THE ETAS MODEL FOR DAILY FORECASTING OF ITALIAN SEISMICITY IN CSEP EXPERIMENT

Anna Maria Lombardi and Warner Marzocchi

Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Rome, Italy.

Revised version submitted to Annals of Geophysics

2010

Abstract

This paper investigates the basic properties of the recent shallow seismicity in Italy through stochastic modeling and statistical methods. Assuming that the earthquakes are the realization of a stochastic point process, we model the occurrence rate density in space, time and magnitude by means of an Epidemic Type Aftershock Sequence (ETAS) model. By applying the maximum likelihood procedure, we estimates the parameters of the model that best fit the Italian instrumental catalog, recorded by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) from April 16th 2005 to June 1st 2009. Then we apply the estimated model on a second independent dataset (June 1st 2009- Sep 1st 2009). We find that the model performs well on this second database, by using proper statistical tests. The model proposed in the present study is suitable for computing earthquake occurrence probability in real time and to take part in international initiatives such as the Collaboratory Study for Earthquake Predictability (CSEP). Specifically we have submitted this model for the daily forecasting of Italian seismicity above Ml4.0.

1. Introduction

There is a growing consensus to accept the existence of an intrinsic stochasticity of the earthquake generating process (see Vere-Jones, 2006, for a review on the use of stochastic models for earthquake occurrence); this view has promoted the formulation of different stochastic models acting on different spatio-temporal scales (Kagan & Knopoff, 1981; Kagan & Jackson, 2000; Ogata, 1988; 1998; Helmstetter *et al*., 2006; Faenza *et al*., 2003; Rhoades & Evison, 2004; Gerstenberger *et al.*, 2005; Marzocchi & Lombardi, 2008; Lombardi *et al*., 2006; 2007; 2010). Each model describes one or more different coexisting physical processes (tectonic loading, coseismic stress interactions, postseismic deformation, aseismic processes, and so on), which have more or less relevance for earthquake occurrence, depending on maturity in the seismic cycle. Here, we focus our attention on daily forecasts. For this class of forecasts, stochastic models describing the phenomenon of earthquake clustering are becoming widely accepted in the seismological community (e.g., Reasenberg & Jones, 1989, 1994; Gerstenberger *et al.*, 2005; Marzocchi & Lombardi, 2009).

Specifically we describe a short-term earthquake forecasting model that we have submitted to the EU-Italy Collaboratory Studies for Earthquake Predictability (CSEP) experiment. The forecast method uses earthquake data only, with no explicit use of tectonic, geologic, or geodetic information. The method is based on the observed regularity of earthquake occurrence rather than on any physical model. The basis underlying this earthquake forecasting method is the popular concept of an epidemic process: every earthquake is a potential triggering event for subsequent earthquakes (Ogata 1988, 1998; Console *et al.* 2003; Helmstetter *et al.,* 2006; Lombardi & Marzocchi, 2007). We apply a version of the ETAS model to seismicity recorded in Italy in recent years. For a first retrospective test, we apply a well-know procedure that consists in fitting the model to the early part of the Italian earthquake catalog and then testing it on the most recent part of the data set. The real time forecasting performance of the model has been successfully checked on the occasion of the recent L'Aquila earthquake (Central Italy; April 6^{th} 2009, Mw 6.3; see Marzocchi & Lombardi, 2009).

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

The ETAS Model (Kagan & Knopoff, 1981; Kagan, 1991, Ogata, 1988, 1998) is a stochastic point process of particular relevance for modeling coseismic stress-triggered aftershock sequences. Its formulation followed from the observation that aftershock activity is not always predicted by a single modified Omori function (Omori, 1894; Utsu, 1961) and that seismicity can include conspicuous secondary aftershock production. Therefore this model assumes that each aftershock has some magnitude-dependent ability to perturb the rate of earthquake production and therefore to generate its own Omori-like aftershock decay. Since the first time-magnitude formulation proposed by Ogata (1988), many others time-magnitude-space versions have been published in the literature, mostly based on empirical studies of past seismicity (Ogata, 1998; Zhuang *et al*., 2002; Console *et al*., 2003; Helmstetter *et al*., 2006; Lombardi & Marzocchi, 2007]. These approaches describe the seismicity rate of a specific area as the sum of two contributions: the "*background rate*" and the "*rate of triggered events*". The first refers to seismicity not triggered by previous events in the catalog; the second is associated with stress perturbations caused by previous earthquakes of the catalog.

The ETAS model defines the total space-time conditional intensity $\lambda(t,x,y,m/\mathcal{H}_t)$ (i.e. the probability of an earthquake occurring in the infinitesimal space-time volume conditioned to all past history) by equation:

$$\lambda(t,x,y,m/\mathcal{H}_t) = \left[\nu u(x,y) + \sum_{t_i<t} \frac{K}{(t-t_i+c)^p} e^{\alpha(M_i-M_c)} \frac{c_{d,q}}{[r_i^2+d^2]^q} \right] \beta e^{-\beta(m-M_c)} \qquad (1)$$

where $\mathcal{H}_t = \{(t_i,x_i,y_i,M_i);\ t_i < t\}$ is the observation history up the time t; M_c is the completeness magnitude of the catalog; ν is the rate of background seismicity for the whole area; K, c and p are the parameters of the modified Omori Law describing the decay in time of short-term triggering effects; α determines how the triggering capability depends on the magnitude of an earthquake; the parameters d and q characterize the spatial probability density function (PDF) of triggered events and $c_{d,q} = \frac{q-1}{\pi}[d^{2(q-1)}]$ is the relative normalization constant; r_i is the distance between location (x,y) and the epicenter of i-th event (x_i,y_i); the function $u(x,y)$ is the spatial PDF of background events; finally, $\beta=b\cdot ln(10)$ is the parameter of the well-known Gutenberg-Richer Law (Gutenberg & Richter, 1954), that is assumed to hold for all magnitudes and invariant in space. Specifically, the model assumes that large earthquakes are indistinguishable from the smaller ones, and therefore they have the same distribution.

The most recent versions of the ETAS model (Ogata & Zhuang, 2006; Helmstetter *et al.*, 2006) are characterized by the introduction of a further term that takes into account the correlation between the aftershock area and the magnitude of triggered events. Some preliminary results show that this correlation may be negligible for Italy (see Marzocchi & Lombardi, 2009). So we decide to use the version of ETAS model described by eq. (1), in which the spatial decay of triggered activity is independent of the magnitude of the triggering shock. A deeper analysis on this topic will be presented and discussed in future works.

The parameters $(\nu,K,c,p,\alpha,d,q,\beta)$ of the model, for the events within a time interval $[T_{start},T_{end}]$ and a region R can be estimated by maximizing the Log-Likelihood function (Daley & Vere-Jones, 2003), given by

$$\log L(\nu,K,c,p,\alpha,d,q,\beta) = \sum_{i=1}^{N} \log \lambda(t_i,x_i,y_i,m_i/H_{t_i}) - \int_{T_{start}}^{T_{end}} \int_{R} \int_{M_c}^{M_{max}} \lambda(t,x,y,m/H_t)\,dt\,dx\,dy\,dm \qquad (2)$$

where M_{max} is the expected maximum magnitude for the region R. The parameters of the model are estimated by means of the iteration algorithm developed by Zhuang *et al.* (2002). By using a suitable kernel, this procedure provides, in addition to the model parameters, an estimation of the PDF $u(x,y)$ for background events. The background rate is given by

$$\upsilon u(x,y) = \frac{1}{T}\sum_{j} p_j K_{d_j}(r_j) \qquad (3)$$

where T is the length of time recovered by the dataset, p_j is the probability that the j-th event is not triggered by previous shocks in the catalog and K_{d_j} is a Gaussian kernel function with a spatially variable bandwidth. Similarly the rate of triggered events is given by

$$c(x,y) = \frac{1}{T}\sum_{j}(1-p_j) K_{d_j}(r_j) \qquad (4)$$

Several physical investigations show that static stress changes decrease with epicentral distance as r^{-3} (Hill *et al.*, 1993; Antonioli *et al.*, 2004), therefore in the present study we impose $q=1.5$. This choice is also justified by the trade-off between parameters q and d that may cause different pairs of q and d values to provide almost the same likelihood of the model (Kagan & Jackson, 2000).

3. Testing the Model

The gold standard for evaluating scientifically earthquake forecasting models is through the comparison of forecasts and true value in prospective experiments (see, e.g., Field, 2007; Schorlemmer *et al.*, 2007; Luen & Stark, 2008; Zechar *et al.*, 2009). Nevertheless, it may be conceivable to evaluate the model also through retrospective experiments, for instance, dividing the available dataset in two parts: a first part of dataset, hereinafter *learning* dataset, can be used to set up the model and a second, the *testing* dataset, to check its reliability (Kagan & Jackson, 2000). The verification of forecasting capability of the model can be achieved by a comparison of observations and forecasts. Such a testing enables us to verify if the model is significantly good performing, and, eventually, to identify the features allowing a better forecasting. In successive subsections we describe the statistical tests used in the present study to check our model retrospectively.

