

EVIDENCE FOR INDUCED SEISMICITY FOLLOWING THE 2001 SKYROS MAINSHOCK

A.K. Adamaki,¹ G. M. Tsaklidis,¹ E. E. Papadimitriou,¹
and V. G Karakostas.¹

¹ Department of Geophysics, School of Geology, Aristotle University of Thessaloniki,
GR54124 Thessaloniki - Greece, aadama@geo.auth.gr, ritsa@geo.auth.gr, vkarak@geo.auth.gr

² Mathematics Department, Aristotle University of Thessaloniki, GR54124 Thessaloniki – Greece,
tsaklidi@math.auth.gr

Abstract

Estimation of the seismicity rate changes caused by a major earthquake is based upon the assumption that the earthquake occurrence can be described by stochastic processes. Three stochastic models are applied to the data, i.e. the homogeneous Poisson model, the non-homogeneous Poisson model with two different rate functions, and the Autoregressive model AR(2). The two latter models seem to be adequate to properly simulate the earthquake production in a given area. The identification of the model which best fits the data, enables the estimations of the seismicity rate changes and the numbers of the earthquakes following a specific main shock.

Key words: Seismicity rates, induced seismicity, Skyros aftershock sequence

1. Introduction

Many researchers have focused on applying statistics and the probability theory to earthquake sequences, in order to study the temporal and spatial distribution of triggered activity following large earthquakes. A useful tool in statistics for modeling and analyzing spatial data is a point process, which can be used as a model for random events in time. Many point process models, where each point represents the time and location of an event, have been proposed (Vere-Jones, 1992; Ogata, 1999; among others) and several attempts have been made to estimating the changes in earthquake production that have been caused by a specific event. The special case of seismicity rate decreases (i.e. seismicity shadows) was also examined in previous studies (Marsan, 2003; Marsan and Nalbant, 2005).

The simplest approach to this goal is to consider a stationary process taking into account only the background seismicity (Toda et al., 1998, 2002), which however leads to underestimation of the rate change (Marsan, 2003; Felzer et al., 2003). Then, stationarity cannot be generally assumed and the declustering of the catalog is a common way to remove nonstationarity from the data (Matthews and Reasenber, 1988; Kilb et al., 2000; Gombert et al., 2001; Wyss and Wiemer, 2000). On the other hand, aftershocks comprise a major portion of an earthquake catalog including important information about the rate changes. Therefore other researchers have tried to model the aftershocks rather than remove them, e.g. Marsan (2003) tested several models such as the autoregressive model, a sum of N power-laws, proposed by Utsu (1970) as a generalization of the model presented in Ogata and Shimazaki (1984) and Woessner et al. (2004) and the ETAS model (Ogata, 1988).

The purpose of the present work is to study the temporal structure of triggered sequences at short time scales after a large earthquake. Therefore certain statistical models are applied on data from the 2001 Skyros sequence, i.e. the homogeneous Poisson model, the non-homogeneous Poisson model and the autoregressive model AR(2).

2. Methods

In order to analyze the data, firstly the homogeneous and non-homogeneous Poisson model is considered:

Homogeneous and Non – homogeneous Poisson model: A Poisson process with constant rate λ , i.e. a time-independent rate, is known as a homogeneous Poisson process. In this case, the waiting times of the point process are exponentially distributed with a mean $\mu=1/\lambda$. The expected number of earthquakes in any interval of length t equals to λt , and the probability that there are exactly n occurrences in this interval is given by

(1)

By replacing the constant λ with a function $\lambda(t)$, which gives the rate of earthquakes at time t , the process becomes a non-homogeneous Poisson process. In this case, the number of earthquakes in any interval is also time dependent, and the mean rate in an interval $[t, t+\Delta t]$ is given by

(2)

The probability that there are exactly n occurrences within the specific interval is then given by:

(3)

Two functions expressing $\lambda(t)$ are tested in the present study, namely

(4)

and

(5)

both selected because of their property of allowing the rate decaying as time passes; they differ in that $\lambda(t)$ given by equation (4) decays rapidly to zero while $\lambda(t)$ given by (5) may decay smoothly to 0 (by a suitable choice of the parameters).

