Earthquake potential in and around China: Estimated from past earthquakes

Yufang Rong and David D. Jackson

Department of Earth and Space Sciences, University of California, Los Angeles, California, USA

Received 9 April 2002; revised 22 May 2002; accepted 3 June 2002; published 22 August 2002.

[1] We present an earthquake potential estimate for magnitude 5.4 or larger earthquakes for China. The potential is expressed as the rate density, or the probability per unit area, magnitude and time, and it is based on smoothed seismicity. We constructed a special earthquake catalog, which combines previous catalogs covering different times. After deriving regression relationships, we estimated moment magnitudes for some events. Then we used the special catalog for constructing our smoothed seismicity model and testing it retrospectively. Using maximum-likelihood we estimated the completeness threshold, the b-value, and the corner magnitude. We adopted a kind of Gutenberg-Richter distribution with modifications at higher value and upper magnitude limit derived from the special catalog. From smoothed seismicity, we estimated local "a-values". We have begun a "prospective" test, finding that earthquakes since the beginning of 2000 are quite compatible with the model. INDEX TERMS: 7223 Seismology: Seismic hazard assessment and prediction; 7215 Seismology: Earthquake parameters; 7230 Seismology: Seismicity and seismotectonics

1. Introduction

[2] In this paper we describe a long-term earthquake potential estimate for magnitude 5.4 or larger earthquakes in China. The earthquake potential is expressed as the probability of an earthquake occurring within a specified time, place and magnitude window.

[3] We assume that earthquake potential is proportional to a smoothed version of past seismicity. The smoothed seismicity method has been successfully applied to Southern California [Jackson et al., 1995], the northwest and southwest Pacific [Jackson and Kagan, 1999; Kagan and Jackson, 2000] and Italy [Console and Murr, 2001], with test results showing the forecasts good fit to actual earthquakes. Here the earthquake potential is estimated based on seismicity from 1950 to 1999. Because our forecast is formally stated, rigorous statistical tests can be applied to it, and we have tested the forecasting model against earthquakes occurring in 2000 and 2001.

[4] We assume the magnitude distribution follows the modified Gutenberg-Richter (G-R) relation described in terms of scalar seismic moment M [Jackson and Kagan, 1999],

\[ G(M) = (M_c/M)^b \exp[(M - M_c)/M_c] \]  \hspace{1cm} (1)

\[ m = \frac{2}{3} \log_{10} M - 10.7 \]  \hspace{1cm} (2)

[Jackson and Kanamori, 1979]. The lower threshold magnitude \( m_l \) and the corner magnitude \( m_c \) are then defined as the magnitudes obtained by converting \( M_c \) and \( M_c \), respectively, using (2). For magnitudes smaller than \( m_l \), the moment distribution (1) implies a magnitude distribution very close to a standard Gutenberg-Richter distribution with \( b = 1.5 \). \( b \).

[5] Authors of the works cited above applied the smoothed seismicity model to plate boundary regions. By contrast, this work extends the method to intra-plate regions, but our ultimate goal is to provide more information about future earthquakes for risk reduction.

2. Data

[6] Since we model the rate-density of future earthquakes as proportional to a smoothed version of past seismicity, the smaller the completeness magnitude and the longer the recording of the catalogs, the better the constructed model.

[7] Three catalogs are used here to construct the model: two local Chinese ones and the Harvard CMT catalog. The first local catalog covers the interval of January 1978 through December 1998 [http://www.eq.igp.ac.cn/bndesn/index.html]. Various magnitudes may be given: \( M_t \) (surface-wave magnitude); \( M_{7} \) (surface-wave magnitude from long period seismographs commonly used in China); \( M_L \) (local magnitude); \( m_B \) (Broad-band body-wave magnitude); \( m_s \) (body-wave magnitude); \( M_{7} \) (surface-wave magnitude given by the National Earthquake Information Service (NEIS)); and \( M_b \) (body-wave magnitude given by NEIS). The second local catalog ([Peishan Chen, Division of Technology and Administration of Chinese Seismological and Geomagnetic Networking, Chinese Seismological Bureau, personal communication] covers two intervals: January 1900 through December 1977, and January 1999 through December 1999. The magnitudes are given by \( M_t \), \( M_{7} \), or/and \( m_B \). The Harvard CMT catalog begins in 1976 and provides moment tensor solutions for global earthquakes. We estimated moment magnitudes using (2). The

