Statistical analysis of the Central-Europe Seismicity

Licia Faenza[1], Sebastian Hainzl[2] and Frank Scherbaum[1]

[1] Institute of Earth Sciences,University of Potsdam, Karl-Liebknecht Str. 24, 14476 Potsdam - Golm, Germany.

[2] GeoForschungsZentrum, Telegrafenberg, 14473 Potsdam, Germany

Accepted *2008, in Tectonophysisc*; in original form *27/07/2007*

Abbreviated title:Statistical analysis of the Central-Europe Seismicity

Corresponding author: Licia Faenza

Institute of Earth Sciences

University of Potsdam

Karl-Liebknecht Str. 24

14476 Potsdam - Golm

Germany.

Tel: +49 331 977 5846, fax: +49 331 977 5700

Now: at Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy. e-mail: licia.faenza@ingv.it

Abstract

The aim of this paper is to characterize the spatio-temporal distribution of Central-Europe seismicity. Specifically, by using a non-parametric statistical approach, the proportional hazard model, leading to an empirical estimation of the hazard function, we provide some constrains on the time behavior of earthquake generation mechanisms. The results indicate that the most conspicuous characteristics of $M_W 4.0+$ earthquakes is a temporal clustering lasting a couple of years. This suggests that the probability of occurrence increases immediately after a previous event. After a few years, the process becomes almost time independent. Furthermore, we investigate the cluster properties of the seismicity of Central-Europe, by comparing the obtained result with the one of synthetic catalogs generated by the epidemic type aftershock sequences (ETAS) model, which previously have been successfully applied for short term clustering. Our results indicate that the ETAS is not well suited to describe the seismicity as a whole, while it is able to capture the features of the short-term behaviour. Remarkably, similar results have been previously found for Italy using a higher magnitude threshold.

Keywords: Earthquake Distribution, Earthquake Forecast, Spatio-temporal statistical analysis, Cluestre, Central Europe.

1 Introduction

The statistical modeling of earthquakes is a fundamental ingredient for earthquake forecasting and for understanding earthquake generation mechanisms. In principle, statistical analysis could be used for testing a variety of hypotheses, such as the existence of seismic cycles or clustering of main events within the spatio-temporal occurrence.

Despite a large number of investigations (Vere-Jones, 1970; Nishenko, 1985; Ellsworth *et al.*, 1998; Ogata, 1998; Kagan and Jackson, 2000; Posadas *et al*, 2002; Parson, 2005, and many other reference therein), so far the scientific community could not reach a consensus on the general properties of the spatio-temporal earthquake distribution (see as reference of earthquake statistic the Working Group on California Earthquake Probability, 2003). Remarkably, the differences between the models considered (i.e., Brownian Passage Time, Poison and ETAS) are not only of statistical nature, but they imply partially opposing physical mechanism for earthquake occurrence. The testing of different hypotheses can not be easily done because of the rarity of large earthquakes. However, presently, huge efforts are spent to develop and implement standard procedures to test the forecasting ability for seismicity (Schorlemmer *et al.*, 2007).

The aim of this paper is to give some insight on the spatio-temporal occurrence of earthquakes in Central-Europe. In a recent paper (Faenza *et al.*, 2003), a non-parametric and multivariate method has been applied in order to estimate the spatio-temporal distribution of earthquakes in Italy . This method drastically reduces the *a priori* assumption on the temporal domain and allows to find a base-line hazard function, which trend versus time provides

information on the earthquake generation mechanism.

The hazard function specifies the instantaneous occurrence rate at time $T = t$, given that no earthquakes occurs before that time. There is a unequivocal relationship between the hazard function and other statistics (i.e., the density function, the survivor function and the cumulative function), and it defines without ambiguity the time distribution of the point process. Moreover, its trend in time shows if the statistic indicates any kind of cycling behaviour (increasing trend), cluster behaviour (decreasing trend) or a random behaviour (constant trend), as already discussed in Sornette and Knopoff (1997) and Faenza *et al.* (2003). So far, in geosciences, the stress release model (Vere-Jones, 1987; and many other application of it) and the ETAS model (Ogata, 1988, 1998) are based on the study of the hazard rate function.

The application of this methodology to the Italian seismicity (Faenza *et al.*, 2003; Cinti *et al.*, 2004) and to the global (Faenza et al.,in press) displays clustering behaviour for large events. In spite of the first order similarity with the time behaviour of aftershock sequences, described by the ETAS model (Ogata, 1988), it was found the empirical data show a longer clustering time that the ETAS activity (Faenza *et al.*, 2004). These results may suggest the existence of different physical mechanism for aftershocks and large (independent) events, and/or that the ETAS model is not able to capture some important feature of seismicity.

In this paper, we want to investigate if a similar behaviour can be found with smaller events ($M4.0+$) in a low seismicity region (Central-Europe). In particular, we will study to what extend the ETAS model is able to reproduce the empirical hazard function for this data set.

2 A statistical model for the spatio-temporal distribution of earthquakes

In this work, we analyze from a statistical point of view the spatio-temporal distribution of earthquakes in Central-Europe. For this purpose, we first apply a statistical model called Proportional Hazard Model (PHM). Once the model is set up, we will compare the results with that of an ETAS model calibrated to the Central-Europe seismicity. This is done with the purpose to test whether the ETAS model is able to represent the real seismicity or whether deviation exists which can give us additional insight into the temporal occurrence of earthquakes.

2.1 Proportional Hazard Model

The Proportional Hazard Model (PHM) was first introduced by Cox (1972) and by Kalbfleisch and Prentice (1980). Here, we only briefly review the aspects that make this model appealing for earthquake distribution. For a more detailed description of the technical approach and the mathematical formulation we refer to Kalbfleisch and Prentice (1980), Faenza *et al.* (2003), and Faenza (2005). To set up the model, two types of random variables are considered: the inter-event time (IET), i.e., the time interval between two consecutive events, and the censoring time (CT), that is the time between the most recent event in the catalog and the end of the catalog. The censoring time are important in time dependent studies, where the probability of occurrence change with the time elapsed from the most recent event. Let us consider a set of N random variables and, for each of them, s explanatory variables (or covariates), that

is, for the j-th element of N, there is a row vector $\mathbf{z}$ of dimension s bearing any kind of information (quantitative and qualitative) that could influence the occurrence of the events. The main problem consists of assessing the relation between the distribution of the random variables and the vector of covariates $\mathbf{z}$. In PHM, the hazard function of a generic time t since the last event is:

$$\lambda(t; \mathbf{z}) = \lambda_0(t) \exp(\mathbf{z}\boldsymbol{\beta}) \tag{1}$$

where $\lambda_0(t)$ is an arbitrary unspecified base-line hazard function and $\boldsymbol{\beta}$ is a column vector of dimension s that gives the weight of each covariate. The hazard function $\lambda(t; \mathbf{z})$ is therefore composed of two parts, one with the temporal dependence ($\lambda_0(\cdot)$), and the other with the information about the process carried by other factors ($\exp(\mathbf{z}\ \boldsymbol{\beta})$).

There are three main aspects that make PHM appealing for earthquake distributions.

Firstly, this model is non-parametric for the temporal domain because it does not assume any specific form for base-line hazard function $\lambda_0(\cdot)$; therefore we do not impose any *a priori* assumption, regarding the temporal domain, on the earthquake occurrence process. In other words, we do not choose any arbitrary temporal distribution for fitting the events. Secondly, the method allows the integration of different factors, such as geophysical, geological and tectonic information, in the study of the earthquake forecasting. As a consequence, it is possible to merge information coming from different disciplines in the modeling of earthquake spatio-temporal distribution. Lastly, the technique is stable because it allows us to consider at the same time all the available data coming from different regions. This is possible through the vector of covariate, $\mathbf{z}$, that is attached to each one of the random variables and which can contain

any spatial/tectonic information on the subregion where IETs and/or CTs are sampled.

There is a key assumption below the model. The relationship between the temporal behavior of the random variable and the information variable $\mathbf{z}$ is addressed via a multiplicative factor. In equation 1 the covariates act multiplicatively on the hazard function and they do not depend on time. This means that the shape of the base-line $\lambda_0(\cdot)$ versus time is always the same for each area apart for a multiplicative factor that depends on the covariates. So, from a physical point of view, the mechanism of earthquakes occurrence, described by $\lambda_0(\cdot)$, is the same for different areas; only the parameters of the system can vary (i.e., $\exp(\mathbf{z}\boldsymbol{\beta})$). This assumption reflects the fact that the information variables can not modify the physics of earthquake occurrence, changing, the general shape of the hazard function, but they can only rescale its temporal trend.

