Temporal Distribution Model and Occurrence Probability of M≥6.5 Earthquakes in North China Seismic Zone

Weijin Xu (wjxuwin@163.com)
Institute of Geophysics China Earthquake Administration https://orcid.org/0000-0002-5843-0498
Wu Jian
China Earthquake Disaster Prevention Center
Mengtan Gao
Institute of Geophysics China Earthquake Administration

Research Article

Keywords: Temporal distribution model, Brownian passage time (BPT) distribution, Lognormal distribution, Probability of earthquake occurrence.

Posted Date: October 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2159311/v1

Version of Record: A version of this preprint was published at Natural Hazards on August 9th, 2023. See the published version at https://doi.org/10.1007/s11069-023-06124-5.

Abstract

The temporal distribution of earthquakes provides important basis for earthquake prediction and seismic hazard analysis. The relatively limited records of strong earthquakes have often made it difficult to study the temporal distribution models of regional strong earthquakes. However, there are hundreds of years of complete strong earthquake records in North China Seismic Zone, providing abundant basic data for studying temporal distribution models. Using the data of M ≥ 6.5 earthquakes in North China as inputs, this paper estimates the model parameters using the maximum likelihood method with exponential, Gamma, Weibull, Lognormal and Brownian passage time (BPT) distributions as target models. The optimal model for describing the temporal distribution of earthquakes is determined according to Akaike information criterion (AIC), determination coefficient R^2 and Kolmogorov-Smirnov test (K-S test). The results show that Lognormal and BPT models perform better in describing the temporal distribution of strong earthquakes in North China. The mean recurrence periods of strong earthquakes (M ≥ 6.5) calculated based on these two models are 8.1 years and 13.2 years, respectively. In addition, we used the likelihood profile method to estimate the uncertainty of model parameters. For the BPT model, the mean and 95% confidence interval of recurrence interval μ is 13.2 (8.9–19.1) years, and the mean and 95% confidence interval of α is 1.29 (1.0-1.78). For the Lognormal model, the mean value and 95% confidence interval of v is 2.09 (1.68–2.49), the mean value exp (v) corresponding to earthquake recurrence interval is 8.1 (5.4–12.1) years. In this study, we also calculated the occurrence probability of M ≥ 6.5 earthquakes in North China Seismic Zone in the future, and found that the probability and 95% confidence interval in the next 10 years based on the BPT model is 35.3% (26.8%-44.9%); the mean value and 95% confidence interval of earthquake occurrence probability based on the Lognormal distribution is 35.4% (22.9%-49.7%); the mean probability and 95% confidence interval based on the Poisson model is 53.1% (41.1%-64%). The results of this study may provide important reference for temporal distribution model selection and earthquake recurrence period calculation in future seismic hazard analysis in North China.

1 Introduction

The temporal distribution models of earthquakes represent an important part of seismological research. For a long time, it has been widely assumed in probabilistic seismic hazard analysis and earthquake prediction that earthquakes follow the Poisson distribution in time (Cornell, 1968; Gardner and Knopoff, 1974; Schwartz et al., 1984; Frankel, 1995; Console et al., 2003; Petersen et al., 2014). Also known as the time-independent model, the Poisson model holds that the occurrence of earthquakes does not change with time, and represents the theoretical basis for probabilistic seismic hazard analysis in some countries at present (Cornell, 1968; Petersen et al., 2014; Gao 1995; Pan et al., 2013).

Aside from the time-independent model, there are also time-dependent models for describing the occurrence of earthquakes. Many studies have proved that time-dependent models perform better in describing the temporal distribution characteristics of earthquakes in certain areas (Utsu, 1984; Nishenko et al., 1987; Ellsworth, 1995; Ogata, 1991; Tripathi, 2006; Sharma et al., 2010). Seismologists have put forward such statistical models as Gamma, Lognormal, Weibull and Brownian passage time (BPT) functions (Utsu, 1984; Matthews et al., 2002; Tripathi, 2006; Pasari et al., 2015, 2018; Bajaj et al., 2019). Contrary to the Poisson model, the time-dependent models hold that the occurrence of earthquakes varies with time, exerting an important influence on the results of seismic hazard.

In addition, the Epidemic Type Aftershock Sequence (ETAS) model has also been used to study the spatio-temporal distribution of earthquakes in recent years (Ogata, 1988; Ogata, 1998; Zhuang et al., 2005; Ogata and Zhuang, 2006), and has been proposed or used by scientists as a means for probabilistic seismic hazard analysis (e.g., Xu et Wu, 2017; Pei et al., 2021; Šipčić et al., 2022). However, for seismic hazard analysis, more attention is paid to the temporal distribution of main shocks (Michael, 2011; Daub et al., 2012; Shearer and Stark, 2012; Beroza, 2012). In fact, in the ETAS model, background earthquakes are sometimes still regarded as following the Poisson distribution (Ogata, 1988; Zhuang et al., 2005; Lombardi and Marzocchi, 2007). In applying the ETAS model to probabilistic seismic hazard analysis, Šipčić et al. (2022) also identified the characteristics of the temporal changes of background seismicity as a key area for future research, which suggests that studying the temporal distribution of main shocks will also contribute to the effective use of the ETAS model.

Research shows that strong earthquakes ($M \geq 6.5$) will generate fault rupture of a certain scale, which has a considerable impact on seismic hazard analysis (Schwartz and Coppersmith, 1984; Frankel et al., 2002). For a long time, the study of the temporal distribution characteristics of strong earthquakes has been constrained by limited strong earthquake records. Fortunately, there are hundreds of years of complete strong earthquake records in North China, providing excellent basic data for studying temporal distribution models. The present study will focus on the temporal distribution models for strong earthquakes ($M \geq 6.5$) in North China Seismic Zone. The North China Seismic Zone boasts the most abundant seismological records in China, with records of strong events spanning hundreds of years. Since events in this region are subject to similar seismic structures and geodynamic environments, Chinese seismologists often take North China Seismic zone as a statistical unit for understanding seismicity characteristics (Gao et al., 2013; Xu and Gao, 2015). In this study, based on the catalogs of $M \geq 6.5$ events in North China Seismic Zone, and using Poisson (exponential distribution), Gamma, Lognormal, Weibull and BPT as target models, we regressed the parameters of each model by using the maximum likelihood method and selected the optimal models for describing the temporal distribution of seismicity using the Akaike information criterion (AIC), the coefficient of determination R^2 and the Kolmogorov-Smirnov test (K-S test). Based on the optimal models, we calculated the occurrence probability of $M \geq 6.5$ events in future in the study area. The results may provide theoretical basis for the selection of temporal distribution models and the calculation of seismicity parameters in seismic hazard analysis in North China Seismic Zone.