3.1 Residuals Analysis

A common diagnostic technique for stochastic point processes is based on transformation of the time axis t into a new scale τ by the increasing function

$$\tau = \Lambda(t) = \int_{T_{start}}^{t} dt' \int_{R} dxdy \int_{M_c}^{M_{max}} dm\ \lambda(t',x,y,m/\mathcal{H}_{t'}) = \int_{T_{start}}^{t} \left[\nu + \sum_{t_i<t'} \frac{k e^{a(M_i-M_c)}}{\left(t'-t_i+c\right)^p} \right] dt' \qquad (5)$$

where T_{start} is the starting time of the observation history H_t (Ogata, 1988). The random variable τ represents the expected number of occurrences in time period $[T_{start}, t]$, into whole region R and with magnitude above M_c. If a model with conditional intensity $\lambda(t,x,y,m/\mathcal{H}_t)$ describes well the temporal evolution of the process, the transformed data $\tau_i = \Lambda(t_i)$, known in statistical seismology with the name of *residuals*, are expected to behave like a stationary Poisson process with the unit rate (Ogata, 1988). Therefore the values $\Delta\tau_i = \tau_{i+1}-\tau_i$ are independent and exponentially distributed (with mean equal to 1) random variables. We check this hypothesis for residual of our ETAS model by means of two nonparametric tests: the Runs test, to verify the reliability of the independence property, and the one-sample Kolmogorov-Smirnov (KS1) test, to check the standard exponential distribution (Gibbons & Chakraborti, 2003; Lombardi & Marzocchi, 2007). We use both tests because the KS1 test is ineffective to check the presence of a memory in the time series. Hence, any discrepancy of residuals by Poisson hypothesis, identified by just one or both tests, is a sign of inadequacy of ETAS model to explain all basic features of analyzed seismicity. This check analysis is similar to the *N*-test, currently used by RELM/CSEP testing centers (Kagan & Jackson, 1995; Schorlemmer *et al.*, 2007), but it avoids the time binning that may lead to biases in the results of the testing phase (see, e.g., Lombardi & Marzocchi, 2010).

3.2 Cumulative Reliability Diagram

The reliability diagram is a common diagnostic technique used to measure the consistency of a forecast model with the observations. Roughly speaking, a probability forecast is reliable if the event actually happens with an observed frequency that is consistent with the forecast. More specifically, a reliability diagram consists of a plot of observed relative frequencies against predicted probabilities (Wilks, 2005). Reliability measures sort the forecast/observations pairs (F_j $/O_i$) into groups, according to the value of forecast variable, and characterize the conditional distributions of the observations given the forecasts. In particular a way to identify visually departures from reliability is to plot the cumulative conditional observed frequency $p(O_i|F_j)$ against the cumulative predicted probability F_j; this gives a Cumulative Reliability Diagram (CRD). The perfect reliability is represented by the diagonal line.

We use this type of analysis to check the predicted spatial distribution on observed seismicity. Specifically we apply a case of dichotomous events, i.e. observations are limited to 2 possible outcomes, the occurrence (O_1) or nonoccurrence (O_2) of an earthquake. To define the forecasting cumulative probabilities F_j, the area under analysis is partitioned in a non-overlapping and exhaustive set of cells C_i; for each cell we compute the proportion of events f_i expected by the forecasting model. These values f_i, by definition between 0 and 1, are sorted in ascending order and are grouped into N bins B_j $(j=1...N)$, that form a partition of the unit interval composed by overlapping increasing subintervals. These bins are characterized by a set of forecasting probabilities F_j that define the probability to have at least one event in B_j

$$I_j = \left\{ i; f_i \in B_j \right\} \qquad \sum_{i \in I_j} f_i \leq F_j \tag{6}$$

The most intuitive choice is to take F_j equally spaced. If the distribution of the forecasts is non-uniform, then choosing the bins so that the sets I_j are equally populated (i.e with the same number of events f_i) can be a good alternative. The values F_j are compared with the cumulative observed frequencies

$$P(O_1 \mid F_j) = \sum_{i \in I_j} \frac{N_i}{N} \tag{7}$$

where N_i is the observed number of shocks into the cell C_i and N is the total number of events. In the case of perfect reliability the conditional probability $p(O_1|F_j)$ is equal to F_j.

4. The INGV Database

Italy is characterized by a generally high seismicity, with observed magnitudes up to about 7.5. The long tradition of seismological studies in Italy produced many efforts for seismic data collection, therefore today Italy can boast of careful seismic instrumental catalogs (Castello *et al.*, 2005; Schorlemmer *et al.* 2010; http://iside.rm.ingv.it/), besides of a tested experience in compiling historical databases (Boschi *et al.*, 2000). The most complete instrumental catalog of italian seismicity is the seismic bulletin of Istituto Nazionale di Geofisica e Vulcanologia (INGV) (http://iside.rm.ingv.it). The Italian seismic network changed significantly in the last years. Specifically the 16 April 2005 marks the date of remarkable changes of the seismic Italian network (Bono & Badiali, 2005; see also Schorlemmer *et al.*, 2010) and of data processing. Given the large difference of INGV bulletin before and after this date, we decide to set up our model on parameters of events collected from April 16th 2005 to June 1st 2009. The earthquakes from June 1st 2009 and Sep 1st 2009 are instead used for a first retrospective test of the model (testing dataset). In agreement with CSEP requirements, we select events above 30 kms of depth occurred in the collection area, as defined by CSEP experiment.

A correct understanding of the physical processes controlling the rate of earthquake production depends on the quality of the available seismic catalog. Specifically, a critical issue that has to be addressed before performing any investigation is the assessment of completeness of dataset. Here we verify the completeness magnitude (M_c) (lowest magnitude at which a negligible number of the events are not detected) and its variations with time. The algorithms are freely available together with the software package ZMAP (Wiemer, 2001). The analysis of whole catalog by Maximum Likelihood method (Shi & Bolt, 1982) provides a value of M_c (local magnitude) equal to 2.0 (see Figure 1a). The analysis of the spatio-temporal variation of completeness magnitude shows clear changes of M_c with time (see Figure 1b) and space (see Figure 1c). We perform these analyses by using a minimum number of events equal to 100 and a radius equal to 50 km. . In particular, M_c reaches about 2.5 soon after the occurrence of recent L'Aquila earthquake (April 6th 2009, Mw6.3; see Figure 2b). This value seems to be a reliable completeness threshold for most part of national territory (see Figure 2c). These results are also in agreement with Schorlemmer *et al.* (2010) which identify $M_c=2.5$ as a reasonable magnitude threshold for most of Italian territory. The only exception is for the southern part of Apulia and the western part of Sicily, showing a higher completeness magnitude (see also Schorlemmer *et al.*, 2010 for details). Considering the small size of these areas, we decide to select for the present study the events above magnitude 2.5 recorded into the INGV bulletin (2100 events for learning and 179 for testing databases). Figure 2 shows the distribution of selected seismic activity for both learning (Figure 2a) and testing (Figure 2b) databases, together with the boundaries of collection area defined by CSEP laboratory.

5. Application and testing of the ETAS model on Italian seismicity

We apply the ETAS model to Italian seismicity recorded into learning database, described in previous section. Following the procedure proposed by Zhuang *et al.* (2002) we estimate the model parameters together with the spatial distribution of background seismicity ($u(x,y)$). Table 1 lists the inferred values of model parameters together with their standard errors and the associated log-likelihood values. The total percentages of triggered and spontaneous events identified by the model are 46% and 54% respectively. In Figure 3 we show two maps: the first represents the distribution of the time-independent background rate ($\nu u(x,y)$, see eq. (3)), the second the distribution of the clustering ratio $r(x,y)$, i.e the ratio between triggered and total rates, for the whole learning period. The clustering ratio is obtained by the formula

$$r(x,y) = \frac{c(x,y)}{\int_{T_1}^{T_2}\int_{M_c}^{M_{max}} \lambda(t,x,y,m/\mathcal{H}_t)\,dtdm} \qquad (8)$$

where $c(x,y)$ and $\lambda(t,x,y,m/\mathcal{H}_t)$ are defined by eq. (1) and (4), respectively. By comparing the two maps shown in Figure 3, we find that the spatial distribution of triggering capability is not a proxy for the seismogenetic potential. For example, the southern part of peninsular Italy shows a lower triggering rate respect to other zones (see Figure 3b), whenever this area is one of most active of whole region (see Figures 2 and 3a). The estimated Omori Law decay predicts that the probability of triggering one or more events with magnitude above 2.5 for an earthquake of magnitude 3.0 is below 1% after about 5-6 hours. The corresponding times for a triggering event of magnitude 5.0 and 7.0 are 2-3 days and about 1 month, respectively (see Figure 4a). We stress that these probabilities refer to direct triggering effects. . The secondary triggered events are not included in this calculation. As regards the spatial decay of the triggering capability, an event has a 50% of probability to trigger one or more events within 2km from its epicenter and about 40% at a distance larger than 10km, regardless its magnitude (see Figure 4b).