The goal consists of estimating $\boldsymbol{\beta}$ and the non-parametric form of $\lambda_0(\cdot)$ in equation 1. The vector of coefficients gives the relative importance of each covariate; $\lambda_0(\cdot)$ gives important insights on the physics of the process. They are estimated using a Maximum Likelihood Estimation strategy, the detail of which can be found in Kalbfleisch and Prentice (1980) and Faenza (2005).

The evaluation of the hazard function is based on the empirical survivor function. There is a biunivocal relationship between the hazard function and the survivor function

$$S(t) = \exp(-\int_0^t \lambda(u)du). \tag{2}$$

For PHM, the relationship becomes

$$S(t;\mathbf{z}) = \exp[-\int_0^t \lambda_0(u)\exp(\mathbf{z}\boldsymbol{\beta})du] = S_0(t)^{\exp(\mathbf{z}\boldsymbol{\beta})}. \tag{3}$$

A simple way to display the trend of $\lambda_0(\cdot)$ is through the comparison of the empirical survivor function $S_0(\cdot)$ and the survivor function of a Poisson process. We apply a double logarithmic transformation

$$u(t) = \ln\{-\ln[S_0(\cdot)]\} \tag{4}$$

in order to obtain a RV, $u(t)$, asymptotically normally distributed (e.g. Kalbfleisch and Prentice, 1980; Faenza *et al.*, 2003). The transformation (4) applied to the survivor function of a Poisson process gives $u_p(t) = \ln\lambda + \ln(t)$, where λ is the mean of the distribution. Then, we can define the residuals $\epsilon(t)$ as

$$\epsilon(t) = u(t) - u_p(t); \tag{5}$$

The function $\epsilon(t)$ versus t shows the departures of the empirical survivor function from the theoretical Poisson distribution as a function of elapsed time. By looking at equation (4), and at the relation between the survivor function and the hazard function in equation (3), it is easy to figure out that the trend of $\epsilon(t)$ has a shape comparable to the trend of $\lambda_0(t)$.

Checking the PHM model

An important step of every modeling is the validation of the model, and checking if the assumptions behind it are in agreement with the real data. For this purpose, we perform a validation test on an independent data set, i.e., data that have not been considered at any step of setting up the model. This is done by dividing the available data in two parts, one used to set up the model (the *learning* phase), and the other to check the model (the *validation* phase). Each IET of the two data sets is transformed in order to form residuals (see Kalbfleisch and Prentice, 1980; and Faenza *et al.*, 2003). This transformation is a sort of statistical standardization and it changes the random non-negative

point process into a Poissonian process with rate 1 (see, for instance, Ogata, 1988). Therefore, if the model is appropriate, the residuals are expected to behave like an exponential distribution with $\lambda = 1$. The comparison between the cumulative residuals and the theoretical exponential curve is checked through a one-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971). This provides a goodness-of-fit test for the model.

2.2 ETAS Model

The Epidemic Type Aftershocks-Sequences (ETAS) model is a stochastic marked point process representing the occurrence of earthquakes of size larger than, or equal to, a threshold magnitude M_0, in the region and the period under consideration (Ogata, 1988). As in other triggering models, it is based on the principle that earthquakes are clustered in time and space because of occurrence of aftershocks; but, unlike those, it solves the debated problem to find the best way to identify clusters and to classify events (as mainshocks, aftershocks or foreshocks). In fact, although the overall seismicity is considered as the superposition of a background activity and of seismicity induced by previous earthquakes, its application to real data does not request any discrimination of events. A complete description of this parametric model and its formulation can be found in Ogata (1988, 1998).

In this paper, we used the 7-parameter space-time ETAS model; given the epicentral coordinates $(x_i; y_i)$, the magnitudes (m_i) and the occurrence times (t_i), the conditional intensity function can be written as

$$\lambda(t,m,x,y) \;=\; \left[\mu g(x,y) + \sum_{i:t_i<t} \frac{K e^{\alpha(m_i-M_0)}}{(t-t_i+c)^p} \frac{(q-1)d^{2(q-1)}}{\pi[(x-x_i)^2+(y-y_i)^2+d^2]^q}\right] \beta e^{-\beta(m-M_0)} \tag{6}$$

where μ, K, c, α and p are the 'standard' ETAS parameters (Ogata 1988; 1998); $g(x,y)$ is the spatial density function of background events; $\beta = b\ln(10)$, with b being the parameter of the Gutenberg-Richter relation. For the spatial distribution of triggered aftershocks, we used the normalized function of Console *et al* (2003), imposing however that the parameter d scales with magnitude m, according to the empirical relation of Well and Coppersmith (1994): $d^2 = d_0^2 10^{0.91(m-M_0)}$ where d_0 is a radius of a circular earthquake rupture with magnitude M_0.

3 Application and Results

3.1 The study of the real seismicity with PHM

For our analysis, we use the Grünthal and Wahlström (2003) catalogue. It is a revised M_w catalogue for Central-Europe, which spans the time window 1300-1993 with a threshold magnitude of M_w3.5. To cover the time window 1994-2004, we use the International Seismological Center (ISC) on-line catalogue (figure 1). In the latter catalogue, we convert all magnitude scales to M_w according to the relations given by Grünthal and Wahlström (2003).

We analyze the events with M_w4.0+ with shallow hypocenter (depth $< 50km$) in the time 1960-2004 (360 events). The completeness of the data set has been checked through the cumulative number of events and the Gutenberg-

Richter relation (figure 2). A constant linear increasing trend stands for a constant rate of earthquake production over time, and it is used to detect the catalogue completeness, under the hypothesis of stationarity. The slope of the line before 1980 and after 1990 suggests that the seismic rate is constant in these two time window. During 1981-1989 the activity is lower. Similar non-stationarity in the event production over decade has been already recognized for other regions (Selva and Marzocchi, 2005; Lombardi and Marzocchi, 2007). We have checked that this decrease is not an artefact by changing the dimension of the area (excluding Romania, or the Apennines, or the Alps, or the Balkan, respectively); increasing the depth up to 80 km; and decreasing or increasing the threshold magnitude. We found the same trend in all cases. These tests indicate that this reduced activity in the time 1981-1989 is not due to some in-homogeneity in terms of magnitude allocations and/or density of array in the regional seismic catalogue. The picture shows also that the Gutenberg-Richter has heterogeneity in the magnitude allocations. In the magnitude range $4.0-4.7$ it has a slop $b = 1.07 \pm 0.04$ and in the range $4.8-6.3$ the slope is $b = 1.30 \pm 0.07$.

The analysis of the whole Central-Europe may be strongly biased since, due to the different tectonic domains, the events in this whole region can be hardly considered as a homogeneous sample. In order to account for the inhomogeneities in the spatial distribution of the events, we divide the surface into areas of equally areal extension. Each node of the grid is the center of a circle and, in order to cover the whole surface, its radius R is set equal to the mean value of the half diagonal of the cell. In this paper, we present the results for a grid with a radius of $R = 50km$ (figure 2), but other radii and

grids have been used to check the stability of the results.

In a following step, we select the areas that contain at least one earthquake; for each circle the IETs among the earthquakes inside the circle, one CT relative to the time elapsed since the most recent event and a vector of covariate $\mathbf{z}$ are evaluated. We analysed 123 areas. In the context of this application, $\mathbf{z}$ is a three dimensional vector composed of *i)* the logarithm of the occurrence rate; *ii)* the magnitude and *iii)* the depth of the events from which we calculate the IET and the CT. As a consequence, the coordinates of the vector $\boldsymbol{\beta}$ are co-related to the logarithm of the rate of occurrence (β_1), the magnitude (β_2) and the depth (β_3) of the seismic activity. The occurrence rate is calculated as the number of independent events in each cell divided by the length of the used catalogue (45 years). To identify the number of independent events we used two distinct strategies. In the fist approach, we mark the independent events in the whole catalogue by using the Reasenberg approach, with a standard setting of parameters($Q = 10$, $P = 0.99$,$\tau_0 = 2$day, $\tau_{max} = 10$days] (Reasenberg, 1985). Then, in the PHM application, for each cell we simply count the number of events marked as independent. In the second approach, in each cell we apply the statistical declustering technique developed by Hainzl *et al.* (2006). This method provides the fraction of independent events for each cell. The results of these two methodologies are not substantially different in terms of the hazard function, probability map and synthetic test (see below). Using the two approach allows comparison of the results and, therefore, adds to the confidence to the conclusions. In the following, we show the results based on the Reasenberg declustering algorithm.