In each case a specific region must be selected, with dimensions a few times larger than the main rupture length. Determining the study area, the threshold magnitude must be defined to ensure the completeness of the data set. The next step is to define the duration of the earthquake catalog, which depends on the purpose of the specific study. This means that, naming T_0 the time of occurrence of the main shock, the catalog expands within the interval $[T_0, T]$ where T is located several days after T_0 .

Continuing with the data processing, the time intervals between consecutive events occurring at times t_i (t_i are measured from the beginning of the observation period, i.e. T_0) are computed and tested using the chi-square test; the null hypothesis is that the inter-arrival times are exponentially

distributed with a mean μ . If the null hypothesis is true, then the number of occurrences follows a Poisson distribution with a rate λ , which is constant during the tested interval (homogeneous Poisson process).

Next, since the rate of the aftershock sequence over the first few days after the main shock is characterized by strong changes, (different) constant rates are assumed over sub-intervals of the time period examined. To deal with that, the time interval $[T_0, T]$ is separated into short intervals $[T_i, T_{i+1}]$, each one tested for its homogeneity as a Poisson process with a rate parameter λ_i . Each λ_i is computed taking into consideration the number of events in the respective time interval, which is not necessarily equal in all sub-intervals. The entire set $\{\lambda_i\}$ is then used as the input data to fit the selected equation that best describes their evolution within $[T_0, T]$. In the present study the equation for the function $\lambda(t)$ which is fitted to the data is of the form $\lambda(t) = a e^{-bt}$, and the parameters a and b are estimated by the least squares' method.

Finally, the non-homogeneous Poisson process is assumed to have a rate function of the form $\lambda(t) = a t^b$, the parameters a and b being estimated by the maximum likelihood method. This intensity function can be an increasing or decreasing function, according to whether $b > 1$ or $b < 1$ respectively. Based on the form of $\lambda(t)$, the aforementioned process is named a process with a Weibull rate function or simply a Weibull process. The special case of $b = 1$ is the exponential case, which has a constant hazard rate and is characterized by the memoryless property. The maximum likelihood estimators for a and b are given by

$$(6)$$

Autoregressive Model: Another way to model the sequence of the earthquakes is to consider a random variable $Z(t), t \geq 0$, representing the number of earthquakes at any time t . The set $\{Z(t), t \geq 0\}$, constitutes a time series, i.e. a family of stochastic processes. The model which is fitted to the data is the autoregressive model of order $p, p \in \mathbb{N}^+$, abbreviated as $AR(p)$. In statistics, signal processing e.t.c., $AR(p)$ is often used to model and predict various types of natural phenomena. The $AR(p)$ model assumes that

$$(7)$$

where ϕ_i are the unknown parameters and α_t is a white noise function with a zero mean and variance σ^2 . The time series is considered to be stationary. Then the parameters ϕ_i can be estimated using several methods, e.g. the Yule-Walker equations which constitute a set of linear equations relating the unknown parameters with the autocorrelations. Then:

$$(8)$$

where the coefficients $\rho_i, i = 1, 2, \dots, p-1$ stand for the autocorrelations which give the correlations of the values of the process against time-shifted versions of themselves.

3. Models application to the 2001 Skyros sequence

An investigation is now performed to the 2001 Skyros sequence, with the main shock of $M_w = 6.4$ to have occurred on July 26, 2001 Skyros. Based on the data and results for this sequence from Karakostas et al. (2003), a region is selected as mentioned before, i.e. a few times larger than the rup-

Table 1. The values of the means μ of the exponential distributions of the inter-arrival times, and the λ -rates of the respective Poisson distributions.

time period	No of data	statistical X^2	μ	λ
$(-600, T_0)$	19	3.9660	31.5454	0.0317
$(T_0, 4)$	116	2.1857	0.0340	29

ture length. The catalog used was found to be complete for events with $M_c \geq 3.5$. The events are separated into two sets, for the time intervals; $[T_0 - t_B, T_0]$ and $[T_0, T_0 + t_A]$, where T_0 is the time of occurrence of the main shock and t_B and t_A denote selected time periods before and after the main shock, respectively.

3.1 Homogeneous Poisson process

The homogeneous Poisson process assumes constant rates, λ_B and λ_A , before and after T_0 , respectively. It is selected that $t_B=600$ days and $t_A=4$ days and the homogeneous Poisson model is fitted to the recorded data in order to estimate the number of events in each time interval. The goodness of fit is tested by means of the chi-square test. Within $[T_0 - t_B, T_0]$ 19 events occurred preceding T_0 whereas 116 events followed T_0 , occurring into the interval $[T_0, T_0 + t_A]$. The inter-arrival times of the earthquakes are then tested for being exponentially distributed. The values of the statistical X^2 , the parameter μ of the exponential distribution of times between subsequent events, and the estimated mean λ (per day) of the two homogeneous Poisson processes, before and after the main shock, are summarized in Table 1.

600 days before the main shock there are 19 data, giving a rate of 0.0317 events per day, with a mean inter-arrival time of length 31.5454 days. On the other hand, 4 days after the earthquake there are 116 data, which lead to a rate of 29 events per day and a mean inter-arrival time of length 0.034 days. As it is shown in Table 1, those time intervals were tested by means of the chi-square test, and the statisticals X^2 are smaller than the critical value which is equal to 5.02 at the $\alpha=0.05$ significance level (for two degrees of freedom).

The following step is to compute the probability of having specific numbers of events in the two different time intervals, using equation (1), where λ takes the values λ_B and λ_A . The duration t of the time period $(0, t)$ appearing in equation (1) can either be smaller or larger than t_B and t_A . If it is assumed that the process does not change for times outside the intervals $[T_0 - t_B, T_0]$ and $[T_0, T_0 + t_A]$, then t can be larger than t_B and t_A (extrapolation). Then the probability derived via (1) refers to a homogeneous Poisson process with the same parameter λ as the one estimated within the corresponding interval. The results of this procedure can also be presented using the survival function which is derived through the Poisson cumulative distribution function given by

$$(9)$$

and representing the probability of having up to N earthquakes during the procedure with rate λ . In

Figure 1 the survival functions $S(x)$ estimated for the time periods before and after

the main shock are shown; the plots are derived in Matlab by connecting the points $S(x)$, $x=0, 1, \dots$. Then, the probabilities of having more (or less) than any number of events can be derived by

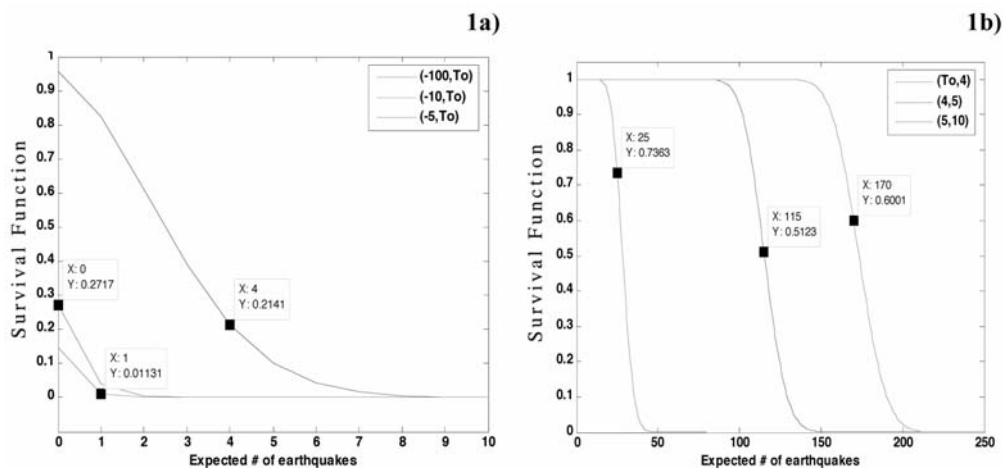


Fig. 1: Estimated survival curves of the number of events in the time periods (a) before and (b) after the main shock, using the homogeneous Poisson model.

means of the survival functions, e.g. in Figure 1b the probability of having more than 17 earthquakes in the interval [1, 2) is 0.989 (18 events are observed), the probability of having more than 25 earthquakes during the fifth day is 0.736 (18 events are observed) and the probability of having more than 170 earthquakes in the time interval (5, 10) is 0.600 (36 events are observed). It is obvious that the homogeneous Poisson model does not fit well the data, especially for large values of t . This observation leads to a need of replacing the constant rate λ with $\lambda(t)$.

3.2 Non-homogeneous Poisson process

Dividing the aftershock sequence into subsets corresponding to short time intervals, each one is tested for its homogeneity as a Poisson process (using the chi-square test). The length of the intervals is not necessarily the same, whereas there are different values, λ_i , for the rates of each interval. The values of means μ_i (of the exponentially distributed inter-arrival times) and mean λ_i are shown in Table 2. In Figure 2 the values of λ_i are plotted versus time t (in days) after T_0 (main shock occurrence time), which is mentioned as 0-time. Because of the fact that λ is decaying with time t , formula (4) can be adopted. In order to estimate the values of the parameters a and b in (4), the least-squares method is applied and the results are given in Table 3. Now, for the first three days after T_0 , the time interval is divided into four subintervals and the respective four-different-values of λ_i are found. Fitting the curve $\lambda(t)=$ to the data is shown in Figure 3.

If the estimated $\lambda(t)$ can be accepted also for larger intervals than the first three days since the main shock, then it comes out that the number of earthquakes during those intervals will not change significantly, as the function $\lambda(t)$ decays to zero rapidly as t increases. It should be mentioned that assuming that the process does not change out of the time borders which were set for the parameter estimation, is a rather arbitrary admission. In Figure 4b the theoretical cumulative number of events along with the observed data is presented.

In Figure 4a the estimated probabilities of having a number of events in different time intervals are presented, using the survival function. Equation (3) can be used in this case to estimate the Poisson probability, where the expected number of events in any interval $(t; t+\Delta t)$ is $\lambda(t; t+\Delta t)$, given by (2);

Table 2. The values of the parameters μ of the exponential inter-arrival times, and the rates λ of the Poisson distributions in subsequent time intervals.

Time period (in days)	μ	Λ
$(T_{\alpha_0}, 0.12)$	0,0046	208,3333
$(0.12, 0.36)$	0,0096	104,1667
$(0.36, 1)$	0,0217	43,75
$(1, 3)$	0,0695	14,5
$(3, 5)$	0,0643	13,5
$(5, 9)$	0,1390	7,5
$(9, 16)$	0,2167	4,5714
$(16, 36)$	0,5936	1,65

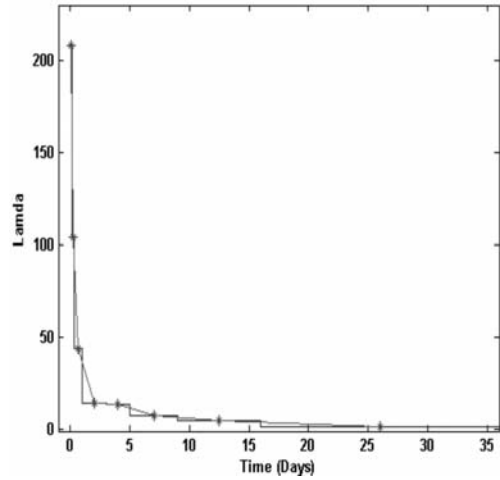


Fig. 2: Rates λ_i plotted versus time (in days).

Table 3. Estimated rate function $\lambda(t) = \exp(\alpha + b \cdot t)$, using the least squares method.

	Value	Standard Error
α	5,49772	0,10446
b	-3,06949	0,66913
Adj. R² = 0,96145		

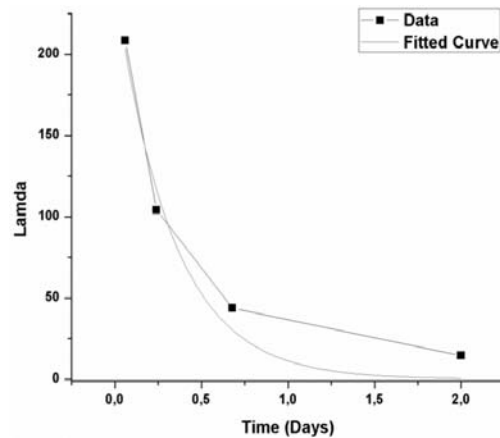


Fig. 3: Observed data λ_i and fitted function $\lambda(t) = e^{\alpha + bt}$.

e.g. the probability of having more than 50 earthquakes within the interval $[T_0, 1]$ is 0.9993 (78 events are observed), and the probability of having more than 90 earthquakes in the interval $[T_0, 5]$ is 0.111 (134 events are observed). Obviously, the model underestimates the number of events as time increases, as it is shown in Figure 4b.

3.3 Non-homogeneous Poisson process with a Weibull rate function

In this case equation (5) is used to compute the expected number of events in any time interval as given in (2) and (3). In Table 4 that follows, the estimated values of parameters a and b for the Weibull rate functions are shown. These parameters are derived from equations (6) using the maximum likelihood method. The values presented in Table 4 come up from different sets of data each one being a subset of the one that follows. For example, if the data come up from observing only

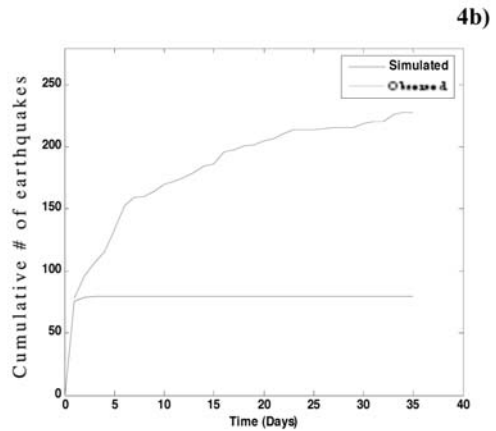
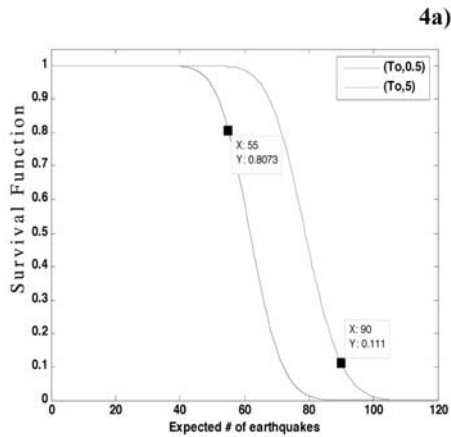


Fig. 4: a) Survival functions of the number of events in time periods after the main shock, using the non-homogeneous Poisson model with $\lambda(t) = e^{a+bt}$. b) Cumulative density function of the theoretical number of the aftershocks (via the non-homogeneous Poisson model), and the cumulative number of observed data.

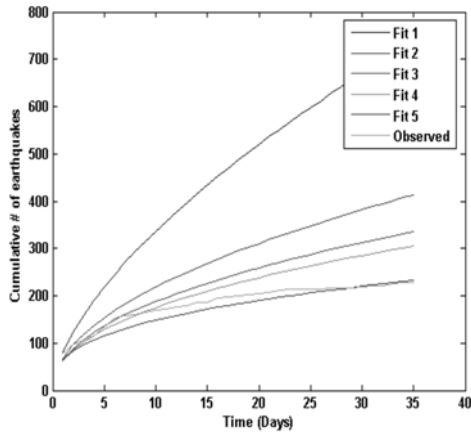


Fig. 5: Cumulative number of earthquakes for the Weibull-rates case for different data set.