2 Seismic Catalogs

We obtained the data of historical strong earthquakes from *The Catalogue of Chinese Historical Strong Earthquakes* (23rd century B.C.~1911 A.D.) (Department of Earthquake Disaster Prevention, National

Earthquake Administration, 1995), the data between 1912 and 1990 from *The Catalogue of Chinese Earthquakes* (1912 ~ 1990) (Department of Earthquake Disaster Prevention, National Earthquake Administration, 1999), and the data after 1990 from *The Catalogue of Earthquakes in China and Adjacent Areas after 1990*. (Lv et al., 2010; Xu and Gao, 2014). Figure 1 shows the distribution of epicenters of M ≥ 6.5 earthquakes in North China Seismic Zone.

As this study focuses on the temporal statistical characteristics of main shocks, it is necessary to remove aftershocks from the catalogs. At present, many methods are available for seismic declustering, such as the traditional space-time window (Gardner and Knopoff, 1974), the stochastic declustering approach based on the ETAS model (Zhuang et al., 2002), and the nearest neighbor distance method (Baiesi and Paczuski, 2004). Among them, the space-time window of Gardner and Knopoff (1974) is widely used in seismicity analysis and seismic hazard analysis (Shearer and Stark, 2012; Daub et al., 2012; Petersen et al., 2014; Xu and Gao, 2015). In this paper, the space-time window method was used to remove aftershocks from the catalogs in the Chinese mainland, in which the space window parameters calculated by Chen et al. (2019) were adopted.

The completeness of catalogs represents an important factor affecting the analysis results. Huang et al. (1994) analyzed the completeness and reliability of the catalogs of historical earthquakes in North China (M_S≥4.8). Through a comparison with the characteristics of earthquake damage recorded by modern instruments, they showed that the catalogs of historical earthquakes in North China are complete and reliable for the purpose of seismological study. The complete records of earthquakes started from 1480. Xu and Gao (2015) used more statistical methods to investigate the completeness of the historical catalogs in North China, and concluded that the records of M_S≥4.8 earthquakes in this region are complete since 1500. The catalogs of historical earthquakes in China provide crucial data for disaster prevention, and are widely used in earthquake prediction, probabilistic seismic hazard analysis and compilation of seismic hazard maps (Huang et al., 1994; Lv et al., 2010; Pan et al., 2013; Gao,1996, 2003, 2015; Xu, 2019). Therefore, for this study, complete records are available for M_S≥6.5 earthquakes in North China since 1500. We obtained 37 M ≥ 6.5 events, which exceeds the requirement of having at least 25 events in order to distinguish temporal distribution models, as proposed by Matthews et al. (2002).

3 Models And Methods

3.1 Statistical models of earthquake recurrence intervals

In this study, we analyze the statistical characteristics of earthquake recurrence intervals, defined as the time intervals between successive events. In engineering seismology, earthquake recurrence interval is also called return period. Its reciprocal, which is the frequency of earthquake occurrence per unit of time, is a crucial parameter for probabilistic seismic hazard calculation and earthquake prediction.

Commonly used models in statistical seismology at present include exponential (Poisson), Gamma, Lognormal, Weibull and BPT (Utsu, 2002; Matthews et al. 2002; Bajaj and Sharma, 2019). In this study, we use the above five models to analyze the statistical characteristics of recurrence intervals in North China (Table 1).

Table 1 Statistical distribution models

Model name	Probability density function and cumulative distribution function of statistical models and their parameters
Exponential	$f(x)=\frac{1}{\mu}e^{\frac{-x}{\mu}}$, $F(x)=1-e^{\frac{-x}{\mu}}$, where μ is the mean value.
Weibull model	$f(x)=\frac{\beta}{\alpha}(x/\alpha)^{\beta-1}e^{-(x/\alpha)^{\beta}}$, $F(x)=1-e^{-(x/\alpha)^{\beta}}$. In the equation, α is the scale parameter and β is the shape parameter.
Gamma model	$f(x)=\frac{x^{(\lambda-1)}\exp^{(-x/\lambda)}}{\lambda^{k}\Gamma k}$, $F(x)=\frac{1}{\Gamma k}\int_0^{x/\lambda}u^{k-1}\exp^{-u}du$, where $\Gamma k=\int_0^{\infty}t^{k-1}\exp^{-t}dt$ is Gamma function, k and λ are shape and scale parameters respectively.
Lognormal model	$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}\exp^{\frac{-(\ln(x)-\mu)^2}{2\sigma^2}}$, $F(x)=\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)$, where Φ is the cumulative probability distribution function of normal distribution, μ and σ are the mean and standard deviation of logarithmic values of x respectively.
Brownian passage time model	$f_{BPT}(t)=\sqrt{\frac{\mu}{2\pi\alpha^2t^3}}\exp\left(-\frac{(t-\mu)^2}{2\alpha^2\mu t}\right)$, $F(t)=P(T\le t)=\int_0^t f_{BPT}(\tau)d\tau=\Phi[u_1(t)]+e^{\frac{2}{\alpha^2}}\Phi[-u_2(t)]$, where μ is the mean of recurrence interval, $\alpha=\sigma/\mu$, and σ is the standard deviation of recurrence interval. $u_1(t)=\alpha^{-1}(t^{1/2}\mu^{-1/2}-t^{-1/2}\mu^{1/2})$, $u_2(t)=\alpha^{-1}(t^{1/2}\mu^{-1/2}+t^{-1/2}\mu^{1/2})$ $\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}e^{-x^2/2}dx$, Φ is the cumulative probability function of standard normal distribution.

3.2 Model parameter estimation and goodness-of-fit (GOF)

The maximum likelihood method was used to estimate the parameters in the above statistical models. The maximum likelihood method was developed to estimate the range of potential parameters by using the data per se (Fisher, 1922), and raised for the first time such a question: given a random sample and its distribution model, which parameters are most likely to produce the sample? In other words, the parameters estimated by the maximum likelihood method can maximize the occurrence probability of the current random sample. As the intervals between earthquakes in this study are definite, the maximum likelihood method may be used to estimate the parameters and their uncertainties. The aforementioned five models all have probability density functions, and the likelihood function is the joint probability density function of random variable *x*, which can be written as:

$$L(\theta) = \prod_{i=1}^{N} f(x_i|\theta)$$

where θ represents one or more model parameters, N is the sample length of the random variable, and f is the probability density function of the statistical model. The natural logarithm of the likelihood function is obtained and the coefficient θ is derived. The parameter values in the model can be obtained by solving the likelihood equation (group).

In this study, the distribution of earthquake recurrence interval is unknown. To know whether they conform to a statistical model, we used the K-S test to test the difference between the observed values of recurrence intervals and their theoretical distributions. The statistic of the K-S test represents the biggest difference in cumulative distribution probability between observed and theoretical values:

$$D_N = \max|O_i - E_i|$$

where O_i is the cumulative probability of observed values, and E_i is the cumulative probability of the theoretical model. Smaller D_N results in better goodness-of-fit.

To select the optimal model, we used AIC and BIC to determine the goodness-of-fit. AIC and BIC are respectively defined as:

$$AIC = 2k - 2\ln(L)$$

where k is the number of model parameters; L is the likelihood function; N is the sample size. Generally, the model with the smallest AIC value is chosen as the optimal model from available alternatives.

In addition, the coefficient of determination R^2 was used to characterize the goodness-of-fit between the observed data and the theoretical distribution models. See Gibbons and Chakraborti (2003) for a detailed description of relevant test methods.

4 Results

4.1 Temporal distribution models and parameters

Table 2 shows the parameter values of each model regressed using the catalogs of North China Seismic Zone and GOF parameter values. For strong earthquakes of magnitude 6.5 or above, Table 2 shows that the AIC and K-S test D_N values of Lognormal and BPT models are close, with smaller values than those of other models. From the coefficient of determination R^2, it can be seen that the R^2 of BPT and Lognormal models are relatively larger, which means a higher probability for earthquake recurrence intervals to follow these two models. The cumulative distribution curves in Fig. 2 shows that the BPT and Lognormal models are in better agreement with empirical data. From the above analysis, it can be concluded that Lognormal and BPT distributions can better describe the temporal statistical distribution characteristics of earthquakes in North China Seismic Zone, and are suitable for describing the temporal distribution of M ≥ 6.5 earthquakes in North China Seismic Zone.

Table 2
Regression parameters and statistical test values of seismicity models in North China Seismic Zone

Model	Model parameters		-ln*L*	AIC	R^2	K-S test D_N value
Exponential distribution	μ = 13.2 (9.8–18.9)	-	128.9	259.9	0.978	0.088
Weibull distribution	α = 13.4	β = 1.0	128.9	261.8	0.974	0.094
Lognormal distribution	μ = 2.1	σ = 1.0	126.7	257.4	0.995	0.053
Gamma distribution	k = 1.2	λ = 11.4	128.7	261.4	0.967	0.100
Brownian passage time	μ = 13.2	α = 1.29	126.9	257.8	0.992	0.070

The maximum likelihood method was used to estimate the values of the model parameters. The limited length of seismic event samples leads to the uncertainty of the statistical parameters. In actual seismic hazard analysis, seismologists pay much attention to the range of such uncertainties, which are often estimated based on normal distribution. In real cases, however, the model parameters may not follow normal distribution. In this paper, we calculate the uncertainty range of parameters of BPT and Lognormal distribution models according to the likelihood surface with the model parameters and the method of likelihood profiles proposed by Biasi et al. (2015).

The confidence interval of a single parameter in the BPT and Lognormal models can be obtained by plotting a profile on the likelihood surface (Biasi et al., 2015). On the likelihood surface, a profile running parallel to the vertical and horizontal axes and passing through the point of the maximum likelihood value is plotted, so as to produce the curve of the likelihood values corresponding to single parameters. Using the confidence interval calculation method introduced by Biasi et al. (2015), the range of each parameter in a certain confidence interval can be obtained. Figure 3 presents the likelihood profile plots of parameters μ and α in the BPT model. Above the gray dotted line is the 95% confidence interval of parameters. The mean and 95% confidence interval of recurrence interval μ is 13.2 (8.9–19.1) years, and the mean and 95% confidence interval of α is 1.29 (1.0-1.78). Figure 4 shows the likelihood profile plots of parameters v and σ of the Lognormal model, where the mean value and 95% confidence interval of v is 2.09 (1.68–2.49), the mean value exp (v) corresponding to earthquake recurrence interval is 8.1 (5.4–12.1) years, and the mean value and 95% confidence interval of σ is 1.01 (0.78–1.38). For the Lognormal model, the mean value and expectation of earthquake recurrence are different; the expectation value is $E=\exp\left(\nu+\sigma^2/2\right)$ = 13.5 years, which is similar to the mean value of the earthquake recurrence interval calculated based on the BPT model.

4.2 Probability of occurrence of strong earthquakes

The probability of strong earthquakes, which often cause serious casualties and damage to buildings, has been an important source of concern among scientists. In the above, we found that both the BPT and Lognormal distributions can describe the temporal distribution characteristics of M ≥ 6.5 earthquakes in North China relatively well. In this study, we calculated the future occurrence probability of M ≥ 6.5 events in the study area based on the above two distribution models. For comparison, we also calculated the occurrence probability based on the Poisson model.

The probability of earthquake occurrence in the future can be calculated when the elapsed time and recurrence intervals of large earthquakes and their uncertainties are known. Let T_e be the time before the last event occurred, that is, the elapsed time, then the conditional probability of at least one earthquake occurring in the future ΔT time is (Matthews et al., 2002):

$$P\left(\Delta T|T_e\right) = \frac{\int_{T}^{T_e+\Delta T} f\left(t\right)dt}{\int_{T_e}^{\infty} f\left(t\right)dt} = \frac{F\left(T_e+\Delta T\right)-F\left(T_e\right)}{1-F\left(T_e\right)}$$

6

where $F\left(T_e\right)=\int_0^{T_e} f\left(t\right)dt$ is the cumulative distribution function of recurrence interval. $f\left(t\right)$ is the probability density function of the seismicity temporal distribution model introduced above.

As can be seen from the parameter values of each model in Table 2, the occurrence probability of M ≥ 6.5 events in North China in the next 1-100 years is calculated by using Eq. (6) (Fig. 5). The probability of earthquake occurrence calculated based on different models is different (Table 3), with the probability calculated based on the exponential distribution model being the largest and the probability calculated based on BPT model being the smallest. Among them, the probability of M ≥ 6.5 events in North China in the next 10 years calculated based on the BPT model is 35.3%, and the 95% confidence interval is 26.8%-44.9%. The average probability and 95% confidence interval calculated based on the Lognormal distribution is 35.4% (22.9%-49.7%), and the average probability and 95% confidence interval calculated based on the Poisson model is 53.1% (41.1%-64%). It is clear the occurrence probability values based on the BPT and Lognormal models are similar, and significantly smaller than that based on the Poisson model.

Table 3
Occurrence probability of M ≥ 6.5 earthquakes (mean value and 95% confidence interval) in North China in the future

Model	Probability of M ≥ 6.5 earthquakes in the future			
	1 year	5 years	10 years	50 years
Exponential	0.073 (0.052–0.097)	0.315 (0.233-0.400)	0.531 (0.411–0.640)	0.977 (0.929–0.994)
Lognormal	0.045 (0.027–0.070)	0.201 (0.125–0.297)	0.354 (0.229–0.497)	0.835 (0.653–0.943)
BPT	0.044 (0.032–0.059)	0.199 (0.148–0.261)	0.353 (0.268–0.449)	0.858 (0.741–0.936)

We also calculated the conditional probabilities in the next 1 and 10 years for different elapsed times. Figure 6 presents the variation of conditional probability with elapsed time. For the 1-year conditional probability, the calculated values based on the BPT and Lognormal models show the following characteristics: When the elapsed time is short, the calculated occurrence probability is smaller than that of the Poisson model; with the increase of elapsed time, the calculated occurrence probability begins to exceed that of the Poisson model; when the elapsed time is relatively long (relative to the recurrence interval), the calculated occurrence probability is again smaller than that of the Poisson model. For the 10-year conditional probability, when the elapsed time of the earthquake is short, the occurrence probability calculated based on the BPT and Lognormal models is greater than that of the Poisson model, and when the elapsed time of the earthquake is greater than the recurrence interval, the earthquake occurrence probability calculated based on the two models is less than that of Poisson model. The black vertical dashed line in Fig. 6 is the current time. It can be seen that the occurrence probability calculated based on the BPT and Lognormal models is about half of the value calculated based on the Poisson model.