The first important result is that the only element of the vector $\mathbf{z}$ which is

statistically significant is the rate of occurrence, since the weight $\beta_1 = 1.1 \pm 0.1$. While β_2 and β_3 do not differ from zero. This means that the rate of occurrence is the only covariate, among the analyzed ones, able to modify the statistic of earthquake occurrence.

In figure 3, we show the residuals defined as the double logarithm transformation of the empirical PHM survivor function minus the double logarithm transformation of the Poisson survivor function (Kalbfleisch and Prentice, 1980; Faenza *et al.*; 2003, and Faenza, 2005). The trend versus time of the residuals mimics the behaviour of the hazard function. The decreasing trend means that the earthquakes tend to be more clustered than a simple Poisson process. The negative trend lasts for a few years after an event; then it becomes almost flat as expected for a Poisson process. Here, we note that figure 3 is the plot of $\lambda_0(\cdot)$. Since we found that only the rate of occurrence is able to modify the hazard function; for each area $\lambda(t; \mathbf{z})$ will be rescaled depending on the value of its rate. Since β_1 is positive, areas with higher rates will have higher hazard functions.

To check the validity of our hypothesis, we applied the residual analysis to the *learning* and *validation* data set. In figure 4, the goodness-of-fit for the two data set is reported. We used the time interval 1960-1990 (232 events) for the *learning* phase, and the one 1990-2004 (128 events) for the *validation*. The figure plots in plate (a) and (b) the same theoretical curves (an exponential distribution with rate equals 1) and the real transformed residual, for the *learning* (figure 4a) and *validation* (figure 4b) data sets. We quantify the goodness-of-fit with a one-sample Kolmogorov-Smirnov test (see Gibbons, 1971). The same theoretical curve fits very well both data sets. Plate (c) plots the two-sample

Kolmogorov-Smirnov test for of the *learning* and *validation* data sets. The levels of significance (α) at which the null hypothesis of equal distributions is rejected are reported in figure. They are > 0.99; we can than conclude that PHM is a suitable way to represent $M_w 4.0+$ Central-Europe seismicity.

As a final step, we calculate from the empirical survivor function $S(t; \mathbf{z})$ a probability map of occurrence of the next shallow $M_w 4.0+$ earthquakes. The conditional probability of earthquake occurrence in the next τ years given the CT (t) as the time elapsed since the most recent event

$$P(t, \tau; \mathbf{z}) = \frac{S(t; \mathbf{z}) - S(t + \tau; \mathbf{z})}{S(t; \mathbf{z})}. \tag{7}$$

In figure 5a, we report the map for the probability of occurrence in the next $\tau = 5$ years. The areas with the higher probability of occurrence are the Northern Apennines and Eastern Alps. Not surprisingly, these are the areas with the higher seismic rate.

Since the seismic hazard in Romania is mostly driven by deep events, we repeat the analysis outlined so far including also deep events. The coupling between shallow and deep events is in general different from the one of shallow seismicity only. PHM allows the integration in the same statistical analysis of heterogeneous events, through the vector of covariate $\mathbf{z}$, see section 2.1. Deep events are mainly localized in the Vrancea region, Romania. We analyze 626 shallow plus deep events for the same time-magnitude window as before.

The result does not substantially differ from the one with only shallow events. Once again, the only variable which is able to modify the statistic of earthquake occurrence is the occurrence rate with $\beta_1 = 0.7 \pm 0.3$. Remarkably, the depth is not statistically relevant since β_3 is not different from zero. The hazard function has a decreasing trend and than a constant shape as the one

from shallow events (figure 3 a and b). So the temporal clustering remains a stable feature for representing the seismicity. While the length of the cluster is almost the same, the inclusion of deep events leads to a marked decrease in the intensity of the cluster and to a slower decay rate. The probability map changes, since there is a significant increase of the rate in the Romania area (figure 5b).

To quantify the forecast ability of PHM in Central Europe seismicity, we applied the Molchan diagram (Molchan, 1990; Kossobokov, 2006), comparing the results of PHM and Poisson model. The Molchan diagram plots the proportion of missed events (ν) against the fraction of time-spacefraction on alarm (τ). The forecasts are quantified for all possible threshold probabilities of alert p_0 ($[0:1]$). The extreme cases of $\tau = 0$ for $p_0 = 1$ and $\tau = 1$ for $p_0 = 0$ can be seen as the optimistic and pessimistic point of view, respectively. In case of purely random forecasts, the diagonal consists of the diagonal joining the points (0-1) - (1-0). We forecasted the time January 1996 - December 2004, 9 years in total. The interval time of alert Δt_j is 6 months, and $j = 1, 18$ to obtain the total forecast time, 9 years. For the spatial distribution, we consider active only the cells in which at least one event occurred in the time 1960-1995. Events in the time period 1996-2004 that occurred outside these circles are considered lost and are not used in the forecast. We find that 9 out of 93 events occurred outside the active cells, therefore we can conclude that 90% of the events occurred in the same areas as in the past; this stands for a spatial clustering of earthquake occurrence; while 10% of events occurred randomly in space. N_j defines the number of events occurred in time Δt_j, with $N_{tot} = \sum_j N_j = 84$. For each active zone i of the grid, we calculate the

probability p_{ji} for an event in the next δt_j; to calibrate the models (PHM and Poisson) we used the data available before the Δt_j under consideration. If p_{ji} is higher than p_0, the cell is in alarm for the jth time of alert. If an event occurred in the cell i in Δt_j, the event was predicted. For each Δt_j we will have a fraction ν_j of missed events. The calculation is done for p_0 from 0 to 1. Figure 6 summarizes the results for PHM versus Poisson, for all Δt_j. The value of ν is the average of the $18th$ $\nu_j, j = 1, \cdots, 18$. The curve of a better predictive model lies below the one of the a less predictive model, therefore we can conclude that PHM has a higher predicative power than a Poisson model. The figure shows also a departure of the Poisson model from the random forecast (i.e., the diagonal) which probably stem from the fact that the earthquakes occurrence is not homogeneous in space. Similar results for the Molchan diagram are found for Δt_j equals to 1 year or 3 months.

3.2 ETAS synthetic test

3.2.1 Synthetic catalogue generation

For the detection of the 7 parameters of the ETAS model (equation 6) we use the shallow events with M_w3.5+ from 1978 to 1994 (Grünthal and Wahlström, 2003), 162 events in total. Here, we choose a lower magnitude threshold since aftershock properties are better confined for larger magnitude intervals. For the joint inversion we follow Ogata *et al.* (1993). We fix the parameters linked to the spatial function; $q = 1.5$ for consistency with the theory of elasticity at infinite distance; $d_0 = 0.40km$ for $M_{w0} = 3.5$ from Well and Coppersmith (1994); we have therefore 5 free parameters. We obtain $\mu[\mu_{5perc}, \mu_{95perc}] = 0.0202[0.0188, 0.0202]$ per day; $K[K_{5perc}, K_{95perc}] =$

$0.018[0.017, 0.023]$; $c[c_{5perc}, c_{95perc}] = 0.009[0.008, 0.017]$ per day;
$p[p_{5perc}, p_{95perc}] = 0.98[0.97, 1.01]$, $\alpha[\alpha_{5perc}, \alpha_{95perc}] = 0.54[0.22, 0.64]$. For the error estimation, we repeat the parameters estimations for 1000 catalogues, where Gaussian errors are added randomly to magnitude (with a standard deviation of $\sigma = 0.1$) and location ($\sigma = 5km$) to the original catalogue.