A preliminary check on the goodness of the inferred ETAS model is done by applying the residual analysis on the learning dataset used to set-up the ETAS model. We find that the residuals pass the KS1 test (p-value 0.8), but the Runs test rejects the hypothesis of no-correlation (p-value 0.007). The cumulative distribution of residuals (Figure 5a) shows a clear deviation from the expected Poisson behavior soon after the occurrence of M_w 6.3 L'Aquila earthquake (April 6 2009). If we take out the l'Aquila sequence by the learning period, the ETAS model passes the Runs test (p-value 0.07). We argue that this result is probably due to the spatial variation of some parameters. In other words, at local scale the model could be significantly different with respect to the same model calibrated using the whole Italian territory. For example, Marzocchi & Lombardi (2009) reported an α-value of 1.5 for the L'Aquila region that increases to 2.0 when M_c=2.5 is considered; this value is certainly larger than the 1.3 found here for the whole Italian territory (see table 1).

In order to test the forecasting performance of the ETAS model, we analyze the residuals and plot the cumulative reliability diagram on testing dataset. By using the KS1 test we cannot reject the null hypothesis that values $\Delta\tau_i=\tau_{i+1}-\tau_i$ are exponentially distributed (with mean equal to 1) (the p-value is equal to 0.14). The Figure 5b show the cumulative number of residuals τ_i versus transformed time τ (solid line) together with the expected linear scaling predicted by a Poisson distribution (that is, the cumulative number of residuals should lie along the bisector). Similarly, the Runs test does not reject the independence hypothesis of $\Delta\tau_i$ (the p-value is equal to 0.81), implying that the hypothesis of uncorrelation of residuals cannot be rejected. This result is corroborated by Figure 5c, in which we plot the variables $U_{k+1}=1-exp(\Delta\tau_{k+1})$ versus U_k for the testing dataset. If $\Delta\tau_k$ are i.i.d exponential random variables with unit mean, the statistics U_k are i.i.d. uniform random variables on [0,1). Assuming that a possible correlation is likely to show up in neighboring intervals, the plot of U_{k+1} versus U_k should recover uniformly the figure panel (Ogata, 1988).

The cumulative reliability diagram of spatial distribution on events collected by testing dataset shows a reliable forecasting (see Figure 6). To define the forecasting probabilities F_j we compute the expected fraction of events f_i by ETAS model, for each cell C_i of the testing grid defined by CSEP laboratory. The values f_i are computed as the ratio between the expected numbers of events in the cell C_i and in whole region R. Specifically we use the formula

$$f_i = \frac{\int_T \int_{C_i} \int_M \lambda(t,x,y,m/\mathcal{H}_t)dtdxdydm}{\int_T \int_R \int_M \lambda(t,x,y,m/\mathcal{H}_t)dtdxdydm} \qquad (8)$$

where T is the testing period, R is the testing area defined by CSEP laboratory, M is the magnitude range [2.5, 9.0], and $\mathcal{H}_t$ is the occurrence history, starting by April 16 2005 (i.e. including the learning period). Then we regroup these values in 10 bins B_j, identified by increasing values of probabilities F_j. The error bars are defined so that the sets I_j (see eq. 6) are equally populated. In Table 2 we report the values of probabilities F_j and $p(O_1|F_j)$ (i.e. the observed frequencies of events in bin B_j), as defined in eq. (7). They are plotted in Figure 6. The error bars indicate the 95% confidence interval of values $p(O_1|F_j)$. These last are obtained by applying the reliability analysis on 1000 synthetic catalogs. These have the same duration of testing period of INGV bulletin and are simulated in agreement with ETAS model, including the real learning period into the past history. The reliability diagram shows that the pairs $[F_j, p(O_1|F_j)]$ are well fitted by diagonal that indicates a perfect reliability. Moreover they are in agreement with variation expected by the model. All these results show that the model estimated on learning dataset is in agreement with the following seismicity. This result is also corroborated by the observation that the parameters estimated from the entire catalog are not statistically different by parameters listed in Table 1.

The model formulated and tested above allows us to compute forecasts in the framework of CSEP experiment. Predictions are in a form of daily probability of occurrence for at least one earthquake with $M_l \geq 4.0$, within a cell of 0.1°x0.1°, in Italy. These are obtained by integrating for each cell C_i and for each forecasting period T_j the intensity function of ETAS model (eq. (1)). The forecast rates above M_l 4.0 are obtained by rescaling the rate of earthquakes above M_l 2.5, in agreement with the Gutenberg-Richter relation. The eq. (1) shows that a time-dependent modeling as the ETAS model imposes to take into account also the triggering effect of seismicity occurred before and expected during the forecast interval. So we include in the past history all real seismicity with magnitude above M_l 2.5 and depth above 30 km, occurred up to the starting time of the forecasting time window. Moreover we simulate 1000 different stochastic realizations for the forecasting time window, by using the thinning method proposed by Ogata (1998) and the intensity function formulated in equation (1). Then we average predictions coming from each of these synthetic catalogs.

6. Discussion and Conclusions

In this paper we have adopted a version of ETAS model to describe the recent shallow seismicity occurred in Italy. The main motivation of this study was to submit our model to EU-Italy CSEP laboratory for 1-day forecasts. To achieve this goal we have proposed a model representing the main average properties of Italian seismicity. The reliability of this model has been successfully checked, at local scale, in a real-time forecasting experiment, on occasion of the occurrence of recent L'Aquila destructive earthquake (Marzocchi & Lombardi, 2009).

One finding of the present paper is that the generalization of local models to the whole Italian territory may be problematic for different reasons. First, the completeness magnitude varies with space (Schorlemmer *et al.*, 2010); in this paper we have adopted M_c=2.5 that is probably optimistic for some zones. In fact, the M_c for the whole territory is about 2.9 (see Figure 1c and Schorlemmer *et al.*, 2010). We are conscious of this limit, but we preferred to adopt a value of M_c that is reliable for most (not all) of Italian territory. The area with M_c>2.5 covers only a very small part of the whole region. The use of a larger completeness magnitude causes a strong reduction of dataset with a consequent increase of uncertainty of the model. Maybe more important, it has been recognized that smaller earthquakes have a decisive role in the triggering process (Helmstetter, 2003; Felzer *et* *al.*, 2002; Helmstetter *et al.*, 2004); therefore, a too high value of M_c might cause an erroneous identification of the triggered part of seismicity.

Second, some of the ETAS parameters may vary with space. This means that some parameters estimated for the whole territory and for a small region may be significantly different. Local variations may occur only as consequence of the occurrence of large earthquakes. For example, the model proposed here for the whole Italian territory is not able to reproduce correctly the time evolution of the first part of 2009 L'Aquila sequence (see Figure 3a). As anticipated before, we argue that this discrepancy is probably due to features of the local seismicity that cannot be extrapolated for the whole territory. In particular the seismicity of L'Aquila is characterized by a larger α-value with respect to the whole Italian seismicity described by our ETAS model. The α parameter is crucial to quantify the dependence of triggering effect by magnitude of parent earthquake. The failure of the model to describe the starting phase of L'Aquila sequence suggests that possible inconsistencies could occur in forecasting future seismicity. This problem may call for the development of more complicated models that take into account local features of seismic activity.

We argue that other parts of the model could be improved in the future. In the following, we report only some possible hints in this direction. First, the model could be enhanced by adopting a modified magnitude distribution, to explicitly allow for the decrease of detection soon after a large earthquake (Kagan 1991, Helmstetter et al., 2006; Lennartz *et al.*, 1998). Second, the background rate and the basic clustering proprieties of aftershocks sequences are assumed to be stationary in time. Such an assumption is mostly motivated by the short learning dataset adopted. Longer datasets may permit to capture departures from stationarity such as long-term time evolution of the seismicity (e.g., Lombardi & Marzocchi, 2007; Marzocchi & Lombardi, 2008). Moreover, other time-dependent processes acting on short time scales, like fluid injection, may have a significant impact on short-term spatio-temporal evolution of seismicity and therefore it may be necessary to include them into the ETAS model (Ogata & Hainzl 2005; Lombardi et al., 2006; 2010). Third, the ETAS model proposed here assumes that all earthquakes are equal. Possible distinctive precursory activity that anticipates large shocks is not considered in this parametrization. Finally, the present model does not incorporate tectonic/geologic information. Their inclusion may represent one possible future direction of investigation to improve the forecasting of large shocks. For example, the Gutenberg-Richter law is used everywhere indistinctively; this means that a magnitude 8 is considered possible everywhere. It is argued that geological information may provide in the future a more appropriate frequency-magnitude law that varies in space.

References

Antonioli, A., M.E. Belardinelli and M. Cocco (2004): Modeling Dynamic Stress Changes Caused by an Extended Rupture in an Elastic Stratified Half Space, *Geophys. J. Int.*, **157**, 229-244.

