Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes

Anna Maria Lombardi
Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy.

Warner Marzocchi
Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy.

Short title: NONSTATIONARITY IN SEISMICITY
Abstract. The purpose of this paper is to characterize the statistical distribution of worldwide largest earthquakes. We analyze the distribution of worldwide shallow events with $M_s$ 7.0+ since 1900, by following the Occam’s razor philosophy; we start from the simplest possible model (stationary Poisson process), and we inspect more complicated models only if the data show significant departures from the simplest one. The results show two important characteristics; first, worldwide $M_s$ 7.0+ earthquakes tend to cluster in time and space, with features similar to smaller events. Second, for some seismic regions there is evidence in favor of long-term fluctuations of the earthquake rate. These results support the hypothesis of universality, suggesting that an ETAS model with a background varying with time can be considered reliable to describe the seismicity distribution over a wide space-time-magnitude window. From a practical point of view, our findings suggest that the paradigm that seismic zones are stationary systems, implicitly assumed in seismic hazard assessment, should be regarded with caution.
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, and/or 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. In this paper, 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, the probabilistic
concept of stationarity, suggesting some methods to detect significant rate changes.
From a more geological point of view, some evidence of nonstationary behavior is also
detected for individual faults by paleoseismological studies [Hurbert-Ferrari et.al, 2005;
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.

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]

$$M_w = \frac{\log(M_0) - 9.1}{1.5}$$  \hspace{1cm} (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.

**Worldwide Zonation**

In this paper, 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 paper 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 1), 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/) global seismic events in the period 1977-2004, with depth \( \leq 70 \) km and \( M_w \geq 5.5 \) (11148 earthquakes). Since, considering the Aki convention [Aki and Richards, 2002], large differences of strike \((\phi)\) and dip \((\delta)\) 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]:

\[
\begin{align*}
\phi' &= \phi; & \delta' &= \delta \quad (0 \leq \phi < 180) \\
\phi' &= \phi - 180; & \delta' &= 180 - \delta \quad (180 \leq \phi < 360).
\end{align*}
\]

Then we compare focal mechanism parameters for neighboring FE zones and group zones for which the difference between strike, dip and rake angles is \( \leq 45^\circ \). The 6 new regions obtained by this method (see Figure 1b) are listed in Table 1, together with the 17 original FE zones that they comprise. The name of these new 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 in Flinn and Engdahl [1965] (for example: R1 coincides with the FE zone 1, R5 is the FE zone 5 and so on).
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

\[ AICc(K) = -2\log L + 2K \frac{N}{N - K - 1} \]  

(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 \( AIC = -2\log L + 2K \) [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 A], while the KS1 checks the hypothesis that IETs are exponentially distributed [see Appendix B]. 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 (i.e., the instantaneous conditional rate of occurrence at \( t \)) depends on the "elapsed" times since the last event (as for renewal models) or on information about all past events (as in ETAS models).
Testing the stationary Poisson model

Table 2 reports the results of statistical tests, described in Appendixes A and B, 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 $\sim 0.003$) that 3 out of 30 independent tests are rejected at a 0.01 significance level by chance. 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 processes having different parameter $\lambda_i$ 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 2 we report the empirical cumulative distribution function of the IETs for the global catalog. 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.

Testing the stationary ETAS model

The main characteristics of the ETAS model [Ogata, 1988; 1998] are reported in Appendix C. 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 \((\nu, k, c, p, \alpha, d, q)\); see Appendix C 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-Flatcher-Powell method [Flatcher 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.

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 Appendix C). 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. 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, $R_{\text{max}}$ (see Appendix C), 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 $R_{\text{max}}$, 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 $\nu$ is the overall rate of occurrence for non induced events, i.e., the background seismic rate (see Appendix C). For stationary ETAS Model 1, $\nu$ is constant and it is $6.7 \pm 0.3 \text{yr}^{-1}$ (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 $\tau_i$ by the relation

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

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 Appendix C). If the model describes well the temporal evolution of seismicity, the transformed data $\tau_i$ 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 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.

The presence of significant autocorrelation in some sequences (see results of the Runs test in Table 2 and in Table 4) may suggest the presence of long-term modulation
of earthquake rates, at least in some zones. In Figure 3, 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 analyses, 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 $\nu$. 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 $\pi_i$ 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 $\lambda(t_i, x_i, y_i/\mathcal{H}_i)$, both computed by the stationary ETAS model (see Appendix C). Then we assign each event $i$ to the declustered catalog if $\pi_i$ is larger than 0.5. 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 $\pi_i$ close to 0 or 1 (see Figure 4), and only 66
events have $\pi_i$ 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 $\pi$ drastically reduces biases due to possible aftershock 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.

**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.

Change points of the background seismic rate $\nu(t)$ are identified by applying the
\textit{K-mean cluster analysis} algorithm (Appendix D). 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 D). Details and problems related to this method can be found in Appendix D. 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.

By the \textit{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 5 we report the number of groups identified by the \textit{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 D) 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 \textit{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 5. We find that the nonstationary model is significantly better (s.l.<0.01) for 3 (R12, R19*, R25*) of the 4 zones considered.

The time/magnitude plot, the boundaries of the subsets identified by the \textit{K}-mean
algorithm, and the normalized cumulative distributions of residuals for the Poisson, ETAS, and NETAS models are shown in Figure 5 (region R12), Figure 6 (region R19*), and Figure 7 (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.

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 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*.

Discussion and Conclusions

The analysis reported in this paper 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 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 [e.g., Selva and Marzocchi, 2005], and/or with shorter characteristic times that were interpreted in terms of static stress changes [Marsan and Nalbant, 2005], a rapid evolution of magmatic intrusions [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 paper 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 paper 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, 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 find any clear 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 6 and 7) 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 in this paper.
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 preferred 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 (\(\sim\) 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 paper 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 in this paper is the increase of seismicity detected between 1930 and 1960 in region R25* (Central Asia) (see Figure 7). 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 7).

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.
Appendix A: 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 $x_1, x_2, \ldots, x_n$, with an unknown distribution $F_X$, arose. Then the null hypothesis $\mathcal{H}_0$ that we wish to test is

$$\mathcal{H}_0 : F_X(x) = F_0(x) \text{ for all } x$$  \hspace{1cm} (5)

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]:

$$KS1 = \max_x |F_0(x) - S_n(x)|$$  \hspace{1cm} (6)

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].

Appendix B: 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:

\[
P_R = \sum_{i=3,5,...}^{R} \left( \frac{(n-1)(p-1)}{(n+p)} \right) + \sum_{i=2,4,...}^{R} 2 \frac{(n-1)(p-1)}{(n+p)}
\]

(7)

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

\[
\bar{R} = \frac{2np}{n + p} \quad \sigma_R^2 = \frac{2np(2np - n - p)}{(n + p)^2(n + p - 1)}
\]

(8)

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

\[
Z = \frac{R - \bar{R}}{\sigma_R}
\]

(9)

[see Gibbons and Chakraborti, 2003].

Appendix C: Epidemic Type Aftershock Sequences (ETAS)

Model

The ETAS model [Ogata, 1988; 1998] is a stochastic point process based on the well-known modified Omori law [Omori, 1894; Utsu, 1961] that models the coseismic stress triggered activity. 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.

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

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

(10)

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; \( \alpha \) 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, because we cannot evaluate effect of directivity in short-term triggering due
to high epicentral error and low number of aftershocks in each sequence. Then, we
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:

\[
f_{ind}^i(x,y) = \frac{(q-1)q^{2(q-1)}}{\pi} \left[ 1 - \left( 1 + \frac{R_{max}^2}{d^2} \right)^{-q+1} \right]^{-1} \left( \frac{1}{r_{(x,y)(x_i,y_i)}} + \frac{1}{d^2} \right)^{q-1}.
\] (11)

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 (10) and (11). Therefore, the total intensity function of the model is

\[
\lambda(t,x,y/\mathcal{H}_t) = \mu(x,y) + \sum_{i < t} \lambda_{i}^{ind}(t,x,y)
\] (12)

where \(\mathcal{H}_t\) is the observation history up time \(t\).

We set \(\mu(x,y) = \nu u(x,y)\), where \(\nu\) is a positive-valued parameter that represents
the poissonian rate of worldwide background events and \(u(x,y)\) is the PDF of their
locations. For the nonstationary ETAS model considered here (NETAS), the background
rate depends on time, i.e., \(\nu(t) \neq \nu\). The parameters \((\nu, k, c, p, \alpha, d, q)\) of the model
can be estimated by maximizing the Log-Likelihood function [Daley and Vere-Jones,
2003]. Given occurrence times of earthquakes \(t_i, i = 1, ..., N\), their magnitudes \(m_i\) and
epicentral coordinates \((x_i, y_i)\), the Log-Likelihood \((\log[L])\) of the ETAS model, in an
interval time \([T_1, T_2]\) and in a region \(\mathcal{R}\), is given by

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

\[\text{(13)}\]

[Daley and Vere-Jones, 2003].

**Appendix D: \(K\)-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*

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

\[\text{(14)}\]

where \(N\) is the number of data vectors \(u_j\) that we are grouping, and \(\mu_k\) 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 $\sigma^2$, 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:

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

(15)

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:

$$
\text{AICc}^* = \text{AICc} + \text{pen}(N, \mathcal{K} - 1) = -2\log L(\mathcal{D}|\mathcal{M}, \sigma^2) + 2p\mathcal{K} \frac{N}{N - p\mathcal{K} - 1} + \text{pen}(N, \mathcal{K} - 1).
$$

(16)

where $p$ is the dimension of $\hat{\theta}_k$ and $\text{pen}(N, \mathcal{K} - 1)$ 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 1/ the detection of the most likely value of $\mathcal{K}$, above which clustering quality does not improve, and 2/ 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 $K$ exponential distributions with rate $\lambda_k$ (then $p = 1$), for a specified configuration $M$ with $K$ clusters the likelihood of $D$ is:

$$\log[L(D|M, \sigma^2)] = \sum_{k=1}^{K} \left( N_k \log(\lambda_k) - \lambda_k T_k \right)$$

(17)

where $N_k$ is the number of interevent times in the $k$-th cluster and $T_k$ is sum of their lengths. We estimate $\lambda_k$ 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 $\text{pen}(N, K - 1)$ 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 $K$-mean procedure, imposing that the number of groups is $K$. 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 $K = 2, \ldots, 10$.

**Acknowledgments.** We thank Y.Kagan and the associate editor for their constructive comments that greatly improved the paper. Special thanks to I. Main for his helpful
suggestions.
References


and Hall, 1966.


Hainzl S. and Y.Ogata, Detecting fluid signals in seismicity data through statistical earthquake


393-403, 1999.


Marzocchi W., Selva J., Piersanti A. and E. Boschi, On the long-term interaction among


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.


Stein R.S., G.C.P. King and J. Lin, Change in failure stress on the southern San Andreas fault


Zhuang J., Ogata Y. and D. Vere-Jones, Stochastic Declustering of Space-Time Earthquake

A. Lombardi, Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Roma, Italy. (e-mail: lombardi@ingv.it)

W. Marzocchi, Istituto Nazionale di Geofisica e Vulcanologia, Via D. Creti 12, 40128 Bologna, Italy. (e-mail: marzocchi@bo.ingv.it)

Received; revised; accepted.
Table Captions

**Table 1**: Regions defined regrouping 17 Flinn-Engdahl zones by Kostrov (1974) method (see text for details).

**Table 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).

**Table 3**: 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.

**Table 4**: 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 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.

**Table 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. Arrows identify regions for which stationary hypothesis is rejected at 0.01 significance level. Last two columns show results of residuals analysis.
Figure Captions

Figure 1: 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).

Figure 2: Empirical cumulative distribution of IETs $\Delta t_i = t_i - t_{i-1}$ for PS92 catalog (reds dots) along with the exponential distribution expected for a stationary Poisson model (continuous blue line) on logarithmic scale.

Figure 3: 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).

Figure 4: Histogram of probability $\pi_i$ 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 $\pi_i$ larger than 0.5 are assigned to declustered catalog.

Figure 5: 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.

Figure 6: As for Figure 5, but relative to region R19*.
Figure 7: As for Figure 5, but relative to region R25*.
Table 1. New regions defined grouping 17 Flinn-Engdahl zones

<table>
<thead>
<tr>
<th>New zones</th>
<th>FE zones</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2*</td>
<td>2-3-4</td>
</tr>
<tr>
<td>R14*</td>
<td>14-15</td>
</tr>
<tr>
<td>R17*</td>
<td>17-18</td>
</tr>
<tr>
<td>R19*</td>
<td>19-20-21</td>
</tr>
<tr>
<td>R22*</td>
<td>22-23</td>
</tr>
<tr>
<td>R25*</td>
<td>25-26-27-28-47</td>
</tr>
</tbody>
</table>
Table 2. Test of the stationary Poisson hypothesis.

<table>
<thead>
<tr>
<th>Region</th>
<th># events</th>
<th>$\alpha$ Runs test</th>
<th>$\alpha$ KS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>42</td>
<td>0.68</td>
<td>0.05</td>
</tr>
<tr>
<td>R2$^*$</td>
<td>21</td>
<td>0.69</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R5</td>
<td>39</td>
<td>0.90</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R6</td>
<td>17</td>
<td>0.04</td>
<td>0.15</td>
</tr>
<tr>
<td>R8</td>
<td>56</td>
<td>0.84</td>
<td>0.1</td>
</tr>
<tr>
<td>R11</td>
<td>15</td>
<td>0.14</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R12</td>
<td>35</td>
<td>0.33</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R14$^*$</td>
<td>96</td>
<td>0.71</td>
<td>0.01</td>
</tr>
<tr>
<td>R16</td>
<td>29</td>
<td>0.51</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R19$^*$</td>
<td>125</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>R22$^*$</td>
<td>54</td>
<td>0.39</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R24</td>
<td>12</td>
<td>0.05</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R25$^*$</td>
<td>54</td>
<td>0.01</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R29</td>
<td>11</td>
<td>0.42</td>
<td>$&gt;0.2$</td>
</tr>
<tr>
<td>R30</td>
<td>19</td>
<td>0.88</td>
<td>0.1</td>
</tr>
<tr>
<td>Global Catalog</td>
<td>698</td>
<td>0.85</td>
<td>0.01</td>
</tr>
</tbody>
</table>
### Table 3. Parameters of the ETAS model

<table>
<thead>
<tr>
<th></th>
<th>ETAS 1 ($q = 1.5$)</th>
<th>ETAS 2 ($q \neq 1.5$)</th>
<th>Poisson ($K = 0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># par.=6</td>
<td># par.=7</td>
<td>#par.=1</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$6.7 \pm 0.3$</td>
<td>$6.7 \pm 0.3$</td>
<td>$7.8 \pm 0.3$</td>
</tr>
<tr>
<td>$K$</td>
<td>$(4 \pm 1) \times 10^{-3}$</td>
<td>$(4 \pm 1) \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>$1.1 \pm 0.1$</td>
<td>$1.1 \pm 0.1$</td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td>$(2 \pm 1) \times 10^{-4}$</td>
<td>$(2 \pm 1) \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$1.2 \pm 0.2$</td>
<td>$1.2 \pm 0.2$</td>
<td></td>
</tr>
<tr>
<td>$d$</td>
<td>$26 \pm 3$</td>
<td>$40 \pm 10$</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>$1.5$</td>
<td>$1.9 \pm 0.3$</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-10957.2</td>
<td>-10956.9</td>
<td>-11248.3</td>
</tr>
<tr>
<td>AICc</td>
<td>21926.4</td>
<td>21927.9</td>
<td>22498.6</td>
</tr>
</tbody>
</table>
Table 4. Test of the ETAS model (Residuals).

<table>
<thead>
<tr>
<th>Region</th>
<th># events</th>
<th>$\alpha$ Runs test</th>
<th>$\alpha$ KS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>42</td>
<td>0.85</td>
<td>0.08</td>
</tr>
<tr>
<td>R2*</td>
<td>21</td>
<td>0.76</td>
<td>0.24</td>
</tr>
<tr>
<td>R5</td>
<td>39</td>
<td>0.91</td>
<td>0.32</td>
</tr>
<tr>
<td>R6</td>
<td>17</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>R8</td>
<td>56</td>
<td>0.88</td>
<td>0.51</td>
</tr>
<tr>
<td>R11</td>
<td>15</td>
<td>0.41</td>
<td>0.16</td>
</tr>
<tr>
<td>R12</td>
<td>35</td>
<td>0.37</td>
<td>0.39</td>
</tr>
<tr>
<td>R14*</td>
<td>96</td>
<td>0.67</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>R16</td>
<td>29</td>
<td>0.47</td>
<td>0.24</td>
</tr>
<tr>
<td>R19*</td>
<td>125</td>
<td>0.02</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>R22*</td>
<td>54</td>
<td>0.39</td>
<td>0.16</td>
</tr>
<tr>
<td>R25*</td>
<td>54</td>
<td>0.01</td>
<td>0.48</td>
</tr>
<tr>
<td>R30</td>
<td>19</td>
<td>99.1</td>
<td>0.14</td>
</tr>
<tr>
<td>Global Catalog</td>
<td>698</td>
<td>0.75</td>
<td>0.10</td>
</tr>
<tr>
<td>Region</td>
<td># groups</td>
<td>AICc st. Poisson</td>
<td>AICc* nonst. Poisson</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>background</td>
<td>background</td>
<td>α</td>
</tr>
<tr>
<td>R5</td>
<td>5</td>
<td>133.0</td>
<td>131.2</td>
</tr>
<tr>
<td>R12</td>
<td>4</td>
<td>130.1</td>
<td>120.5</td>
</tr>
<tr>
<td>R19*</td>
<td>9</td>
<td>177.6</td>
<td>161.4</td>
</tr>
<tr>
<td>R25*</td>
<td>3</td>
<td>154.3</td>
<td>146.5</td>
</tr>
</tbody>
</table>

Table 5. Tests of NETAS model
Residual index

\[ \Delta \tau_i \]

R19*

R25*
Poisson Model

ETAS Model

NETAS Model

cumulative number of residuals

time [years]

transformed time \( \tau \)

magnitude

cumulative number of earthquakes

0 50 100
0 20 40 60 80 100 120
1900 1930 1960 1990
7.5 6.5 8.5 9.5 7.0 8.0 9.0

cumulative number of residuals

transformed time \( \tau \)

a) b) c) d)
The diagrams illustrate the cumulative number of earthquakes and the cumulative number of residuals over time for different models:

- **Poisson Model** (b): Show the cumulative number of residuals against transformed time, indicating a linear relationship with a slight deviation from linearity.

- **ETAS Model** (c): Displays the cumulative number of residuals for the ETAS model, showing a pattern similar to the Poisson model.

- **NETAS Model** (d): Similar to the ETAS model, the NETAS model shows a pattern in the cumulative number of residuals against transformed time.

These models help in understanding the seismic activity and residuals over historical time periods.