For the generation of the synthetic ETAS catalogues, an inverse method is used according to Felzer *et al.* (2002). In each realization, events are generated sequentially: first time, then magnitude and epicenter coordinate. For the spatial distribution of independent events ($g(x, y)$), a Gaussian filter with correlation distance equal to $20km$ is applied (Frankel, 1995). The declustering method of Zhuang *et al.* (2002) is applied to detect independent events; 69% events are marked as independent. A similar result has been founded with the declustering procedure of Hainzl *et al.*, 2006. The cascade process is set for events with M_w3.5+, but later, we analyze only events with M_w4.0+. For the selection of the magnitude, the b of the Gutenberg-Richter empirical low is evaluated by using a Maximum Likelihood Estimation method (Marzocchi and Sandri, 2003), and we get $b = 0.91 \pm 0.07$.

3.2.2 PHM applied to synthetic catalogues

We generate 500 synthetic catalogues of time duration 45 years, and PHM is applied to all of them separately, following the same procedure outlined in the section 3.1.

In this synthetic study, the vector $\mathbf{z}$ is a two dimensional vector made by the logarithm of the occurrence rate and the magnitude. The occurrence rate is calculated as before, by using the two methodologies separately. The results

shown here are the ones using the Reasenberg method, but we can draw the same conclusion by using the Hainzl *et al.* method.

As before we find that the rate of occurrence is the only statistically significant parameter, with an average $\beta_1 = 1.3 \pm 0.1$.

Figure 7 compares the residuals (see above) for the real catalogue (RC) and the synthetic ones (SCs). Also in cases of SCs a negative trend is shown. This means that earthquakes tend to be clustered. This is not surprising, since the ETAS model is constructed imposing aftershock clustering. However, even if the RC and the SCs show the same structure of the hazard function, with a negative trend that becomes almost flat for larger times, the clustering time is different for the RC and SCs. Figure 7 shows that the flattening occurred in synthetic catalogues earlier than in the real one which occurs at about 2 years, which is determined by the algorithm of Mulargia and Tinti (1985) (see below).

A quantitative test of the difference between the real and the synthetic clustering properties can be performed through a two-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971) on the empirical survivor functions (figure 8). In figure 9, we report the cumulative of the significance level α for the rejection of the null hypothesis; it is small enough to reject the null hypothesis; this means that the ETAS model is not able to capture the complete picture of real seismicity. This means that for 85% of ETAS randomly generated samples the Kolmogorov-Smirnov test would reject with 95% confidence the real series as being drawn from the same distribution.

Going into the detail of the behaviour of the hazard function of the SCs, it seems that its first part (short term) fits the real hazard function really

well, while for longer times a substantial difference occurs. The real hazard function has two different trends: it decreases till reaching the background. By using a non-parametric technique developed by Mulargia and Tinti (1985), we detect the time at which the two behaviours change. The changing-point algorithm (Mulargia and Tinti, 1985) is a technique to see if two parts of a sample belong to different populations. Applying the algorithm to the real survivor function, we get a change point at 2.1 years (see figure 7). Applying the two-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971) to the first part ($t \leq 2.1$ years) of the real and synthetic survivor function, we find a good agreement between them, as shown in figure 9.

As a further test to verify the agreement between the ETAS model and the short term part of the RC, we fit the real survivor function with the one coming from an ETAS model:

$$S(t) = \exp[-\int_0^t \lambda(u)du] = \exp\{\lambda_0 t - \frac{k'}{1-p}[(c+t)^{1-p} - c^{1-p}]\} \qquad (8)$$

where k' is the average productivity factor. The purpose of this last fitting procedure is to see if the p of the real hazard function is similar to the one found in the ETAS inversion. In fact, p is the parameter which drives the temporal decay in the hazard function. The fit gives a value of $p = 0.57$ considering the long term, while it gives a value $p = 0.92$ if we consider only the events in the short term.

Figure 10 reports a graph of the goodness-of-fit of the PHM applied to the SCs. As in the real case, we used the first 30 years of catalogue for *learning* and the second part for *validation*. The goodness-of-fit is quantitatively evaluated through a one-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971). The significance levels at which the null hypothesis of equal distribution is rejected

is > 0.95 in all simulations.

4 Summary and Conclusions

In this paper, we perform a statistical analysis of the spatio-temporal distribution of Central-Europe seismicity. We have applied a non-parametric technique in order to characterize the features of the spatio-temporal earthquake occurrence. In particular, we have compared these non-parametric results with that of an ETAS type seismicity.

Our analysis shows that clustering dominates the $M_w 4.0+$ earthquake activity in Central-Europe in the first two years. Subsequently, the events tend to occur as Poisson random events, since the hazard function shows a constant shape. The physical reason for the short-term clustering is likely to be linked to stress interactions between faults (Stein, 1994).

Remarkably, a similar result has been found in previous studies on large time and magnitude scale by Kagan (1991); Kagan and Jackson (1991); Kagan and Jackson (2000); Cinti *et al.* (2004) and Faenza *et al.*, (submitted). Therefore, it seems that the clustering is a scale independent aspect with respect to the magnitude of the events (from low-medium to large) and spatial domain (for a regional to a global scale).

The inclusion of deep events, mostly located in the Vrancea (Romania) regions, yields a transform of the hazard function in agreement with the observation that such earthquakes have significantly less aftershocks lasting for much shorter times than comparable events at shallow depth (Trifu and Rasulian, 1991; Enescu *et al.*, 2005).

The comparison with ETAS-type seismicity underlines that the space-time

ETAS model is able to capture the behaviour of real seismicity on the short time scale, i.e., for a couple of years. This short term behaviour can be linked to the aftershocks activity. While, in the real seismicity, the clustering stays longer; the ETAS activity returns to the constant behaviour (Poisson) within a shorter time. A possible reason can be addressed to visco-elastic relaxation processes, introducing a long term memory effect which is not accounted for in the ETAS parameterization.

Acknowledgments: This work was supported by the Deutsche Forschungsgemeinschaft (SCHE280/14-2) and the EU-project SAFER. We thank Dr. Grünthal for providing the data. We would like to thank V. Kossobokov and one anonymous reviewer for the punctual and precise comments that improve the clarity of the manuscript.

Bibliography

Cinti, F.R., Faenza, L., Marzocchi, W., Montone, P., 2004. Probability map of the next M≥ 5.5 earthquakes in Italy. Geochem. Geophys. Geosys. 5, Q11003, doi:10.1029/2004GC000724.

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

Console R., Murru M. and Catalli, F., 2006. Physical and stochastic models of earthquake clustering, Tectonophysics, 417, 141-153.

Cox, D.R., 1972. Regression models and life tables (with discussion). J. R. Stat. Soc. B., 34, 187-220.

Enescu, B., Ito, K., Radulian, M., Popescu, E., Bazacliu, O., 2005. Multifractal and chaotic analysis of Vrancea (Romania) intermediate-depth earthquakes: investigation of the temporal distribution of events, Pure Appl. Geophys., 162, 249-271.

Ellsworth, W.L., Matthews, M.V., Nadeau, R.M., Nishenko, S.P., Reasenberg, P.A., Simpson, R.W., 1998. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities. In: Proceedings of the second point meeting of the UJNR panel on earthquake research, pp. 135-149, Geographical Survey Institute.

Faenza, L., Marzocchi, W., Boschi, E., 2003. A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an appli-

cation to the Italian catalogue. Geophys. J. Int. 155, 2, 521-531.

Faenza, L., Marzocchi, W., Lombardi, A.M., Console, R., 2004. Some insights into the time clustering of large earthquakes in Italy. Annals of Geophysics, 47, 5, 1635-1640.

Faenza, L., 2005. *Analysis of the spatio-temporal distribution of large earthquakes*, Ph.D. Thesis, Università degli Studi di Bologna, *Alma Mater Studiorum*, Bologna, Italy. Published on Earth Print: http://hdl.handle.net /2122/3160

Faenza, L., Marzocchi, W., Serretti, P., Boschi, E., 2007. On the spatio-temporal distribution of M 7.0+ worldwide seismicity. Tectonophysics, in press.

Felzer, K.R., Becker, T.W., Abercrombie, R.E., Ekström, G.& Rice, J., 2002. Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.2 Landers earthquake, *J. Geophys. Res.*, **107**, B9, 2190, doi: 10.1029/2001JB000911.

Frankel, A., 1995. Mapping Seismic Hazard in the Central and Easter United States, *Seism. Res. Let.*, **66**, 8-21.

