THE DOUBLE BRANCHING MODEL FOR EARTHQUAKE FORECAST APPLIED TO THE JAPANESE SEISMICITY

Anna Maria Lombardi

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

(e-mail: annamaria.lombardi@ingv.it)

Warner Marzocchi

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

(e-mail: warner.marzocchi@ingv.it)

Abstract

The purpose of this work is to apply the Double Branching Model (DBM) to forecast moderate-large Japanese seismicity. The proposed model is time-dependent, since it assumes that each earthquake can generate or is correlated to other earthquakes, through physical mech-

anisms acting at different spatio-temporal scales. The model is set up through two sequential steps. In the first step, we estimate the well-established short time clustering. Then, we analyze and characterize the declustered catalog through a second order branching process. The inclusion of the second branching is motivated by the statistically significant departure of the declustered catalog from a time-independent model. From a physical point of view, this new branching accounts for possible long-term earthquake interactions. Some recent applications of this model at global and regional scales (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009; 2010) have shown that earthquake occurrences tend to have two main time features: a short-term clustering up to months-few years and a longer time modulation of decades (up to few centuries). Here we apply the DBM to the instrumental Japanese database, collected by the Japan Meterological Agency (JMA) ($M \geq 5.0$). The purpose of this application is twofold. First, we check the existence of two time branchings previously found in other regions. Second, we provide forecasts to be evaluated by the Japanese CSEP (Collaboratory for the Study of Earthquake Predictability) testing center.

1 Introduction

Earthquake forecasting has a key role in the geophysical investigation. It has direct implications for planning risk mitigation actions, and it yields important contributions for a better understanding of earthquake generation process. Presently, a large variety of models is available; these models are based on different physical and stochastic components and they cover quite different forecasting time windows, from 1 day to years and decades (see, e.g., Kagan and Knopoff, 1981; Kagan and Jackson, 2000; Ogata, 1988; 1998; Helmstetter et al., 2006; Rhoades and Evison, 2004; Gerstenberger et al., 2005; Marzocchi and Lombardi, 2008; 2009; Lombardi and Marzocchi, 2009; Console et al., 2010).

Despite the efforts devoted to build models, their reliability has been only partially checked (mostly by the same modelers), often using past data and different statistical methodologies. Moreover, very few attempts have been made to compare the forecasting capabilities of different models. The use of different and inhomogeneous procedures leads to an inherent difficulty to judge what is the best performing model, or more generally, to evaluate relative forecasting performances. Only recently, one important international effort, the Collaboratory for the Study of Earthquake Predictability (CSEP; www.cseptesting.org), has been set to create a common platform for testing

and comparing forecasting/prediction models. This initiative is a generalization of the experiment carried out in California, named the Regional Earthquake Likelihood Models (RELM, www.relm.org; Schorlemmer et al., 2007). Specifically, CSEP has established different testing regions and testing center for evaluating and comparing forecasting/prediction models on different forecasting time windows (Schorlemmer et al., 2010). Recently, Japan joined the CSEP initiative establishing a testing center and a testing region (Tsuruoka et al, 2008).

The goal of the present paper is twofold. First, we describe the implementation of a recently proposed model, named the Double Branching Model (DBM hereinafter), to forecast earthquakes in the Japan testing region. Second, the comparison of the parameters of the model estimated for Japan and other regions of the world allows us to get some new insights on the nature of the earthquake occurrence process. The DBM takes into account long-term modulation of earthquakes occurrence, beside of the short-term clustering of earthquakes. In other words, compared to the classical ETAS (Epidemic Type-Aftershocks Sequences) model (Ogata, 1998), we relax the assumption of long-term stationary seismic background that has been questioned by many recent studies (Kagan and Jackson, 1991; Rhoades and Evison, 2004; Lombardi and Marzocchi, 2007; Marzocchi and Lombardi, 2008; Marzocchi et al., 2009). These studies emphasizes the existence of significant long-term

time modulation of the earthquake occurrence, probably due to fault interaction and stress perturbations on spatio-temporal scales much larger than the ranges interested by aftershock sequences. Other possible departures from a stationary seismic background on a time scale of few days (Hainzl and Ogata, 2005; Lombardi et al., 2006; 2010) are not taken into account by DBM. Notably, the DBM has shown better earthquake forecasting performances for large earthquakes at both worldwide (Marzocchi and Lombardi, 2008) and regional (Lombardi and Marzocchi, 2009; 2010) scales, with respect to models with a time-independent background rate. The forecast method uses earthquake data only, with no explicit use of tectonic, geologic, or geodetic information. The basis underlying this earthquake forecasting method is the popular concept of epidemic process: every earthquake is regarded as a potential triggering event for subsequent earthquakes (Ogata 1988, 1998; Helmstetter et al., 2006; Lombardi and Marzocchi, 2007) on different spatio-temporal scales. The method does not deal with single earthquake prediction, but quantifies the chance of an earthquake by estimating the mean rate of future seismicity.

2 The Double Branching Model

In this study we apply the stochastic model proposed by Marzocchi and Lombardi (2008), consisting of a sequential application of two branching processes, in which any earthquake can trigger a family of later events. The main goal of our model is to account for interaction between events, due to different physical processes and involving largely different spatio-temporal domains. In the first step of our modeling we apply a version of well-known ETAS Model (Ogata, 1998), in order to describe the short-term clustering of seismic events in space and time. The second step of our procedure consists in re-applying a branching process to filtered database that is obtained by using the ETAS-derived declustering procedure. Notably, this second branch works at larger space-time scales compared to smaller domains involved by the short-term clustering, that is removed after the first step. The overall seismicity rate of Double Branching Model is given by

$$\lambda(t,x,y,m/\mathcal{H}_t) = \left[\nu u(x,y) + \sum_{t_i<t}\left(\frac{K_1 e^{\alpha_1(M_i-M_{min})}}{(t-t_i+c)^p}\frac{C_{d_1,q_1}}{(r_i^2+d_1^2)^{q_1}}\right) + \right.$$
$$\left. + \sum_{t_i<t} w_i\left(K_2 e^{\alpha_2(M_i-M_{min})} e^{-\frac{(t-t_i)}{\tau}}\frac{C_{d_2,q_2}}{(r_i^2+d_2^2)^{q_2}}\right)\right]\beta e^{\beta(m-M_{min})} \qquad (1)$$

where $\mathcal{H}_t = \{t_i, M_i, (x_i, y_i), \quad t_i < t\}$ is the observation history up time t and M_{min} is the minimum magnitude of catalog. The parameter ν indicates the overall background rate and $u(x,y)$ is the probability density function (PDF) of locations of spontaneous events. K_1, c and p are the parameters

of the modified Omori Law defining the temporal decaying of short-term triggering effect. The long-term triggering effect is described in time by an inverse exponential function with a characteristic time τ. This parameterization aims at reproducing the temporal evolution of the postseismic stress variations. Usually, the latter are modeled by a sum of exponential decays, mimicking different relaxation modes [Piersanti et al., 1995; Pollitz et al., 1992]; in our model we assume that one relaxation mode is predominant. Parameters α_1 and α_2 define, respectively, the dependence (assumed of exponential type) of short and long-term triggering effect with the magnitude of exciting event. The spatial decays of short and long-term stress variations are described by two inverse power PDF, with parameters (d_1, q_1) and (d_2, q_2), respectively (C_{d_1,q_1} and C_{d_2,q_2} are the normalization constants and r_i marks the distance between a general location (x, y) and the epicenter of the i-th earthquake (x_i, y_i)). For all events the magnitude distribution is assumed in agreement with a Gutenberg-Richter law (Gutenberg and Richter, 1954) with a parameter $\beta = b \cdot ln(10)$. Finally w_i is the probability that the i-th event is not coseismically triggered and it is calculated by using the ETAS model. Specifically by equation (1) we can compute the probabilities that the i-th event is short-term triggered (π_i^I), is long-term triggered (π_i^{II}) or is most related to tectonic loading (π_i^{III}). These probabilities are given by:

$$\pi_i^I = \frac{\sum_{t_j<t_i}\left(\frac{K_1 e^{\alpha_1(M_j-M_{min})}}{(t_i-t_j+c)^p}\frac{C_{d_1,q_1}}{(r_{ij}^2+d_1^2)^{q_1}}\right)}{\lambda(t_i,x_i,y_i,m_i/\mathcal{H}_{t_i})}$$

$$\pi_i^{II} = \frac{\sum_{t_j<t_i}\left(K_2 e^{\alpha_2(M_j-M_{min})}e^{\frac{-t-t_i}{\tau}}\frac{C_{d_2,q_2}}{(r_{ij}^2+d_2^2)^{q_2}}\right)}{\lambda(t_i,x_i,y_i,m_i/\mathcal{H}_{t_i})}$$

$$\pi_i^{III} = \frac{\nu u(x_i,y_i)}{\lambda(t_i,x_i,y_i,m_i/\mathcal{H}_{t_i})} \tag{2}$$

where r_{ij} is the distance between the epicenters of i-th and j-th events. The weights w_i of equation (1) are therefore given by:

$$w_i = 1 - \pi_i^I \tag{3}$$

To estimate the parameters of the model we use the iteration algorithm developed by Zhuang et al. (2002); the method is based on the Maximum Likelihood Method and on a kernel estimation of total seismic rate. Further details on the model and on estimation of its parameters can be found in Marzocchi and Lombardi (2008).

3 Application of Double Branching Model to Japanese catalog

Japan is one of most active seismic region of the world. It experiences more than 100 earthquakes at year with magnitude larger than 5.0 and more than 1-2 earthquakes with magnitude above 7.0. In order to estimate the model parameters we follow the guidelines given by the CSEP laboratory. Specifically we use the data collected by the JMA catalog since January 1 1965 up to December 31 2008 in the region $[110^o - 160^o W, 15^o - 50^o N]$ (background region), with magnitude above 5.0 and depth less than 100 km (5648 events). In Figure 1 we show the map of seismicity together with the boundaries of the forecasting region; this area defines the CSEP natural laboratory and it is used to compare and test the submitted models.

Following the procedure proposed by Zhuang et al. (2002), we estimate the model parameters together with the spatial distribution of not triggered seismicity ($u(x, y)$; see eq. (1)) by mean of Maximum Likelihood Method. Table 1 lists the inferred values of model parameters together with their errors. The values of parameters that controlling the short-term triggering are in agreement with values found in most tectonic region. The temporal decaying of long-term interaction has a characteristic time τ equal to about 30 yrs. The limited temporal window covered by JMA catalog (44 years)

could rise some doubts about the reliability of estimated τ-parameter. In any case we stress that the value of τ estimated in the present study is in agreement with what found in previous studies, both at global and regional scale (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009; 2010). In these studies we used datasets covering larger time windows (from one to several centuries), and then more suitable for our investigations. In order to check the reliability of the estimation of τ, we have verified that smaller values of τ do not provide a better fit of data (in terms of log-likelihood); then, we have investigated the stability of the parameter τ by changing the minimum magnitude. The procedure adopted to estimate the best model does not provide significantly different value of τ on earthquakes above M5.5 and M6.0 (about 900 and 300 events, respectively).

In Figure 2 we show the histograms of probabilities π_i^I and π_i^{III} of being short-term triggered and tectonically driven, respectively, for all events of learning dataset. From Figure 2a, we note that most of events have a probability π_i^I close to 0 and 1, revealing a well-defined identification of short-term triggered events. The histogram of probabilities π_i^{III} (Figure 2b) shows a more uncertain recognition of long-term triggered effects, although the distribution remain bimodal. In Figure 3 we compare the short-term and long-term decays of triggering functions. While at short time scales, each earthquake has a magnitude-dependent ability to trigger further events, at longer time

scales the capability to trigger other earthquakes appears to be independent by the magnitude ($\alpha_2 = 0$). In Figure 3 we can see that the short-term triggering effect given by a parent event with magnitude M6.0 is dominant for the first year; afterwards, the long-term magnitude-independent triggering becomes more important. For M 7.0, the short-term triggering prevails for the first 10 years. We underline that in Figure 3 we plot the probabilities of direct triggering, without taking into account secondary triggering effects.

In Figure 4 we show the distributions of background seismicity rate ($\nu u(x, y)$, panel a), of short term triggering rate (first sum of equation (1), panel b)) and of long-term triggering rate (second sum of equation (1), panel c)) in the forecast region. The main contribution to overall seismicity is given by tectonic loading (35%; panel a) and short-term triggering rate (50%; panel b), but the effect of long-term triggering is not negligible (15% of the overall rate). The long-term triggered seismicity appears to be more diffuse than short term triggered events. This is due to different distances involved by two triggering mechanisms. The viscoelastic relaxation, that we hypothesize to be a possible cause of long-term interactions, decays less rapidly than co-seismic effects with distance (see Marzocchi et al., 2003). Moreover the limited duration of the JMA catalog causes a lower spatial resolution of the long-term triggering respect to the analogous and stronger short-term effect.

4 Checking the Model

In order to make a very preliminary check of the forecasting capability of our model, we show in Figure 5 the map of predicted number of events for the period Jan 1 2009 - December 31 2009. We plot also the locations of events with $M \geq 5.0$ and at depth of 100 km or less that occurred in the same period inside the CSEP background-region and collected by the CMT (Centroid Moment Tensor, http://www.globalcmt.org/) database (67 events). All events occurred in cells with relatively high forecast rates. We remark that the forecasted rates are computed without taking into account the triggering effect of real events occurred during the year 2009. Including this effect in forecasting calculations might improve further these results.

A more careful checking of the performance of the DBM can be done by a comparison with the simpler ETAS model. A common diagnostic technique for stochastic point processes, called Residual Analysis (Ogata, 1988), is to transform the time axis t to a new scale $\tilde{t}$ by the increasing function

$$\tilde{t} = \Lambda(t) = \int_{T_{start}}^{t} dt' \int_{\mathcal{R}} dxdy\lambda(t', x, y/\mathcal{H}_{t'}) \tag{4}$$

where T_{start} is the starting time of observation history, $\mathcal{R}$ is the region under study and $\lambda(t, x, y/\mathcal{H}_t)$ is the conditional intensity of the model, parameterized by maximum likelihood parameters. $\Lambda(t)$ is the expected number of

events since time T_{start} up to time t, given the occurrence history $\mathcal{H}_t$. If the model describes well the temporal evolution of seismicity, the transformed data $\tilde{t_i} = \Lambda(t_i)$ (residuals) are expected to behave like a stationary Poisson process with the unit rate (Papangelou, 1972; Ogata, 1988). We apply the one-sample Kolmogorov-Smirnov test (KS1) (Gibbons and Chakraborti, 2003) on residuals of the ETAS model ($K_2 \equiv 0$, see eq. (1)) of the JMA catalog, used to set up the DBM model. We find that the Poisson hypothesis for the variable $\tilde{t_i}$ can be rejected at a significance level of 0.04. This means that the ETAS Model probably does not capture all basic features of seismicity collected into JMA catalog. On the other hand we find also that the log-likelihood of the ETAS model is larger then the log-likelihood of more sophisticated DBM, showing that the DBM does not improve the fit of the data respect to the ETAS model. This result is in disagreement with what found in other regions (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi 2009; 2010).

We argue that the scarce fit of the ETAS model and the lack of improvement with the DBM might be due to two different factors. First, probably there may be a bias into the distance between earthquakes; in fact, both ETAS and DBM consider only epicentral distances, neglecting the depth, whereas the latter can reach up to 100 km. Second, offshore and deep earthquakes may have different features compared to crustal inland earthquakes;

this difference may come up from a different resolution in monitoring (Nanjo et al., 2010); or may reflect a real physical difference between these two kind of earthquakes.

In the light of what just said, we decide to focus our analysis also to inland seismicity above 30 km of depth. Specifically we estimate the DBM on events occurred inside the mainland region, as defined by the Japanese CSEP laboratory (see www.cseptesting.org). The DBM parameters are listed in Table 2. The most striking result is a faster temporal decaying of the long-term interactions respect to other regions. The characteristic time τ, equal to about 8 years, is significantly smaller than values, all close to 30 years, estimated at local and global scale for the shallow seismicity (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009, 2010). Both the ETAS model and the DBM pass the KS1 test on residuals, but the DBM improves the likelihood on data.

In order to check if DBM significantly improves the performance of the more simple ETAS model, we follow the strategy proposed by Marzocchi and Lombardi (2009). Specifically we compute the information gain per event (IGpe), given by the difference of log likelihood of two models, DBM and ETAS, divided the number of events (N) into database

$$IGpe = \frac{LogL_{DBM} - LogL_{ETAS}}{N} \tag{5}$$

The JMA catalog provides $IGpe = 0.15$. To quantify the significance of this result, we compare the IGpe obtained for the JMA catalog (IGpe*) and for two sets of 1000 synthetic catalogs, simulated by using the ETAS model and the DBM. This comparison allows the forecasting performances of the two models, ETAS and DBM, to be tested. Specifically we assume as true the model used to simulate the synthetic datasets and we check if IGpe* can be seen as a random realization from the IGpe distribution obtained for the model under testing. We find that the ETAS model is rejected, being IGpe* above all 1000 values obtained by synthetic catalogs (see Figure 6). On the other hand, the IGpe distribution obtained for the synthetic DBM catalogs includes IGpe*. Therefore we conclude that the difference of log-likelihoods is significant and that DBM improves the performance of ETAS model.

An objection to the last conclusion should be that, to compare two models, we would have to resort to measures which penalize models with more degrees of freedom. This argument requests the use of measures as the Akaike Information Criterion, that in our case is given by

$$\Delta AIC = 2(LogL_{DBM} - LogL_{ETAS}) - 2(k_{DBM} - k_{ETAS}) \tag{6}$$

where k_{DBM} and k_{ETAS} are the numbers of free parameters for DBM and ETAS model, respectively. Therefore a better indicator of the improvement in predictability is given by

$$\frac{\Delta AIC}{2N} = IGpe - \frac{(k_{DBM} - k_{ETAS})}{N}. \tag{7}$$

To compare real and simulated values of $\frac{\Delta AIC}{2N}$, we can simply translate the values of $IGpe$ shown in Figure 6 by the factor $-\frac{(k_{DBM}-k_{ETAS})}{N}$. So the conclusions on the significance of the improvement of DBM do not change.

In order to check further our results, we compare the forecasting performances of ETAS and DBM using a dataset that has not used to calibrate the models (cf. Marzocchi and Lombardi, 2009). This goal can be achieved by dividing the available dataset in two parts: a first (in chronological meaning) part of dataset, hereinafter *learning* dataset, can be used to set up the model and a second, the *testing* dataset, to check its reliability. In this case the forecasts have zero degrees of freedom, since each model uses information available before the starting time of each test day. Therefore comparing the model likelihoods is sufficient. We set the *learning* dataset as the part of the JMA catalog spanning the time interval 1965-2005 and the *testing* dataset

as the subset of events occurred from 2006 to 2008. We do not find any significant change into parameters of both ETAS model and DBM. In Figure 7 we show the comparison of the real and simulated values of the IGpe, which remarkably confirms the significance of the improvement of DBM, respect to ETAS model.

5 Forecasting Maps

The model formulated and tested in previous sections allows us to compute forecasts of future seismicity. For the sake of example, we show a map of probability of occurrence for at least one earthquake with $M \geq 5.0$ and depth lower than 100km, within a zone of $0.1° \times 0.1°$, in the next 5 (January 2010 - December 2014) years in Japan. We stress that the magnitude range and the spatial boundaries are in agreement with choices adopted for the CSEP Japan experiment, which has mostly motivated this study. To produce forecasting calculations the DBM requires to take into account the triggering effect of seismicity occurred both before and during the forecast interval. Since this last is unknown, we simulate 10000 different stochastic realizations along the forecasting time window. For this purpose we use 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. Results are shown in Figure 8. The DBM identifies as most dangerous zones the north-western coast of the testing region. One of most hazardous zones on the mainland is the region near Tokyo city and the Tokai region, in which a strong earthquake is expected in short on the basis of different models (Rikitake, 1999; Mogi, 2003).

6 Discussion and Conclusions

The main goal of the present paper has been to describe the DBM applied to Japanese seismicity. This study has been mainly motivated by the participation to the CSEP experiment for the Japanese testing region. From a seismological point of view, the results obtained in the present study basically confirm the main finding of previous analyses (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009; 2010). Large earthquakes in Japan tend to cluster in time and space at different time scales. Besides the short-term clustering described by the ETAS model, we have found also a significant time clustering longer than typical aftershock sequences. Notably, we have found that the DBM has a different forecasting performance on shallow and deep seismicity. Specifically the DBM has a poor reliability on seismicity with a depth up to 100 km, whereas it works significantly better than ETAS model for shallow seismic events (up to 30 km of depth). The

scarce fitting with DBM and ETAS model is probably due to the use of epicentral distances; neglecting depth may alter significantly the real distance between earthquakes. Another possibility is that offshore and deeper earthquakes have different features compared to crustal inland earthquakes. The characteristic time of the second branching for the crustal earthquakes ($\tau \sim 8$ years) seems to be smaller than the characteristic time found in other regions (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009, 2010). We explain this shorter time length as due to the high seismic background for Japanese seismicity. In fact, the time decay of the long-term interaction will fade sooner into the background seismicity when the latter is higher. In any case the value of τ is in agreement with Lombardi and Marzocchi (2007), which founded a significant variation of background seismicity in Japan about every 10 years.

In all our previous analyses (Marzocchi and Lombardi, 2008; Lombardi and Marzocchi, 2009, 2010), as well as in the present study, we find a low value for α_2, not statistically significant from zero. This implies that the postseismic triggering capability of an event is independent by its magnitude. By a physical point a view, we explain this finding with the not-suitability of the available data to provide the actual value of α_2. In fact, whereas the coseismic stress transfer is a phenomenon spanning all magnitude scales, the postseismic effects are likely mostly caused by the strongest events [Piersanti

et al., 1997; Pollitz et al., 1998]. The magnitudo range recovered by JMA and by all previously analyzed catalogs is rather small and the proportion of giant events ($M \geq 8.0$) is negligible. This could be the origin of the indeterminateness of the α_2-value. By a statistical point of view, we have shown that a further explanation for the uncertain estimate of the α_2-value could be the inefficiency of data to reveal its actual value (Lombardi and Marzocchi, 2009). Specifically we have shown that the probability to estimate a null value for α_2 is not negligible also for a class of synthetic catalogs with α_2 significantly different from 1.0 and a size comparable with the available real datasets. This is clear evidence that the limited number of data of a catalog might prevent to find a positive value of α_2 significantly different from zero.

The first version of the DBM submitted for CSEP Japanese laboratory is focused on providing earthquakes forecast until 100 km depth. The results of this analysis has encouraged us to submit a new version of DBM focused only on the crustal inland earthquakes. We expect that this second version of the DBM should work better than classical ETAS models. The results of the CSEP experiment in the Japanese testing region will provide us interesting insights on this topic.

7 References

Console R., Murru M., G. Falcone, 2010. Probability gains of an epidemic-type aftershock sequence model in retrospective forecasting of M¿5.0 earthquakes in Italy, *J. of Seismology*, **14**, 9-26.

Gerstenberger M., S. Wiemer, L.M. Jones, P.A. Reasenberg, 2005. Real-time forecasts of tomorrow's earthquakes in California, *Nature*, **435**, 328-331.

Gutenberg B., C.F. Richter, 1954. Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.*, **34**, 185-188.

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

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

Kagan, Y.Y., Jackson, D.D., 1991. Seismic gap hypothesis, ten years after, *J. Geophys. Res.*, **96**, 21419-21431.

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

Lombardi, A.M., Marzocchi, W., 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 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., W. Marzocchi, 2009. Double Branching model to forecast the next M$\geq$ 5.5 earthquakes in Italy, *Tectonophysics*, **475**, 514-523, doi:10.1016/j.tecto.2009.06.014

Lombardi, A.M., M. Cocco and W. Marzocchi, 2009. 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.

Lombardi A.M. and W. Marzocchi, 2010. The Double Branching Model (DBM) applied to long-term forecasting Italian seismicity (Ml$\geq$5.0) in CSEP experiment, in press on Annals of Geophysics.

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

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

Marzocchi, W., Selva J., Cinti F.R., Montone P., Pierdominici S., Schivardi R., Boschi E., 2009, On the occurrence of large earthquakes: New insights from a model based on interacting faults embedded in a realistic tectonic setting, *J. Geophys. Res.*, **114**, B01307, doi:10.1029/2008JB005822.

Mogi K., 1987, Recent seismic activity in the Tokai (Japan) region where a large earthquake is expected in the near future, *Tectonophys.*, **138**, 255-268.

Nanjo K.Z., T. Ishibe, H. Tsuruoka, D. Schorlemmer, Y. Ishigaki, and N. Hirata, 2010. Analysis of the Completeness Magnitude and Seismic Network Coverage of Japan, *Bull. Seismol. Soc. Am.*, **100**, doi: 10.1785/0120100077

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

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

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

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

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

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

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

Rikitake T., 1999. Probability of a great earthquake to recur in the Tokai district, Japan: reevaluation based on newly-developed paleoseismology, plate tectonics, tsunami study, micro-seismicity and geodetic measurements., *Earth Planets Space*, **51**, 147-157.

Schorlemmer D., Gerstenberger, M. C., Wiemer, S., Jackson, D.D., Rhoades, D.A., 2007. Earthquake Likelihood Model Testing, *Seism. Res. Lett.*, **78**, No. 1, 17-29.

Schorlemmer D., Zechar J.D., Werner M.J., Field E.H., Jackson d. D., Jordan T.H. and the RELM Working Group, First Results of the Regional Earthquake Likelihood Models Experiment, submitted on *Pure and Applied Geophysics.*

Tsuruoka, H., Hirata, N., Schorlemmer, D., Euchner, F., Jordan, T. H.,

2008. CSEP Earthquake Forecast Testing Center for Japan, American Geophysical Union, Fall Meeting 2008.

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

8 Table Captions

Table 1: Maximum Likelihood parameters (with relative errors) of the Double Branching model (see equation(1)) for the events of the JMA catalog above 100 km of depth (Mc = 5.0; Jan 1 1965 Dec 31 2008, 5648 events).

Table 2: Maximum Likelihood parameters (with relative errors) of the Double Branching model (see equation(1)) for the events of the JMA catalog above 30 km of depth (Mc = 5.0; Jan 1 1965 Dec 31 2008, 1935 events).

9 Figure Captions

Figure 1: Map of seismic events collected in the JMA catalog used in the present study (Jan 1 1965 Dec 31 2008, M5.0; 5648 events). The symbol sizes are scaled with magnitude. The shadow area identifies the testing region, used by CSEP laboratory to test and compare the models.

Figure 2: Histograms of a) probabilities π_i^I to be short-term triggered and b) probabilities π_i^{III} of belonging to background seismicity for events collected into JMA catalogs.

Figure 3: Probabilities to trigger an event above $M5.0$ on short term scale (Omori function) given by a parent event of magnitude 6.0 and 7.0. These probabilities are compared with the magnitude-independent triggering function on long term scale, which has an exponential decay (see text for

details) .

Figure 4: Maps of a) the spatial distribution of tectonic-driven seismicity u(x,y), b) the short-term triggered rate and c) the long-term triggered rate (see eq. (1) and the text for details).

Figure 5: Map of seismic rates (number of events in a cell of $0.1x0.1$) predicted by Double Branching Model for the period January 1 2009-December 31 2009, inside the CSEP testing region. Black circles mark the locations of 67 events occurred in the same period collected by the CMT dataset.

Figure 6: Plot of IGpe for the whole JMA catalog (1965-2008; depth $\leq$ 30km; $M \geq 5.0$; vertical solid line) and for synthetic catalogs obtained by the ETAS model and the DBM.

Figure 7: The same of Figure 6 but for the *testing* JMA catalog (2006-2008; see text for details).