5. Conclusions And Discussions

The abundant records of historical strong earthquakes in North China make it possible to investigate the corresponding temporal distribution models. This study provides a reliable and effective method for determining the optimal models for the recurrence intervals of M ≥ 6.5 earthquakes in North China. Using exponential, Gamma, Weibull, Lognormal and BPT models as target models, we regressed the model parameters by the maximum likelihood method, and evaluated the goodness-of-fit using the K-S test, the Chi-square test, and AIC and BIC values. The results show that the BPT and Lognormal models outperform other models in describing the temporal distribution of M ≥ 6.5 earthquakes in North China. Regarding the uncertainty of earthquake recurrence intervals, we used the likelihood profile method introduced by Biasi et al. (2015) to calculate parameter uncertainties of the BPT and Lognormal models. For the BPT model, the calculated recurrence interval and 95% confidence interval is 13.2 (8.9–19.1) years. For the Lognormal distribution model, the mean value and 95% confidence interval of v is 2.09 (1.68–2.49), and the mean value exp (v) corresponding to the recurrence interval is 8.1 (5.4–12.1) years. For the Lognormal model, the mean value and expectation of earthquake recurrence are different, and the expectation value is $E=\exp\left(\nu+\sigma^2/2\right)$ = 13.5 years, which is similar to the mean value of the recurrence intervals of the BPT model.

Based on the BPT, Lognormal and Poisson models, we also calculated the probability of future M ≥ 6.5 events in North China: the probability and 95% confidence interval in the next 10 years based on the BPT model is 35.3% (26.8%-44.9%); the mean value and 95% confidence interval of earthquake occurrence probability based on the Lognormal distribution is 35.4% (22.9%-49.7%); the mean probability and 95% confidence interval based on the Poisson model is 53.1% (41.1%-64%). The BPT and Lognormal models yield similar probability values, while that based on the Poisson model is significantly greater.

This study shows that M ≥ 6.5 events in North China conform better to the time-dependent BPT and Lognormal statistical models. Normally, for the time-dependent model, the calculated occurrence probability should be greater than that of Poisson model in case of a long elapsed time (relative to the return period). However, the occurrence rate of earthquakes calculated based on the time-dependent models in this study is lower than that calculated by the Poisson model, which is associated with the uncertainty of recurrence intervals. The coefficient of variation of earthquake recurrence intervals calculated based the on BPT model is 1.29 (1.0-1.78). For the BPT model, Matthews et al. (2002) showed that the earthquake sequence is dominated by noise when α is greater than $2^{-1/2}$. When the elapsed time is relatively long (relative to the earthquake recurrence interval), the waiting time for the next event based on the BPT model will exceed that of the Poisson model, i.e., the occurrence probability of the former will be smaller than that of the latter.

The results of this study may serve as important reference for seismic hazard analysis and earthquake prediction in North China.

Declarations

Data Availability Statement

The China earthquake catalogs used in this study were obtained from the China Earthquake Networks Center (CENC) (http://www.csndmc.ac .cn/wdc4seis@bj/intro/cenc/intro.jsp, last accessed December 2021) and the China Seismograph Network (CSN) (http:// www.csndmc.ac.cn/wdc4seis@bj/earthquakes/csn_catalog _p001.jsp, last accessed December 2021).

Acknowledgments

The authors thank the editor and anonymous reviewers for their valuable comments.

Funding

This study was supported in part by the Special Fund of Institute of Geophysics, China Earthquake Administration (Grant Number DQJB22Z03).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Xu Weijin is responsible for the idea and the writing of this manuscript. Wu Jian is responsible for the data analysis of this study. Gao Mengtan provides theoretical guidance for this manuscript.

References

1. Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. Physical Review E, 69(6), 066106. https://doi.org/10.1103/physreve.69.066106
2. Bajaj S. and M. L. Sharma. (2019). Modeling Earthquake Recurrence in the Himalayan Seismic Belt Using Time-Dependent Stochastic Models: Implications for Future Seismic Hazards. Pure Appl. Geophys. 176, 5261–5278
3. Ben-Naim E., E. G. Daub and P. A. Johnson. 2013. Recurrence Statistics of Great Earthquakes. Geophysical Research Letters, VOL. 40, 3021–3025, doi:10.1002/grl.50605, 2013
4. Beroza Gregory C. (2012). How many great earthquakes should we expect? PNAS, vol. 109, no. 3, 651–652
5. Biasi Glenn P., Robert M. Langridge, Kelvin R. Berryman, Kate J. Clark, and Ursula A. Cochran. (2015). Maximum-Likelihood Recurrence Parameters and Conditional Probability of a Ground-Rupturing Earthquake on the Southern Alpine Fault, South Island, New Zealand. Bull. Seismol. Soc. Am. 105(1): 94–106.
6. Bufe Charles G. and Perkins David M. 2005, Evidence for a Global Seismic-Moment Release Sequence. Bulletin of the Seismological Society of America, 95(3): 833–843
7. Bufe, C. G. and Perkins, D. (2011), The 2011 Tohoku earthquake: Resumption of temporal clustering of Earth's megaquakes. *Seismol. Res. Lett.*, 82, 455.
8. Cheng, J., Rong, Y., Magistrale, H., Chen, G. and Xu, X., 2020. Earthquake rupture scaling relations for mainland China. Seismological Research Letters, 91(1), pp.248-261.
9. Console, R., M. Murru, and A. M. Lombardi. 2003. Refining earthquake clustering models. *J. Geophys. Res.*, 108(B10), 2468, doi:10.1029/ 2002JB002130.
10. Cornell, C. A. (1968), Engineering seismic risk analysis, Bull. Seismol. Soc.Am., 58, 1583–1606.
11. Daub Eric G., Ben-Naim Eli, Guyer Robert A., and Johnson Paul A., 2012, Are megaquakes clustered? Geophys Res Lett, 39: L06308, doi:10.1029/2012GL051465.