Gibbons, J.D., 1971. Non-parametric Statistical Inference. 306 pp., McGraw-Hill, New York.

Grünthal, G. & Wahlström, R., 2003. An Mw based earthquake catalogue for central, northern and northwestern Europe using a hierarchy of magnitude

conversions, *J. Seimol*, **7**, 507-531.

Hainzl, S., Beauval, C. & Scherbaum F., 2006. Estimating background activity based on inter-event-time distribution, *Bull. seismol. Soc. Am.*, **96**, 313-320.

Kagan, Y.Y., 1991. Likelihood analysis of earthquake catalogues. Geophys. J. Int. 106, 135-148.

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

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

Kalbfleisch, J.D., Prentice, R.L., 1980. The Statistical Analysis of Failure Time Data. John Wiley, New York.

Kossobokov, V.G., 2006. Testing earthquake prediction methods:¡¡The West Pacific short-term forecast of earthquakes with magnitude MwHRV $\geq$ 5.8¿¿. Tectonophysics, 413, 25-31.

International Seismological Center, On-line Bulletin, http://www.isc.ac.uk, Internatl. Seis. Cent., Thatcham, United Kingdom, 2001

Marzocchi, W., Sandri, L., 2003. A review and new insights on the estimation of the *b*-value and its uncertainty. Annals of Geophysics, 46, 6, 1271-1282.

Lombardi A.M., Marzocchi W., 2007. Evidence of clustering and non-stationarity in the time distribution of large worldwide earthquakes. J. Geophys. Res., 112,

B02303, doi:10.1029/2006JB004568.

Molchan G.M., 1990. Strategies in strong earthquake prediction. Phys. Earth Plan. Int. 61, 84-98.

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

Nishenko, S.P., 1985. Seismic potential for large and great interplate earthquakes along the Chilean and Southern Peruvian margins of South America: a quantitative reappraisal. J. Geophys. Res. 90, 3589-3615.

Ogata, Y., 1988. Statistical models for earthquake occurrence and residual analysis for point processes, *J. Am. Stat. Assoc.*, **83**, 9-27.

Ogata, Y., Matsuúra, R.S., Katsura, K, 1993. Fast likelihood computation of the epidemic type aftershock-sequence model, *Geophys. Res. Lett.*, **20**, 2143-2146.

Ogata, Y., 1998. Space-time point-process models for earthquake occurrence, *Ann. Inst. Statis. Math.*, **50**, 379-402.

Parsons, T., 2005. Significance of stress transfer in time-dependent earthquake probability calculations. J. Geophys. Res., 110, B05S02, doi:10.1029/2004JB003190.

Posadas, A., Hirata, T., Vidal, F., 2002. Information theory to characterize the spatio-temporal patterns of seismicity in the Kanto region. Bull. Seismol. Soc. Am. 92, 600-610.

Reasenberg, P., 1985. Second-order moment of central California seismicity, 1969-1982. J. Geophys. Res., 90, 5479-5495.

Schorlemmer, D., Gerstenberger, G.C., Wiemer, S., Jackson, D.D., 2007. Earthquake likelihood model testing. Seismological Research Letters, 78, 1, 17-28.

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

Sornette, D., Knopoff, L., 1997. The paradox of the expected time until the next earthquakes. Bull. seism. Soc. Am., 87, 789-798.

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

Trifu, C.I., Radulian, M., 1991. Frequency-magnitude distribution of earthquakes in Vrancea: Relevance for a discrete model. J. Geophys. Res., 96(B3), 4301-4311, 10.1029/90JB02221.

Vere-Jones, D, 1970. Stochastic models for earthquake occurrence (with discussion). J.R.Stat.Soc., B31, 1-62.

Vere-Jones, D, 1978. Earthquake prediction: a statistician's view. J. Phys. Earth, 26, 129-146.

Wells, D.L., Coppersmith, K.J., 1994. New Empirical Relationships among Magnitude, Rupture Length, Rupture Area, and Surface Displacement. Bull.

seism. Soc. Am., 84, 974-1002.

Working Group on California Earthquake Probabilities, 2003. Earthquake Probabilities in the San Francisco Bay Region (U.S. Geologic Survey, Denver), USGS Open File Report 03-214.

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

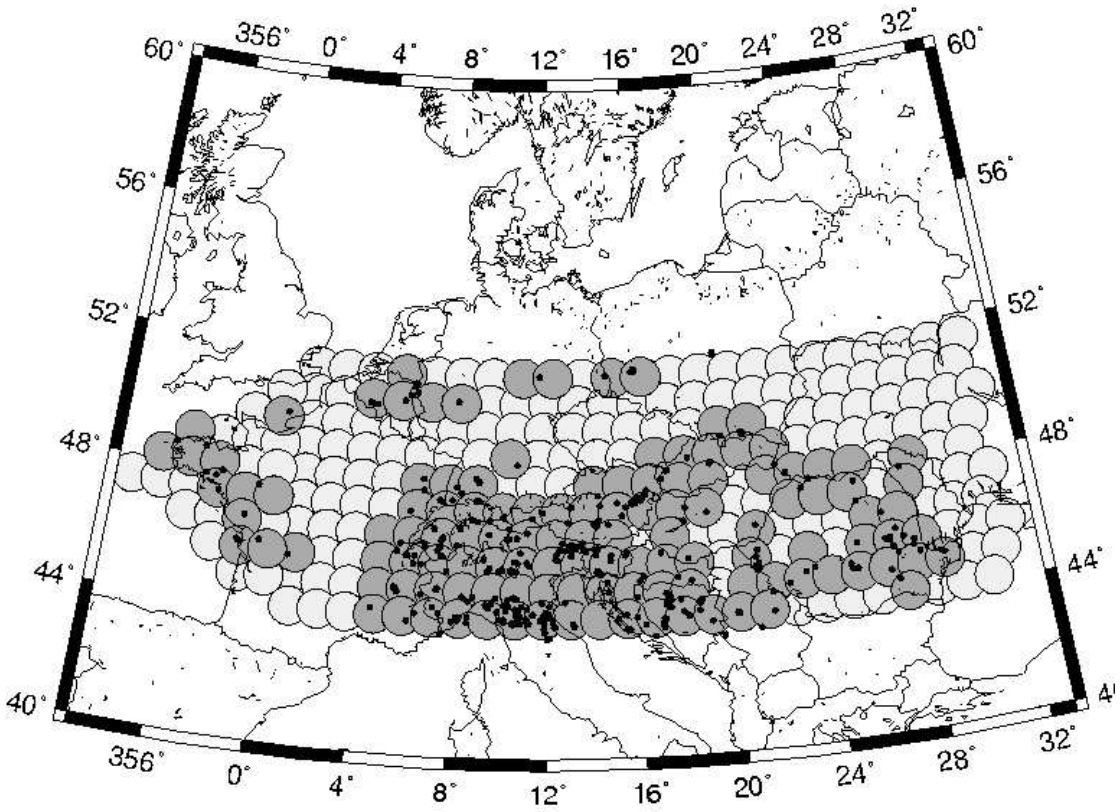

Figure 1: Target area: The dark circles refer to the selected areas, while the black dots are the earthquakes used in the analysis, with M_w4.0+ since 1960 (Grünthal and Wahlström, 2003; ICS).