Copyright 2002 by the American Geophysical Union.
0094-8276/02/2002GL015297S05.00
completeness levels of all the catalogs vary with time and region. Kagan [1999] suggested that the Harvard CMT catalog is complete at $m_t = 5.8$ and $m_t = 5.6$ for earthquakes after 1977 and 1982, respectively. The local Chinese catalogs have more accurate locations, lower magnitude of completeness, and longer records, but the Harvard CMT catalog is believed to have more accurate magnitudes. To benefit from the merits of all three catalogs, we converted all other magnitudes to moment magnitudes. To benefit from the merits of all three Harvard CMT catalog is believed to have more accurate magnitude of completeness, and longer records, but the Chinese catalogs have more accurate locations, lower

### Table 1

<table>
<thead>
<tr>
<th>Magnitude scales (M)</th>
<th>Number of events</th>
<th>a (σm)</th>
<th>b (σm)</th>
<th>c (σm)</th>
<th>Standard deviation σm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_r$</td>
<td>328</td>
<td>0.205</td>
<td>0.020</td>
<td>-1.745</td>
<td>0.240</td>
</tr>
<tr>
<td>$M_L$</td>
<td>165</td>
<td>0.133</td>
<td>0.027</td>
<td>-0.740</td>
<td>0.306</td>
</tr>
<tr>
<td>$M_L$</td>
<td>147</td>
<td>0.208</td>
<td>0.078</td>
<td>-1.955</td>
<td>0.878</td>
</tr>
<tr>
<td>$m_B$</td>
<td>276</td>
<td>0.301</td>
<td>0.046</td>
<td>-2.579</td>
<td>0.542</td>
</tr>
<tr>
<td>$m_B$</td>
<td>212</td>
<td>0.484</td>
<td>0.092</td>
<td>-4.208</td>
<td>1.011</td>
</tr>
<tr>
<td>$m_B$</td>
<td>191</td>
<td>0.042</td>
<td>0.018</td>
<td>0.210</td>
<td>0.206</td>
</tr>
<tr>
<td>$m_B$</td>
<td>279</td>
<td>0.403</td>
<td>0.077</td>
<td>-3.439</td>
<td>0.860</td>
</tr>
</tbody>
</table>

The Taiwan area is not included in this study because Taiwan is located at a plate boundary where seismicity, folding, and thrusting are extremely active. However, for Mainland China and surrounding regions, continental deformation is a key in generating big earthquakes. Since our forecast model depends on smoothing technique, which differs for different tectonic areas, including the Taiwan region would make our smoothing model not applicable.

A starting time $t_s$ of good quality records and a minimum magnitude of complete recording, $m_t$, are two important parameters for studying the magnitude frequency distributions and estimating earthquake potentials. By plotting the annual numbers of earthquakes larger than 5.1 with time, we found a fairly clear division at year 1950. Such plots show that annual numbers of events before 1950 have a gradually increasing trend, and the average annual number before 1950 is clearly smaller than that after 1950. The yearly rates are quite stable after 1950. Therefore, we take $t_s = 1950$. To estimate $m_t$, we employ the method of Wiemer and Wyss [2000]. The estimation of $m_t$ is based on the assumption that, for a given volume, a simple power law can approximate the frequency-magnitude distribution. At a 90% confidence level, $m_t = 5.4$.