12. Department of Earthquake Disaster Prevention, China Earthquake Administration. (1995). The Catalogue of Chinese Historical Strong Earthquakes (23th century BC—1911 AD) (in Chinese). Beijing: Seismological Press.
13. Department of Earthquake Disaster Prevention, China Earthquake Administration. (1999). The Catalogue of Modern Earthquakes in China (A. D. 1912~A. D. 1990, MS≥4.7) (in Chinese). Beijing: China Science and Technology Publishing House
14. Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics, Phil. Trans. Roy. Soc. Lond. 222, 309–368.
15. Frankel, A. (1995), Mapping seismic hazard in the central and eastern United States. *Seismological Research Letters*, 66(4), 8–21.
16. Frankel, A. D., et al. (2002), Documentation for the 2002 update of the National Seismic Hazard Maps, U.S. Geol. Surv. Open-File Rep., 2002-420.
17. Gao, M T. (1996), Probability model of earthquake intensity based on Poisson distribution. *Earthquake Research in China*, 12(2), 195-201.
18. Gao, M. T. (2003), New national seismic zoning map of China. *Atca Seismologica Sinica*, 25(6), 630-636.
19. Gao, M. T. (2015), A handbook of GB18306–2015 Seismic Ground Motion Parameter Zonation Map of China. *Beijing: China Standard Press.*
20. Gardner, J. K. and L. Knopoff. 1974. Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? *BSSA*, 64(5):1363-1367.
21. Geist E. L. and T. Parsons. (2011). Assessing historical rate changes in global tsunami occurrence. Geophysical Journal International, 187(1): 497–509.
22. Gibbons, J. D. and Chakraborti, S. (2003), Non-parametric Statistical Inference, 4th ed., 645 pp., CRC Press, Boca Raton, Fla.
23. Hu Y X. 1990. Comprehensive probability method in seismic hazard analysis. Beijing: Seismological Press, 1-8.
24. Huang, W. Q., Li, W. X., Gao, A. J. (1994), Analysis on integrity of China historical earthquake catalogue-take North China for example. *Acta Seismologica Sinica*, 16(3), 273-280.(in Chinese)
25. Kulkarni Ram, Wong Ivan, Zachariasen Judith, Goldfinger Chris, and Lawrence Martin. 2013, Statistical Analyses of Great Earthquake Recurrence along the Cascadia Subduction Zone. Bulletin of the Seismological Society of America, 103(6): 3205–3221.
26. LI Zheng-Fang, ZHOU Ben-Gang, RAN Hong-Liu. 2012. Strong earthquake risk assessment of eastern segment on the East Kunlun fault in the next 100 years based on paleo-earthquake data[J]. Chinese Journal of Geophysics (in Chinese), 2012, 55(09): 3051-3065, doi: 10.6038/j.issn.0001-5733.2012.09.023
27. 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.
28. Lv Y. J., Wang J., Wang S. Y., Peng Y. J., Xie Z. J., Gao A. J., Liu D. Y., 2010. Seismic catalogs for seismic ground motion parameters zonation map of China. Earthquake Press, Beijing, China, 2010.
29. Matthews MV, Ellsworth WL, Reasenberg PA. 2002. A Brownian model for recurrent earthquakes. Bull Seism Soc Am. 92(6):2233–2250.
30. Mega Mirko S., Allegrini Paolo, Grigolini Paoloet al. 2003. Power-Law Time Distribution of Large Earthquakes. Phys. Rev. Lett. 90(18): 188501-1-4
31. Michael AJ,2011, Random variability explains apparent global clustering of large earthquakes. Geophys Res Lett 38:L21301.
32. Nishenko, S. P., & Buland, R. (1987). A generic recurrence interval distribution for earthquake forecasting. Bulletin of the Seismological Society of America, 77(4), 1382–1399.
33. Ogata, Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.*, 83, 9–27.
34. Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, Ann. Inst. Stat. Math. 50, no. 2, 379–402.
35. Ogata, Y. and K. Abe. 1991. Some statistical features of the long-term variation of the global and regional seismic activity. *Int. Stat. Rev.*, 59, 139–161.
36. Ogata, Y., and J. Zhuang (2006). Space-time ETAS models and an improved extension, Tectonophysics, 413, nos. 1/2, 13–23, doi:10.1016/j.tecto.2005.10.016.
37. Parsons, T., and Velasco, A. A. (2011), Absence of remotely triggered large earthquakes beyond the main shock region. *Nat. Geosci.*, 4, 312–316. doi 10.1038/ngeo1110.
38. Parsons Tom and Geist Eric L. 2012. Were Global M≥8:3 Earthquake Time Intervals Random between 1900 and 2011? Bulletin of the Seismological Society of America, 102(4):1583–1592.
39. Pasari S, Dikshit O. 2015a. Distribution of earthquake interevent times in northeast India and adjoining regions. Pure Appl Geophys. 172:2533–2544.
40. Pasari, S., and O. Dikshit (2018). Stochastic earthquake interevent time modelling from exponentiated Weibull distributions, Nat. Hazards 90, 823–842.
41. Pasari, S. (2019). Inverse Gaussian versus lognormal distribution in earthquake forecasting: Keys and clues, J. Seismol. 23, 537–559.
42. Pei W, Zhou S, Zhuang J, Xiong Z, Piao J. 2021. Application and discussion of statistical seismology in probabilistic seismic hazard assessment studies. Science China Earth Sciences, https://doi.org/10.1007/s11430-021-9824-0
43. Petersen Mark D., Morgan P. Moschetti, Peter M. Powers et al. 2014. Documentation for the 2014 Update of the United States National Seismic Hazard Maps. USGS Open-File Report 2014–1091.
44. RAN Hong-Liu, HE Hong-Lin. Research on the magnitude and recurrence interval of characterized earthquakes with Magnitude≥6.7 along the northwestern portion of the Xianshuihe fault zone in western Sichuan, China[J]. Chinese Journal of Geophysics (in Chinese), 2006, 49(1): 153-161.

45. Salditch Leah, Stein Seth, Neely James et al., 2020, Earthquake supercycles and Long-Term Fault Memory. Tectonophysics, Tectonophysics 774 (2020) 228289.
46. Schwartz, D. P., and K. J. Coppersmith, 1984, Fault Behavior and Characteristic Earthquakes: Examples From the Wasatch and San Andreas Fault Zones. JGR, 89(B7): 5681-5698.
47. Shearer Peter M. and Stark Philip B. 2012. Global risk of big earthquakes has not recently increased. PNAS, 109(3): 717–721.
48. Šipčić, N., M. Kohrangi, A. N. Papadopoulos, W. Marzocchi, and P. Bazzurro (2022). The Effect of Seismic Sequences in Probabilistic Seismic Hazard Analysis, Bull. Seismol. Soc. Am. XX, 1–16, doi: 10.1785/0120210208
49. Tripathi JN. 2006. Probabilistic assessment of earthquake recurrence in the January 26, 2001 earthquake region of Gujarat, India. J Seismol. 10:119–130.
50. Utsu T. 1984. Estimation of parameters for recurrence models of earthquakes. Bull Earthq Res Inst Univ Tokyo.59:53–66.
51. Utsu, T., 2002: Statistical features of seismicity. In: Lee, W. H. K., H. Kanamori, P. C. Jennings, and C. Kisslinger, (Eds.), International Handbook of Earthquake & Engineering Seismology, 81A, Academic Press, 719-732.
52. Xu W. J. & Gao M. T. (2014). Statistical analysis of the completeness of earthquake catalogues in China Mainland. *Chinese J. Geophys*, 57(9), 2802-2812.
53. Xu W J, Wu J. 2017. Effect of temporal – spatial clustering of aftershock on the analysis of probabilistic seismic hazard. Chinese J. Geophys. (in Chinese), 60(8): 3110-3118.
54. Zaliapin, I., and Y. Ben-Zion (2013), Earthquake clusters in southern California I: Identification and stability, J. Geophys. Res. Solid Earth, 118, 2847–2864, doi:10.1002/jgrb.50179.
55. Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association, 97(458), 369–380. https://doi.org/10.1198/016214502760046925
56. Zhuang, J., C.-P. Chang, Y. Ogata, and Y.-I. Chen (2005), A study on the background and clustering seismicity in the Taiwan region by using point process models,J. Geophys. Res.,110, B05S18, doi:10.1029/2004JB003157.