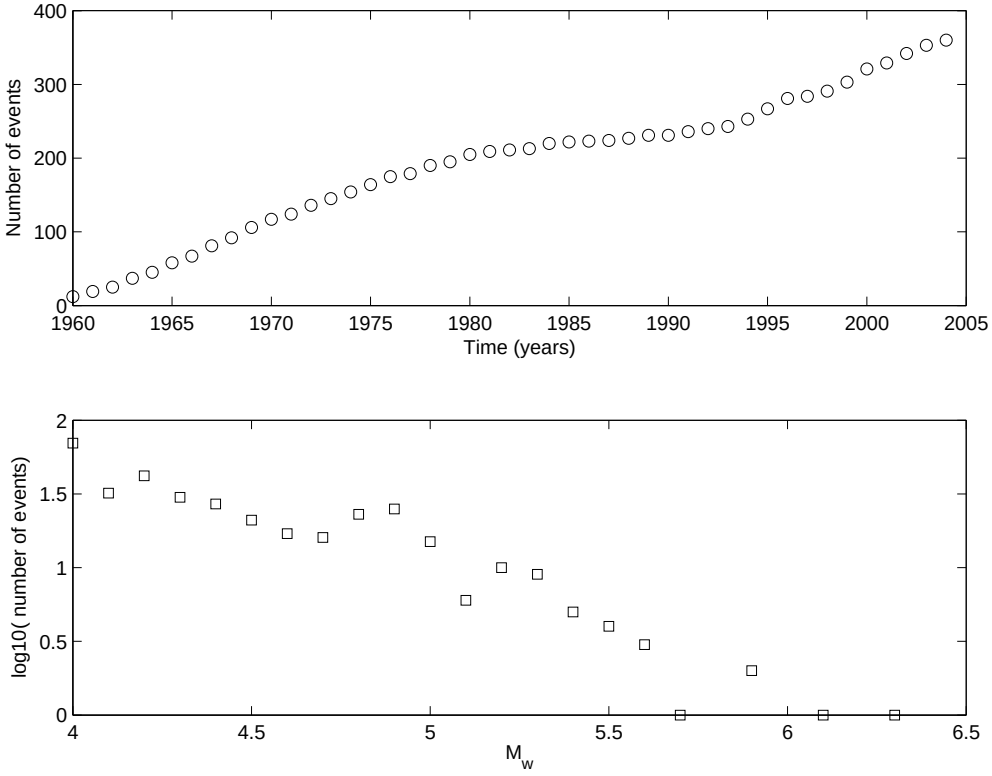

Figure 2: Top panel: Cumulative number of events as a function of time. The constant linear trend stands for a constant rate of earthquake production. Bottom panel: Frequency-magnitude distribution for the of the time 1960-2004. The straight line suggests the completeness of the data set.

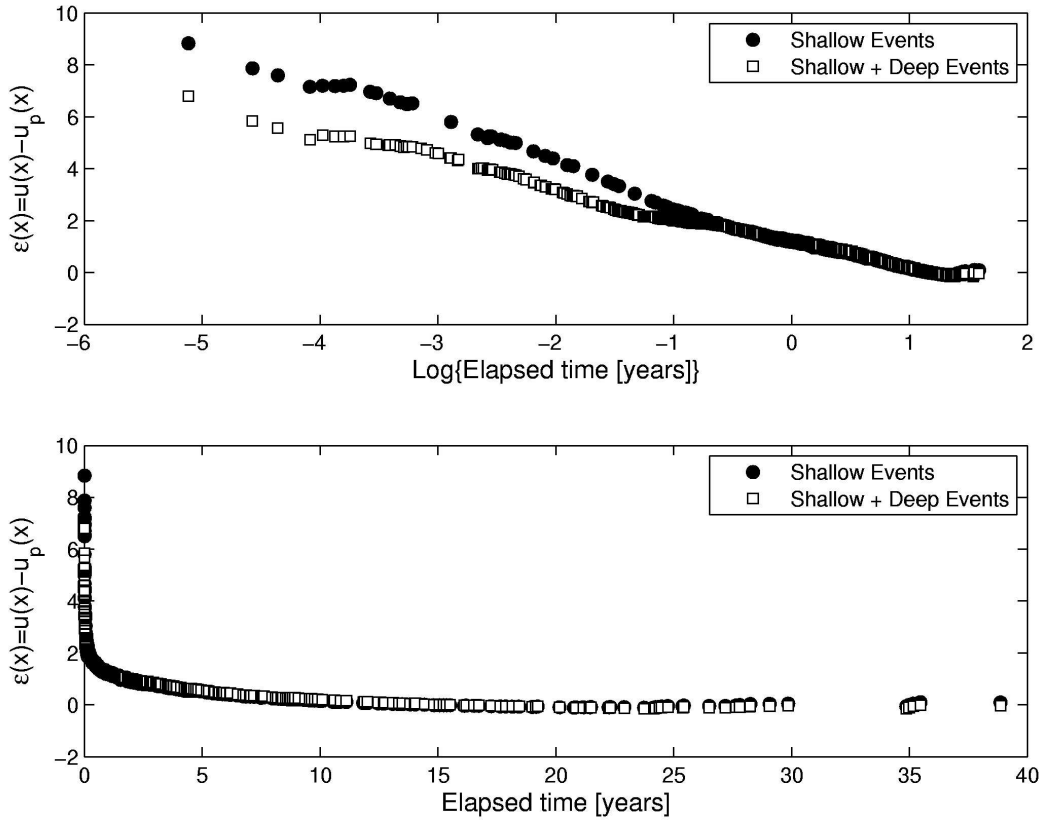

Figure 3: Plot of the residuals for the real catalogue as a function of the time elapsed since the most recent event (top panel, in logarithm 10 for the time; bottom panel, linear time scale. The residuals mimic the time behaviour of the $\lambda_0(\cdot)$. The result is shown for shallow seismicity alone (depth $\leq 50km$) and for shallow and deep events together.

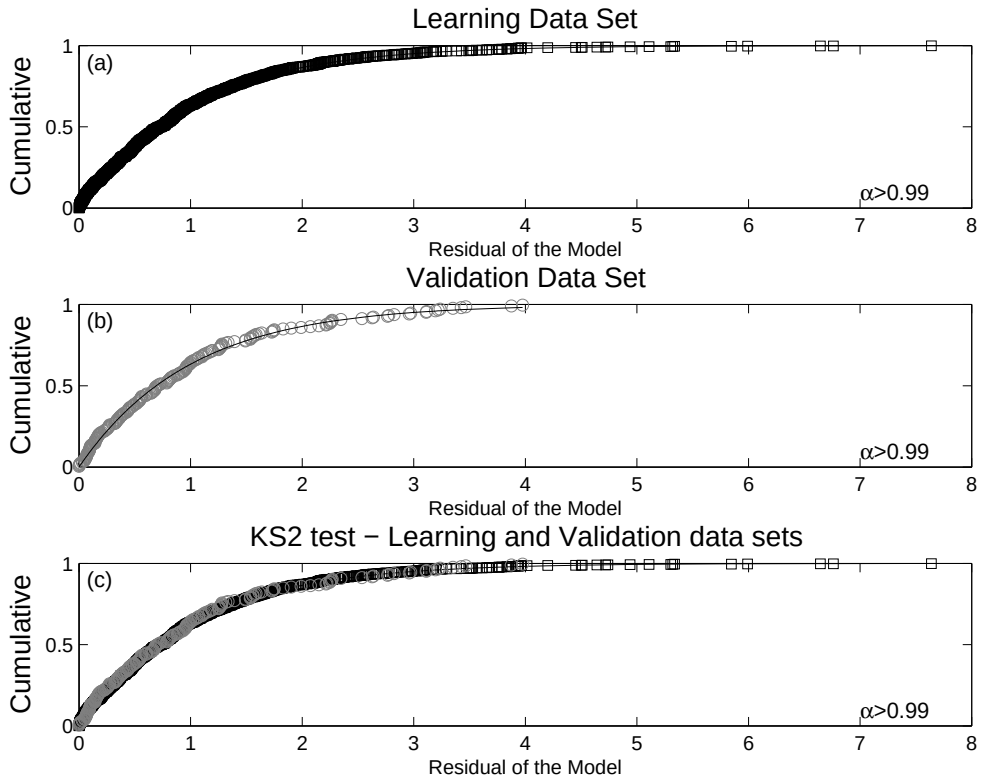

Figure 4: (a) Empirical (squares) and theoretical (solid line) cumulative functions for the *learning* data set. (b) The same as for (a), but now for the *validation* data set. (c) Two-sample Kolmogorov-Smirnov test for the *learning* empirical (black squares) and *validation* empirical (gray circles) cumulative functions. The plot shows also the theoretical curve.

