SPATIOTEMPORAL PATTERNS OF EARTHQUAKES AND THEIR IMPLICATIONS FOR EARTHQUAKE HAZARDS

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment of

the Requirements for the Degree

Doctor of Philosophy

---

by

YUXUAN CHEN

Dr. Mian Liu, Dissertation Supervisor

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

SPATIOTEMPORAL PATTERNS OF EARTHQUAKES AND THEIR IMPLICATIONS FOR EARTHQUAKE HAZARDS

Presented by Yuxuan Chen, a candidate for the degree of Doctor of Philosophy, and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Mian Liu

Professor Francisco Gomez

Professor Eric Sandvol

Professor Shih-Kang Chao

ACKNOWLEDGEMENTS

I want to express my great appreciation to my advisor Professor Mian Liu. Thank you for your instructions and encouragement. I have learned how to become a scientist from you.

I want to thank my committee members, Professor Francisco Gomez, Eric Sandoval, and Shih-Kang Chao for providing constructive suggestions on my dissertation research. I also want to thank all teachers in the University of Missouri who offered me valuable courses and advice.

Finally, I want to express my gratitude to my family and friends for their love and support.

TABLE OF CONTENTS

Chapter

List of Figures

ABSTRACT

This work focuses on characterizing spatiotemporal patterns of earthquakes, their possible causes, and their implications for seismic hazard assessment. I studied both local and global earthquakes in the view of complex fault systems. Specifically, I studied the background seismicity and long-lived aftershock activities in intraplate North China and the Central and Eastern United State (CEUS), and characterized the correlation between strain rate and seismicity and evaluated the prediction power of strain rate in different tectonic settings.

I found that periodic or quasiperiodic earthquake recurrence on individual faults, as predicted by the elastic rebound model, is not common in nature. Instead, most earthquake sequences are complex and variable, and often show clusters of events separated by long but irregular intervals of quiescence. The common earthquake clustering may be caused by earthquake-induced viscoelastic relaxation and fault interaction. Most earthquake sequences are burstier than the Poisson model, implying a higher probability of repeating events soon after a large earthquake. Possible long-lived aftershocks are found in intraplate North China and the CEUS. Background seismicity in intraplate regions may vary with time, highlighting the complexity of intraplate seismicity. Mistakenly identifying long-lived aftershocks as background earthquakes may overestimate seismic hazard in intraplate regions. The correlation between strain rate and seismicity varies between different tectonic settings and is time-dependent. Good strain rate-seismicity correlations are found in plate boundary regions and during seismically active periods, while no correlations are found in stable continents and during inactive periods. All these variations need to be considered in hazard assessment.

Chapter 1: Introduction

An earthquake, the shaking of earth caused by a sudden release of energy, has potential to cause hazards, including but not limited to: ground shaking, surface rupture, landslide, soil liquefaction, and tsunamis. Large earthquakes can cause huge economic loss and massive casualties. For example, the 2011 *M* 9.1 Tohoku earthquake caused around $200 billion US dollar in economic loss, and the 1976 *M* 7.8 Tangshan earthquake killed more than 240,000 people (Chen, 1988; Satake, 2014). Assessing and mitigating earthquake hazard require a sufficient understanding of spatiotemporal patterns of earthquakes, while current knowledge about them is limited or even inaccurate, so the temporal and spatial patterns of earthquakes, their possible causes, and their implications for seismic hazard assessment are the focuses of this dissertation.

Plate tectonic theory predicts that large earthquakes occur mainly in plate boundaries where tectonic loading concentrates. However, plate boundaries are usually not formed by single fault planes but systems of fault branches and segments. The fault systems share the loading from relative plate motions. Within the fault systems, faults and fault segments interact with each other. Under this circumstance, the isolated single fault plane assumed in the elastic rebound model is not satisfied, and studies of single-fault-based earthquake recurrence are inadquate for characterizing the spatiotemporal patterns of earthquakes. Moreover, large earthquakes also occur in plate interiors, where slow far-field tectonic loading is shared by widespread interacting faults (M. Liu & Stein, 2016). Different from interplate earthquakes, intraplate earthquakes seems to roam between widespread faults and less likely to be quasiperiodic (M. Liu et al., 2011). Therefore, some commonly used concepts, such as seismic cycles, recurrence intervals, characteristic earthquakes, and seismic gaps, may be inadequate or even inaccurate. The probability seismic hazard assessment based on these concepts need to be rethought.

1.1 Temporal patterns of earthquakes

The recurrence time of earthquakes is an important parameter for hazard assessment and therefore a focus of earthquake studies (Atwater et al., 2003; Molnar, 1979; Youngs & Coppersmith, 1985). At first, it may be helpful to have a brief review about the elastic rebound model, the base of current understanding of earthquake physics. The elastic rebound model is proposed in the aftermath of the 1906 San Francisco earthquake (Reid, 1910). In this model, stress gradually accumulates on a frictionally locked fault plane by the relative crustal motion across the fault. When the accumulated stress overcomes the fault strength, the fault ruptures in an earthquake, suddenly releasing the accumulated energy. As a result, stress on the fault drops, and the strained crust near the fault plane springs back to its undeformed shape (elastic rebound). With time, the process repeats. Nearly half a century after the introduction of the elastic rebound model, Plate tectonics provides the driving mechanism in the elastic rebound model, stating that the plate boundary faults are steadily loaded by the relative plate motion. Assuming at each time an earthquake ruptures the whole fault plane or fault segment (the characteristic earthquake model) (Schwartz & Coppersmith, 1984), the strength of the fault plane or fault segment is fixed, and along the fault or fault segment the stress drops uniformly to some background value, then the steady tectonic loading would lead to periodic or quasi-periodic occurrence of earthquakes.

Periodic or quasi-periodic earthquakes, however, are uncommon in nature. The Parkfield section of the San Andreas Fault in California is perhaps the best-known example for quasi-periodic earthquakes. Between 1857 and 1966, six M_w 6.0 event occurred there with a ~22-year recurrence interval (W. H. Bakun & Lindh, 1985), but the next event, which took place in 2004, was overdue, challenging the simple elastic rebound model (W. Bakun et al., 2005; Jackson & Kagan, 2006). In most fault systems in varies tectonic settings, large earthquakes are clustered in time, with long and variable quiescent intervals (Salditch et al., 2019). Paleoearthquake study by Sieh et al. (1989) at Pallett Creek, on the San Andreas Fault in California, found clusters of large earthquakes with interevent times in several decades, but the length of the dormant periods between earthquake clusters is two to three centuries. In the Great Basin of western US, paleoseismic data show clusters of events separated by long quiescent periods on individual faults or fault segments (Wallace, 1987). At the Africa-Eurasia plate boundary off west Algeria, Ratzov et al. (2015) analyzed a long record of turbidites and found that three clusters of earthquakes with durations of ~300-600 years separated by two long quiescent periods of ~1.6 thousand years. A similar pattern is found along the Dead Sea Transform Fault based on a sixty-thousand-year record of seismites (Agnon, 2014). In Australia, rich morphogenic evidence of faulting across the continent commonly shows that a few large earthquakes within thousands of years are separated by much longer periods (10^4-10^6 years) of quiescence on a single fault or proximal faults (Clark et al., 2014; Clark et al., 2012).

Rather than the time-dependent model based on the elastic rebound model (Matthews et al., 2002), the time-independent Poisson model (Cornell, 1968) is commonly used as the model for earthquake occurrence in the Probabilistic Seismic Hazard Assessment (PSHA) (Fujiwara et al., 2006; Petersen et al., 2014). The time-independent model assumes that earthquake occurrence follows the Poisson process, so each earthquake is independent, and the probability of earthquake occurrence does not vary with time (i.e. time-independent) (Cornell, 1968). Therefore, it is worthy to statistically check whether large earthquakes occur periodically, randomly, or burstily, explore the underlying physics, and discuss the implications for earthquake hazard assessment. The main results and discussions are presented in Chapter 2.

Another special temporal pattern of earthquakes focused in this dissertation is possible long-lived aftershocks in intraplate regions. In slowly deforming intraplate regions, aftershock sequences may be much longer than that at plate boundaries (Stein & Liu, 2009). Some aftershocks may be mistakenly identified as elevated background seismicity and therefore may lead to an overestimation of seismic hazard (Toda & Stein, 2018) and raise unwarranted social concerns and panic (M. Liu & Wang, 2012). The possible long-lived aftershock activities are still in debate in North China and the CEUS (Ebel et al., 2000; Jiang et al., 2013; Y. Liu et al., 2020; Page & Hough, 2014; Zhong & Shi, 2012), so I separated background earthquakes and aftershocks in North China and the CEUS by using a statistical method called the nearest-neighbor method, and discussed the uncertainties and implications for seismic hazards, which are detailed in Chapter 3 and 4.

1.2 Spatial patterns of earthquakes

Earthquakes occur on faults. Plate tectonics describes the major spatial pattern of earthquakes: most events occur at plate boundary faults. However, plate tectonics theory is only a rough approximation. In fact, plate boundaries are not several faults with simple geometry but complex fault systems sharing tectonic loading. Because tectonic loading rate is quite steady, large earthquakes repeatedly ruptured on plate boundary faults. In contrast, plate interiors are slowly loaded from the far field, and large earthquakes seem to roam between different faults and are not necessary to recur on the same faults (Calais et al., 2016; M. Liu et al., 2011).

Spatial distributions of earthquakes are found to correlate well with observed strain rate fields in plate boundary regions (Kreemer et al., 2002; Shen et al., 2007; Zeng et al., 2018), but poor or no correlation between strain rate and seismicity are found in intraplate regions, such as North China and stable North America (Calais et al., 2016; Kreemer et al., 2018; M. Liu & Wang, 2012). Therefore, I did a systematic statistical analysis on the correlation between strain rate and seismicity in different tectonic settings, and also evaluated the power of strain rate as a spatial predictor of earthquakes in Chapter 5.

For intraplate regions, because of low strain rate, the spatial patterns of intraplate earthquakes may also be controlled by other factors. The ancient rift zones are one of the factor that has been suggested to be associated with intraplate seismicity (Gangopadhyay & Talwani, 2003; Johnston & Kanter, 1990). Schulte and Mooney (2005) did a systematic study on this possible correlation with an updated global earthquake catalogue from year 495 to 2003. Their result show that most of $M \geq 7$ earthquakes occurred within rifts and continental margins, but only 52% of event are associated with rifted crust for M

≥ 4.5 earthquakes. M. Liu and Stein (2016) also point out that the Mid-continent rift in central and eastern U. S. is essentially aseismic, so rift zones are neither necessary or sufficient for intraplate earthquakes. Mooney et al. (2012) explored the correlation between global intraplate seismicity and the seismic velocity of the lithosphere. They found that $5 \leq M_w \leq 7$ intraplate earthquakes tend to occur at rifted continental margins and younger crust surrounding the ancient cratons, which are defined by the S-wave velocity anomalies from seismic tomography in a global scale. Locally, Assumpção et al. (2014) also found that S-wave velocity anomalies at 100 km depth in Brazil are correlated with the epicenter distributions of earthquakes and interpreted it as the results of stress concentration due to lithospheric thinning and cratonic edge effect. Except edge of cratons and rifts, large intraplate earthquakes do occur at other unexpected places. For example, the 1966 $M7.2$ Xingtai earthquake and 1976 $M7.8$ Tangshan earthquake occurred within North China Plain and no active tectonics or major active faults were recognized before their occurrence (M. Liu & Wang, 2012). Therefore, rifts and seismic velocity anomaly are possible indicators for the spatial distribution of intraplate earthquakes, which are studied in Chapter 3.

1.3 Phenomenological laws in earthquake occurrence

Faulting is a highly non-linear process controlled by various physical mechanisms at different spatiotemporal scales and is also affected by many external components (e.g., stress perturbation from nearby faults or fault segments). As a result, seismicity shows fractal structures in magnitude, time, and space distributions, as expressed in the

Gutenberg–Richter law, the Omori’s law, and the fractal dimension of earthquake locations (Hirata, 1989; Kagan, 1994; Turcotte, 1997), which are introduced below.

1.3.1 Size distribution: the Gutenberg–Richter law

The Gutenberg-Richter law is an empirical relationship describing the power-law (fractal) magnitude distribution of earthquakes in a defined region and time interval:

$$\log_{10} N = a - bM \tag{1.1}$$

where *N* is the total number of earthquakes with magnitude $\geq M$, a and b are constants (Gutenberg & Richter, 1944, 1954). The parameter *a* depends on the area and time-window of investigation and describes the productivity, while the parameter b describes the relative size distribution of earthquakes (El-Isa & Eaton, 2014). The b-value for global earthquakes is close to 1 and typically in the range of 0.8-1.2. The GR law is valid both regionally and globally, and is used in calculating recurrence time intervals of earthquakes with different magnitudes, mapping subsurface magmatic chambers, and investigating induced seismicity (El-Isa & Eaton, 2014).

The value of b in this study is calculated by the maximum likelihood method (Aki, 1965):

$$b = \frac{1}{\ln(10)\,(\bar{M} - M_{min})} \tag{1.2}$$

where $\bar{M}$ is the mean magnitude of the sample and M_{min} is the minimum magnitude or magnitude of completeness (M_c).

1.3.2 The Omori's law and Aftershock duration

A large earthquake is usually followed by many smaller earthquakes (aftershocks). Their occurrence rate typically decreases with time and follows the modified Omori's law (Omori, 1894; Utsu, 1961):

$$n(t) = \frac{k}{(c+t)^p} \tag{1.3}$$

where $n(t)$ is the frequency of earthquakes per unit time interval at time t after the mainshock (aftershock rate); k, p, and c are constants. The value of c is typically positive close to zero, and p close to 1 (Utsu & Ogata, 1995). In this study, these three parameters and their uncertainties are estimated by the maximum likelihood method (Ogata, 1983).

The aftershock duration can be estimated based on the Omori's law. Aftershock duration is typically defined as the time required for seismicity rate after an mainshock to return to the background seismicity rate before the next mainshock (Dieterich, 1994; Toda & Stein, 2018). Assuming background seismicity rate r_b is time-independent, the aftershock duration t_a can be estimated by

$$t_a = \left(\frac{k}{r_b}\right)^{\frac{1}{p}} - c \tag{1.4}$$

1.3.3 Fractal dimension of earthquake locations

The spatial distributions of earthquakes have been proved to have a fractal structure (Hirata, 1989; Wu et al., 2017):

$$C(r) \propto r^{d_f} \tag{1.5}$$

where d_f is the fractal dimension of an spatial distribution of epicenters (longitude and latitude) or hypocenter (longitude, latitude, and depth), $C(r)$ is the correlation integral introduced by Grassberger and Procacia (1983), and r is the correlation distance. $C(r)$ is defined as

$$C(r) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} H\left(r - \left\|\overrightarrow{X_i} - \overrightarrow{X_j}\right\|\right) \tag{1.6}$$

where N is the number of events in the data set, $\left\|\overrightarrow{X_i} - \overrightarrow{X_j}\right\|$ is the epicentral distance between event i and j, and H is the step function ($H(x) = 0$ for $x \leq 0$ and $H(x) = 1$ for $x > 0$). $N_c = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} H\left(r - \left\|\overrightarrow{X_i} - \overrightarrow{X_j}\right\|\right)$ is the cumulative number of pairs (i, j) whose distance is less than the correlation distance r.

The fractal dimension d_f is a measure of earthquake clustering in space (Dimitriu et al., 1993; Yadav et al., 2011). The value of d_f close to 0 may be interpreted as all earthquakes tend to cluster into one point. The value of d_f close to 1 indicates the dominance of line sources. The value of d_f close to 2 indicates planar fractured surface being filled-up. Therefore, the smaller the fractal dimension, the stronger the earthquake clustering in space.

The fractal dimension d_f varies with space and time (De Rubeis et al., 1993; Dimitriu et al., 1993). In the Tohoku region, Japan, the estimated fractal dimension of earthquake epicenters is $d_f = 1.69$ (Hirata, 1989). In Southern California, the estimated fractal dimension of earthquake epicenters is $d_f = 1.6$ (Alvaro Corral, 2003).

1.4 Probability models used in earthquake studies

One way to characterize the temporal patterns of earthquakes is to fit the interevent times in earthquake catalogs with probability distribution functions. Five probability models (Poisson, gamma, Weibull, lognormal, and Brownian passage time) are used in this dissertation and introduced below. All of them were used in previous studies on the interevent-time distribution of earthquakes.

The Poisson model is a one parameter model (the mean μ) and commonly used in probabilistic hazard analyses (Cornell, 1968; Field et al., 2014; Field et al., 2009; Fujiwara et al., 2006; WGCEP, 1988, 1995). The gamma distribution is a two-parameter distribution defined by the shape parameter a and the scale parameter b. It has been used for global and local catalogs with different magnitude range (Alvaro Corral, 2004; Hainzl et al., 2006). The Weibull distribution is a similar two-parameter distribution used in previous studies (Abaimov et al., 2008; Hasumi et al., 2009). The lognormal distribution is a two-parameter distribution defined by the mean μ and the standard deviation σ of the variable's natural logarithm, and the BPT model is a two-parameter distribution defined by the mean μ and the aperiodicity α. They also have been used in earthquake hazard analysis (Field et al., 2015; Field et al., 2009; WGCEP, 1988, 1995, 2003).

Table 1.1: Properties of Possible Interevent-time distributions

Name of distribution	Probability Density Function $f(\tau)$	Parameters	Properties of Parameters
Exponential (Poisson model)	$\frac{1}{\mu}e^{-\frac{\tau}{\mu}}$	$\mu > 0$	$\mu = \text{mean}(\tau)$
Gamma	$\frac{1}{b^a\,\Gamma(a)}\tau^{a-1}e^{-\frac{\tau}{b}}$	$a > 0$ $b > 0$	a: shape b: scale
Weibull	$\frac{a}{b}\left(\frac{\tau}{b}\right)^{a-1} e^{-\left(\frac{\tau}{b}\right)^a}$	$a > 0$ $b > 0$	a: shape b: scale
Lognormal	$\left(\sqrt{2\pi}\sigma\tau\right)^{-1} e^{-\frac{(\ln\tau-\mu)^2}{2\sigma^2}}$	$\mu \in (-\infty, \infty)$ $\sigma > 0$	$\mu = \text{mean}(\ln\tau)$ $\sigma = \text{std}(\ln\tau)$
Brownian passage time	$\left(\frac{\mu}{2\pi\alpha^2\tau^3}\right)^{1/2} e^{-\frac{(\tau-\mu)^2}{2\alpha^2\mu\tau}}$	$\mu > 0$ $\alpha > 0$	$\mu = \text{mean}(\tau)$ $\alpha = \frac{\text{std}(\tau)}{\text{mean}(\tau)}$

* τ is interevent time between two successive events. $\Gamma(a)$ is the gamma function $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$. The std means the standard deviation.

The Poisson model is a special case of the gamma and Weibull distributions. When the shape parameter $a = 1$ and the scale parameter $b = \mu$, gamma and Weibull distributions are the same as the Poisson model (Table 1.1). When $a < 1$, the gamma and Weibull distributions are burstier than the Poisson model (Figure 1.1a). When $a > 1$, the gamma and Weibull distributions are less burstier than the Poisson model (Figure 1.1b).

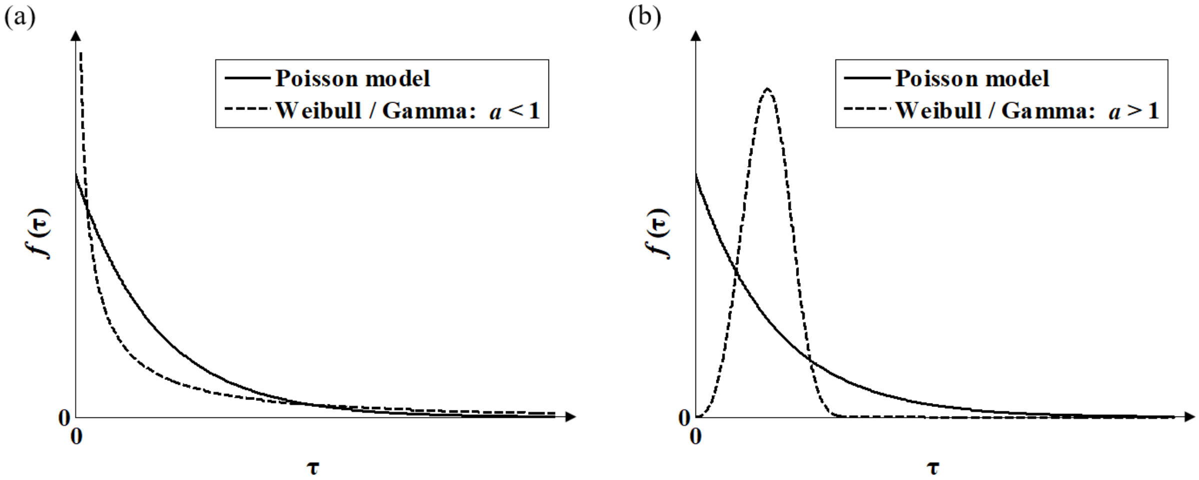


Figure 1.1. Comparison between the Poisson distribution with the Weibull and gamma distributions when the shape parameter a is smaller or greater than one.

1.5 Structure of this dissertation

Chapters 2 presents my study on the temporal patterns of large earthquakes charactering by the Devil's staircase, a fractal property of complex systems. In Chapter 3, I narrow my scope on the spatiotemporal patterns of earthquakes in North China. Chapter 4 presents my study on the possible long-lived aftershocks in the CEUS. In Chapter 5, Correlations between stain rate and seismicity in different tectonic settings are explored. Finally, the main conclusions from this dissertation are summarized in Chapter 6, which also includes suggestions for future research.

References

Abaimov, S., Turcotte, D., Shcherbakov, R., Rundle, J., Yakovlev, G., Goltz, C., & Newman, W. (2008). Earthquakes: Recurrence and Interoccurrence Times. *Pure and Applied Geophysics, 165*(3-4), 777-795.

Agnon, A. (2014). Pre-instrumental earthquakes along the Dead Sea rift. In *Dead Sea transform fault system: reviews* (pp. 207-261): Springer.

Aki, K. (1965). Maximum likelihood estimate of b in the formula log N= a-bM and its confidence limits. *Bull. Earthq. Res. Inst., Tokyo Univ., 43*, 237-239.

Assumpção, M., Ferreira, J., Barros, L., Bezerra, F., FranÇa, G. S., Barbosa, J. R., et al. (2014). Intraplate seismicity in Brazil. *Intraplate Earthquakes. Cambridge University Press, New York*, 50-71.

Atwater, B. F., Tuttle, M. P., Schweig, E. S., Rubin, C. M., Yamaguchi, D. K., & Hemphill-Haley, E. (2003). Earthquake recurrence inferred from paleoseismology. *Developments in Quaternary Sciences, 1*, 331-350.

Bakun, W., Aagaard, B., Dost, B., Ellsworth, W., Hardebeck, J., Harris, R., et al. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature, 437*(7061), 969.

Bakun, W. H., & Lindh, A. G. (1985). The Parkfield, California, earthquake prediction experiment. *Science, 229*(4714), 619-624.

Calais, E., Camelbeeck, T., Stein, S., Liu, M., & Craig, T. (2016). A new paradigm for large earthquakes in stable continental plate interiors. *Geophysical Research Letters, 43*(20).

Chen, Y. (1988). *The great Tangshan earthquake of 1976: an anatomy of disaster*: Pergamon.

Clark, D., McPherson, A., & Allen, T. (2014). Intraplate earthquakes in Australia. In *Intraplate Earthquakes* (Vol. 8, pp. 49): Cambridge University Press, New York.

Clark, D., McPherson, A., & Van Dissen, R. (2012). Long-term behaviour of Australian stable continental region (SCR) faults. *Tectonophysics, 566*, 1-30.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America, 58*(5), 1583-1606.

Corral, A. (2003). Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Physical Review E, 68*(3), 035102.

Corral, A. (2004). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letters, 92*(10), 108501.

De Rubeis, V., Dimitriu, P., Papadimitriou, E., & Tosi, P. (1993). Recurrent patterns in the spatial behaviour of Italian seismicity revealed by the fractal approach. *Geophysical Research Letters, 20*(18), 1911-1914.

Dimitriu, P., Papadimitriou, E., Papazachos, B., & Tsapanos, T. (1993). Global study of the distribution of earthquakes in space and in time by the fractal method. *Proc. 2nd Cong. Hellenic Geophys. Union, Florina, 1*, 164-174.

Ebel, J. E., Bonjer, K.-P., & Oncescu, M. C. (2000). Paleoseismicity: Seismicity evidence for past large earthquakes. *Seismological Research Letters, 71*(2), 283-294.

El-Isa, Z. H., & Eaton, D. W. (2014). Spatiotemporal variations in the b-value of earthquake magnitude–frequency distributions: Classification and causes. *Tectonophysics, 615*, 1-11.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., et al. (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model. *Bulletin of the Seismological Society of America, 104*(3), 1122-1180.

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., et al. (2015). Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3). *Bulletin of the Seismological Society of America*, *105*(2A), 511-543.

Field, E. H., Dawson, T. E., Felzer, K. R., Frankel, A. D., Gupta, V., Jordan, T. H., et al. (2009). Uniform California earthquake rupture forecast, version 2 (UCERF 2). *Bulletin of the Seismological Society of America, 99*(4), 2053-2107.

Fujiwara, H., Kawai, S., Aoi, S., Morikawa, N., Senna, S., Kobayashi, K., et al. (2006). National seismic hazard maps of Japan. *Bull. Earthq. Res. Inst. Univ. Tokyo, 81*, 221-232.

Gangopadhyay, A., & Talwani, P. (2003). Symptomatic features of intraplate earthquakes. *Seismological Research Letters, 74*(6), 863-883.

Grassberger, P., & Procacia, I. (1983). Measuring the strangeness of strange attractors. *Physica. D, 9*(1-2), 189-208.

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America, 34*(4), 185-188.

Gutenberg, B., & Richter, C. F. (1954). Seismicity of the earth and related phenomena. *Princeton*.

Hainzl, S., Scherbaum, F., & Beauval, C. (2006). Estimating background activity based on interevent-time distribution. *Bulletin of the Seismological Society of America, 96*(1), 313-320.

Hasumi, T., Akimoto, T., & Aizawa, Y. (2009). The Weibull–log Weibull distribution for interoccurrence times of earthquakes. *Physica A: Statistical Mechanics and its Applications, 388*(4), 491-498.

Hirata, T. (1989). A correlation between the b value and the fractal dimension of earthquakes. *Journal of Geophysical Research: Solid Earth, 94*(B6), 7507-7514.

Jackson, D. D., & Kagan, Y. Y. (2006). The 2004 Parkfield earthquake, the 1985 prediction, and characteristic earthquakes: Lessons for the future. *Bulletin of the Seismological Society of America, 96*(4B), S397-S409.

Jiang, C., Wu, Z., & Zhuang, J. (2013). ETAS model applied to the Earthquake-Sequence Association (ESA) problem: the Tangshan sequence. *CHINESE JOURNAL OF GEOPHYSICS-CHINESE EDITION*, *56*(9), 2971-2981.

Johnston, A. C., & Kanter, L. R. (1990). Earthquakes in stable continental crust. *Scientific American, 262*(3), 68-75.

Kagan, Y. Y. (1994). Observational evidence for earthquakes as a nonlinear dynamic process. *Physica D: Nonlinear Phenomena, 77*(1-3), 160-192.

Kreemer, C., Hammond, W. C., & Blewitt, G. (2018). A robust estimation of the 3-D intraplate deformation of the North American plate from GPS. *Journal of Geophysical Research: Solid Earth*, *123*(5), 4388-4412.

Kreemer, C., Holt, W. E., & Haines, A. J. (2002). The global moment rate distribution within plate boundary zones. *Plate boundary zones, 30*, 173-190.

Liu, M., & Stein, S. (2016). Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards. *Earth-Science Reviews*, *162,* 364-386.

Liu, M., Stein, S., & Wang, H. (2011). 2000 years of migrating earthquakes in North China: How earthquakes in midcontinents differ from those at plate boundaries. *J Lithosphere, 3*(2), 128-132.

Liu, M., & Wang, H. (2012). Roaming earthquakes in China highlight midcontinental hazards. *Eos, Transactions American Geophysical Union, 93*(45), 453-454.

Liu, Y., Zhuang, J., & Jiang, C. (2020). Background Seismicity before and after the 1976 Ms 7.8 Tangshan Earthquake: Is Its Aftershock Sequence Still Continuing? *Seismological Research Letters*.

Matthews, M. V., Ellsworth, W. L., & Reasenberg, P. A. (2002). A Brownian model for recurrent earthquakes. *Bulletin of the Seismological Society of America, 92*(6), 2233-2250.

Molnar, P. (1979). Earthquake recurrence intervals and plate tectonics. *Bulletin of the Seismological Society of America, 69*(1), 115-133.

Mooney, W. D., Ritsema, J., & Hwang, Y. K. (2012). Crustal seismicity and the earthquake catalog maximum moment magnitude (Mcmax) in stable continental regions (SCRs): Correlation with the seismic velocity of the lithosphere. *Earth and Planetary Science Letters, 357,* 78-83.

Ogata, Y. (1983). Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure. *Journal of Physics of the Earth, 31*(2), 115-124.

Omori, F. (1894). *On the after-shocks of earthquakes* (Vol. 7): The University.

Page, M. T., & Hough, S. E. (2014). The New Madrid seismic zone: Not dead yet. *Science, 343*(6172), 762-764.

Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D., et al. (2014). *Documentation for the 2014 update of the United States national seismic hazard maps*. Retrieved from

Ratzov, G., Cattaneo, A., Babonneau, N., Déverchère, J., Yelles, K., Bracene, R., & Courboulex, F. (2015). Holocene turbidites record earthquake supercycles at a slow-rate plate boundary. *Geology, 43*(4), 331-334.

Reid, H. F. (1910). The Mechanics of the Earthquake, The California Earthquake of April 18, 1906, Report of the State Investigation Commission. *Vol. 2*.

Salditch, L., Stein, S., Neely, J., Spencer, B. D., Brooks, E. M., Agnon, A., & Liu, M. (2019). Earthquake supercycles and Long-Term Fault Memory. *Tectonophysics*, 228289.

Satake, K. (2014). *The 2011 Tohoku, Japan, earthquake and tsunami*: Cambridge University Press.

Schulte, S. M., & Mooney, W. D. (2005). An updated global earthquake catalogue for stable continental regions: reassessing the correlation with ancient rifts. *Geophysical Journal International*, *161*(3), 707-721.

Schwartz, D. P., & Coppersmith, K. J. (1984). Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5681-5698.

Shen, Z.-K., Jackson, D. D., & Kagan, Y. Y. (2007). Implications of Geodetic Strain Rate for Future Earthquakes, with a Five-Year Forecast of M5 Earthquakes in Southern California. *Seismological Research Letters, 78*(1), 116-120.

Sieh, K. E., Stuiver, M., & Brillinger, D. (1989). A more precise chronology of earthquakes produced by the San Andreas fault in southern California. *Journal of Geophysical Research, 94,* 603-623.

Stein, S., & Liu, M. (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature, 462*(7269), 87-89. http://www.scopus.com/inward/record.url?eid=2-s2.0-70449124788&partnerID=40&md5=37b8160bace8d6e8842ff1be818de349

Toda, S., & Stein, R. S. (2018). Why Aftershock Duration Matters for Probabilistic Seismic Hazard Assessment. *Bulletin of the Seismological Society of America, 108*(3A), 1414-1426.

Turcotte, D. L. (1997). *Fractals and chaos in geology and geophysics*: Cambridge university press.

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

Utsu, T., & Ogata, Y. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth, 43*(1), 1-33.

Wallace, R. E. (1987). Grouping and migration of surface faulting and variations in slip rates on faults in the Great Basin province. *Bulletin of the Seismological Society of America*, *77*(3), 868-876.

WGCEP. (1988). *Probabilities of large earthquakes occurring in California on the San Andreas fault* (88-398). Retrieved from http://pubs.er.usgs.gov/publication/ofr88398

WGCEP. (1995). Seismic hazards in southern California: Probable earthquakes, 1994 to 2024. *Bulletin of the Seismological Society of America*, *85*(2), 379-439.

WGCEP. (2003). *Earthquake probabilities in the San Francisco Bay Region, 2002-2031* (2003-214). Retrieved from http://pubs.er.usgs.gov/publication/ofr03214

Wu, H., Liu, H., Xu, W., & Wang, X. (2017). Fractal dimension and b value of the aftershock sequence of the 2008 M S 8.0 Wenchuan earthquake. *Natural Hazards, 88*(1), 315-325.

Yadav, R., Papadimitriou, E., Karakostas, V., Shanker, D., Rastogi, B., Chopra, S., et al. (2011). The 2007 Talala, Saurashtra, western India earthquake sequence: tectonic implications and seismicity triggering. *Journal of Asian Earth Sciences, 40*(1), 303-314.

Youngs, R. R., & Coppersmith, K. J. (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological Society of America, 75*(4), 939-964.

Zeng, Y., Petersen, M. D., & Shen, Z. K. (2018). Earthquake potential in California-Nevada implied by correlation of strain rate and seismicity. *Geophysical Research Letters*, *45*(4), 1778-1785.

Zhong, Q., & Shi, B. (2012). Aftershock duration of the 1976 Ms 7.8 Tangshan earthquake and implication for seismic hazard estimation. *Acta Seismol Sin, 34*(4), 494-508.

Chapter 2: Complex temporal patterns of large earthquakes: Devil's Staircases

*All content in this chapter is modified based our publication:

Chen, Y., Liu, M., & Luo, G. (2020). Complex Temporal Patterns of Large Earthquakes: Devil's Staircases. *Bulletin of the Seismological Society of America, 110*(3), 1064-1076. https://doi.org/10.1785/0120190148.

In this chapter, I examined the temporal patterns of large shallow earthquakes. I show that, for fault systems of various scales, most earthquake sequences share common characteristics that can be described by the Devil's Staircase, a fractal property of complex dynamic systems. I characterized these sequences statistically, explored the influencing tectonic factors, and discussed the implications for earthquake hazard assessment.

2.1 Temporal patterns of large earthquakes: Devil's Staircase

I focus on large shallow earthquakes ($M_w \geq 6$ and depth ≤ 60 km) because of the hazard that they pose and because they usually rupture multiple fault segments or faults, hence clearly deviate from simple elastic rebound models. The data I used are the ISC-GEM Global Instrumental Earthquake Catalogue (1904-2016) (Giacomo et al., 2018; Storchak et al., 2013), a historical earthquake catalog for North China from Cheng et al. (2017) and paleoseismic data for ruptures on different segments of the southern San Andreas compiled by Williams et al. (2019). The reason I analyze the historical and paleoseismic records is because instrumentally recorded large earthquakes in continental interiors and on individual faults or fault segments are often too few for vigorous statistical analysis.

Figure 2.1 shows the temporal patterns of large earthquakes in fault systems of different scales. Globally, $M \geq 8.5$ earthquakes show an irregular temporal pattern, with a long quiescent period (1965-2005) separating two active periods of clustered events (Figure 2.1a). Similar patterns emerge for large events in tectonically active regions, such as Japan and California, with long periods of no or few events followed by clusters of several events within relatively short periods (Figure 2.1b). In midcontinents, similar patterns are found, but the quiescent intervals between earthquake clusters tend to be a few times longer than those in tectonically active regions (Figure 2.1c). Such temporal patterns are found even on large individual faults (Figure 2.1d).

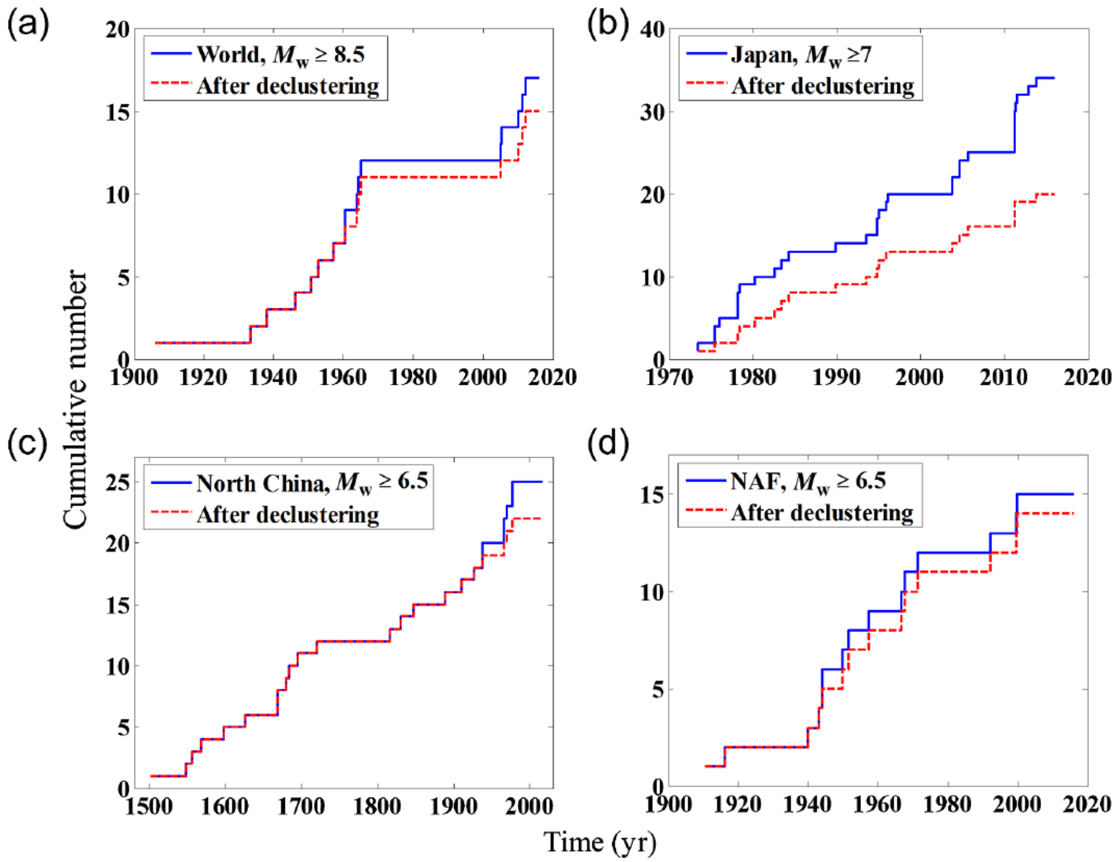


Figure 2.1. Temporal patterns of large earthquakes in (a) the world, (b) Japan, (c) North China, and (d) the North Anatolian Fault (NAF). Solid lines are for the whole catalogs, and dash lines are results after declustering, which has no clear effects on early records that have few large aftershocks.

Some of the earthquake clustering is likely due to dependent earthquakes (i.e., foreshocks and aftershocks). I used the method by Gardner and Knopoff (1974) to decluster catalogs. Earthquakes are grouped into different clusters according to time window T(days) and spatial window L(km) among them, which are defined as

$$\log_{10} T = \begin{cases} 0.032M + 2.7389, & \text{if } M \geq 6.5 \\ 0.5409M - 0.547, & \text{otherwise} \end{cases} \tag{2.1}$$

$$\log_{10} L = 0.1238M + 0.983 \tag{2.2}$$

where M is the magnitude of an mainshock. For every cluster, the largest earthquake is identified as the mainshock, and other events within the space-time window are removed. I declustered the catalogs to see the impact of dependent earthquakes on the temporal patterns. Because large foreshocks or aftershocks are as important as mainshocks in terms of hazard, the whole catalogs are used for most of our statistical analysis in this study.

Figure 2.1 shows that declustering removes some events from the clusters, but the general patterns are unchanged: earthquakes are clustered within relatively short periods, which are separated by longer and variable intervals of quiescence. Such patterns are known mathematically as the “Devil’s Staircase” (Mandelbrot, 1982; Turcotte, 1997). The Devil’s Staircase is a fractal property of complex dynamic systems and can be constructed from the Cantor set. The Devil’s Staircase is commonly found in nature, including depositional sequences and the reversal of Earth’s magnetic field (Bak, 1996; De Michelis & Consolini, 2003; Simkin & Roychowdhury, 2014; Turcotte, 1997). A fractal property is scale invariant. In Figure 2.2, I show earthquake sequences with different lower cut-off magnitudes from the global catalog. They all show similar patterns of clusters of events separated by longer periods of inactivity.

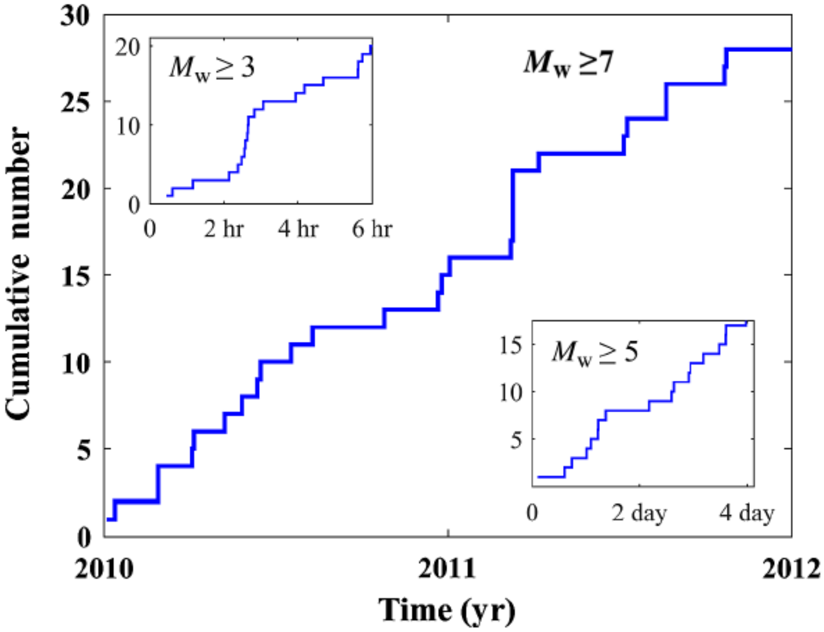


Figure 2.2. Temporal patterns of $M \geq 7$ earthquakes in the world. The insets show similar patterns for smaller events ($M \geq 5$ and $M \geq 3$, records start in 2010).

2.2 Characterizing the temporal patterns

One way to characterize the temporal patterns of earthquakes is to fit the interevent times in earthquake catalogs with probability distribution functions (Figure 2.3). Since all large earthquakes ($M_w \geq 6$) are potentially hazardous, I used the whole catalogs for regions and faults that have a dozen or more events for statistical analysis (Table 2.1). I fit the interevent time data with probability models using the maximum likelihood method. I tested five probability models (Poisson, gamma, Weibull, lognormal, and Brownian passage time (BPT)). All of them were used in previous studies on the interevent-time distribution of earthquakes and have been introduced in Chapter 1.4 (Abaimov et al., 2008; Cornell, 1968; Corral, 2004; Field et al., 2014; Field et al., 2015; Field et al., 2009; Fujiwara et al., 2006; Hainzl et al., 2006; Hasumi et al., 2009; WGCEP, 1988, 1995, 2003).

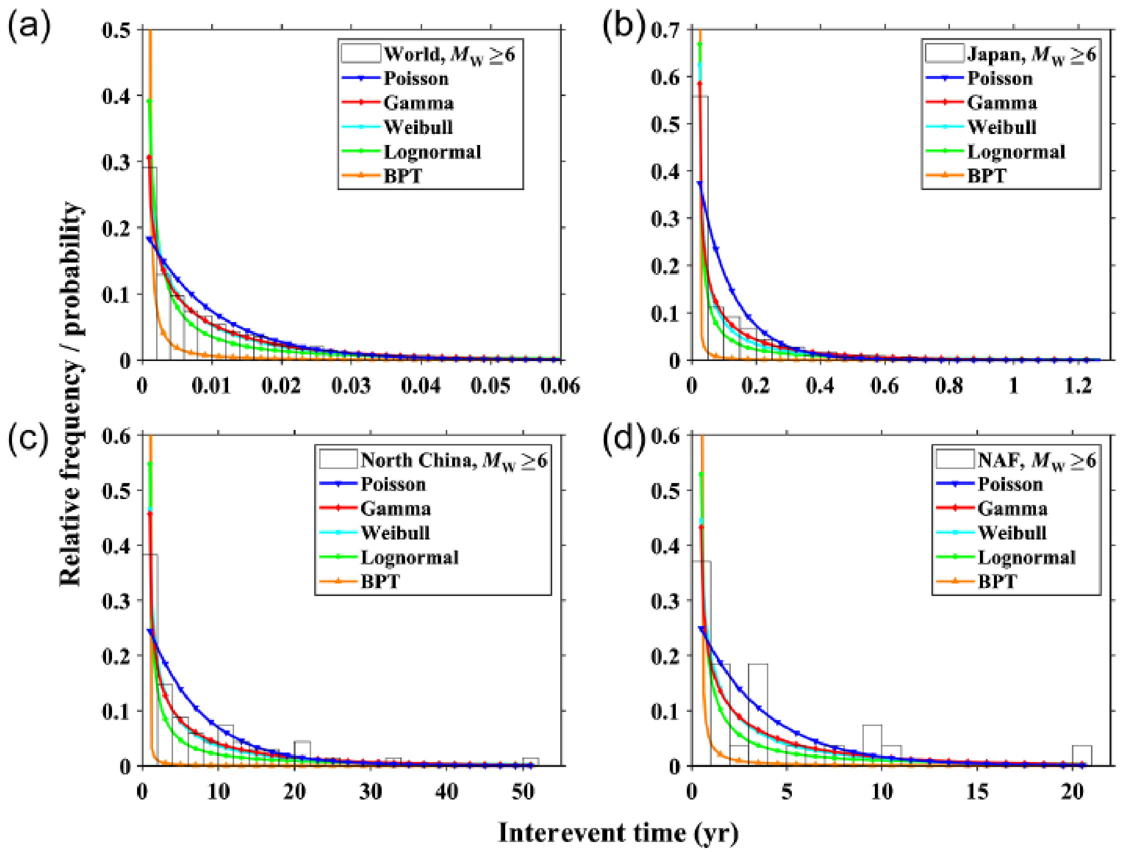


Figure 2.3. Comparison of the relative frequency (i.e. frequency divided by the total number of events) histograms of the distribution of interevent times for different fault systems with probabilities (with same lengths of interevent time windows as the data) predicted by five probability models (curves). The closer a curve is to the centers of the tops of the histogram boxes, the better the curve fits the data. (a) World, (b) Japan, (c) North China, and (d) North Anatolian Fault (NAF). The values of best-fitting parameters are shown in Table 2.4-2.7.

The Poisson model assumes that, although the mean interval between events is known for a sequence, the exact time of each event to occur is random (the Poisson process). The interevent time distribution of such a sequence follows an exponential distribution. The Poisson model is simple (a one-parameter model) and commonly used in probabilistic hazard analyses (Cornell, 1968; Field et al., 2014; Field et al., 2009; Fujiwara et al., 2006; WGCEP, 1988, 1995). It is a special case of the more generalized gamma and Weibull distributions (Chapter 1.4).

I investigated the distribution of interevent times of all the seismic sequences I studied. The relative frequency histograms show high frequency for events falling within short interevent times, and the frequency decreases rapidly with longer interevent times (Figure 2.3). The pattern can be generally fit by all the five probability models (Figure 2.3). I did Kolmogorov-Smirnov (KS) tests to statistically compare the fitting of these probability models with data and found that the gamma model fits the data best (detailed in Chapter 2.7). Both the gamma and Weibull models fit better than the Poisson model, whereas the lognormal and BPT models fit worse, as can also be seen in Figure 2.3. Both the gamma and Weibull models have higher probability for short interevent times than the Poisson model. In other words, the data have tighter clusters, or are burstier, than the prediction of the Poisson model.

The variation of the interevent times can be measured by the coefficient of variation (COV), called the aperiodicity, which is defined by $\frac{\sigma_\tau}{\mu_\tau}$, the ratio of the standard deviation of interevent times (σ_τ) to the mean of interevent times (μ_τ) (Goes, 1996; Kagan & Jackson, 1991; Salditch et al., 2019). For a sequence generated by a Poisson process, the COV value is 1. To measure the deviation from the Poisson model, we use a normalized COV, called the burstiness parameter (B) (Goh & Barabási, 2008), which is defined as

$$B = \frac{\frac{\sigma_\tau}{\mu_\tau} - 1}{\frac{\sigma_\tau}{\mu_\tau} + 1} = \frac{\sigma_\tau - \mu_\tau}{\sigma_\tau + \mu_\tau}. \tag{2.3}$$

The value of B ranges from −1 to 1. $B = -1$ corresponds to a perfectly periodic sequence, because its $\sigma_\tau = 0$ $(\mathrm{COV} = 1)$. When $\sigma_\tau \gg \mu_\tau$, $\mathrm{COV} \to \infty$ and $B \to 1$, which corresponds to the most bursty sequence. $B = 0$ corresponds to a sequence produced by an ideal Poisson process with $\sigma_\tau = \mu_\tau$ $(\mathrm{COV} = 0)$ (Figure 3.4a), for which the temporal activity pattern is random. Thus, a sequence is "bursty" when $0 < B < 1$ (Figure 2.4b) and "quasiperiodic" when $-1 < B < 0$ (Figure 2.4c).

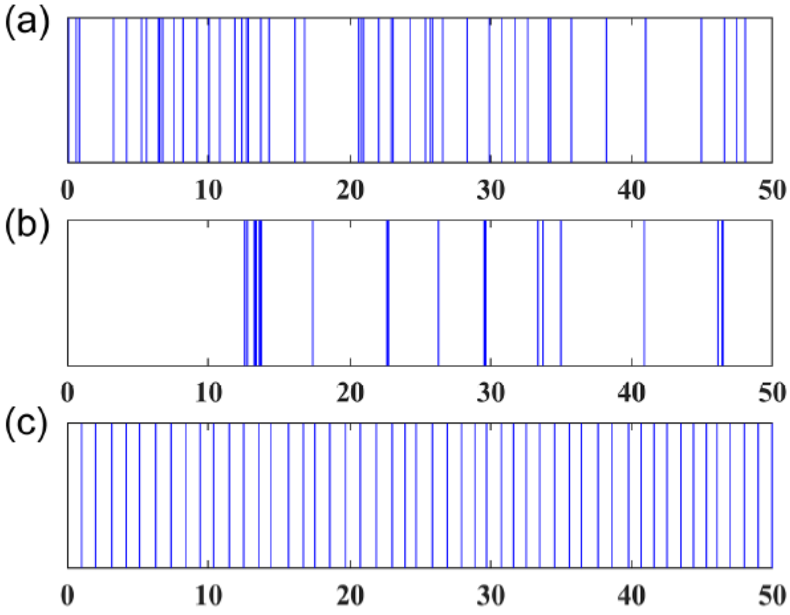


Figure 2.4. (a) A sequence of events generated by a Poisson model with $\mu = 1$. (b) A bursty sequence generated by the Weibull interevent time distribution with $a = 0.3$, $b = 2$. (c) An anti-bursty sequence generated by the Gaussian interevent time distribution with the mean $m = 1$ and the standard deviation $\sigma = 0.1$.

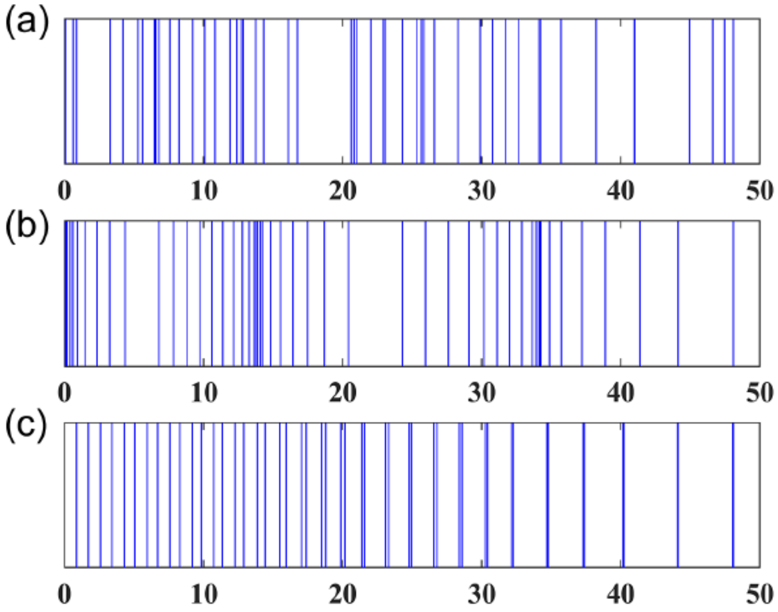


Figure 2.5. A bursty pattern can emerge through memory. The bursty pattern in (b) is obtained by shuffling the Poisson signals of (a) to increase the memory effect (the short (long) interevent times tend to follow short (long) ones). A more regular pattern, with negative memory (short (long) interevent times tend to be followed by long (short) ones), is obtained by the shuffling procedure (c). Note that sequences (a), (b), and (c) have identical interevent time distributions.

The sequences with the same B values can have a different order of events (Figure 2.5). This difference can be described by the memory coefficient m, defined as the correlation coefficient of consecutive interevent time values over a sequence. That is, given all pairs of consecutive interevent times (τ_i, τ_{i+1}) (Goh & Barabási, 2008),

$$m = \frac{1}{n_\tau - 1} \sum_{i=1}^{n_r - 1} \frac{(\tau_i - \mu_1)(\tau_{i+1} - \mu_2)}{\sigma_1 \sigma_2} \quad , \tag{2.4}$$

where n_τ is the number of interevent time measured from the sequence, μ_1 (μ_2) and σ_1 (σ_2) are the sample mean and sample standard deviation of the τ_i's (τ_{i+1}'s) ($i = 1, \ldots, n_\tau - 1$), respectively. The memory coefficient ranges from -1 to 1. For sequences with $m > 0$, a short (long) interevent time tends to be followed by a short (long) one. A sequence with $m < 0$ tends to have a short (long) interevent time followed by a long (short) one.

I calculated the burstiness parameters and memory coefficients for earthquake sequences of faults ($M_w \geq 6.5$) and regions ($M_w \geq 6$) where available catalogues have more than a dozen events. The results are presented in Tables 2.1 and 2.2. The burstiness parameters are close to but larger than 0 for all faults and regions studied before declustering (Tables 2.1 and 2.2), except that it is slightly negative for the Great Sumatran Fault. Thus, the earthquake sequences are close to but burstier than ones produced by ideal Poisson processes. Removing dependent earthquakes decreases the values of burstiness parameters. The memory coefficients are complicated: they tend to be positive for regional sequences (Table 2.1) but negative for sequences for individual faults (Table 2.2).

I also calculated the burstiness parameters and memory coefficients for ruptures on five segments of the southern San Andreas compiled by Williams et al. (2019) from paleoseismic data (Table 2.3). For ruptures on each individual segment, the burstiness parameters are negative, or quasiperiodic as concluded by Williams et al. (2019). However, these segments are not isolated from each other. Some ruptures, including the 1857 Fort Tejon earthquake, ruptured multiple segments. When these segments are viewed as a whole, the burstiness parameter is slightly positive (i.e., bursty), similar to results for large individual faults based on instrumental and historical catalogs (Table 2.2). The memory coefficients are negative or close to zero, also similar to the results of other large individual faults (Table 2.2).

Table 2.1: Statistical parameters of $M_w \geq 6$ earthquakes in different tectonic regions

	Maximum time interval (years)	Mean time interval (years)	Burstiness Parameter	Memory Coefficient	Total Events	Removed Events	Time Period
World	0.12 (0.15)	0.010 (0.016)	0.10 (0.02)	0.07 (0.02)	5269	2054	1964-2016
Japan	1.27 (2.26)	0.11 (0.24)	0.23 (0.08)	0.25 (0.04)	840	465	1926-2016
Taiwan	3.29 (3.66)	0.73 (1.39)	0.05 (-0.11)	-0.11 (-0.36)	71	33	1964-2016

East African Rift	14.57 (14.57)	1.66 (3.46)	0.31 (0.06)	0.00 (-0.02)	26	13	1964-2016
California	7.84 (7.84)	1.58 (2.11)	0.08 (-0.01)	0.35 (0.20)	53	13	1932-2016
New Zealand	6.40 (6.40)	1.21 (1.49)	0.09 (0.02)	-0.02 (-0.13)	61	11	1942-2016
Tibet	2.35 (2.35)	0.46 (0.66)	0.06 (-0.08)	0.21 (0.26)	111	33	1964-2016
Xinjiang	4.50 (4.50)	1.37 (1.95)	0.03 (-0.14)	-0.23 (-0.26)	38	11	1964-2016
North China	51.41 (51.41)	7.11 (8.63)	0.12 (0.04)	0.31 (0.25)	69	12	1500-2016

*Results in the parentheses are obtained after declustering using the Gardner-Knopoff method (Gardner & Knopoff, 1974). The ISC-GEM catalog for earthquake magnitudes larger than 6 is complete since 1964 (Michael, 2014). The local earthquake catalogs of Japan, California, and New Zealand start in 1926, 1934, and 1942, respectively, so are likely complete for $M \geq 6$ events. The North China catalog is complete for $M \geq 6$ events since 1291 (Huang et al., 1994). We chose to use the catalog since 1500, to be conservative.

Table 2.2: Statistical parameters of $M_w \geq 6.5$ earthquakes on individual faults

	Maximum time interval (years)	Mean time interval (years)	Burstiness Parameter	Memory Coefficient	Total Events	Removed Events	Time Period
Great Sumatran Fault	18.69 (18.69)	6.23 (6.23)	-0.05 (-0.05)	-0.15 (-0.15)	16	0	1904-2016
North Anatolian Fault	23.92 (23.92)	6.38 (6.88)	0.07 (0.04)	-0.06 (-0.13)	15	1	1904-2016
Sagaing Fault	34.47 (34.47)	10.05 (12.56)	0.01 (-0.10)	0.23 (-0.03)	11	2	1904-2016
Xianshuihe Fault	76.72 (76.72)	23.23 (23.23)	0.02 (0.02)	-0.34 (-0.34)	12	0	1700-2016

*Results in the brackets are obtained after declustering using the Gardner-Knopoff method (Gardner & Knopoff, 1974).

Table 2.3: Statistical parameters of ruptures on five different segments of southern SAF using the best-estimated time based on data compiled by Williams et al. (2019)

	Maximum time interval (years)	Mean time interval (years)	Burstiness Parameter	Memory Coefficient	Total Events
SAF Wrightwood	166	94.5	-0.36	0.03	15
SAF Carizzo	352	203.8	-0.38	-0.42	6
SAF Mission Creek	271	212.8	-0.58	-0.47	5
SAF Pallet Creek	305	141.1	-0.27	-0.36	9
SAF Big Bend	385	128.2	-0.07	0.05	11
All segments	116	29.4	0.003	-0.14	46

2.3 Possible tectonic factors and causes

The Devil's Staircase pattern of large earthquakes is characterized by clusters of events separated by longer and irregular intervals of quiescence. Here we examine how these features may be related to tectonic factors.

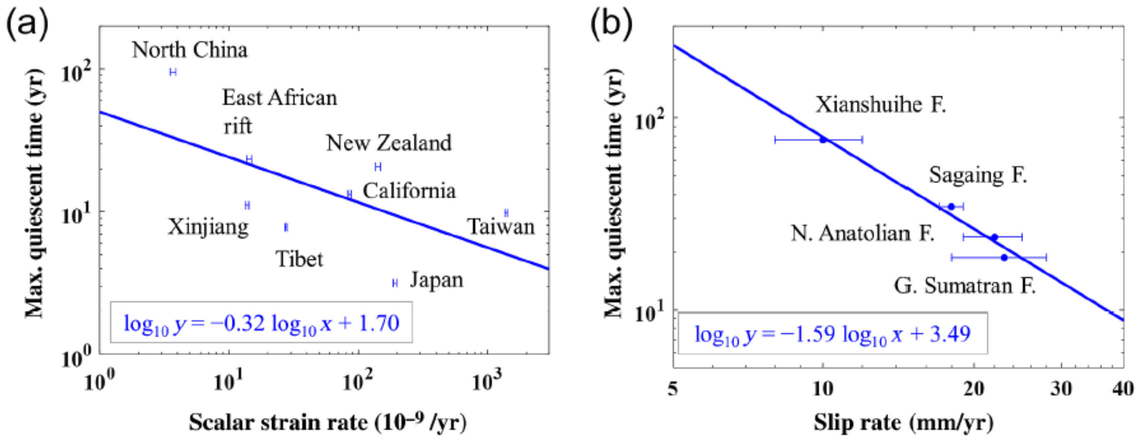


Figure 2.6. (a) The longest quiescent intervals for $M \geq 6.5$ earthquakes in various tectonic regions vs. the magnitude of strain rates. The scalar strain rate data are from Kreemer et al. (2014). The scalar strain rate is defined as $\dot{E} = \sqrt{\dot{\varepsilon}_{\phi\phi}\dot{\varepsilon}_{\phi\phi} + \dot{\varepsilon}_{\lambda\lambda}\dot{\varepsilon}_{\lambda\lambda} + 2\dot{\varepsilon}_{\lambda\phi}\dot{\varepsilon}_{\lambda\phi}}$, where ϕ and λ are longitude and latitude. (b) The longest quiescent intervals for $M \geq 6.5$ earthquakes on individual faults vs. the slip rates. Sources of slip rates: Genrich et al. (2000) for the Great Sumatran fault; Straub et al. (1997) and McClusky et al. (2000) for the North Anatolian fault, Vigny et al. (2003) for the Sagaing Fault, and Shen et al. (2005) for the Xianshuihe Fault. Solid lines are least-square fitting. The insets show the regression formulas.

The length of quiescent intervals between earthquake clusters seems to be inversely related to tectonic loading rates (strain rates for tectonic regions or slip rates for individual faults) (Figure 2.6): lower loading rates correlate to longer quiescent intervals. The longest quiescent intervals also increase with the mean recurrence intervals, which is an indicator of average tectonic loading rate (Figure 2.7). This is consistent with the very long intervals between earthquake clusters in Australia and other stable continents (Clark et al., 2014), although we did not include those sequences in our analysis because of limited events in these sequences.

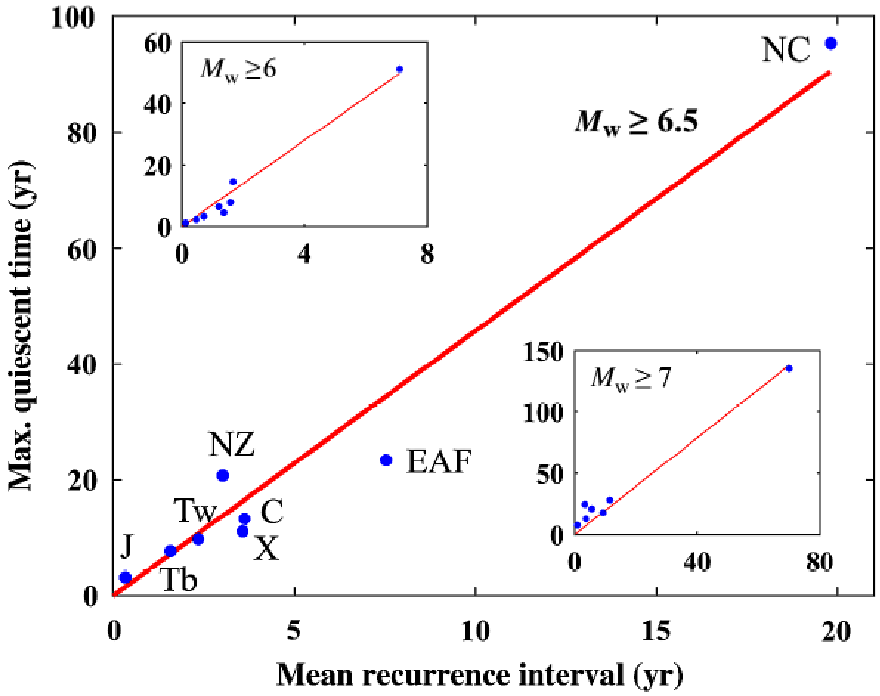


Figure 2.7. Relationship between mean recurrence interval and maximum quiescent time of $M_w \geq 6.5$ earthquakes in different tectonic regions (J: Japan; Tb: Tibet; Tw: Taiwan; NZ: New Zealand; X: Xinjiang region, China; C: California; EAF: East African Rift; and NC: North China). The insets show similar patterns for different cut-off magnitudes ($M_w \geq 6$ and $M_w \geq 7$).

For the active periods of clustered events, the tectonic control is more complex. The average number of events in each cluster seems to be larger in tectonically active regions with higher loading rates and shorter mean recurrence time (Figure 2.8a-b). The average lengths of the seismically active periods, however, are not clearly correlated with tectonic loading rate (Figure 2.8c-d).

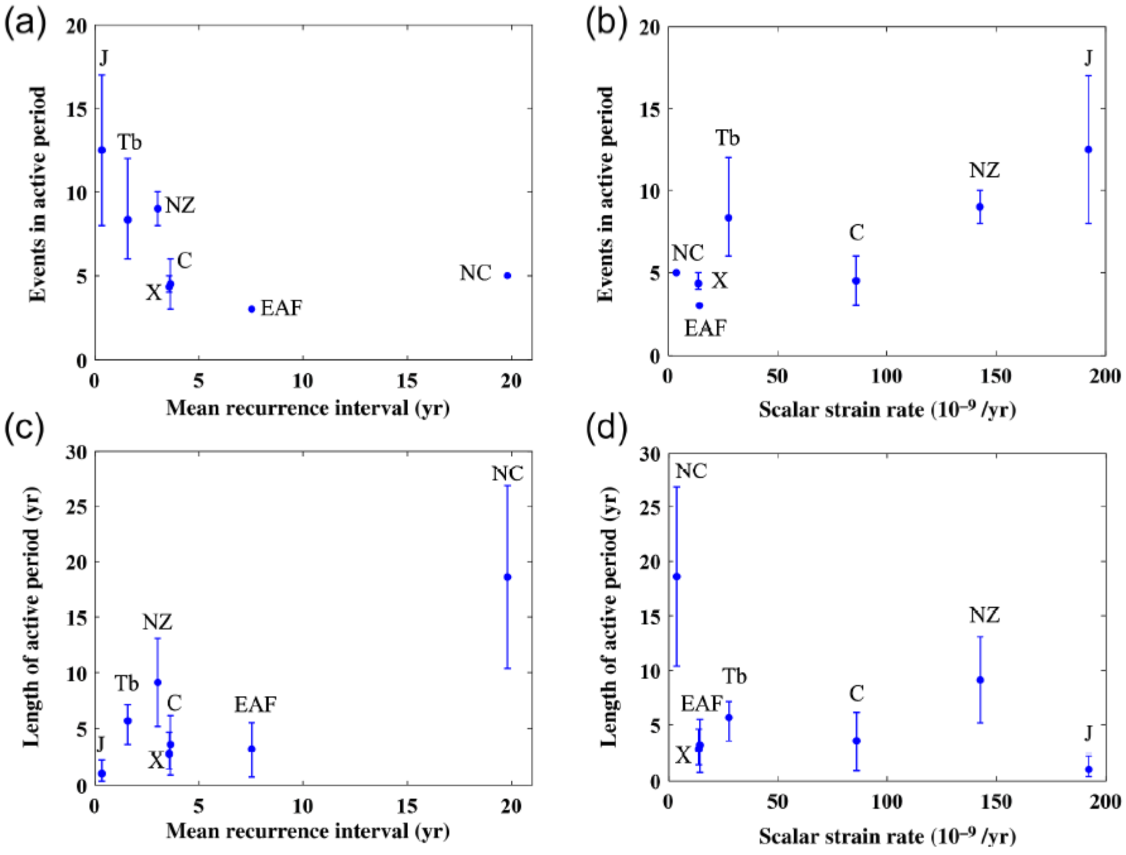


Figure 2.8. Relationship between features of earthquake clusters and tectonic factors for $M_w \geq 6.5$ earthquakes in different regions. (a-b) The number of events in active periods (clusters) vs. mean recurrence intervals and regional strain rate, respectively. (c-d) The length of active periods vs. mean recurrence intervals and regional strain rate, respectively. Abbreviations for tectonics regions are explained in Figure 2.7.

The physical causes of temporal clustering of earthquakes may be multiple, including both earthquake-induced changes of frictional property and stress transfer. Viscoelastic stress relaxation and fault interaction are two relatively well-known mechanisms of earthquake-induced stress transfer (Freed & Lin, 1998; Li et al., 2009; Luo & Liu, 2012; Stein & Liu, 2009). To explore how these two mechanisms may contribute to the Devil’s Staircase patterns of earthquakes, my co-author Dr. Gang Luo developed a simple numerical model with one or multiple faults within a region (Figure 2.9a), using a visco-elastoplastic finite element code Dr. Gang Luo developed (Luo & Liu, 2010, 2012, 2018). Details of the numerical scheme are presented in these references. The model work is mainly done by Dr. Gang Luo, so I omit the details about the model and focus on the results. The details of the model are introduced in Chen et al. (2020).

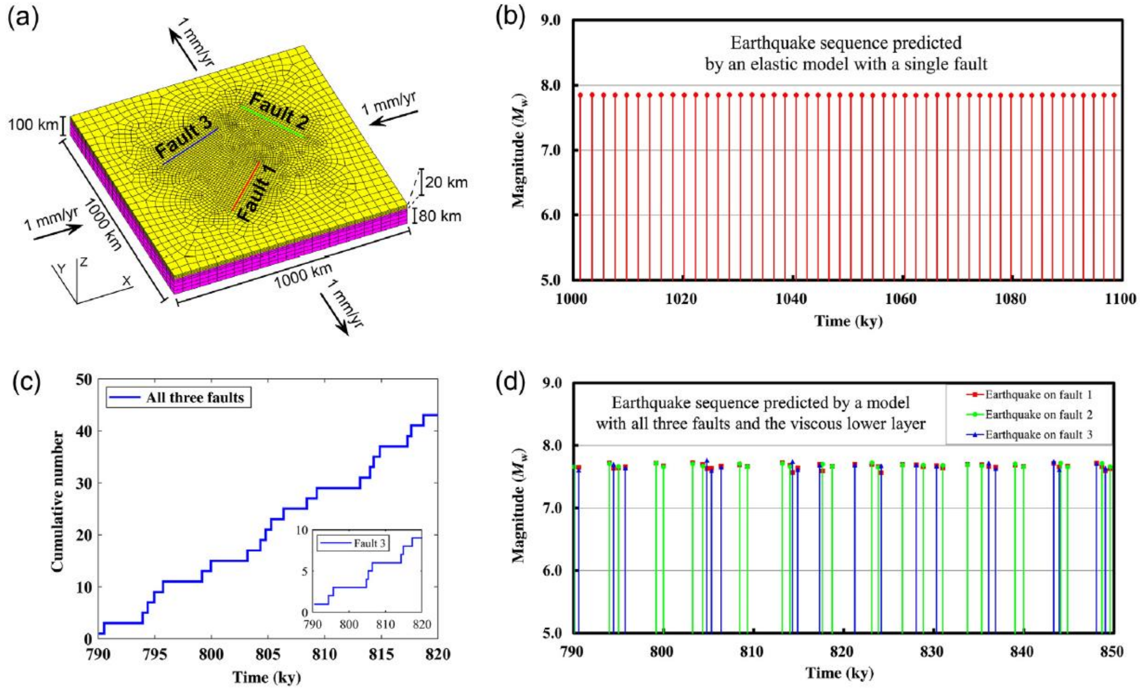


Figure 2.9. Numerical simulation of intraplate earthquakes. (a) Model setting and numerical mesh. The seismogenic upper crust is elastoplastic, sitting on top of a viscoelastic layer (lower crust and upper mantle). The model domain is loaded by the imposed velocity (vectors) on the sides. (b) Earthquake sequence predicted by a model that includes only one fault (Fault 1) and without the viscoelastic lower layer. Time is

since the beginning of the numerical experiment. (c) Temporal patterns of earthquakes for the fault system (all three faults). Inset is the pattern on Fault 3. (d) Earthquake sequence predicted by a model with the three faults and both an elastoplastic upper layer and a viscoelastic lower layer.

When only one fault is included in the model and the entire model domain is assumed to be elastic, steady loading imposed on the sides of the model domain produces repeated failure of the fault with regular recurrence times (Figure 2.9b), as would be predicted by the elastic rebound model. However, with three arbitrarily oriented faults included in the model and the model domain includes both an elastoplastic (seismogenic) upper crust and a viscoelastic lower crust, the failure patterns show clusters of events separated by longer intervals of inactivity (Figure 2.9d). In this case, stress on each fault is perturbed by failures of other faults and by viscous relaxation of the lower crust that transfers stress back to the upper crust (Freed & Lin, 2001; Li et al., 2009). The resulting temporal patterns of earthquakes, either for the whole system (three faults) or on a single fault, have the features of the Devil’s Staircase (Figure 2.9c). Thus, stress changes from viscous relaxation and fault interaction are likely important factors contributing to the clustering of earthquakes.

2.4 Implications for hazard analysis

The Devil’s Staircase patterns of large earthquakes have important implications for earthquake hazard assessment. Firstly, the mean recurrence time, a key parameter for seismic hazard analysis, can vary drastically depending on which portion of the sequence the catalog represents. This can be a serious concern in hazard assessment, because catalogs for large earthquakes are often too short to reflect their complete temporal pattern, and it is difficult to know whether the few events in a catalog occurred within an earthquake cluster or spanned both clusters and quiescent intervals. For the same reason, we need to be cautious when assessing an event is “overdue” just because the time measured from the previous event has passed some “mean recurrence time” based on an incomplete catalog, as discussed by Salditch et al. (2019).

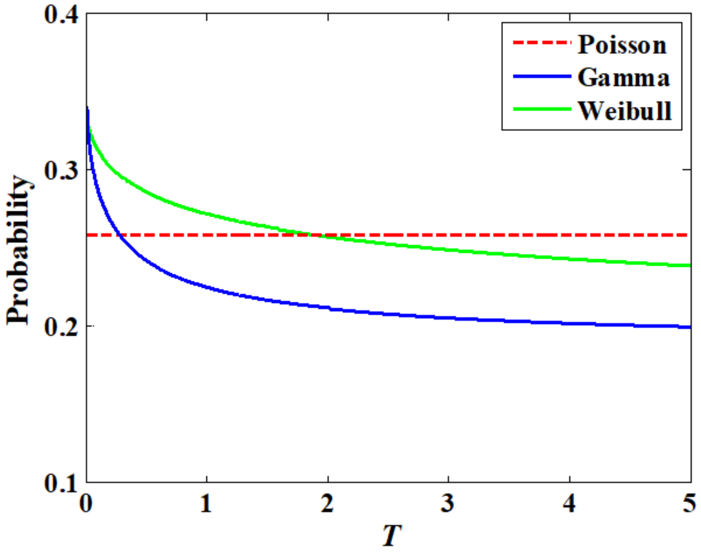


Figure 2.10. Comparation of conditional probabilities between the Poisson, gamma and Weibull models. The prediction interval $\Delta T = 0.3$. All models have a mean of 1. The gamma model and Weibull model have a coefficient of variation of 1.2.

Secondly, probability seismic hazard analysis usually uses the conditional probability, which is defined by

$$\mathrm{P}(T \leq \tau \leq T + \Delta T \mid \tau > T) = \frac{\int_{T}^{T+\Delta T} f(\tau)\, d\tau}{\int_{T}^{\infty} f(\tau)\, d\tau}, \tag{2.5}$$

where τ is the interevent time, T is the time interval since the last earthquake in a catalog, ΔT is the prediction interval, and $f(\tau)$ is the probability density function (PDF) of the interevent time. It gives the probability of the next earthquake in the interval $(T, T + \Delta T)$ when knowing that no earthquake occurred before this interval since the last earthquake. For the Poisson model, which is commonly used in seismic hazard analysis (Field et al., 2015; WGCEP, 1988, 1995, 2003), the conditional probability is a constant (Chapter 2.7.4), so the Poisson model is also called as time-independent model. In other words, each event in the sequence is independent of other events. However, our results suggest that most earthquake sequences, especially when dependent events are not excluded, are burstier than a Poisson sequence and may be better fit by the gamma or Weibull distributions. The conditional probability of both the gamma and Weibull models is higher than of the Poisson model for a small T but decreases as T increase (Figure 2.10). In other word, the probability of repeating events soon after a large earthquake is higher than that predicted by the commonly used Poisson model. This is clear from the histograms of earthquakes sequences (Figure 2.3). These repeating events could be aftershocks or events triggered by stress transferred from the preceding events. Recent examples are plenty, including the 1999 Hector Mine earthquake (M_w 7.1) following the 1992 Lander earthquake (M_w 7.3) in southern California (Freed & Lin, 2001), and the 2013 Lushan earthquake (M_w 6.6) following the 2008 Wenchuan earthquake (M_w 7.9) on the Longmenshan fault in eastern Tibetan Plateau (Liu et al., 2014). Kagan and Jackson

(1999) shows that doublets of large shallow earthquakes, with partly overlapped rapture zones and significantly shorter interevent time than the time needed for strain accumulation, are pervasive worldwide.

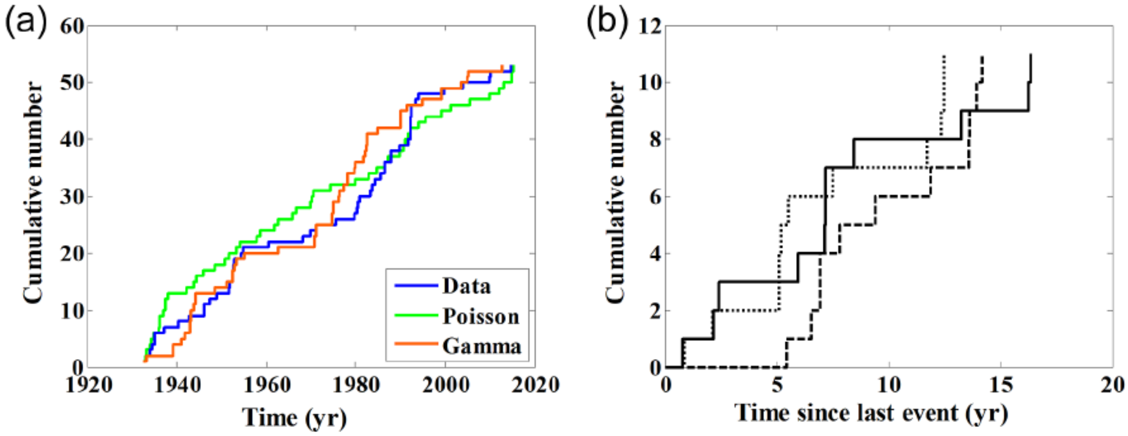


Figure 2.11. (a) Comparison of the temporal patterns of large ($M_w \geq 6$) earthquakes in California with those predicted by the best-fitting probability functions. (b) Three forecasted sequences for future $M_w \geq 6$ events in California after the last event in the catalog (2014/08/24), based on the best-fitting gamma probability distribution (Table 2.8).

Finally, the Devil's Staircase patterns are characteristic of complex dynamic systems, which are nonlinear systems composed of many components (here faults and fault segments) that interact with each other to produce nonlinear system behaviors. Small changes in some components can lead to big changes in the system. Short-term fault behavior in such a fault system is much more difficult to predict than faults experiencing cyclic loading and release. Nonetheless, one can try to forecast the system behavior of such fault systems. The best-fitting probability distribution function of a sequence can be inverted to generate pseudorandom interevent times that obey the same probability distribution (Devroye, 1986). For example, Figure 2.11a shows the instrumentally recoded $M_w \geq 6$ earthquakes in California and the sequences generated by the best-fitting Poisson and gamma probability distributions. The gamma model fits the general features of the data better than the Poisson model.

Inverting the best-fitting gamma distribution function, we can forecast the temporal feature of future $M_w \geq 6$ earthquake sequence in California (Figure 2.11b). The forecast is non-unique; nonetheless, it indicates the statistical behavior of the regional fault system. Three forecasted sequences are shown here. They suggest that the future sequence of earthquakes will likely occur in clusters separated by relative long and variable quiescent intervals, i.e., Devil's Staircase.

2.5 Discussion

In earthquake studies, much effort has been devoted to establishing the recurrence times of large earthquakes on a given fault to assess the probability of the next event occurring within a certain range of times (Biasi et al., 2015; Biasi et al., 2002; McCalpin & Nishenko, 1996). The elastic rebound model explains the basic physics of earthquake recurrence and predicts periodic or quasi-periodic earthquakes.

Establishing the temporal patterns of large earthquakes on individual faults, however, is difficult because large earthquakes are infrequent, and the catalogs are usually too short and incomplete. Quasi-periodic recurrence of large earthquakes has been reported on the Alpine Fault, in the south-central Chile subduction zone, on the southern San Andreas Fault, and on the intraplate Loma Blanca fault (Berryman et al., 2012; Moernaut et al., 2018; Scharer et al., 2010; Williams et al., 2017). In the first two cases, large earthquakes occurred on simple fault structures with a steadily high loading rate, so the elastic rebound model might work (Berryman et al., 2012). In the last two cases, although the earthquake sequences appear to be quasi-periodic, they also contain notable clusters (Weldon et al., 2004; Williams et al., 2017). When regional fault systems are considered, the earthquake sequences generally deviate from quasi-periodic patterns. Instead, they are bursty, with clusters of events separated by relative long and irregular intervals of quiescence as Devil's Staircases.

The complexities of the temporal patterns of earthquakes may be partly attributed to data limitations. Instrumental catalogs may not be long enough to show statistically robust temporal patterns. For large earthquakes in continental interiors where large earthquakes are less frequent than in plate boundary zones, historic and paleoseismic data may be needed, but they come with the associated uncertainties. Some of the earthquake clustering can certainly be attributed to aftershocks or foreshocks, and we have shown that declustering would reduce the burstiness in earthquake sequences. However, large aftershocks or foreshocks are as important as mainshocks when hazard is concerned, therefore should be included in hazard analysis.

On the other hand, there are good reasons not to expect large earthquakes to be quasi-periodic, because they violate the key premises of the elastic rebound model. Firstly, the loading rate may not be constant, even for plate boundary faults (Benedetti et al., 2013; Friedrich et al., 2003; Ratzov et al., 2015). Although the rates of relative plate motion are steady during the past few million years (DeMets et al., 1994), plate boundary faults are usually not a single fault plane but a system of fault branches and segments.

The San Andreas Fault, for example, consists of a complex system of subparallel faults in southern California that shares the loading from the relative Pacific-North American plate motion. Large earthquakes on one fault of this fault system could affect stress and loading rates on the other faults (Dolan et al., 2007; Luo & Liu, 2012). For intraplate faults, the loading rates are lower and more variable than for plate boundary faults, because tectonic loading from plate boundaries is collectively accommodated by a widespread network of faults (Li et al., 2009; Liu & Stein, 2016). On each individual fault, the loading rate is likely variable, affected by previous earthquakes on the fault, earthquakes on other faults in the system, and transient local stress perturbations such as erosion (Calais et al., 2010).

Secondly, the elastic rebound model assumes cyclic strain accumulation and release on a given fault plane, but large earthquakes often rupture multiple and variable fault segments (e.g., the 2001 M_w 7.8 Kunlun earthquake, China) and faults (e.g., the 2016 M_w 7.8 Kaikoura earthquake, New Zealand), making the concept of periodic release of strain energy on a defined fault plane inadequate. When seismicity in a regional network of faults is considered, the simple elastic rebound model does not apply.

The nonlinear and complex temporal patterns of large earthquakes have long been observed (Clark et al., 2012; Sieh et al., 1989; Xu & Deng, 1996). Some called these sequences supercycles (Salditch et al., 2019). We have shown that the sequences of large earthquakes on fault systems of different scales show similar features of the Devil’s Staircase, with clusters of events separated by long and variable intervals of quiescence. This suggests that large ruptures of faults behave like nonlinear complex dynamic systems, as noted in previous studies (Calais et al., 2010; Li et al., 2009; Liu & Stein, 2016).

This is not surprising, as fault systems in nature are known to be complex systems (Turcotte & Malamud, 2002). One evidence is the Gutenberg-Richter frequency-magnitude relationship, a power-law (fractal) distribution that is valid both regionally and globally (Turcotte, 1997). Complex systems are nonlinear dynamic systems composed of many components (here faults and fault segments) that interact with each other, producing nonlinear system behaviors. The clustering of earthquakes could arise from such interactions. Fault interaction includes transfer of static stress (the Coulomb stress) from a ruptured fault (or fault segment) to neighboring faults (or fault segments), as well as perturbation of regional loading conditions by local fault ruptures. It also includes stress transfer from the ductile lower crust to the upper crust by viscoelastic stress relaxation. Large earthquakes may also be triggered far from aftershock zones by dynamic stresses when seismic waves propagate. Fault interaction in a complex dynamic system is the key to understanding spatiotemporal variations of large earthquakes, especially those in continental interiors (Liu & Stein, 2016).

For a complex dynamic system of faults, the prediction or forecasting of fault ruptures is more difficult than expected from existing models based on elastic rebound. Nonetheless, we could try to characterize their temporal patterns to learn about the system behavior. We have shown that bursty sequences can be better fit by the gamma distribution function than the commonly used Poisson model. One implication is that soon after a large earthquake, the chance of having another one in the system is higher than that predicted by models assuming Poissonian occurrences of earthquakes, as suggested by numerous recent large earthquake sequences.

2.6 Conclusions

I found that the temporal patterns of large earthquakes, on individual faults or regional fault systems, show clusters of events separated by relative long and variable intervals of seismic quiescence. Such patterns are characteristic of the Devil’s Staircase, a fractal property of nonlinear complex systems. For these earthquake sequences, the mean recurrence intervals, often estimated from records of the few most recent events, can vary significantly depending on whether these events are clustered in a relatively short active period or cover a spectrum of clusters and quiescent intervals.

The lengths of the quiescent intervals between clusters are inversely related to tectonic loading (or slip) rates. They are usually a few times longer than the periods of clustered events in tectonically active regions, but can be thousands of years or longer in stable continental interiors. The earthquake clusters likely result from earthquake-induced stress transfer, including fault interaction and viscoelastic relaxation. The clustered events could include aftershocks and foreshocks. Because large aftershocks and foreshocks are as important as mainshocks in terms of seismic hazard, they should be included in seismic hazard analysis.

The burstiness of earthquake sequences can be statistically characterized using the burstiness parameter and the memory coefficient. The burstiness parameters for most earthquake sequences I studied have slightly positive values, meaning that these sequences are bustier than the Poisson process. The interevent times of most earthquake sequences can be better fitted by the gamma probability distribution than by the Poisson model. Models assuming a bursty gamma distribution (with positive burstiness parameter) predicts higher probabilities than models assuming a Poisson process for repeating events soon after a large earthquake, as indicated by numerous recent sequences of large earthquakes.

The Devil’s Staircase distribution of large earthquakes implies that large fault ruptures, which often involve multiple segments or faults, behave as nonlinear complex systems. Prediction or forecasting of large earthquakes in such systems are much more difficult than for earthquakes on an isolated fault with simple structures under fast and steady tectonic loading. Thus, studying large earthquakes requires a system approach, rather than focusing only on stress accumulation and release on individual faults.

2.7 Supplementary materials

In this supplement, we describe Kolmogorov–Smirnov test, Best-fitting parameters for earthquake sequences in Figure 2.3, Best-fitting parameters for California earthquake sequences, and Conditional probabilities used in Figure 2.10.

2.7.1 Kolmogorov–Smirnov test

Kolmogorov–Smirnov (KS) test is used to check the goodness of fit for different probability distributions. The KS statistic is defined as

$$D_{ks} = \max_{x} |F_{obs}(x) - F(x)| \tag{2.6}$$

where x is a random variable, $F_{obs}(x)$ is the empirical cumulative distribution function of the data, and $F(x)$ is the cumulative distribution of a hypothetic distribution (Massey Jr, 1951). The KS statistic measures the maximum difference between an empirical and a hypothetic distribution. The smaller the KS statistic is, the better the distribution fits the data.

2.7.2 Best-fitting parameters for earthquake sequences in Figure 2.3

Table 2.4: Results of best-fitting parameters of different distributions for data in Figure 2.3a (global $M_w \geq 6$ earthquakes)

Distribution	First Parameter	Second Parameter	KS statistic
Exponential	$\mu = (9.87 \pm 0.27) \times 10^{-3}$	-	0.111
Gamma	$a = 0.59 \pm 0.02$	$b = (1.67 \pm 0.08) \times 10^{-2}$	0.027
Weibull	$a = 0.72 \pm 0.02$	$b = (8.16 \pm 0.32) \times 10^{-3}$	0.040
Lognormal	$\mu = -5.67 \ \pm 0.05$	$\sigma = 1.99 \pm 0.04$	0.113
BPT	$\mu = 9.87 \times 10^{-3}$	$\alpha = 12.6$	0.606

* The uncertainties of parameters are 95% confidence limits. For the BPT model, the uncertainties of parameters are not evaluated because the errors are too big.

Table 2.5: Results of best-fitting parameters of different distributions for data in Figure 2.3b ($M_w \geq 6$ earthquakes in Japan)

Distribution	First Parameter	Second Parameter	K-S statistic
Exponential	$\mu = 0.11 \pm 0.01$	-	0.287
Gamma	$a = 0.33 \pm 0.025$	$b = 0.33 \pm 0.05$	0.048
Weibull	$a = 0.45 \pm 0.03$	$b = (5.23 \pm 0.82) \times 10^{-2}$	0.083

Lognormal	$\mu = -4.32 \pm 0.21$	$\sigma = 3.04 \pm 0.15$	0.113
BPT	$\mu = 0.11$	$\alpha = 43.4$	0.627

* The errors of parameters are 95% confidence limits. For the BPT model, the uncertainties of parameters are not evaluated because the errors are too big.

.

Table 2.6: Results of best-fitting parameters of different distributions for data in Figure 2.3c ($M_w \geq 6$ earthquakes in North China)

Distribution	First Parameter	Second Parameter	KS statistic
Exponential	$\mu = 7.11 \pm 1.69$	-	0.161
Gamma	$a = 0.40 \pm 0.11$	$b = 17.83 \pm 8.28$	0.119
Weibull	$a = 0.54 \pm 0.11$	$b = 4.81 \pm 2.21$	0.131
Lognormal	$\mu = 0.31 \pm 0.76$	$\sigma = 3.18 \pm 0.54$	0.212
BPT	$\mu = 7.11$	$\alpha = 91.20$	0.792

* The errors of parameters are 95% confidence limits. For the BPT model, the uncertainties of parameters are not evaluated because the errors are too big.

.

Table 2.7: Results of best-fitting parameters of different distributions for data in Figure 2.3d ($M_w \geq 6$ earthquakes on the NAF)

Distribution	First Parameter	Second Parameter	KS statistic
Exponential	$\mu = 3.49 \pm 1.32$	-	0.200
Gamma	$a = 0.45 \pm 0.20$	$b = 7.78 \pm 5.54$	0.094
Weibull	$a = 0.59 \pm 0.18$	$b = 2.46 \pm 1.66$	0.112

Lognormal	$\mu = -0.19 \pm 0.99$	$\sigma = 2.64 \pm 0.72$	0.183
BPT	$\mu = 3.49$	$\alpha = 20.93$	0.688

* The errors of parameters are 95% confidence limits. For the BPT model, the uncertainties of parameters are not evaluated because the errors are too big.

2.7.3 Best-fitting parameters for California earthquake sequences

Table 2.8: Results of best-fitting parameters of different distributions for $M_w \geq 6$ California earthquakes (1932-2016)

Distribution	First Parameter	Second Parameter	KS statistic
Exponential	$\mu = 1.58 \pm 0.43$	-	0.168
Gamma	$a = 0.44 \pm 0.14$	$b = 3.58 \pm 1.85$	0.130
Weibull	$a = 0.58 \pm 0.13$	$b = 1.14 \pm 0.56$	0.139
Lognormal	$\mu = -1.01 \pm 0.76$	$\sigma = 2.79 \pm 0.54$	0.216
BPT	$\mu = 1.58$	$\alpha = 27.91$	0.716

* The errors of parameters are 95% confidence limits. For the BPT model, the uncertainties of parameters are not evaluated because the errors are too big.

2.7.4 Conditional probabilities used in Figure 2.10

For an exponential distribution, its conditional probability is

$$\mathrm{P}(T \leq \tau \leq T + \Delta T \mid \tau > T) = \frac{\int_T^{T+\Delta T} \frac{1}{\mu} e^{-\frac{\tau}{\mu}} \, d\tau}{\int_T^{\infty} \frac{1}{\mu} e^{-\frac{\tau}{\mu}} \, d\tau}$$

$$= \frac{\left(-e^{-\frac{T+\Delta T}{\mu}}\right) - \left(-e^{-\frac{T}{\mu}}\right)}{0 - \left(-e^{-\frac{T}{\mu}}\right)}$$

$$= 1 - e^{-\frac{-\Delta T}{\mu}} = \text{constant.} \tag{2.7}$$

For a gamma distribution, its conditional probability is

$$\mathrm{P}(T \le \tau \le T + \Delta T \mid \tau > T) = \frac{\int_T^{T+\Delta T} \frac{1}{b^a\,\Gamma(a)} \tau^{a-1} e^{-\frac{\tau}{b}}\, d\tau}{\int_T^{\infty} \frac{1}{b^a\,\Gamma(a)} \tau^{a-1} e^{-\frac{\tau}{b}}\, d\tau}$$

$$= \frac{\gamma\left(a, \frac{T+\Delta T}{b}\right) - \gamma\left(a, \frac{T}{b}\right)}{\Gamma(a) - \gamma\left(a, \frac{T}{b}\right)}$$

$$= 1 - \frac{\Gamma(a) - \gamma\left(a, \frac{T+\Delta T}{b}\right)}{\Gamma(a) - \gamma\left(a, \frac{T}{b}\right)} \tag{2.8}$$

where $\gamma(a, x)$ is the lower incomplete gamma function, which is defined as $\gamma(a, x) = \int_0^x t^{a-1}\, e^{-t} dt$. When $T \to \infty$, the equation (2.8) becomes

$$\lim_{T\to\infty} \mathrm{P}(T \le \tau \le T + \Delta T \mid \tau > T) = 1 - \lim_{T\to\infty} \frac{\Gamma(a) - \gamma\left(a, \frac{T+\Delta T}{b}\right)}{\Gamma(a) - \gamma\left(a, \frac{T}{b}\right)}$$

$$= 1 - \lim_{T\to\infty} \frac{\frac{d}{dT}\gamma\left(a, \frac{T+\Delta T}{b}\right)}{\frac{d}{dT}\gamma\left(a, \frac{T}{b}\right)}$$

$$= 1 - \lim_{T\to\infty} \frac{(T+\Delta T)^{a-1} e^{-\frac{T+\Delta T}{b}}}{T^{a-1} e^{-\frac{T}{b}}}$$

$$= 1 - e^{-\frac{-\Delta T}{b}} = \text{constant.}$$

For a Weibull distribution, its conditional probability is

$$\mathrm{P}(T \le \tau \le T + \Delta T \mid \tau > T) = \frac{\int_T^{T+\Delta T} \frac{a}{b}\left(\frac{\tau}{b}\right)^{a-1} e^{-\left(\frac{\tau}{b}\right)^a}\, d\tau}{\int_T^{\infty} \frac{a}{b}\left(\frac{\tau}{b}\right)^{a-1} e^{-\left(\frac{\tau}{b}\right)^a}\, d\tau}$$

$$= \frac{\left(-e^{-\left(\frac{T+\Delta T}{b}\right)^a}\right) - \left(-e^{-\left(\frac{T}{b}\right)^a}\right)}{0 - \left(-e^{-\left(\frac{T}{b}\right)^a}\right)}$$

$$= 1 - \frac{e^{-\left(\frac{T+\Delta T}{b}\right)^a}}{e^{-\left(\frac{T}{b}\right)^a}} \tag{2.9}$$

When $T \to \infty$, the equation (2.9) becomes

$$\lim_{T\to\infty} \mathrm{P}(T \le \tau \le T + \Delta T \mid \tau > T) = 1 - \lim_{T\to\infty} \frac{e^{-\left(\frac{T+\Delta T}{b}\right)^a}}{e^{-\left(\frac{T}{b}\right)^a}}$$

$$= 1 - \lim_{T\to\infty} e^{-\left[\left(\frac{T+\Delta T}{b}\right)^a - \left(\frac{T}{b}\right)^a\right]}$$

$$= \begin{cases} 0, & a < 1 \\ 1 - e^{-\frac{-\Delta T}{b}}, & a = 1 \\ 1, & a > 1 \end{cases}$$

References

Abaimov, S., Turcotte, D., Shcherbakov, R., Rundle, J., Yakovlev, G., Goltz, C., & Newman, W. (2008). Earthquakes: Recurrence and Interoccurrence Times. *Pure and Applied Geophysics, 165*(3-4), 777-795.

Bak, P. (1996). *How nature works: the science of self-organized criticality*: Springer Science & Business Media.

Benedetti, L., Manighetti, I., Gaudemer, Y., Finkel, R., Malavieille, J., Pou, K., et al. (2013). Earthquake synchrony and clustering on Fucino faults (Central Italy) as revealed from in situ 36Cl exposure dating. *Journal of Geophysical Research, Solid Earth*, *118*(9), 4948-4974.

Berryman, K. R., Cochran, U. A., Clark, K. J., Biasi, G. P., Langridge, R. M., & Villamor, P. (2012). Major earthquakes occur regularly on an isolated plate boundary fault. *Science, 336*(6089), 1690-1693.

Biasi, G. P., Langridge, R. M., Berryman, K. R., Clark, K. J., & Cochran, U. A. (2015). Maximum-likelihood recurrence parameters and conditional probability of a ground-rupturing earthquake on the Southern Alpine Fault, South Island, New Zealand. *Bulletin of the Seismological Society of America*, *105*(1), 94-106.

Biasi, G. P., Weldon, R. J., Fumal, T. E., & Seitz, G. G. (2002). Paleoseismic event dating and the conditional probability of large earthquakes on the southern San Andreas fault, California. *Bulletin of the Seismological Society of America*, *92*(7), 2761-2781.

Calais, E., Freed, A. M., Van Arsdale, R., & Stein, S. (2010). Triggering of New Madrid seismicity by late-Pleistocene erosion. *Nature, 466*(7306), 608-611. http://www.scopus.com/inward/record.url?eid=2-s2.0-77955170178&partnerID=40&md5=fd033c8bfe596c82450a2a2d669f6394

Chen, Y., Liu, M., & Luo, G. (2020). Complex Temporal Patterns of Large Earthquakes: Devil’s Staircases. *Bulletin of the Seismological Society of America*, *110*(3), 1064-1076.

Cheng, J., Rong, Y., Magistrale, H., Chen, G., & Xu, X. (2017). An M w-based historical earthquake catalog for Mainland China. *Bulletin of the Seismological Society of America*, *107*(5), 2490-2500.

Clark, D., McPherson, A., & Allen, T. (2014). Intraplate earthquakes in Australia. In *Intraplate Earthquakes* (Vol. 8, pp. 49): Cambridge University Press, New York.

Clark, D., McPherson, A., & Van Dissen, R. (2012). Long-term behaviour of Australian stable continental region (SCR) faults. *Tectonophysics*, *566*, 1-30.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America, 58*(5), 1583-1606.

Corral, A. (2004). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letters, 92*(10), 108501.

De Michelis, P., & Consolini, G. (2003). Some new approaches to the study of the Earth's magnetic field reversals. *Annals of Geophysics*, *46*(4), 661-670.

DeMets, C., Gordon, R. G., Argus, D. F., & Stein, S. (1994). Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motion. *Geophys. Res. Lett.*, *21*, 2191-2194.

Devroye, L. (1986). *Sample-based non-uniform random variate generation.* Paper presented at the Proceedings of the 18th conference on Winter simulation.

Dolan, J., Bowman, D., & Sammis, C. (2007). Long-range and long-term fault interactions in Southern California. *Geology*. http://dx.doi.org/10.1130/G23789A.1

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., et al. (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—

The time-independent model. *Bulletin of the Seismological Society of America, 104*(3), 1122-1180.

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., et al. (2015). Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3). *Bulletin of the Seismological Society of America, 105*(2A), 511-543.

Field, E. H., Dawson, T. E., Felzer, K. R., Frankel, A. D., Gupta, V., Jordan, T. H., et al. (2009). Uniform California earthquake rupture forecast, version 2 (UCERF 2). *Bulletin of the Seismological Society of America, 99*(4), 2053-2107.

Freed, A. M., & Lin, J. (1998). Time-dependent changes in failure stress following thrust earthquakes. *Journal of Geophysical Research, B, Solid Earth and Planets, 103*(10), 24,393-324,409.

Freed, A. M., & Lin, J. (2001). Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer. *Nature (London), 411*(6834), 180-183.

Friedrich, A. M., Wernicke, B. P., Niemi, N. A., Bennett, R. A., & Davis, J. L. (2003). Comparison of geodetic and geologic data from the Wasatch region, Utah, and implications for the spectral character of Earth deformation at periods of 10 to 10 million years. *Journal of Geophysical Research: Solid Earth, 108*(B4).

Fujiwara, H., Kawai, S., Aoi, S., Morikawa, N., Senna, S., Kobayashi, K., et al. (2006). National seismic hazard maps of Japan. *Bull. Earthq. Res. Inst. Univ. Tokyo, 81*, 221-232.

Gardner, J., & Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America, 64*(5), 1363-1367.

Genrich, J., Bock, Y., McCaffrey, R., Prawirodirdjo, L., Stevens, C., Puntodewo, S., et al. (2000). Distribution of slip at the northern Sumatran fault system. *Journal of Geophysical Research: Solid Earth, 105*(B12), 28327-28341.

Giacomo, D. D., Engdahl, E. R., & Storchak, D. A. (2018). The ISC-GEM Earthquake Catalogue (1904–2014): status after the Extension Project. *Earth System Science Data, 10*(4), 1877-1899.

Goes, S. D. (1996). Irregular recurrence of large earthquakes: an analysis of historic and paleoseismic catalogs. *Journal of Geophysical Research: Solid Earth, 101*(B3), 5739-5749.

Goh, K.-I., & Barabási, A.-L. (2008). Burstiness and memory in complex systems. *EPL (Europhysics Letters), 81*(4), 48002.

Hainzl, S., Scherbaum, F., & Beauval, C. (2006). Estimating background activity based on interevent-time distribution. *Bulletin of the Seismological Society of America, 96*(1), 313-320.

Hasumi, T., Akimoto, T., & Aizawa, Y. (2009). The Weibull–log Weibull distribution for interoccurrence times of earthquakes. *Physica A: Statistical Mechanics and its Applications, 388*(4), 491-498.

Huang, W.-Q., Li, W.-X., & Cao, X.-F. (1994). Research on the completeness of earthquake data in the Chinese mainland (I)—North China. *Acta Seismologica Sinica, 7*(3), 351-359.

Kagan, Y. Y., & Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International, 104*(1), 117-133.

Kagan, Y. Y., & Jackson, D. D. (1999). Worldwide doublets of large shallow earthquakes. *Bulletin of the Seismological Society of America, 89*(5), 1147-1155.

Kreemer, C., Blewitt, G., & Klein, E. C. (2014). A geodetic plate motion and Global Strain Rate Model. *Geochemistry, Geophysics, Geosystems, 15*(10), 3849-3889.

Li, Q., Liu, M., & Stein, S. (2009). Spatial-temporal complexity of continental intraplate seismicity: Insights from geodynamic modeling and implications for seismic hazard estimation. *Bulletin of the Seismological Society of America*, *99*(1), doi: 10.1785/0120080005.

Liu, M., Luo, G., & Wang, H. (2014). The 2013 Lushan earthquake in China tests hazard assessments. *Seismological Research Letters*, *85*(1), 40-43.

Liu, M., & Stein, S. (2016). Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards. *Earth-Science Reviews*, *162*, 364-386.

Luo, G., & Liu, M. (2010). Stress evolution and fault interactions before and after the 2008 Great Wenchuan earthquake. *Tectonophysics, 491*(1-4), 127-140, doi:110.1016/j.tecto.2009.1012.1019.

Luo, G., & Liu, M. (2012). Multi-timescale mechanical coupling between the San Jacinto Fault and the San Andreas Fault, southern California. *Lithosphere*, doi: 10.1130/L1180.1131.

Luo, G., & Liu, M. (2018). Stressing rates and seismicity on the major faults in eastern Tibetan Plateau. *Journal of Geophysical Research: Solid Earth, 123,* https://doi.org/10.1029/2018JB015532.

Mandelbrot, B. B. (1982). *The fractal geometry of nature* (Vol. 1): WH freeman New York.

Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association, 46*(253), 68-78.

McCalpin, J., & Nishenko, S. (1996). Holocene paleoseismicity, temporal clustering, and probabilities of future large (M> 7) earthquakes on the Wasatch fault zone, Utah. *Journal of Geophysical Research: Solid Earth*, *101*(B3), 6233-6253.

McClusky, S., Balassanian, S., Barka, A., Demir, C., Ergintav, S., Georgiev, I., et al. (2000). Global Positioning System constraints on plate kinematics and dynamics in the eastern Mediterranean and Caucasus. *Journal of Geophysical Research: Solid Earth, 105*(B3), 5695-5719.

Michael, A. J. (2014). How complete is the ISC-GEM global earthquake catalog? *Bulletin of the Seismological Society of America, 104*(4), 1829-1837.

Moernaut, J., Van Daele, M., Fontijn, K., Heirman, K., Kempf, P., Pino, M., et al. (2018). Larger earthquakes recur more periodically: New insights in the megathrust earthquake cycle from lacustrine turbidite records in south-central Chile. *Earth and Planetary Science Letters, 481*, 9-19.

Ratzov, G., Cattaneo, A., Babonneau, N., Déverchère, J., Yelles, K., Bracene, R., & Courboulex, F. (2015). Holocene turbidites record earthquake supercycles at a slow-rate plate boundary. *Geology, 43*(4), 331-334.

Salditch, L., Stein, S., Neely, J., Spencer, B. D., Brooks, E. M., Agnon, A., & Liu, M. (2019). Earthquake supercycles and Long-Term Fault Memory. *Tectonophysics*, 228289.

Scharer, K. M., Biasi, G. P., Weldon, R. J., & Fumal, T. E. (2010). Quasi-periodic recurrence of large earthquakes on the southern San Andreas fault. *Geology, 38*(6), 555-558.

Shen, Z.-K., Lü, J., Wang, M., & Bürgmann, R. (2005). Contemporary crustal deformation around the southeast borderland of the Tibetan Plateau. *Journal of Geophysical Research: Solid Earth, 110*(B11).

Sieh, K. E., Stuiver, M., & Brillinger, D. (1989). A more precise chronology of earthquakes produced by the San Andreas fault in southern California. *Journal of Geophysical Research, 94*, 603-623.

Simkin, M., & Roychowdhury, V. (2014). Stochastic modeling of a serial killer. *Journal of theoretical biology, 355*, 111-116.

Stein, S., & Liu, M. (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature, 462*(7269), 87-89. http://www.scopus.com/inward/record.url?eid=2-s2.0-70449124788&partnerID=40&md5=37b8160bace8d6e8842ff1be818de349

Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H., et al. (2013). Public release of the ISC–GEM global instrumental earthquake catalogue (1900–2009). *Seismological Research Letters, 84*(5), 810-815.

Straub, C., Kahle, H. G., & Schindler, C. (1997). GPS and geologic estimates of the tectonic activity in the Marmara Sea region, NW Anatolia. *Journal of Geophysical Research: Solid Earth, 102*(B12), 27587-27601.

Turcotte, D. L. (1997). *Fractals and chaos in geology and geophysics*: Cambridge university press.

Turcotte, D. L., & Malamud, B. D. (2002). 14 earthquakes as a complex system. *International Geophysics, 81*, 209-IV.

Vigny, C., Socquet, A., Rangin, C., Chamot-Rooke, N., Pubellier, M., Bouin, M. N., et al. (2003). Present-day crustal deformation around Sagaing fault, Myanmar. *Journal of Geophysical Research: Solid Earth, 108*(B11).

Weldon, R., Scharer, K., Fumal, T., & Biasi, G. (2004). Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work. *GSA today, 14*(9), 4-10.

WGCEP. (1988). *Probabilities of large earthquakes occurring in California on the San Andreas fault* (88-398). Retrieved from http://pubs.er.usgs.gov/publication/ofr88398

WGCEP. (1995). Seismic hazards in southern California: Probable earthquakes, 1994 to 2024. *Bulletin of the Seismological Society of America, 85*(2), 379-439.

WGCEP. (2003). *Earthquake probabilities in the San Francisco Bay Region, 2002-2031* (2003-214). Retrieved from http://pubs.er.usgs.gov/publication/ofr03214

Williams, R. T., Davis, J. R., & Goodwin, L. B. (2019). Do large earthquakes occur at regular intervals through time? A perspective from the geologic record. *Geophysical Research Letters*.

Williams, R. T., Goodwin, L. B., Sharp, W. D., & Mozley, P. S. (2017). Reading a 400,000-year record of earthquake frequency for an intraplate fault. *Proceedings of the National Academy of Sciences, 114*(19), 4893-4898.

Xu, X., & Deng, Q. (1996). Nonlinear characteristics of paleoseismicity in China. *Journal of Geophysical Research: Solid Earth*, *101*(B3), 6209-6231. http://dx.doi.org/10.1029/95JB01238

Chapter 3: Spatiotemporal patterns of intraplate earthquakes: insight from North China

In this chapter, I studied the earthquakes in intraplate North China, a reactivated craton, to get some insight into the spatiotemporal patterns of intraplate earthquakes and their implications for seismic hazard assessment. This chapter is partly based on my publication: Chen, Y., Liu, M., & Wang, H. (2021). Aftershocks and background seismicity in Tangshan and the rest of North China. *Journal of Geophysical Research: Solid Earth*, *126*, e2020JB021395. https://doi.org/10.1029/2020JB021395.

3.1 Geological settings of North China

North China, or the geologically defined North China block, includes the North China Plain and the mountain ranges and the Ordos Plateau to the west. It is an Archaean craton that was reactivated during the Mesozoic, with maximum lithospheric thinning under the North China Plain. Today, North China is a region of active intraplate seismicity (M. Liu et al., 2014), with twenty-two $M \geq 7$ earthquakes recorded since 70 BCE , including five $M \geq 8.0$ events (Figure 3.1) (Gu et al., 1983). Beside the 1976 Great Tangshan earthquake, North China was devastated by the 1966 M_s 7.2 Xingtai earthquake, the 1969 M_b 7.4 Bohai earthquake, and the 1975 M_s 7.3 Haicheng earthquake in the past century (Figure 3.1). These earthquakes all occurred in places where no large earthquakes were recorded in the previous 2000 years (M. Liu & Wang, 2012), highlighting the challenges of assessing seismic hazard in slowly deforming continental interior (M. Liu & Stein, 2016).

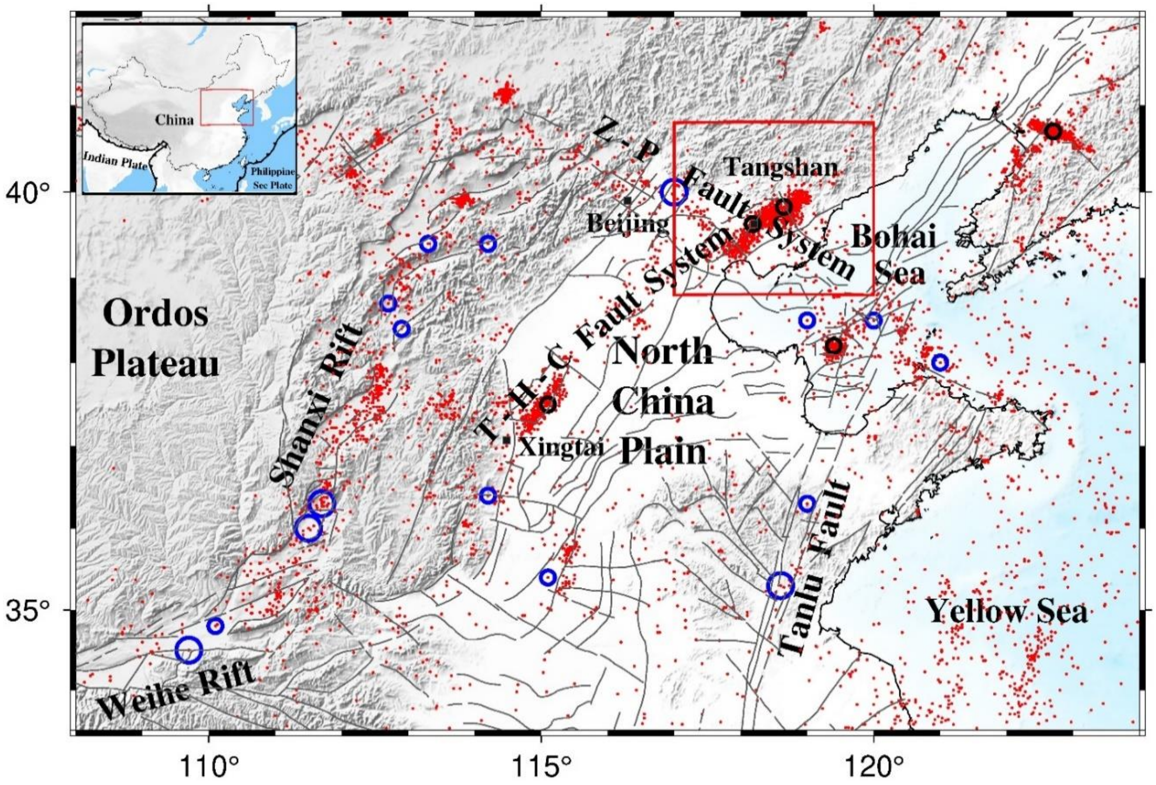


Figure 3.1. Seismicity in North China. Blue circles are $M \geq 7$ earthquakes since 70 BCE and before 1960. Large blue circles are $M \geq 8$ earthquakes. Black circles are $M \geq 7$ earthquakes from 1960 to July 2019. Red dots are $M \geq 3$ earthquakes from 1960 to July 2019. The gray lines are faults. Z-P fault system: the Zhangjiakou-Penglai fault system; T-H-C fault system: the Tanshan-Hejian-Cixian fault system. The red rectangle shows the epicentral region of the 1976 M_S7.8 Tangshan earthquake.

3.2 Data and declustering method

The earthquake catalog (1970-2020/07/12, $M \geq 2.5$) of the Tangshan region (117°-120°E, 38.8°-40.8°N) used in this chapter includes two parts: earthquakes between 1970 and 2019 from the China Seismic Array Data Management Center (http://www.chinarraydmc.cn/products/index) and recent earthquakes (2020/01/01-2020/07/12) from the China Earthquake Networks Center (http://news.ceic.ac.cn). The earthquake catalog of North China (108°-124°E, 33.5°-42°N) includes historical earthquakes (70 BCE-1969, $M \geq 4$) from Gu et al. (1983) and instrumental earthquakes (1970-2019/08/01, $M \geq 3$) from the China National Earthquake Data Center (http://data.earthquake.cn). In North China, the seismic records of $M \geq 2.6$ earthquakes are suggested to be complete since 1970 (Mignan et al., 2013). However, within a short period after the 1976 Great Tangshan earthquake, many small earthquakes were not recorded and the magnitude of completeness (M_c) is as high as 4 (Jiang et al., 2013). Reduced detection capability shortly after a large earthquakes is a global problem (Iwata, 2008). Furthermore, some stations in the epicentral region were damaged by the Tangshan mainshocks (Jiang et al., 2013).

Various declustering methods have been developed to separate aftershocks from background earthquakes (van Stiphout et al., 2012). Classic declustering algorithms include the window method proposed by Gardner and Knopoff (1974) and an updated method by Reasenberg (1985). The window method identifies aftershocks based on magnitude-dependent inter-event distances in time and space. The specific spatial and temporal windows for aftershocks are empirically predetermined based on earthquakes in California (Gardner & Knopoff, 1974). However, aftershock durations vary with tectonic settings; they tend to be longer in slowly deforming continents than those in rapidly loaded plat boundaries (Stein & Liu, 2009; Toda & Stein, 2018), so the window method may not be suitable for North China. The Reasenberg (1985) declustering algorithm identifies aftershocks according to spatial and temporal interaction zones. The space and time extents of an interaction zone are chosen according to the stress distribution near a mainshock and Omori’s law of decaying aftershocks. The algorithm links observed earthquakes with underlying physics (stress), but the values of the parameters are somewhat arbitrarily chosen and can cause significant difference in the declustered results (Toda & Stein, 2018). Besides the spatiotemporal window methods, probability declustering methods have been developed. These methods use stochastic models to estimate the probabilities that each earthquake is either a mainshock or an aftershock (Marsan & Lengline, 2008; Zhuang et al., 2002). These methods use catalog-constrained declustering parameters but need to assume the underlying probability distribution and the background seismicity, which is often treated as being constant.

In this chapter, I declustered the catalogs of earthquakes in North China using the nearest-neighbor (NN) method (Baiesi & Paczuski, 2004; Zaliapin et al., 2008). The NN method is based on the Gutenburg-Richter frequency-magnitude relationship (Gutenberg & Richter, 1954) and fractal property of spatial distribution of seismicity. It statistically identifies clustered events (treated as aftershocks here) based on the distance in the space-time-magnitude domain between these events. If the events are too close to be expected for independent background events that distribute in the seismic region following the Gutenburg-Richter law, they are taken as clustered (i.e., correlated) events.

In the NN method, the distance η_{ij} in the space-time-magnitude domain between a given earthquake j occurring at t_j and a past earthquake i occurring at t_i $(t_i < t_j)$ is defined as

$$\eta_{ij} = \begin{cases} t_{ij}\left(r_{ij}\right)^{d_f} 10^{-bm_i}, & t_{ij} > 0 \\ \infty \qquad\qquad\quad , & t_{ij} \leq 0 \end{cases} \tag{3.1}$$

where $t_{ij} = t_j - t_i$ is the inter-occurrence time in years, r_{ij} is the surface distance between the earthquake epicenters in kilometers, m_i is the magnitude of earthquake i, d_f is the fractal dimension characterizing the distribution of the epicenters, and b is the slope parameter of the Gutenberg-Richter law. The value of b in our study is calculated by the maximum likelihood method (Equation 1.2) (Aki, 1965). The fractal dimension d_f used in the NN method is estimated by the correlation integral method (Chapter 1.3.3) (Grassberger & Procacia, 1983). The nearest-neighbor distance (NND, shown as η_{ij}^* or η) is defined as the distance between event j and its nearest neighbor i^*, which is the closest event to event j among all past events. From a catalog, the NND values η can be calculated and a threshold value η_0 can be found from the distribution of NND (see below), which is used to separate clustered events from background events. When $\eta \leq \eta_0$, event j is taken as an aftershock of its nearest neighbor i^*; when $\eta > \eta_0$, it is treated as a background earthquake.

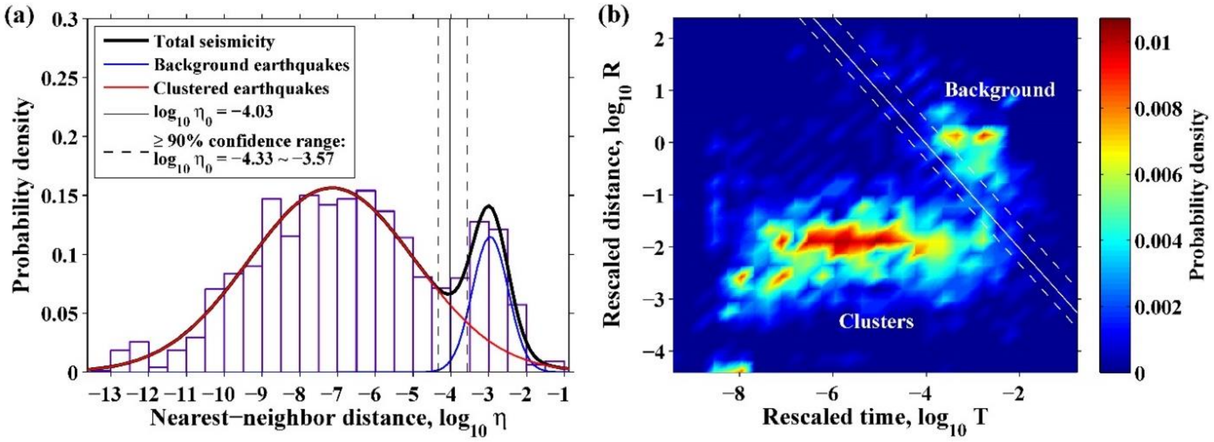


Figure 3.2. Nearest-neighbor method applied to $M \geq 4$ earthquakes since 1960 in North China: **(a)** 1D density distribution of η, with estimated Gaussian densities for clustered

(red) and background (blue) components. The threshold distance η_0 is where the two components are best separated, here at $\log_{10}\eta_0 = -4.03$ (the solid vertical line). The dashed vertical lines show the ≥ 90% confidence range (see discussion in text). **(b)** 2D joint distribution of rescaled distance and time (R, T). The solid line ($\log_{10} T + \log_{10} R = -4.03$) is the preferred threshold between background and clustered earthquakes. The dashed lines are the ≥ 90% confidence range shown in (a). The results are calculated with $b = 1$ and $d_f = 1.5$ (Figure 3.23, 3.24), and are not sensitive to variations of these values within reasonable ranges (Figure 3.25).

The NND of earthquakes ($M \geq 4$) in North China (Figure 3.2a) shows a bimodal Gaussian distribution, similar to earthquakes in other regions (Peresan & Gentili, 2018; Zaliapin et al., 2008). The same pattern also shows for earthquakes predicted by an epidemic type aftershock sequence (ETAS) model that mixes a homogeneous Poissonian background field with an aftershock series, with aftershocks group to the left (lower NND) and background events to the right (higher NND) (Zaliapin et al., 2008). The separation threshold η_0 can be calculated according to a Gussian mixture model with the background and clustered modes (Hicks, 2011; Zaliapin & Ben-Zion, 2016). In North China, the calculated threshold is $\log_{10}\eta_0 = -4.03$ (the solid vertical line in Figure 3.2a).

Because of the overlapping of the two Gaussian distributions, using the separation threshold $\log_{10}\eta_0 = -4.03$ could misidentify some background events as aftershocks and some aftershocks as background events. Assuming the data represent the true underlying probability distribution, I can quantify this uncertainty by using the concept of confidence level. Here, the preferred separation gives us 92.2% confidence level (the fraction of the total area under the red curve in Figure 3.2a, measured from the left to the right, truncated by the separation line) that aftershocks would be identified. Similarly, this separation gives 98.8% confidence of identifying background events.

The vertical dashed lines in Figure 3.2a show the range of ≥ 90% confidence levels for aftershocks and background earthquakes, respectively. The left boundary ($\log_{10}\eta_0 = -4.33$) corresponds to 90% confidence level for aftershock identification, and the right boundary ($\log_{10}\eta_0 = -3.57$) corresponds to 90% confidence level for identifying background events.

The NND can be decomposed into the corresponding rescaled (magnitude-normalized) time (T_{ij}) and distance (R_{ij}) (Zaliapin et al., 2008):

$$T_{ij} = t_{ij}10^{-\frac{bm_i}{2}}, \qquad R_{ij} = r_{ij}^{d_f}10^{-\frac{bm_i}{2}}. \tag{3.2}$$

with $\eta_{ij} = T_{ij}R_{ij}$, or equivalently $\log_{10}\eta_{ij} = \log_{10} T + \log_{10} R$. In the T-R plot (Figure 3.2b), the separation boundary between background earthquakes and aftershocks is a straight line: $\log_{10}\eta_0 = \log_{10} T + \log_{10} R$. The clustered events are in the domain of lower *T* and *R* values. The negative correlation between *T* and *R* in Figure 3.2b may be interpreted this way: for a distant event (larger R) to be an aftershock, it needs to occur shortly after the mainshock (smaller T), and vise versa. Note the band of high probability density around $\log_{10} R = -2$ in Figure 3.2b, which differs from a centrally clusterd distribution of a single synthetic aftershock series following the Omori's law (Figure 1 of

Zaliapin et al., 2008) and probably reflects the proximity of most aftershocks to rupture zones in nature.

3.3 Tangshan earthquakes: Aftershocks vs. Background Seismicity

At 3:42 a.m. (UTC+8) on July 28, 1976, a magnitude (M_s) 7.8 earthquake hit Tangshan, an industrial city in North China just 150 km east of Beijing, causing a surface rupture more than 47 km long (Figure 3.3) (Guo et al., 2017). About 15 hours later, a second shock, of M_s 7.1, occurred approximately 40 km to the east-northeast on a different fault and ruptured surface more than 6 km long (Guo et al., 2017). These two events, together referred to as the Great Tangshan earthquake, obliterated the city of Tangshan and killed more than 240,000 people, making it the deadliest earthquake of the 20th century (Y. Chen et al., 1988). More than 40 years have passed, but the memory of the devastation remains fresh. Thus, a series of moderate ($M \geq 4.5$) earthquakes in recent years near the 1976 Tangshan earthquake rupture zone, including an M_s 5.1 event on July 12, 2020 (Figure 3.3), have raised social concern and scientific debate: are these events aftershocks of the Great Tangshan earthquake? Or are they indicators of stress buildup for future large earthquakes in Tangshan or somewhere in North China?

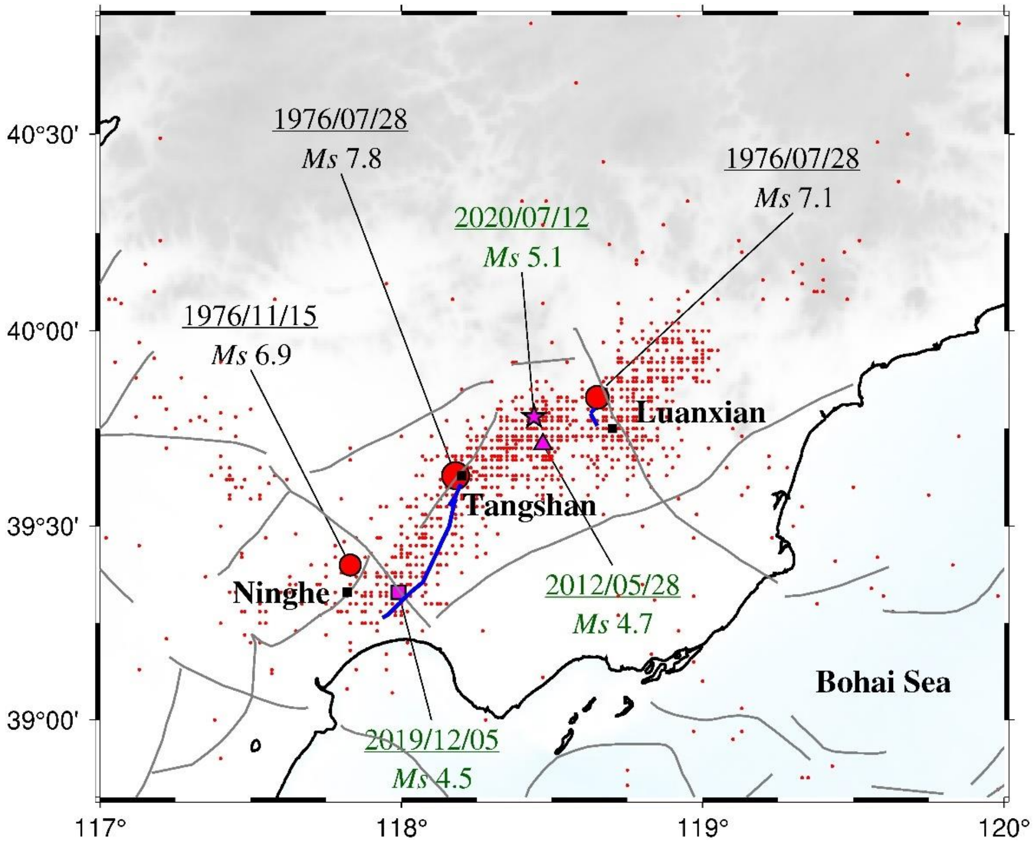


Figure 3.3. Seismicity in the epicentral region of the 1976 Tangshan earthquake (red rectangle in Figure 3.1). Red dots are $M \geq 3$ earthquakes from 1970 to July, 2020. Pink symbols are moderate earthquakes in recent years. The gray lines are faults. The blue lines are the surface ruputures of the 1976 M_s 7.8 Tangshan earthquake and M_s 7.1 Luanxian earthquake from Guo et al. (2017).

Establishing reliable background seismicity is the basis for probabilistic seismic hazard analysis (Frankel, 1995; Petersen et al., 2018), but a challenge for intraplate seismicity. In slowly deforming intraplate regions, large earthquakes are less frequent and aftershock sequences are generally much longer than that at plate boundaries (Stein & Liu, 2009). Some aftershocks may be mistakenly identified as elevated background seismicity and therefore lead to an overestimation of seismic hazard (Toda & Stein, 2018). This issue underlies the debate about the recent earthquakes in the Tangshan region: are they aftershocks of the 1976 Great Tangshan earthquake, or are they background seismicity? Zhong and Shi (2012), by fitting the Omori decay exponents of seismicity in the Tangshan region and using the formula of aftershock duration from the rate-and-state models (Dieterich, 1994), suggested that the duration of the 1976 Tangshan aftershock sequence is around 70-140 years. Therefore, the recent moderate earthquakes in the Tangshan region would likely be aftershocks. M. Liu and Wang (2012) reached similar conclusion by fitting seismicity in Tangshan with the Omori's decay law and noting the elevated seismicity and strain rates in the Tangshan region relative to the rest of North China. However, as argued by Y. Liu et al. (2020), the parameters used in the calculations by Zhong and Shi (2012) are associated with large uncertainties. Using epidemic-type aftershock sequence (ETAS) models with a complete earthquake catalog of $M \geq 4$ earthquakes, Jiang et al. (2013) and Y. Liu et al. (2020) have suggested that the recent moderate earthquakes in Tangshan are more likely background earthquakes. However, the ETAS model assumes a constant background seismicity rate, which may not hold, as acknowledged by Y. Liu et al. (2020) and suggested by other studies (Zaliapin & Ben-Zion, 2020).

