing worldwide, 14.06.2007, Available from: http://www.wnd.com/news/article.asp?ARTICLE_ID=40584/
- Mandeville M. W. (June 2007). “Eight Charts Which Prove That Chandler's Wobble Causes Earthquakes, Volcanism, El Nino and Global Warming”, 14.06.2007, Available from: http://www.michaelmandeville.com/polarmotion/spinaxis/vortex_correlations2.htm
- Parzen, E. (1962). *Stochastic Processes*, Holden-Day, ISBN 0-8162-6664-6, San Francisco, USA
- Raftery A. E. & Akman V. E. (1986). Bayesian analysis of a Poisson process with a change point, *Biometrika*, Vol.73, pp.85–89
- Rotondi R. & Garavaglia, E. (2002). Statistical Analysis of the Completeness of a Seismic Catalogue, *Natural Hazards*, Vol.25, No. 3, pp. 245-258

- Rydelek, P. A. & Sacks I. S. (1989). Testing the completeness of earthquake catalogues and the hypothesis of self-similarity, *Nature*, Vol. 337, pp.251–253
- Schorlemmer, D. & Woessner, J. (2008). Probability of Detecting an Earthquake, *Bulletin of the Seismological of America*, Vol.98, No.5, pp. 2103-2117
- Smith, A. F. M. (1975).A Bayesian approach to inference about a change point in a sequence of random variables, *Biometrika*,Vol.62, pp. 407-416
- U.S. Geological Survey web page 16. 03. 2011, Available from:
<http://earthquake.usgs.gov/regional/world/historical.php>
- Yang, T. Y. & Kuo, L. (2001). Bayesian binary segmentation procedure for a Poisson process with multiple change points, *Journal of Computational and Graphical Statistics*, Vol. 10, pp. 772–785
- Yao, Q. (1993). Tests for change-points with epidemic alternatives. *Biometrika*, Vol.80, pp. 179-191
- Yigiter, A. & İnal, C. (2010). Estimation of change points in a homogeneous Poisson process with an application to earthquake data, *Pakistan Journal of Statistics*, Vol. 26, No. 3, pp. 525-538

IntechOpen



Earthquake Research and Analysis - Statistical Studies, Observations and Planning

Edited by Dr Sebastiano D'Amico

ISBN 978-953-51-0134-5

Hard cover, 460 pages

Publisher InTech

Published online 02, March, 2012

Published in print edition March, 2012

The study of earthquakes plays a key role in order to minimize human and material losses when they inevitably occur. Chapters in this book will be devoted to various aspects of earthquake research and analysis. The different sections present in the book span from statistical seismology studies, the latest techniques and advances on earthquake precursors and forecasting, as well as, new methods for early detection, data acquisition and interpretation. The topics are tackled from theoretical advances to practical applications.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ayten Yiğiter (2012). Change Point Analysis in Earthquake Data, Earthquake Research and Analysis - Statistical Studies, Observations and Planning, Dr Sebastiano D'Amico (Ed.), ISBN: 978-953-51-0134-5, InTech, Available from: <http://www.intechopen.com/books/earthquake-research-and-analysis-statistical-studies-observations-and-planning/change-point-analysis-in-earthquake-data>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen