Università degli Studi di Bologna
Facoltà di Scienze Matematiche Fisiche e Naturali
Dottorato di Ricerca in Geofisica – XIX Ciclo

Development and Application of Stochastic Models of Earthquake Occurrence

Ph.D. Thesis of:
Anna Maria Lombardi

Supervisor:
Dr. Warner Marzocchi

Co-supervisor:
Prof. Paolo Gasperini

Director:
Prof. Michele Dragoni

Settore scientifico disciplinare: GEO10

March 2007

Contents

Introduction

In the study of relations between seismicity dynamics and underlying physical processes, a leading role, highlighted also by recent development of fault interaction models [*Stein*, 1999], is played by the question of how seismicity rates evolve with time. Several statistical methods have been developed in the past to serve this purpose. The general framework for statistically estimating seismicity rate changes is formalized by the Theory of Point Processes [*Daley and Vere-Jones*, 2003]. The temporal dynamic in a interval $[T1, T2]$ of such processes is fully described by the mean rate (called also *intensity* or *hazard*) $\lambda(t)$, that is ultimately related to expected number of events in a certain time period. Two main points determine basic properties of such as modeling: 1) the dependence from past history (*time memory*) and 2) the *stationarity*. As regard the first point the memory concerns the influence that the occurrence of an event has on future seismic rate and is modeled by introducing in expression of $\lambda(t)$ the occurrence times of past events. The second point means that the main statistical descriptors of data (for example the mean rate) are invariant for different temporal non-overlapping ranges of the same size. Specifically in probability theory a stochastic process X_t is called *stationary* if, for all n, $t_1 < t_2 < \ldots < t_n$, and $h > 0$, the joint distribution of $[X(t_1 + h), \ldots, X(t_n + h)]$ does not depend on h [e.g., *Cox and Lewis*, 1966; *Daley and Vere-Jones*, 2003]. This means that the statistical description of the process is invariant with respect to shifts of the starting time; then the stochastic behavior of a stationary process is the same, no matter when the process is observed. The stationary and memory properties are absolutely not correlated: we can have point processes of which temporal trend is driven by all possible combination of stationary/nonstationary behaviour and lack/presence of memory of the past history (see Table1).

Moreover the term ”nonstationary” has not to be mistaken with the misleading term

	MEMORYLESS	**WITH MEMORY**
STATIONARY	**Stationary Poisson** rate: λ $\mathbb{P}[N(s,t+s)=n]=\frac{e^{-\lambda t}(\lambda t)^n}{n!}$	rate: $\lambda(t/H_t)$ H_t past history $\mathbb{P}[N(s,t+s)=n]=\frac{e^{-\int_s^{t+s}\lambda(x/H_x)dx}\left(\int_s^{t+s}\lambda(x/H_x)dx\right)^n}{n!}$
NONSTATIONARY	**Nonstationary Poisson** rate: $\lambda(t)$ $\mathbb{P}[N(s,t+s)=n]=\frac{e^{-\int_s^{t+s}\lambda(x)dx}\left(\int_s^{t+s}\lambda(x)dx\right)^n}{n!}$	

Table 1: Classification of stochastic point processes basing on memory and stationary properties. $N(s,t+s)$ indicates the number of events into interval time $(s,t+s)$. Its probability distribution is univocally identified by the rate of the process $\lambda(t)$.

"time-dependent". This actually refers to the fact that the hazard rate $\lambda(t)$ depends on the time t. All the time-dependent processes so far proposed in seismology (ETAS, Brownian Passage Time, Weibull, etc...) are stationary, because the parameters of the models do not vary with time.

In occurrence of seismic events the problem of memory is ultimately related to triggering or, more generally, to modulation of seismic activity by earthquakes, in consequence of relaxation of the tectonic strain. This is a complex phenomenon that involves earthquakes of different size, over spatio-temporal scales that can be much larger than the rupture length and duration of initial triggering earthquake. The difficulty in estimating the changes in earthquake production caused by a given earthquake, as well as in describing their evolution in space and time, lies in making sure whether observed variations in seismicity are effectively due to a shock or not. While this can be trivial on short spatio-temporal scale, the problem becomes much more difficult at longer scales.

The existence of time memory is particularly obvious after moderate-large shallow earthquakes for which the seismicity rate of the region increases for a certain time period. These triggered events are usually called aftershocks if their magnitude is smaller than the first event. However the definition of an aftershock contains unavoidably a degree of arbitrariness: the qualification of an earthquake as "aftershock" requires the specification of time and space windows, that are often more based on common sense that on hard science.

There is an intense research activity and a heated debate on the possibility that events interact on spatio-temporal scales well wider than those interested by aftershocks occurrence. In the simple view of a single isolated fault with a constant stress rate, adopted in seismic hazard assessment, earthquakes occur periodically by rupturing the whole fault, with a period equal to the ratio of the stress drop divided by the rate of stress loading [*Working Group*, 2002 and references therein]. Actually there are many evidences that such faults interact, causing significative changes in seismic rate [see *King and Cocco*, 2000 and reference therein]. How such changes in seismicity depend on the relative locations of the faults or on the time between the earthquakes occurrence is still a widely open question.

In terms of stochastic and physical modeling, the memory is taken into account to describe the well-recognized short-term triggering proprieties of seismicity. The two most important types of short-term earthquake clustering are *mainshock-aftershock sequences* and earthquake *swarms*. The main physical process thought responsible of aftershocks occurrence is the static stress changes [*Stein*, 1999], but also other mechanisms, as locally induced fluid flows [*Nur and Booker*, 1972] or dynamic stress variations [*Gomberg et al.*, 1998], are been adduced to explain the occurrence of these patterns of seismicity. The occurrence rate $\lambda(t)$ for aftershocks is generally described by the modified Omori Law $\lambda(t) \sim (c + \Delta t)^{-p}$ where Δt is the time elapsed since the mainshock [*Utsu et al.*, 1995]. On the other side the occurrence of seismic swarms is mainly ascribed to intrusion of magma (in volcanic zones) or fluids and to following redistribution of stress, caused by reduction of the resistance of faults [*Kisslinger*, 1975; *Noir et al.*, 1997]. In contrast to mainshock-aftershocks sequences, earthquake swarms are not characterized by a dominant earthquake; their temporal evolution, complex and locally variable, cannot be described by any simple relation comparable to the Omori law. They appear to be different in their temporal features and energy release from stress triggered aftershocks sequences. Whereas the power-law decay rate of an aftershock sequence reflects the process of stress relaxation following a large magnitude earthquake, the short temporal patterns of near-equal magnitude events in most earthquake swarms appear driven by magmatic processes or pore-fluid movements within the crust.

Actually the earthquake swarm activity and tectonic earthquake clusters share some common features [*Hainzl and Fisher*, 2002]. In particular embedded aftershock sequences, recognized in seismic swarms, according to the Omori law, point out an important role of stress triggering also for such as pattern of seismicity [*Hainzl*, 2004]. Specifically if a large earthquake occurs during a swarm, the activity following this earthquake may be regarded as its aftershocks. Frequency of such triggered earthquakes does not decrease regularly, because it is actually a mixture of the swarm events and aftershocks triggered by the large earthquake. This effect cannot be ignored when we examine the correlation between some external physical processes and the occurrence of earthquakes. On the other side fluid flow can play a leading role also in triggering seismicity of typical mainshock-aftershocks

sequences [*Nur and Booker*, 1972; *Antonioli et al.*, 2005]. The recognition of these common features for so different patterns of seismicity has two main consequences. The first is that it justifies the use of the same modeling to describe time evolution of mainshock-aftershocks sequences and of seismic swarms. The second is that the possible variety of physical processes directly linked to seismic capability of a region could produces a nonstationary behaviour of shocks occurrence. This problem is rarely properly discussed in application of stochastic models. But to take into account the chance of a nonstationary trend is crucial in any time analysis, especially in a statistical one: standard statistical techniques used in estimating parameters or testing any hypothesis are often largely invalid in cases where the set of variables is not entirely stationary.

The role of probability in the study of earthquake occurrence is of primary importance. Computation of a probability of triggering over various space and time scales should improve our understanding of how earthquakes interact with each other. Moreover some statistical tools permit to recognize departures by stationary assumption of earthquakes occurrence and to interpret possible nonstationarity in terms of underlying physical processes. A well-established tool to explore all these issues is the Epidemic-Type Aftershock Sequences (ETAS) model [*Ogata*, 1988; 1998]. This is a stochastic point process incorporating the empirically observed characteristics of stress triggered activity: its main peculiarity is that each earthquake has some magnitude-dependent ability to trigger its own Omori law type aftershocks. In particular the ETAS model describes the total seismic rate as the sum of two contributions: the ”background rate”, that refers to activity which is not triggered by precursory events and is forced by external physical processes, and the rate of events internally triggered by stress variations of previous earthquakes. We stress that this model ascribes a well defined meaning to the term ”background”, lacking of an objective and univocal definition (Suffice it to say, for example, that Coulomb stress triggering models [*Toda et al.*, 1998] call background rate the *a priori* rate of all events, aftershocks included, occurred in a certain time interval; this is compared to the rate of aftershocks production, following occurrence of a strong event, to take into account global rate changes caused by static stress triggering).

The ETAS model, formulated to describe typical mainshock-aftershocks occur-

rence, can reproduce the main characteristics of the swarm as well [*Hainzl and Ogata*, 2005]. In this case background rate refers to activity forced by pore pressure changes more than by the stress loading.

The problem of identification of nonstationarity can be solved by taking into account significative variations in parameters of the ETAS model, more directly linked to physical processes responsible of seismicity. In Chapter 1 we apply a general stochastic temporal ETAS modeling to characterize and to interpret the time evolution of two swarms: the Izu Islands (Japan) seismic swarm, occurred in 2000, and the 1983-1984 swarm occurred in the Phlegrean Fields (Italy). The method is developed along the line suggested by *Hainzl and Ogata* [2005] and it accounts for a possible nonstationary behaviour of the process by introducing time variations of parameters that may provide more stringent constrains on the nature of the process. The two swarms analyzed here are very different: whereas the highly energetic Izu swarm is clearly linked to magma motion (the swarm was accompanied by five phreatic eruptions of volcano Miyakejima), the involvement of magma chamber for more moderate Phlegrean Fields swarm is still a question in debate. The study of time variations of most meaningful parameters of the ETAS model can be an interesting tool to interpret information coming from seismicity in terms of underlying physical processes.

The nonstationary ETAS model seems to be an interesting investigation tool also for tectonic seismic sequences that show a time evolution hard to interpret in terms of stress relaxation. As example we present the application to a complex seismic sequence occurred in central Italy in 1997-1998 (Chapter 2). The coherent and significative variations of some parameters of the ETAS model can be interpreted as an evidence that fluid flow was among the driving processes of the sequence.

If the discussion on short-term time distribution of earthquakes mainly concerns the stationary of the process, on longer spatio-temporal scale the debate on basic time features of seismogenetic process is much more open. Because of the lack of enough long complete seismic catalogs recording small magnitude events and considering the obvious implications for seismic hazard assessment, the studies on long-term temporal evolution of seismicity basically concern moderate-strong

events. Despite some decades of effort there is still no conclusive assertion, on both theoretical and practical grounds, about long-term dynamics of seismic activity. The achievement of an agreement on these issues is complicated by some problems: the too short time period recovered by catalogs to test adequately any hypothesis, the lack of an unambiguous definition of term ”large earthquake”, that is related to seismic capability and tectonic structure of the region, the uncertainty of geologic and geodetical data, from which most hazard and long-term forecasting models are derived. All these inefficiencies strongly affect formulation of long-term models, that are often more based on subjective belief than on checks on data. For example, apart from the model used, the seismic hazard calculations are mostly based on two assumptions: the lack of long-term and long range interactions between events (faults are mainly considered as isolated systems) and the stationary of seismogenetic process [*Cornell*, 1968; *Working Group*, 2002]. In regard with the memory in large earthquake occurrence, the frequency of events for a source is modeled by renewal processes that impose the independence by previous occurrence history (Poisson process) [*Kagan and Jackson*, 1994] or, at most, the dependence by time of last event, only (see [*Working Group*, 2002] and references therein). The non poissonian renewal processes are mostly in agreement with the supposed periodicity of fault slips [*Ellsworth et al.*, 1999; *Nishenko and Buland*, 1987] or are derived by theoretical assumptions as the seismic gap hypothesis [*McCann et al.*, 1979] or the time-predictable model [*Shimazaki et Nakata*, 1980]. On the other side some paleoseismological studies [*Hurbert-Ferrari et.al*, 2005; *Ritz et.al*, 1995; *Friedrich et al.*, 2003; *Weldon*, 2004, 2005] have lately questioned the reliability of these models, by suggesting a clustering or a nonstationary behavior on single seismogenic structures. Moreover some statistical studies [*Kagan and Jackson*, 1994; 1999; 2000; *Faenza et al.*, 2003; *Rhoades and Evison*, 2004] point out the importance of the interactions between faults respect to behaviour of a single source. The problem if the seismic sources are or not isolated systems is made clear also by some evidence of coupling between faults [*Chéry et al.*, 2001a, 2001b; *Mikumo et al.*, 2002; *Pollitz*, 1992; *Pollitz et al.*, 1998; *Pollitz et al.*, 2003; *Rydelek and Sacks*, 2003; *Corral*, 2004, 2005; *Santoyo et al.*, 2005; *Thatcher*, 1983; *Piersanti et al.*, 1995; 1997; *Piersanti*, 1999; *Kenner and Segall*, 2000], mostly explained by postseismic viscoelastic interaction. These results could suggest, as a more reasonable

methodology for hazard calculations, to consider larger areas composed by multiple sources. As a matter of fact the Poisson paradigm is still implicitly accepted in many practical applications, mainly related to seismic hazard assessment; the real effect of this a priori assumption in a system with possible strong local interactions are still not clear.

One of the main questions about the distribution of high magnitudes earthquakes is the degree of similarity with smaller events in short-term time behavior. This issue involves the complex problem of the "universality" of earthquake distribution [*Bak et al.*, 2002], that means the dependence of main features of seismicity on magnitude-spatio-temporal scales. Dealing with the short-term triggering, it is not definitively ascertained if the physical processes governing clustering are independent by magnitude window considered. By applying the ETAS model to worldwide strong ($M \geq 7.0$) seismicity of the last century, we find that the estimated parameters are consistent with the values computed for moderate-small events in tectonic sequences (Chapter 3), showing the space and magnitude scale-invariance of main features of elastic stress triggering. Moreover we find that in some regions a nonstationary version of the ETAS model, obtained by modeling "background seismicity" by a piecewise constant Poisson process, better describes the time evolution of data respect to the classical ETAS model. The simplicity of the model, chosen in alternative to stationary Poisson process, and the paucity of data do not permit to further investigate the main temporal features of background seismicity and make difficult the interpretation of our finding by a physical point of view. Considering the high magnitude threshold of dataset used in our analysis, the background seismicity is in this context mainly related to global dynamics of surface plate tectonics. Therefore the identified time variation of mean background rate could reflect an irregularity in local tectonic loading on decadal timescale. However this interpretation does not seem to be confirmed by global geodetic studies, showing a substantial stationarity in tectonic motion on time scale of decades [*Sella et al.*, 2002]. The use of the memoryless nonstationary Poisson process (see Table 1) in our model does not permit to include the other more likely cause of recognized time variation in mean seismic rate: the long-term triggering of seismicity or, more generally, the existence of a memory of the past besides the short-term interactions. This is another aspect of the above mentioned "universality", more related to spatio-temporal scaling prop-

erties of earthquake occurrence, that identifies the interaction between events as a driving and general (i.e. at each spatio-temporal scale) feature of time evolution of seismic activity. To understand which is (if there is), among memory of past history and nonstationary behaviour, the predominant feature of long-term (and long-range) earthquake occurrence could help to improve present forecasting models and hazard methodologies.

We deal with this difficult subject in Chapter 4, by modeling the ”background” seismicity of moderate-strong events in different magnitude-spatio-time windows by the Fractionally Integrated Autoregressive Moving-Average (ARFIMA) time series analysis [*Granger and Joyeux*, 1980; *Hosking*, 1981]. The ARFIMA processes provide a flexible class of models able to represent a nonstationary behaviour as well as a long-memory trend of a time series. Therefore they are a useful investigation tool able to provide information on the main temporal features of earthquake occurrence, to use in an acquainted long-term modeling of seismicity. The results of application of ARFIMA modeling to our datasets seem to exclude that a nonstationarity behaviour is the cause of recognize departure by Poisson hypothesis, favouring the long-term memory as the more likely feature, at least on a temporal scale up to some centuries. The resulting scenario coming from this finding shows a seismicity externally forced by an almost stable global dynamic, with internally originated fluctuations, mainly raising by the long-term memory of the system. The recognized stationarity of strong events occurrence reassures on seismic hazard potentiality, showing a system of which statistical properties remain constant in time and therefore of which future rate can be ”predicted” with some confidence from an adequate sample of past records.

On the basis of these results we propose a new long-term time-dependent model (Chapter 5). This new model is based on the ”self-exciting” modeling [*Daley and Vere-Jones*, 2003], the same that has produced the ETAS model, and is directed towards identification of interactions between events not usually recognized as member of the same sequence. The application on worldwide and historical italian seismicity shows very interesting results. We find that this new model improves description of data respect to the Poisson model, showing a systematic and locally variable influence of events on following remote (in space and time) seismicity. This finding can be interpreted in terms of long-term stress transfer of strong events and can be

adduced to support the postseismic relaxation theory [*Pollitz*, 1992; *Piersanti et al.*, 1995]. The conversion of the couplings found into a well defined changes in probability occurrence open new prospects for the time-dependent hazard assessment and the long-term forecasting practice.

Chapter 1

Testing stationarity hypothesis for seismic swarms

1.1 Introduction

The two most important types of earthquake clustering, mainshock-aftershocks sequences and seismic swarms, are usually assumed to result from different physical processes. If stress triggering is identified as the most important mechanism for aftershocks sequences [*Stein*, 1999], seismic swarm are thought to be mainly triggered by an intrusion of fluids reducing the resistance of faults [*Kisslinger*, 1975; *Noir et al.*, 1997]. Of consequence seismic swarms are thought to differ significantly in temporal clustering and energy release from aftershocks sequences [*Scholz*, 2002]. In contrast to these, those do not contain a dominant earthquake and their temporal evolution cannot be described by any simple law, as the Omori law.

Actually also earthquakes induced by fluids themselves produce local stress field changes. Each earthquake within the swarm redistributes stress, which may in turn influence the subsequent swarm evolution, especially if the crust is in a critical state. Therefore most of the complexity of a seismic swarm spatio-temporal distribution is probably linked to the contribution of different source processes that, in general, can be related both to the presence of magma or fluids and to seismic interactions. Then the earthquake swarm activity can also share some common features with tectonic earthquake clusters. In particular the recognized embedded aftershock sequences,

according to Omori law, point out an important role for stress triggering also for seismic swarms [*Hainzl and Fisher*, 2002; *Hainzl*, 2004].

Until few years ago, the most remarkable observational feature about the nature of a seismic swarm was the occurrence of low-frequency events that has been usually considered as evidence of the presence of fluids in the generating process [*Chouet et al.*, 1994; *Chouet*, 1996; *Neuberg*, 2000]. Only recently, some researchers [*Toda et al.*, 2002; *Hainzl and Ogata*, 2005] gave new important insights to interpret a seismic swarm. Specifically, *Toda et al.* [2002] describe the spatio-temporal evolution of a seismic swarm at the Izu Islands through the co-seismic stress variations induced by a constantly growing dyke emplaced at the beginning of the swarm. Conversely, *Hainzl and Ogata* [2005] show that the epidemic-type aftershock sequence (ETAS) model is an appropriate tool to extract the primary fluid signal from the complex seismicity patterns. They studied a seismic swarm at Vogtland through a stochastic ETAS model with the additional feature of a background seismicity varying through time. Basically, they identify earthquakes due to seismic interaction and then estimate the temporal variation of the background seismicity that is interpreted in terms of a temporal variation of the source process energy (i.e., a magma/fluid migration).

Here, we apply a general stochastic nonstationary ETAS modeling to characterize and to interpret the physical time evolution of two swarms: the Izu Islands (Japan) seismic swarm, occurred in 2000, and the 1983-1984 swarm occurred in the Phlegrean Fields (Italy). The method is developed along the line suggested by *Hainzl and Ogata* [2005], but it accounts for possible time variations of other parameters that may provide more satisfactory fit and more stringent constrains on the nature of the process. Specifically, we compare the performance of two different ETAS models (including the one proposed by *Hainzl and Ogata*, [2005]) and investigate the time behavior of some important parameters of the model. This allows the time variation of the model parameters to be interpreted in terms of the spatio-temporal evolution of the seismic swarms under study and to yield physical constrains of the driving process. In the following, we neglect the space variables of the events in stochastic modeling, assuming that the whole seismic area is homogeneous from a statistical point of view. This assumption is justified by the small dimension of the active areas under study.

1.2 Stochastic modeling of a seismic swarm

We model a seismic swarm by using a temporal stationary *Epidemic-Type Aftershocks Sequences* (ETAS) time model [*Ogata*, 1988; 1998], and a generalization of this model that allows time variation of the parameters to be accounted for. Here, with the term *stationary* we mean a process whose parameters do not vary through time.

In order to include secondary aftershocks activity, ETAS stochastic model describes the short-time clustering features of earthquakes as superposition of modified Omori functions [*Utsu*, 1961] shifted in time. The total occurrence rate at a time t is given by the sum of triggering rates of all preceding events and of a time-independent background rate ν [see *Ogata*, 1988, 1998]:

$$\lambda(t) = \nu + \sum_{t_i<t} \frac{K}{(t - t_i + c)^p} e^{\alpha(m_i - M_{min})}. \tag{1.1}$$

The parameter K measures the productivity of the aftershock activity, whereas α defines relation between triggering capability and magnitude m_i of a triggering event. The parameter c measures incompleteness of catalog in the earliest part of each cluster, caused by lowering in detectability of stations after a strong event [*Kagan*, 2004]. The parameter p controls the temporal decay of triggered events. M_{min} is the completeness magnitude. Estimation of the model parameters $\{\mu, K, c, p, \alpha\}$ is carried out by maximizing the log-likelihood [*Daley and Vere-Jones*, 2003]. Given the occurrence times of collected earthquakes $\{t_i, i = 1, ..., N\}$, the log-likelihood ($LogL$) of the time ETAS model, in an interval time $[T_1, T_2]$, is given by

$$LogL(\nu, K, c, p, \alpha) = \sum_{i=1}^{N} \log \lambda(t_i/\mathcal{H}_{t_i}) - \int_{T_1}^{T_2} \lambda(t/\mathcal{H}_t)dt \tag{1.2}$$

[*Daley and Vere-Jones*, 2003]. To find parameters that maximize this, we use the Davidon-Fletcher-Powell optimization procedure [*Fletcher and Powell*, 1963], which provides also a numerical approximation of errors.

The rationale to consider also nonstationary ETAS modeling is based on the fact that stationary ETAS model is not always able to fully explain temporal pattern of real seismicity, especially for seismic swarms. In many volcanic areas, seismic ac-

tivity is strongly controlled both by fluid intrusion as by stress triggering and shows mixed occurrence of mainshock-aftershocks sequences and magma-related swarms. As a consequence, background as well as inducted activity could be strongly characterized by nonstationarities; for example, *Matsu'ura* [1983] shows that temporal evolution of triggered seismicity cannot be always described by a single Omori law.

In particular, we generalize the ETAS model by considering the time variations of ν and p ($\nu(t)$ and $p(t)$, respectively). We take into account only these two parameters because they are the most clearly linked to the time evolution of a possible magma/fluids source and, therefore, they are the most likely candidates for significant time variations. Specifically, the p-value is found to be positively correlated with crustal temperature, which controls stress release and therefore aftershocks decay [*Mogi*, 1967; *Kisslinger and Jones*, 1991], while time variation of ν is usually associated to the time evolution of the energy of the source [see *Hainzl and Ogata*, 2005].

To explore the nonstationarities of our datasets we follow a procedure similar to that used by *Hainzl and Ogata* [2005]. We fit the ETAS model in a moving non-overlapping time window τ. The choice of τ is a balance that accounts for two opposite requirements: the need to have a short time window to follow the details of the time evolution of the process, and the need to have a large time window including enough data for the calculations. In each time interval we estimate the (joined) variation of the parameter(s) that change with time, setting all the other parameters to the values found for the whole sequence. In computing log-likelihood we take in account all past occurrence history, to include probability that an earthquake is triggered by an event occurred in a previous interval time.

1.3 The case of the 2000 Izu Islands seismic swarm

1.3.1 The 2000 seismic swarm

A seismic swarm occurred at the Izu Islands (150 km south of Tokyo) in June-August 2000. The swarm is composed by more than 5000 $M \geq 3$ events, and 5 $M \geq 6$ shocks. The energy released during the swarm was about one order of

magnitude larger than the one released at Long Valley during the seismic unrest occurred at the beginning of the eighties [*Toda et al.*, 2002]. In order to identify objectively the start and the end of seismic swarm, we look for change points in the sequence of the number of events per day for the whole year 2000.

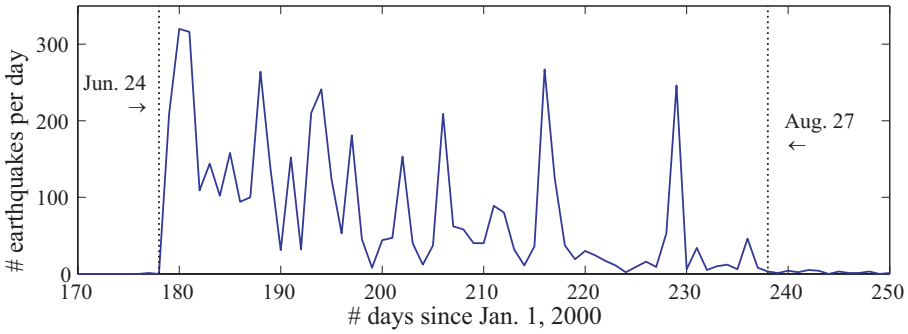

Figure 1.1: Number of earthquakes per day at the Izu Islands for the year 2000. The vertical dotted lines indicate the start and the end of the swarm as suggested by the change point analysis.

For this purpose we use the nonparametric method suggested by *Mulargia and Tinti* [1985] (see Appendix 1A). The procedure is based on the Kolmogorov-Smirnov two samples test, and it provides satisfactory answers in more cases: when the number of regimes (distinct temporal patterns) is unknown, the regimes follow different statistical distributions, and the regimes may involve a relatively small sample size. The method assumes that the behavior of the time series is piecewise stationary, but exhaustive simulations on synthetic data sets have shown it to be efficient also for systems with smooth variations [*Mulargia et al.*, 1987]. In Figure 1.1 we report the results of the analysis with the arrows indicating the two most significant (significance level $\ll 0.01$) change points identified, corresponding to June 24 (start of the swarm) and to August 27 (end of the swarm).

1.3.2 Discussion of the results

As mentioned before, we model the Izu Islands seismic swarm by using three stochastic models: 1) a stationary ETAS model (model I), 2) an ETAS model with only the background $\nu(t)$ changing with time (model II) [see *Hainzl and Ogata*, 2005], and 3) an ETAS model with both $\nu(t)$ and $p(t)$ varying through time (model

III). For our dataset we choose $M_{min} = 3.0$ [*Toda et al*., 2002].

Parameter	**value**
ν	7.7 ± 0.8 day^{-1}
k	0.005 ± 0.002 day^{p-1}
p	2.4 ± 0.3
c	0.023 ± 0.005 day
α	0.45 ± 0.06

Table 1.1: Estimated ETAS Parameters of the Izu Eartquake Swarm

In Table 1.1 we report the values and the relative errors of parameters estimated for the stationary case (Model I). The high p-value indicates a sharp decaying after-shocks activity, in agreement with previous studies on the Izu area seismicity [*Utsu et al*, 1995; see also *Toda et al.*, 2002].

To fit nonstationary models, the length of τ is set to one day. Note that $\tau = 1$ day is suitable to study the process evolution of the Izu Islands seismic swarm that has characteristic time of few days [*Toda et al*., 2002].

Model	$\boldsymbol{\tau}$	**AIC**
I		-43156
II	1 day	-43138
II	5 days	-44044
III	1 day	-43210
III	5 days	-44186

Table 1.2: Values of AIC of the Izu Swarm for Model I,II and III.

In order to verify the stability of the results we also carry out all the calculation for $\tau = 5$ days. In order to select which is the "best" of three models, we calculate the AIC values for all of them [*Akaike*, 1974]. The AIC statistic is defined by

$$AIC(K) = -2LogL + 2K \tag{1.3}$$

where K is the number of parameters, and $LogL$ is the log-likelihood of the model, given by (1.2), computed for the best parameters. In comparing models with different numbers of parameters, addition of the quantity $2K$ roughly compensates for the additional flexibility which the extra parameters provide. The lower value of the AIC identifies the model that better represents the data.

In Table 1.2 we report the values of AIC for three models. These indicate that model III is the best one to describe the data regardless the value of τ, and that a time window of 5 days seems the most appropriate to describe the process evolution.

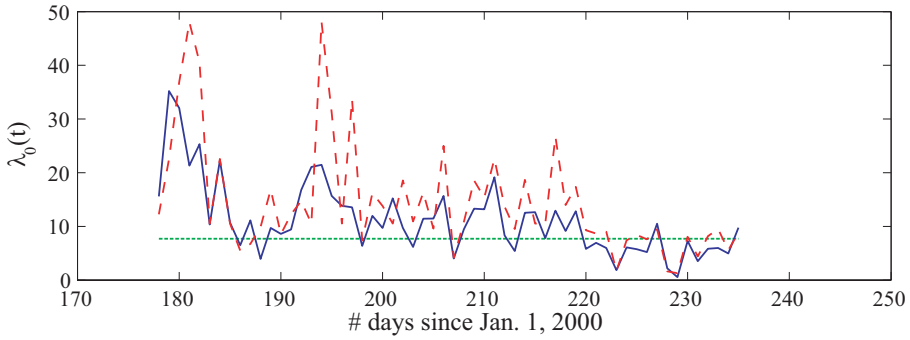

Figure 1.2: The background seismicity $\nu(t)$ estimated for the three models and $\tau = 1$ day: the green horizontal dotted line is the constant value of ν for model I; the red dashed line represents $\nu(t)$ for model II [*Hainzl and Ogata*, 2005]; the blue solid line is $\nu(t)$ for model III.

In Figure 1.2 we report ν estimated for the three models and $\tau = 1$ day. This result, together with the estimated errors for ν, indicate that the only two significant peaks in Figure 1.2 are the ones occurred at the beginning of the sequence and about two weeks later; the other fluctuations are comparable with the errors associated to ν. The plots of Figure 1.2 highlight two important issues: first, the evolution of $\nu(t)$ is not a simple proxy for the time evolution of the seismic rate (see Figures 1.1 and 1.2). For instance, the difference is substantial for the end of the swarm where large shocks induced a high number of events, but the background $\nu(t)$ is low. Second, the evolution of the background for model II and III has important differences. This evidence, together with the AIC results, suggest that a model with only background varying with time [*Hainzl and Ogata*, 2005] may not be able to reproduce correctly the time evolution of the source process energy.

Model III implies that the parameters $\nu(t)$ and $p(t)$ underwent significant time vari-

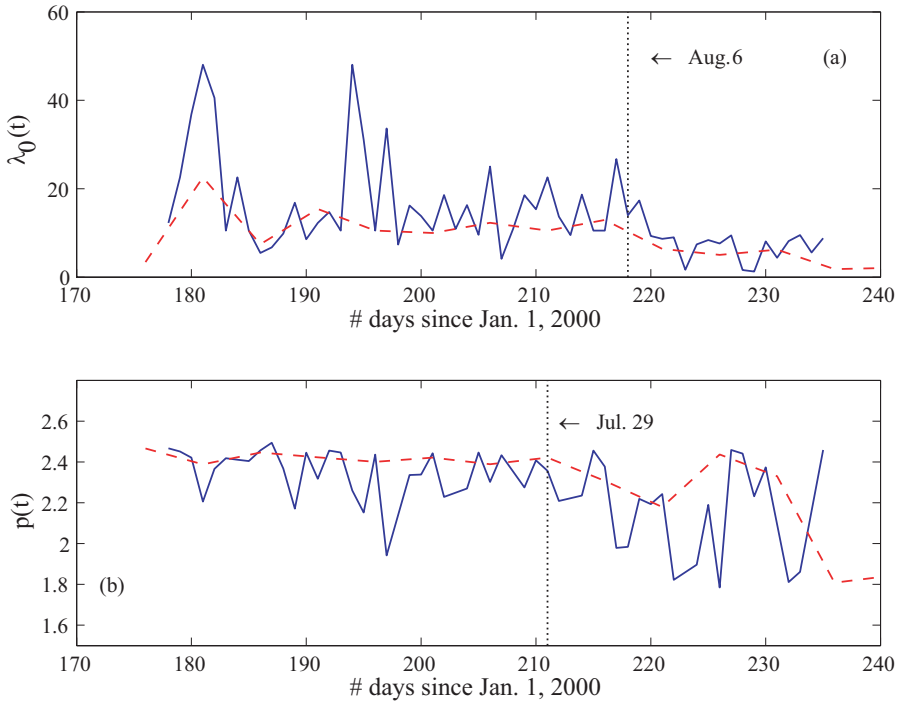

Figure 1.3: (a) $\nu(t)$ of model III for $\tau = 1$ day (blue solid line) and $\tau = 5$ days (red dashed line); the vertical dotted line indicates the significant (significance level<0.01) change point found. (b) the same as for (a), but relative to the parameter $p(t)$

ations. In Figure 1.3 we report the time evolution of $\nu(t)$ and $p(t)$, and the results of the change point analysis [*Mulargia and Tinti*, 1985; *Mulargia et al.*, 1987] on these sequences. Remarkably, both time series show comparable trends, with the most significant change point (significance level < 0.01) at nearly the same time. In particular, $\nu(t)$ has a change point at August 6, where there was a significant decrease of the background activity (the average drops from 17 ± 2 to 7 ± 1 events per day); as regards $p(t)$, it experienced a significant change point at July 29, when a larger spreading and a concomitant decrease of the average (from 2.35 ± 0.02 to 2.16 ± 0.05) was observed. The results are stable for τ of 1 and 5 days.

The results of stochastic modeling yield important clues to interpret the evolution of a generic seismic swarm. From a pure phenomenological point of view, the coherent variations found for $\nu(t)$ and $p(t)$ suggest that possible phase transitions of the system can be detected by monitoring simultaneously the time evolution of the parameters of the model. In this respect, we argue that the kind of transition may

provide some hints on the nature of the process; for instance, strongly fluctuating coherent parameters and/or abrupt changes of the parameters in short time intervals may be more likely due to magma/fluids intrusions rather than to tectonic processes.

A more detailed interpretation of the results in terms of the physics of the process requires the definition of the physical meaning associated to the parameters ν and p. As mentioned before, we assume that a time variation of the background activity reflects changes in the energy of the source process [see, i.e., *Hainzl and Ogata*, 2005], and that modifications of $p(t)$ indicate fluctuation of the average temperature in the system. Under this perspective, the analysis of the Izu Islands seismic swarm suggests that the fluid/magmatic activity lasted until the end of July and occurred in two main outbursts that may represent different episodes of fluid intrusion. Since the beginning of August, the source energy as well as the average temperature of the system diminish rather suddenly, and the rate of seismicity becomes mainly governed by mainshock-aftershocks interaction.

This scenario highlights a more irregular time evolution of the source process compared to the model proposed by *Toda et al.* [2002] where the evolution of the swarm was due to a single vertical dyke that propagate to its full length in the first week, and then opened continuously for seven weeks. Conversely our results are compatible with the ones found by *Ozawa et al.* [2004] that inferred a more complex time evolution of the Izu seismic swarm, with two main intrusions at the beginning of the swarm and at mid of July, and a sinking of the intrusion after the beginning of August. Note that, our interpretation implies that the $p(t)$ evolution is due to variations of the system temperature instead of stressing rate as suggested by *Toda et al.* [2002]. Even though our model cannot discriminate between these two hypotheses, we note that the stressing rate hypothesis was supported only by checking the apparent duration of the aftersock sequences (that is ultimately related to the background level), while we investigate on the parameter p whose time variations are usually considered more related to the temperature of the system [*Mogi*, 1967; *Kisslinger and Jones*, 1991].

In order to substantiate our interpretation of the temporal evolution of Izu Islands seismic swarm, we plot an independent observable, that is the time evolution of the hypocenters spreading of the background seismicity for τ equal to 1 and 5 days.

For such a comparison to be meaningful, we assume that magma intrusions are usually characterized by strongly localized seismicity, while background activity has a larger spreading being due to the average increase of stress in the whole area [*Toda et al.*, 2002]. Spreading of background seismicity is estimated by computing the average of the relative spatial distances for all pairs of events belonging to the Izu Islands seismic swarm de-clustered catalog because we want to highlight the spatial distribution of earthquakes related to the source process rather than the spatial distribution of their aftershocks.

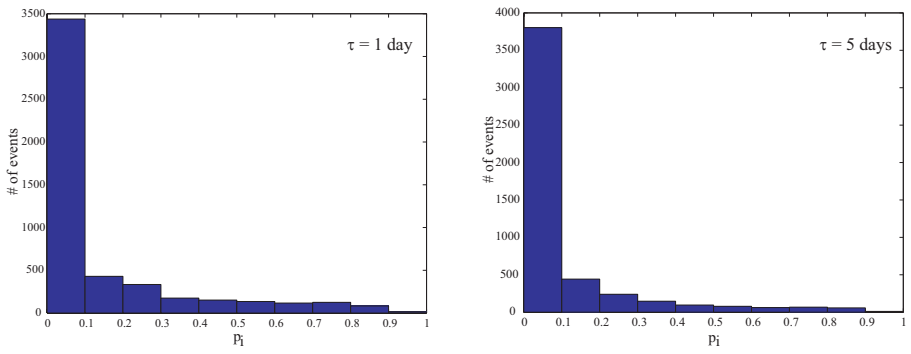

Figure 1.4: Histogram of probability p_i (probability of belonging to background seismicity) for events of Izu swarm, computed by model III for $\tau = 1$ and $\tau = 5$ days.

The de-clustering is obtained by applying the random procedure proposed by *Zhuang et al.* [2002] based on ETAS modeling, with parameters estimated for model III. The probability p_i that an event i belongs to the background activity is calculated as the ratio between the background rate $\nu(t_i)$ and the total occurrence rate $\lambda(t_i/\mathcal{H}_{t_i})$, both computed by model III. We de-cluster the catalog by selecting all the events with $p_i > 0.5$. The percentage of background events compared to the total number of earthquakes is about 10%(see Figure1.4). Remarkably, also in the hypocenters spreading sequence of the de-clustered catalog, we find the main significant change point (significance level < 0.01) at July 26 (Figure 1.5).

The main feature discernible in Figure 1.5 is that before the change point the hypocenters of the background seismicity are more clustered (i.e., they have a lower average distance), with a slight tendency toward an increase of spreading as a function of time. Notably, the lower spreading is linked to the beginning of the swarm, where the highest peaks of the background (see Figure 1.3a) suggest the highest

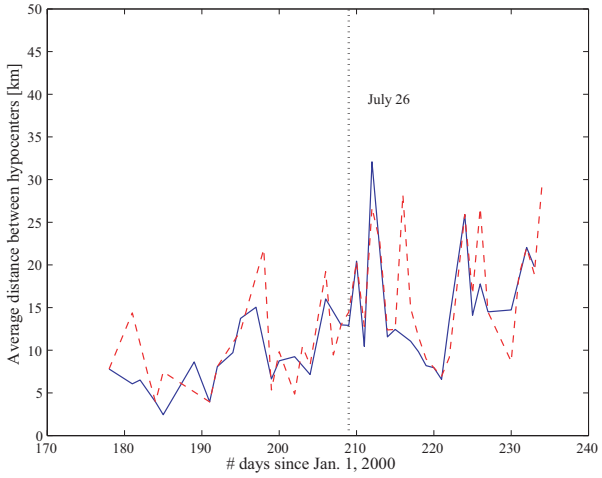

Figure 1.5: Spreading of the hypocenters as a function of time for the background seismicity estimated by model III. The blue solid line and the red dashed line are relative to τ equal to 1 and 5 days, respectively. The vertical dotted line represents the significant (significance level < 0.01) change point found.

magma/fluid activity. This evidence supports the hypothesis of a magma/fluids driving process for this time period. After the change point, the hypocenters spreading becomes larger when the energy and average temperature of the system diminish, suggesting that the main process for earthquake occurrence is the stress increase in the whole area. The change points found for $\nu(t)$, $p(t)$, and the hypocenter spreading all occur in a time range of about 10 days. This range can be due to different factors, such as a not instantaneous physical change in the source process, and/or to a limited resolution of the method of analysis.

1.4 The case of the 1983-1984 Phlegraean Fields seismic swarm

1.4.1 Geophysical setting of the area and the 1983-1984 seismic swarm

"Phlegrean Fields" is an active volcanic region, covering an area of about 400 km^2, located west of the city of Naples and centered on Pozzuoli town. Geological data

show that, in this zone, a large eruption, occurred at least 35,000 yr ago, caused a collapse of the area with the formation of a caldera [*Rosi et al.*, 1983]. The many eruptions occurred in the last 11,000 yr caused the opening of at least 22 volcanic vents. The last of these eruption, the only historically recorded, occurred in 1538. This was accompanied by a pronounced ground uplift and formed a small volcano (Mt. Nuovo) 140m high. After this event, the area was affected by alternation of phases of subsidence and uprising [*De Natale and Zollo*, 1986]. During the last 40 years, this area has been the site of two periods of unrest (that is a multitude of anomalous phenomena indicative of possible eruptive reactivation of a dormant volcano) indicated by strong uplift, intense earthquake swarms, increased fumarolic output and changed thermal-fluid chemistry. These are related to a typical phenomenon, called *bradyseism*, characterizing many caldera around the world, caused by a secular ground deformation that generally produces a very low seismicity. In the first unrest period, from 1970 to 1972, a slight subsidence occurred and about three thousand shallow, small events were recorded by local network. Here we analyze seismicity occurred in the second unrest period, from 1982 to 1984. Beginning from mid-1982 the ground started to rise at a very high rate (2mm per day on average) and an anomalous increase of the seismic activity was observed after few months, from the beginning of 1983 until December 1984 [*Del Pezzo et al.*, 1984]. The anomalous ground deformation phenomenon (about 1.8m of total uplift) was accompanied by a high seismic activity with more than 15,000 small magnitude earthquakes ($0<$Ml<4).

The initial interpretation on the 1982-1984 Phlegrean Caldera unrest, during the crisis, suggested the possible existence of a pressure source at 3km depth and identified in all conditions the typical precursors of a potential eruptive crisis. Notwithstanding the evidence of a crisis in progress, no actual alert for an impeding eruption was issued. In fact, no clear signs of magma rise, as an upward migration of earthquake hypocenters or relevant gravity changes, were recognized. After the crisis many papers dealt with the 1982-1984 unrest and provided models for the observed physical and chemical changes, but up till now there is no still a clear and unanimous explanation for the timing of this period of uplift. The main source of the crisis are identified both in internal processes, such as magma injection or fluid expansion, as in an external process, such as large regional earthquake or regional subsidence.

In particular, the main matter of debate concerns the involvement of magma motion. Many geophysicists emphasize the difficulty to explaining the observed uplift by pressure buildup within the residual magma chamber and seem inclined to attribute the uplift to different processes, as fluid expansion [*Bonafede*, 1990, 1991; *De Natale et al.*, 1991a]. This issue already rose immediately after the crisis. *Martini* [1989] raised doubts that increased pressure within a shallow magmatic body could be the source for large vertical movement at Phlegrean Fields. As evidence, he cited the absence of a significant magmatic component to gases emitted within the caldera. This result was used to argue against expansion of a magma body as the cause for recent seismicity and uplift and to ascribed all changes in gas concentrations to changes in a hydrothermal system. This last conclusion was also put forward later, with similar arguments, by *Todesco et al.* [1988]. An alternative explanation for the earthquake activity during the 1982-1984 unrest was provided by *De Natale et al.* [1995], which pointed out a directed relation with the ground motion deformation. They support the hypothesis in which the seismic activity during the two unrest episodes at Phlegrean Fields occurred along weakness zones of the ring fracture system and was generated by the local stress field associated with the ground deformation. As regard the fundamental triggering mechanism of the unrest, *De Natale et al.* [1991a] mention that the November 1980 Irpinia M6.9 earthquake, whose epicentral area was over 100km distant from Phlegrean Caldera, could have generated sufficient regional stress to affect the area involved by the crisis, possibly creating new fractures in the proximity of the magma chamber. Through these, thermal energy would have been transferred by fluid convection to shallower aquifer, thereby creating fluid overpressure and the uplift. An argument adduced to contradict this assertion is that no regional or local high-energy earthquake occurred before the 1970-1972 unrest, which had the same dynamics source and apparent mechanism as the 1982-1984 crisis.

Also if data processing after the crisis led the majority of geophysicists and geochemists to favor a model in which the unrest did not due to any direct involvement of the magma, a triggering mechanism related to overpressure in the magma chamber is not totally ruled out. The limited extent of the deformed area, that requires a source depth of less than 2-3 km, and the inferred minimum depth of the magma chamber of 3.5-4 km, led to *De Natale et al.* [1991b] to hypothesize that

the source of the pressure was the heating of shallow aquifers by an increasing heat flow from magma chamber. *Allard et al.* [1991] identify a clear magmatic character in composition of Solfatara (a hydrothermal system inside the Phlegrean Fields) gases, inferring a magmatic contribution to the Phlegrean Fields fumaroles. An further evidence for a probably overpressure within the chamber during the crisis was advanced by *De Natale et al.* [1993]. They observed that the majority of 1982-1984 earthquakes hypocenters extend down to the proximity of the magma-chamber depth, along fractures of the rim of the inner part of caldera, corresponding to the zone involved by the maximum uplift and by anomalous fumarolic degassing. Also *Dvorak and Berrino* [1991] related the episodes of rapid uplift and shallow earthquakes swarms of the 1982-1984 unrest to the growth of a resurgent dome, caused by the intrusion of a hotter mafic magma into a body of higher silicic content that forms a zoned magma chamber. In summary all these last instances could demonstrate that the Phlegrean Caldera magma chamber was involved in the most recent unrest and that the 1970-1972 and 1982-1984 crises could represent steps in a prolonged process of volcanic reactivation, like that which probably preceded the 1538 Mt. Nuovo eruption [*Barberi and Carapezza*; 1996].

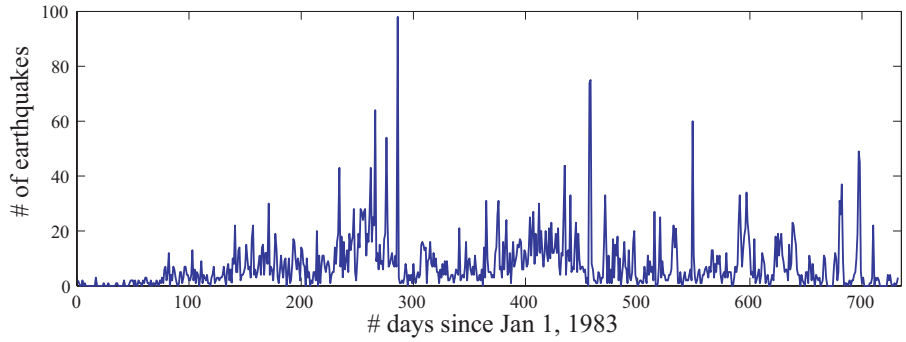

Figure 1.6: Number of earthquakes per day at the Phlegraean Fields for the years 1983-1984.

To interpret the driving mechanism of 1982-1984 unrest, we analyze 2 yrs of seismic activity, during the period January 1983-December 1984. The data set collects more than 5,600 events with magnitudes ranging between 0.8 and 4.0. In Figure 1.6 we report the time history of earthquake activity given in number of events per day. The largest two earthquakes, with magnitude M4.0, occurred on October 4 1983

and April 1 1984.

1.4.2 Discussion of results

To analyze the Phlegraean Fields 1983-1984 seismic swarm, we consider the same three stochastic models, used to study the Izu Islands activity.

Parameter	**value**
ν	0.6 ± 0.1 day^{-1}
k	0.05 ± 0.002 day^{p-1}
p	1.07 ± 0.01
c	0.0004 ± 0.0001 day
α	0.4 ± 0.1

Table 1.3: Estimated ETAS Parameters of the 1983-1984 Phlegrean Fields Swarm

In Table 1.3 we report the parameters (with relative errors) estimated for the stationary ETAS model. Respect to the Izu area the smaller p-value indicates a slower decaying aftershocks activity that could suggests a lower temperature of a system, in which magma motion is not a driving process. The low value of α is in agreement with previous findings for earthquake swarm activity. In contrast with mainshock-aftershocks sequences which are mainly driven by stress triggering and are characterized by higher values of α, the recognized low α-values for swarms show a negligible dependence of seismic rate evolution from magnitude of previous events. This result points out the presence of different external driving mechanisms for this type of activity.

To fit nonstationary models, in which we assume only the background ν (model II) or both ν and p (model III) varying with time, the length of τ is set to 10 and 30 days. In Figure 1.7 we report ν estimated for the three models and $\tau = 10$ days. From this analysis we can infer two main results. The plots of Figures 1.6 and 1.7 highlights a general agreement between the time evolution of $\nu(t)$ and that of total seismic rate. Only at the end of the swarm there is a rapid increase of background that does not correspond to as much evident increase of seismic rate. Second, the evolution of the background for model II and III has not important differences, suggesting

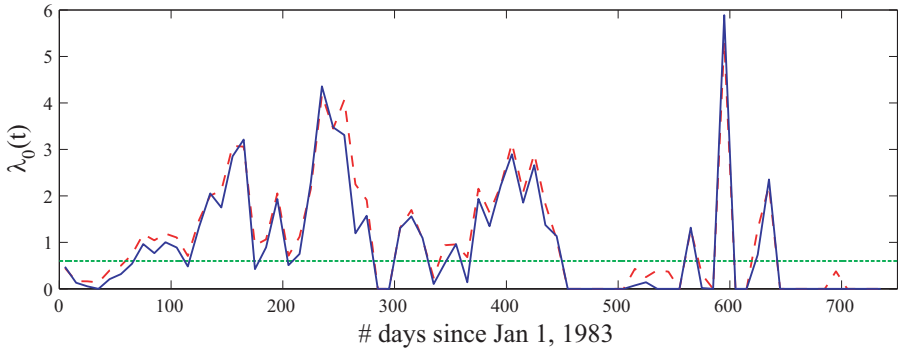

Figure 1.7: The background seismicity $\nu(t)$ estimated for the three models and for $\tau = 10$ days: the green horizontal dotted line is the constant value of ν for model I; the red dashed line represents $\nu(t)$ for model II; the blue solid line is $\nu(t)$ for model III.

that including time variation of p-value in the model does not noticeably improve description of data. This result is partly confirmed by AIC calculations, reported in Table1.4. These indicate that model III is the best one to describe the data regardless the value of τ, but the moderate difference between AIC values raises further doubts on significance of time variation of p-value.

Model	$\boldsymbol{\tau}$	**AIC**
I		-21168
II	10 days	-21276
II	30 days	-21308
III	10 days	-21296
III	30 days	-21344

Table 1.4: Values of AIC of the Phlegrean Caldera Swarm for Model I,II and III.

In order to corroborate our interpretation, we carry out the change point analysis [*Mulargia and Tinti*, 1985; *Mulargia et al.*, 1987] on $\nu(t)$ and $p(t)$ (see Figure 1.8). Remarkably, $\nu(t)$ time series shows two significant change points (significance level < 0.01) at April 30 1983 and March 25 1984. These distinguish three subintervals in two years of activity collected in our dataset. The first represents the initiation phase in which there is a gradual increasing of seismic activity, whereas the last corresponds to final transient phase, in which, after the occurrence of M4.0 event

at April 1 1984, the seismicity returns to normal activity. The central subinterval represents a stationary period containing the most and more energetic part of seismicity.

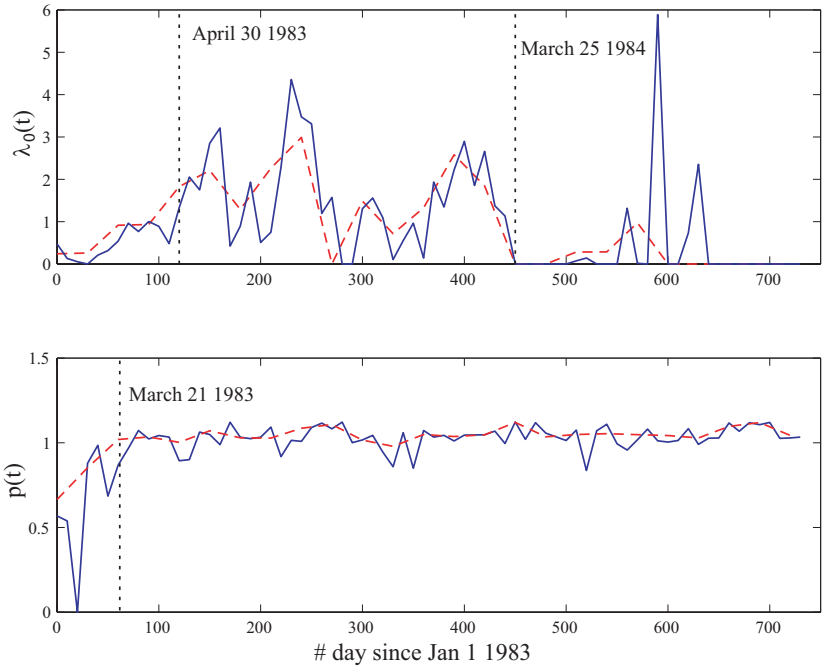

Figure 1.8: (a) $\nu(t)$ of model III for $\tau = 10$ day (blue solid line) and $\tau = 30$ days (red dashed line); the vertical dotted line indicates the significant (significance level < 0.01) change point found. (b) the same as for (a), but relative to the parameter $p(t)$

As regards $p(t)$, it experienced no significant change point, but one at March 21 1983; this corresponds to end of the initial transient phase, after which p-value becomes almost time-invariant. The substantial stability of $p(t)$ may be closely linked to the underlying physical processes. The lack of significative variations of p-value, being around a rather low value ($1.0 - 1.1$), suggests that system temperature is almost steady and that no magma motion was involved. The results are stable for τ of 10 and 30 days.

Applying the *Zhuang et al.* [2002] declustering procedure with parameters estimated by model III we find that the background activity, represented by events with $p_i > 0.5$, is a very low percentage (about 2%) of total seismicity (see Figure 1.9), especially in the second year of activity.

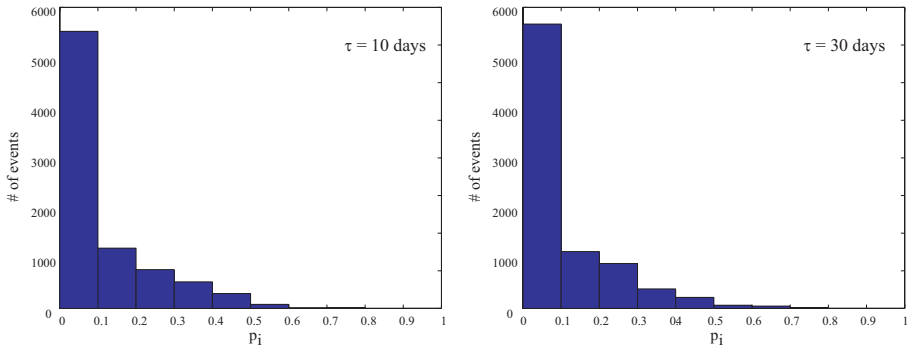

Figure 1.9: Histogram of probability p_i (probability of belonging to background seismicity) for events of Phlegraean Fields swarm, computed by model III for $\tau = 10$ and $\tau = 30$ days.

This result points out a decisive rule of triggering effect in time evolution of seismic swarm, particularly after the second change point recognized in ν sequence. These observation justify the suitability of ETAS model to describe the temporal clustering of seismic swarm under study. It is able to reproduce the seismicity induced by external fluid intrusion as well as the activity triggered by stress transfers and to separate the external fluid signal by the stress triggering effect in observed data. In the case for swarm activity in the Phlegrean Fields region we find that the external force, responsible of most seismic activity in the initial phase, is not related to any magma motion nor to processes causing high temperature and is probably due to the fluid intrusion. The fluid signal, represented by sequence ν, persists in the whole first year of activity and then decreases again the time, apart from a short peak in the ending of the swarm. From March-April 1983 the stress triggering, identified with Omori-like aftershock sequences, dominates the whole activity, especially in the second year of activity. The substantial stationarity of p-value highlights none variation of temperature of the system under study.

1.5 Final remarks

We have analyzed the 2000 Izu Islands and the 1983-1984 Phlegrean Fields seismic swarms through a stochastic nonstationary ETAS modeling. We have found that a nonstationary ETAS model with background activity varying through time describes the observations better than a stationary model. Remarkably, the fluctu-

ations of the background activity are coherent and in agreement with the temporal earthquake distribution. This correlation gives important clues on the nature of the source process of the seismic swarm and on its temporal evolution. As regards the Izu Islands seismic swarm evolution, we have found evidence of a magma source that evolves through outbursts of activity superimposed to lower frequency fluctuations. The source energy diminishes before the end of the swarm that is lastly dominated by mainshock-aftershocks sequences. The analysis of Phlegrean Fields swarm seems to rule out any involvement of magma chamber. Our results show evidence for a fluid intrusion as initial forcing process, mainly responsible of the seismicity in the first 400 days, and for the stress triggering as the dominant process of the remaining activity.

The main findings of our study seems to show a relationship between the temporal variation of p-value of Omori law and the physical processes driving seismic swarms. The values of $p(t)$ computed for Izu Islands swarm and the coherent variations of $\nu(t)$ and $p(t)$, supports hypothesis of phase transition of the system, from a period in which the activity is mainly driven by magma motion to a phase ruled by stress triggering. The low and stationary value of p for Phlegrean Fields swarm highlights the steadiness of the temperature of the system. Then, it seems that the presence of significative p-value variations could be a very important signal for identifying the presence of magma motion. This suggestion should be further verified by future analysis of seismic crises in other volcanic areas.

These results indicate that stochastic modeling of seismic swarm occurrence may yield important insights in constraining the physics of the source process (i.e., magma/fluids or tectonics) and in characterizing its temporal evolution. From a practical point of view, stochastic modeling may be used to develop a new tool for tracking in almost real time the evolution of a magma/fluids source.

1.6 Appendix 1A: A nonparametric method for change points detecting in a time series

The search for change points in a time series [*Mulargia and Tinti,* 1985; *Mulargia et al.,* 1987] is based on Kolmogorov-Smirnov two sample nonparametric statistics.

This is defined by

$$J3 = \left(\frac{mn}{d}\right) \max_{-\infty<x<+\infty} |G_n(x) - F_m(x)| \tag{1.4}$$

where m and n are the number of observations before and after the change point, respectively, d is the maximum common divisor of m and n, $F_m(x)$ and $G_n(x)$ are the empirical distribution functions of the two subsets X and Y of observations (separated by the possible change point). The $J3$ statistic is related to significance level α of a test of which the null hypothesis H_0 is that the two subsets of observations have the same distribution, i.e.

$$H_0 : I\!P(X < x) = I\!P(Y < x), \quad -\infty < x < +\infty. \tag{1.5}$$

For $m, n \geq 30$ the critical values $C(m, n)$ are

$$C(m, n) = J3\frac{d}{[mn(m + n)]^{0.5}}. \tag{1.6}$$

The probability distribution of $C(m, n)$ is well approximated by the formula:

$$I\!P(C(m, n) < c) = \sum_{i=-\infty}^{+\infty} (-1)^j e^{-2j^2c^2} \quad c > 0. \tag{1.7}$$

This is tabulated in some textbooks as well as the distribution for small n and m (see [*Mulargia and Tinti*, 1985] and references therein).

Starting by hypothesis that a single change point is present in a given set of N data, the algorithm computes the vector $\{C^i(m, n), i = 1, ..., N\}$, obtained assuming that the change point corresponds to datum i. The most likely position j for the change point corresponds to maximum value of critical values vector:

$$j : \quad \max_{i=1,...,N} C^i(m, n) = C^j(m, n). \tag{1.8}$$

This is accepted as a real change point if it is lower of the critical value for the prefixed significance level α. If a change point is identified, a new analysis is carried out for each of two subsets obtained. This recursive algorithm ends if none further change point is found or if the subsets became to small.

Chapter 2

Nonstationary in a tectonic zone: the 1997-1998 Umbria-Marche (Italy) sequence

2.1 Introduction

The most important physical process responsible of short-term clustering of earthquakes consists of stress variations caused by earthquake dislocations [*King and Cocco*, 2001]. The activity rate $\lambda(t)$ of aftershock sequences generally decays according to the modified Omori law $\lambda(t) = k(t + c)^{-p}$, where t is the elapsed time since the mainshock and k, c and p are constants [*Utsu et al.*, 1995]. Among point process models [*Daley and Vere-Jones*, 2003], used to represent statistical features of temporal patterns of shallow sequences, the Epidemic-Type Aftershock Sequences (ETAS) model [*Ogata*, 1988; 1998] seems to best represents the main features of seismicity driven by coseismic stress changes. This describes the triggered seismicity time evolution in agreement with the modified Omori law and takes into account possibility of production of secondary aftershocks.

It is well-known that the coseismic stress field can be modified by further processes as viscoelastic relaxation [*Pollitz*, 1992; *Piersanti et al.*, 1995] or fluid flow [*Nur and Booker*, 1972]. Whereas the viscoelastic relaxation can be decisive on long time scale, fluid flow may play a determinant role in triggering seismicity at a time scale

compatible with aftershocks occurrence [*Nur and Booker*, 1972]. The seimicity driven by fluid intrusion has different characteristics respect to mainshock-aftershocks sequences. It is not characterized by a dominant earthquake and the temporal patterns of near-equal magnitude events appear to have a very rapid decay. In Chapter 1 it has been proved that the ETAS model is an appropriate tool also to extract fluid signal from short-term seismicity patterns (see also [*Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006]). The time description of fluid intrusion is turned into identification of significative changes of parameters more directly linked to underlying driving physical processes.

In this study we deal with this issue to explain the complex Umbria-Marche (Central Italy) seismic sequence occurred in 1997-1998. The presence of fluids in this area has been recognized by previous studies [*Chiodini et al.*, 2000]. Moreover the fluid flow, probably promoted by elastic stress changes of largest magnitude shocks, has been adduced as the main cause of the evident migration of activity of the 1997-1998 sequence [*Antonioli et al.*, 2005]. Here, by applying the method outlined by [*Hainzl and Ogata*, 2005], we demonstrate that the identified variations of parameters of ETAS model are consistent with hypothesis of a fluid flow promotion of seismicity.

2.2 Data set

The data set used in this study is extracted from the catalog "Catalogo della Sismicitá Italiana" (CSI) 1981−2002 [*Castello et al.* 2005]. This collects shallow seismicity occurred from 1 January 1981 to 31 December 2002 in Italy. Specifically we consider the earthquakes occurred in the region affected by 1997-1998 Umbria-Marche seismic sequence [12°-13.5°W,42°-44°N] with magnitude $Ml \geq 2.5$ (1511 events). The our dataset includes 10 events with magnitude $5.0 \leq Ml \leq 6.0$, of which 8 belongs to 1997-1998 Umbria-Marche sequence. This consists of thousands of events that in some tens of days activated a NW-trending prevalently normal fault system. A comprehensive synthesis of this seismic sequence can be found in [*Chiaraluce et al.*, 2003].

The sequence is characterized by a clear migration phenomenon (Figure 2.1) from

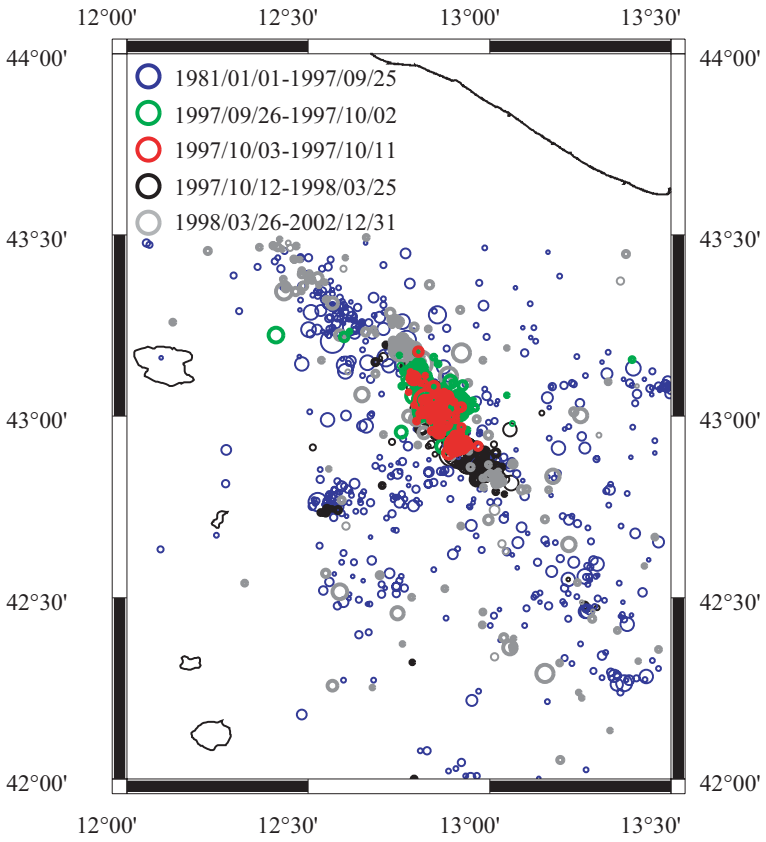

Figure 2.1: Map showing the seismicity occurred from January 1 1981 to December 31 2002 in region affected by the 1997-1998 Umbria-Marche sequence. The events are color-coded by time intervals defined by the occurrence of the mainshocks of the sequence

NW to SE with progressive activation of adjacent fault segments. The two largest events of the sequence (Ml=5.6 and Ml=5.8) struck the Colfiorito area on September 26th (within some hours and few km of distance from each other). These was followed, at the beginning of October, by two other events with comparable magnitude (Ml=5.0 and Ml=5.4). Then seismicity began to migrate towards SE, where two other main shocks with magnitude larger than 5.0 occurred at about the half of October, in Sellano region (October 12 Ml=5.1, October 14 Ml=5.5). The last two main events (Ml=5.4 and Ml=5.3) struck after some months, on March 1998, the region near Gualdo Tadino, north of Colfiorito.

In Figure 2.2 we show the mean number of events per year together with a time/magnitude plot. The exceptional (respect to previous 15 years) amount of released en-

ergy during the sequence under study suggests a possible nonstationary trend, sign of involvement of other physical processes besides the coseismic stress triggering.

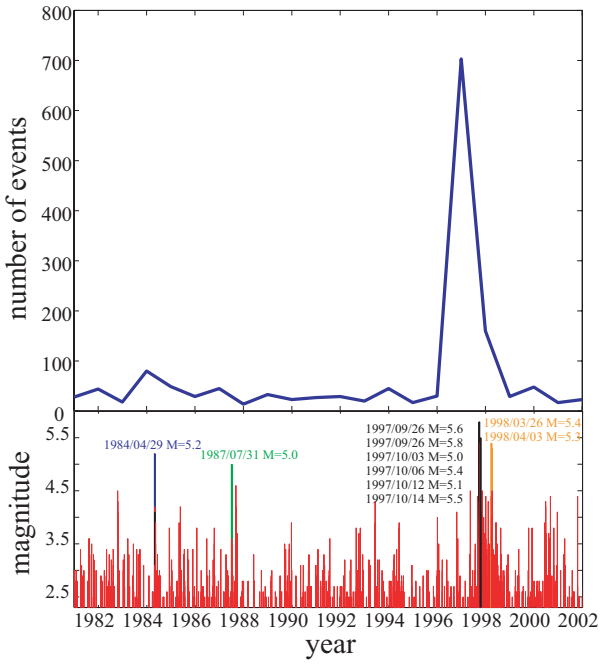

Figure 2.2: Histogram (top) and time/magnitude plot (bottom) of events occurred from January 1 1981 to December 31 2002 in region affected by the 1997-1998 Umbria-Marche sequence

2.3 The Spatio-Temporal Epidemic Type Aftershock Sequences (ETAS) Model

The ETAS Model is a stochastic point process, based on the well-known modified Omori law [*Omori*, 1894; *Utsu*, 1961], that models the coseismic stress triggered activity responsible of aftershocks occurrence. Its formulation follows from the observation that aftershock activity is not always predicted by a single modified Omori function and that it can include conspicuous secondary aftershock production. Therefore, this model assumes that every aftershock can trigger its further aftershocks and that the occurrence rate at time t is given by a superposition of the modified Omori law functions, shifted in time. To first time-magnitude formulation, proposed by *Ogata* [1988], many others time-magnitude-space versions followed,

mostly based on empirical studies of past seismicity [*Ogata*, 1998; 2006; *Console et al.*, 2003]. All of these describes seismicity as the sum of a time-independent activity, due to not-inducted earthquakes ("background") and of the triggering seismicity of all preceding events. The last improved extension proposed by *Ogata and Zhuang* (2006) defines the rate of aftershocks at a time t and at a location (x, y), induced by i-th event occurred at a time $t_i < t$ and with magnitude M_i and epicenter (x_i, y_i), is given by

$$\lambda_i^{ind}(t, x, y) = \frac{K}{(t - t_i + c)^p} e^{\alpha(M_i - M_{min})} f_{ind}^i(x, y) \tag{2.1}$$

where M_{min} is the minimum magnitude of the catalog and $f_{ind}^i(x, y)$ is the probability density function (PDF) of occurrence for a triggered event. The parameter K measures the productivity of the aftershocks activity; α estimates the triggering capability of events with magnitude m_i; the parameter c measures the incompleteness of the catalog in the earliest part of each cluster, caused by the decrease in aftershocks detectability after a strong event [*Kagan*, 2004]; the parameter p controls the temporal decay of triggered events.

Both physical investigations [*Dieterich*, 1994; *Shaw*, 1993; *Hill et al.*, 1993] and statistical studies [*Ogata*, 1998; *Kagan and Jackson*, 2000; *Console et al.*, 2003; *Helmstetter et al.*, 2006] show that the stress induced by an event decreases with the distance r from its epicenter, by an inverse power law. In particular, some physical models predict that static stress changes, caused by a point source or a circular crack in an elastic medium, decrease rapidly with increasing r, as $1/r^3$ [*Hill et al.*, 1993]. These finding suggest choice of a power law function for f_{ind}^i.

We impose an isotropic influence of an inducing event on area surrounding its location and relate probability of occurrence at a location (x, y) to distance $r_{(x,y)(x_i,y_i)}$ from epicenter (x_i, y_i) of triggering earthquake, up to a maximum distance R_{max}. By these assumptions $f_{ind}^i(x, y)$ is given by:

$$f_{ind}^i(x, y) = \frac{(q-1)d^{2(q-1)}}{\pi} \left[1 - \left(1 + \frac{R_{max}^2}{d^2}\right)^{-q+1}\right]^{-1} \frac{1}{(r_{[(x,y),(x_i,y_i)]}^2 + d^2 e^{\gamma(M_i - M_{min})})^q}. \tag{2.2}$$

This new version of the model is characterized by the introduction of the term

$e^{\gamma(M_i - M_{min})}$ that takes into account the correlation between the logarithm of the aftershocks area and the mainshock magnitude, proposed by *Utsu and Seki* [1955].

The total occurrence rate is the sum of the triggering rate of all preceding events and a time-independent rate $\mu(x, y)$, due to non-induced earthquakes ("background" activity). The stationary ETAS model ascribes as primary cause of seismicity a poissonian "impulse", of which the rate changes with seismogenic capability of the examined zone, but not with time. The sudden stress variations, caused by occurrence of these events, generate the aftershock activity, that is well described by the clustering features of equations (2.1) and (2.2). Therefore, the total space-time conditional intensity (i.e. the probability of an earthquake occurring in the infinitesimal space-time volume conditioned to all past history) $\lambda(t, x, y/\mathcal{H}_t)$ by equation:

$$\lambda(t, x, y/\mathcal{H}_t) = \nu u(x, y) + \sum_{t_i<t} \frac{K e^{\alpha(M_i - M_{min})}}{(t - t_i + c)^p} \frac{c_{d,q,\gamma}}{(r^2_{[(x,y),(x_i,y_i)]} + d^2 e^{\gamma(M_i - M_{min})})^q}. \tag{2.3}$$

where $\mathcal{H}_t = \{(t_i, x_i, y_i, M_i); t_i < t\}$ is the observation history up time t, ν is a positive-valued parameter that represents the poissonian rate of background events, $u(x, y)$ is the PDF of location of background events and $c_{d,q,\gamma} = \frac{(q-1)}{\pi}(d^2 e^{\gamma(M_i - M_{min})})^{q-1}$ is the normalization constant of the PDF for location of a triggered event.

The parameters $(\nu, k, c, p, \alpha, d, q, \gamma)$ of the model, for data in an interval time $[T_1, T_2]$ and in a region $\mathcal{R}$, can be estimated by maximizing the Log-Likelihood function [*Daley and Vere-Jones*, 2003], given by

$$\log L(\nu, K, c, p, \alpha, d, q) = \sum_{i=1}^{N} \log \lambda(t_i, x_i, y_i/\mathcal{H}_{t_i}) - \int_{T_1}^{T_2} \int_{\mathcal{R}} \lambda(t, x, y/\mathcal{H}_t) dt dx dy. \tag{2.4}$$

To estimate the parameters for our dataset, we apply the iteration algorithm developed by *Zhuang et al.* [2002]; this provides also an estimation of distribution for location of background events $u(x, y)$ by a suitable kernel method. Since some physical investigations show that static stress changes decrease with epicentral distance as r^{-3} [*Hill et al.*, 1993], we impose $q = 1.5$. This choice is also justified by the recognized trade-off between parameters q and d that may cause different pairs

of q and d values to provide almost the same likelihood of the model [*Kagan and Jackson*, 2000].

Parameter	**value**
ν	0.053 ± 0.003
K	0.027 ± 0.002
p	1.26 ± 0.02
c	0.04 ± 0.01
α	1.55 ± 0.05
d	0.8 ± 0.1
q	$\equiv 1.5$
γ	0.39 ± 0.05
Log-likelihood	-12001.5

Table 2.1: Estimated ETAS Parameters for the 1981-2002 Umbria-Marche seismicity

In Table 2.1 we report the values of estimated parameters together with the errors and the log-likelihood value. The p and α values are similar to previous findings for other tectonic areas [*Ogata*, 1992]. The scaling of aftershocks area against magnitude of triggering event is rather weak, also if it is not totally negligible. This result is probably due to epicentral errors, to limited range of magnitude recovered by catalog and to the strong anisotropy of the triggering effect (see Figure 2.1) that does not match with the isotropic power law chosen in our model.

The suitability of the ETAS model is tested by a *residual analysis* [*Ogata*, 1988]. The occurrence times t_i of earthquakes are transformed into new values τ_i by relation

$$\tau_i = \int_{T_{start}}^{t_i} dt \int_{\mathcal{R}} dxdy\lambda(t, x, y/\mathcal{H}_t) \tag{2.5}$$

where T_{start} is the start time of observation history. If the ETAS model well describes the temporal evolution of seismicity, the transformed data τ_i are expected to behave like a stationary Poisson process with the unit rate [*Ogata*, 1988].

Figure 2.3a shows the cumulative numbers of transformed times τ_i. By the Kolmogorov-Smirnov test [*Gibbons and Chakraborti*, 2003] (see Appendix 2A) we cannot

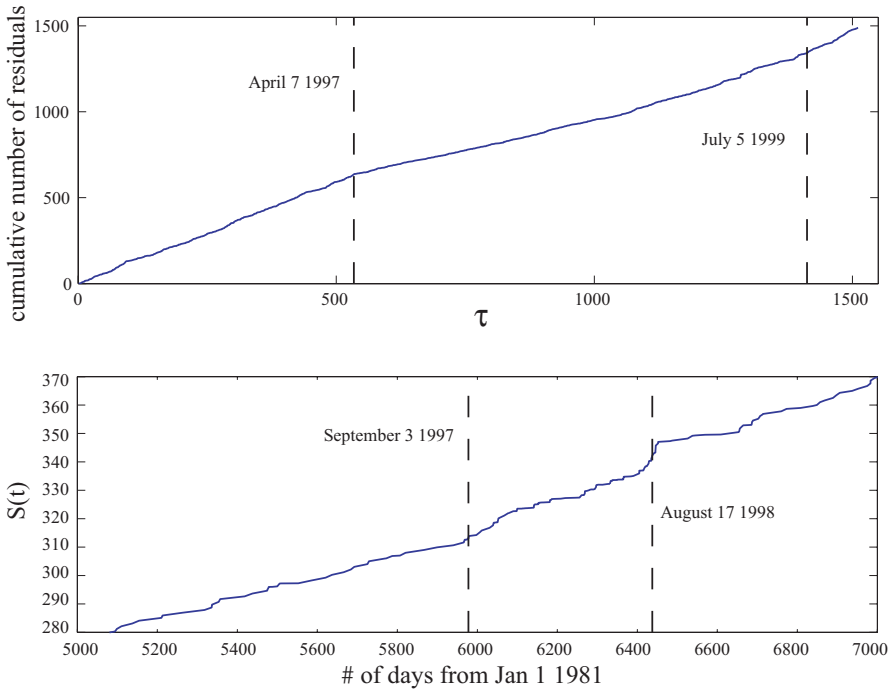

Figure 2.3: Cumulative number of transformed times τ_i (top) and the cumulative background seismicity (bottom) obtained by the estimated ETAS model. Vertical dotted lines mark change points identified by procedure proposed by *Mulargia and Tinti* [1985]

reject the null hypothesis that values $\Delta\tau_i = \tau_{i+1} - \tau_i$ are exponentially distributed (with mean equal to 1) at 95% confidence level, but the empirical distribution of $\Delta\tau_i$ shows evident change points. We identify them by the change point analysis algorithm developed by *Mulargia and Tinti* [1985]. We identify two significative change points at April 7 1997 and July 5 1999 (see Figure 2.3a); these mark off a time period, including the 1997-1998 Umbria-Marche sequence, in which seismicity is not well described by the estimated ETAS model. To understand the origin of these anomalies we investigate the suitability of the stationary hypothesis for the background seismicity. A direct way to discuss how the background rate changes with time is to calculate the cumulative background seismicity defined by

$$S(t) = \sum_{t_i<t} \phi_i \tag{2.6}$$

where ϕ_i is the background probability for the i-th event given by

$$\phi_i = \frac{\nu u(x_i, y_i)}{\lambda(t_i, x_i, y_i/\mathcal{H}_{t_i})} \tag{2.7}$$

[*Zhuang et al.*, 2005].

The change point procedure identifies a different trend of background rate during the period from September 3 1997 to August 17 1998 (Figure 2.3b).

All previous results make clear the exceptionality of the seismicity occurred during the Umbria-Marche sequence. Comparing the number of observed and expected (by estimated ETAS model) number of events (Figure 2.4) we see that the total number of observed events is larger than the theoretical values. Therefore the ETAS model, fitted on the whole seismicity of this zone, underestimates the real number of events for the 1997-1998 sequence, characterized by an explosive aftershock activity. To better understand the causes of this anomalous behaviour and to carry out more analyses, we select a subset of our catalog including the sequence. In order to identify objectively the start and the end of this interval time, without include subjectivity related to ETAS calculations, we look for change points in the sequence of the number of events per day of the whole dataset.

The two most significant (significance level < 0.01) change points identified correspond to May 3 1997 (start) and August 27 1998 (end). The selected interval time contains 850 events. The identified period of activity does not varies if the change point analysis is carry out computing the number of events in intervals of 12 hours.

We remark that all change point analyses carried out here identify the beginning of an anomalous behaviour of seismicity some months before the occurrence of the two first mainshocks of the sequence (September 26 1997).

2.4 Analysis of the 1997-1998 Umbria-Marche sequence

The ETAS model can be used in order to identify the background as well as the triggered component of seismicity, since the values of parameters are related to driving physical processes responsible of the seismic activity of a region. Specifically the background rate ν refers to activity forced by external physical processes as the

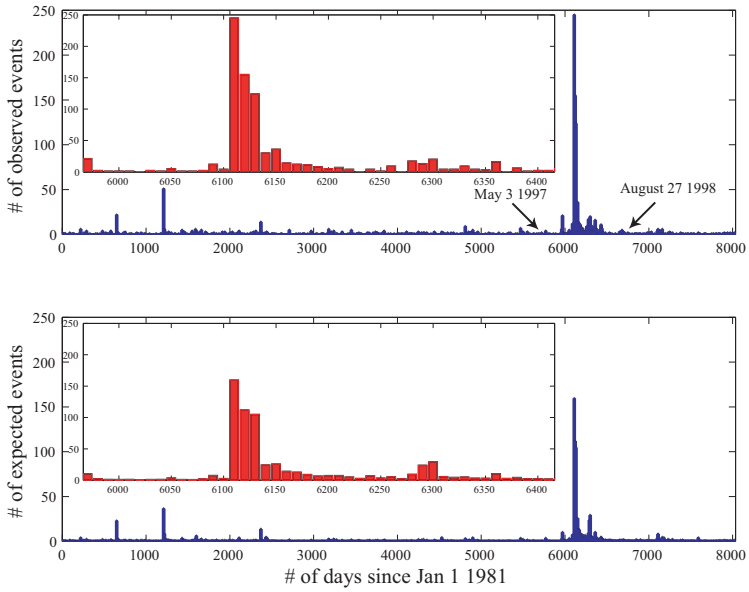

Figure 2.4: Histograms of the observed (top) and expected (bottom) number of events in Umbria-Marche region from January 1 1981 to December 31 2002. The duration of each interval time is of 10 days. The arrows mark the identified change points. Into the panel results relative to period bounded by change points are shown.

tectonic loading or fluids intrusion. The parameter p controls the temporal decay of triggered events. Its variation is related to heat flow, degree of structural heterogeneity in the fault zone, stress and crustal temperature [*Mogi*, 1962; *Kisslinger and Jones*, 1991; *Utsu*, 1995].

In order to recognize the underlying physical processes responsible of the Umbria-Marche sequence we look for possible time variations for ν and p parameters, that are the most clearly linked to physical mechanisms of the system. In particular we fit the ETAS model on the 850 events selected by the change point analysis, in a moving non-overlapping time window with size $\tau = 10g$. We estimate the joined variation of ν and p values and set all the other parameters to the values found for the whole sequence. These last are not significantly different by those estimated for the whole catalog (see Table 2.1).

The choice of the size of the time window τ is dictated by some requests: to follow details of evolution of 1997-1998 sequence and to have enough data in each interval time for the whole period under study (see Figure 2.4). The main goal of

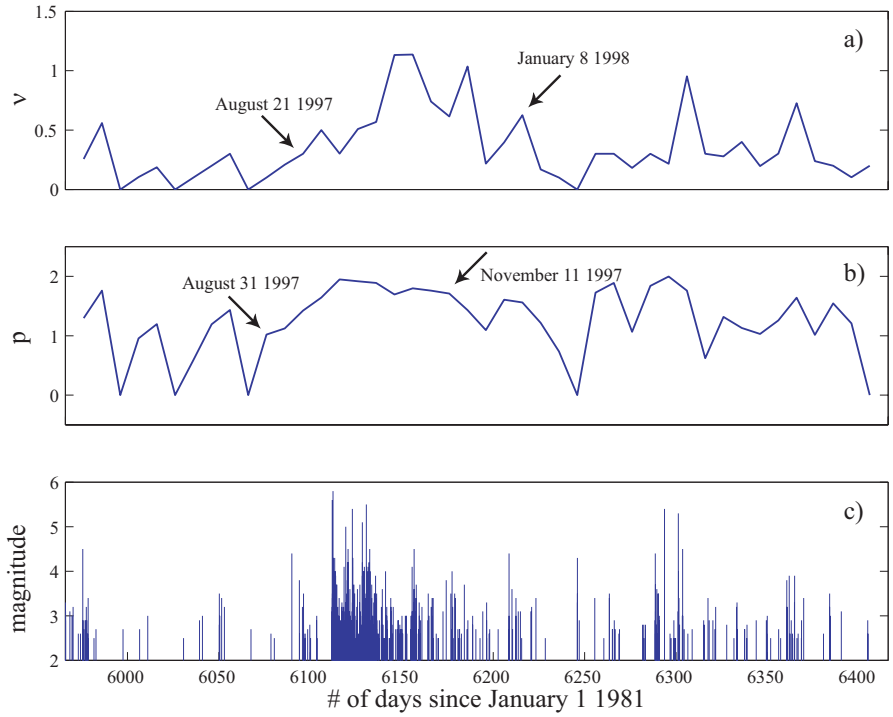

Figure 2.5: (a) $\nu(t)$ of model III for $\tau = 10$ days; the arrows indicates the significant (significance level$<$0.01) change points found, (b) the same as for (a), but relative to the parameter $p(t)$, (c) time/magnitude plot

our analysis is to verify if there is a coherent variation of the background rate and of p-value. A such evidence would be a clear sign that the high level of seismicity of the sequence is not only due to an increasing of triggering capability, but is also externally forced. This could be interpreted as an evidence of fluid intrusion.

As Figure 2.5 shows, we find a large systematic variation in the external forcing ν. The starting time of an increasing of the background rate is identified at August 21 1997; a second change point, at January 8 1998, marks the end of this ”anomalous” behaviour. A coherent variation of p-value is found between the end of August and November 1997, pointing out a variation of time-decaying of triggering effect in a fluid-saturated medium.

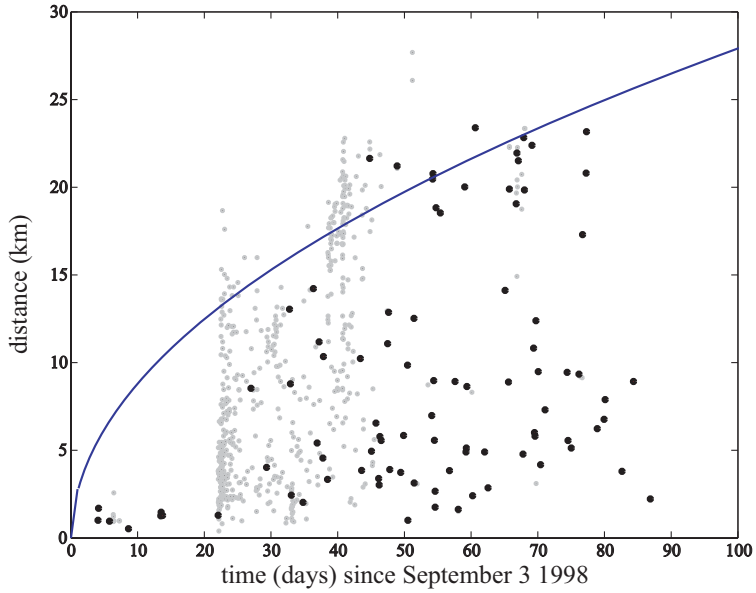

Figure 2.6: Spatio-temporal distribution of events for the Umbria-Marche sequence. The distances from the first event are plotted as a function of the occurrence times. The blue line refers to the theoretical position of the propagating pore pressure front in the case of a hydraulic diffusivity of $D = 0.90m^2/s$. Light and dark points refer to clustered and background events, respectively

In order to substantiate interpretation of our finding in Figure 2.6 we show the spatio-temporal distribution of the complete and declustered activity. The distance from the first event (M4.4 occurred at September 3 1997) is plotted as a function of the occurrence time. We compare the data points with the theoretical prediction for fluid induced earthquake activity . The distance r of the propagating pore pressure front from the point source is given by $r = \sqrt{4\pi Dt}$ [*Shapiro et al.*, 1997], where t is the time and D the hydraulic diffusivity. Once the pore pressure change reaches a point r at time t, it may trigger an event in that location at a certain time after the perturbation. Therefore, in a (t, r) plot, seismicity is expected to occupy the space beneath the theoretical curve. The fit of a such curve is rather poor for the complete catalog, but, in agreement with [*Antonioli et al.*, 2005], the declustered seismicity distribution matches with the pore-pressure triggering front envelope predicted by an isotropic diffusivity of $90m^2/s$.

2.5 Discussion and conclusive remarks

We discuss the hypothesis that one of driving processes of seismicity observed during the Umbria-Marche sequence is intrusion of fluids. The presence of fluid flow is checked by exploring the time dependence of some parameters of the ETAS model. The coherent variation of parameters examined here points out that the explosion of seismic activity of the 1997-1998 Umbria-Marche sequence has not be ascribed only to an increase of triggering effect, due to redistribution of the stress following main-shocks occurrence. The significative increase of background seismicity is a clear sign of a externally forced impulse, that can be interpreted in terms of fluid intrusion. The time period bounded by change points of $\nu(t)$ and $p(t)$ (see Figure 2.5) does not include the two strong events occurred at March 1998 in Gualdo Tadino region. This is an evidence of significance of our analysis. Even if these shocks have a magnitude comparable with previous mainshocks, they occurred in an area not involved by fluid flow; therefore the time evolution of aftershock production is coherent with parameters estimated by the whole catalog. Clearly it cannot be determined unquestionably whether the nonstationary background activity results from tectonics or pore pressure changes. However we belief that the recognized large variations on a short timescale are unlikely imputable to tectonic causes.

The procedure applied here seems to be a promising method to unveil fluid signals, even in the case of poor hypocenter information and for patterns of seismicity largely containing short-term triggered events.

2.6 Appendix 2A: The one–sample Kolmogorov–Smirnov test

The one–sample Kolmogorov–Smirnov test is used to test if a specified continuous function F_0 is the distribution function from which a random sample $x_1, x_2, ..., x_n$, with an unknown distribution F_X, arose. Then the null hypothesis $\mathcal{H}_0$ that we wish to test is

$$\mathcal{H}_0 : F_X(x) = F_0(x) \text{ for all x} \tag{2.8}$$

The Glivenko-Cantelli Theorem [*Gibbons and Chakraborti*, 2003] states that the empirical distribution function of the sample, $S_n(x)$, provides a point estimate of $F_X(x)$, for all x. Therefore, for large n, the deviations between the true and empirical distribution functions, $|F_X(x) - S_n(x)|$ should be small for all values of x.

The statistic KS1 used in the Kolmogorov-Smirnov one-sample test is defined as [*Gibbons and Chakraborti*, 2003]:

$$KS1 = \max_x |F_0(x) - S_n(x)| \tag{2.9}$$

The critical value of the exact distribution of KS1, under hypothesis H_0, has been tabulated for various values of *n* [see *Gibbons and Chakraborti*, 2003].

Chapter 3

Some insight on time distribution of strong events occurrence: evidence for departures by stationary Poisson Model

3.1 Introduction

The time independence of large earthquake occurrence is one of the most commonly accepted paradigms in seismology. This important feature of the earthquake generating process is the main rationale behind the choice to consider (more or less tacitly) large earthquakes occurring inside a specific region to be distributed as a stationary Poisson process [e.g., *Vere-Jones*, 1970]. The only suggested discrepancy from this model is the presence of a few "doublets" [*Kagan and Jackson*, 1999], that are events very close in time (from days up to few years) and space (tens of km).

The time independence (Poisson) hypothesis has many interesting and important implications both from a physical and practical point of view. As a first important remark, this hypothesis implies that there is a significant difference between the time distribution of large and smaller earthquakes, for which the interaction between events is predominant. In fact, the time distribution of medium-small earthquakes

is strongly governed by triggering effects, that lead to clusters of seismic activity well described by Epidemic Type Aftershocks Sequence (ETAS) models [*Ogata*, 1988; 1998]. On the other hand, the independence hypothesis presupposes that the evidence of coupling between large earthquakes reported in literature [*Chéry et al.*, 2001a, 2001b; *Kagan and Jackson*, 1999; *Mikumo et al.*, 2002; *Pollitz et al.*, 2003; *Rydelek and Sacks*, 2003; *Santoyo et al.*, 2005] can be considered exceptions rather than a rule. In terms of the earthquake generating process, the independence hypothesis for large earthquakes suggests that the triggering involves a spatial scale significantly lower than the usual distance between large earthquakes, and that the earthquakes time distribution is not universal because physical processes which generate large and small earthquakes are different.

Actually, a robust validation of the time independence hypothesis is usually hard to achieve (as well as to disprove) because of the scarce number of large earthquakes for each seismotectonic area. In other words, with few events it is difficult to reject any hypothesis, Poisson included. Then, in spite of some evidence of long-term time dependence between strong earthquakes [*Chéry et al.*, 2001a, 2001b; *Kagan and Jackson*, 1991, 1999; *Mikumo et al.*, 2002; *Pollitz et al.*, 2003; *Rydelek and Sacks*, 2003; *Faenza et al.*, 2003; *Rhoades and Evison*, 2004; *Corral*, 2004, 2005; *Santoyo et al.*, 2005], there are not yet robust statistical tests to reject the Poisson hypothesis for the largest earthquakes worldwide, nor commonly accepted alternative statistical distribution for these events. For this reason, the stationary Poisson paradigm is implicitly accepted in many practical applications, such as in the formulation of probabilistic seismic hazard assessment methodologies based on Cornell's method [*Cornell*, 1968], and in evaluating earthquake prediction/forecasting models [e.g., *Kagan and Jackson*, 1994; *Frankel*, 1995; *Varotsos et al.*, 1996; *Gross and Rundle*, 1998; *Kossobokov et al.*, 1999; *Marzocchi et al.*, 2003a].

Notably, also other recent time-dependent approaches based on the supposed "recurrent" behavior of each seismogenic fault [*Working Group*, 2002, and references therein] are still based on the assumption of negligible interaction between large earthquakes; in fact, the probability of an event for each seismogenic structure is estimated only accounting for the "recurrence" time of the fault, without considering possible effects of surrounding large events [*Working Group*, 2002]. We stress that the use of different time distributions for faults [*McCann et al.*, 1979; *Nishenko*

and Buland, 1987; *Ellsworth et al.*, 1999; *Sieh et al.*, 1989; *Kenner and Simons*, 2005; *Hurbert-Ferrari et. al*, 2005] and seismic areas [*Cornell and Winterstein*, 1988; *Gardner and Knopoff*, 1974; *Kagan and Jackson*, 1994; 2000; *Jackson and Kagan*, 1999; *Papazachos et al.*, 1997; *Faenza et al.*, 2003] is certainly not surprising. Hypothetically, it is possible to have a very regular behavior on individual faults and a Poisson distribution over the whole region. Here, we do not deepen this important issue because the available earthquake catalogs allow us only to address the study of seismogenic zones.

One particularly interesting aspect that we consider in characterizing the statistical distribution of largest worldwide earthquakes is the presence of possible nonstationarity, i.e., changes in earthquake rate. Remarkably, in fact, almost all (antithetical) statistical models so far proposed assume stationarity for the seismogenetic process. Until now, only very few papers have faced this issue. *Marzocchi et al.* [2003b] and *Selva and Marzocchi* [2005] found a seismic rate change of decades in the Southern California seismicity. *Rhoades and Evison* [2004], even though they do not mention the nonstationarity issue, try to explain long-term fluctuations in seismicity, recognizing patterns of earthquakes that are precursory to strong events. *Marsan and Nalbant* [2005] take more explicitly into account, from a theoretical point of view, sometime misinterpreting, the probabilistic concept of stationarity, suggesting some methods to detect significant rate changes. From a more geological pointy of view, some evidence of nonstationary behavior is also detected for individual faults by paleoseismological studies [*Hurbert-Ferrari et.al*, 2005; *Ritz et.al*, 1995; *Wallace*, 1987; *Friedrich et al.*, 2003].

3.2 Data Set

Here, we analyze the distribution of worldwide large earthquakes. For this purpose, we use the *Pacheco and Sykes* [1992] (from now on PS92) catalog. This contains epicentral coordinates, origin time, surface magnitude, M_s, and seismic moment, M_0, of events from the period 1900-1990, with $M_s \geq 7.0$ and depth $d \leq 70$ km. The total number of events is 698; 547 of them occurred in the Pacific Ring (see 3.1).

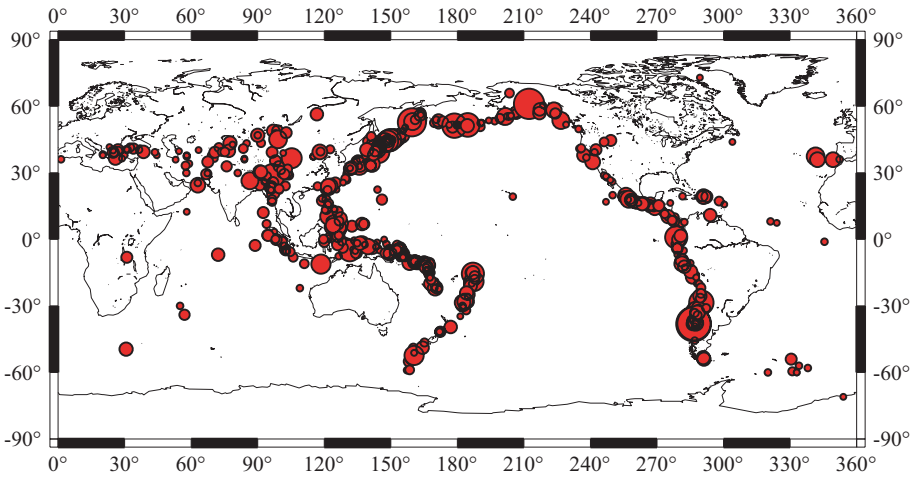

Figure 3.1: Map of locations of PS92 earthquakes. The size of circles is proportional to magnitude of events.

For PS92, all surface magnitudes and some seismic moment values are collected from the literature. For events with no published value of seismic moment, this is estimated from the M_0/M_s relationship obtained by *Ekström and Dziewonski* [1988] for large earthquakes. Surface magnitude and seismic moment are not uniformly recorded, because of human-induced changes and mixing of different inversion procedures. For this reason, the authors apply some corrections to surface magnitude in order to provide an homogeneous estimate at all times [see *Pacheco and Sykes*, 1992, for more details].

In our study we use these corrected M_s values but, to avoid the problem of saturation of the surface magnitude scale, we prefer to consider the moment magnitude M_w for events with $M_s > 8.0$. These values are obtained by seismic moment using the relation of *Hanks and Kanamori* [1979]

$$M_w = \frac{\log(M_0) - 9.1}{1.5} \tag{3.1}$$

where M_0 is measured in Nm. Since all but two of events with $M_s > 8.0$ have independently (i.e from literature) determined seismic moments, saturation of surface-wave magnitudes should not affect M_w estimation. For events with $M_s \leq 8.0$ the M_w values are very close to corrected M_s values provided by catalog. The largest computed moment magnitude value is $M_w = 9.6$, for the 1960 Chile earthquake:

this is the strongest earthquake in the history of recorded events.

3.3 Worldwide Zonation

In this study, we deal with the spatial distribution of large earthquakes by means of a subdivision of the Earth surface in zones that are homogeneous with regard to seismic activity and orientation of the predominant stress field. For this purpose, we use the Flinn-Engdahl worldwide regionalization [*Flinn and Engdahl*, 1965; *Flinn et al.*, 1974]. This choice guarantees that the results obtained in this study are not biased by some sort of "ad hoc" selection of the areas, for instance by using the same earthquake catalogs twice, first to define regional boundaries, and then to estimate seismicity parameters in identified areas.

The Flinn-Engdahl (FE hereinafter) regionalization consists of 50 seismic zones (see Figure 3.2), and it is however too detailed for our analysis; for instance, some regions are too small and contains a very low number of events. In order to avoid this problem we explore the possibility to define a new regionalization, merging some FE zones into larger tectonically homogeneous zones.

We group original FE regions by computing a representative mean focal mechanism for each one of the 50 zones, using the Cumulative Moment Tensor Method introduced by *Kostrov* [1974]. This consists of summing all moment tensors of earthquakes in a given area and then of computing the best double couple for this cumulative tensor. We use source parameters contained in the CMT catalog ([*Dziewonski and Woodhouse*, 1983]; http://www.seismology.harvard.edu/projects/CMT/) of global seismic events occurred in the period 1977-2004, with depth ≤ 70 km and magnitude $M_w \geq 5.5$ (11148 earthquakes). Since, considering the Aki convention [*Aki and Richards*, 2002], large differences of strike (ϕ) and dip (δ) angles can be obtained for two close planes, we apply to regional mean double couples the biunique angle transformation proposed by *Selva and Marzocchi* [2004]:

$$\begin{aligned} \phi' &= \phi; \qquad \delta' = \delta \qquad (0 \leq \phi < 180) \\ \phi' &= \phi - 180; \qquad \delta' = 180 - \delta \qquad (180 \leq \phi < 360). \end{aligned} \tag{3.2}$$

Then we compare focal mechanism parameters for neighboring FE zones and group

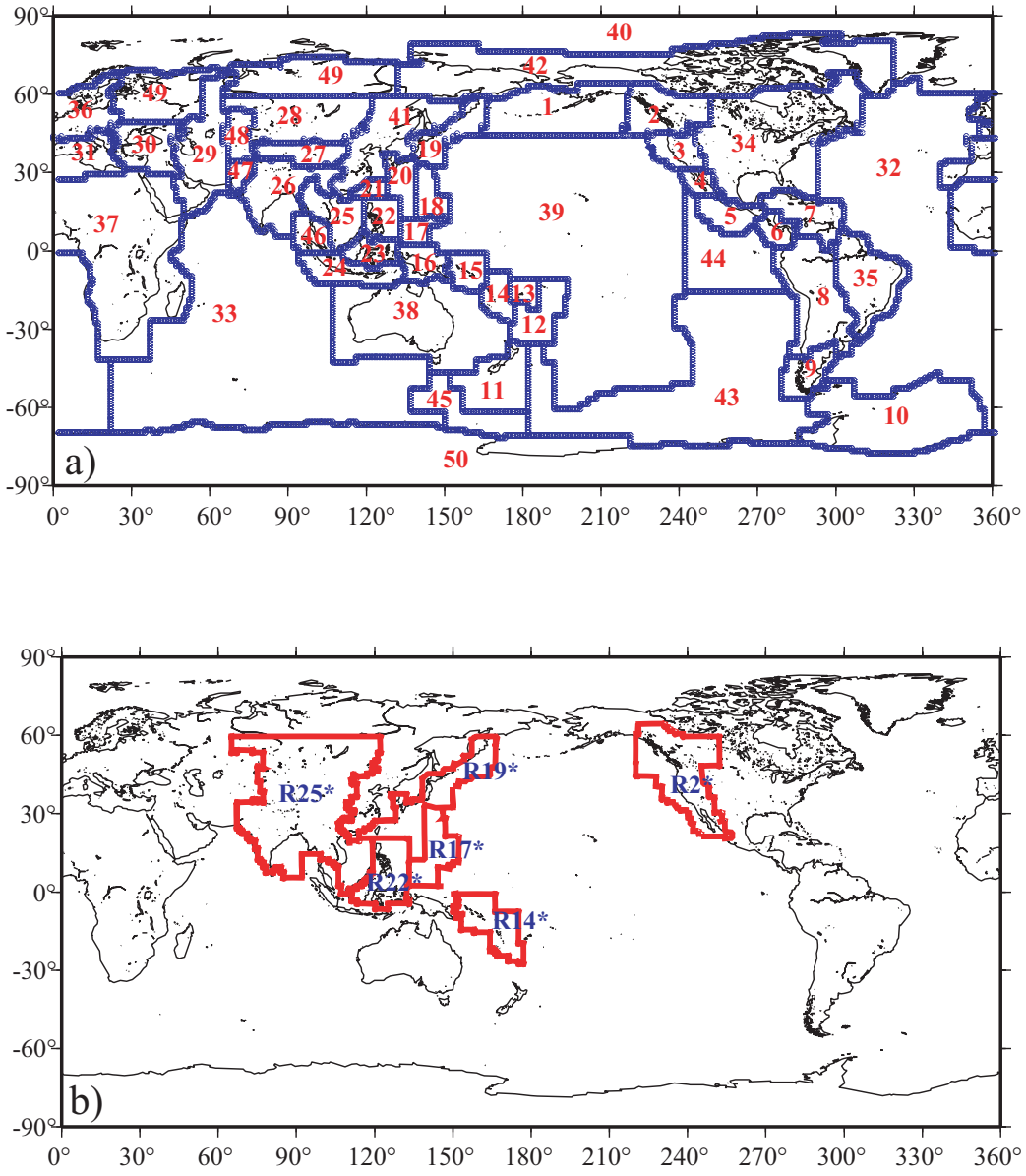

Figure 3.2: Map of 50 Flinn-Engdahl seismic regions [*Flinn and Engdahl*, 1965] (upper panel) and map of new zones defined regrouping 17 Flinn-Engdahl regions (lower panel) by *Kostrov* [1974] method (see text for details).

zones for which the difference between strike, dip and rake angles is $\leq 45°$.

The 6 new regions obtained by this method (see Figure 3.2b) are listed in Table 3.1, together with the 17 original FE zones that they comprise. The name of these new

New zones	FE zones
R2*	2-3-4
R14*	14-15
R17*	17-18
R19*	19-20-21
R22*	22-23
R25*	25-26-27-28-47

Table 3.1: Regions defined regrouping 17 Flinn-Engdahl zones by Kostrov (1974) method (see text for details).

zones is marked with an asterisk in the following. The other regions correspond to original FE zones and they will be identified in the following with the same number assigned by *Flinn and Engdahl* [1965] (for example: R1 coincides with the FE zone 1, R5 is the FE zone 5 and so on).

3.4 Methodology and Results

We investigate the time distribution of large earthquakes by following the minimalist Occam's razor philosophy. In a nutshell, we start by considering the simplest possible model, i.e., the least informative and/or the one with the least number of parameters; then, if the data show significant departures from this hypothesis, we consider more complicated models whose characteristics are indicated by the discrepancies found. The reliability of each model is statistically tested, and the relative ability of each model to describe the data is estimated by calculating the corrected version of the Akaike Information Criteria (AICc) [*Hurvich and Tsai*, 1989]. The AICc statistic is defined by

$$AICc(K) = -2LogL + 2K\frac{N}{N-K-1} \tag{3.3}$$

where N is the sample size, K is the number of parameters, and L is the likelihood of the model. The corrected version of the Akaike Criterion removes the deficiency of AIC, defined as $AIC = -2LogL + 2K$ [*Akaike*, 1974], to choose an over-parameterized model, when the number of parameters is high relative to sample

size (more than 30%).

In our case, the simplest possible model is the stationary Poisson process. This assumes that the inter-event times between events (IETs hereinafter) are independent and exponentially distributed. We check these two properties by means of two non-parametric tests, namely the Runs test and the one-sample Kolmogorov-Smirnov test (KS1) [*Gibbons and Chakraborti*, 2003]. The Runs test verifies the reliability of the independence of the earthquakes occurrence [see Appendix 3A], while the KS1 checks the hypothesis that IETs are exponentially distributed [see Appendix 2A]. The use of both is necessary because the goodness-of-fit tests are insensitive to possible memory in the data (the order of the sequence is not considered in KS1, as well as for any goodness-of-fit test). In other words, this means that it is possible to find a good fit with a stationary Poisson process even if the process is nonstationary and/or autocorrelated.

The concept of stationarity is particularly important here, and it deserves a careful definition. In probability theory a stochastic process X_t is called *stationary* if, for all n, $t_1 < t_2 < \ldots < t_n$, and $h > 0$, the joint distribution of $[X(t_1+h), \ldots, X(t_n+h)]$ does not depend on h [e.g., *Cox and Lewis*, 1966; *Daley and Vere-Jones*, 2003]. This means that the statistical description of the process is invariant with respect to shifts of the starting time; then the stochastic behavior of a stationary process is the same, no matter when the process is observed. In particular, a point process is stationary if it possesses stationary increments, i.e., the distribution of the number of events that occur in any interval of time $[t, t + h)$ depends on the length of time interval h (and not on t), and possibly on external variables (for example magnitude of events). In order to further clarify the concept, it is useful to comment on the difference between a nonstationary process and the so-called "time-dependent" processes used in seismology (ETAS, Brownian Passage Time, Weibull, etc...). All the time-dependent processes so far proposed in seismology are stationary, because the parameters of the models do not vary with time. The misleading use of the term "time-dependent" actually refers to the fact that the hazard rate $\lambda(t)$ depends on time t. Moreover we again stress that the stationary has not be mistaken with the "memory" of a process, indicative of dependence on the "elapsed" times since the last event (as for renewal models) or on information about all past events (as in ETAS models).

3.4.1 Testing the stationary Poisson model

Table 3.2 reports the results of statistical tests, described in Appendixes 2A and 3A, applied to the stationary Poisson hypothesis. The results clearly indicate that such a hypothesis is rejected at a significance level $\alpha < 0.01$ for three regions. Despite the fact that we carry out two tests for 15 regions, it is very unlikely (probability ~ 0.003) that 3 out of 30 independent tests are rejected at a 0.01 significance level by chance.

Region	**# events**	**α Runs test**	**α KS1**	
R1	42	0.68	0.05	
R2*	21	0.69	> 0.2	
R5	39	0.90	> 0.2	
R6	17	0.04	0.15	
R8	56	0.84	0.1	
R11	15	0.14	> 0.2	
R12	35	0.33	> 0.2	
R14*	96	0.71	0.01	$\leftarrow$
R16	29	0.51	> 0.2	
R19*	125	0.03	0.01	$\leftarrow$
R22*	54	0.39	> 0.2	
R24	12	0.05	> 0.2	
R25*	54	0.01	> 0.2	$\leftarrow$
R29	11	0.42	> 0.2	
R30	19	0.88	0.1	
Global Catalog	698	0.85	0.01	$\leftarrow$

Table 3.2: Runs Test and KS1 Test applied to IETS of the PS92 catalog to verify the Poisson model. Tests are performed both on global catalog as on data relative to regions with more than 10 events. Last two columns show the significance level at which we reject the null hypothesis of Poisson model. Arrows indicate regions for which almost one of two test is rejected at 0.01 significance level. Asterisks identify the regions obtained by regrouping some original Flinn-Engdahl zones (see text).

Moreover, the Poisson hypothesis is also rejected for the whole catalog; this is particularly important, because it has been proved that the sum of n Poisson pro-

cesses having different parameter λ_i gives another Poisson process with rate $\lambda_{tot} = \sum_{i=1}^{n} \lambda_i$ [*Cox and Lewis*, 1966]. In other words, the significant departure from the Poisson hypothesis on the global catalog means that earthquakes in one or more tectonic areas are not distributed as a Poisson process.

In Figure 3.3 we report the empirical cumulative distribution function of the IETs for the global catalog.

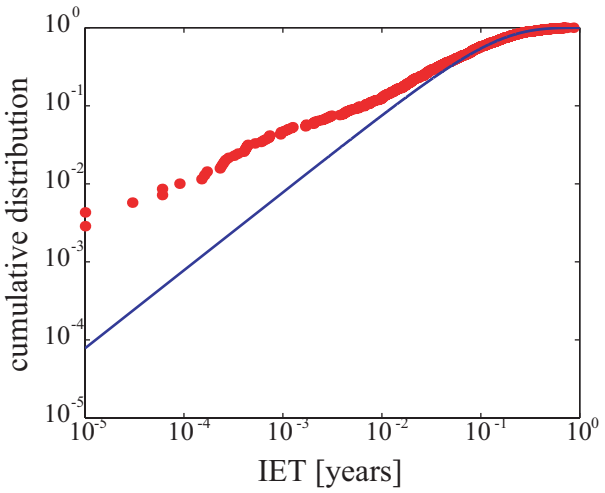

Figure 3.3: Empirical cumulative distribution of IETs $\Delta t_i = t_i - t_{i-1}$ for PS92 catalog (reds dots) along with the exponential distrubution expected for a stationary Poisson model (continuous blue line) on logarithmic scale.

As we can see from the figure, the main discrepancy from the theoretical Poisson distribution is an excess of small IETs. This suggests the presence of a more clustered distribution than Poisson. For this reason, we decide to test the reliability of a more complicated model, the stationary ETAS model, that is suitable to describe a cluster process.

3.4.2 Testing the stationary ETAS model

The main characteristics of the ETAS model [*Ogata*, 1988; 1998] are reported in previous Chapter (Section 2.3). Here, we only emphasize that the choice to check its reliability has two main rationales: to account for the clustering found in the previous section, and to check if a distribution that fits well the seismicity in a quite

different time-space-magnitude window (i.e., aftershocks sequence) is also able to reproduce satisfactorily the main features of the worldwide large earthquakes time distribution. The latter is the cornerstone for universality [*Bak et al.*, 2002].

To estimate the parameters of the stationary ETAS model ($\nu, k, c, p, \alpha, d, q, \gamma$; see Section 2.3 for more details of these parameters) we use the iteration algorithm developed by *Zhuang et al.* [2002], which simultaneously provides an estimation of the distribution for location of background events $u(x, y)$, by a kernel method. At each step, we identify parameters that optimize the log-likelihood by using the Davidon-Fletcher-Powell method [*Fletcher and Powell*, 1963]. This procedure provides also a good approximation of the inverse Hessian matrix, that is the variance-covariance matrix of the estimated parameters.

	ETAS 1 ($q = 1.5$)	**ETAS 2 ($q \neq 1.5$)**	**Poisson ($K = 0$)**
	# par.=6	# par.=7	#par.=1
ν	6.7 ± 0.3 year^{-1}	6.7 ± 0.3 year^{-1}	7.8 ± 0.3 year^{-1}
K	$(4 \pm 1)10^{-3}$ year^{p-1}	$(4 \pm 1)10^{-3}$ year^{p-1}	
p	1.1 ± 0.1	1.1 ± 0.1	
c	$(2 \pm 1)10^{-4}$ year	$(2 \pm 1)10^{-4}$ year	
α	1.2 ± 0.2	1.2 ± 0.2	
d	26 ± 3 Km	40 ± 10 Km	
q	1.5	1.9 ± 0.3	
γ	~ 0.0	~ 0.0	
Log-likelihood	-10957.2	-10956.9	-11248.3
AICc	21926.4	21927.9	22498.6

Table 3.3: Parameters of the ETAS model. Maximum Likelihood parameters, Loglikelihood and AICc values, for Poisson Model and for two versions of stationary ETAS model (obtained imposing $q = 1.5$, or estimating also q) for PS92 catalog.

The parameters q and p are of particular interest here; they govern the decay in space and time, respectively, of the coupling between events. Since some studies show that static stress changes decrease with epicentral distance as r^{-3}, we test this hypothesis by checking if q is significantly different by 1.5 (see Section 2.3). In this way we also check indirectly the trade-off between parameters q and d that may cause different pairs of q and d values to provide almost the same likelihood [*Kagan*

and Jackson, 2000]. The maximum likelihood parameters and relative errors for the whole catalog, obtained both imposing $q = 1.5$ (ETAS Model 1) and optimizing q (ETAS Model 2), are reported in Table 3.3. In both case γ is not significantly different by 0. We also list the maximum likelihood rate, obtained by modeling the dataset as a stationary Poisson process. Notably, values of all parameters and log-likelihoods of ETAS Models 1 and 2 are very close, and the parameter q of Model 2 is not significantly different from 1.5. This indicates the stability of results. In our calculations we set the maximum interaction distance between events, R_{max} (see Section 2.3), to 1000 km [see also *Kagan and Jackson*, 2000]. However, some trials show that estimated values of parameters do not depend on the chosen value of R_{max}, if this is larger than 300 Km. Note that this value is comparable to what was found by *Huc and Main* [2003], and it seems to limit to a few hundreds of kilometers any significant effects of co-seismic perturbation.

We compare performance of the Poisson and two ETAS models through the statistic AICc. The lower AICc value shows that ETAS Model 1 is the best model and than q is not significantly different from 1.5. Moreover the AICc statistics indicate that the ETAS model is a better representation of the data than a Poisson process and hence that the clustered activity in our database is not negligible.

The parameter ν is the overall rate of occurrence for non induced events, i.e., the background seismic rate (see Section 2.3). For stationary ETAS Model 1, ν is constant and it is 6.7 ± 0.3 yr^{-1} (607 background events in a time interval of 92 years covered by the catalog). Other estimated parameters predict that the coseismic triggering effect becomes negligible after a few months the occurrence of an inducing event and at 300 km of distance from its location.

We stress that the use of constant parameters of the ETAS model assures the stationarity of the process. In fact, a cluster process such as the ETAS model is stationary if the cluster centers process is stationary and the distribution of the cluster members depends only on their positions relative to the cluster center [*Daley and Vere-Jones*, 2003].

The ability of the ETAS model to follow the dynamics of the time series is tested through the analysis of residuals [see *Ogata*, 1988]. The occurrence times t_i of earthquakes are transformed into new values τ_i by the relation

Region	# events	α Runs test	α KS1	
R1	42	0.85	0.08	
R2*	21	0.76	0.24	
R5	39	0.91	0.32	
R6	17	0.10	0.10	
R8	56	0.88	0.51	
R11	15	0.41	0.16	
R12	35	0.37	0.39	
R14*	96	0.67	$<$0.01	$\leftarrow$
R16	29	0.47	0.24	
R19*	125	0.02	$<$0.01	$\leftarrow$
R22*	54	0.39	0.16	
R25*	54	0.01	0.48	$\leftarrow$
R30	19	99.1	0.14	
Global Catalog	698	0.75	0.10	

Table 3.4: Test of the ETAS model (Residuals). Results of residuals analysis for stationary ETAS model, performed on global catalog and on data relative to regions with more than 10 events. The tests are the same used in Table 3.2. Arrows indicate regions for which almost one of two tests is rejected at 0.01 significance level. Asterisks identify regions obtained by regrouping some original Flinn-Engdahl zones.

$$\tau_i = \int_{T_{start}}^{t_i} dt \int_{\mathcal{R}} dxdy\lambda(t, x, y/\mathcal{H}_t) \tag{3.4}$$

where T_{start} is the start time of observation history, $\mathcal{R}$ is the examined region and $\lambda(t, x, y/\mathcal{H}_t)$ is the conditional intensity of the ETAS model (see Section 2.3). If the model describes well the temporal evolution of seismicity, the transformed data τ_i are expected to behave like a stationary Poisson process with unit rate [*Ogata*, 1988]. The goodness of fit of the stationary ETAS model to each region is evaluated by testing the null hypothesis that values $\Delta\tau_i = \tau_{i+1} - \tau_i$ are exponentially distributed (with mean equal to 1). The tests used are the Runs and KS1 test as before. In Table 3.4 we report the results for each zone considered. From the results we can see that three areas, R14*, R19* and R25*, show significant departure from the null hypothesis, implying that the stationary ETAS model is not able to describe well the

data in those regions.

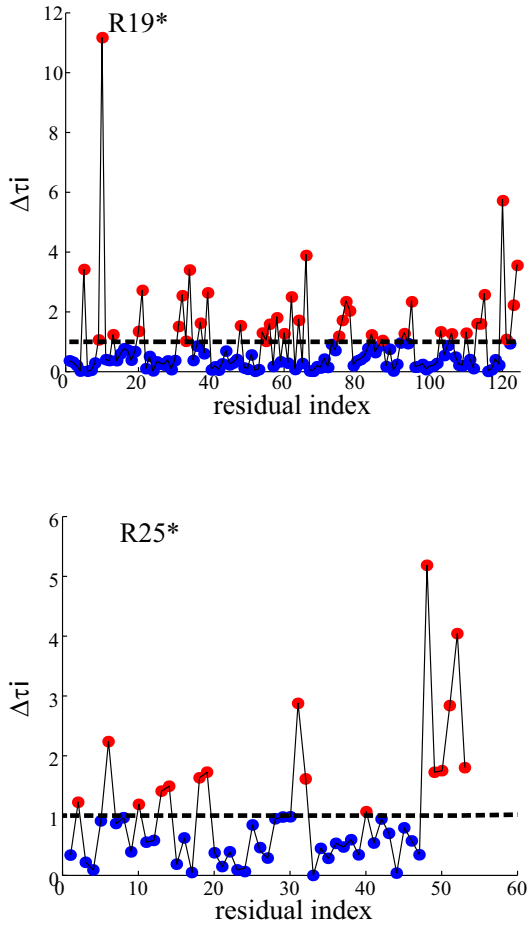

Figure 3.4: Sequence of values $\Delta\tau_i$ for regions R19* and R25* where we reject the stationary ETAS model by Runs test of residuals. Blue and red dots mark values below and above the average, respectively. The dashed black line represent the average (equal to 1).

The presence of significant autocorrelation in some sequences (see results of the Runs test in Table 3.2 and in Table 3.4) may suggest the presence of long-term modulation of earthquake rates, at least in some zones. In Figure 3.4, we show the sequence of $\Delta\tau_i$ in regions for which we reject the stationary ETAS model by the Runs test on residuals (R19*, R25*). These plots highlight persistences, i.e., almost uninterrupted subsequences all above or below the mean rate, suggesting that the discrepancies found could be due to nonstationarity.

Since the low number of events in each region does not allow very elaborate anal-

yses, we only investigate if the average seismic rate varies with time. This type of nonstationarity is known in literature as *weak nonstationarity* [*Daley and Vere-Jones*, 2003]. Moreover, since the paucity of induced events (about 10% of all events) in the database does not allow more investigation on time variability of triggering capability into seismotectonic zones, we only explore the nonstationary behavior of the background rate ν. In particular, we check if the departure from the stationary ETAS model is due to failure of the stationary hypothesis for the Poissonian background seismicity, by analyzing the time distribution of the declustered catalog.

The declustering of datasets is performed by applying the *Zhuang et al.* [2002] procedure, developed for ETAS modeling. This technique estimates the probability π_i that the i-th event belongs to background activity by the ratio between the background rate $\mu(x_i, y_i)$ and the total occurrence rate $\lambda(t_i, x_i, y_i/\mathcal{H}_{t_i})$, both computed by the stationary ETAS model (see Section 2.3). Then we assign each event i to the declustered catalog if π_i is larger than 0.5.

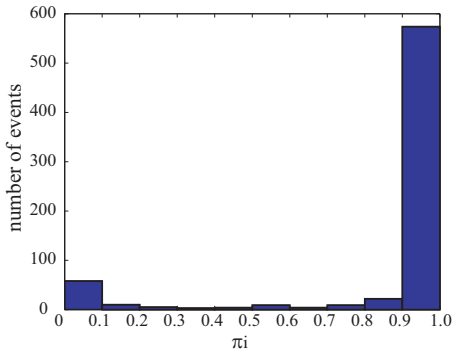

Figure 3.5: Histogram of probability π_i of belonging to background seismicity for events of Pacheco and Sykes (1992) catalog. These values are computed by stationary ETAS model. Events with a probability π_i larger than 0.5 are assigned to declustered catalog.

By this algorithm we identify 618 background events, and 80 aftershocks. The choice of probability threshold is not critical because 632 events (i.e., more than 90%) have a probability π_i close to 0 or 1 (see Figure 3.5), and only 66 events have π_i in the range $(0.1, 0.9)$. Note that, although any declustering procedure produces non-unique identifications (because of the intrinsic nature of aftershocks), this sharp bimodal distribution of π drastically reduces biases due to possible after-

shock misidentifications. We also stress that the declustering of the PS92 catalog is performed taking as fixed all parameters of the ETAS space-time model. This assumption implies that the physical mechanism of earthquake generation is spatio-independent, therefore the details of the time and space distribution are unchanged. Only the seismic background can differ in distinct tectonic regions.

3.4.3 Testing the nonstationary ETAS model (NETAS)

To test for nonstationarity we compare the stationary ETAS model with a new model in which background rate changes with time ($\nu(t)$). This model, called in the following *NETAS* (Nonstationary ETAS) model, is formulated by modeling background seismicity as a piecewise stationary Poisson process. The choice to split the sequence into stationary subsets is only a mathematical expedience to simplify the problem. We anticipate that the identification of nonstationarity with this model implies significant variations of the earthquake rate, but it cannot discriminate if these variations truly occur through discrete jumps, or through slow fluctuations. In fact, the available dataset does not allow this distinction to be made. From a physical point of view, the use of NETAS could be necessary to account for long-term interaction between earthquakes and/or for time variations in tectonic processes.

Region	**# groups**	**AICc st. Poisson background**	**AICc* nonst. Poisson background**	**s.l.**	**Runs test α**	**KS1 α**	
R5	5	133.0	131.2	0.1	0.82	0.03	
R12	4	130.1	120.5	$<$0.01	0.41	0.04	$\leftarrow$
R19*	9	177.6	161.4	$<$0.01	0.69	0.22	$\leftarrow$
R25*	3	154.3	146.5	0.01	0.57	0.05	$\leftarrow$

Table 3.5: Results of the NETAS model for the areas where NETAS perform better than ETAS. The number of groups are identified by means of a K-mean algorithm. In third and fourth columns we report the penalized AICc values (AICc*) obtained by using a stationary and nonstationary (piecewise stationary) Poisson model for the background. In fifth column, we report the significance level at which stationary model is rejected. Last two columns show results of residuals analysis.

Change points of the background seismic rate $\nu(t)$ are identified by applying the *$\mathcal{K}$-mean cluster analysis* algorithm (Appendix 3B). This procedure is commonly applied to group objects of similar kinds into respective categories. It uses two rules to classify IETs into groups that represent different stationary Poisson models (see Appendix 3B). Details and problems related to this method can be found in Appendix 3B. Here, we only stress that the use of the AICc* criterion makes negligible the possibility of overfitting the data, i.e., of providing too many change points.

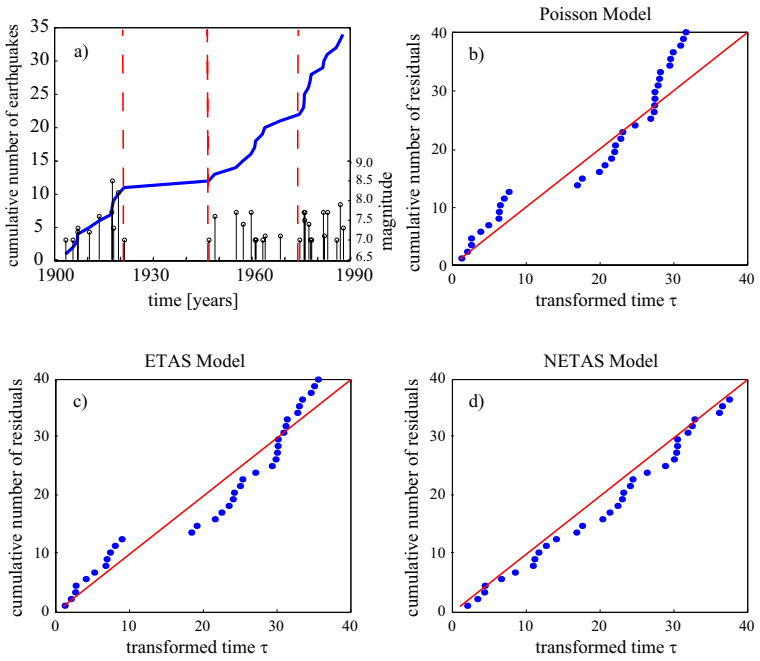

Figure 3.6: Results of analysis for region R12. a) Cumulative number of background events (blue line) with magnitudes (black stem plot) and change points identified by K-mean algorithm (red line). Comparison of normalized cumulative number of residuals (blue dots) for b) stationary Poisson, c) ETAS and d) NETAS models.

By the $\mathcal{K}$-mean procedure we identify significant change points in 4 (R5, R12, R19*, R25*) of 13 analyzed zones (the ones with more than 10 events). In Table 3.5 we report the number of groups identified by the $\mathcal{K}$-mean algorithm and AICc* values (corrected version of the AICc statistic that takes into account the increase of the degrees of freedom of the model due to the presence of change points; see Appendix 3B) for stationary and nonstationary Poisson models of the background. For the stationary background model the penalty term is zero, and then the AICc* value

is equal to AICc value. To estimate the statistical significance of the difference between the AICc* values of the two models, we apply the $\mathcal{K}$-mean algorithm to 1000 simulated stationary Poissonian catalogs with the same rate as the examined region.

Then, we compare the real difference with the distribution of the differences obtained from simulated data, to estimate the probability that the real difference can be obtained by chance, if background seismicity is a stationary Poisson process. Results are reported in the fifth column of Table 3.5. We find that the nonstationary model is significantly better (s.l.$<$0.01) for 3 (R12, R19*, R25*) of the 4 zones considered.

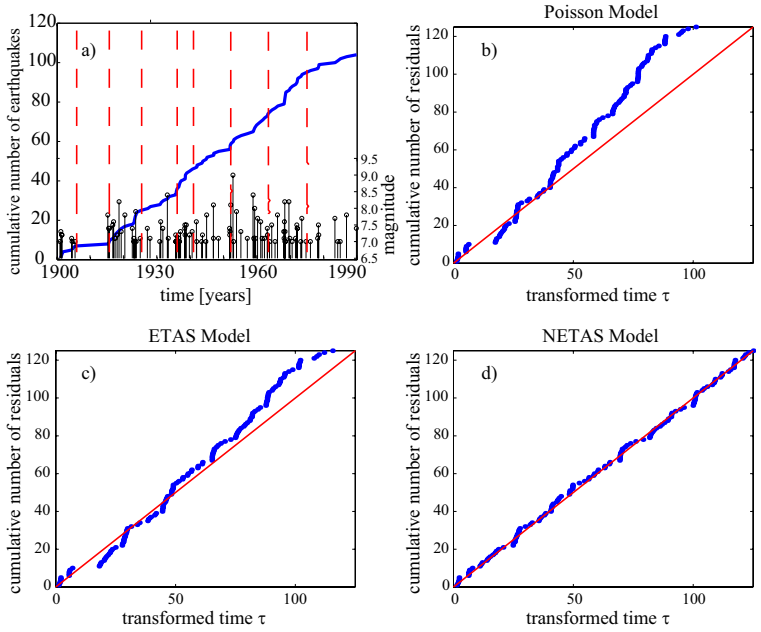

Figure 3.7: As for Figure 3.6, but relative to region R19*.

The time/magnitude plot, the boundaries of the subsets identified by the $\mathcal{K}$-mean algorithm, and the normalized cumulative distributions of residuals for the Poisson, ETAS, and NETAS models are shown in Figure 3.6 (region R12), Figure 3.7 (region R19*) and Figure 3.8 (region R25*). As a first remark, we note that the improvement of the goodness-of-fit by taking into account nonstationarity is substantial. Second, there is no evidence in favor of a systematic relation between rate changes and magnitudes of events. Finally, the change points do not occur simultaneously inside

the three regions, as we could expect if their origin would be due to man-made changes in global recording practice.

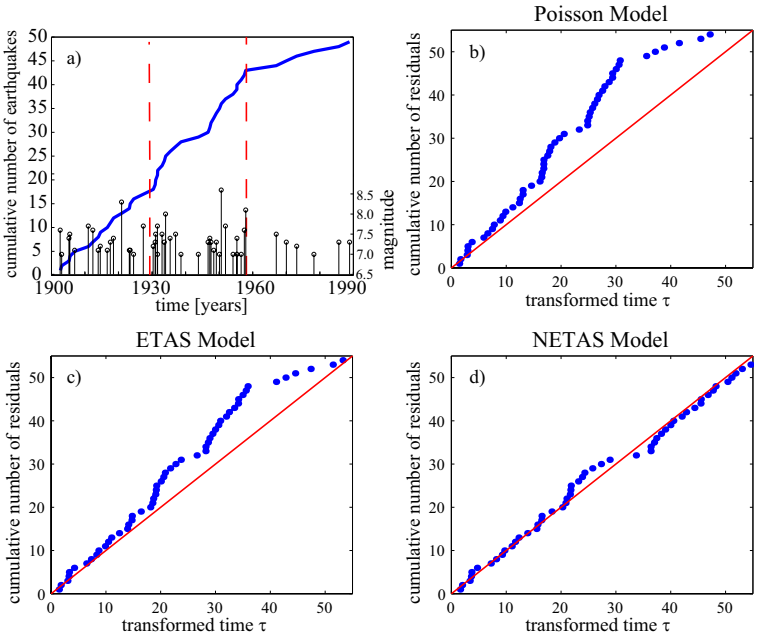

Figure 3.8: As for Figure 3.6, but relative to region R25⋆.

Suitability of the NETAS model to fit the data is judged by applying the Runs and Kolmogorov-Smirnov tests on residuals (last two columns of Table 3.5). By the Runs test we find that the hypothesis of stationarity of the residuals is not rejected for all zones. The null (Poisson process) hypothesis is never rejected by the Kolmogorov-Smirnov test at a significance level of 0.01, even though some weak discrepancies from this hypothesis (significance level of about 0.05) appear for regions R12 and R25⋆.

3.5 Discussion and Conclusions

The analysis reported above highlights two main characteristics of the time distribution of strong earthquakes: 1) the stationary ETAS model is better than the Poisson model to describe the time distribution of worldwide large earthquakes; 2) the background of the ETAS model is not always a stationary Poisson model, but it

can present time variations of decades (or longer) with different trends.

As regards the first point, the better suitability of the stationary ETAS model than the Poisson model, on both regional and global scale, points out that short-term triggering activity due to co-seismic effects [*King et al.*, 1994; *King and Cocco*, 2000; *Stein et al.*, 1997; *Toda et al.*, 1998] is a typical feature also of strong earthquakes occurrence. Remarkably, the estimated parameters of the stationary ETAS model are consistent with values computed for moderate-small events (as for aftershocks sequences) in tectonic zones [e.g., *Ogata*, 1999]. Note that this result is far from obvious, and it shows that the physical process governing short-term clustering is independent of the time-space-magnitude window considered.

As regards the second point, further investigations on a regional scale show that a stationary ETAS model does not represent all characteristics of strong earthquake occurrence. In particular, for some regions we find significant evidence of decadal-scale time variations. This evidence indicates an earthquake generating process that is time-dependent and maybe not memoryless, because it presents time variations of the rate, on a different spatio-temporal scale respect to short-term triggering. Note that this long-term nonstationarity is a novelty. Until now, the little evidence there is in favor of nonstationary behavior of seismic sequences is at regional scale (Chapter 2) [see also e.g., *Selva and Marzocchi*, 2005], and with shorter characteristic times that were interpreted in terms of static stress changes [*Marsan and Nalbant*, 2005], of a rapid evolution of magmatic or fluids intrusions (Chapter 1) [see also *Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006], or short-term changes in stress relaxation [*Ouillon and Sornette*, 2005]. At the same time, it is worth noting that some very recent paleoseismological studies report independent evidence of nonstationarity on single seismogenic structures [*Ritz et.al*, 1995; *Wallace*, 1987; *Friedrich et al.*, 2003; *Hubert-Ferrari et al*, 2005].

In general, the two results found in this analysis corroborate the hypothesis of universality of physical laws for earthquakes generation. In fact, the finding that an ETAS model with the background slowly varying with time can be considered a reliable distribution for the seismicity in a wide time-space-magnitude window is one of the cornerstone of the universality hypothesis [*Bak et al.*, 2002]. Note that the time evolution of the background is not usually recognized in other worldwide

(NEIC, CMT) [*Huc and Main*, 2003; *McKernon and Main*, 2005] or regional [e.g. *Hainzl et al.*, 2006] catalogs simply because the time variations found here have characteristic times that are comparable, if not longer, than the time length of such catalogs.

The two main findings of this study are almost in accordance with results previously reported by *Kagan and Jackson* [1991] who found a long-term clustering at global scale also for largest earthquakes. In both papers, the independence hypothesis is rejected, but there are important differences in the interpretation and modeling of the time variations found. The long-term clustering proposed by *Kagan and Jackson* [1991] implies a time-dependent stationary process where the probability of earthquake occurrence depends only on some characteristics of previous events (memory of the past), for instance, the time of occurrence and location. Moreover, each clustering model presupposes only an increase in seismicity, namely the occurrence of an event can only promote other events. On the contrary, we do not impose any long-term relationship between earthquake rate variations and information about previous events (time, magnitude or location). Our NETAS model defines a long-term nonstationary behavior of seismicity, but dependence on past history is described only by short-term clustering. Moreover our finding suggest the possibility of a decrease in seismicity, not explained by stationary long-term clustering models.

As regards the causes of nonstationarity, a first possibility is that variations in occurrence rate found in some regions can be due to inhomogeneities in collecting data or to modification of the quality of seismic network and/or macroseismic information. As a matter of fact, datasets that cover large space-time windows, such as PS92, include both instrumental as macroseismic data, making very hard the recognition of possible spurious effects. However, although we cannot firmly rule out this possibility, we think this explanation is unlikely, because the time variations found are not (always) compatible with typical trends induced by a decreasing under-reporting due to improving the quality and the coverage of seismic network. For instance, we should observe a significant increase of seismicity at the time when a new denser network is installed. In our case, instead, we observe temporal trends of earthquake occurrence (see Figures 3.7 and 3.8) that highlight a significant decrease of seismic rate in recent decades, for two zones (R19*, R25*). For region R12 the increase of

the rate in recent decades can be due to an improvement of historical information, although this likely does not justify the doubling of the rate (it increases from 0.4 to 0.9). At the same time, the loss of information is not the only cause of the lack of events in the period 1920-1950, in which a significant decrease of the rate probably occurs. Moreover, global man-made changes in recording earthquake practice would have led to almost simultaneous change points in different regions, different from what is reported here.

As regards the physical causes of nonstationarity, we identify two main potential factors: i) the presence of tectonic rates that vary with time; ii) long-term stress interactions at regional or global scale. Note that both factors act by varying loading stress into a region; what makes them different is the physical source that causes such variations.

Although the present state of the art does not allow us to rule out any one of them, some evidence may indicate a prefered mechanism. In particular, the hypothesis that the driving mechanism for nonstationarity in large earthquake occurrence is time varying plate motion is maybe the less likely. As a matter of fact, the tectonic motion appears surprisingly stable, with comparable velocities over time intervals which span 5 order of magnitude, i.e., from a few years up to millions of years [*Sella et al.*, 2002; *DeMets et al.*, 1994]. We remark that this stability is far from obvious, and it implies the presence of very low (non-seismic) fluctuations of the tectonic motion, and perhaps of the related mean regional tectonic rate. On this subject, we note that the comparison between GPS and geological estimation of tectonic motions is usually done far from seismogenetic areas [*Sella et al.*, 2002], because of the significant influence of seismic cycle effects, i.e., such as the post-seismic slip (~ 1 year) and/or deep crust - mantle viscoelasticity effects (20-30 years).

The post-seismic effects induced by great earthquakes are the driving mechanism behind the second possible factor reported above. The occurrence of any earthquake induces a perturbation in the stress field at any point on the earth's surface. Generally speaking, there are three different types of perturbations: the dynamical stress variations (DSV) due to the passage of the seismic waves, the co-seismic stress variations (CSV) due to the elastic residual deformation of the lithosphere, and the post-seismic stress variations (PSV) due to the visco-elastic readjustment

of the lower-crust and/or asthenosphere and mantle. From an observational point of view, these three perturbations are characterized by different attenuation of the effects as a function of distance from the epicenter, and different characteristic times. The DSV lasts only a few minutes (at maximum), and its maximum amplitude attenuates with distance slowly, compared to the CSV and PSV [e.g. *Gomberg et al.*, 1998]. The CSV is approximately instantaneous (being due to the elastic rebound) and it does not vary with time; its maximum perturbation decreases drastically with distance [see e.g. *Stein et al.*, 1992; *King et al.*, 1994; *Stein et al.*, 1994]. The PSV reaches its maximum effect after a few decades or centuries [e.g. *Thatcher*, 1983; *Pollitz*, 1992; *Pollitz et al.*, 1998; *Piersanti et al.*, 1995; 1997; *Piersanti*, 1999; *Kenner and Segall*, 2000], depending on the viscosity of the lower crust and mantle, and it decays with distance less rapidly than the CSV. In other words, CSV are prominent at small distances from the source (the classical aftershock sequences), while delayed effects due to the asthenosphere and/or lower crust relaxation are relatively more important at larger distances.

Until now, there is well established evidence only in favor of co-seismic effects [*Stein et al.*, 1997; *Toda et al.*, 1998; *Wyss and Wiemer*, 2000] and of some clear dynamic triggering [*Gomberg*, 1996; *Marsan*, 2003; *Marsan and Nalbant*, 2005; *Felzer and Brodsky*, 2006; *Main*, 2006]. As regards PSV, many authors reported retrospective analyses [*Pollitz and Sacks*, 1992; 1997; *Chéry et al.*, 2001b; *Freed and Lin*, 2001; *Casarotti and Piersanti*, 2003; *Selva and Marzocchi*, 2005], and evidence for long-term coupling between earthquakes and volcanic eruptions [*Linde and Sacks*, 1998; *Marzocchi*, 2002; *Marzocchi et al.*, 2002], but the real effectiveness of such a coupling is still a matter of debate. Here, we only remark that the time variations found for some zones in this study are compatible with the characteristic times of PSV [see, i.e., *Kenner and Segall*, 2000]. Moreover, we also remark that this kind of nonstationary behavior, with long-term clustering and periods of low seismicity (similar to ”seismic gap”), was explicitly predicted by a numerical model that simulates the PSV effects in a realistic tectonic setting [*Marzocchi et al.*, 2003b].

One of the more clear examples of long-term clustering of large events found here is the increase of seismicity detected between 1930 and 1960 in region R25* (Central Asia) (see Figure 3.8). This period of very high seismicity has been correlated with

postseismic regional stress evolution due to viscoelastic relaxation of the lower crust and upper mantle [*Chéry et al.*, 2001b; *Pollitz et al.*, 2003]. This physical process implies a long-term memory, not modeled by a (stationary or nonstationary) Poisson model, which is likely responsible for the nonlinear trend of the cumulative number of events, in the time period from 1930 to 1950 (see Figure 3.8).

To sum up, the present knowledge of the earthquake generating process does not allow the process responsible for the nonstationarity that we have found to be unambiguously identified. We speculate that the most likely candidate could be some sort of long-term interaction due to viscous relaxation of the asthenosphere and/or lower crust. However, different possible catalog deficiencies prevent us from firmly ruling out man-made effects, for instance due to changes in earthquake recording practice, as a possible explanation of the nonstationarity found.

From a practical perspective, the results reported here highlight that the stationary Poisson hypothesis that stands behind most of the seismic hazard assessment should be considered with care. In this case, the detection of the cause of nonstationarity in the seismic activity becomes imperative to significantly improve long-term forecasting and seismic hazard assessment.

3.6 Appendix 3A: The Runs test

The runs test can be used to test if a process is stationary and not autocorrelated (null hypothesis). In any time sequence of real numbers, a run is an uninterrupted subsequence of consecutive numbers with the same sign, immediately preceded and followed by numbers with opposite sign. In particular we can consider the randomness of runs about the mean of the sample (positive sign for each data larger than the mean of the sample, negative run otherwise). If we have N data, of which p are positive and n are negative ($n + p = N$), the probability P_R of obtaining a number of runs lesser or equal to observed number R, under the null hypothesis, is:

$$P_R = \sum_{i=3,5,\ldots}^{R} \frac{\binom{n-1}{\frac{i-1}{2}}\binom{p-1}{\frac{i-3}{2}} + \binom{n-1}{\frac{i-3}{2}}\binom{p-1}{\frac{i-1}{2}}}{\binom{n+p}{n}} + \sum_{i=2,4,\ldots}^{R} 2\frac{\binom{n-1}{\frac{i}{2}-1}\binom{p-1}{\frac{i}{2}-1}}{\binom{n+p}{n}} \qquad (3.5)$$

If p and n are both > 10, the distribution of R can be approximated by a normal distribution with mean $\bar{R}$ and variance σ_R^2 equal to

$$\bar{R} = \frac{2np}{n+p} \qquad \sigma_R^2 = \frac{2np\,(2np-n-p)}{(n+p)^2(n+p-1)}. \tag{3.6}$$

In this case we can carry out a Z-test, which consists in testing a standard normal distribution for the statistic

$$Z = \frac{R-\bar{R}}{\sigma_R} \tag{3.7}$$

[see *Gibbons and Chakraborti*, 2003].

3.7 Appendix 3B: $\mathcal{K}$-means Cluster Analysis Algorithm

The $\mathcal{K}$-means iterative algorithm is one of the simplest procedures that solve the clustering problem. It is a non-hierarchical clustering algorithm that permits compact clusters to be extracted by using global knowledge about data structure. It has many variants, but the first one was published by *MacQueen* [1967]. This procedure follows a simple way to classify a given Q-dimensional data set, through a certain number of $\mathcal{K}$ clusters fixed a priori. The main idea is to define initial $\mathcal{K}$ centroids, one for each cluster, and to associate each point of a given data set to the nearest centroid. At each step, new $\mathcal{K}$ centroids are re-calculated as barycenters of clusters resulting from the previous step. The $\mathcal{K}$-centroids change their location step by step until no more significant variations are obtained. This algorithm aims at minimizing the *within-class variance*

$$\sigma^2 = \frac{1}{N}\sum_{k=1}^{\mathcal{K}}\sum_{j\in C_k}\|u_j-\mu_k\|^2 \tag{3.8}$$

where N is the number of data vectors u_j that we are grouping, and μ_k is the centroid of k-th cluster C_k. If dimensions of the u_j vectors have very different ranges, it is advisable to normalize the data, in order to prevent one feature from dominating the clustering procedure.

There are two main problems related to application of the $\mathcal{K}$-means algorithm. The first one is that there are usually several stable partitions, depending on the initial configuration of centroids. The second is to decide the number $\mathcal{K}$ of clusters: increasing $\mathcal{K}$ reduces the within-class variance σ^2, but also increases the risk of overfitting. We resolve these problems by using a *model-based criteria*.

Let us assume a general dataset $\mathcal{D} = \{u_j\}$ that comes from a mixture of $\mathcal{K}$ groups of random variables with a probability density function $f_{\hat{\theta}_k}$, where $\hat{\theta}_k$ is vector of parameters estimated for cluster C_k. In this case, we can invoke the AICc model selection. The log-likelihood of $\mathcal{D}$, for a specified configuration $\mathcal{M}$ with $\mathcal{K}$ clusters is:

$$\text{Log}[L(\mathcal{D}|\mathcal{M},\sigma^2)] = \log\Big(\prod_j f(u_j|\mathcal{M},\sigma^2)\Big) = \log\Big(\prod_j f_{\hat{\theta}_{k_j}}(u_j)\Big). \tag{3.9}$$

where k_j is the cluster to which u_j is assigned. Since potential change points are selected by data, we must consider a $corrected$ version of the AICc, given by:

$$\begin{aligned}\text{AICc}^\star &= \text{AICc} + pen(N,\mathcal{K}-1) = \\ &= -2\text{Log}L(\mathcal{D}|\mathcal{M},\sigma^2) + 2p\mathcal{K}\tfrac{N}{N-p\mathcal{K}-1} + pen(N,\mathcal{K}-1).\end{aligned} \tag{3.10}$$

where p is the dimension of $\hat{\theta}_k$ and $pen(N,\mathcal{K}-1)$ is the contribution of $\mathcal{K}-1$ change points as adjusted parameters. This last term depends on the number of events in the dataset. The AICc* statistic permits the detection of the most likely value of $\mathcal{K}$, above which clustering quality does not improve, and the identification of the configuration of centroids which maximizes the likelihood.

In our case, we have a dataset composed by IETs, that is $\mathcal{D} = \{\Delta t_j\}$. If we assume that they come from a mixture of $\mathcal{K}$ exponential distributions with rate λ_k (then $p = 1$), for a specified configuration $\mathcal{M}$ with $\mathcal{K}$ clusters the likelihood of $\mathcal{D}$ is:

$$\text{Log}[L(\mathcal{D}|\mathcal{M},\sigma^2)] = \sum_{k=1}^{\mathcal{K}}\Big(N_k\log(\lambda_k) - \lambda_k T_k\Big) \tag{3.11}$$

where N_k is the number of interevent times in the k-th cluster and T_k is sum of their lengths. We estimate λ_k as the ratio between N_k and T_k (maximum likelihood estimator).

Each IET between two adjacent groups is assigned to the group with the lowest rate, if the probability of no event in this interval time (by the Poisson model) is larger than 5%. Otherwise it is considered as a further group, with rate equal to the value which provides a probability of occurrence equal to 5%.

Among possible initialization strategies, we simply use a random subset of the data as initial centers. Therefore we choose the final partition that has the lowest AICc* value. To compute the penalty term $pen(N, \mathcal{K}-1)$ we use an empirical procedure. We simulate 10000 stationary poissonian catalogs with the same rate as the real dataset (different for each zone). For each catalog we select the best nonstationary model defined by the $\mathcal{K}$-mean procedure, imposing that the number of groups is $\mathcal{K}$. Then we compute the penalty term as the average (over all simulated catalogs) of differences between the AICc values for stationary and nonstationary models. This procedure is applied for all $\mathcal{K} = 2, ..., 10$.

Chapter 4

Long-term Memory and Nonstationarity in strong earthquake occurrence: comparison of two hypotheses

4.1 Introduction

The modeling of spatio-temporal distribution of moderate-large earthquakes forms the basis for time-dependent seismic hazard assessment and, more generally, for forecasting calculations. The huge social impact of risk mitigation has promoted formulation of different and also contradictory models [*Kagan and Jackson*, 1991; *Ellsworth et al.*, 1999; *Nishenko and Buland*, 1987; *Michel and Jones*, 1998; *McCann et al.*, 1979; *Shimazaki and Nakata*, 1980] directed towards the understanding of main features of dangerous earthquakes occurrence. Then, as a matter of fact, if short-term spatio-temporal clustering of events in sequences is now widely accepted, quite different statistical models, on time periods of decades or centuries, are still commonly used in forecasting and hazard studies [*Working Group*, 2002]. For this reason, the stationary Poisson paradigm is still implicitly accepted in many practical applications, such as in the formulation of probabilistic seismic hazard assessment methodologies based on Cornell's method [*Cornell*, 1968], and in eval-

uating earthquake prediction/forecasting models [e.g., *Kagan and Jackson*, 1994; *Frankel*, 1995; *Varotsos et al.*, 1996; *Gross and Rundle*, 1998; *Kossobokov et al.*, 1999; *Marzocchi et al.*, 2003a].

In this context the role of stochastic modeling and of statistical tools is of primary importance. Detection of possible significative rate changes and, more generally, of significative departures by hypothesis of time independence of large earthquake distribution can help to improve our understanding on driving mechanisms of seismicity.

Two main reasons can be given for justifying departures by hypothesis of Poissonian distribution: (1) the history of occurrence exhibits persistences (i.e. long-term dependence) [*Kagan and Jackson*, 2000] and (2) the main properties of the seismic rate of an area has endogenous variations with time (weakly nonstationarity) [Chapter III; see also *Lombardi and Marzocchi*, 2007]. The first point concerns the problem of the *memory* of the process, i.e. the influence that the occurrence of an event has on future seismic rate. In terms of stochastic modeling the dependence on past history is represented by introducing the occurrence times before t in expression of the mean rate $\lambda(t)$. To avoid any misunderstanding, we remark that for a stochastic model the dependence on the past history (memory) is not indicated by the "time dependence" property. Specifically the rate $\lambda(t)$ of a "time-dependent" process changes with time (i.e. it is a function of t), but is not necessarily function of past occurrence times. For example, the occurrence rate of a nonstationary Poisson process depends on time t, but not on past occurrence history (see Table 1 of Introduction). On the other side the ETAS model [*Ogata*, 1988; 1998] is an example of a process of which time evolution is strongly affected by past events: the occurrence times t_i before t are explicitly introduced in the expression of $\lambda(t)$, so that the past history drives the probability of future occurrence for a certain time period. A well known example of a temporal trend influenced by the memory of the past, well recognized in seismology at short spatio-temporal scale, is the clustering, that imposes the occurrence of an event promote future seismicity. However we stress that the memory in a process could be modeled in other way, for example by a decrease of the rate after the occurrence of an event, as the seismic gap hypothesis [*McCann et al.*, 1979] or, in some cases, interaction faults models based on Coulomb stress changes [*Stein*, 1999; *Piersanti et al.*, 1995] predict for earthquakes.

As regard the stationarity, we remark that this is a critical issue of probabilistic modeling, involving that the statistical descriptors of process under study (for example the mean rate) are invariant for different temporal ranges [*Daley and Vere-Jones*, 2003]. This hypothesis is generally assumed in seismic forecasting and hazard assessment, but it is rarely tested and adequately discussed. Actually it is crucial in any statistical analysis: if we are interested in estimating parameters or testing any hypothesis in cases where the set of variables is not entirely stationary, standard techniques are often largely invalid and, therefore, inappropriate.

At short time scales the memory in earthquake occurrence is a well recognized feature, mainly justified with faults interaction and with stress variations produced by each earthquake on area surrounding its location. In Chapters I and II the problem of nonstationarity recognition in sequences has been token into account and interpreted in terms of underlying driving physical processes [see also *Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006]. On longer spatio-temporal scale the debate on these issues is much more open and the achievement of an agreement is complicated by some problems: the too small number of data to test adequately any hypothesis, the lack of an unambiguous definition of term ”large earthquake”, that is related to seismic capability and tectonic structure of the region, the uncertainty of geologic and geodetical data, from which most hazard models are derived. All these inefficiencies affect also formulation of long-term models, that are more based on subjective belief than on checks on data. To understand which is (if there is) among memory of past history and nonstationary behaviour the predominant feature of long-term (and long-range) earthquake occurrence can help to improve present forecasting models and hazard methodologies.

Extensive research was carried out to develop suitable statistical methods for estimating and modeling both the intensity of long-range dependence and nonstationarity, especially in time series analysis [*Box and Jenkins*, 1976]. For many time series, autoregressive moving-average (ARMA) models serve as a parsimonious way to summarize the autocovariance structure of the data [*Box and Jenkins*, 1976; *Brockwell and Davis*, 1987]. One limitation of such model is that autocovariances die off relative quickly and then they fit short-memory time series, for which an event affects the level of the series for a relatively short time. Moreover they are

implicitly assumed to be stationary. To fill this gap, the ARMA modeling has been firstly extended to integrated ARMA (ARIMA) processes [*Box and Jenkins*, 1976; *Brockwell and Davis*, 1987], that take into account nonstationary behaviour, and then to fractionally integrated ARMA (ARFIMA) models [*Granger and Joyeux*, 1980; *Hosking*, 1981], directed towards the inclusion of long-memory persistence, too. These last models are more suitable for data that have a slowly decaying autocorrelation function and then they have come to play an increasing role in large number of applications, as longer time series have become available.

We stress that here, the ARFIMA models are considered more as an investigation tool than as a modeling strategy to use for hazard and forecasting calculations. We use them to recognize the presence of long-term memory or a nonstationary trend in earthquake occurrence of a prefixed area. Results coming from this analysis can be used for a more acquainted modeling of seismicity and to improve performance of hazard assessment.

4.2 Data Sets

In this study, we deal with the distribution of moderate-large earthquakes, by separately analyzing data collected into two worldwide and one regional catalogs. The main rationale behind the choice of investigating more datasets is to compare behaviour of time histories recovering different domains. Moreover the analysis of worldwide datasets can highlight variability of main features of earthquake occurrence due to different tectonic characteristics.

The first dataset that we consider is the *Pacheco and Sykes* [1992] (hereinafter PS92) catalog. This contains epicentral coordinates, origin time, surface magnitude, M_s, and seismic moment, M_0, of 698 events occurred in the period 1900-1990, with $M_s \geq 7.0$ and depth $d \leq 70$ km. The values of M_s can be considered homogeneous in time, because the authors apply some corrections to original estimates in order to compensate the lack of uniformity in recording [see *Pacheco and Sykes*, 1992, for details]. In our study we use these corrected M_s values but, to avoid the problem of saturation of surface magnitude scale, we prefer to consider the moment magnitude M_w for events with $M_s > 8.0$. These values are obtained by seismic moment

using the relation of *Hanks and Kanamori* [1979]. Since all but two of events with $M_s > 8.0$ have independently (i.e from literature) determined seismic moments, saturation of surface-wave magnitudes should not affect M_w estimation. For events with $M_s \leq 8.0$ the M_w values are very close to corrected M_s values provided by catalog.

The second global dataset is the Preliminary Determination of Epicenters (hereinafter NEIC) catalog, collected by the National Earthquake Information Service (NEIC/USGS) (www.neic.cr.usgs.gov/neis/epic/epic.html). The NEIC catalog is the most complete instrumental data-set of the last thirty years. *Kagan* [2003], in his analysis on global earthquake catalogs, identifies about 50 $M_s \geq 6$ aftershocks present in the NEIC catalog but missing in the CMT (Centroid Moment Tensor) dataset ([*Dziewonski and Woodhouse*, 1983]). About 30 earthquakes of these are not temporally proximate to others CMT earthquakes. The magnitude scale considered is the maximum (M_{max}) among different magnitude values reported. This choice permits to reduce problems coming from saturation of magnitude and is in agreement with NEIC valuation method used in compiling catalog (eighty columns format). We select events occurred from 1 January 1974 to 31 December 2003, with depth $\leq 70km$ and magnitude $M_{max} \geq 6.0$ (3197 events). For most events (about 50%) M_{max} is the surface magnitude (M_s), whereas moment (M_w) and body (M_b) magnitude are considered for 30% and 13% of events, respectively. The magnitude of remaining events belongs to minor (local M_l, duration M_d, energy M_e) or to unknown scales. Clearly some of these magnitude classes have not been uniformly recorded in time. For example, M_w recording practically begin at about 1990, with development of organized network as Centroid Moment Tensor (CMT) system. Moreover the events for which the magnitude scale is unknown mostly occurred in the first decade of time period examined here.

The third catalog that we consider is the Italian parametric seismic catalogue (CPTI) published by the *Working Group CPTI* [2004]. It reports earthquakes, collected independently in different national datasets, with intensity $I_0 \geq V/VI$, occurred from the second century B.C. to 2002. This catalogue is obtained by applying a declustering procedure to an original dataset, consisting in removing all events occurred

in a 90-days and 30-kms spatio-temporal window respect to a shock with a larger magnitude. In order to have a complete catalogue we consider the events with $Mw \geq 5.5$ since 1600 (203 events). This magnitude threshold represents the value above which the earthquakes in the Italian territory are considered dangerous.

4.3 Methodology

4.3.1 Declustering procedure

Our study is directed towards the recognition of long-term and long-range properties of earthquake occurrence. In order to remove the short-term triggering effect, we apply a declustering procedure to PS92 and NEIC catalogs. As above mentioned, the CPTI dataset is already declustered, then we decide to not carry out further filtering procedures. Specifically we consider the declustering stochastic algorithm, used in previous Chapters, proposed by *Zhuang et al.* [2002] and based on the ETAS model [*Ogata*, 1988; 1998]. This modeling describes the total seismic rate as the sum of two contributions: the background rate, that refers to activity which is not triggered by precursory events and is forced by external physical processes, and the rate of events internally triggered by stress variations of previous earthquakes, in agreement with the modified Omori-law [*Utsu*, 1961]. The "background" seismicity is modeled by a stationary Poisson process, of which rate changes with seismogenetic capability of examined zone. In particular we consider an isotropic power law function to model the triggering effect of each event on area surrounding its epicenter. Therefore the total intensity function of model is given by

$$\lambda(t, x, y/\mathcal{H}_t) = \mu(x, y) + \sum_{t_i<t} \frac{K}{(t - t_i + c)^p} e^{\alpha(M_i - M_{min})} \frac{C_{d,q}}{(r^2_{[(x,y),(x_i,y_i)]} + d^2)^q} \quad (4.1)$$

where $\mathcal{H}_t = \{t_i, M_i, (x_i, y_i), \quad t_i < t\}$ is the observation history up time t, M_{min} is the minimum magnitude of catalog and $C_{d,q}$ is normalization constant of triggering spatial function. We set $\mu(x, y) = \nu u(x, y)$, where ν is a positive-valued parameter and $u(x, y)$ is the probability density function (PDF) of locations of background events. The probability π_i that the i-th event, occurred at location (x_i, y_i), belongs to

background activity is estimated as the ratio between the background rate $\mu(x_i, y_i)$ and the total occurrence rate $\lambda(t_i, x_i, y_i/\mathcal{H}_{t_i})$.

To estimate parameters of the stationary ETAS model $(\nu, k, c, p, \alpha, d, q)$ we use the iteration algorithm developed by *Zhuang et al.* [2002] and basing on the Maximum Likelihood Method [*Dalay and Vere-Jones*, 2003]. This procedure also provides a PDF for location of background events $u(x, y)$, by means of a kernel method. Since some physical investigations show that static stress changes decrease with epicentral distance as r^{-3} [*Hill et al.*, 1993], we impose $q = 1.5$. This choice is also justified by the recognized trade-off between parameters q and d, that is different pairs of q and d values provide almost the same likelihood of the model [*Kagan and Jackson*, 2000].

Parameter	**PS92**	**NEIC**
ν	6.7 ± 0.3 year^{-1}	84 ± 2 year^{-1}
K	$(4 \pm 1)10^{-3}$year^{p-1}	$(2.3 \pm 0.3)10^{-3}$year^{p-1}
p	1.1 ± 0.1	1.22 ± 0.02
c	$(2 \pm 1)10^{-4}$ year	$(1.2 \pm 0.3)10^{-4}$ year
α	1.2 ± 0.2	1.6 ± 0.1
d	26 ± 3 km	12 ± 0.5 km
q	$\equiv 1.5$	$\equiv 1.5$
Log-likelihood	-10957.2	-37555.0

Table 4.1: Maximum Likelihood Parameters of the ETAS model for PS92 and NEIC catalogs

We apply the *Zhuang et al.* [2002] procedure to PS92 and NEIC catalogs. The values of parameters, with relative errors, and the maximum likelihood value are reported in Table 4.1. We assign to background seismicity events for which the probability p_i is larger than 0.5. We stress that the choice of this threshold does not affect following results, because about 10% of all events have a probability p_i ranging between 0.1 and 0.9. By this procedure we select 618 and 2563 events for PS92 and NEIC catalogs, respectively.

4.3.2 Definition of the time series

The goal of this study is to provide a testable method to investigate the stationary and long-term memory in occurrence of moderate-large earthquakes. In order to consider the spatial variability of main features of seismogenetic process, due to different tectonic domains, we define a discrete time series $\{R_j^k, k = 1...N\}$ for each node (X_j,Y_j) of a suitable spatial grid. The term R_j^k is the occurrence rate in time period $I_k = [(k-1)\tau, k\tau)$ relative to node (X_j, Y_j). The choice of interval time τ is a balance that accounts for different opposite requirements: it must be large enough to include data needed for calculations, but also small enough to have a sufficient number of data in the time series. To avoid any dependence by subjective decisions, as for examples the definition of prior seismic zoning that dictates the boundaries of seismic activity, we estimate the seismicity rate for each node by smoothing the surrounding seismicity by suitable kernel functions. Much more critical than the choice of kernel function type is the choice of an appropriate bandwidth. This controls the degree of smoothness of data: a too low value of this parameter can produce an overfitting with consequent preservation of very small seismicity features; on the other side a too high bandwidth causes a strong redistribution of data that can lead to blurring of local seismicity properties. Here we use a Gaussian kernel estimate method, similar to procedure proposed by *Zhuang et al.* [2002]. The bandwidth changes with both the spatial occurrence density and the released energy in the region. For each node the mean occurrence rate R_j^k is given by

$$R_j^k = \sum_{i=1...N; t_j \in I_k} \frac{1}{2\pi c_i^2} exp\left(- \frac{r^2_{[(X_j,Y_j);(x_i,y_i)]}}{2c_i^2} \right) \tag{4.2}$$

where $r^2_{[(X_j,Y_j);(x_i,y_i)]}$ is the distance between the node (X_j, Y_j) and the epicenter of i-th event (x_i, y_i). The choice of the local bandwidth c_i is driven by some aspects. It has to be larger of the epicentral error of the event, to avoid a too detailed smoothing respect to quality of data; moreover it must take into account the fault length L_i, to consider the dimension of the source. This last is estimated by magnitude M_i using the empirical relation

$$L_i = \sqrt{10^{1.21M_i - 5.05}} \tag{4.3}$$

in agreement with [*Bath and Duda*, 1964]. A general simple empirical rule, that provides a reliable value of c_i for each event, in agreement with the above requests and independently by quality of data collected in different catalogs, is $c_i = 3L_i$. By this relation the bandwidth c_i, for an event with magnitude ranging from 5.5 to 8.5, varies from few tens to about one thousand kilometers. This is in agreement with assumptions of *Kagan and Jackson* [2000] which, in their analysis of moderate-large seismicity ($M \geq 5.8$) of last tens of years, impose an interaction distance smaller than 1000 km. For larger events ($M \geq 8.5$), c_i has an order of magnitude from one to few thousand of kilometers. By this choice we include the contribution of remote strong events in computation of the seismic rate of a region. The values of bandwidth for these events is in agreement with *Marzocchi et al.* [2003], which to estimate the stress perturbation in a region include contribution of strong events ($M \geq 7.0$) occurred up to 5000 km of distance.

4.3.3 ARFIMA modeling

An essential part in the selection of a suitable model for a given set of observations is to identify dependence structure trough time, deviations from stationarity, and features which constitutes the source of the randomness. We investigate all these topics for the time distribution of large earthquakes by the ARFIMA models [*Box and Jenkins*, 1976; *Brockwell and Davis*, 1987]. The main characteristics of them are reported in Appendix 4A. Here we only stress that, by using these models, we are providing a simple description of what is in fact an extremely complicated situation. The main rationale behind the choice to check their reliability is that they provide a flexible, estimable and interpretable class of models, able to recognize the main features of the system under study. Information coming from this type of analysis can be used in a more sofisticated modeling.

A problem related to the ARFIMA modeling is to define the order of the best model and to avoid overfitting and parameter redundancy [*Brockwell and Davis*, 1987]. Therefore, some care is needed in estimating procedures, to achieve parsimony in parameterization. Specifically, for time series of only a few hundred or less of ob-

servations, the estimated autocorrelation function (ACF) from data generated by an ARMA(p,q) model would be indistinguishable from that of a time series generated by a process AR(p) ($q \equiv 0$). Only with time series of more observations enough information would be available to obtain a good estimate of additional parameters and then the need to account for a more elaborate model could arise. However we stress that the estimators of parameters of ARFIMA models need not be seen as purely "parametric", in the sense of depending crucially on a correct parametrization of the process. The autoregressive structure, which extracts statistical information from the data, is capable of fitting a very general class of processes. Then useful estimates can also be obtained from mis-specified models. In this sense estimating procedures may be thought of as "semi-parametric". The long-memory component is captured via fractional parameter r, and the presence of nonstationarity via the analysis of a variable number of ARMA terms. This last may be increased with sample size, to detect increasingly subtle short-memory features as sample information accumulates.

All methods proposed in the literature to determine the order of an ARMA(p,q) process, such as those based on autocorrelations and partial autocorrelations, cannot be used to determine the order of an ARFIMA(p,d,q) process. In fact when the fractional parameter r is significantly non null, the autocorrelations decay to zero at a slower rate of convergence than the ARMA models; this makes impossible to recognize the short-memory components (see Appendix 4A).

The estimation of parameters for a specified ARFIMA(p,d,q) model is carried out in two levels, in agreement with the fast and accurate method proposed by *Haslett and Raftery* [1989]: an outer univariate unimodal optimization in r over the interval [0,.5], using Brent's algorithm [*Brent*, 1973], and an inner nonlinear least-squares optimization in the AR and MA parameters to minimize white noise variance of residuals. The justification of this two-step procedure is that if X_t is an ARFIMA(p,d=r+i,q) model, then the process $Y_t = (1-B)^r X_t$, obtained by filtering X_t through the fractional difference operator $(1-B)^r$, is an ARIMA(p,i,q) model. The *Haslett and Raftery* [1989] methodology can be used in the non-stationarity region as well; then it is useful in all applications where stationarity is not normally known a priori to apply. Other estimators, in particular the Yule-Walker [*Box and Jenkins*, 1976; *Brockwell and Davis*, 1987], are convenient for obtaining theoretical

properties in the stationary region, only. The procedure proposed by *Haslett and Raftery* [1989] is applied to our time series by using the MINPACK subroutines written by Chris Fraley [*More et al.*, 1980].

The best ARFIMA model is identified by calculating the corrected version of the Akaike Information Criteria (AICc) [*Hurvich and Tsai*, 1989] for the filtered (by r) series Y_t. The AICc statistic is defined by

$$AICc(K) = -2LogL + 2(K)\frac{N}{N-K-1} \tag{4.4}$$

where N is the sample size, $K = p+q+1$ is the number of parameters, and L is the likelihood of the model, computed by recursive optimization procedure proposed by *Haslett and Raftery* [1989] on the process Y_t.

To select the optimal moving average autoregressive orders, we consider that, once the model parameters have been estimated, there should be no structure in the residuals ϵ_t. In particular, we are interested in finding whether the residuals are independent and identical distributed (iid) random variables. Considering above comments on parsimony in parameterization, we firstly consider $p \leq 5$ and $q \leq 1$. Higher ARMA components are taken into account only if this procedure does not provide a model with iid residuals for more than 5% of non null time series. The white noise distribution for residuals is checked by the Ljung-Box (LB) version of the Portmanteau test [*Box and Jenkins*, 1976; *Ljung and Box*, 1978], of which details are described in Appendix 4B. A large value of statistic Q indicates that the sample autocorrelations of residuals are too large to accept null hypothesis. This suggests that there is significant dependence among the residuals and then that we need to look for a more complex time series model.

Once the best model is identified, we check on stationarity of the time series, that is on the presence of unit roots in autoregressive polynomial (see Appendix 4A). Among the procedures for testing unit roots hypothesis available in literature, we consider the Augmented Dickey-Fuller Test (ADF) [*Dickey and Fuller*, 1979; *Said and Dickey*, 1984] (see Appendix 4C). For both ADF and LB tests a significance level of 5% is assumed.

4.4 Testing the ARFIMA algorithm

The estimation of the fractional differencing parameter r is fraught with difficulties. All known methods produce relatively large estimation errors, specially for relatively small samples ($N \sim 100$). Another problem arising with the ARFIMA models is that nonstationary ARIMA processes and nearly nonstationary ARMA processes (such as autoregressive processes having a root of the AR polynomial close to unit circle) have almost indistinguishable sample autocorrelation properties from those of stationary long-memory ARFIMA processes. Then, in order to substantiate interpretation of results of analysis of real data, we test the procedure described in previous subsection on simulated time series. The main rationale behind this check is to verify that the possible identification of long-term persistence is not the result of a misspecification of a nonstationary trend. Specifically we simulate four classes of synthetic datasets, obtained imposing a short/long-term memory and a stationary/nonstationary behaviour. The first two groups include 1000 stationary time series with 100 observations, simulated by the short-term memory model ARMA(1,0) and the long-term memory model ARFIMA(1,0.3,0). For both simulation we consider $\phi_1 = 0.5$. The other two tests are performed on as many time series simulated by two different nonstationary models: an ARIMA(1,1,0) short-term and an ARFIMA(1,1+0.3,0) long-term memory model. For both these last simulations ϕ_1 is equal to $1.$, of course. We assume that residuals are a normal white noise process. To test the goodness of the algorithm respect to the size of residuals, we perform different tests, by changing σ_ϵ from 0.1 to 1000. In Table 4.2 we report some information on empirical distribution of parameter r (the 2.5th, 50th and 97.5 percentiles) and the proportion of time series for which a nonstationary trend is identified.

The first result is that the large estimation error on r can produce a misleading interpretation about the presence or not of long-term memory. Specifically the algorithm strongly undervalues this parameter, especially in the stationary case. However this inefficiency does not significantly influence the recognition of nonstationary trend. The algorithm is able to distinguish stationary and nonstationary time series with a good confidence level ($\sim$ 90%). Moreover the not recognized nonstationary time series have a value of statistic T_{LB} very close (the difference is less than 0.1) to

	ARMA(1,0) ϕ_1=0.5; r=0.0		**ARFIMA(1,0.3,0)** ϕ_1=0.5; r=0.3		**ARIMA(1,1,0)** ϕ_1=1.0; r=0.0		**ARFIMA(1,1+0.3,0)** ϕ_1=1.0; r=0.3	
σ_ϵ	perc. r	% NONST	perc. r	% NONST	perc. r	% NONST	perc. r	% NONST
0.1	(0.0, 0.0, 0.38)	0%	(0.0, 0.0, 0.33)	8%	(0.0, 0.04, 0.25)	91%	(0.0, 0.16, 0.5)	90%
1.0	(0.0, 0.0, 0.38)	0%	(0.0, 0.0, 0.34)	6%	(0.0, 0.03, 0.24)	91%	(0.0, 0.17, 0.5)	89%
10.0	(0.0, 0.0, 0.38)	0%	(0.0, 0.0, 0.36)	7%	(0.0, 0.03, 0.24)	91%	(0.0, 0.18, 0.5)	89%
100.0	(0.0, 0.0, 0.39)	0%	(0.0, 0.0, 0.36)	7%	(0.0, 0.02, 0.25)	92%	(0.0, 0.16, 0.5)	89%
1000.0	(0.0, 0.0, 0.39)	0%	(0.0, 0.0, 0.37)	7%	(0.0, 0.02, 0.25)	92%	(0.0, 0.17, 0.49)	90%

Table 4.2: Test of the algorithm on 1000 syntetic catalogs for four different ARFIMA models. The 2.5th, 50th and 97.5th percentiles of estimated values of r and the proportion of time series recognized as nonstationary are reported.

critical value (-2.89). Then, in real data analysis, a check on T_{LB} statistic can help to corroborate interpretation of results. None influence of residuals variance σ_ϵ is found.

4.5 Results

The procedure described and tested in previous sections is applied to our datasets.

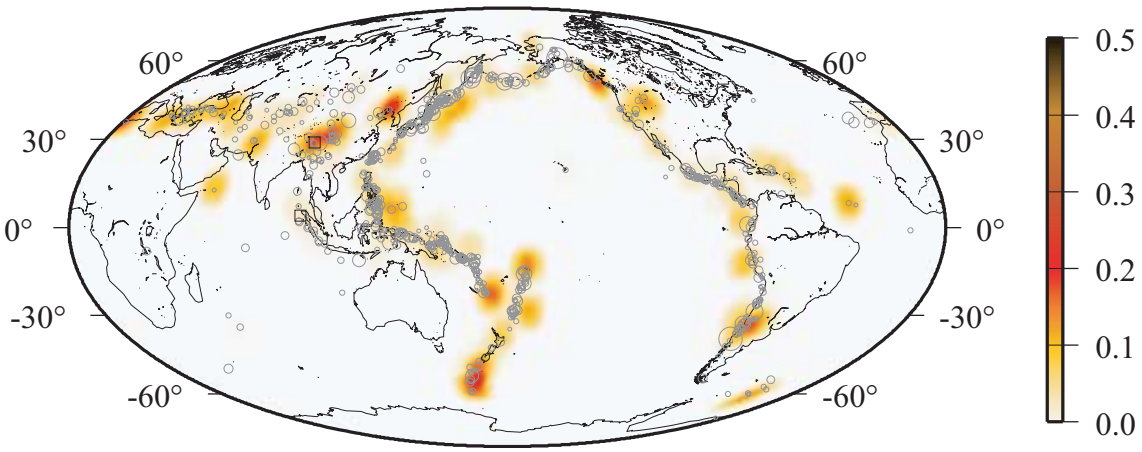

Figure 4.1: Map of the fractionally integrated parameter r of the best ARFIMA model for PS92 catalog. Gray circles indicate locations of events. The black squares mark locations of nodes for which some details of analysis are discussed.

To define time series for PS92 catalog, we set the grid side equal to 300km (5666 nodes) and $\tau = 1$ yr. Therefore in each node of the grid we have a time series of 90 observations. The maximum lag of estimated autocorrelations m, used in calculations, is equal to 20. All non null time series are described by an ARFIMA process, with $p \leq 5$ and $q \leq 1$.

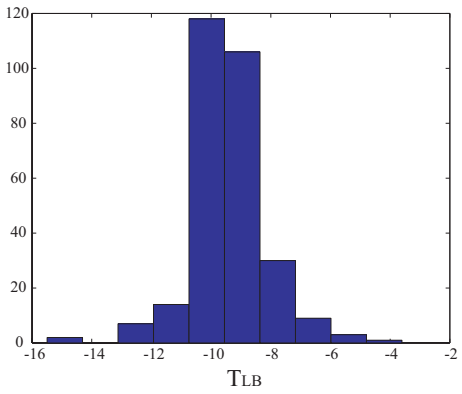

Figure 4.2: Histogram of the T_{LB} values for PS92 time series with the autoregressive order p larger than 0. The critical value for T_{LB} statistic, above which a time series can be considered nonstationary at confidence level of 95%, is -2.89 (sample size 90).

For all time series we accept the null hypothesis of uncorrelation of residuals by the Portmanteau test at 95% confidence level. The estimated values of fractional parameter r highlights significative long term temporal dependence (see Figure 4.1). Specifically we find a clear evidence of long term memory of seismic rate in Mongolia, in Turkey region and in most parts of Pacific Ring, especially in California, in Chile, in Southern-Eastern Asia and in area surrounding New Zealand. In order to investigate if any region shows a nonstationary behaviour, we apply the ADF test (see Appendix 4C). By this we reject the nonstationary hypothesis at 95% confidence level in all zones. The values of statistic T_{LB}, much below the critical value (equal to -2.89 for a sample with 90 observations), remove any doubt on the truth of this deduction (see Figure 4.2). The values of $\sum_{i=1}^{p} \psi_i$, that could be equal to or larger than 1 for a nonstationary time series (see Appendix 4A), not exceed 0.5 and further corroborate results of the ADF test.

The time series relative to NEIC catalog are identified by setting $\tau = 0.5yr$ and by

using the same grid of PS catalog.

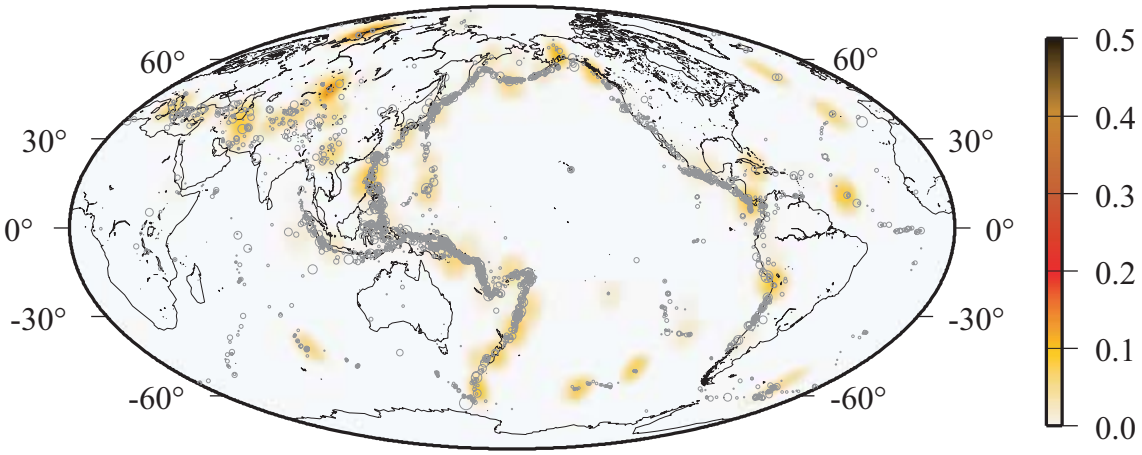

Figure 4.3: Map of the fractionally integrated parameter r of the best ARFIMA model for NEIC catalog. Gray circles indicate locations of events.

For each node the time series has 60 observations, then we consider the maximum lag m for ACF calculations equal to 15. All NEIC time series are described by an ARFIMA process, with $p \leq 5$ and $q \leq 1$, but the autoregressive order is low ($p \leq 2$) in most zones. The estimated values of r are generally rather low and show a very lame evidence of long-term persistence in the Western coast of the Pacific Ring and in interior of the Asia (see Figure 4.3).

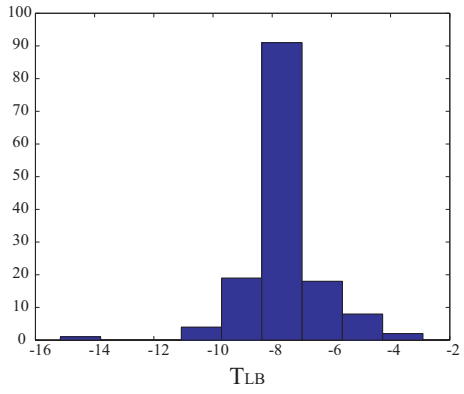

Figure 4.4: Histogram of the T_{LB} values for NEIC time series with the autoregressive order p larger than 0. The critical value for T_{LB} statistic, above which a time series can be considered nonstationary at confidence level of 95%, is -2.91 (sample size 60).

The ADF test does not identify a nonstationary trend in any zone (see Figure 4.4) at 95% confidence level (the critical value for a sample with 60 observations is -2.91.) The values of $\sum_{i=1}^{p} \psi_i$ do not exceed 0.5.

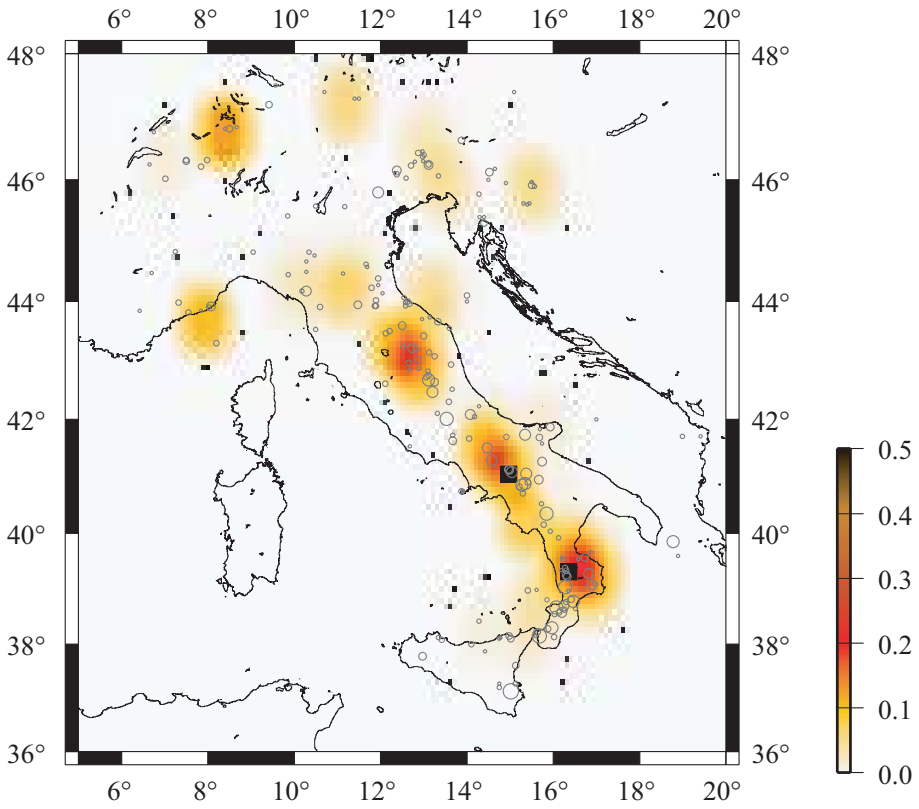

Figure 4.5: Map of the fractionally integrated parameter r of the best ARFIMA model for CPTI catalog. Gray circles indicate locations of events. The squares mark locations of nodes for which some details of analysis are discussed.

To explore the time distribution of historical moderate-large seismicity in Italy (CPTI catalog) we set $\tau = 5yr$ and $m = 20$. The grid size is equal to 50kms and for each node the time series has 81 observations. The estimated values of r highlight two zones in which seismicity is affected by a long-term memory (see Figure 4.5). These was interested by two of the most strong earthquakes of the last century: the Mw6.9 Irpinia earthquake occurred in 1980 and the 1997 Mw6.0 Colfiorito earthquake.

The ADF test does not highlight any nonstationarity (see Figure 4.6): the values of

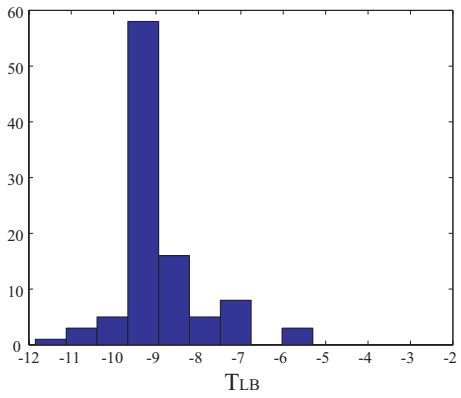

Figure 4.6: Histogram of the T_{LB} values for CPTI time series with the autoregressive order p larger than 0. The critical value for T_{LB} statistic, above which a time series can be considered nonstationary at confidence level of 95%, is -2.9 (sample size 81).

statistic T_{LB} are below the critical value -2.90 and the values of $\sum_{i=1}^{p} \psi_i$ does not exceed 0.4.

To further verify that above results do not come out form a misspecification of a nonstationary ARIMA model with a stationary long-term memory ARFIMA model, we apply again the procedure to all time series by imposing $r = 0$. Some variations are found in estimating autoregressive order, of course, but none nonstationary is recognized in any time series.

4.6 Discussion and Final Remarks

The causes of departures by time independent hypothesis of large earthquake occurrence, recognized at global scale in Chapter III and in some studies [*Kagan and Jackson*, 1991; *Kagan and Jackson*, 1994; *Lombardi and Marzocchi*, 2007], can be explained by two potential factors. The first is that the seismic capability of a region is almost memoryless but its basic properties can change with time, following variations of tectonic rate. By a statistical point of view this is interpreted as a nonstationarity time-dependent behaviour (in agreement with a nonstationary Poisson process; see Table I of Introduction), that is as time variation of basic descriptors (for example the mean rate) of the occurrence history. The process exhibits a time

independent local (short-term) behaviour around a mean level that changes with time (see Figure 4.7a).

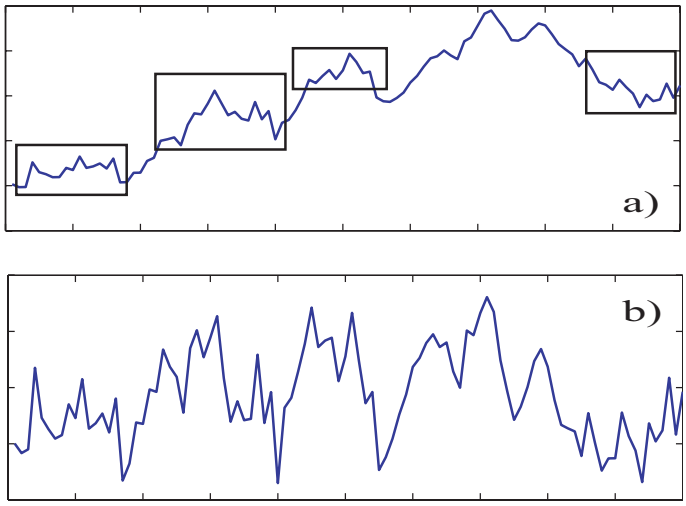

Figure 4.7: Examples of a nonstationary (top) and a long-term memory (bottom) time series.

The second cause is the presence of long-term memory: the seismic rate shows a long-term autocorrelated behaviour, by a longer time respect to typical length of sequences. Then the recognized fluctuations of mean rate are due to a modulation by past history that is not fully explainable with short-term memory (see Figure 4.7b). The present state of the art does not allow us to rule out any one of these possible causes and to identify the driving mechanism of large earthquake occurrence. The class of ARFIMA models presented here provides a general framework for representing seismic time series that could display long-term persistence. Estimation of long-memory models by autoregressive approximation is feasible even in quite small samples. Moreover it is consistent across a range of stationary and nonstationary values of parameters, so that estimation does not require knowledge of an appropriate transformation to apply.

In order to better explain methodology and results, we show in details analysis performed for two PS92 time series. Specifically we consider nodes with coordinates [35.76°S, 151.132°W] (Mongolia) and [3.98°S, 94.81°W] (Indonesia region), of which locations are marked in Figure 4.1 (black squares). We remark that the properties of the time series relative to these nodes are very different. The seismicity of

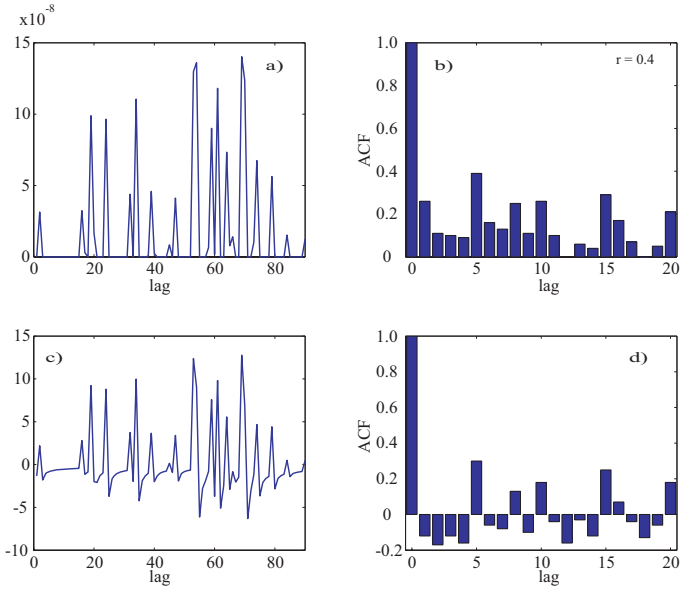

Figure 4.8: Details of analysis for the PS92 time series relative to node [35.76 °S, 151.132°W] (Mongolia). a) Original time series b) ACF for original time series c) Filtered (by r) time series d) ACF for filtered time series.

Indonesia region has not long-term memory features ($r = 0$), whereas the Mongolia is characterized by a higher fractionally parameter r, showing more persistent long-term interactions (see Figures 4.1). For node located in Mongolia (see Figure 4.8) we remark the influence of long-term memory on the ACF of the original time series (Figure 4.8b), above all evident for the ever-positive values. Neither the filtered (by parameter r) time series (Figure 8c) nor the relative ACF (Figure 4.8d) show evidence for a nonstationary behaviour.

Figure 4.9 shows the plots and the ACF both for original as for filtered time series relative to node located in Indonesia. The ACF for the original (Figure 9b) time series does not highlight long-term memory nor nonstationarity. The behaviour of time series could suggest a periodic trend, justified also by the higher value of the ACF for lag 14. However any statistical significance is found to confirm this speculation.

The ARFIMA modeling of PS92 and NEIC catalogs highlights substantially the same regions with a long-term memory trend. The agreement of results for two

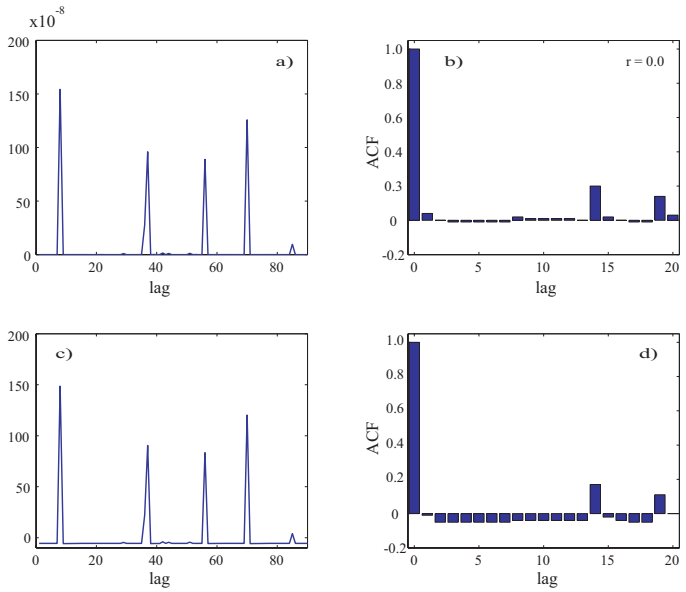

Figure 4.9: Details of analysis for the PS92 time series relative to node [3.98 °S, 94.81°W] (Indonesia region). a) Original time series b) ACF for original time series c) Filtered (by r) time series d) ACF for filtered time series.

so different datasets, a basically macroseismic catalog (PS92) and an instrumental (NEIC) catalog, relative to different seismicity level and different temporal domains, strengthens our belief on significance of our procedure. We stress that the regions identified by our algorithm are not the areas with the higher number of events. This is above all clear in Central America, a highly seismogenetic region, where no evidence of a long-term memory is found. The long-term memory effect for NEIC time series seems to be much more weak respect to PS92 dataset. This result could be ascribed to lower magnitude of events and/or to shorter time period recovered by NEIC dataset (30 years) respect to PS92 dataset. This could mean that the long-term memory is a phenomenon above all related to strong seismicity, with a characteristic time (of decades) comparable with the length of time period recovered by NEIC catalog, that could be inadequate to fully highlights it.

The applications of ARFIMA procedure to global seismicity provides interesting cues to a more conscious modeling of moderate-strong seismicity. The interpretation of our results is not without difficulties. To avoid a misunderstanding, we stress

that, in terms of earthquake occurrence, the ”memory” of ARFIMA modeling does not necessary imply a dependence of the rate by past occurrence time history, but simply its modulation in time. The identification of a significative long-term persistence in autocorrelations means that the rate at time t is correlated with seismicity rate of following interval times $t+i$, also for large lags i, more than a short-memory process. However the recognition of none nonstationarity, in time series recovering very different temporal ranges (from few decades to centuries), denotes a substantial stability of tectonic motion, at least on temporal scales considered in the present study. This implies very low fluctuations of tectonic loading that shows itself in a steady mean regional seismic rate. This result is in agreement with the geodetic analysis of *Sella et al.* [2002] that show a stable plate velocity on the time scale of thousands of years. For the same reason it seems unlikely that the recognized long-term autocorrelation is due to an ”internal” time modulation of tectonic motion. More probably it could be an evidence of long-term interactions between events. The recognized long-term memory would come out by influence that events have on following seismicity.

In terms of earthquakes occurrence, this result could be interpreted as an evidence of postseismic stress variations due to viscoelastic readjustment of the lower-crust and of the mantle. The recognition of long-term memory corroborates results obtained in the past both by physical as by statistical studies. For instance the interior of the Asia seismicity is one of the more clear examples of long-term (in time and space) fault interaction [*Chéry et al.*, 2001b; *Pollitz et al.*, 2003]. Specifically the Mongolia region has experienced a cluster of some exceptionally large earthquakes in about 60 years (from 1905 to 1967), that has been ascribed to postseismic stress perturbations. Similar features have been recognized in California region [*Selva and Marzocchi,* 2005; *Freed and Lin,* 2001] or in Chile [*Casarotti and Piersanti,* 2005].

Departures by assumption of long-time independence in occurrence of strong events have been recognized in previous Chapter [see also *Lombardi and Marzocchi,* 2007]. By a suitable statistical tests on the declustered PS92 catalog, we have shown evidence for a nonstationary behaviour of occurrence rate in some regions. To justify the different interpretations between that and the present studies (i.e. the strong earthquake occurrence is a nonstationary or a long-term memory process), we stress

that in Chapter 3 the occurrence rate of a region was modeled by a nonstationary Poisson process. This means that the independence by past history (lack of memory) was automatically assumed (see Table 1 in Introduction) and, consequently, the variations of seismic rate were ascribed to a nonstationary behaviour. On the contrary, in the present study, we do not formulate any hypothesis on the main features of the seismic occurrence, but we get these by a suitable analysis.

The finding about the historical seismicity of Italy opens up new prospects for seismic hazard assessment of this region. The commonly recognized long memory features is in contrast with the usual procedure adopted in seismic hazard calculations, based on time-independence of large earthquake occurrence.

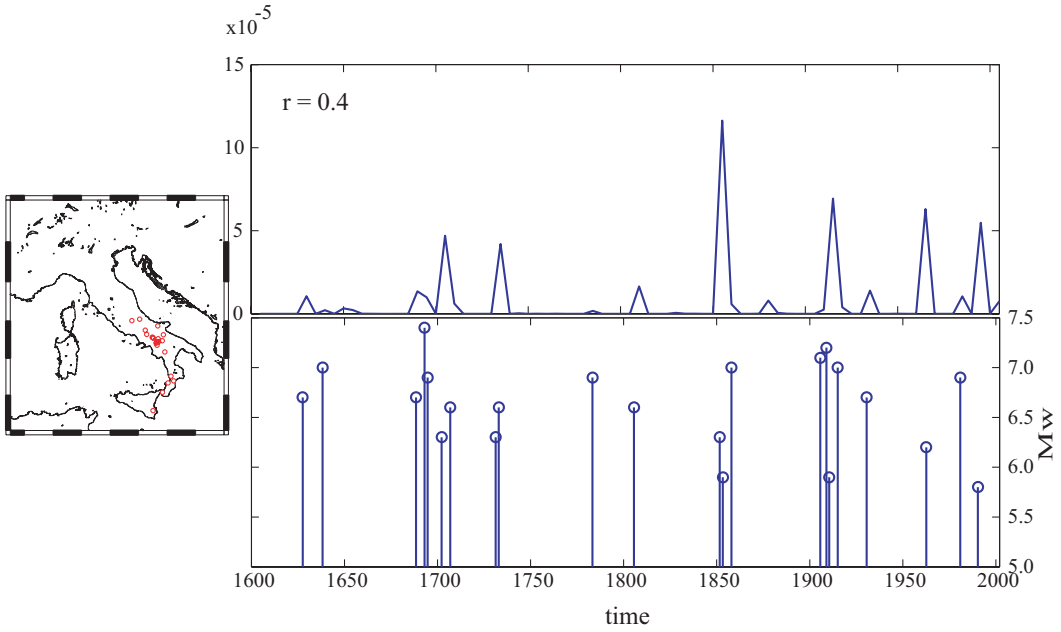

Figure 4.10: Details of analysis for the CPTI time series relative to node [41.04 °N, 14.97°W] (Irpinia region). Time series history and locations of events occurred at distance lower than $2c_i$ by node are shown.

To better understand the meaning of our results, we show time histories relative to two nodes (of which locations are marked in Figure 4.5): the first, with coordinates [41.04°N, 14.97°W], is located in Irpinia region (see Figure 4.10), that has been interested by an earthquake with Mw=6.9 in 1980; the second, in Calabria region,

has coordinates [38.85°N, 16.25°W] and is near the epicenter of one of the most strong earthquakes occurred in Italy in last centuries (1905, Mw=7.1) (see Figure 4.11).

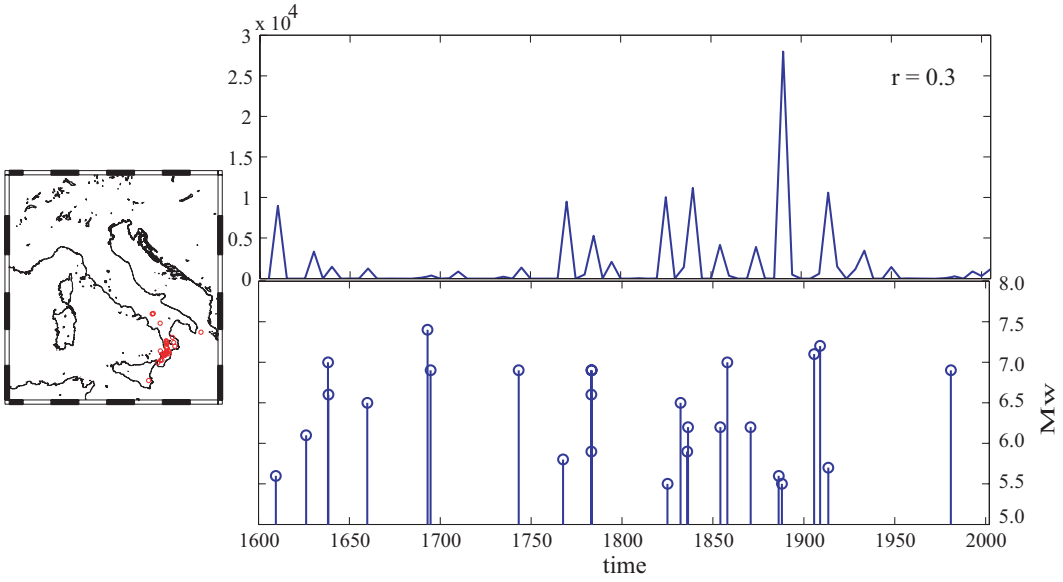

Figure 4.11: Details of analysis for the CPTI time series relative to node [38.85 °N, 16.25°W] (Calabria region). Time series history and locations of events occurred at distance lower than $2c_i$ by node are shown.

By comparing time histories and time series R_j^k for two nodes, we can see that seismicity occurred in surrounding areas appear to have a clustering trend that could justify the recognized long-term memory. Similar features have been already identified on the same catalog, by a nonparametric stochastic modeling, by *Faenza et al.* [2003].

As final consideration we remark that the finding reported here should be regarded only as a first attempt towards a new time-dependent modeling of large earthquake occurrence. Here we have only deepen which are the main long-term features of moderate-strong seismicity on different spatio-temporal domains, considering information coming from seismic catalogues. Our calculations seem rule out any change in occurrence rate explainable by a nonstationary trend, at least on tempo-

ral scales recovered by dataset used in our analysis. The evidence of a memory in the earthquake occurrence, not explainable with short-term triggering, could be the starting point for formulation of a new and more acquainted long-term modeling of seismicity in a region.

4.7 Appendix 4A: ARFIMA MODELS

Let $\{X_t\}$ be an observable time series and ϵ_t be a white noise process, i.e. a sequence of iid random variables $\{\epsilon_1, \epsilon_2, ..., \epsilon_N\}$, with mean zero and variance σ_ϵ^2. The *autoregressive and moving-average* process (ARMA) of order (p,q), [*Box and Jenkins*, 1976; *Brockwell and Davis*, 1987], has the form

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q}. \tag{4.5}$$

If the process is stationary, the parameter μ is the mean about which the process varies. Otherwise μ could have no specific meaning except as a reference point for the "level" of the process. It will be assumed throughout (unless the mean is believed a priori to be zero) that the data have been "mean-corrected" by subtraction of the sample mean, so that the overall constant μ is equal to zero. Let B denote the *backwards shift operator* $BX_t = X_{t-1}$; hence $B^m X_t = X_{t-m}$. The previous formula can be written

$$\Phi(B) X_t = \Theta(B) \epsilon_t \tag{4.6}$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ represent the polynomials $\Phi(z) = 1 - \phi_1(z)... - \phi_p(z^p)$ and $\Theta(z) = 1 - \theta_1(z)... - \theta_q(z^q)$.

To ensure stationarity, the roots of $\Phi(B) = 0$ must lie outside the unit circle [*Box and Jenkins*, 1976]. In the case of stationary model the autocorrelation function (ACF) will quickly "die out", for moderate and large lags k. Thus the ARMA models can be called short-memory models, because a shock affects the level of the series for a relatively short time.

A natural way of obtaining nonstationary processes would be to relax above restriction. If the roots lie inside the unit circle the process exhibits explosive nonsta-

tionary and in particular do not vary about a fixed mean. If i of the zeros of the polynomial $\phi(B)$ are unity, i.e. $1 - \sum_{j=1}^{p} \phi_j = 0$, and the remainder lie outside the unit circle, the series is nonstationary but exhibit homogeneous behaviour of a kind. Specifically the level about which fluctuations occur may be dependent on time, but the broad behaviour of the series, when differences in level are allowed for, may be similar. Then we can express the model (4.6) in the form

$$\Phi(B)X_t = \psi(B)(1-B)^i X_t = \Theta(B)\epsilon_t \tag{4.7}$$

where $\psi(B)$ is a stationary autoregressive operator.

Let ∇ be the *backward difference operator* $\nabla X_t = X_t - X_{t-1} = (1-B)X_t$, we can write the model (4.7) as:

$$\Phi(B)X_t = \psi(B)\nabla^i X_t = \Theta(B)\epsilon_t \tag{4.8}$$

The process defined by (4.8) is called an *autoregressive integrated moving average (ARIMA) process*, of order (p, i, q) and provides a powerful model for describing stationary ARMA ($i = 0$) and homogeneous nonstationary ($i \neq 0$) time series. In practice i is usually 0,1, or at most 2. In short, if we use ARIMA processes for modeling nonstationary time series, the nonstationarity leads to the presence of unit roots in the autoregressive polynomial. In other words, the series X_t is nonstationary, but its i-th difference can be represented by a stationary, invertible ARMA(p-i,q) process.

The standard ARIMA model can be generalized by allowing the order of differencing to take nonintegral values. Let $d = r + i$ be any real number, where i is a non-negative integer and $0 \leq r \leq 0.5$. The fractional difference of the process X_t is defined by the binomial expansion

$$\nabla^r = (1-B)^r = \sum_{j=0}^{\infty} \pi_j B^j \tag{4.9}$$

where

$$\begin{cases} \pi_0 = 1 \\ \pi_j = \prod_{0<k\leq j} \frac{k-1-r}{k}, \quad \text{if } j = 1, 2, \ldots \end{cases}.$$

Then the process X_t follows a *fractionally integrated autoregressive moving average (ARFIMA) process* if $\nabla^d X_t$ is an ARMA(p,q) process [*Granger and Joyeux*, 1980; *Hosking*, 1981]. In other words, after rth filtering, the process $Y_t = (1 - B)^r X_t$ is an ARIMA(p,i,q) process. Alternatively the process obtained by i-th differencing of X_t is a stationary ARFIMA(p,r,q) model. Specifically the two terms i and r of fractionally parameter d are related to nonstationary behaviour and long-term persistence of the time series, respectively. The ARMA and ARIMA models can be thought of as particular cases of ARFIMA models having $d = 0$ and $r = 0$, respectively. When $r \neq 0$, but all the roots of $\Psi(B)$ lie outside the unit circle ($i = 0$), X_t is stationary but with much more slowly decreasing ACF than for an ARMA process. *Hosking* [1981] showed that the correlation function ρ_k of an such ARFIMA process is proportional to k^{2r-1} as $k \to \infty$. Consequently, the autocorrelations of the ARFIMA process decay hyperbolically to zero as $k \to \infty$, in contrast to the faster exponential decay of a stationary ARMA process. This means that there is a strong association between observations widely separated in time. Then the ARFIMA processes are said to have "long memory", in contrast with "short memory" stationary ARMA processes, whose ACF converges to 0 rapidly. The higher the value of r, the higher the intensity of long memory displayed by the model. In practice, to obtain a useful estimate of the autocorrelation structure, the time series would have at least N=50 observations and the maximum lag of the estimated autocorrelations would not be larger than $N/4$ [*Box and Jenkins*, 1976].

4.8 Appendix 4B: The Portmanteau Test

The Portmanteau test [*Box and Jenkins*, 1976] can be used for checking the adequacy of an ARFIMA (p,d,q) model. In particular it permits to detect departure by assumption that residuals come from an iid sequence. To fill inefficiencies of the original test, different versions and a new breed of statistics have been proposed. We choice the Ljung-Box version [*Ljung and Box*, 1978; *Brockwell and Davis*, 1987] because it performs better if the autoregressive order is understated. Let $\{\rho_k$, k=1,2,...K$\}$ be the first K sample autocorrelations of the residuals ϵ_t from any ARFIMA (p,d,q) model and let N be the number of values in time series. It is

possible to show that, if the fitted model is appropriate, the statistic

$$Q = N(N+2)\sum_{k=1}^{K}\frac{\rho_k^2}{N-K} \tag{4.10}$$

is approximately distributed as $\chi^2(K-p-q)$. We therefore reject the iid hypothesis at level α if $Q > \chi^2_{1-\alpha}(K-p-q)$, where $\chi^2_{1-\alpha}(K-p-q)$ is the $1-\alpha$ quantile of the chi-squared distribution with $K-p-q$ degrees of freedom.

4.9 Appendix 4C: The Augmented Dickey-Fuller Test

A crucial point of time series analysis by ARIMA(p,i,q) models is to detect the degree of differencing i. This means to test for the presence of one or more unit roots in the autoregressive polynomial, in order to decide whether or not a time series should be differenced to achieve the stationarity. A systematic approach on this topic was proposed by [*Dickey and Fuller*, 1979].

Let us consider an AR(p) model (q=0) (see equation (4.5)). It can be rewritten as the following regression model:

$$\nabla X_t = \phi_1^\star X_{t-1} + \sum_{i=1}^{p-1}\phi_{i+1}^\star \nabla X_{t-i} + \epsilon_t \tag{4.11}$$

where $\phi_1^\star = \sum_{i=1}^{p}\phi_i - 1$ and $\phi_j^\star = -\sum_{i=j}^{p}\phi_i$. Since the autoregressive polynomial will have a unit root if $\sum_{i=1}^{p}\phi_i = 1$, the presence of such a root is formally equivalent to the null hypothesis $\phi_1^\star = 0$. The Augmented Dickey-Fuller Test is based on estimation of the coefficient $\phi_1^\star$ and the corresponding statistic [*Brockwell and Davis*, 1987]:

$$T_{LB} = \frac{\phi_1^\star}{SE(\phi_1^\star)} \tag{4.12}$$

where the estimated standard error $SE(\phi_1^\star)$ of $\phi_1^\star$ is given by:

$$SE(\phi_1^\star) = \sqrt{\frac{(n-p-2)^{-1}\sum_{j=p+1}^{n} \nabla X_t - \phi_1^\star X_{t-1} - \phi_2^\star \nabla X_{t-1} - ... - \phi_p^\star \nabla X_{t-p+1}}{\sum_{j=p+1}^{n} X_{t-1}^2}}. \tag{4.13}$$

Dickey and Fuller derived the limit distribution of statistic (4.12) under the unit root assumption $\phi_1^\star = 0$, from which a test of the null hypothesis $H_0 : 1 - \sum_{i=1}^{p} \phi_i = 0$ can be constructed. The quantiles of the limit distribution of T_{LB} have been computed by *Fuller* [1976] (see Table 4.3).

Sample size	**Probability of a smaller value**		
N	0.01	0.05	0.1
25	-3.72	-2.99	-2.63
50	-3.57	-2.92	-2.60
100	-3.50	-2.89	-2.58
250	-3.46	-2.87	-2.57
500	-3.44	-2.87	-2.57

Table 4.3: Critical values of the ADF test.

Specifically the augmented Dickey-Fuller test reject the null hypothesis of a unit root at level 0.05 if $T_{LB} < -2.89$. *Said and Dickey* [1984] showed that the Dickey-Fuller procedure, which was originally developed for autoregressive representations of known order, remains valid asymptotically for a general ARIMA process in which autoregressive and moving-average orders are unknown.

Chapter 5

Towards a new long-term time-dependent stochastic modeling of earthquake occurrence

5.1 Introduction

Results presented in Chapters 3 and 4 suggest that a time-dependent model could better describes, respect to stationary Poisson process, the long-term time distribution of moderate-strong seismicity of a region. In particular the recognized departures by the Poisson model seem to be due to the influence that events have on following seismicity (long-term memory). This finding suggests that the time-dependent rate $\lambda(t)$ of a reliable point process modeling the long-term occurrence of events, would must depend on past occurrence times t_i, too (see Table 1 in Introduction).

Traditional probabilistic seismic hazard and long-term forecasting models are rarely time-dependent, except perhaps for description of aftershock sequences or single source behaviour. However phenomena such as main shock clustering and stress shadows following large earthquakes imply that they should be. Such time-dependent effects are becoming more widely recognized both by physical [*Chéry et al.*, 2001a, 2001b; *Mikumo et al.*, 2002; *Pollitz*, 1992; *Pollitz et al.*, 1998; *Pollitz et al.*, 2003; *Rydelek and Sacks*, 2003; *Corral*, 2004, 2005; *Santoyo et al.*, 2005; *Thatcher*, 1983;

Piersanti et al., 1995; 1997; *Piersanti*, 1999; *Kenner and Segall*, 2000] as by probabilistic [*Kagan and Jackson*, 1994, 2000; *Faenza et al.*, 2003; *Rhoades and Evison*, 2004] analyses. Moreover many historical catalogues include examples of pairs or short sequences of large events, separated by distances of some hundreds of kilometers or greater, which occur within time periods which are short by comparison with the overall mean time between events [*Kagan and Jackson*, 1999]. Nevertheless, despite of massive literature on earthquake forecasting and hazard assessment, there are very few procedures based on models from which conditional intensities or probabilities can be calculated at long time-scale, and that, at the same time, have been fairly tested on data.

Whereas the physical process adduced as the main cause of short-term earthquake clustering is the static stress change [*Stein*, 1999; *King and Cocco*, 2001], on longer timescales or at greater distances viscoelastic readjustments may be important in triggering other main shocks, outside the traditional aftershock zone [*Pollitz*, 1992; *Piersanti et al.*, 1995]. Several methods, mainly based on stress transfer, have been developed to estimate the change in probability of an earthquake elsewhere, caused by occurrence of an event [*Toda et al.*, 1998; *Parson et al.*, 2000]. However in complex faults networks the stress interactions over time can be very intricate and the following probability estimates very uncertain.

Here we propose a new stochastic model, based on the ”self-exciting” modeling [*Daley and Vere-Jones*, 2003]. Despite of its simplicity and lack of strong physical based constraints, it fits moderate-strong seismicity on long time scale better than the stationary Poisson model, opening new perspectives towards the time-dependent hazard assessment.

5.2 The model

In the early 1970s, *Hawkes* [1971] introduced a family of point processes, called ”self-exciting” models, which became first examples of models of general utility for the description of seismic catalogues. The rate $\lambda(t)$ of this class of models is given by

$$\lambda(t) = f(t) + \sum_{t_i<t} g(t - t_i) \tag{5.1}$$

where $f(t)$ and $g(t)$, in seismological context, represent the background rate and clustering density, respectively, and are usually given in parametric form. The process has a branching process interpretation. Any single event can be thought of as the ”parent” of a family of later events, its ”offspring”, which ultimately die out. Early applications to seismic data are in [*Hawkes and Adamopoulos*, 1973; *Kagan Knopoff*, 1987], but this model has been greatly improved and extended by *Ogata* [1988; 1998] in formulation of the ETAS model.

Until now models of this kind have been used to describe short-term clustering of earthquake occurrence. The common use of the Omori law, as parametric form of the function $g(t)$, restricts the forecasting capability of these model to short-term triggering. On longer time scale the predictive power is based on the form of background rate, usually assumed constant, in agreement with a stationary Poisson process. However this class of models are based on an intuitive motivation that could be extended to larger spatio-temporal scale respect to that affected by short-term triggering. The ”self exciting” processes involve two components: the locations of clusters and the locations of elements within a cluster; if we drop assumption that these last coincide with aftershocks, we could capture interactions between events, mainly due to transfer of stress, involving larger space-time scales.

In anticipation of a global model able to capture basic features of overall seismicity, we propose a model suitable for ”declustered” catalogs. The main reason behind this choice is to test if the seismicity of a region, filtered by short-term interactions, shows statistically significative subtler features that may suggests a time-dependent trend. Moreover the analysis of declustered records is in agreement with seismic hazard practice. We stress that declustering procedures is not so crucial for historical catalogs as for instrumental ones. Whereas the latter have a large proportion of short-term triggered events, the former rarely contain more than few events into a sequence.

To establish some constraints on time evolution of the damping factor, that controls the long-term stress transfer, we compare performance of two functions. The

first is an inverse exponential distribution $e^{-t/\tau}$, where t is the elapsed time from the occurrence of the earthquake that generates stress variations. This choice is a simplification of model proposed by [*Piersanti et al.*, 1995; 1997] to describe temporal evolution of postseismic stress variations. The relaxation time τ depends on the viscosity of the mantle, for which very different values, ranging from 5x10^{17} Pa·s [*Pollitz et al.*, 1998] to 5x10^{20} Pa·s [*Piersanti*, 1999] have been proposed. The second function is a power function $1/(t+\eta)^{\rho}$, similar to Omori law.

To describe the spatial decay of stress variation, with distance r from epicenter of perturbing event, we chose an inverse power law probability distribution function $c_{d,q}/(r^2+d^2)^q$. The dependence of hazard with the magnitude M of exciting event is assumed of exponential type, i.e. proportional to $e^{\alpha M}$. In agreement with *Kagan and Jackson* [1994], we explore anisotropy of triggering effect along the strike direction, by multiplying the spatial distribution function by the orientation function $1+\delta\cos^2(\phi_i(x,y)$, where $\phi_i(x,y)$ is the angle subtended at (x_i,y_i) between the fault strike and the location (x,y). This relation predicts that the maximum perturbation is in the direction of the earthquake rupture.

Therefore the time-dependent conditional rate of earthquake occurrence for the first model (Model I) is given by

$$\lambda_1(t,x,y/\mathcal{H}_t) = \nu u(x,y) + \sum_{t_i<t} e^{\alpha(M_i-M_{min})} K e^{-\frac{(t-t_i)}{\tau}} \frac{c_{d,q}}{(r^2+d^2)^q} \qquad (5.2)$$

and for the second model (Model II) by

$$\lambda_2(t,x,y/\mathcal{H}_t) = \nu u(x,y) + \sum_{t_i<t} e^{\alpha(M_i-M_{min})} \frac{K}{(t-t_i+\eta)^{\rho}} \frac{c_{d,q}}{(r^2+d^2)^q}. \qquad (5.3)$$

where r is the distance between location (x,y) and epicenter of i-th event (x_i,y_i) and M_{min} is the minimum magnitude of the catalog. In agreement with the branching modeling, our model ascribes the occurrence of strong events to the superposition of two effect: the tectonic loading, assumed stationary and with a constant spatially variable rate, and the long-term stress transfer produced by earthquakes.

To estimate the parameters of our models we following the method outlined by *Zhuang et al.* [2002] for the ETAS model. It is based on the Maximum-Likelihood

Method and on the estimation of the total seismic mean rate ($\hat{m}(x,y)$). Specifically this is obtained by smoothing observed seismicity with a Gaussian kernel function K_{d_j} , having a variable bandwidth d_j for each event j (see [*Zhuang et al.*, 2002] for details). The probability ρ_j that the j-th event belongs to "background" activity is given by

$$\rho_j = \nu u(x_j, y_j)/\lambda_2(t_j, x_j, y_j). \tag{5.4}$$

Therefore the "background" seismicity rate $\hat{\mu}(x,y)$ and the clustering ratio $\hat{c}(x,y)$, that is the ratio between the spatial triggering intensity divided by the total intensity, at location (x,y) can be given by

$$\begin{aligned} \hat{\mu}(x,y) &= \frac{1}{T}\sum_j \rho_j K_{d_j}(x-x_j, y-y_j) \\ \hat{c}(x,y) &= \frac{\hat{m}(x,y) - \hat{\mu}(x,y)}{\hat{m}(x,y)} \end{aligned} \tag{5.5}$$

where T is the length of the observation period.

5.3 Application to PS92 catalog

The first application of time-dependent models proposed is carried out on world-wide large earthquakes, occurred from 1900 to 1990, with $M_s \geq 7.0$ and depth $d \leq 70$ km, collected by the *Pacheco and Sykes* [1992] catalog. The total number of events is 698. To avoid the problem of saturation of the surface magnitude scale, we consider the moment magnitude M_w for events with $M_s > 8.0$; this is obtained by seismic moment, using the relation of *Hanks and Kanamori* [1979].

The focal parameters for 588 of events collected into PS92 catalog have been estimated by [*Selva and Marzocchi*, 2004]; the fault plane is selected for 537 events. We use these results to apply the time-dependent models proposed and specifically to estimate the anisotropy of stress variation along the strike direction. For events without any estimation of focal mechanism parameters, we consider an isotropic stress transfer (δ=0). In order to filter the seismicity by short-term triggering effect, we decluster the catalog using the method proposed by *Zhuang et al.* [2002] and

based on ETAS modeling (see Chapter 3). The remaining events are 618.

Parameter	**Poisson model**	**Time-dependent isotropic triggering model I ($\delta \equiv 0$)**	**Time-dependent anisotropic triggering model I ($\delta \neq 0$)**
ν	6.9 ± 0.3 year^{-1}	3.8 ± 0.3 year^{-1}	2.8 ± 0.2 year^{-1}
K		0.016 ± 0.003	0.00011 ± 0.00002
τ		52 ± 15 year	33 ± 5 year
α		~ 0.0	~ 0.0
d		63 ± 7 Km	110 ± 10 Km
q		$\equiv 1.5$	$\equiv 1.5$
δ		$\equiv 0.0$	470 ± 60
Log-likelihood	-10627.6	-10505.7	-10449.1

Table 5.1: Maximum Likelihood Parameters of the Poisson process and of our time-dependent, isotropic and anisotropic, triggering model I for declustered PS92 catalog

By applying the Maximum Likelihood procedure we find that Model II is not significantly better than Poisson Model. On the other side the Model I considerably improves description of data respect to Poisson Model. Some runs of the algorithm used makes clear a trade-off between parameters related to spatial distribution of triggered events (d, q, δ): different combinations of these give about the same likelihood value. Then we choose to fix the parameter q equal to 1.5.

To explore significance of supposed anisotropic behavior of triggering effect, we compare results obtained in two different runs, in which δ is fixed to 0 (isotropic triggering) or is included in the set of parameters to be estimated. The values of parameters obtained and the relative errors are reported in Table 5.1. The supposed anisotropy of stress transfer is supported by larger log-likelihood of anisotropic model. In addition to the significative role of the parameter δ, these results show two other important features of global strong seismicity. The first is that the time τ, that drives time decay of triggering effect, is about equal to 30 years. Interpreting this parameter as the relaxation time of the postseismic stress transfer, its estimation can help to establish some constraints on the mean value of the astenosphere viscosity. The second result to stress is the negligible dependence of the rate of

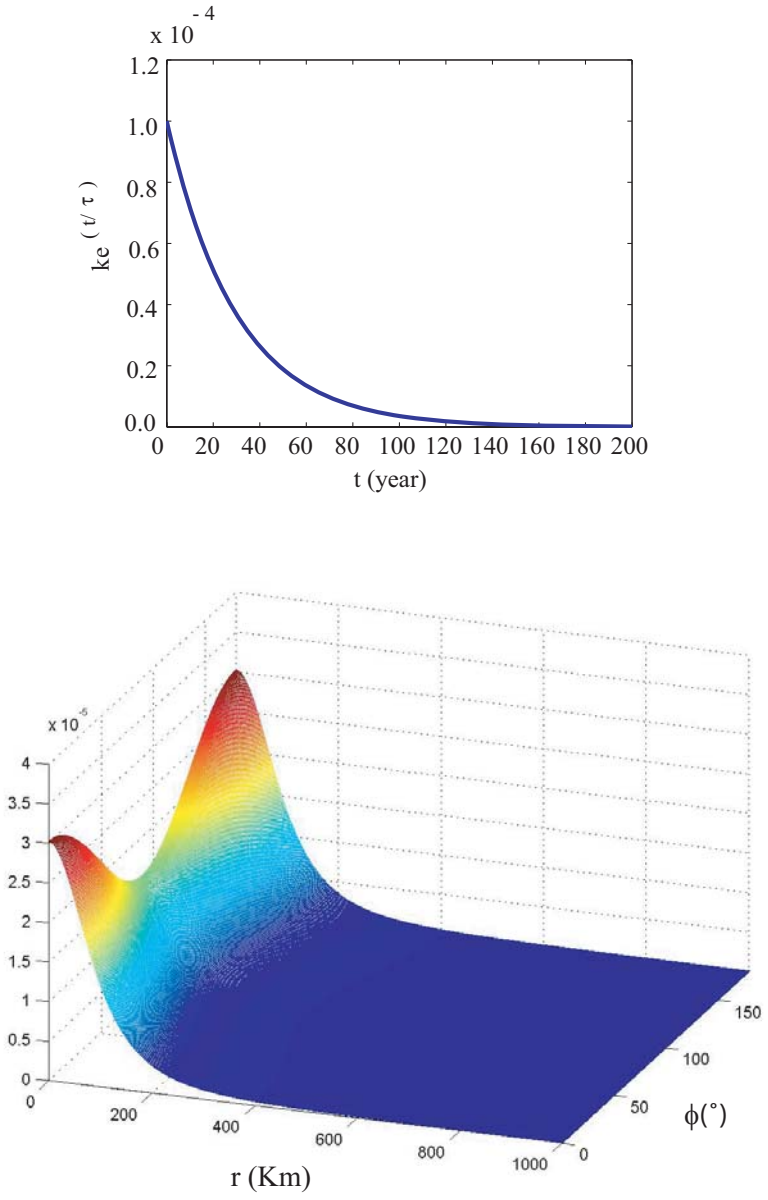

Figure 5.1: Time (top) and space (bottom) decay of the triggering effect described by time-dependent model I for PS92 catalog

induced events by the magnitude of the triggering earthquake (α=0). This result can be due to small range of magnitude recovered by the PS92 catalog and however deserves more investigation. The time and space decay of triggering rate (see Figure 5.1) shows that the occurrence of an event can affect following seismicity, for about one century and up to some hundreds of kilometers of distance. Moreover the

anisotropy of triggering effect becomes almost negligible for distance larger than about 200km.

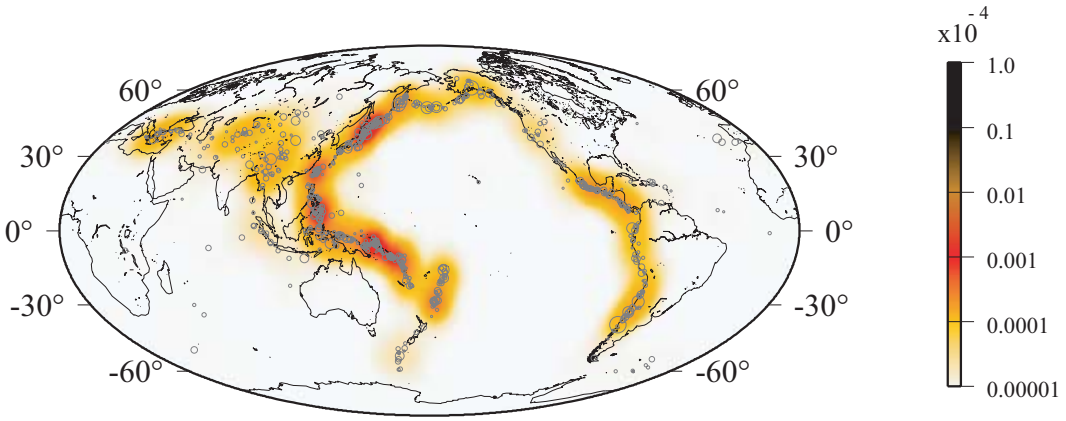

Figure 5.2: Map of background seismicity rate $\hat{\mu}(x, y)$ for PS92 catalog

The maps of background seismicity rate $\hat{\mu}(x, y)$ and of the clustering ratio $\hat{c}(x, y)$ show that the tectonic loading and the stress transfer are not correlated physical processes (Figures 5.2 and 5.3). This means that in some regions the seismogenetic capability is lower that in others, but, once an event occurs, there is a high probability to have other events in a suitable space-temporal window. Remarkably the higher values of background rate are obtained for the Western coast of the Pacific ring. Lower values are obtained for as active areas, as, for example, Central America, Chile and New Zealand region.

The zone in which the triggering rate is larger than the background rate is the interior of Asia. The seismicity of this area is adduced as one of the more clear examples of long-term (in time and space) fault interaction [*Chéry et al.*, 2001b; *Pollitz et al.*, 2003]. In other significative zones, as the Pacific Ring or the Mediterranean Basin, the contribution of triggering effect is about 40%.

Given an earthquake catalog, we can compute the probability that an event i had promoted a following j-th event, occurred at distance r, by:

$$\rho_{ij} = \frac{e^{\alpha(M_i - M_{min})} \dfrac{K}{(t_j - t_i + \eta)^{\rho}} \dfrac{c_{d,q}}{(r^2 + d^2)^q}}{\lambda_2[(t_j, x_j, y_j)/\mathcal{H}_t]}. \tag{5.6}$$

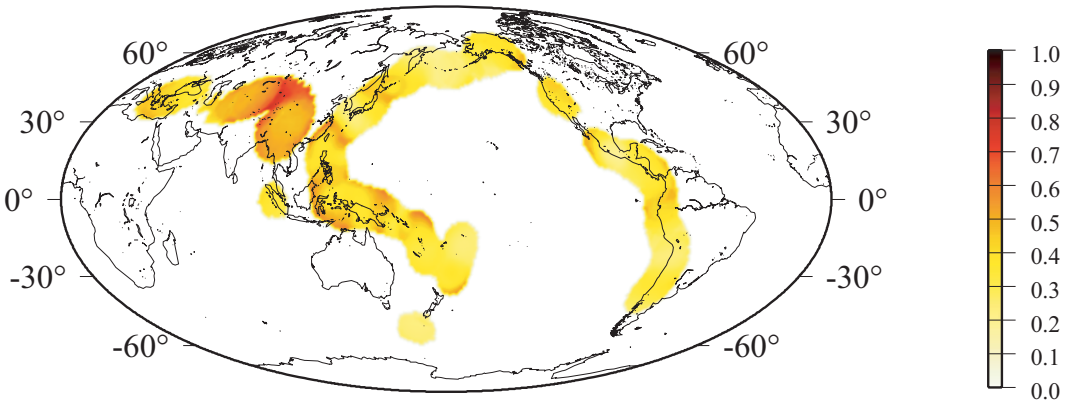

Figure 5.3: Map of the clustering ratio $\hat{c}(x, y)$ for PS92 catalog

Remarkably, by using this equation, we can quantify the probability of interaction between events that are been adduced as single examples of coupling in different previous studies. The increase of seismicity detected between 1905 and 1967 in the interior of Asia is one of the more clear examples of clustering of large events. This period of very high seismicity level has been correlated with postseismic regional stress evolution due to viscoelastic relaxation of the lower crust and upper mantle [*Chéry et al.*, 2001b; *Pollitz et al.*,2003]. Our findings are in agreement with these results and identify significative relations between events occurred in this area, in which the highest value of clustering ratio is recognized. Moreover it is known that the North Anatolian fault seismicity shows period of great activity followed by long quiescent times and that an earthquake cluster, started in 1939, producing several strong events until the 1999 Duzce event [*Stein et al.*, 1997; *Barka*, 1999]. In support of this thesis, we find very high probability of triggering between events occurred in this zone after 1939, up to 70%. Similar results point out the Aleutian arc and the Kurile-Kamchetka trench as regions affected by strong interactions between events, confirming assertions of *Pollitz et al.*, [1998].

5.4 Application to CPTI catalog

The second test, to verify reliability of long-term time dependent models proposed, against the stationary Poisson model, is carried out on the Italian historical seis-

micity. Specifically we consider events occurred in Italy since 1600 to 2002, with magnitude $Mw \geq 5.5$, collected by the Italian parametric seismic catalogue (CPTI) [*Working Group CPTI*, 2004] (203 events). Since this catalog does not provide information on focal mechanism of events, we consider an isotropic function for triggering effect ($\delta \equiv 0$).

The estimation of parameters for Model I has convergence problems and however any combination of them does not give a better likelihood respect to Poisson Model.

Parameter	**Poisson Model**	**Time-dependent isotropic Model II ($\delta = 0$)**
ν	0.50 ± 0.04	0.27 ± 0.04 year^{-1}
K		0.016 ± 0.004 year^{p-1}
p		0.51 ± 0.07
c		~ 0.0 year
α		0.5 ± 0.3
d		10.0 ± 2.0 Km
q		$\equiv 1.5$
Log-likelihood	-3074.3	-3010.7

Table 5.2: Maximum Likelihood Parameters, relative to the Poisson and time-dependent model II, for CPTI catalog.

On the other side, by applying the Maximum Likelihood Method, we find that the time-dependent Model II performs significantly better than Poisson model. The values of parameters and of the maximum log-likelihood are reported in Table 5.2 for both Poisson and time-dependendent model. The value of parameter q is not significantly different by 1.5 and some trials show also in this case a trade-off of this parameter with d [*Kagan and Jackson*, 2000]. The value of c is found to be not significantly different by 0. We remark that in this context this parameter loses meaning, generally assigned for short-term triggering, of representing incompleteness of aftershocks detection soon after mainshock occurrence.

The estimated time-dependent model predicts that an event can perturb area surrounding its locations, up to some tens of kilometers, for more than one century (see Figure 5.4). The scaling of this perturbation with magnitude of inducing event, defined by parameter α, is rather poor. The time decay of triggering effect has

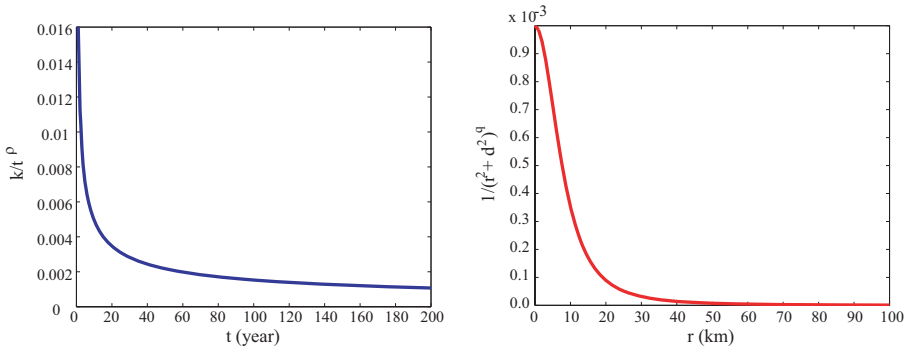

Figure 5.4: Time (left) and space (right) decay of the triggering effect described by time-dependent model II for CPTI catalog

similar features with the temporal function estimated for the PS92 catalog (see Figure 5.1). Also if described by two different analytical functions, both distributions predict a strong influence on following seismicity for some decades after the occurrence of an event, up to about one century. The difference of characteristic size of the perturbed area, surrounding location of triggering event, is due to different magnitude range (and then faults size) recovered by CPTI and PS92 catalogs.

The ability of the time-dependent model to follow the dynamics of the time series is tested through the analysis of residuals [see *Ogata*, 1988].

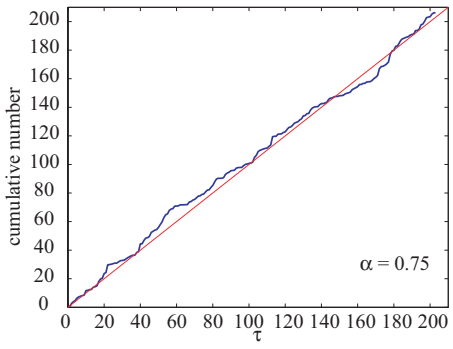

Figure 5.5: Cumulative number of residuals process (blue line). The red line indicates the average of the Kolmogorov-Smirnov statistic. The parameter α is the significance level at which the Poisson hypothesis can be rejected by the Kolmogorov-Smirnov test.

The occurrence times t_i of earthquakes are transformed into new values τ_i by the

relation

$$\tau_i = \int_{T_{start}}^{t_i} dt \int_{\mathcal{R}} dxdy\lambda_2(t, x, y/\mathcal{H}_t) \tag{5.7}$$

where T_{start} is the start time of observation history, $\mathcal{R}$ is the examined region and $\lambda_2(t, x, y/\mathcal{H}_t)$ is the conditional intensity of time-dependent model II given by equation 5.3. If the model well describes the temporal evolution of seismicity, the transformed data τ_i are expected to behave like a stationary Poisson process with unit rate [*Ogata*, 1988]. The goodness of fit of the time-dependent model is evaluated by testing the null hypothesis that values $\Delta\tau_i = \tau_{i+1} - \tau_i$ are exponentially distributed (with mean equal to 1). A deviation from this hypothesis implies the existence of basic features of examined seismicity not captured by the model. The test used is the Kolmogorov-Smirnov test (see Appendix 2A). The cumulative number of points τ_i, versus the transformed time τ, shows a good agreement with standard Poisson model (see Figure 5.5). The Kolmogorov-Smirnov test gives a significance level α at which the null hypothesis can be rejected equal to 0.75, showing an high goodness of fit.

In Figure 5.6 we show an histogram for probability ρ_j of belonging to background seismicity (equation 5.4).

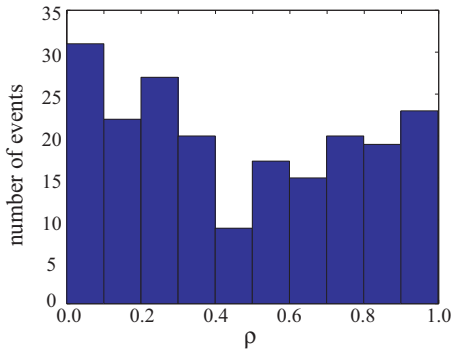

Figure 5.6: Histogram of probability ρ_j of belonging to background seismicity for CPTI catalog

Unlike the short-term triggering, for which a similar histogram appears strongly picked around 0.0 and 1.0 values, in this case values are more uniformly distributed. This result shows a balance between contribution of tectonic loading and triggering effect for occurrence of events.

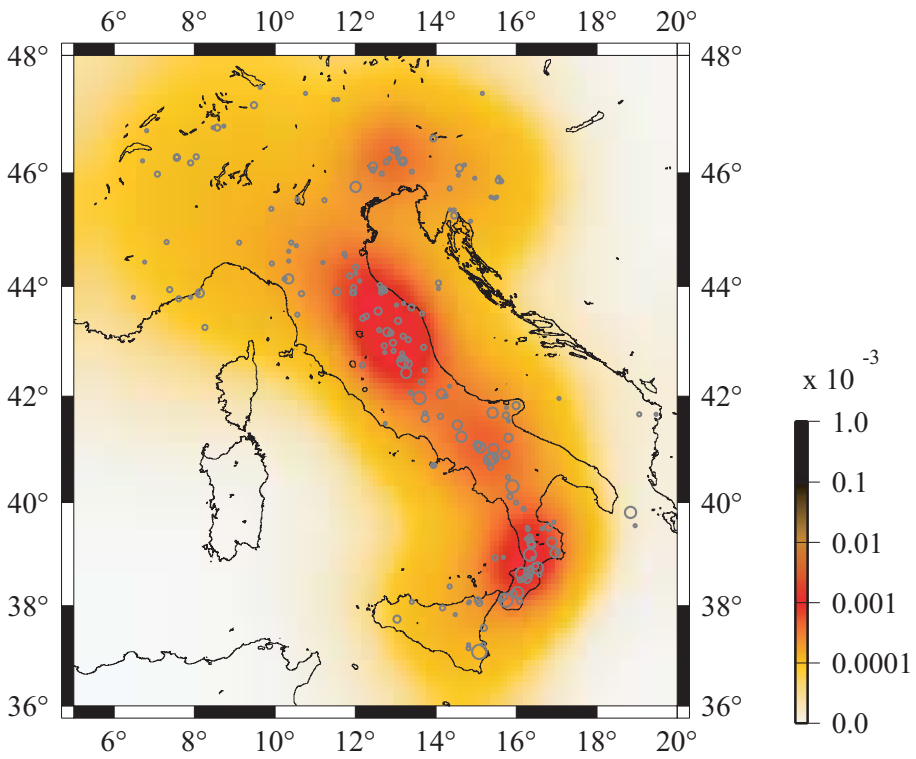

Figure 5.7: Map of background seismicity rate $\hat{\mu}(x, y)$ for CPTI catalog

The background seismicity rate $\hat{\mu}(x, y)$, estimated by equation 5.5, linked to local tectonic loading, has higher values in Central Appennines and Calabria region (see Figure 5.7), that are among the most seismogenetic areas in Italy. Lower values are found in Friuli and in part of Puglia and Campania regions.

As the map of clustering ratio $\hat{c}(x, y)$ shows (Figure 5.8), the contribution of triggering effect is not strictly correlated to tectonic loading. This is above all clear for Calabria region, in which an high value of background rate is estimated, but the clustering ratio is lower than in other zones. We remark that the regions highlighted in Figure 5.8 are in agreement with the areas, shown in Figure 4.5, in which we have recognized a significantly non-null value of fractionally integrated parameter r of ARFIMA model (see Chapter 4). We remind that this parameter is related to long-term memory of temporal trend of the occurrence rate. The agreement of results obtained by two different methods, both focusing on long-term time-dependent memory recognition, further support the reliability of our finding.

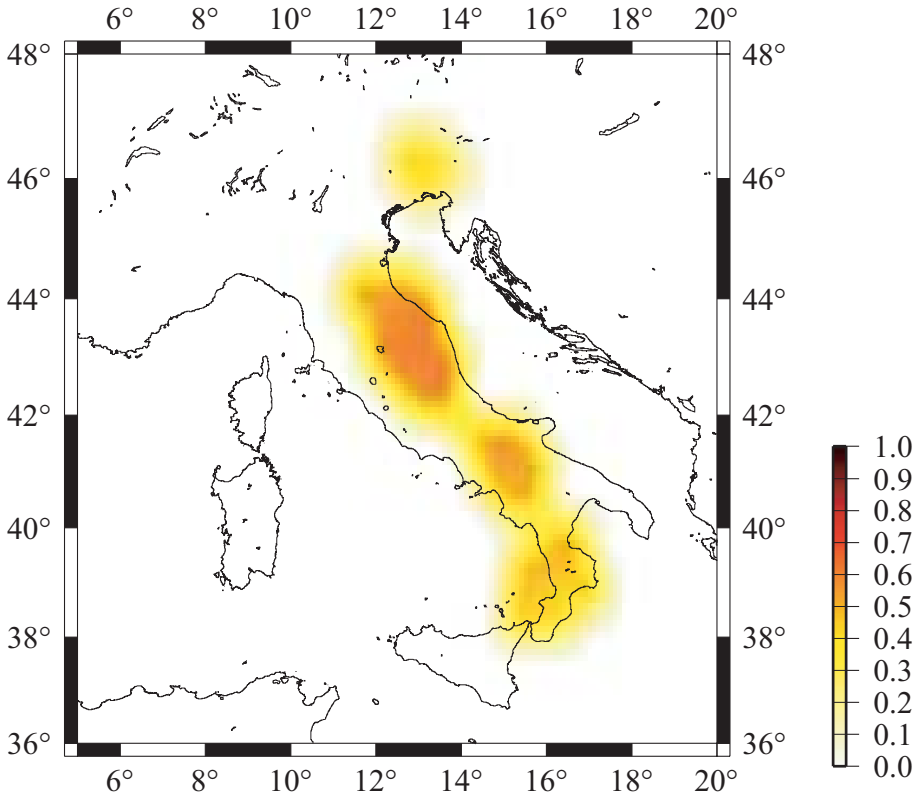

Figure 5.8: Map of the clustering ratio $\hat{c}(x, y)$ for CPTI catalog

Using equation 5.6, we can compute the probability that an event has promoted a following earthquake (see Figure 5.9). Among the events occurred in the last century, we find a strong relation ($>$50%) between the event occurred in Irpinia in 1980 (Mw=6.9) and the 1990 Potenza earthquake (Mw=5.8). Moreover we find a ”small cluster” of three events, occurred in Calabria-Sicilia regions between 1907 and 1909, including the strong earthquake (Mw7.2) of 1908 (the other magnitudes are Mw=5.9 and Mw5.5).

The Mw7.0 1915 Avezzano earthquake is promoted (40%) by the event occurred in Marsica in 1904 (Mw5.7) and, in its turn, in minimum part (about 20%), contributes to occurrence of Sora earthquake of 1922 (Mw5.6). Finally a relation (40%) between the Senigallia earthquakes of 1924 and 1930 (Mw5.6 and Mw5.9) is found.

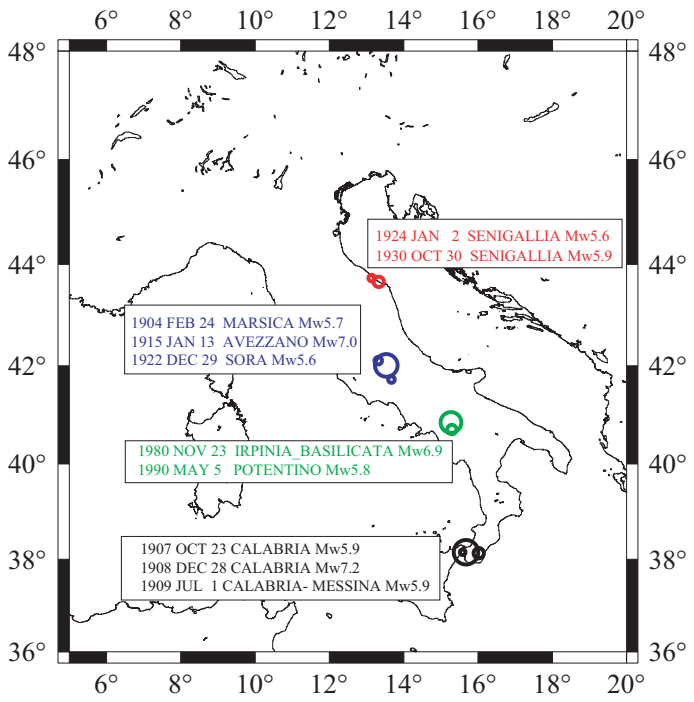

Figure 5.9: Some examples of interaction between events recorded by the CPTI catalog occurred in the last century

5.5 Final remarks

Previous results show that the interactions between events is a driving phenomenon of seismic activity, concerning a spatio-temporal scale well larger of that involved by aftershocks occurrence. The most important consequence is that seismic regions and, even more, single tectonic sources, cannot be considered as "'closed" systems, but are able to interact significantly with other remote regions. This result open a new prospect in seismic hazard assessment and long-term forecasting, because it would imply that a reliable long-term estimate of the probability of earthquake occurrence has to take into account the stress transfer of past large events.

Our models are generic and purely empirical, with no geological or geophysical basis. Independently measured variables as, for example, plate tectonic motions, fault slip rates and other geological parameters, could improve these forecasts.

Our finding has both scientific and practical results. By a theoretical point of view we characterize the statistical relationship between successive earthquakes in a

quantitative way to perform a hypothesis testing. It is widely believed that Coulomb stress increments from past earthquakes control aftershock behaviour and long-term occurrence of large earthquakes. Case studies can provide important examples to support this hypothesis, but model verification requires a thorough statistical treatment. The forecasts will also have important practical benefits in terms of mitigation strategies and hazard estimates.

Overall Conclusions and Future Prospects

The main goal of our work was to investigate two main properties of stochastic processes, the stationary and the memory of the past history, for earthquake occurrence, at different spatio-temporal scale. Whereas at short-term temporal scale the well recognized clustering features (in time and space) do not leave any doubt on the influence that the occurrence of an event has on following seismicity (memory), for longer time-period this is still a debated question. Moreover few studies deal in depth with the problem of invariance of main features of seismogenetic process in time (stationary).

The main findings of our study seems to show a strong nonstationary behaviour of sequences and swarms. Specifically the analysis of volcanic swarms and of complex sequences makes clear the relationship between the temporal variation of some parameters of the ETAS model and the physical processes driving seismicity. In particular the presence of significative p-value variations could be a very important signal for identifying the presence of magma motion. On the other side the time variation of background rate is a sign of coherent changes of external physical processes linked to seismogenetic capability of a region. These results indicate that stochastic modeling of short-term seismicity may yield important insights in constraining the physics of the source process (i.e., magma/fluids or tectonics) and in characterizing its temporal evolution. From a practical point of view, the procedure applied here seems to be a promising method for tracking in almost real time the evolution of a magma/fluids source and to unveil fluid signals, even in the case of poor hypocenter information and for patterns of seismicity largely containing short-term triggered events.

The other our main result on short-term behaviour of seismicity is that the basic features of physical process governing short-term clustering are independent on the time-space-magnitude window considered. The application of stationary ETAS modeling to strong worldwide seismicity provides parameters consistent with values computed for tectonic sequences, supporting the universality hypothesis of physical laws for earthquakes generation.

The analysis of stationarity and dependence on the past history of moderate-strong seismicity had as main goal to improve the knowledge on spatio-temporal distribution of large earthquakes. Actually unlike springing up of both statistical as physical studies about short and intermediate-term variations of seismic rate, the time behaviour of large magnitude seismicity is more hardly investigated, especially at global scale. So, whereas basic features of short-term clustering are known, main characeristics of large seismicity are not still fully understand.

For a long time, and sometimes still today, it was accepted that in a region the strong earthquakes occur independently one from others, in agreement with a Poisson distribution, with an uniform rate over time. This assumption understands some fundamental features of earthquake occurrence, the independence of interevent time interval, the stationary of seismic activity and the lack of memory of the process, and identifies sequences, as only significant departures by ”normal activity”. This point of view has been implicitly accepted in formulation of probabilistic seismic hazard assessment methodologies based on the Cornell’s method. Our finding shows evident departures by stationary Poisson hypothesis and raises serious doubts on reliability of these assumptions. Different statistical analyses seem to rule out a nonstationarity behaviour as the cause of identified dependence on time, identifying as the most likely candidate a sort of long-term interaction, probably due to viscous relaxation of the asthenosphere and/or lower crust. We cannot rule out possibility of a nonstationary behaviour of tectonic motion on longer time scale respect to the one considered in our studies. Clearly the catalogs used here are inadeguate to capture a similar feature.

From a practical perspective, the results reported here highlight that the stationary Poisson hypothesis, that stands behind most of the seismic hazard assessment,

should be considered with care. Specifically the detection of the cause and a suitability modeling of temporal dependence in the seismic activity becomes imperative to significantly improve long-term forecasting and seismic hazard assessment. The time-dependent models proposed here are only the first step towards the achievement of this aim.

References

Akaike H., A new look at the statistical model identification, *IEEE Trans. Autom. Control. 19*, 716-722, 1974.

Aki K. and Richards, P., *Quantitative seismology*, 2nd ed., Sausalito, Calif., University Science Books, 700 pp., 2002.

Allard P., Maiorani A., Tedesco D., Cortecci G. and B. Turi, Isotopic study of the origin of sulfur and carbon in Solfatara fumaroles, Campi Flegrei caldera, *J. Volcanol. Geotherm. Res, 48*, 139-159, 1991.

Antonioli A., Piccinini D., Chiaraluce L. and M. Cocco, Fluid flow and seismicity pattern: evidence from the 1997 Umbria-Marche (central Italy) seismic sequence, *Geophys. Res. Lett., 32*, L10311, doi:10.1029/2004GL022256, 2005.

Bak P., K. Christensen, L. Danon and T. Scanlon, Unified scaling law for earthquakes, *Phys. Rev. Lett., 88*, 1-4, 2002.

Barberi F. and M.L. Carapezza, The problem of volcanic unrest, the Campi Flegrei case history. In: *Monitoring and Mitigation of Volcano Hazards*. Scarpa R (ed.), Tilling RI (ed.), Springer-Verlag. Berlin, Federal Republic of Germany, 771-786, 1996.

Barka A., The 17 August 1999 Izmt earthquake, Science, 285, 1858-1859, 1999.

Bath M. and Duda S.J., Earthquake volume, plain area, seismic strain, deformation and related quantities, *Ann. Geofis., 17*, 353-368, 1964.

Box G.E.P. and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, revised ed., Holden-Day, Merrifield, Va., 1976.

Brent R., *Algorithms for Minimization without Derivatives*, Prentice-Hall, 1973.

Brockwell P.J. and R.A. Davis, *Time Series: Theory and Method*, Springer-Verlag, New York, 1987.

Bonafede M., Axy-symmetric deformation of a thermo-poro-elastic half space: inflation of a magma chamber, *Geophys. J. Int., 103*, 289-299, 1990.

Bonafede M., Hot fluid migration: an efficient source of ground deformation: appli-

cation to the 1982-1985 crisis at Campi Flegrei, Italy, *J. Volcanol. Geotherm. Res., 48*, 187-198, 1991.

Casarotti E. and A. Piersanti, Postseismic stress diffusion in Chile and South Peru, *Earth Planet. Sci. Lett., 206*, 325-333, 2003.

Castello B., G. Selvaggi, C. Chiarabba and A.Amato, CSI Catalogo della sismicitá italiana 1981-2002, vers. 1.0, INGV-CNT, Roma, www.ingv.it/CSI/.

Chéry J., S. Merkel and S. Bouissou, A Physical Basis for Time Clustering of Large Earthquakes, *Bull. Seism. Soc. Am., 91*, 1685-1693, 2001a.

Chéry J., S. Carretier and J.F. Ritz, Postseismic stress transfer explains time clustering of large earthquakes in Mongolia, *Earth Planet. Sci. Lett.,194*, 277-286, 2001b.

Chiaraluce L., Ellsworth W.L., Chiarabba C. and M. Cocco, Imaging the complexity of an active normal fault system: the 1997 Colfiorito (central Italy) case study, *J. Geophys. Res., 108*,(B6) 2294, doi:10.1029/2002JB002166, 2003.

Chiodini G., Frondini F., Cardellini C. and L. Peruzzi, Rate of diffuse carbon dioxide earth degassing estimated from carbon balance of regional aquifers: the case of central Apennine, Italy, *J. Geophys. Res., 105*, 8423-8434, 2000.

Chouet B.A., Long-period volcano seismicity: its source and use in eruption forecasting, *Nature 380*, 309-316, 1996.

Chouet B.A., R.A. Page, C.D. Stephens, J.C. Lahr, J.A. Power, Precursor swarms of long-period events at Redoubt volcano (1989-1990), Alaska. Their origin and use as a forecasting tool, *J. Volcan. Geoth. Res. 62*, 95-135, 1994.

Console R., M. Murru and A.M. Lombardi, Refining earthquake clustering models, *J. Geophys. Res., 108(B10)*, 2468, doi:10.1029/2002JB002130,2003.

Cornell, C. A., Engineering seismic risk analysis, *Bull. Seism. Soc. Am., 58*, 1583-1606, 1968.

Cornell, C. A. and S.R. Winterstein, Temporal and magnitude dependence in earthquake recurrence models, *Bull. Seism. Soc. Am., 78*, (4), 1522-1537, 1988.

Corral A., Long-Term Clustering, Scaling, and Universality in the Temporal Occur-

rence of Earthquakes, *Phys. Rev. Lett., 92*, 108501, 2004.

Corral A., Time-Decreasing Hazard and Increasing Time until the Next Earthquake, *Phys. Rev. Lett., E71*, 017101, 2005.

Cox D.R., and P. A. Lewis, *The Statistical Analysis of Series of Events*, London, Chapman and Hall, 1966.

Daley D.J. and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer-Verlag, New York, 2-nd ed., Vol. 1, pp.469, 2003.

DeMets, S., R. G. Fordon, D. F. Argus, and S. Stein, Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, *Geophys. Res. Lett., 21*, 2191 2194, 1994.

De Natale G. and A. Zollo, Statistical analysis and clustering features of the Phlegrean Fields earthquake sequence (May 1983-May 1984), *Bull. Seism. Am. Soc., 76*(3), 801-814, 1986.

De Natale G., F. Pingue, Allard P. and A. Zollo, Geophysical and geochemical modeling of the 1982-1984 unrest phenomena at Campi Flegrei caldera (Southern Italy), *J. Volcanol. Geotherm. Res., 48*, 199-222, 1991a.

De Natale G., A. Ferraro and J. Virieux, A probability method for local earthquake focal mechanism, *Geophys. Res. Lett., 18*(4), 613-616, 1991b.

De Natale G., F. Pingue and G. Mastrolorenzo, Modelli vulcanologici del bradisismo flegreo, *Boll. GNGTS*, 507-520, 1993.

De Natale G., Zollo A., Ferraro A. and J. Virieux, Accurate fault mechanism determinations for a 1984 earthquake swarm at Campi Flegrei caldera (Italy) during an unrest episode: Implications for volcanological research, *J. Geophy. Res., 100*, 24,167-24,185, 1995.

Del Pezzo E., G. De Natale and A. Zollo, Space time distribution of small earthquakes in campi Flegrei, Southern Italy, *Bull. Volc., 47*(2), 202-207, 1984.

Dickey D.A. and W.A. Fuller, Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *J. Am. Stat. Ass., 74*, 427431, 1979.

Dieterich J., A constitutive law for rate of earthquake production and its application

to earthquake clustering, *J. Geophys. Res., 99*, 2601-2618, 1994.

Dvorak J.J. and G. Berrino, Recent Ground Movement and Seimic Activity in Campi Flegrei, Southern Italy: Episodic Growth of a Resurgent Dome, *J. Geophys. Res., 96*, 2309-2323, 1991.

Dziewonski A.M. and J.H. Woodhouse, An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981, *J. Geophys. Res., 88*, 3247-3271, 1983.

Ekström G. and A. Dziewonski, Evidence of bias in estimation of earthquake size, *Nature, 332*, 319-323, 1988.

Ellsworth W.L., Matthews M.V., Nadeau R.M., Nishenko S.P., Reasenberg P.A. and R.W. Simpson, A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities, *U.S. Geological Survey, OFR 99-522*, 23 p., 1999.

Faenza L., W. Marzocchi, and E. Boschi, A nonparametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the Italian catalog, *Geophys. J. Int., 155*, 521-531., 2003.

Felzer K.R. and E. E. Brodsky, Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature, 441*, 735-738, 2006.

Fletcher R. and M.J.D. Powell, A rapidly convergent descent method for minimization, *Comput. J. 6*, 163-168, 1963.

Flinn E.A. and E.R. Engdahl, A Proposed Basis for Geographical and Seismic Regionalization, *Rev.Geophys.,3*, 123-149, 1965.

Flinn E.A., Engdahl E.R. and A.R. Hill, Seismic and Geographical Regionalization, *Bull. Seism. Soc. Am., 64*, 771-993, 1974.

Frankel A., Mapping Seismic Hazard in the Central and Eastern United States, *Seism. Res. Lett., 66*(4), 8-21, 1995.

Freed A.M. and J. Lin, Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer, *Nature, 411*, 180-183, 2001.

Friedrich A.M., Wernicke B.P., Niemi N.A., Bennett R.A. and J.L. Davis, Compar-

ison of geodetic and geologic data from the Wasatch region, Utah, and implications for the spectral character of Earth deformation at periods of 10 to 10 million years, *J. Geophys. Res.*, *108*(B4), 2199, doi:10.1029/2001JB000682, 2003.

Fuller W., *Introduction to statistical time series*, John Wiley, New York.

Gardner J.K and L. Knopoff, Is the sequence of earthquakes in Southern California with aftershocks removed, Poissonian?, *Bull. Seism. Soc. Am., 64*(5), 1363-1367, 1974.

Gibbons, J.D. and S. Chakraborti, *Non-parametric Statistical Inference*, 4th ed., rev. and expanded, New York: Marcel Dekker, 645 pp, 2003.

Gomberg J., Stress/strain changes and triggered seismicity following the M_s 7.4 Landers, California, Earthquake, *J. Geophys. Res., 101*, 751-764, 1996.

Gomberg J., N.M. Beeler, M.L. Blanpied and P. Bodin, Earthquake triggering by transient and static deformations, *J. Geophys. Res., 103*, 24411-24426, 1998.

Granger C.W.J. and R.Joyeux, An introduction to long-range time series models and fractional differencing, *J.Time Ser. Anal.,1*, 15-30, 1980.

Gross S., A. Rundle, A systematic test of time-to-failure analysis, *Geophys.J.Int., 133*, 57-64, 1998.

Hawkes A., Point Spectra of Some Mutually Exciting Point Processes, *Journal of the Royal Statistical Society, Series B (Methodological), 33*(3), 438-443,1971.

Hawkes A. and L. Adamopoulos, Cluster models for earthquakes regional comparisons, *Bull. Int. Statist. Inst., 45*(3), 454-461, 1973.

Hainzl S., Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering, *Geophys.J.Int., 159*, 1090-1096, 2004.

Hainzl S. and T. Fisher, Indications for successively triggered rupture growth underlying the 2000 earthquake swarm in Vogtland/NW Bohemia, *J. Geophys Res., 107*, 2338, doi:1029/2002JB001865, 2002.

Hainzl S. and Y. Ogata, Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geoph. Res. 110*, B05S07, doi:10.1029/2004JB003247, 2005.

Hainzl S., Scherbaum F. and C. Beauval, Estimating Background Activity Based on Interevent-Time Distribution, *Bull. Seism. Soc. Am., 96(1)*, 313-320, 2006.

Hanks T.C.and H. Kanamori, A moment magnitude scale, *J. Geophys. Res., 84(B5)*, 2348-2350, 1979.

Haslett J. and A.E. Raftery, Space-time Modelling with Long-memory Dependence: Assessing Ireland's Wind Power Resource (with Discussion), *Appl Stat 38*(1),150, 1989.

Helmstetter A., Kagan Y.Y. and D.D. Jackson, Comparison of short-term and time-independent earthquake forecast models for southern California, *Bull. seism. Soc. Am., 96*(1), 90-106, 2006.

Hill D.P., Reseanberg P.A., Michael A., Arabaz W.J., Beroza G., Brumbaugh D., Brune J.N., Castro R., Davis S., dePolo D., Ellsworth W.L., Gomberg J., Harmsen S., House L., Jackson S.M., Johnston M.J.S., Jones L., Keller R., Malone S., Munguia L., Nava S., Pechmann J.C., Sanford A., Simpson R.W., Smith R.B., Stark M., Stickney M., Vidal A., Walter S., Wong V., Zollweg J., Seismicity Remotely Triggered by the Magnitude 7.3 Landers, California, Earthquake, *Science, 260*, 1617-1623, 1993.

Hosking J.R.M., Fractional differencing, *Biometrika, 68*, 165-176, 1981.

Huc M. and I.G. Main Anomalous stress diffusion in earthquake triggering: Correlation length, time dependence, and directionality, *J. Geophys. Res., 108 (B7)*, 2324, doi:10.1029/2001JB001645, 2003.

Hurbert-Ferrari A., Suppe J., Van Der Woerd J, Wang X. and H. Lu, Irregular earthquake cycle along the southern Tianshan front, Aksu area, China, *J. Geophys. Res., 110*, B06402,
doi:10.1029/2003JB002603, 2005.

Hurvich, C.M. and Tsai, C.-L., Regression and Time Series Model Selection in Small Samples, *Biometrika, 76*, 297-307, 1989.

Jackson D.D. and Y.Y. Kagan, Testable Earthquake Forecasts for 1999, *Seism. Res. Lett., 70*, 393-403, 1999.

Kagan Y.Y., Accuracy of modern global earthquake catalogs, *Phys. Earth Planet.*

Int., 135, 173-209, 2003.

Kagan Y.Y., Short-term properties of earthquake catalogs and models of earthquake source, *Bull. Seism. Soc. Am. 94(4),* 1207-1228, 2004.

Kagan Y.Y. and L. Knopoff, Statistical short-term earthquake prediction, *Science, 236*, 1563-1567, 1987.

Kagan Y.Y. and D.D. Jackson, Long-term earthquake clustering, *Geophys. J. Int., 104*, 117-133, 1991.

Kagan Y.Y. and D.D. Jackson, Long-term probabilistic forecasting of earthquakes, *J. Geophys. Res., 99,* 13685-13700, 1994.

Kagan Y.Y. and D.D. Jackson, Worldwide Doublets of Large Shallow Earthquakes, *Bull. Seismol. Soc. Am., 89,* 1147-1155, 1999.

Kagan Y.Y. and D.D. Jackson, Probabilistic forecasting of earthquakes, *Geophys. J. Int., 143,* 438-453, 2000.

Kenner, S. J., and P. Segall, Postseismic deformation following the 1906 San Francisco earthquake, *J. Geophys. Res., 105,* 13195 13209, 2000.

Kenner S.J. and M. Simons, Temporal clustering of major earthquakes along individual faults due to post-seismic reloading, *Geophys.J.Int., 160,* 179-194,2005.

King, G.C.P., Stein R.S. and J. Lin, Static stress changes and the triggering of earthquakes, *Bull. Seism. Soc. Am., 84*, 935-953, 1994.

King, G.C.P., and M. Cocco, Fault interaction by elastic stress changes: New clues from earthquake sequences, *Adv. Geophys., 44*, 1 38, 2000.

Kisslinger C., Processes during the Matsushiro swarm as revealed by leveling, gravity, and spring-flow observations, *Geology, 3,* 57-62, 1975.

Kisslinger C. and L.M. Jones, Properties of aftershocks in southern California, *J.Geophys. Res. 96*, 11947-11958,1991.

Kossobokov V.G., V.I. Keilis-Borok, L.L. Romashkova, J.H. Healy, Testing earthquake prediction algorithms: Statistically significant real-time prediction of the largest earthquakes in the Circum-Pacific, 1992-1997, *Phys. Earth Planet. Inter., 111* (3-4), 187-196, 1999.

Kostrov V.V., Seismic moment and energy of earthquakes, and seismic flow of rock, *Izv. Acad. Sci. USSR, Phys. Solid Earth., Engl. Trans., 1*, 23-44, 1974.

Ljung G. and G.Box, On a measure of lack of fit in time series models, *Biometrika 65*, 297-303, 1978.

Linde, A.T. and I.S. Sacks, Triggering of volcanic eruptions, *Nature, 395*, 888-890, 1998.

Lolli B. and P. Gasperini, Aftershocks hazard in Italy Part I: Estimation of time-magnitude distribution model parameters and computation of probabilities of occurrence, *J. Seism, 7*, 235-257, 2003.

Lombardi A.M., Marzocchi W. and J. Selva, Exploring the evolution of a volcanic seismic swarm: The case of the 2000 Izu Islands swarm, *Geophys. Res. Lett., 33*, L07310, doi:10.1029/2005GL025157, 2006.

Lombardi A.M. and W.Marzocchi, Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes, *J. Geophys. Res., 112*, B02303, doi:10.1029/2006JB004568., 2007.

MacQueen J.B., Some Methods for classification and analysis of multivariate observations, *Proceedings of 5-th Berkely Symposium on Mathematical Statistics and Probability*,Berkeley, University of California Press, 281-297, 1967.

Main I., Earthquakes: A hand on the aftershock trigger, *Nature, 441*, 704-705, 2006.

Marsan D., Triggering of seismicity at short timescales following Californian earthquakes, *J. Geophys. Res., 108 (B5)*, 2266, doi:10.1029/2002JB001946, 2003.

Marsan D. and S.S. Nalbant, Methods for Measuring Seismicity Rate Changes: A Review and a Study of how the M_w 7.3 Landers Earthquake Affected the Aftershock Sequence of the M_w 6.1 Joshua Tree Earthquake, *Pure Appl. Geophys., 162*, 1151-1185, 2005.

Martini M., The forecasting significance of chemical indicators in areas of quiescent volcanism: examples from Vulcano and Phlegraean Fields (Italy). In: *Volcanic Hazards*, Latter (ed.) J.H., Springer, Berlin Heidelberg New York 625 pp., 1989.

Marzocchi W., Remote seismic influence on the large explosive eruptions, *J. Geo-*

phys. Res., 107 (B1), 2018, doi:10.1029/2001JB000307, 2002.

Marzocchi, W., E. Casarotti, and A. Piersanti, Modeling the stress variations induced by great earthquakes on the largest volcanic eruptions of the twentieth century, *J. Geophys. Res., 107(B11)*, 2320, doi:10.1029/2001JB001391, 2002.

Marzocchi W., L. Sandri, E. Boschi, On the validiation of earthquake-forecasting models: the case of pattern recognition algorithms, *Bull. Seismol. Soc. Am. 93*, 1994-2004, 2003a.

Marzocchi W., Selva J., Piersanti A. and E. Boschi, On the long-term interaction among earthquakes: Some insight from a model simulation, *J. Geophys. Res., 108 (B11)*, 2538, doi:10.1029/2003JB002390, 2003b.

Matsu'ura R.S., Detailed study of the earthquake sequence in 1980 off the east coast of the Izu peninsula, Japan, *J. Phys. Earth. 31*, 65-101, 1983.

McCann W.R., S.P. Nishenko, L.R. Sykes ad J.Krause, Seismic Gaps and Plate Tectonics: Seismic Potential for Major Boundaries, *Pageoph, 117*, 1082-1147, 1979.

McKernon C. and I.G. Main, Regional variations in the diffusion of triggered seismicity, *J. Geophys. Res., 110*, B05S05, doi:10.1029/2004JB003387, 2005.

Michel A.J. and L.M. Jones, Seismicity alert probabilities at Parkfield, California, revisited, *Bull. Seism. Soc. Am., 88*, 117-130, 1998.

Mikumo T., Y. Yagi, S.K. Singh, M.A. Santoyo, Coseismic and postseismic stress changes in subducting plate: possible stress interactions between large interplate thrust and intraplate normal-faulting earthquakes, *J. Geophys. Res. 107*, doi:10.1029/2001JB000446, 2002.

Mogi K., Earthquakes and fractures, *Tectonophysics 5*, 35-55, 1967.

More J.J., Garbow B.S. and K.E. Hillstrom, *Users Guide for MINPACK-1, Technical Report ANL-80-74*, Applied Mathematics Division, Argonne National Laboratory, 1980.

Mulargia F., and S. Tinti, Seismic sample areas defined from incomplete catalogues: an application to the Italian territory, *Phys. Earth Planet. Int. 40*, 273-300, 1985.

Mulargia F., P. Gasperini, and S. Tinti, Identifying different regimes in eruptive

activity: an application to Etna volcano, *J. Volcan. Geoth. Res. 34*, 89-106, 1987.

Neuberg J., Characteristics and causes of shallow seismicity in andesite volcanoes, *Phil. Trans. R. Soc. 358*, 1533-1546, 2000.

Nishenko S.P. and R. Buland, A generic recurrence interval distribution for earthquake forecasting, *Bull. Seism. Soc. Am., 77*, 1382-1399, 1987.

Nur A. and J.R. Booker, Aftershocks caused by pore fluid flow?, *Science, 175*, 885-887, 1972.

Noir J., Jacques E., Bekri S., Adler P.M., Tapponnier P. and G.C.P. King, Fluid flow triggered migration of events in the 1989 Dobi earthquake sequence of Central Afar, *Geophys. Res. Lett., 24*, 2335-2338, 1997.

Ogata Y., Statisticals Models for Earthquake Occurrences and Residual Analysis for Point Processes, *J. Amer. Statist. Assoc. 83(401)*,9-27, 1988.

Ogata Y., Detection of precursory relative quiescence before great earthquakes through a statistical model, *J. Geophys. Res. 97*, 19,845-19,871, 1992.

Ogata Y., Space-Time Point-Process Models for Earthquake Occurrences, *Ann. Inst. Statist. Math. 50(2)*, 379-402, 1998.

Ogata Y., Seismicity Analysis Point-process Modeling: A review, *Pure appl.geophys., 155*, 471-507, 1999.

Ogata Y. and J. Zhuang, Space-time ETAS models and an improved extension, *Tectonophysics, 413*, 13-23, 2006.

Omori, F., On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*,7, 111120, 1894

Ouillon G. and D. Sornette., Magnitude-dependent Omori law: Theory and empirical study, *J. Geophys. Res., 110*, B04306, doi:10.1029/2004JB00331, 2005.

Ozawa S., S. Miyazaki, T. Nishimura, M. Murakami, M. Kaidzu, T. Imakiire and X.Ji, Creep, dike intrusion, and magma chamber deflation model for the 2000 Miyake eruption and the Izu islands earhquakes, *J.Geophys. Res. 109*, B02410, doi:10.1029/2003JB002601, 2004.

Pacheco J.F. and L.R. Sykes, Seismic moment catalog of large shallow earthquakes,

1900 to 1989, *Bull. Seismol. Soc. Am., 82,* 1306-1349, 1992.

Papazachos B.C., Papadimitriou E.E.,Karakaisis G.F. and D.A. Panagiotopulos, Long-term Earthquake Prediction in the Circum-Pacific Convergent Belt, *Pageoph, 149,* 173-217, 1997.

Piersanti A., G. Spada, R. Sabadini, and M. Bonafede, Global postseismic deformation, *Geophys. J. Int., 120*, 544-566, 1995.

Piersanti A., G. Spada and R. Sabadini, Global postseismic rebound of a viscoelastic Earth: Theory for finite faults and application to the 1964 Alaska earthquake, *J. Geophys. Res., 102*, 477-492, 1997.

Piersanti, A., Postseismic deformation in Chile: Constraints on the asthenospheric viscosity, *Geophys. Res. Lett., 26*, 3157 3160, 1999.

Pollitz F.F., Postseismic relaxation theory on the spherical Earth, *Bull. Seismol. Soc. Am., 82*, 422-453, 1992.

Pollitz F.F. and I.S. Sacks, Modeling of postseismic relaxation following the great 1857earthquake, southern California, *Bull. Seismol. Soc. Am., 82*, 454-480, 1992.

Pollitz F.F. and I.S. Sacks, The 1995 Kobe, Japan, earthquake: A long-delayed aftershock of the offshore 1944 Tonankai and 1946 Nankaido earthquakes, *Bull. Seismol. Soc. Am., 87*, 1-10, 1997.

Pollitz F.F., R. Bürgmann and B. Romanowicz, Viscosity of oceanic asthenosphere inferred from remote triggering of earthquakes, Science, 280, 1245-1249, 1998.

Pollitz F.F., M. Vergnolle and E. Calais, Fault interaction and stress triggering of twentieth century earthquakes in Mongolia, *J. Geophys. Res., 108*(B10), 2503, doi:10.1029/2002JB002375, 2003.

Rhoades D.A. and F.F.Evison, Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pageoph, 161*,1, 47-72, 2004.

Rydelek P.A., I.S. Sacks, Triggering and inhibition of great Japanese earthquakes: the effect of Nobi 1891 on Tonakai 1944, Nankaido 1946 and Tokai, *Earth Plan. Sci. Lett. 206*, 289-296, 2003.

Ritz J.F., Brown E.T., Bourls D.L., Philip H., Schlupp A., Raisbeck G.M., Yiou F.,

Enkhtuvshin B., Slip rates along active faults estimated with cosmic-ray-exposure dates: Application to the Bogd fault, Gobi-Altai, Mongolia, *Geology, 23*,(11), 1019-1022, 1995.

Rosi M., Sbrana A. and C. Principe, The Phlegraean Fields: structural evolution, volcaniv history and eruptive mechanism, *J. Volc. Geoth. Res., 17*,, 237-288, 1983.

Said E. and D. A. Dickey, Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order *Biometrika, 71*(3), 599-607, doi:10.2307/2336570, 1984.

Santoyo M.A., S.K. Singh, T. Mikumo, M Ordaz, Space-time clustering of large thrust earthquakes along the mexican subduction zone: an evidence of source stress interaction, *Bull. Seism. Soc. Am. 95*, 1856-1864, doi:10.1785/0120040185, 2005.

Scholz C.H., *The Mechanics of Earthquakes and Faulting*, Cambridge University Press, 2002.

Shaw B.E., Generalized Omori law for aftershocks and foreshocks from a simple dynamics, *Geophys. Res. Lett.,20*, 907-910, 1993.

Sella G.F., Dixon T.H. and A. Mao, A model for recent plate velocities from space geodesy, *J. Geophys. Res., 107*, doi:10.1029/2000JB000033, 2002.

Selva J. and W. Marzocchi, Focal parameters, depth estimation, and plane selection of the worldwide shallow seismicity with $M_s \geq 7.0$ for the period 1900-1976, *Geochem. Geophys. Geosyst, 5*, Q05005, doi:10.1029/2003GC000669, 2004.

Selva J., W. Marzocchi, Variations of Southern California seismicity: empirical evidence and possible physical causes. *J. Geophys. Res., 110*, B11306, doi:10.1029/2004JB003494, 2005.

Sieh K., M. Stuiver and D. Brillinger, A More Precise Chronology of Earthquakes Produced by the San Andreas Fault in Southern California, *J. Geophys. Res., 94*(B1), 603-623, 1989.

Shapiro S.A., E. Huenges and G. Borm, Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site, *Geophys. J. Int., 131*, 15-18, 1997.

Shimazaki K. and T. Nakata, Time-predictable recurrence model for large earthquakes, *Geophys. Res. Lett., 7*, 279-282, 1980.

Stein R.S., The role of stress transfer in earthquake occurrence, *Nature, 402*, 605-609, 1999.

Stein R.S., G.C.P. King and J. Lin, Change in failure stress on the southern San Andreas fault system caused by the 1992 M=7.4 Landers earthquake, *Science, 258*, 1328-1332, 1992.

Stein R.S., G.C.P. King and J. Lin, Stress triggering of the 1994 M=6.7 Northridge, California, earthquake by its predecessors, *Science, 265*, 1432-1435, 1994.

Stein, R.S., Barka, A.A. and J.H. Dieterich, Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering, *Geophys. J. Int., 128*, 594-604, 1997.

Thatcher, W., Nonlinear buildup and earthquake cycle on the San Andreas fault, *J. Geophys. Res., 88*, 5893 5902, 1983.

Toda, S., Stein, R.S., Reaseanberg, P.A., Dieterich, J.H. and Yoshida, A., Stress transferred by the 1995 M_w=6.9 Kobe, Japan, shock: effect on aftershocks and future earthquake probabilities, *J. Geophys. Res., 103*, 24543-24565, 1998.

Toda S., Stein R.S. ad T.Sagiya, Evidence from the AD 2000 Izu islands earthquake swarm that sressing rate governs seismicity, *Nature 419*, 58-61, 2002.

Todesco D., R. Pece and J.C. Sabroux, No evidence of a new magmatic gas contribution to the Solfatara volcanic gases, during the bradyseismic crisis at Campi Felgrei (Italy), *Geophys. Res. Lett. 15*, 1441-1444, 1988.

Utsu T. and A. Seki, Relation between the area of aftershocks region and the energy of the mainshock (in Japanese), *Zisin (J. Seismol. Soc. Japan.) 2nd Ser. ii, 7*, 233-240, 1955.

Utsu T., A Statistical Study on the Occurrence of Aftershocks, *Geophysical Magazine 30*, 521-605, 1961.

Utsu T., Ogata Y. and R.S. Matsu'ura, The Centenary of the Omori Formula for a Decay Law of Aftershock Activity, *J. Phys. Earth 43*, 1-33, 1995.

Varotsos P., Eftaxias K. , Vallianatos F. and M. Lazaridou, Basic principles for evaluating an earthquake prediction method , *Geoph. Res. Lett., 23* (11), 1295-1298, 1996.

Vere-Jones D., Stochastic models for earthquake occurrence, *J.R. Stat. Soc., B32,* 1-62, 1970.

Wallace R.E., Grouping and migration of surface faulting and variations in slip rates on faults in the Great Basin province, *Bull. Seism. Soc. Am., 77*, 868 - 876, 1987.

Weldon R.J., Scharer K.M., Fumal T.E. and G.P.Biasi, Wrightwood and the earthquake cycle: what a long recurrence record tells us about how faults work, *GSA Today, 14*(9), 4-10, 2004.

Weldon R.J., Fumal T.E., Biasi G.P. and K.M.Scharer, Past and Future Earthquakes on the San Andreas Fault, *Science, 308*, 966-967, 2005.

Working Group CPTI, *Catalogo Parametrico dei Terremoti Italiani*, versione 2004 (CPTI04), INGV, Bologna, 2004.

Working Group on California Earthquake Probabilities (WG02) *Earthquakes probabilities in the San Francisco Bay region: 2002 to 2031*, USGS, Open-file Rept. 03-214, 2003;
http://pubs.usgs.gov/of/2003/of03-214.

Wyss M. and S. Wiemer, Change in the Probability for Earthquakes in Southern California due to the Landers Magnitude 7.3 Earthquake, *Science, 290,* 1334-1338, 2000.

Zhuang J., Y. Ogata and D. Vere-Jones, Stochastic declustering of space-time earthquake occurrence, *J. Am. Stat. Assoc., 97*, 369-380, 2002.