Figures

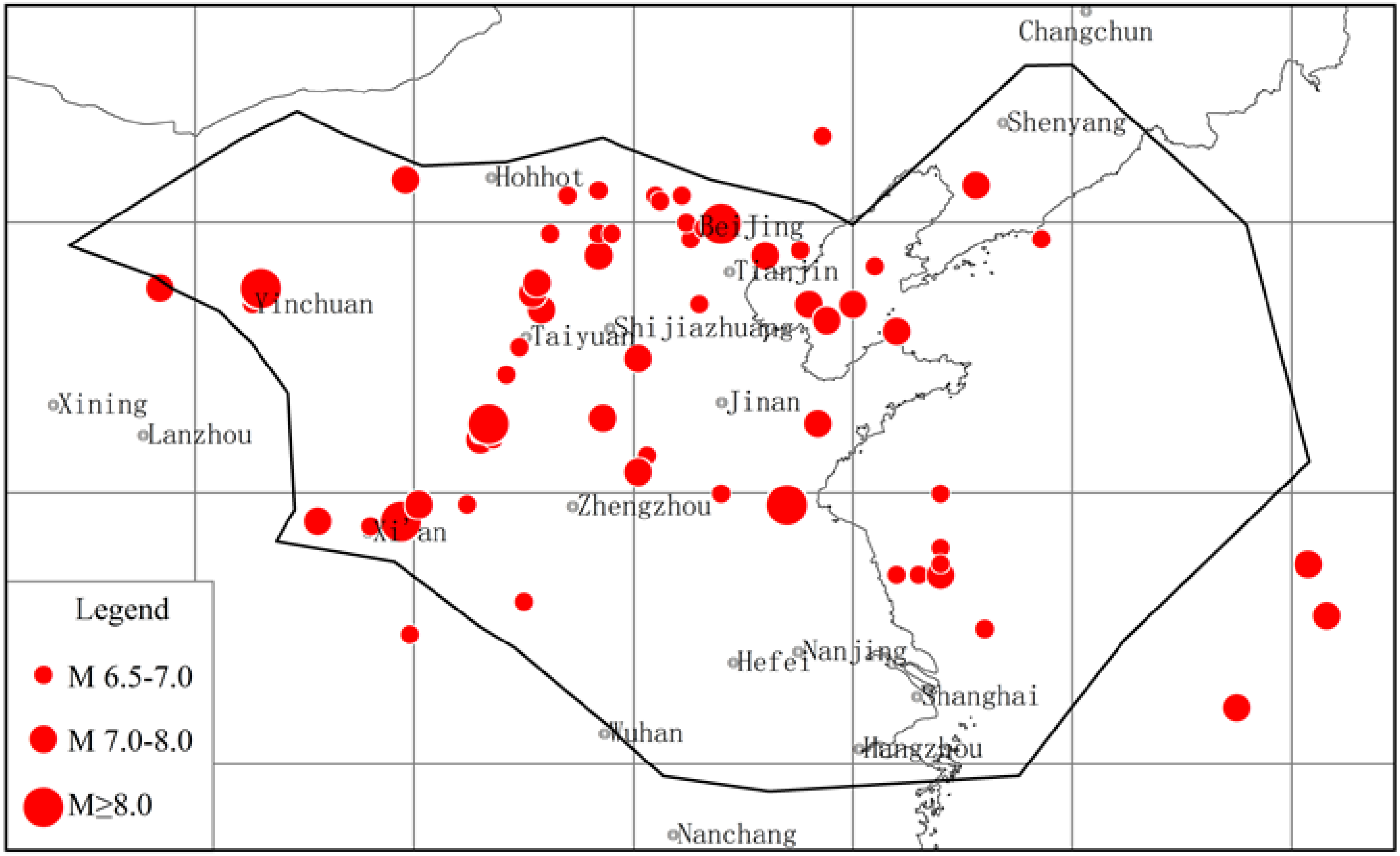

Figure 1

Map of earthquake epicenter distribution in North China seismic region, the back polygon area is the North China seismic region