The β-value and the corner magnitude $m_c$ are two other important parameters for constructing a modified G-R magnitude distribution. Kagan [1999] estimated them for the global earthquake distribution in the Harvard catalog. For shallow global earthquakes, he found 95% confidence limits for β-value and corner magnitudes of 0.657 ± 0.017 (corresponding to $b = 0.986 ± 0.025$) and 8.1–8.7, respectively. The thick solid and dash-dotted curves in Figure 1 show the number of earthquakes larger than a given seismic moment and the number predicted from a modified G-R distribution with $β = 0.657$ and $m_c = 8.37$ (the value Kagan used in the graph in his paper), respectively. The thin dotted curves represent $2σ$ (σ is standard deviation) bounds of the modified G-R model by assuming a Poissonian process. The curves in Figure 1 demonstrate that earthquake size data are reasonably well approximated by the modified G-R distribution at lower magnitudes. The discrepancy at larger magnitudes is mainly caused by two large events: the August 15th, 1950, $m = 8.6$ Indo-China Border, and the December 4th, 1957, $m = 8.0$ Gobi-Altai earthquakes. The results for such discrepancy could be: 1) the magnitudes of past earthquakes are more uncertain due to different techniques and earlier instruments; and 2) the catalog is not long enough. Nonetheless, the likelihood score for the observed catalog falls within the 95% confidence interval for the modified G-R distribution.

### 3. Method and Results

Kagan and Jackson [1994, 2000] and Jackson and Kagan [1999] factored earthquake potential into three functions of location, size and focal mechanism. Here we use only location and size, because we do not know the focal mechanisms for many past earthquakes. Thus, earthquake potential is represented by

$$\Xi(x, m) = f(x) f_M(M), \quad (3)$$

where $\Xi$ is the rate of earthquakes per unit area, seismic moment and time, $x$ is the latitude and longitude, $M$ the scalar seismic moment, $f$ the spatial density function, and $f_M$ the moment distribution. When the focal mechanism data are completed for Chinese earthquakes, we will consider this factor. As in Kagan and Jackson [2000], the long-term earthquake rate $\Xi$ is considered here to be time-independent.
Thus, only the spatial kernel of the smoothing function needs to be optimized.

The spatial kernel with units of earthquakes per unit area and time is a function of $r$, the distance from a past earthquake [Kagan and Jackson, 2000]:

$$f_r = \frac{1}{\pi} \left( \frac{1}{r^2 + r_s^2} \right) \quad (4)$$

where $r_s$ is a scale parameter. Integrated over two spatial dimensions the cumulative kernel function is

$$F(R) = 2 \int_0^R \frac{rdr}{r^2 + r_s^2} = \log \left( 1 + \frac{R^2}{r_s^2} \right) \quad (5)$$

where $R$ is the truncation distance beyond which earthquakes are neglected. The two equations apply to all earthquakes of $m \geq 5.4$ or larger. The total spatial rate for a map point is calculated by summing contributions from all earthquakes closer than $R$ [equation (7) in Kagan and Jackson, 2000]. For some cells in the study region, no $m \geq 5.4$ earthquakes occurred during the period 1950–1999 within distance $R$. However, this does not mean that these zones are immune from earthquakes forever. Thus, we assume a spatially uniform background potential, which integrates to a somewhat arbitrary 1% of the total earthquake potential. Equation (5) has two adjustable parameters: $R$ and $r_s$. We use the smoothed bootstrap technique for choosing the optimal smoothing parameters [Silverman, 1986, p. 145; Kagan and Jackson, 2000]. Here we optimized the parameters $r_s$ and $R$ by past earthquakes at 30 and 300 km, respectively. The seismic moments are modeled to follow the modified G-R relation (1).

0.25° grid. Two contour lines bound the areas with the highest 50% and 95% earthquake potentials, respectively (thin black lines in Figure 2). From Jan. 2000 to Dec. 2001, 21 earthquakes with $m \geq 5.4$ occurred in the study area and are plotted as “beach-balls” in Figure 2. To test our model against these events, we sorted the cells from the highest potential to the lowest. In Figure 3, we plotted the cumulative potentials and the cumulative earthquakes count of the sorted cells. The dash-dotted smooth curve in Figure 3
demonstrates the relation between the fraction of total area and fraction of theoretical earthquake rate. The solid step curve demonstrates how the fractions of total area and total earthquakes relate. As shown by the smoothed curve, the model predicts that 50% of all earthquakes should occur within the “hottest” 9% of the area: that with the highest estimated rate. In fact, 9 out of 21 earthquakes did occur there. Similarly, the model predicts that 95% of events should fall within the “hottest” 46% of the area, as all 21 events did.

To evaluate the forecast, we employed two statistical tests: number and the likelihood [Kagan and Jackson, 1995; Jackson and Kagan, 1999]. In each case, we defined a statistic (a measurable feature of the earthquake catalog), simulated synthetic records for the hypotheses, computed the defined statistic for both the observed and simulated earthquake records, and compared the observed statistic with the simulated values. To test the consistency of a hypothesis, we asked if the observed catalog is consistent with those that satisfy the hypothesis by construction; otherwise we rejected the hypothesis as an explanation for observations. The first test, which we call the “N” test, is based on comparing the total number of earthquakes in the real catalog to those in the synthetic. The second test, which we call the “L” test, is based on comparing the logarithm of the likelihood of a particular set of areas with earthquakes falling into those synthetic catalogs [equation (1) in Kagan and Jackson, 1995; equation (4) in Jackson and Kagan, 1999]. The result of the “N” test shows that in the test period 27.8% of the simulations have fewer than 21 earthquakes, and 64.3% have more than 21 earthquakes. According to the two-tailed test rule, we cannot reject the forecast at the 95% confidence level. The results of the “L” test show that about 44.2% of the simulations have a log likelihood score lower than the CMT catalog. Again, we cannot reject the forecast at the 95% confidence level.

4. Discussion and Conclusions

Our earthquake potential map is for earthquakes with $m \geq 5.4$. For larger magnitudes, we can obtain the potentials from the modified G-R distribution (1) by assuming a uniform $b$-value and uniform magnitude limit, both derived from our special catalog. The smoothed seismicity method is based on two assumptions: earthquakes are prone to occur in a seismically active area and earthquake potential is time-independent. In the future, we will introduce time-dependent factors.

Estimating seismic potential from smoothed seismicity has several advantages over other methods [Kagan and Jackson, 2000]. For example, all the needed data are readily available, and the model can be implemented and tested without subjective judgments about fault geometry, etc. As already mentioned, this procedure has been applied successfully to many inter-plate regions. Can the method be successful in intra-plate environments like our study area?

Results to date suggest that the smoothed seismicity model does work in continental as well as plate-boundary regions. However, the most definitive test of any earthquake potential model is prospective such that probabilities are fully specified before the test. The tests reported above were “pseudo-prospective” in which the methodology and the data used to estimate parameters predated the test period, but the probabilities themselves were calculated after the test period began. While the total number and locations of the events were consistent with the model, a true validation awaits more earthquakes. We plan to account for new earthquakes each year and test the updated model in a fully prospective mode. Our experience [e.g., Jackson and Kagan, 1999; Kagan and Jackson, 1994] suggests that a model can be tested adequately only after a time for which the model predicts a few dozen earthquakes. For magnitude 5.4 and above this time will be just a few years. For larger events, say 6.4 and above, testing the model will take a few decades.

Our forecast model shows that 50% of total earthquake potential falls into only 9% of the most active area including Yunnan and Sichuan provinces, the China-Burma and China-India boundaries, and west of Xinjiang province. In addition, 95% of the total earthquake potential falls into less than 50% of the area, including Tibet, the Beijing area, Hebei provinces, and part of Mongolia, in addition to the areas above. However, most of northeast and southeast China are at little risk of big earthquakes, consistent with historical records.

[20] Acknowledgments. We thank Yan. Y. Kagan for very helpful discussions and comments. We would also thank Prof. Peishan Chen and Dr. Rufeng Liu of Institute of Geophysics, China Seismological Bureau for providing us earthquake catalogs. This research was supported by National Science Foundation grant EAR9725599 and Southern California Earthquake Center (SCEC; Purchase order 004876). SCEC is funded by NSF Cooperative Agreement EAR-8920136 and USGS Cooperative Agreements 1408-0001-A0899 and 1434-HQ-07AG01718. The SCEC contribution number for this paper is 689.

References