Table 4. Estimated parameters of the Weibull distributions

Time period (in days)	estimated a	estimated b
(1 st day)	$9.1338 \cdot 10^{-4}$	0.6257
1 st -2 nd day	$2.5054 \cdot 10^{-4}$	0.5085
1 st -3 rd day	$1.2807 \cdot 10^{-4}$	0.4647
1 st -4 th day	$8.7278 \cdot 10^{-5}$	0.4435
1 st -30 th day	$8.7228 \cdot 10^{-6}$	0.3584

the first day after T_0 , then the expected number of events for the following days is overestimated, as shown in Figure 5. Taking more and more data into account (from the first 2, 3, 4 days etc) the estimation of the oncoming process comes closer and closer to the observation and the model fits the data well.

Assuming that the underlying process is the Weibull process which is known up to the third day after T_0 , equations (2) and (3) are used to compute the possibility of having specific numbers of events into several time intervals, and then the survival function is estimated. The results are summarized in Figure 6, and the probabilities of having more (or less) than any number of events can be derived by means of the survival functions. For example, the probability of having more than 10 earthquakes during the fourth day is 0.8971 (9 events are recorded), and the probability of having

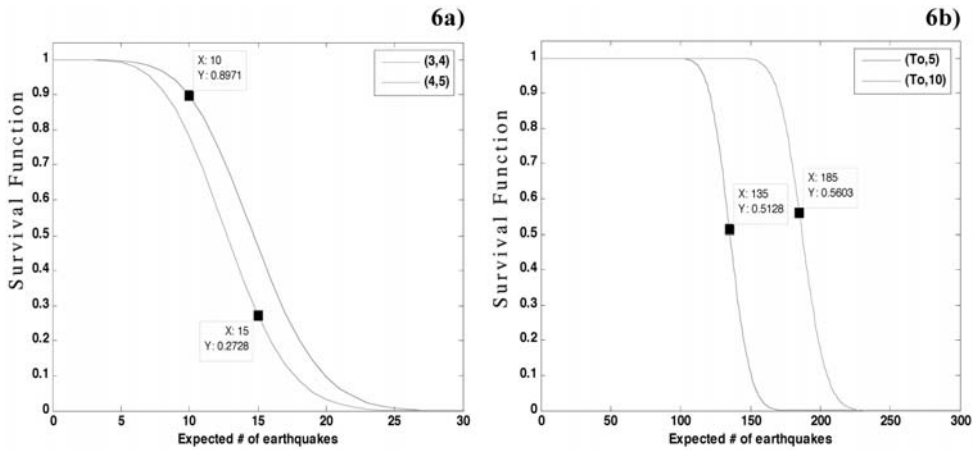


Fig. 6: Survival functions of the number of events in various time periods after the main shock, using the non-homogeneous Poisson model with $\lambda(t)=\alpha^{-b}bt^{b-1}$.

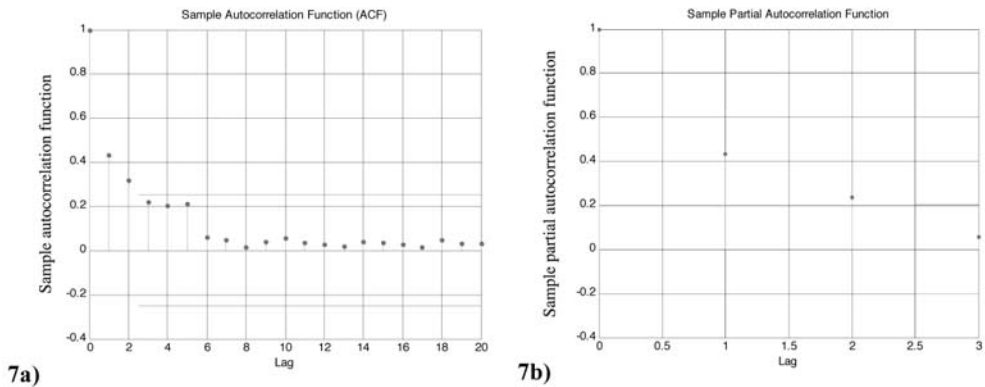


Fig. 7: a) Sample autocorrelation function (ACF) and b) sample partial autocorrelation function (PACF) for the number of events of the time period $[T_0, 20]$.

