Earthquake Engineering, Tenth World Conference © 1992 Balkema, Rotterdam. ISBN 90 5410 060 5

Non-stationarity of historical seismicity in China

R. K. McGuire & G.R.Toro
Risk Engineering, Inc., Golden, Colo., USA

D.Veneziano
Massachusetts Institute of Technology, Cambridge, Mass., USA

C.A.Cornell
Stanford University, Calif., USA

Y.X. Hu, Y.Jin, Z.Shi & M.Gao
State Seismological Bureau, Beijing, People's Republic of China

ABSTRACT: The Chinese catalog of historical seismicity contains multiple large earthquakes in identified earthquake sources, thereby offering the opportunity to study and model non-stationary aspects of seismicity, and to include these non-stationarities in seismic hazard calculations for decision-making regarding levels of earthquake-resistant design. We examine the Chinese catalog, using the most up-to-date interpretations of seismic zones, and fit a stress-release model to represent earthquake cycles and periods of higher and lower probability of occurrence. These models are then used in the predictive sense to calculate the seismic hazard from future occurrences. Results indicate that neglecting nonstationarity introduces moderate errors in the calculated hazard.

1 INTRODUCTION

The Chinese earthquake catalog is one of the longest in the world and offers opportunities to investigate the real effect of non-stationarities in seismicity in estimating earthquake hazards. This paper reports on some of the results of collaborative investigations conducted by the authors to determine (a) what non-stationarities and non-homogeneities are evident in the Chinese catalog, (b) what mathematical models can be fit to the catalog to represent these characteristics, and (c) what difference those models have on calculated seismic hazard. Ultimately the purpose is to judge how important is the modeling of non-stationary, non-homogeneous processes in the context of other uncertainties in the earthquake hazard problem that must be handled.

2 DATA

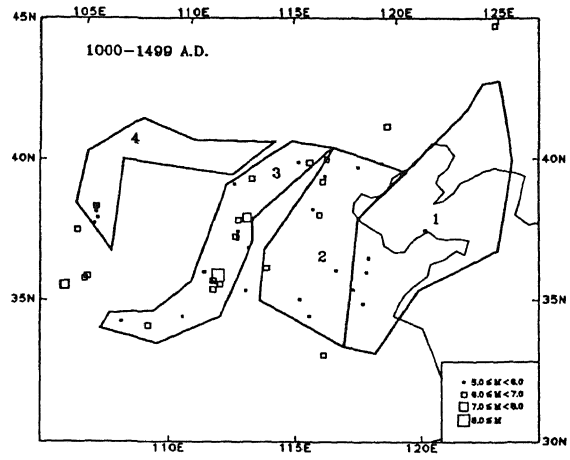
The main source of data for this study is the catalog of seismicity in North China developed by the State Seismological Bureau (SSB) in Beijing. This catalog contains dates, locations, and magnitudes for earthquakes dating from the year 780 B.C., which were obtained from historical documents, especially those maintained by local governments.

The catalog exhibits incompleteness for various regions, times, and magnitudes, as is discussed below. Figure 1 presents maps of seismicity for the periods 780 B.C. to 1499 A.D., and 1500 A.D. to 1987 A.D. Obvious aftershocks were removed from the catalog by informal analysis as well as a more formal mathematical technique that identified smaller events within a magnitude-dependent time and space window from a mainshock.

We divide the regions into four tectonic seismogenic provinces developed by the SSB for the purposes of seismic hazard mapping. These provinces have been further sub-divided into smaller seismic source zones, typically associated with faults or fault segments. Figure 2 shows a map of these provinces and seismic sources. Earthquakes not falling within the four main provinces are not considered in this study.

3 EXPLORATORY ANALYSIS

The purpose of the exploratory analysis is to identify the most important characteristics of the North China earthquake data. Based on observations and experience from other parts of the world, we investigate non-stationarity in the rate of occurrence and b-value, immunity (non-homogeneity in the spatial distribution following a large earthquake), and deviations from an exponential magnitude distribution.



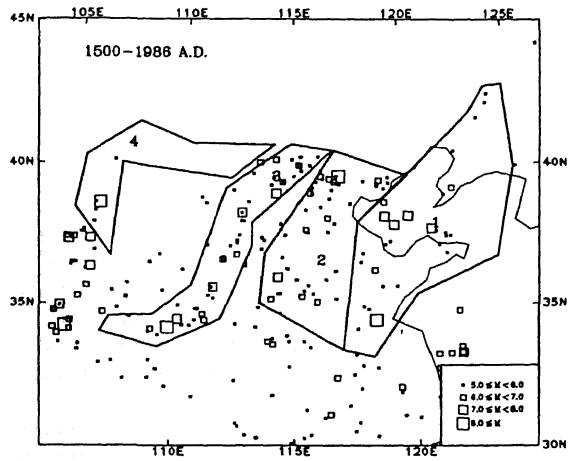


Figure 1. Map of earthquakes in North China and seismogenic provinces.

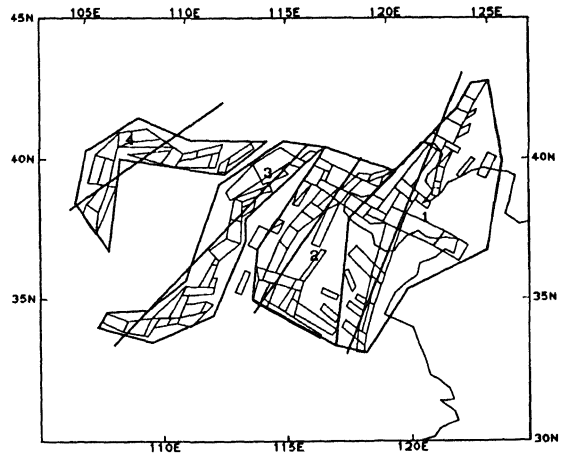


Figure 2. Map of seismogenic provinces and seismic sources.

3.1 Visual inspection: interaction and incompleteness

To investigate interaction and possible incompleteness of the catalog, earthquake events were plotted versus time and location, using as the location measure the distance along the province "axis" (which is also the trend of faulting in each of the provinces, except for Province 4; see Figure 2). Each earthquake was centered at the reported epicenter and a rupture length was estimated for that event. For large earthquakes (e.g. M>7.5) for which an isoseismal map was available, the rupture length was based on the longest length (which was generally parallel to the province axis) encompassed by the intensity VIII isoseismal; this method appears to give reasonable estimates for recent events. For smaller earthquakes the rupture length was based on the relation $RL=10^{-0.41+0.8M}$ reported by Bonilla et al (1984) for strike-slip earthquakes. For all earthquakes the rupture was assumed to be parallel to the axis of the province and to be centered at the epicenter, as defined in the catalog.

Figure3 presents these plots for the four provinces. Province 1 was further divided into 1A (south) and 1B (north), because that province is split by Po Hai Bay where seismicity is incomplete, and the history of the regions to the north and south are quite different. Based on this analysis the regions 1A, 2 and 3 were judged to be complete for M > 5 since 1500. This assumption is consistent with the completeness function derived by Lee and Brillinger (1979), and is similar to the assumptions of Zheng and Vere-Jones (1991) and Vere-Jones and Yonglu (1988). We therefore concentrate on these three provinces and on events since the year 1487 (this gives 500 years of earthquakes for analysis).

3.2 Categorical analysis

With the most complete provinces and data selected for analysis, a categorical analysis was conducted to detect and quantify the following:

1. global non-stationarity,
2. province-to-province variability in non-stationarity, which may be indicative of migration of seismicity among provinces, and
3. variations of b-value during different phases of the seismic cycle, which would be a useful diagnostic tool for other regions.

The procedure used for this analysis was to calculate observed earthquake counts in various location-time-magnitude "cells". Each location corresponds

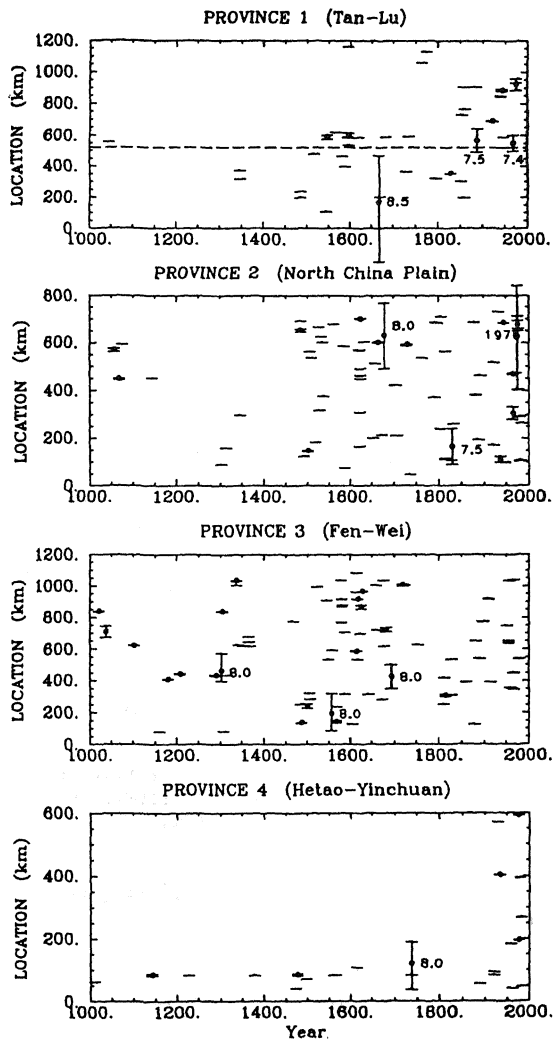


Figure 3. Time-location plots for earthquakes since the year 1000 A.D. The province axes run from SW to NE (see Figure 2). The dashed line in the top panel indicates the boundary between provinces 1a and 1b.

to a province (1a, 2, or 3); each time corresponds to a 100-yr time interval beginning in 1487; magnitude was partitioned into 5 to 6 and greater than 6. A non-parametric log-linear model (e.g. Bishop et al, 1974) was then fit to the cell counts to explain observations and to test the statistical significance of various effects . The form of the model used is as follows:

$$\ln[\lambda(X_i,T_j,M_k)]=\mu_0+\mu_{X_i}+\mu_{M_k}+\mu_{T_j}+\mu_{X_iT_j}+\mu_{X_iM_k}+\mu_{T_jM_k} \qquad (1)$$

where μ_0 is a constant parameter, μ_{X_i}, μ_{T_j}, and μ_{M_k} are parameters that represent the marginal dependencies on location, time, and magnitude. For instance, if the term μ_{T_j} is selected, the model estimates a value of μ_{T_j} for each time interval (under the constraint that these terms add to 0), in order to represent non-stationarity. Similarly, the second-order terms $\mu_{X_iT_j}$, $\mu_{X_iM_k}$, and $\mu_{T_jM_k}$ represent interaction. For instance, the terms $\mu_{X_iT_j}$ represent province-dependent non-stationarity. Not all terms are included in the model; a group of terms (e.g., the XT terms) is included only if this group is statistically significant.

The fitting method used to estimate the μ's maximizes the likelihood function using the iterative-proportional fitting algorithm (see Bishop et al., 1974), with the likelihood function based on the assumption of Poisson counts in the cells. Significance of the calculated parameters is assessed using the likelihood ratio:

$$L^2 = -2\sum_{i\,j\,k} n_{ijk}\ln\left(\frac{\lambda_{ijk}}{n_{ijk}}\right) \qquad (2)$$

which compares the log-likelihood of the model to that of the saturated model (in which the expected count for each cell is equal to the observed count for that cell; i.e, n_{ijk}). This likelihood ratio has a distribution which is approximately χ^2, with degrees of freedom equal to the number of cells minus the number of parameters being estimated. Results from the application of this model, using different terms, are shown in Table 1.

A review of the results presented in Table 1 indicates that, comparing models 1 and 2, there is a significant decrease in L^2 for model 2 compared to model 1, and that this decrease is significant at the α=0.001 level. This implies that non-stationarity in seismicity should be accepted, and that a time term should be introduced. The calculated time term $\exp(\mu_T)$ is plotted versus time for 50-year time steps in Figure 4 (top panel); it indicates that the rate of seismicity varies by a factor of 4, from the lowest rate to the highest, in all provinces. The significance of nonstationarity agrees with previous investigations of the Chinese catalog (e.g. Lee and Brillinger, 1979; McGuire, 1979).Model 3 investigates whether non-stationarity is province dependent, which would include whether or not migration is taking place. Comparison of the L^2 values and degrees of freedom indicates that α=0.61, i.e. there is a 61% probability that the observed XT dependence is caused by chance. Thus we judge this term to be not statistical-

ly significant. Figure 4 (middle panel) plots $\exp(\mu_{XT})$ versus time; note that the amplitude of this term is comparable to the μ_T term, even though μ_{XT} is statistically insignificant.

Table 1. Results from categorical analysis of earthquake counts.

Model	First Order Terms	Second Order Terms	L^2	d.o.f.	Description
1	X,M	--	42.5	26	stationary, same b value for all provinces
1'	X,M	XM	41.7	24	stationary, province-dependent b value
2	X,M,T	--	24.1	22	non-stationary, same value for all provinces
3	X,M,T	XT	17.8	14	non-stationary, different time-dependence for all provinces
4	X,M,T	MT	17.8	18	non-stationary, time-dependent b values
5	X,M,T	XM	23.3	20	non-stationary, province-dependent b value

Model 4 investigates whether there is a time-dependent b-value for the entire region (all provinces considered together). The calculated value of alpha=0.18 indicates a low level of significance, even though the amplitude of this term is comparable to the μ_{T_j} term. Figure 4 (bottom panel) illustrates this term plotted versus time.

Model 5 represents a province-dependent b-value; comparison to model 2 indicates almost the same likelihood, so that this effect is judged to be statistically insignificant. This is confirmed by Model 1', which allows a separate b-value for each province and which does not lead to a significantly different likelihood value from Model 1.

One result from this analysis is that there is very strong evidence of nonstationarity in the historical catalog. This is not surprising, and other researchers have come to the same conclusion. There is weak evidence of temporal variation in b-values, and this effect may result from incompleteness in the data at early times (see figure 10). Finally, there is no evidence of migration of seismicity. A similar analysis was conducted with provinces 2 and 3 alone, using

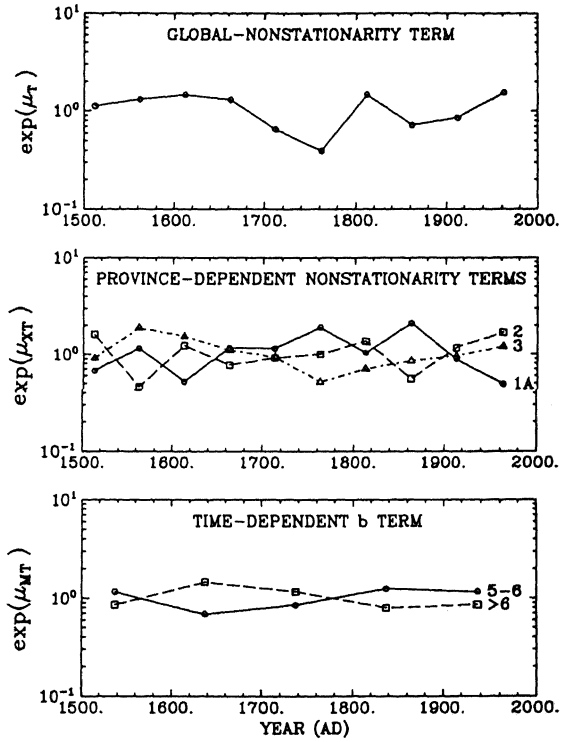


Figure 4. Terms obtained from the categorical analysis of the catalog. T and XT terms were obtained using 50-year intervals; MT term was obtained using 100-year intervas.

both 100- and 50-year time windows, leading to identical conclusions.

4 MODELS OF NON-STATIONARITY

To model the effects seen in the Chinese data, we develop two models of nonstationarity. These are extensions of work conducted by other researchers.

4.1 Stress-accumulation model

Our model of stress accumulation and release is based on the concept of accumulation and release of tectonic energy, which has been proposed by Shi et al (1983) and other Chinese investigators. This model has been applied to North China as well by Vere-Jones and Yonglu (1988) and by Zheng and Vere-Jones (1991). The model takes the form:

$$\lambda(t)=\exp[\alpha+\beta(t-\rho\sum_i 10^{\sigma(M_i-5)})] \qquad (3)$$

where $\lambda(t)$ (the earthquake potential) is the expected rate of earthquake occurrences at time t, the term

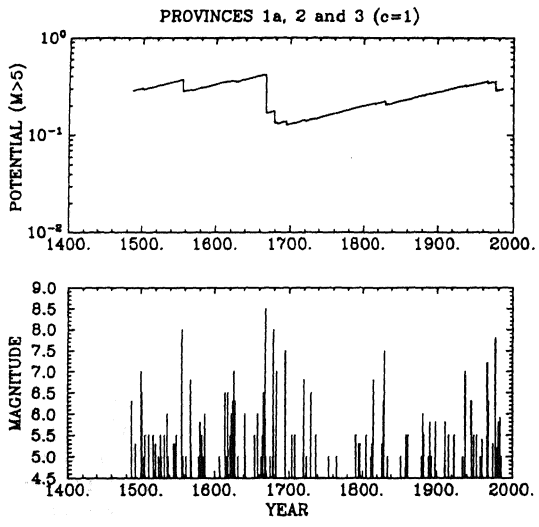


Figure 5. Variation of potential and activity as a function of time from the stress-release model fitted to provinces 1a, 2, and 3.

βt represents accumulation of tectonic input, and the term $\beta\rho\Sigma_i 10^{c(M_i-5)}$ represents tectonic release. The value of parameter c proposed by Vere-Jones and co-workers is 0.75 on the basis of conservation of stress. If, on the other hand, the tectonic driving mechanism is a constant slip rate (e.g., Anderson, 1979), and the moment-magnitude scale is used, the value of c should be 1.5. Based on preliminary runs we obtained a value of c of 0.57, with standard error 0.32, but we judge this value to be too small, so we adopt a value of 1.0. Our reasoning is that the concept of a slip-consistent model make physical sense, but a value of 1.5 would make the model too sensitive to the occurrences of large earthquakes that are few in number and whose magnitudes are imprecise known.

The maximum likelihood method was used to fit equation 3 to data in provinces 1A, 2, and 3, with the value of c fixed at 1.0. The results obtained were:

$$\alpha = -1.25(\pm 13\%)$$

$$\beta = 0.0046(\pm 34\%)$$

$$\rho = 0.062\ (\pm 12\%)$$

with a correlation matrix from the estimates of:

1.00		
-0.39	1.00	
0.84	-0.52	1.00

The fit to the catalog is shown in Figure 5; the stress-accumulation model replicates the overall variations in earthquake rate for these three provinces. To test the significance of these results, the likelihood ratio statistic was used, to test against the hypothesis that the observed catalog was generated by a Poisson process. The indicated α value was 0.002, implying that it is highly improbable that a Poisson sequence would have caused the historical seismicity.

To further test the goodness of fit of this model, artificial catalogs were generated from the model drawing magnitudes from the empirical distribution. Parameters then were fit to each artificial catalog, and the process was repeated to obtain a set of likelihood values from multiple, simulated catalogs. The likelihood from the real catalog was within the range of the likelihoods from the simulations, implying that the model fits the observed data.

Simulation of the 500 years of seismicity following the catalog (1987-2486), including uncertainty in parameters, indicates essentially no change in the expected average rate. Figure 6 shows expected average rates and one-standard-deviation bounds on rates for 25 yr intervals in the future, plotted at the centers of the intervals. The expected average rate for the next 50 years is 0.31 events per year, slightly higher than the 500-year historic rate of 0.26, and there is no observable trend with time. The bounds diverge with time, because of the possibility of a large earthquake occurring which would reduce the earthquake potential for a period subsequent to that event.

There are several observations from these applications. First, the North China catalog exhibits periodicity, but the stress-release model does not contain periodicity (see Figure 12). Second, attempts to fit the model to individual provinces (where it should work best because the effects of stress release should be stronger at a local level) met with mixed success: province 2 gave satisfactory results, but province 3 did not. This implies an inconsistency: if models of this type are correct, their effects should be seen more strongly at the local level than at the regional level. One possible explanation is that there is increased activity (over a time scale longer than that of obvious aftershocks), caused by crustal stress readjustments after major earthquakes. This increased activity masks the reduction in activity caused by a decrease in the average strains in the region surrounding the large earthquake.

An additional issue that was examined was the coupling between high rates of activity and the magnitude distribution. We expected to observe that large earthquakes tend to occur during periods when the seismic potential is high, but there was no evidence of coupling in the data we analyzed. There is

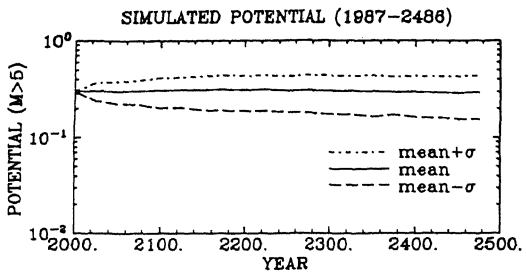


Figure 6. Variation of potential as a function of time from simulations of the stress-accumulation model for the period 1987-2486.

a weak trend toward higher maximum magnitudes during times of high potential, but this trend does not appear to be significant.

The success of the stress-accumulation model to represent historical seismicity leads us to use it as one of the models to examine for seismic hazard analysis.

5 OTHER RESULTS

This project has also investigated local interaction effects (e.g., imunity), exponentiality of the magnitude distribution, and alternative models of non-stationarity. Results will be reported in a future paper.

6 CONCLUSIONS

This application of non-stationary models of seismicity to the Chinese earthquake catalog has led to several conclusions:

1. Although the catalog exhibits non-stationarity when several provinces are taken together, the data from individual provinces are too sparse to derive separate models of seismic cycles for these individual provinces. As a result, non-stationarity must be modeled in the global sense. Application of these models to seismic hazard analysis at the local level must include uncertainty on how the total potential is apportioned among the provinces (and among the sources in each province).

2. Calibration of non-stationary models to historical seismicity in North China over the last 500 years, using provinces with the most complete data, and calculation of earthquake potential for the short-term (next 50 yrs), leads to expected rates of earthquake occurrence that are similar to the long-term average rate of seismicity. This results from the current phase of seismicity in China and would not necessarily be transferable to other regions of the world.

3. In other areas with similar seismicity characteristics the seismic potential could vary by a total factor of 4, depending on the current state in the stress-accumulation process. It does not appear, according to the stress-accumulation model, that earthquake potential increases rapidly over a period of, say decades. The earthquake potential may have large drops over short periods of time, but these drops are very infrequent because they are caused by large earthquakes.

ACKNOWLEDGMENTS

The work reported here was conducted under the U.S.-China Protocol for Cooperative Scientific Research. Support was provided by the U.S. National Science Foundation (under Grant CES-8814817) and by Chinese State Seismological Bureau. Dr. Cliff Astill was project manager for NSF.

REFERENCES

Anderson, J.G. 1979. Estimating seismicity from geological structure for seismic risk studies. Bull. Seis. Soc. Am., 69: 135-158.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. 1974. Discrete multivariate analysis. Cambridge, Ma., MIT Press.

Bonilla, M.G., Mark, R.K., and Lienkaemper, J.J. 1984. Statistical relations among earthquake magnitude, surface rupture length, and surface fault displacement. Bull. Seis. Soc. Am., 74: 2379-2412.

Lee, W.H.K., and Brillinger, D.R. 1979. On chinese earthquake history-an attempt to model an incomplete data set by point process analysis. Pageoph., 117: 1229-1257.

McGuire, R.K. 1979. Adequacy of simple probability models for calculating felt-shaking hazard, using the chinese earthquake catalog. Bull. Seis. Soc. Am., 69: 877-892.

Shi, S., Huan W., Lu, S., and Yan J. 1983. On the characteristics of seismic activity in central and eastern Asia continent. Scientia Sinica, Series B, 26: 438-448.

Vere-Jones, D., and Yonglu, D. 1983. A point process analysis of historical earthquakes from North China. Earthquake Research in China, 2: 165-181.

Zheng, X.-G., and Vere-Jones, D. 1991. Application of stress rekease models to historical earthquakes from North China. Pageoph. 135: 558-576.