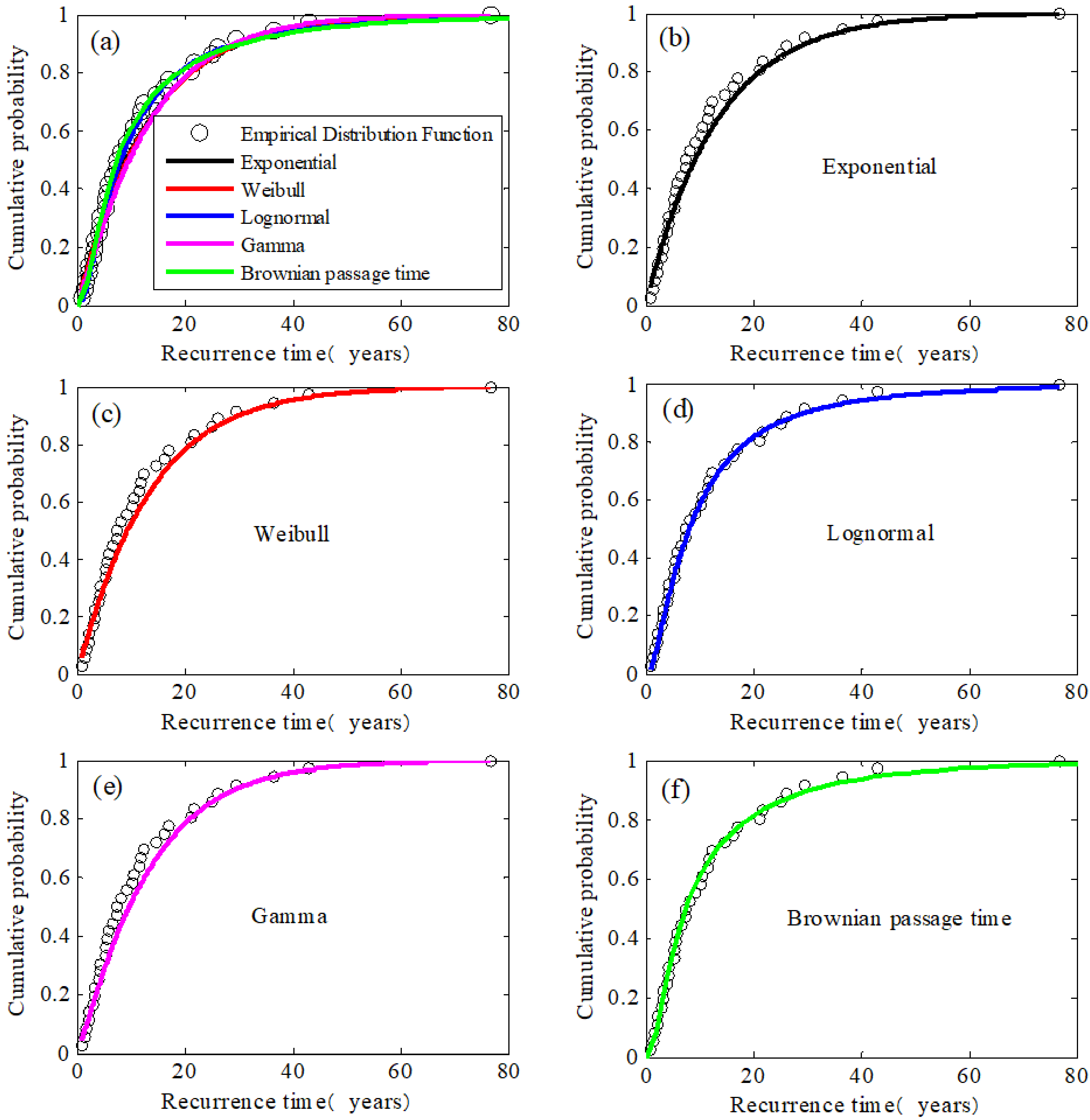

Figure 2

Empirical cumulative distribution of the recurrence intervals of strong earthquakes in North China Seismic Zone and the corresponding cumulative distribution functions of the five models

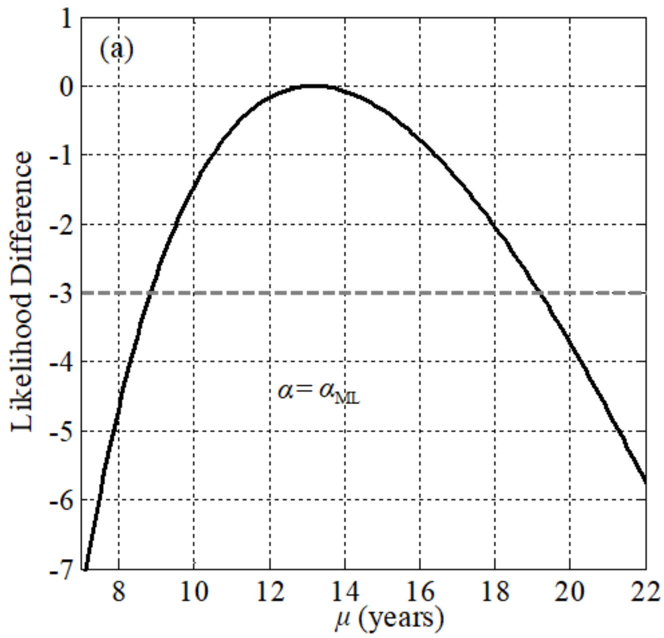

Figure 3

Variation of likelihood values with parameters in the BPT model. (a) is the curve of likelihood function values with μ when parameter $\alpha=\alpha_{ML}$, and (b) is the curve of likelihood function values of the BPT model with different α values when $\mu=\mu_{ML}$

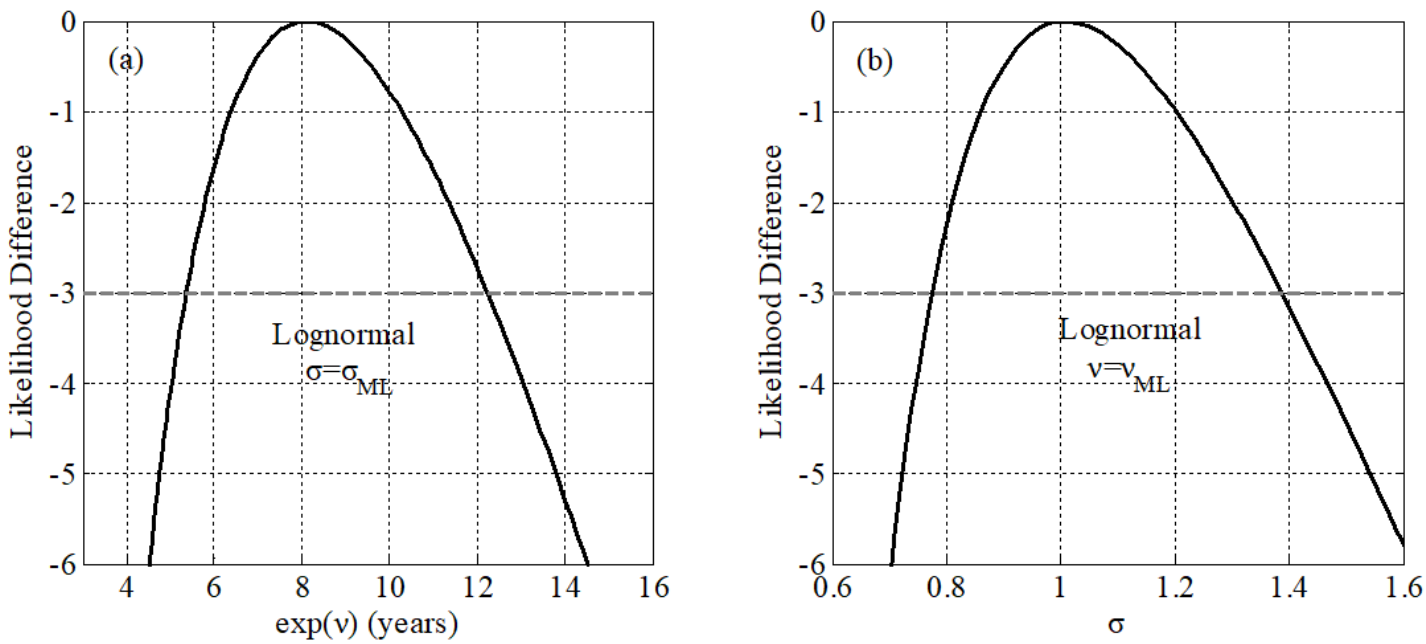

Figure 4

Variation of likelihood values of the Lognormal model with parameters. (a) shows the variation of likelihood function values with exp (v) when parameter $\sigma=\sigma_{ML}$, and (b) shows the variation of likelihood function values with different σ values when $v=v_{ML}$

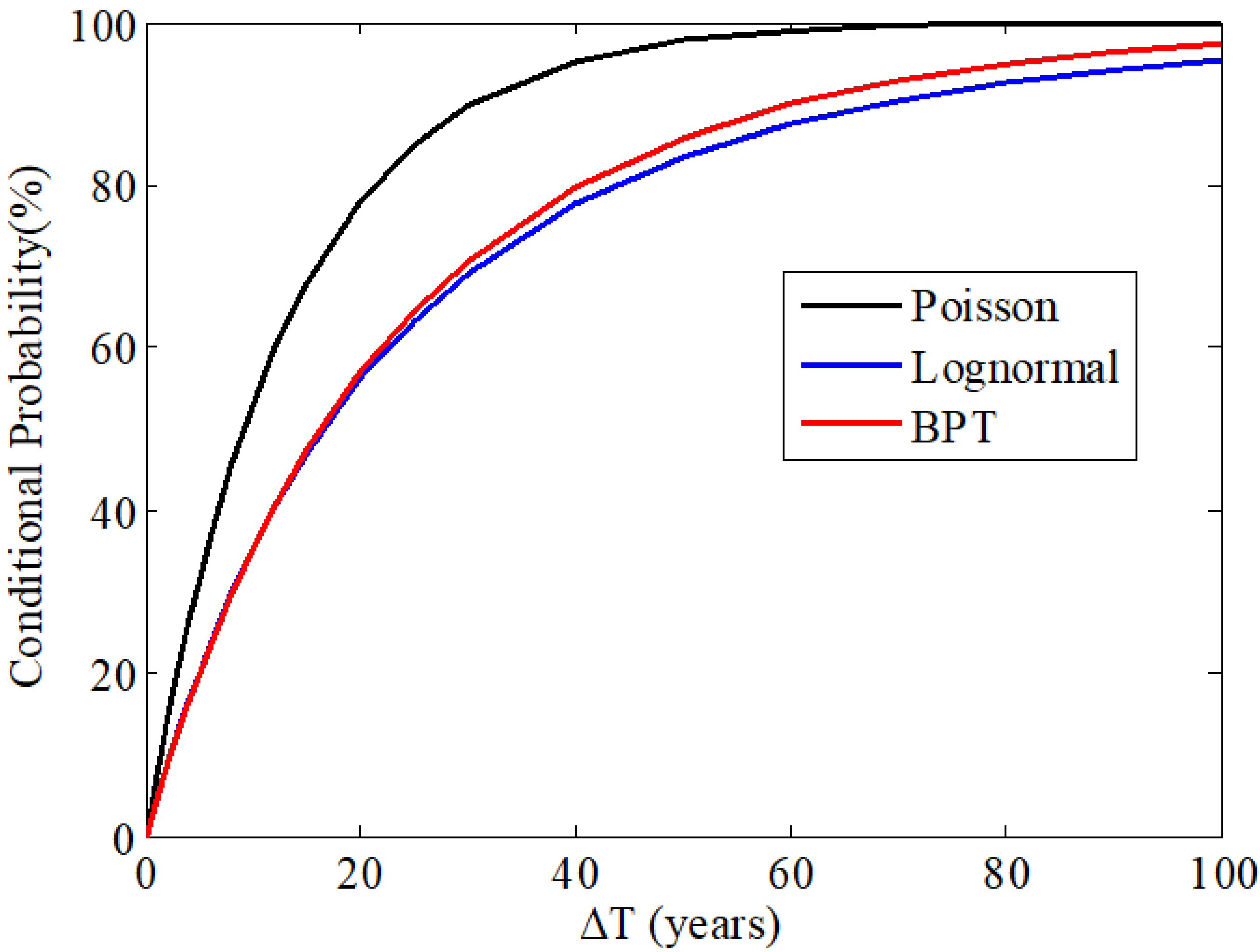

Figure 5

Variation of occurrence probability of M≥6.5 earthquakes in North China with time

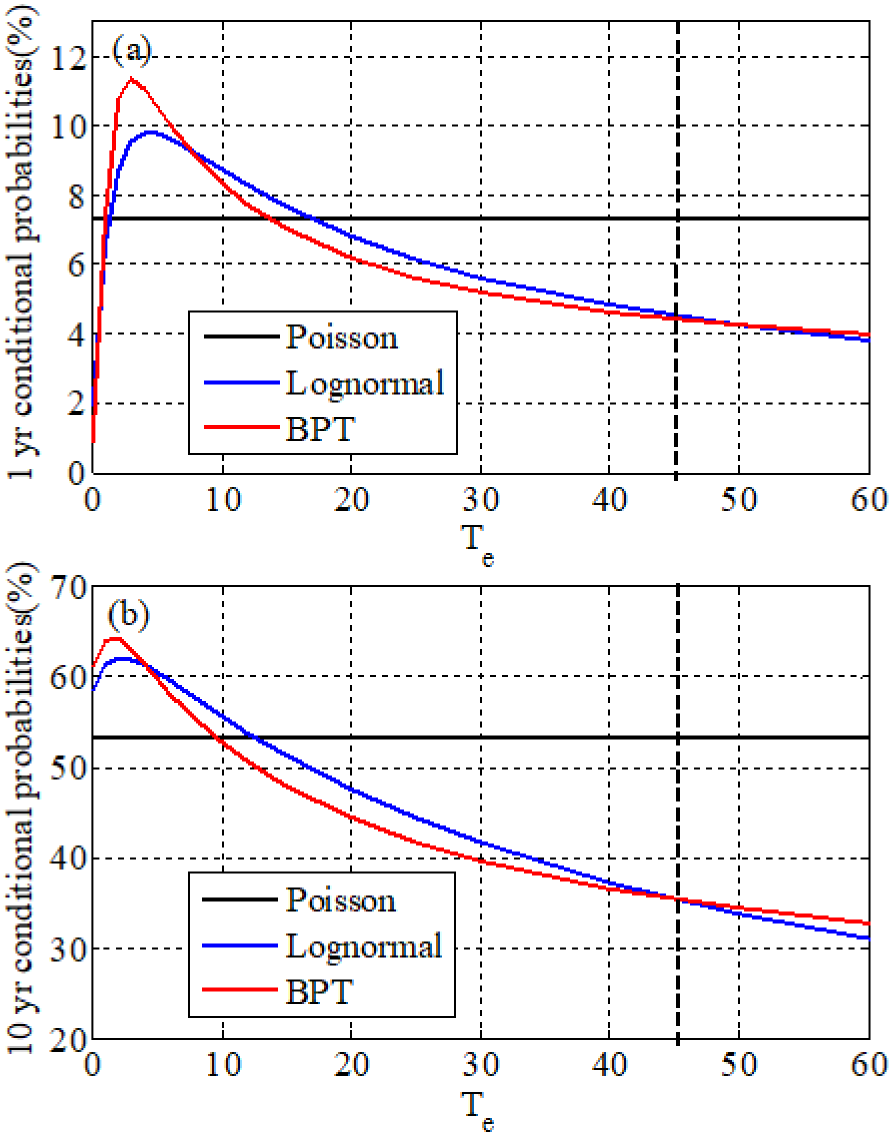

Figure 6

(a) Variation of conditional probability of earthquake occurrence with elapsed time in the next 1 year based on Poisson, BPT and Lognormal models, and (b) variation of conditional probability of earthquake occurrence with elapse timed in the next 10 years. The thick black vertical dashed line represents the current time.