more than 135 earthquakes within $[T_0, 5]$ is 0.5128 (134 events are recorded). It turns out that the model is adequate for describing the earthquake sequence, especially for $1 < t < 10$.

3.4 Autoregressive model AR(p)

The number of earthquakes can be considered as a time series and among the several models dealing with that, the Autoregressive model AR(p) can be chosen and applied to the data, following Marsan (2003). AR(p) models explain a series of observations where the value of an observation depends on the (values of the) p past observations. Here the random variable Z_t of the AR(p) model represents the number of earthquakes at time t. Finding the order p of the model and using the Yule-Walker equations to estimate the unknown parameters, the numbers of earthquakes, Z_t , in different future time intervals can be estimated.

The first step is to estimate the partial autocorrelation coefficients in order to determine the order p

Table 5 and 6. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) and the bounds ($\alpha=0.05$).

ACF	Bounds
0.4267	0.2884
0.3126	-0.2884
0.2079	
0.1995	
0.2041	

PACF	Bounds
0.4288	0.2884
0.2374	-0.2884
0.0465	

of the model. If $p=k$, $k \in \mathbb{N}^+$, is selected and the k^{th} autocorrelation comes out to be statistically insignificant, then the last variable may be omitted and the real order of the model is less than k . In Figure 7 and Table 6 the partial autocorrelation function is shown. In this case the initial assumption is $p=3$. The values z_1 taken into account are the number of earthquakes in subsequent time intervals with duration 2.4 hours each. It turns out that the assumption of the order to be 3 must be rejected ($\alpha=0.05$), and the order ends up to be $p=2$. In the correlogram shown in Figure 7, the autocorrelation coefficients are found to be close to zero for $k \geq 3$. The next step is to estimate the parameters of the AR(2) model, using the Yule-Walker equations given in (8). It is found that $\phi_1=0,4772$ and, $\phi_2=0,1831$, thus

The values of the parameters indicate that the AR(2) is stationary since $\phi_1+\phi_2<1, \phi_2-\phi_1<1, -1<\phi_2<1$.

4. Discussion–Conclusions

In the present study certain statistical models are applied to the 2001 Skyros aftershock sequence, aiming to figure out whether there is a specific statistical approach of the phenomenon which can lead to a simulation of its evolution. The earthquakes' sequence is assumed to be a stochastic process, following the homogeneous or the non-homogeneous Poisson model. The homogeneous Poisson model with constant rate λ doesn't fit well the data, especially in the case of long time periods which were tested, before and after the main shock. This fact leads to the need of applying a non-homogeneous Poisson process with a time dependent rate, $\lambda(t)$, which is considered to obey two certain forms (equations (4) and (5)). In the first case the rate $\lambda(t)$ is rapidly decaying to zero as time t passes, implying that no events are expected after a few days following the main shock. This was not in accordance with the observation, as the aftershock sequence does not actually decay to zero in such a short time interval. In the second case the Weibull-type rate $\lambda(t)$ is considered (equation 5), and the associate parameters are estimated using the maximum likelihood method. It turns out that the non-homogeneous Poisson model fits quite well the data and the results, presented in the previous section, show that this model provides a good approach to the aftershock sequence, compared with the previous model.

Finally, the earthquake sequence is tested via time series analysis, using the Autoregressive model of order 2; that is the number of earthquakes at some time interval was supposed to depend linearly on the number of the earthquakes of the two preceding time intervals. The data analysis shows that the AR(2) model fits adequately the data.