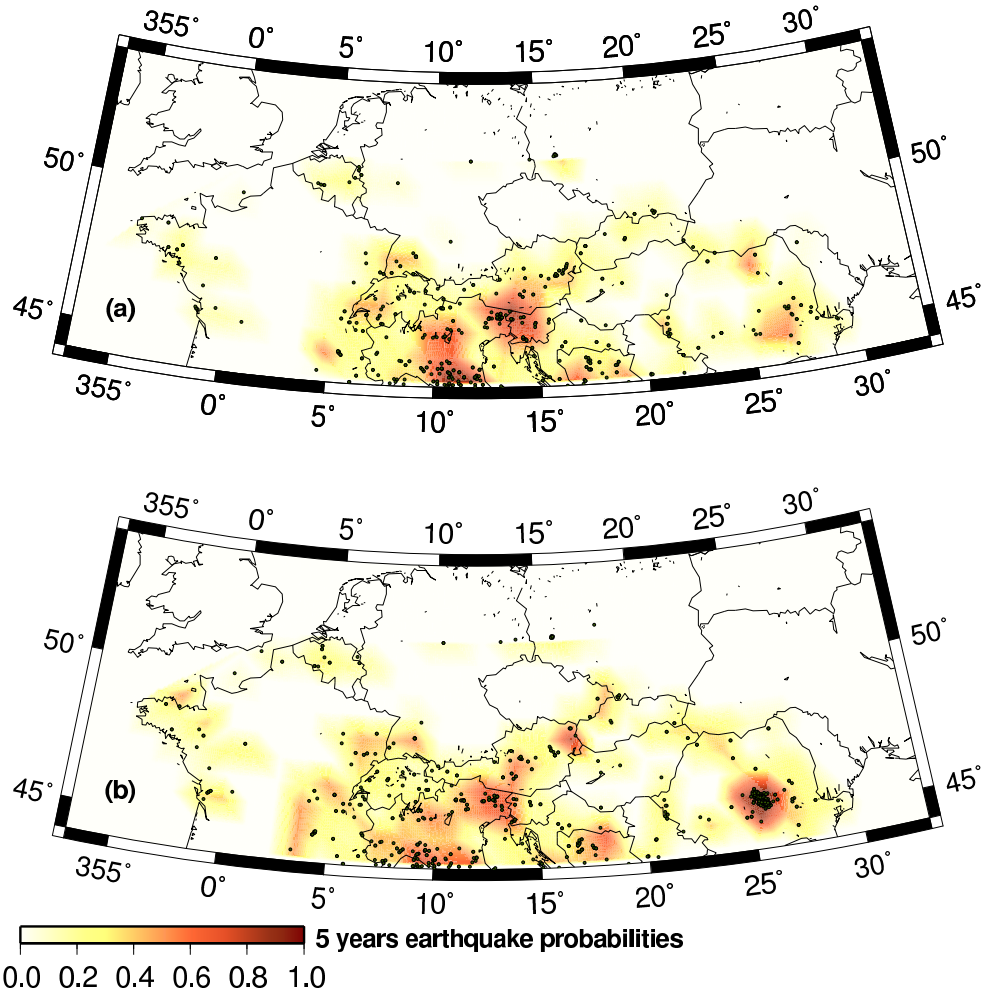

Figure 5: Probability map for M_w4.0+ earthquakes in Central-Europe in the next 5 years: (a) shallow and (b) shallow plus deep events.

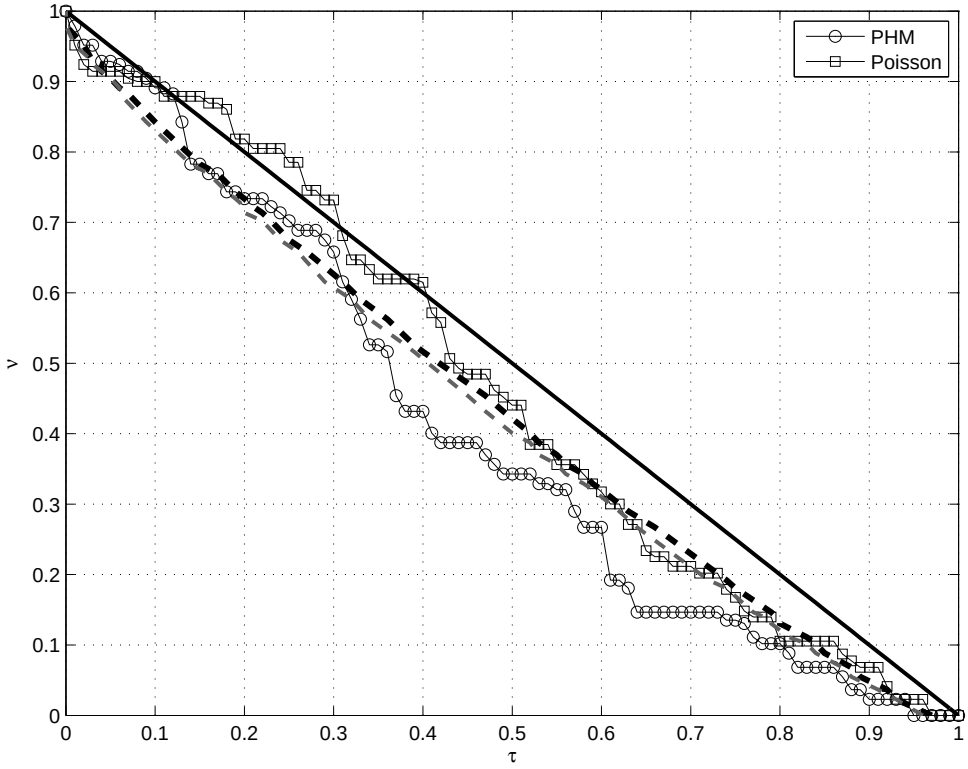

Figure 6: Molchan diagram; it represents the fraction of missed events (ν) against the fraction of time-space alarm (τ). The figure shows also the diagonal as the purely random forecast (solid black line); with its 99% (dashed gray line) and 95% confident limit (dashed dark line).

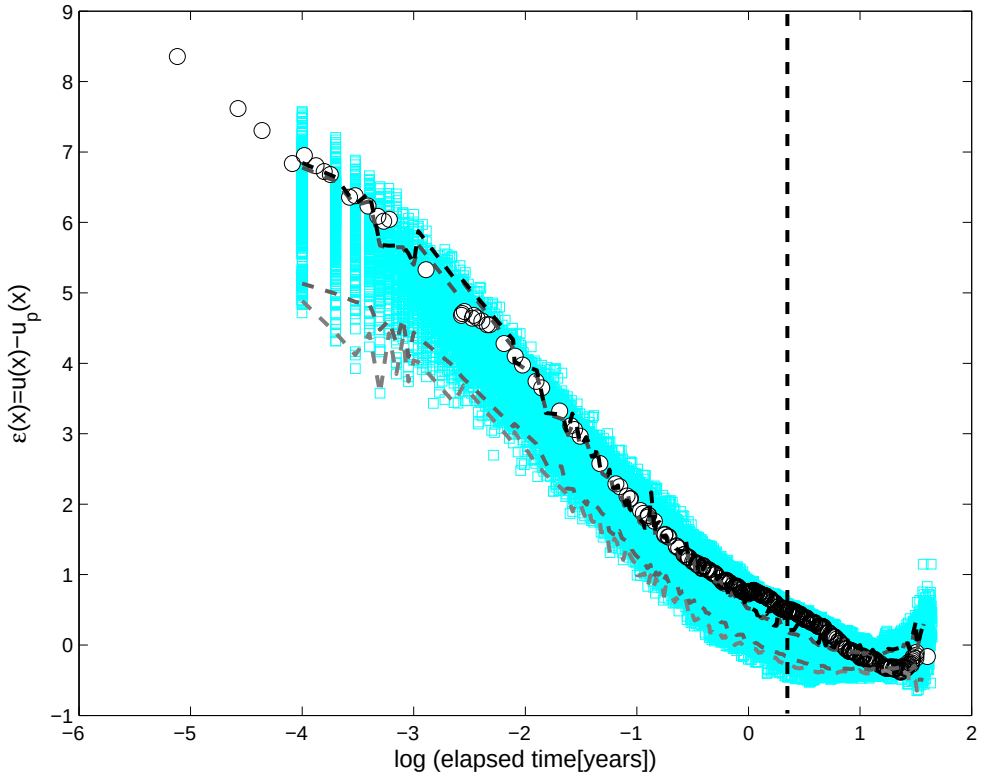

Figure 7: Plot of the residuals of the real catalogue (white circles) and synthetic data (squares) as a function of the elapsed time since the most recent event. The dash black line marks the value of the changing-point between short and long term behaviour.The picture shows also the 1-st, 5-th, 95-th and 99-th per cent percentiles of the distribution, in increasing gray scale.