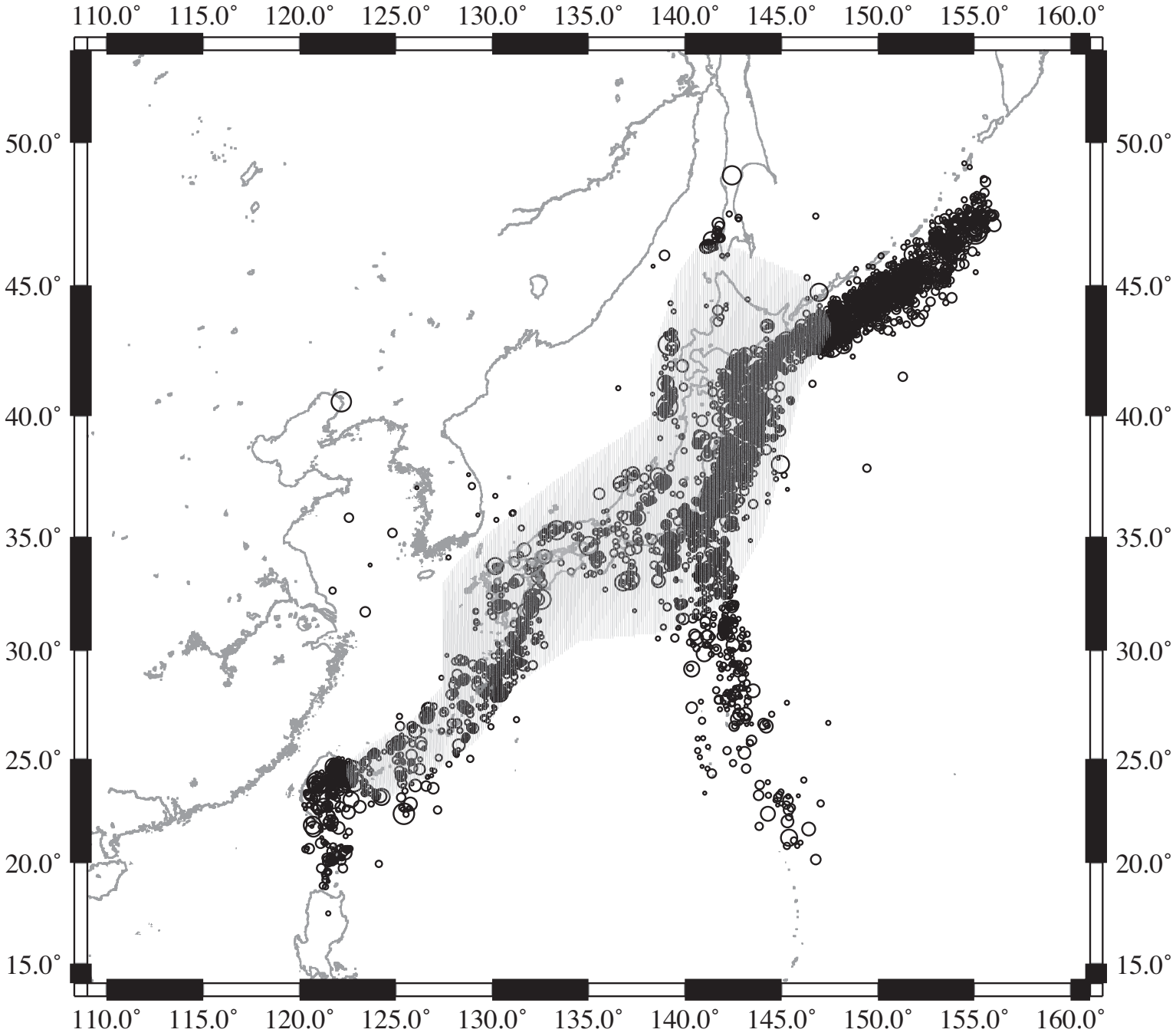
Figure 8: Map of rate of occurrence (number of events per cells) of events with magnitude above M 5.0 for the next 5 years. Blue squares mark the most hazardous areas.

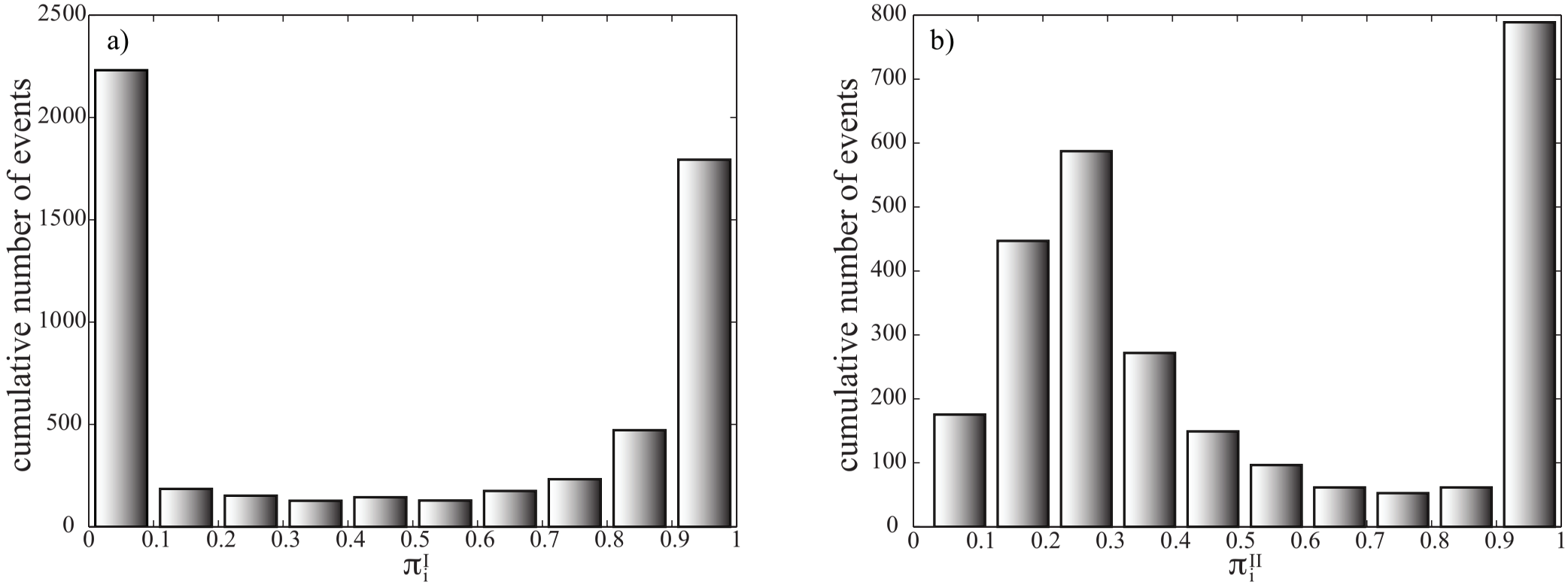
Table 1:

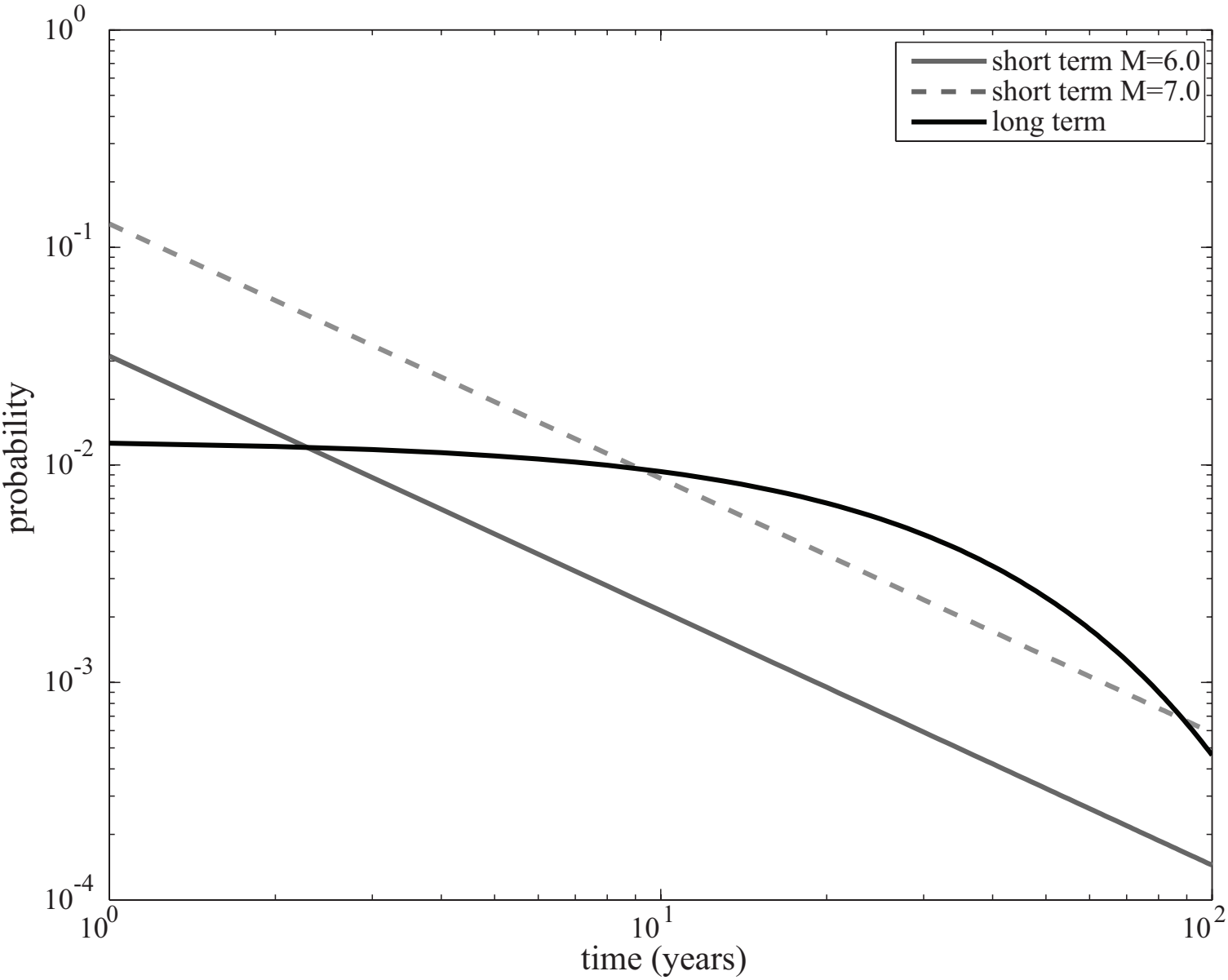
Parameter	Value
ν	$61 \pm 2\ year^{-1}$
K	$7.8 \pm 0.5 \cdot 10^{-3}\ year^{p-1}$
p	1.17 ± 0.01
c	$7.0 \pm 1.0 \cdot 10^{-5}\ year$
α_1	1.40 ± 0.04
d	$4.6 \pm 0.2\ km$
q	$\equiv 1.5$
γ	0.53 ± 0.03
K_2	0.013 ± 0.001
τ	$30 \pm 6\ year$
α_2	~ 0.0
d_2	$82 \pm 7\ km$
q_2	1.5 ± 0.2

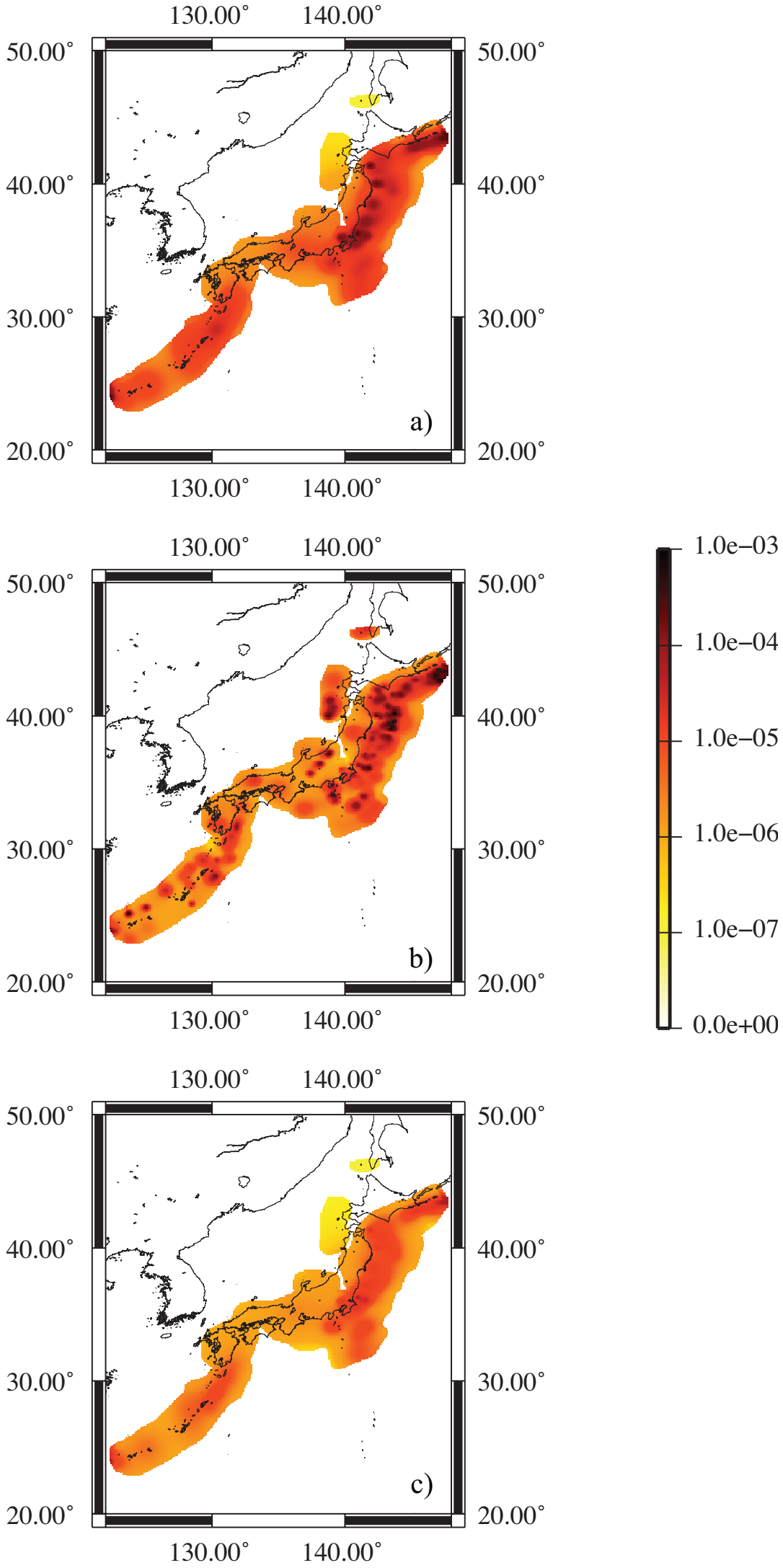
Table 2:

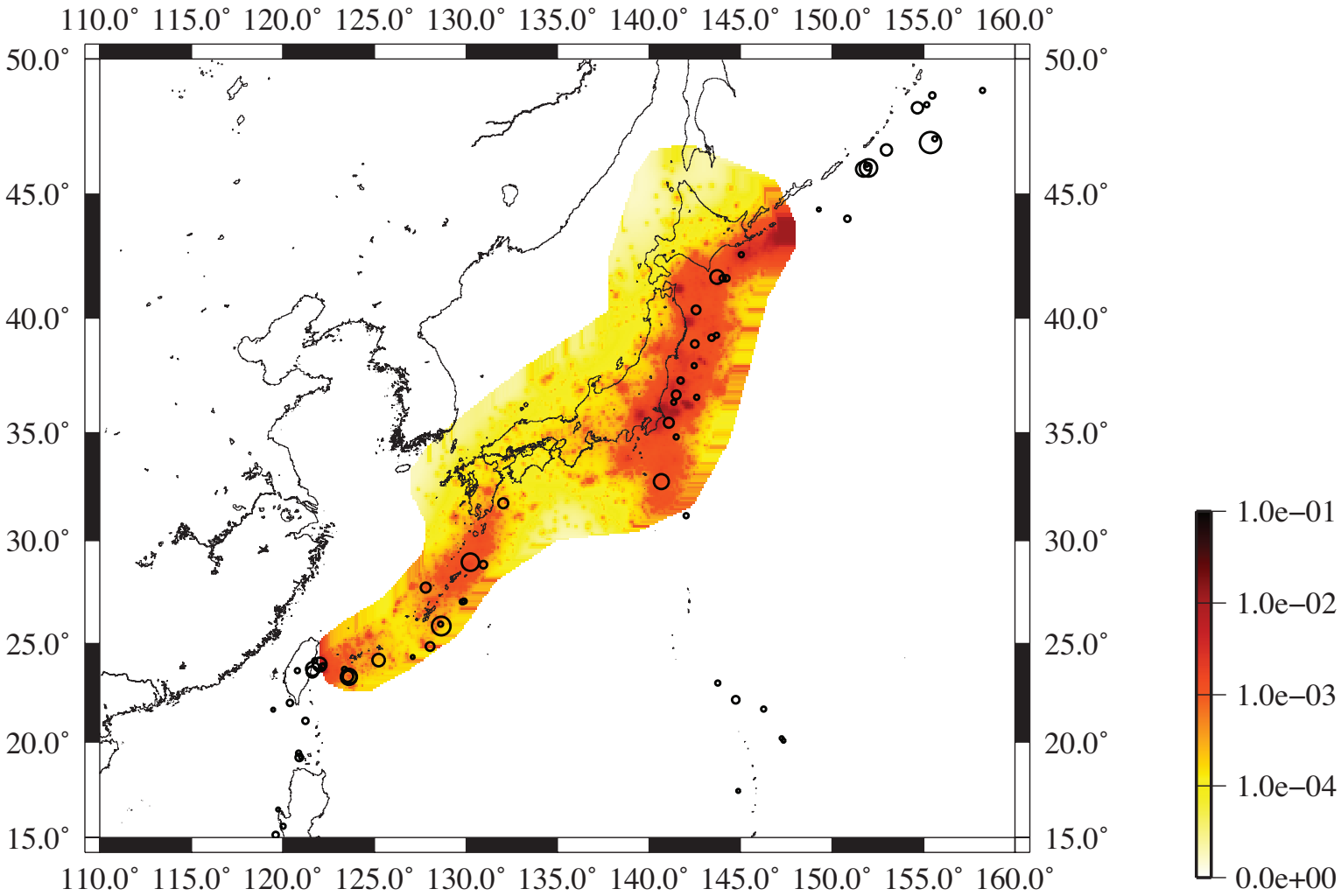
Parameter	Value
ν	$11 \pm 1\ year^{-1}$
K	$1.1 \pm 0.1 \cdot 10^{-2}\ year^{p-1}$
p	1.16 ± 0.01
c	$6.0 \pm 1.0 \cdot 10^{-5}\ year$
α_1	1.20 ± 0.04
d	$4.6 \pm 0.2\ km$
q	$\equiv 1.5$
γ	0.53 ± 0.03
K_2	0.09 ± 0.01
τ	$8 \pm 1\ year$
α_2	~ 0.0
d_2	$24 \pm 4\ km$
q_2	2.0 ± 0.2

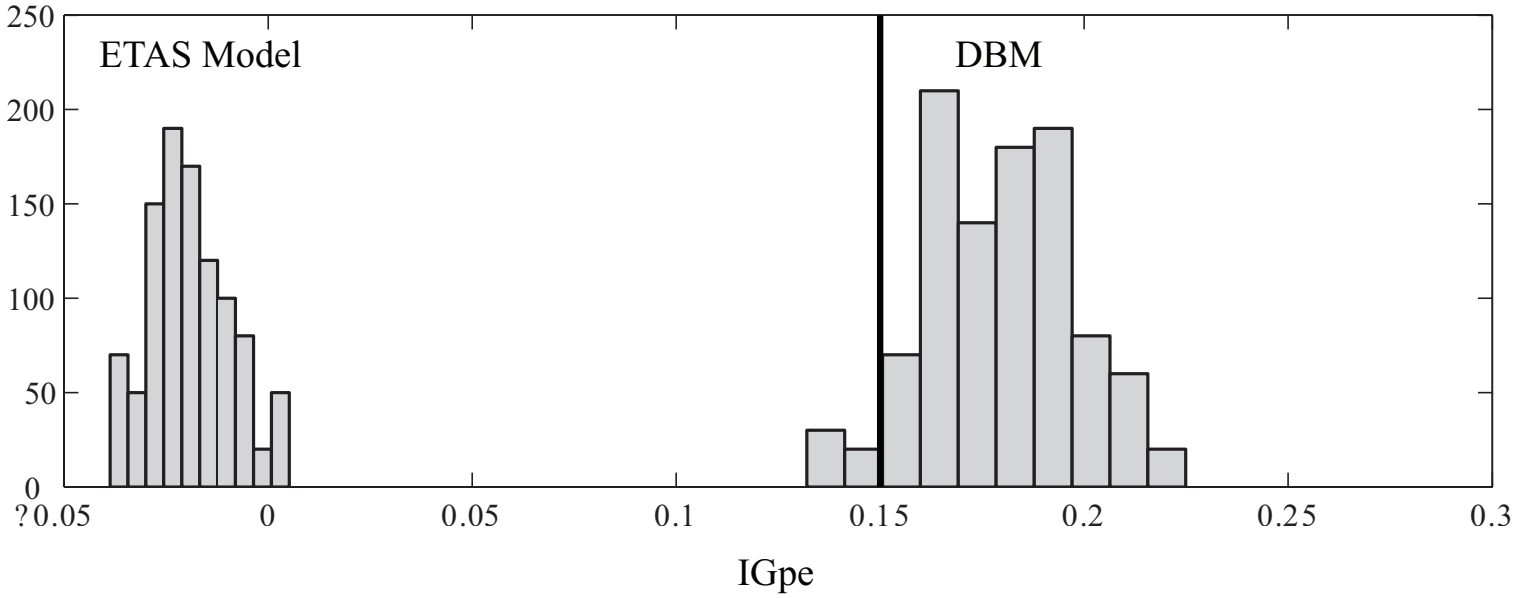

110.0°
115.0°
120.0°
125.0°
130.0°
135.0°
140.0°
145.0°
150.0°
155.0°
160.0°
50.0°
45.0°
40.0°
35.0°
30.0°
25.0°
20.0°
15.0°

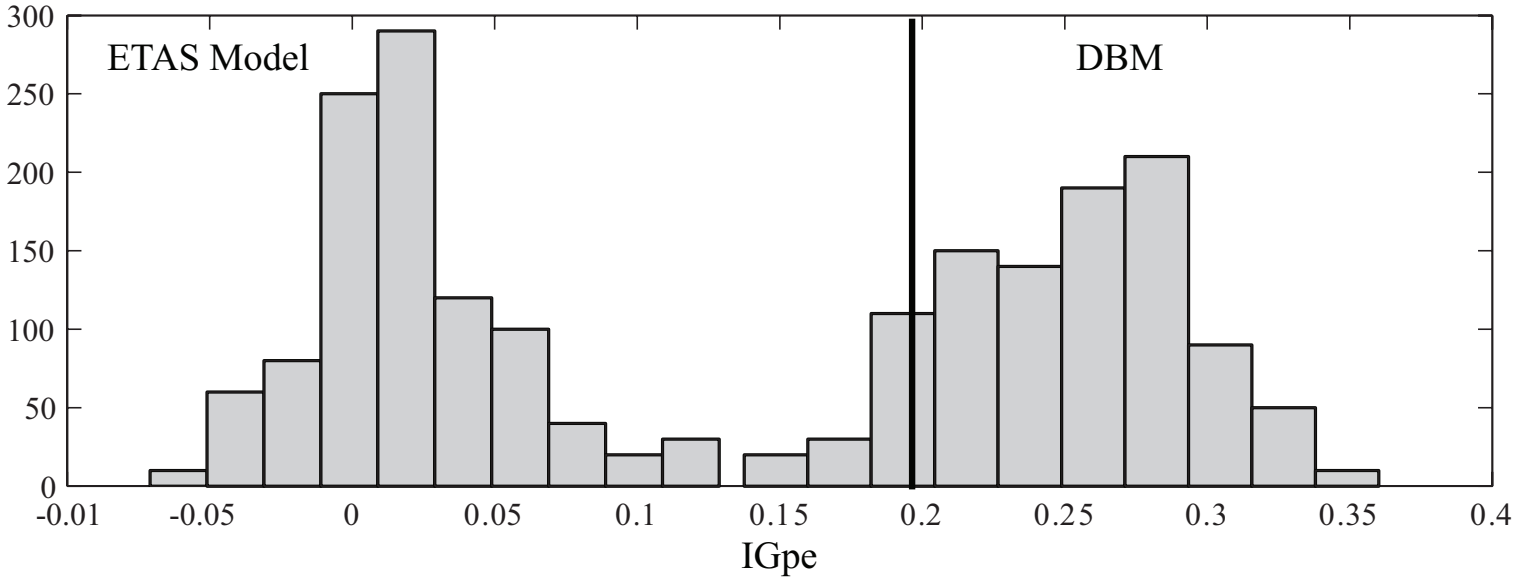

a)
cumulative number of events
2500
2000
1500
1000
500
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
π_i^I
b)
cumulative number of events
800
700
600
500
400
300
200
100
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
π_i^{II}

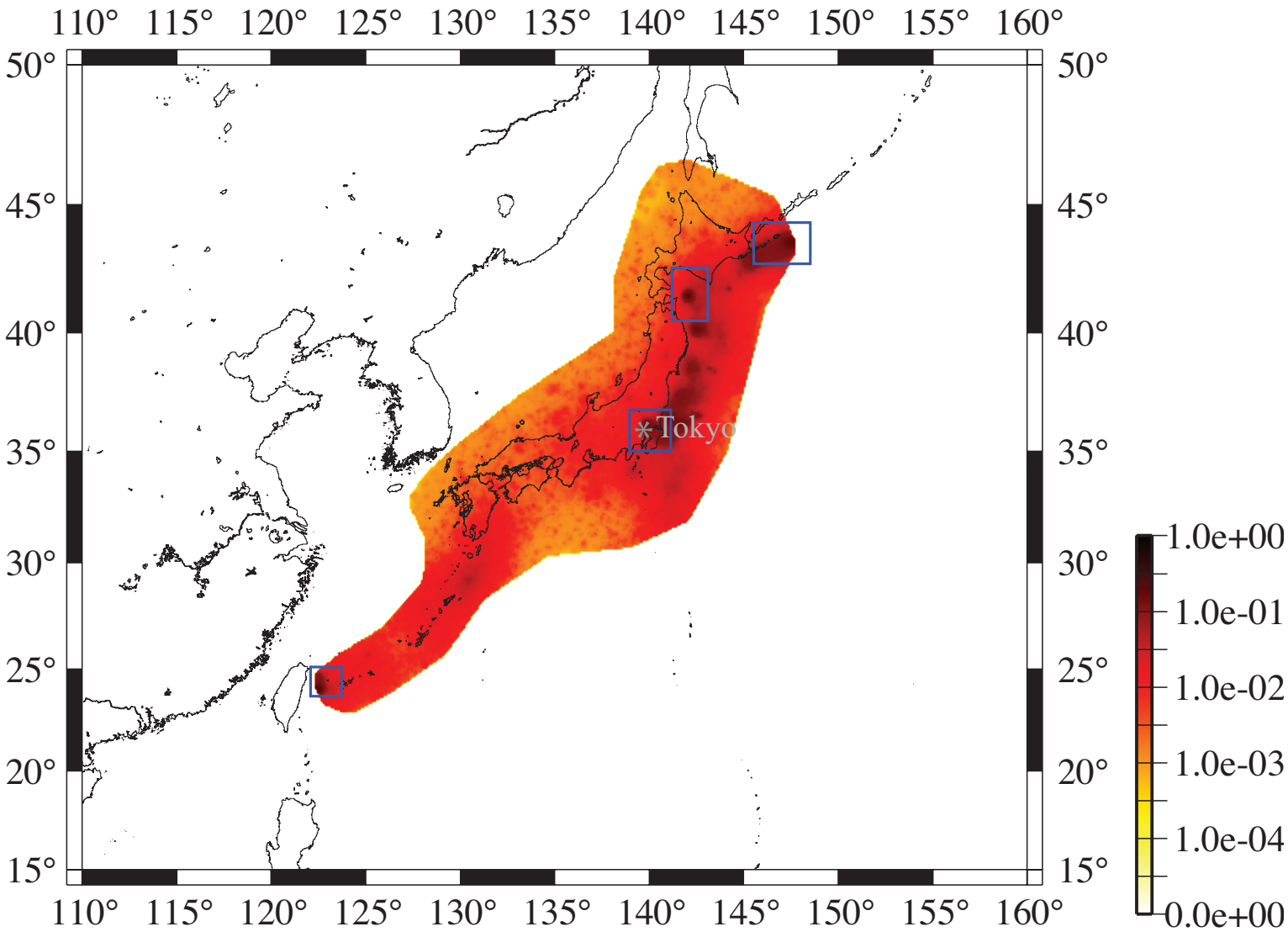

short term M=6.0
short term M=7.0
long term
probability
time (years)
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{0}
10^{1}
10^{2}


130.00°
140.00°
50.00°
40.00°
30.00°
20.00°
a)
b)
c)
1.0e−03
1.0e−04
1.0e−05
1.0e−06
1.0e−07
0.0e+00


110.0° 115.0° 120.0° 125.0° 130.0° 135.0° 140.0° 145.0° 150.0° 155.0° 160.0°
50.0°
45.0°
40.0°
35.0°
30.0°
25.0°
20.0°
15.0°
1.0e−01
1.0e−02
1.0e−03
1.0e−04
0.0e+00


ETAS Model
DBM
250
200
150
100
50
0
?0.05
0
0.05
0.1
0.15
0.2
0.25
0.3
IGpe


ETAS Model
DBM
300
250
200
150
100
50
0
-0.01
-0.05
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
IGpe


110° 115° 120° 125° 130° 135° 140° 145° 150° 155° 160°
50°
45°
40°
35°
30°
25°
20°
15°
Tokyo
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
0.0e+00