Bono, A. and L. Badiali (2005): Pwl personal wavelab 1.0, an object-oriented workbench for seismogram analysis on windows system, *Computers & Geosciences*, **31**, 55-64.

Castello, B., G. Selvaggi, C. Chiarabba and A. Amato (2005): CSI Catalogo della sismicità italiana 1981-2002, vers. 1.0, INGV-CNT, Roma, www.ingv.it/CSI/.

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

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

Field, E. H. (2007): Overview of the working group for the development of regional earthquake likelihood models (relm), *Seismol. Res. Lett.*, **78**, 7-16.

Gerstenberger, M.C., S. Wiemer, L.M. Jones and P.A. Reasenberg (2005): Real-time forecasts of tomorrow's earthquakes in California, *Nature* **435**, 328-331. doi:10.1038/nature03622.

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

Gutenberg, B. and C.F. Richter (1954): Seismicity of the Earth and Associated Phenomena, Princeton, pp. 273.

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

Felzer, K. R., T. W. Becker, R. E. Abercrombie, G. Ekstrom, and J. R. Rice (2002): Triggering of the 1999 M_W 7.1 Hector Mine earthquake by aftershocks of the 1992 M_W 7.3 Landers earthquake, *J. Geophys. Res.*, **107**, 2190, doi:10.1029/2001JB000911.

Helmstetter, A. and D. Sornette (2003): Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.* **108**, 2237, doi:10.1029/2001JB001580.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2004): Importance of small earthquakes for stress transfers and earthquake triggering, *J. Geophys. Res.*, **110**, B05S08, doi:10.1029/2004JB003286.

Helmstetter, A., Y.Y. Kagan and D.D. Jackson (2006): Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bull. Seism. Soc. Am.* **96**, 90-106, doi: 10.1785/0120050067.

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

Kagan, Y.Y. and L. Knopoff (1981): Stochastic synthesis of earthquake catalogs, *J. Geophys. Res.*, **86**, 2853-2862.

Kagan, Y. Y. (1991): Likelihood analysis of earthquake catalogs, *Geophys. J. Int.*, **106**, 135–148.

Kagan, Y.Y. and D.D. Jackson (1995): New seismic gap hypothesis: Five years after, *J. Geophys. Res.*, **100**, 3943-3959.

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

Lennartz S and A. Bunde (2008): Missing data in aftershock sequences: Explaining the deviations from scaling laws, *Phys. Rev. E,* **78**, 041115.

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

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

Lombardi A.M., W. Marzocchi (2010): The assumption of Poisson seismic rate variability in CSEP/RELM experiments. *Bull. Seismol. Soc. Am.*, in press.

Lombardi, A.M., M. Cocco and W. Marzocchi (2010): On the increase of background seismicity rate during the 1997-1998 Umbria-Marche (central Italy) sequence: apparent variation or fluid-driven triggering? *Bull. Seismol. Soc. Am.*, **100**, 1138-1152.

Luen, B. and P.B. Stark (2008): Testing Earthquake Predictions. IMS Lecture Notes Monograph Series. Probability and Statistics: Essays in Honor of David A. Freedman, 302–315. Institute for Mathematical Statistics Press, Beachwood.

Marzocchi, W. and A.M. Lombardi (2008): A double branching model for earthquake occurrence, *J. Geophys. Res.*, **113**, B08317, doi:10.1029/2007JB005472.

Marzocchi, W. and A.M. Lombardi (2009): Real-time forecasting following a damaging earthquake, *Geophys. Res. Lett.*, **36**, L21302, doi:10.1029/2009GL040233.

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

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

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

Omori, F. (1894): On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*, **7**, 111-120.

Reasenberg, P. A. and L. M. Jones (1989): Earthquake hazard after a mainshock in California. *Science* **243**, 1173-1176.

Reasenberg, P. A. and L. M. Jones (1994): Earthquake aftershocks: Update. *Science* **265**, 1251-1252.

Rhoades, D. A. and F. F. Evison (2004): Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pure Appl. Geophys*., **161**, 47-72.

Shi, Y., and B.A. Bolt (1982): The standard error of the Magnitude-frequency b value, *Bull. Seism. Soc. Am*., **72**, 1677-1687.

Schorlemmer D., M.C. Gerstenberger, S. Wiemer, D.D. Jackson and D.A, Rhoades (2007): Earthquake Likelihood Model Testing, *Seism. Res. Lett*., **78**, 17-29.

Schorlemmer, D., F. Mele and W. Marzocchi (2010): A Completeness Analysis of the National Seismic Network of Italy, *J. Geophys. Res*. **115**, B04308, doi:10.1029/2008JB006097

Utsu, T. (1961): A statistical study on the occurrence of aftershocks, *Geophys. Mag*., **30**, 521-605.

Vere-Jones, D. (2006): The development of statististical seismology: my personal experience. *Tectonophysics* **413**, 5-12.

Wilks, D.S. (2005): Statistical Methods in the Atmospheric Sciences, Second Edition, Elsevier, Academic Press, New York, NY., 627 pp.

Wiemer, S. (2001): A software package to analyze seismicity: ZMAP, *Seism. Res. Lett*., **72**, 373-382.

Zechar, J. D., D. Schorlemmer, M. Liukis, J. Yu, F. Euchner, P. J. Maechling, and T. H. Jordan (2009): The Collaboratory for the Study of Earthquake Predictability perspective oncomputational earthquake science, *Concurrency and Computation: Practice and Experience*, doi:10.1002/cpe.1519.

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

Table Captions

Table 1: Maximum Likelihood parameters (with relative errors) and log-likelihood of ETAS model for the learning INGV bulletin (M_c = 2.5; Apr 16 2005 – Jun 1 2009; 2100 events).

Table 2: Cumulative Reliability Diagram of spatial distribution of earthquakes predicted by ETAS model relative to the testing INGV bulletin (M_c = 2.5; Jun 1 2009 – Sep 1 2009; 179 events). The values F_j and $p(O_1|F_j)$ indicate the forecasts and the observed frequencies, respectively.

Figure Captions

Figure 1: Completeness magnitude of INGV bulletin (from April 16th 2005 up to June 1st 2009) obtained by the Maximum Likelihood Method (MLM). a) Frequency magnitude distribution for the whole dataset: the MLM provides $M_c=2.0$; b) M_c as a function of time; c) M_c as a function of space.

Figure 2: Map of seismic events with magnitude above 2.5 and depth smaller than 30 km that occurred in Italy inside the collection area identified by the CSEP experiment (blue solid line; see Schorlemmer et al., 2010b). The symbol sizes are scaled with magnitude. a) Map of events of the learning dataset (April 16th 2005-June 1st 2009; 2100 events) used to set-up the model; b) map of the testing dataset (June 1st 2009- Sep 1st 2009; 179 events) used for a retrospective forecasting test of the model.

Figure 3: Maps of a) the background seismicity rate $\nu u(x,y)$, and b) the ratio between the triggered rate and the total seismic rate of the INGV bulletin learning dataset (April 16th 2005-June 1st 2009; 2100 events).

Figure 4: Spatio-temporal behavior of the triggering probability inferred by the ETAS model. a) Time decay (by the Omori law) of the probability to generate at least one event for different magnitudes. b) Cumulative of the spatial probability distribution of triggering at least one event (see eq. (1)).

Figure 5: Residuals Analysis of the ETAS model on the learning (April 16th 2005-June 1st 2009; 2100 events) and testing INGV bulletin (June 1st 2009- Sep 1st 2009; 179 events). a) Cumulative number of transformed times τ_i (solid line) for the learning period together with the theoretical distribution (dotted line) predicted by a Poisson distribution. b) The same as a), but for the testing period. c) Plot of values $U_{k+1}=1-exp(\tau_{k+1}-\tau_k)$ versus U_k for the testing period.

Figure 6: Cumulative Reliability Diagram of the spatial earthquake distribution predicted by ETAS model for the testing INGV bulletin (M_c = 2.5; Jun 1 2009 – Sep 1 2009; 179 events). Stars mark the pairs F_j / $p(O_1|F_j)$, i.e., the forecasts and the observed spatial distributions. The dotted black line

represents the perfect reliability. Error bars identify the 95% confidence interval of the observed values $p(O_1|F_j)$. The forecast probabilities F_j identify equally populated bins B_j (see text for details).

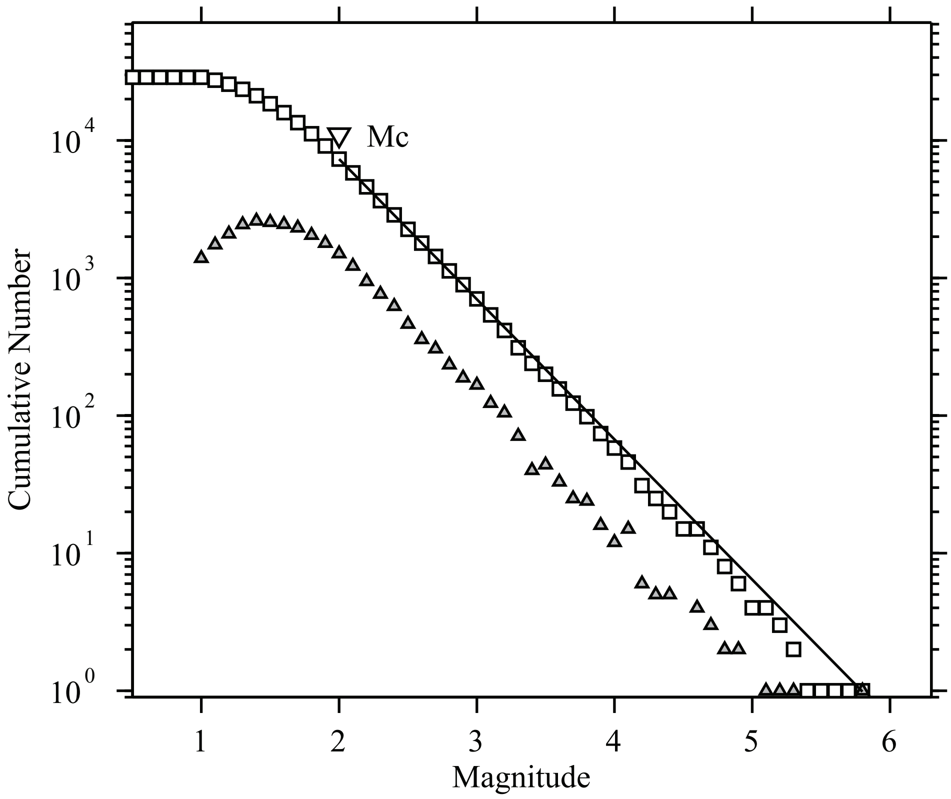
Table1: Parameters of ETAS model for Italian seismicity

(M_c = 2.5; Apr 16 2005 – Jun 1 2009; 2100 events)

Parameter	**Value**
ν	237 ± 8 ($year^{-1}$)
K	0.011 ± 0.001 ($year^{p-1}$)
p	1.16 ± 0.02
c	0.00004 ± 0.00001 (year)
α	1.3 ± 0.1
d	1.10 ± 0.05 (km)
q	≡ 1.5
Log-likelihood	-7808.1

Table2: Values of Cumulative Reliability Diagram

F_j	$p(O_1\|F_j)$
$1.6\cdot10^{-3}$	$1.8\cdot10^{-3}$
$5.6\cdot10^{-3}$	$7.5\cdot10^{-3}$
$1.2\cdot10^{-2}$	$1.4\cdot10^{-2}$
$2.2\cdot10^{-2}$	$1.9\cdot10^{-2}$
$3.5\cdot10^{-2}$	$3.2\cdot10^{-2}$
$5.2\cdot10^{-2}$	$5.6\cdot10^{-2}$
$7.6\cdot10^{-2}$	$8.6\cdot10^{-2}$
$1.1\cdot10^{-1}$	$1.3\cdot10^{-1}$
$1.7\cdot10^{-1}$	$2.2\cdot10^{-1}$
1.0	1.0



Maximum Likelihood Solution
b?value = 1.02 +/- 0.01, a value = 5.9, a value (annual) = 5.28
Magnitude of Completeness = 2

Figure 1a

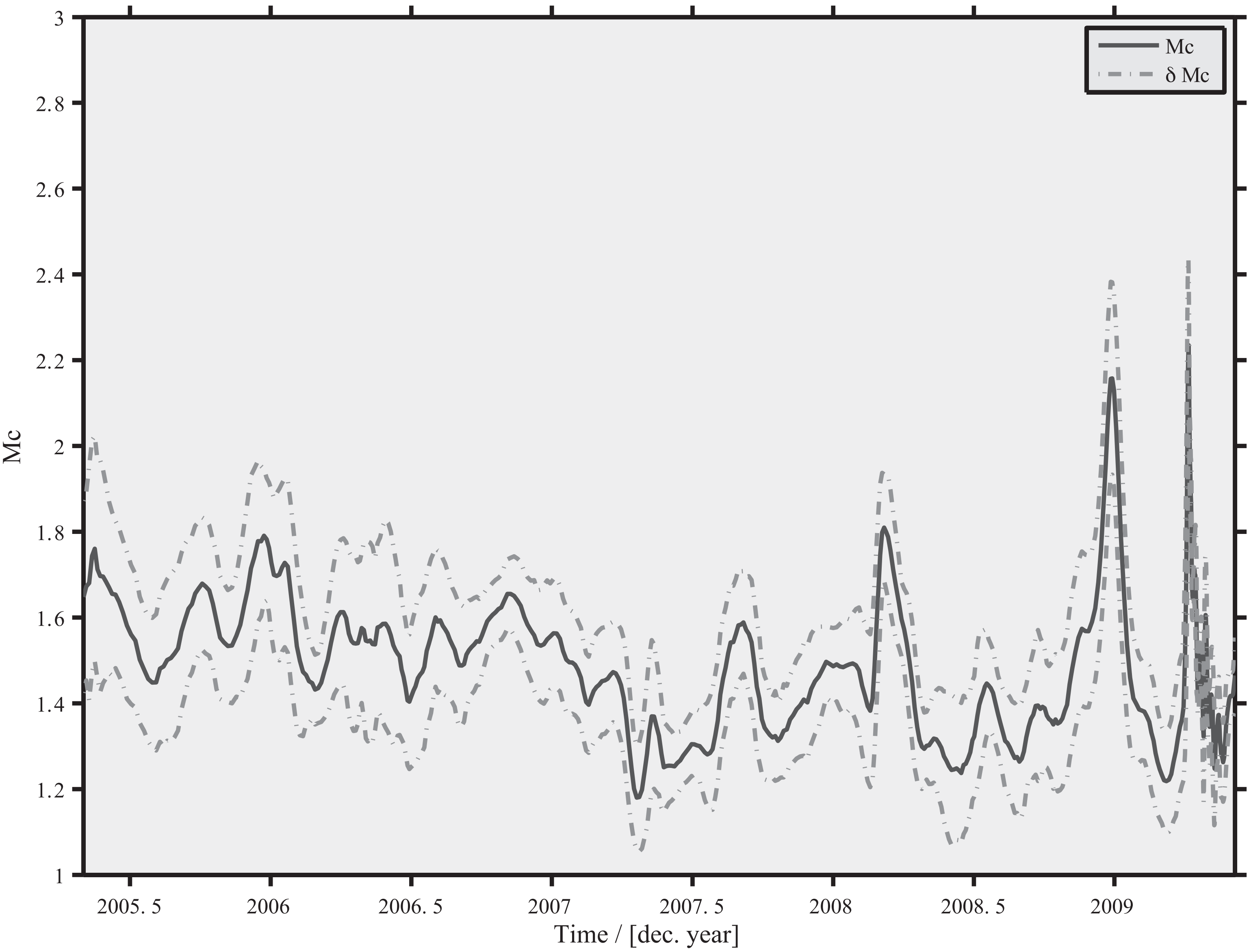

Mc
δ Mc
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
2005. 5
2006
2006. 5
2007
2007. 5
2008
2008. 5
2009
Time / [dec. year]

Figure 1b

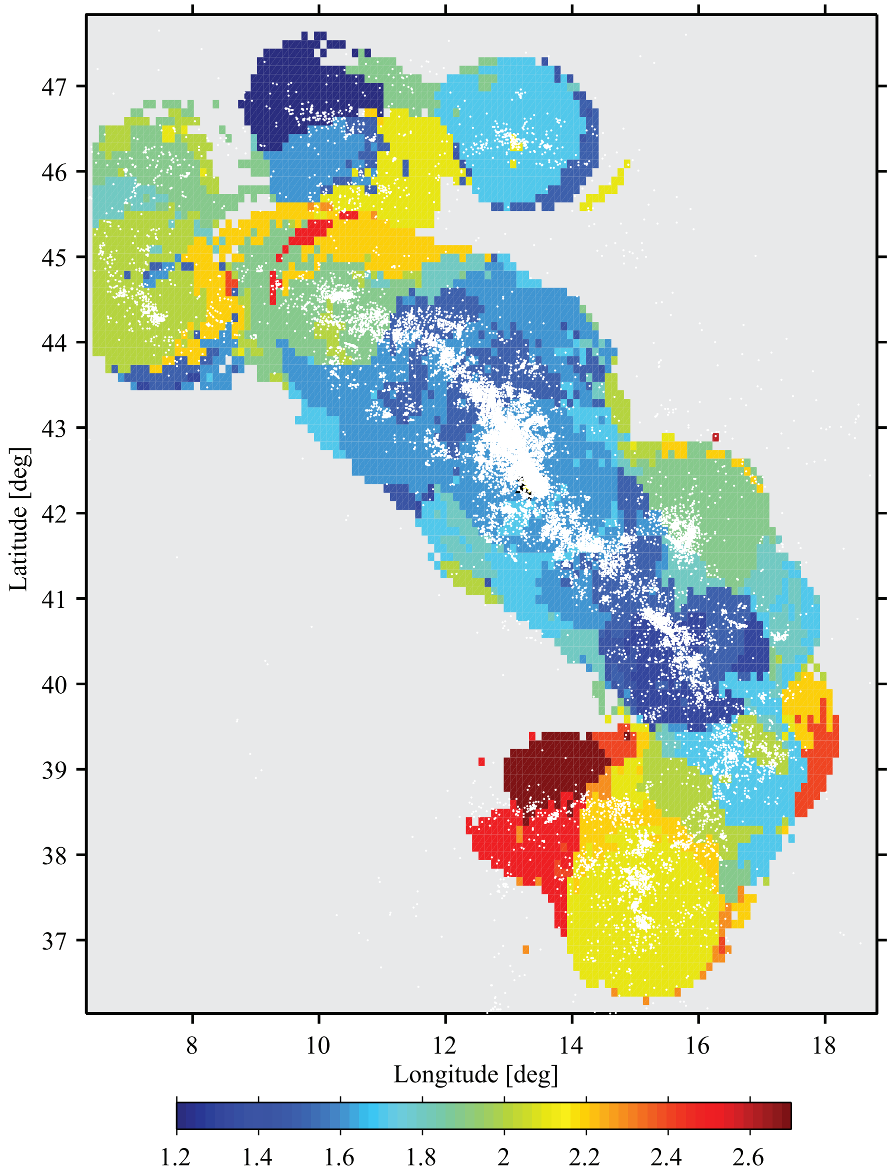

Latitude [deg]
47
46
45
44
43
42
41
40
39
38
37
8
10
12
14
16
18
Longitude [deg]
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6

Figure 1c

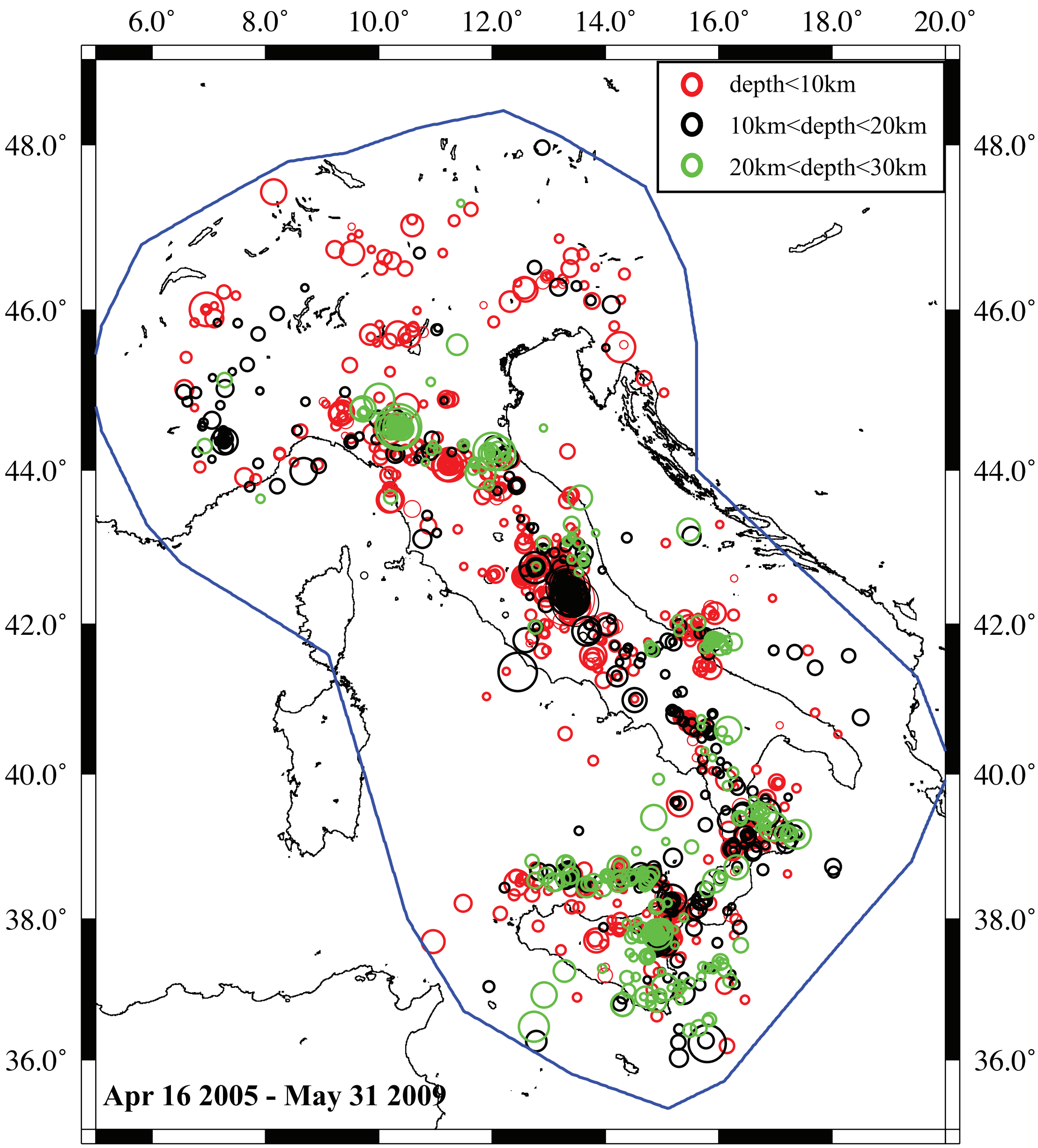

6.0°
8.0°
10.0°
12.0°
14.0°
16.0°
18.0°
20.0°
48.0°
46.0°
44.0°
42.0°
40.0°
38.0°
36.0°
depth<10km
10km<depth<20km
20km<depth<30km
Apr 16 2005 - May 31 2009

Figure 2a

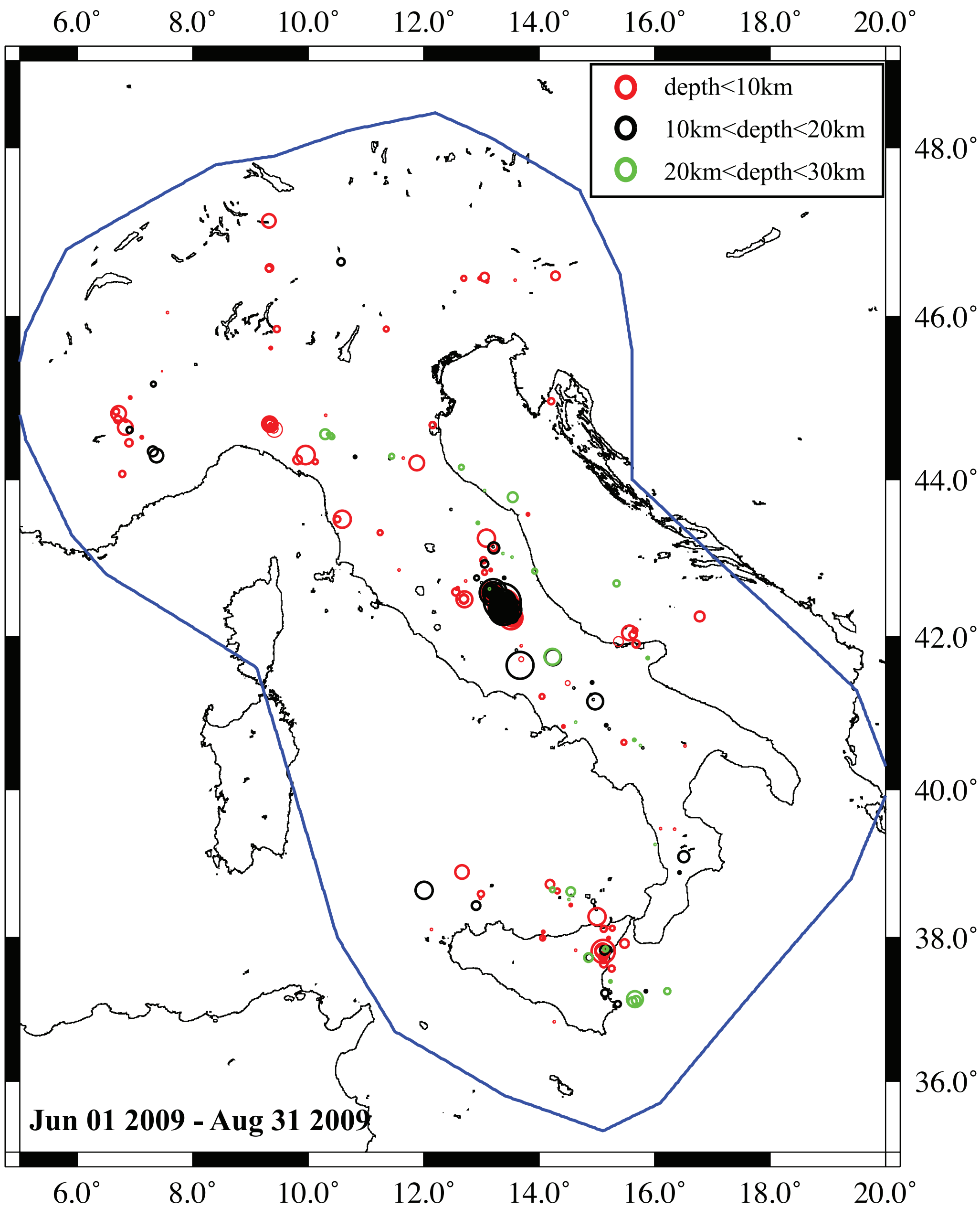
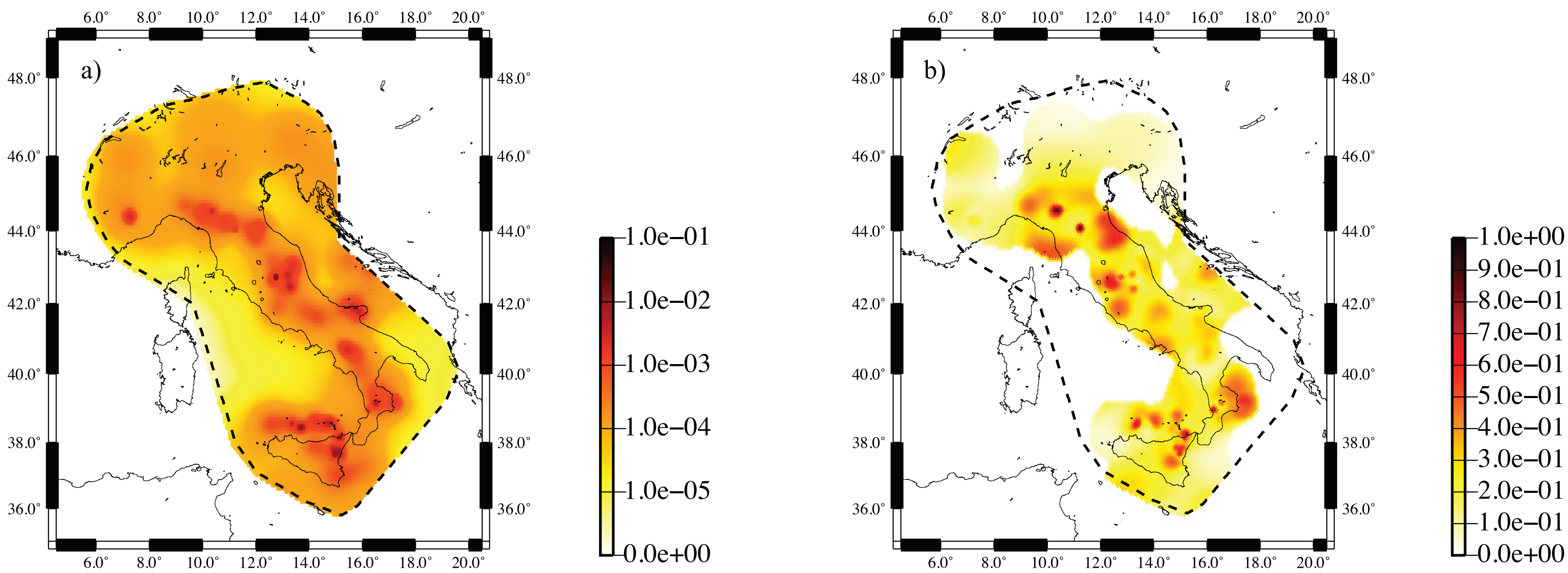

6.0°
8.0°
10.0°
12.0°
14.0°
16.0°
18.0°
20.0°
48.0°
46.0°
44.0°
42.0°
40.0°
38.0°
36.0°
depth<10km
10km<depth<20km
20km<depth<30km
Jun 01 2009 - Aug 31 2009

Figure 2b



a)
b)
6.0°
8.0°
10.0°
12.0°
14.0°
16.0°
18.0°
20.0°
36.0°
38.0°
40.0°
42.0°
44.0°
46.0°
48.0°
1.0e−01
1.0e−02
1.0e−03
1.0e−04
1.0e−05
0.0e+00
1.0e+00
9.0e−01
8.0e−01
7.0e−01
6.0e−01
5.0e−01
4.0e−01
3.0e−01
2.0e−01
1.0e−01
0.0e+00

Figure 3

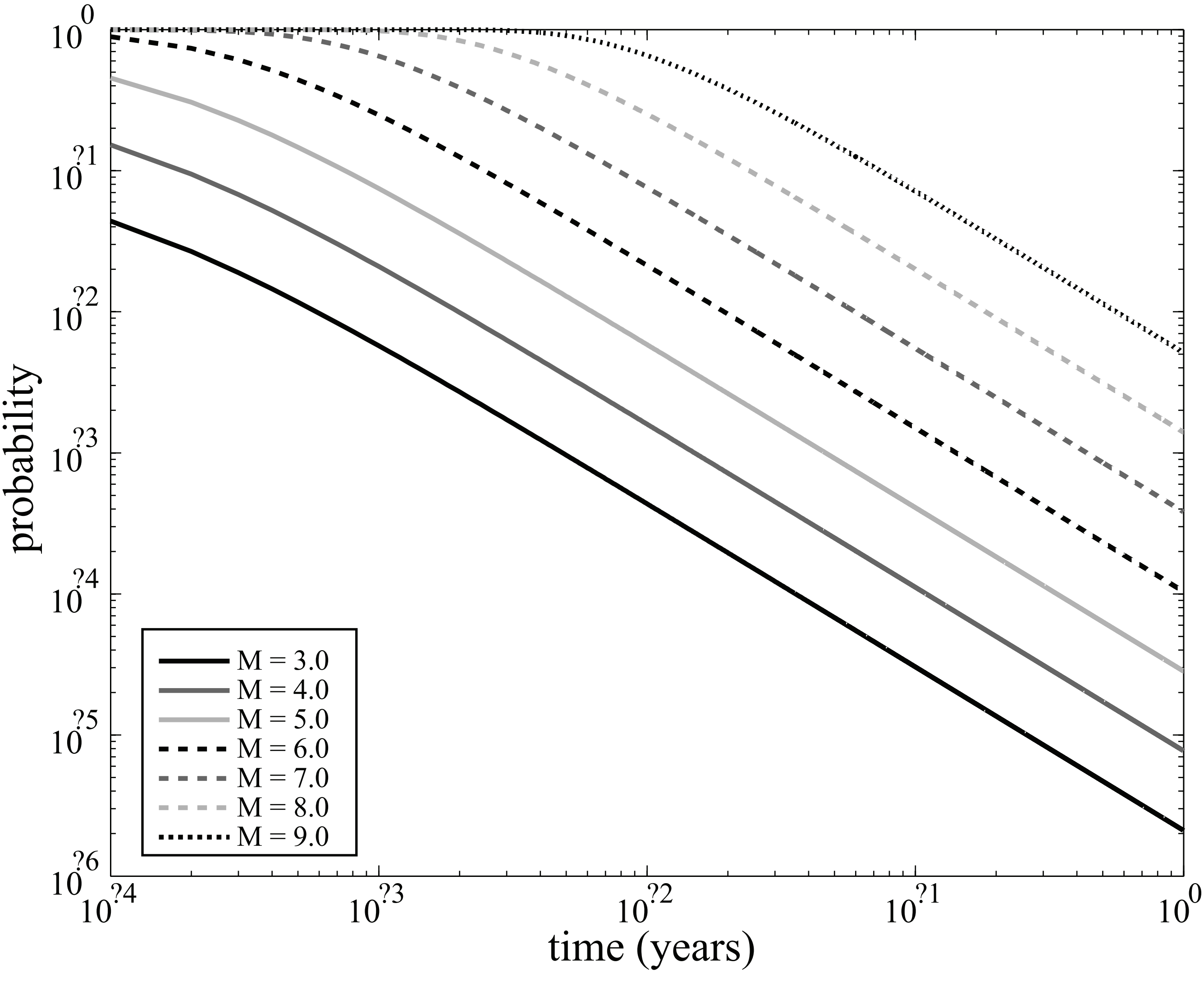


Figure 4a

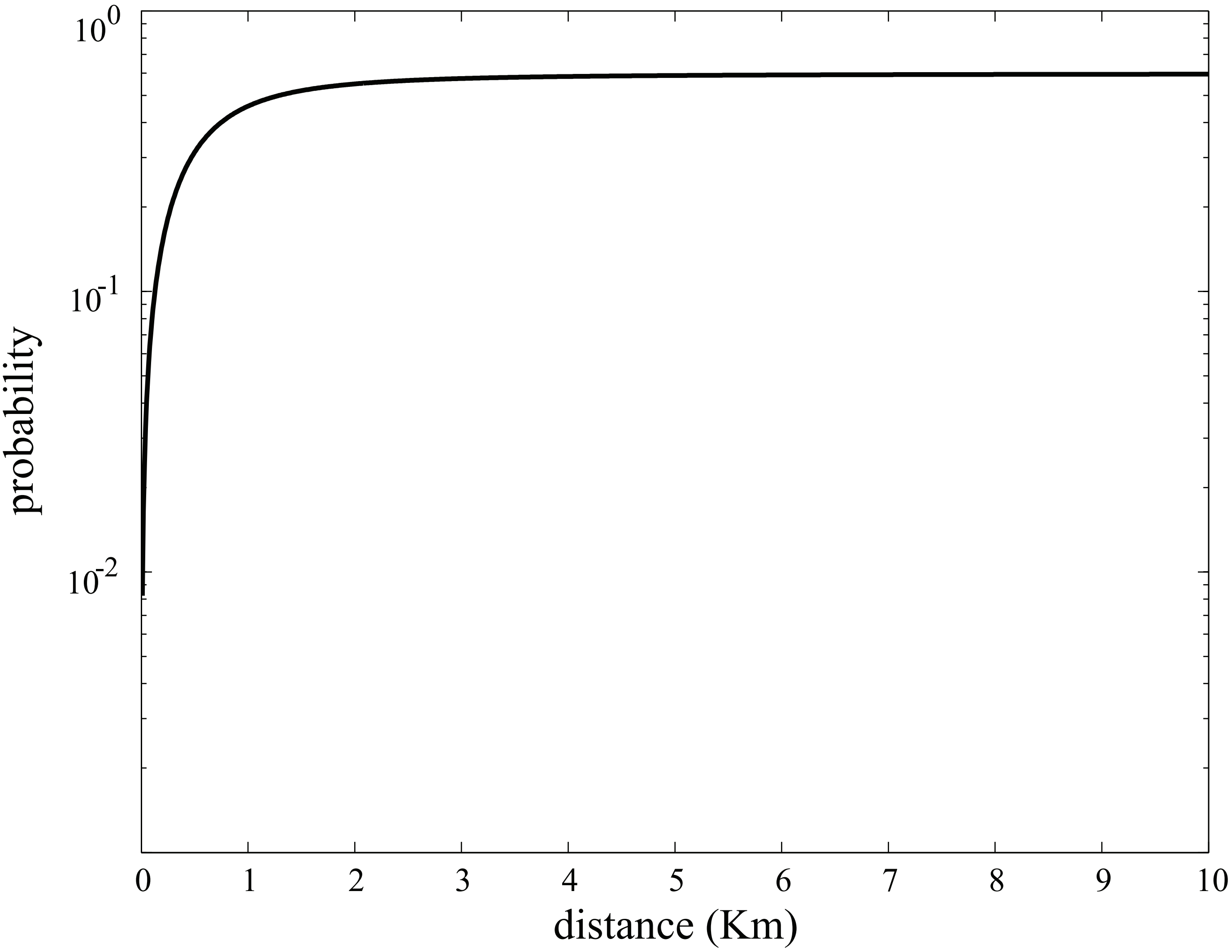


Figure 4b

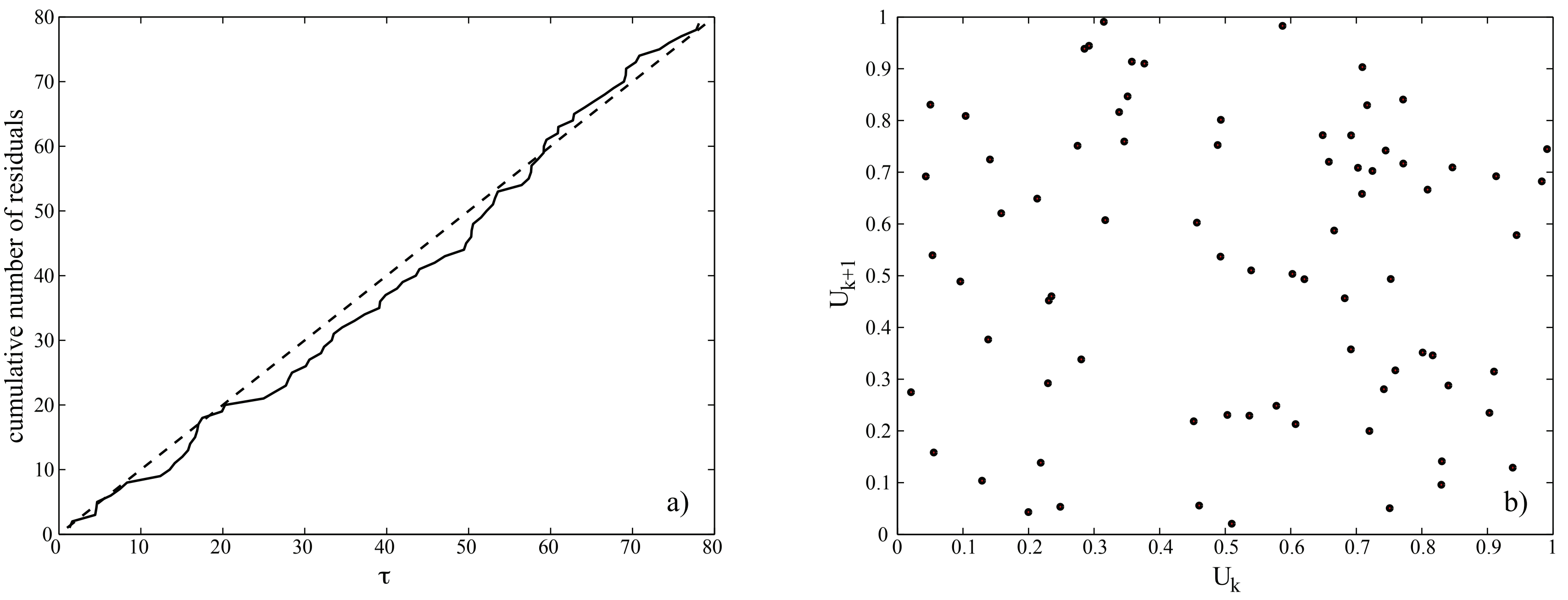
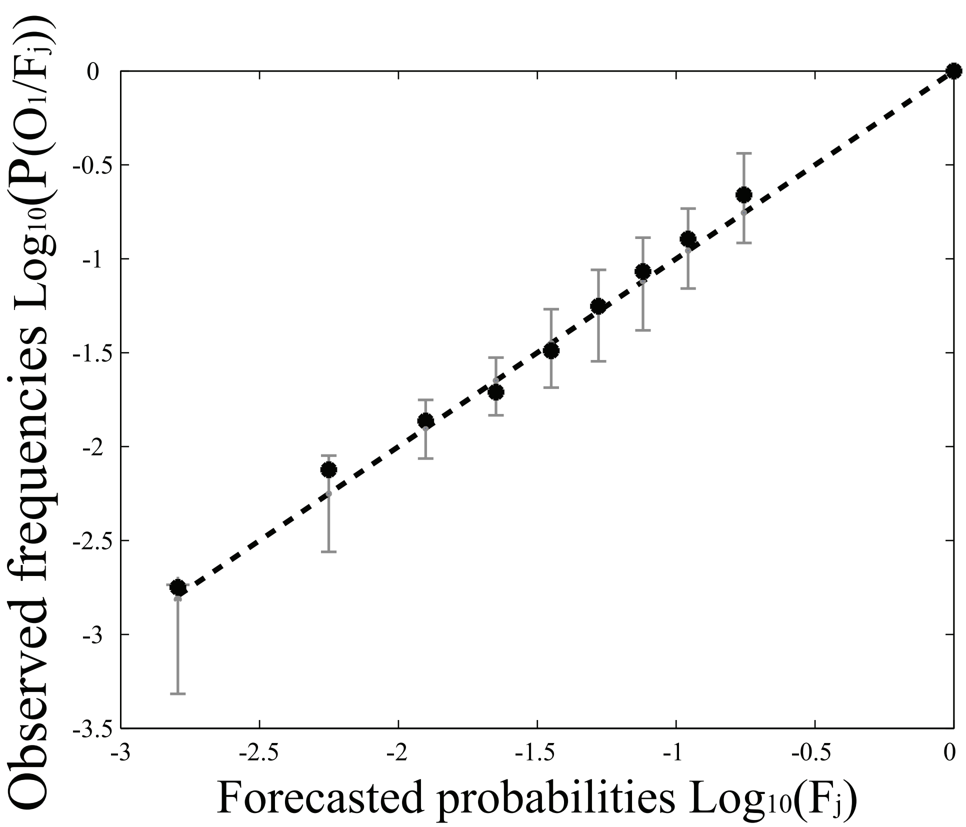


Figure 5



0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
Observed frequencies $Log_{10}(P(O_1/F_j))$
Forecasted probabilities $Log_{10}(F_j)$

Figure 6