5. Acknowledgments

The paper is benefited by thorough revision from George Drakatos. Geophysics Department contribution 753.

References

- Felzer, K.R., Abercrombie, R.E., and Brodsky, E.E. (2003), Testing the stress shadow hypothesis, EOS Trans. AGU 84(46), Fall Meet. Suppl., Abstract S31A-04.
- Gomberg, J., Reasenberg, P.A., Bodin, P., and Harris, R.A. (2001), Earthquake Triggering by Seismic Waves Following the Landers and Hector Mine Earthquakes, *Nature* 411, 462–466.
- Karakostas, V.G., Papadimitriou, E.E., Karakaisis, G.F., Papazachos, C.B., Scordilis, E.M., Vargemezis, G., and Aidona, E. (2003), The 2001 Skyros, Northern Aegean, Greece, earthquake sequence: off-fault aftershocks, tectonic implications, and seismicity triggering, *Geophys. Res. Lett.* 30, doi:10.1029/2002GL015814.
- Kilb, D., Gomberg, J., and Bodin, P. (2000), Triggering of Earthquake Aftershocks by Dynamic Stresses, *Nature* 408, 570–574.
- Marsan, D. (2003), Triggering of Seismicity at Short Timescales Following Californian Earthquakes, *J. Geophys. Res.* 108, doi:10.1029/2002JB001946.
- Marsan, D., and Nalbant, S. S. (2005), Methods for measuring seismicity rate changes: A review and a study of how the Mw 7.3 Landers earthquake affected the aftershock sequence of the Mw 6.1 Joshua Tree earthquake, *Pure Appl. Geophys.*, 162, 1151–1185.
- Matthews, M.V. and Reasenberg, P.A. (1998), Statistical Methods for Investigating Quiescence and other Temporal Seismicity Patterns, *Pure Appl. Geophys.* 126, 357–372.
- Ogata, Y. (1988), Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes, *J. Am. Stat. Ass.* 83, 9–27.
- Ogata, Y. (1999), Seismicity Analysis through Point-process Modeling: A Review, *Pure Appl. Geophys.* 155, 471–507.
- Ogata, Y. and Shimazaki, K. (1984), Transition from Aftershock to Normal Activity: The 1965 Rat Islands Earthquake Aftershock Sequence, *Bull. Seismol. Soc. Am.* 74, 1757–1765.
- Toda, S., Stein, R.S., and Takeshi, S. (2002), Evidence from the AD 2000 Izu Islands Earthquake Swarm that Stressing Rate Governs Seismicity, *Nature* 419, 58–61.
- Toda, S., Stein, R.S., Reasenberg, P.A., Dieterich, J.H., and Yoshida, A. (1998), Stress Transferred by the 1995 Mw=6.9 Kobe, Japan, Shock: Effect on Aftershocks and Future Earthquake Probabilities, *J. Geophys. Res.* 103, 24543–24565.
- Utsu, T. (1970), Aftershocks and Earthquake Statistics (II)-Further Investigation of Aftershocks and other Earthquake Sequences Based on a New Classification of Earthquake Sequences, *J. Fac. Sei., Hokkaido University, Ser. VII*, 3, 197–266.
- Vere-Jones, D., Statistical Methods for the Description and Display of Earthquake Catalogs. In *Statistics in the Environmental and Earth Sciences* (eds. Walden A.T. and Guthorp P.) (Edward Arnold Publisher 1992), pp. 220–246.
- Woessner, J., Hauksson, E., Wiemer, S., and Neukomm, S. (2004), The 1997 Kagoshima (Japan) Earthquake Doublet: A Quantitative Analysis of Aftershock Rate Changes, *Geophys. Res. Lett.* 31, doi:10.1029/2003 GL018858.
- Wyss, M. and Wiemer, S. (2000), Change in the Probability for Earthquakes in Southern California due to the Landers Magnitude 7.3 Earthquake, *Science* 290, 1334–1338.