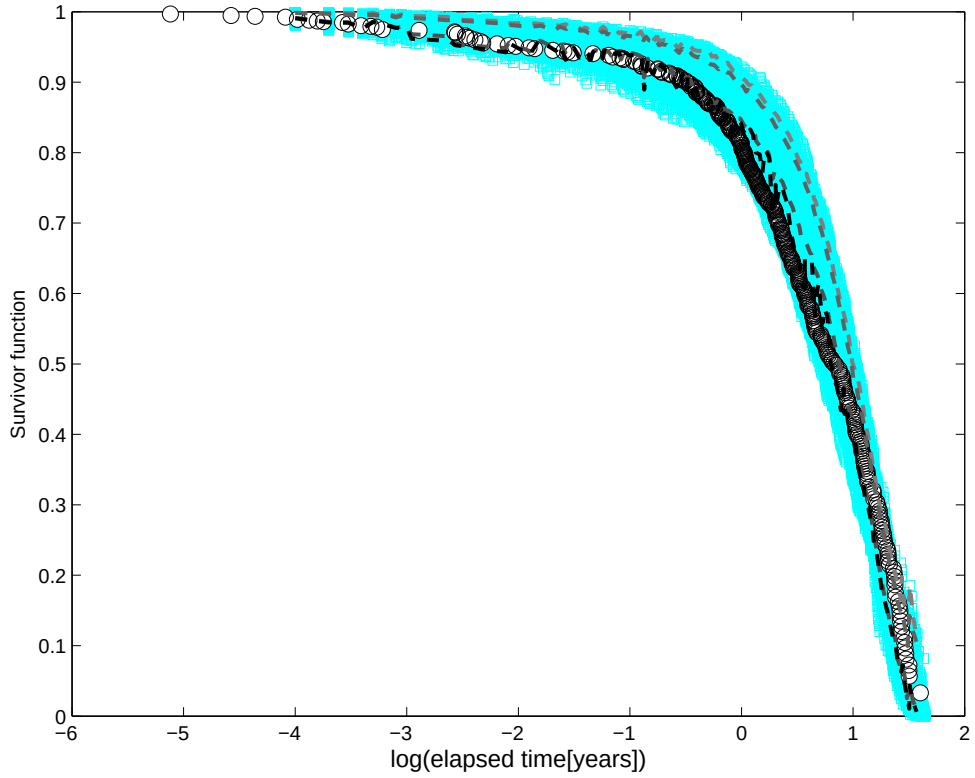

Figure 8: Plot of the empirical survivor function of the real data (white circles) and synthetic one (squares). The picture shows also the 1-st, 5-th, 95-th and 99-th per cent percentiles of the distribution, in increasing gray scale.

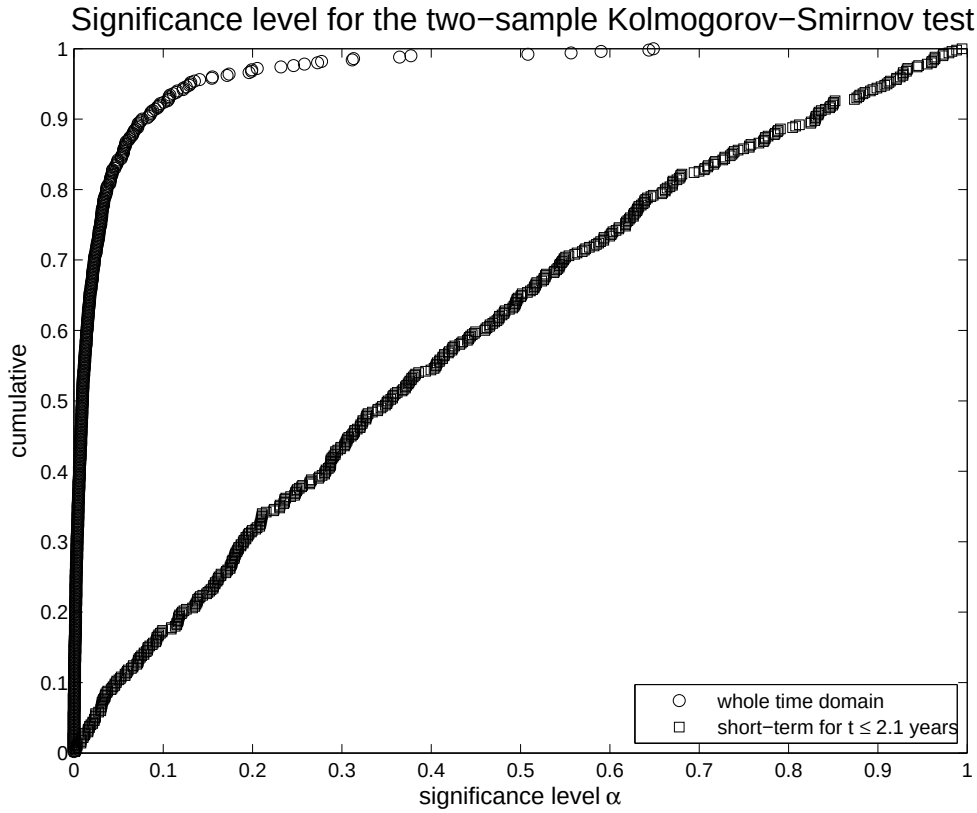

Figure 9: Empirical cumulative function of the significance level α of the two-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971).

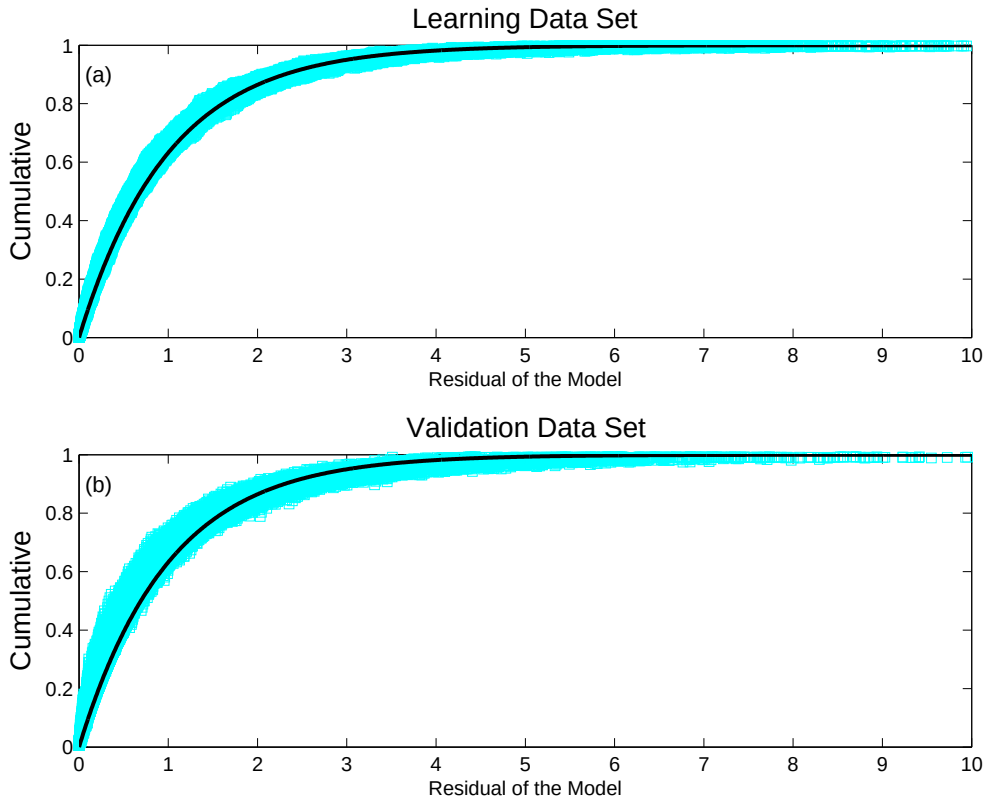

Figure 10: (a) Empirical (squares) and theoretical (solid line) cumulative functions for the *learning* data set. (b) The same as for (a), but now for the *validation* data set.