3.3.1 Results from the NN method

Using the threshold distance constrained by the North China earthquakes, we can attempt to determine if the recent moderate earthquakes in the Tangshan region are aftershocks or background earthquakes (Figure 3.4). We used the catalogs with two different minimum cut-off magnitudes ($M \geq 4$ and $M \geq 3$), the patterns are robust with respect to the minimum cut-off magnitude, as suggested by Zaliapin and Ben-Zion (2013). Most events in the catalog are aftershocks, as indicated by the high event-density cluster in the T-R space (Figure 3.4). The number of background earthquakes are relatively small, and hence the weak event-density in the background domain (above the $\log_{10} \eta_0 = -4.03$ boundary). The M_s 5.1 event of 2020/07/12, as well as two other $M \geq$ 4.5 events since 2012, all fall near the background-clusters boundary within the ≥ 90% confidence range (Figure 3.4), hence it is difficult to definitely identify them as aftershocks or background earthquakes. However, these three events are at the tail of the aftershock mode and has a clear separation from the peak of background events (Figure 3.4), so they are more likely aftershocks.

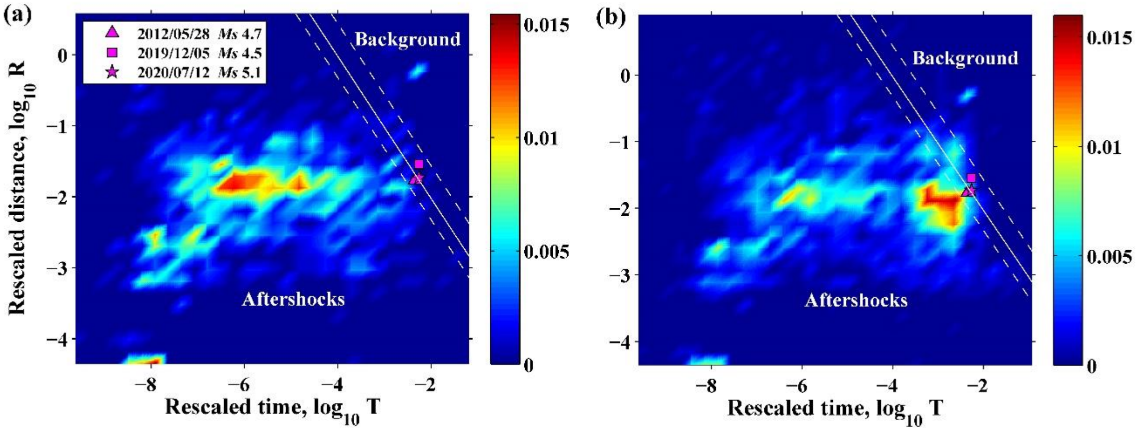


Figure 3.4. Nearest-neighbor method applied to **(a)** $M \geq 4$ and **(b)** $M \geq 3$ earthquakes since 1970 in the Tangshan region. The threshold distance ($\log_{10} \eta_0 = -4.03$, solid lines) and the ≥ 90% confidence range (dashed lines) are based on earthquakes from the entire North China (Figure 3.2). The relatively small probability density in (b) for aftershocks may be attributed to record incompleteness for $M \geq 3$ events shortly after the

mainshock (see text and Figure 3.26). The magenta symbols show the recent moderate earthquakes in the Tangshan region.

3.3.2 Duration of the 1976 Tangshan aftershock sequence

Alternatively, one may estimate the duration of aftershock sequence of the 1976 Great Tangshan earthquake directly from the catalog. If the aftershock sequence had ended before the recent events, then these events are background earthquakes. Aftershock duration is typically defined as the time required for seismicity rate after an mainshock to return to the background seismicity rate before the mainshock (Dieterich, 1994; Toda & Stein, 2018). Some called this “apparent aftershock duration” to emphasize that the underlying physical triggering process may continue after this duration and the true aftershock duration may be longer (Hainzl et al., 2016).

M. Liu and Wang (2012) used the total seismicity rate to suggest that the aftershock sequence of the 1976 Great Tangshan earthquake had likely continued to 2010 (the end of their catalog), because the seismicity rate had not fallen back to the level before the 1976 Great Tangshan earthquake. However, doing so requires subjective choice of the region from where earthquakes are recorded. Furthermore, besides the background earthquakes and aftershocks of the Great Tangshan earthquake, the catalog may include other clusters (Figure 3.26).

Here, I used the background seismicity rate from the results of NN method and the Omori’s law to estimate the aftershock duration of the 1976 Great Tangshan earthquake (Chapter 1.3.2). Based on the background earthquakes identified by the NN method, the background seismicity rates before and after the 1976 *Ms* 7.8 Tangshan mainshock are different (Figure 3.5). Note that $M \geq 3$ (and perhaps some $M \geq 4$) background earthquakes are missing wthin several years after the 1976 mainshock (as indicated by the flat step after the 1976 mainshock in Figures 3.5a and 3.5b). A number of factor could contribute to this missing background seismicity; a likely one is the limitation of the NN method, which may identify background events shortly after large mainshocks as aftershocks (Zaliapin & Ben-Zion, 2020). However, this limitation has little effect on background earthquakes after 1985, when background seismicity starts to accumulate with a linear trend. The background seismic rates after 1985 for both $M \geq 4$ and $M \geq 3$ earthquakes are lower than the ones before the 1976 mainshock (Figure 3.5).

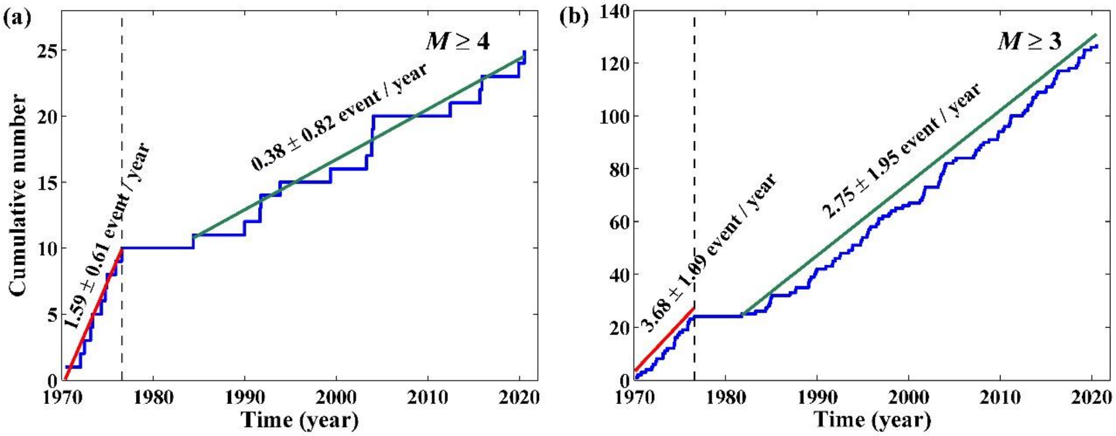


Figure 3.5. Background seismicity rate of **(a)** $M \geq 4$ and **(b)** $M \geq 3$ earthquakes in the Tangshan region. The background earthquakes are identified by the NN method. The blue staircases are the cumulative numbers of earthquakes. The dashed lines indicate the occurrence time of the 1976 M_s 7.8 mainshock. The lines are linear least-square fits to the average background seismicity rates before (red) and after (green) the mainshock, respectively.

I calculated the aftershock durations using Equation 1.4 and the different background seismicity rates before and after the 1976 mainshock (Figure 3.6). For $M \geq 4$ earthquakes in the Tangshan region, although the recent (2008.3-2020.5) seismicity rate is lower than the rate of the identified background earthquakes before the 1976 mainshock, the seismicity continues the Omori's decay and has not reached the rate of the post-mainshock background seismicity (Figure 3.6a). The apparent aftershock duration is near 100 years. For $M \geq 3$ earthquakes, the recent seismicity rate has not decayed to the background rates (Figure 3.6b). The incomplete record of $M \geq 3$ earthquakes shortly after the mainshock causes problems for fitting the the Omori's law, so we fit only the complete data after 1980 (the blue line in Figure 3.6b) to estimate the aftershock duration t_a. The predicted seismicity rate by July 12, 2020 is higher than the background seismicity rate before or after the 1976 mainshock, suggesting that aftershock activity is still on-going for $M \geq 3$ earthquakes in Tangshan. The aftershock duration of $M \geq 3$ earthquakes in Tangshan is 66.9 year or 89.0 years, depending on the background seismicity rate used (before or after the 1976 mainshock), therefore the aftershock activity continues in 2020. Overall, the apparent aftershock duration of the 1976 Great Tangshan earthquake is around 65-100 years, so the current seismicity in the Tangshan region is likely dominated by aftershocks.

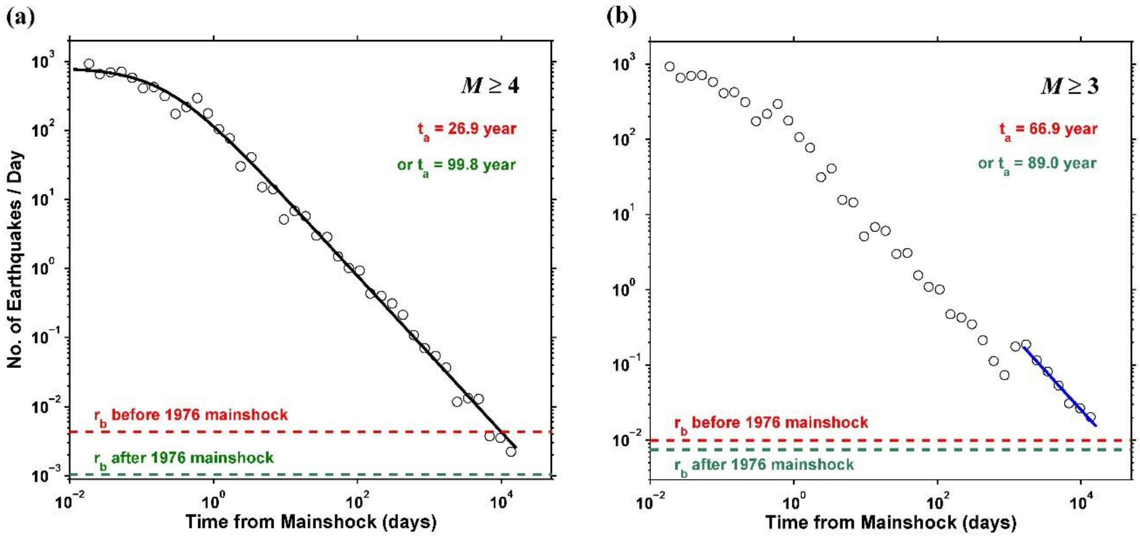


Figure 3.6. Temporal variations of seismicity rate based on **(a)** all $M \geq 4$ or **(b)** all $M \geq 3$ earthquakes after the 1976 M_s 7.8 mainshock. The black curve in (a) is the Omori's law fitting based on Equation 1.3. The blue line in (b) is least-square linear fitting of the complete records after 1980. The dashed lines are the estimated background seismicity rate (r_b) before the mainshock (red) and after the mainshock (green) shown in Figure 3.5, respectively. The end of fitting curves is 2020/07/12, the last date in the catalog. The duration of aftershocks t_a is indicated by the intersection of the fitting line with the background r_b.

3.3.3 Evolution of Coulomb Failure Stress

Aftershocks are supposed to be physically linked to the mainshock. My co-author Dr. Hui Wang calculated the spatiotemporal variation of the Coulomb Failure Stress change (ΔCFS) caused by the Tangshan earthquake sequence in 1976, and I compared the results with the aftershocks identified by our NN model (Figure 3.7). The 1976 Tangshan earthquake sequence includes the M_s 7.8 Tangshan mainshock and the M_s 7.1

Luanxian earthquake, both on 1976/07/28, and the M_s 6.9 Ninghe earthquake on 1976/11/15 (Figure 3.3). The evolution of the ΔCFS were calculated using a layered viscoelastic model (Wang et al., 2006). The major fault parameters and average dislocations of these three major events were from previous studies (B.-S. Huang & Yeh, 1997; Robinson & Zhou, 2005); the viscosity of the lower crust is taken as 1×10^{19} Pa s (H. Xu et al., 2018).

Figure 3.7 shows that the area of increasing ΔCFS enlarged with time due to sequential major events and post-seismic viscoelastic relaxation. 66.3% of the $M \geq 3$ aftershocks of the Tangshan earthquake identified by our NN model are found in areas with ΔCFS > +0.1 MPa. Some aftershocks are distant from the rupture zones (Figure 3.7b-c). They are identified as aftershocks because their occurrence times are close to that of the mainshock, because the NN method uses the distance in time-space-magnitude domain (Equation 3.1). The recent moderate earthquakes, including the 2020 M_s 5.1 Tangshan earthquake, all occurred in regions where ΔCFS increased, and thus may be regarded as triggered events, thus aftershocks (Figure 3.7d).

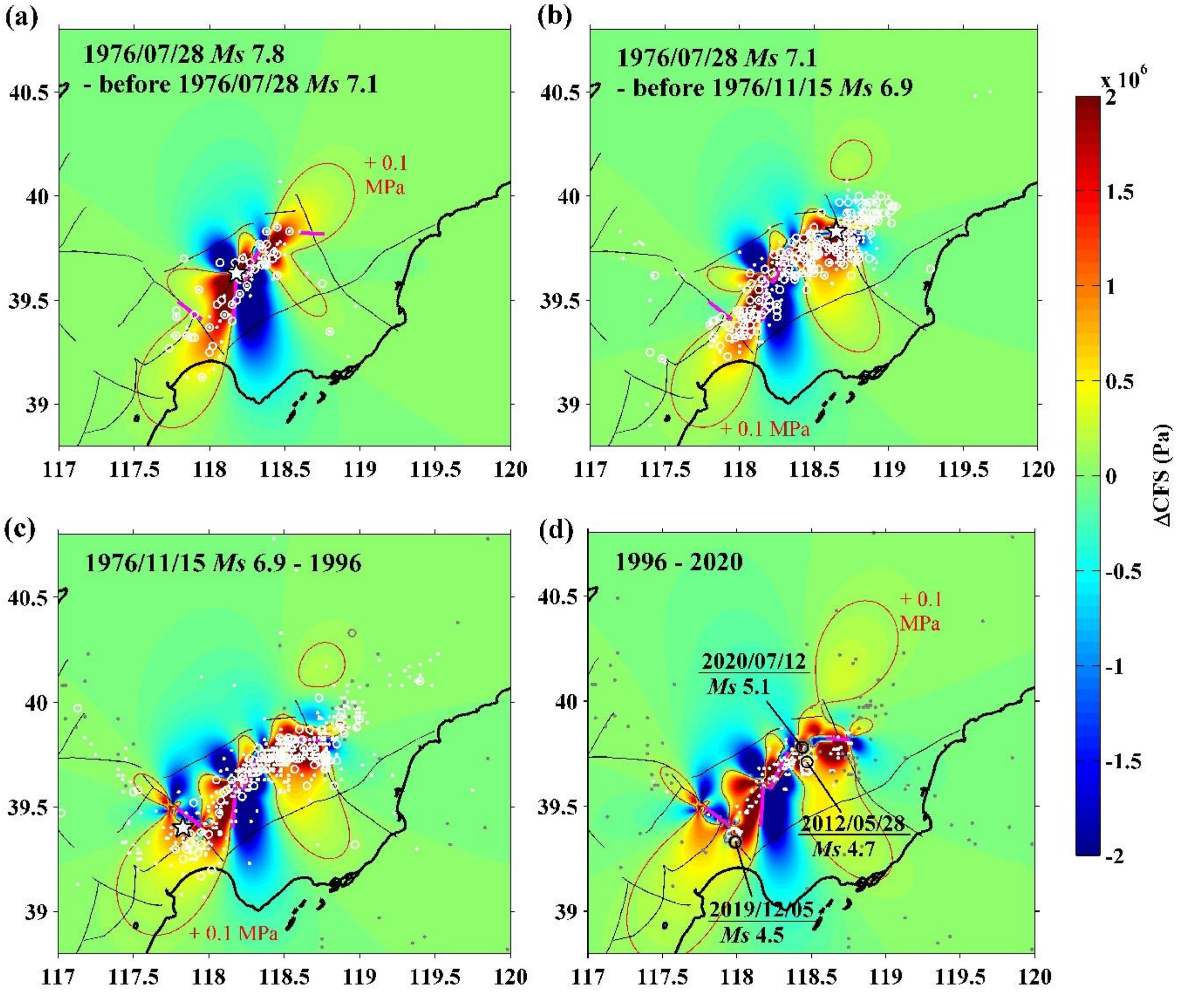


Figure 3.7. Spatiotemporal distribution of the Tangshan aftershocks and temporal variation of the Coulomb failure stress (ΔCFS) produced by 1976 mainshock **(a)** and two major subevents (white stars) in the same year **(b-c)**. The three recent moderate earthquakes are labeled **(d)**. Magenta lines show ruptures of the three major events. White circles ($M \geq 4.5$) and dots ($M < 4.5$) are aftershocks of the 1976 mainshock. Gray circles ($M \geq 4.5$) and dots ($M < 4.5$) are background earthquakes. The red contour lines show ΔCFS = + 0.1 MPa.

3.4 Variable background seismicity in North China

Background seismicity is the basis for probabilistic seismic hazard analysis (Frankel, 1995; Petersen et al., 2018). The spatiotemporal variations of background seismicity in North China may provide some useful insights into the question: Is North China entering a new period of more active seismicity?

As shown in Figure 3.1, several devastating earthquakes struck North China since 1960. Previous studies have noticed that seismic activity in North China seems to fluctuate between highs and lows over periods of a few decades (X. Xu & Deng, 1996). In general, seismic activity tends to cluster in relatively short periods, separated by long and relatively quiescent periods (Yuxuan Chen et al., 2020).

The seismicity rate of all $M \geq 3$ earthquakes in North China (Figure 3.8) shows strong temporal variations due to large earthquakes and their aftershocks. Before January 1976 (Figure 3.8a), the regions of high seismicity rate are mainly near the epicenters of the 1966 M_s 7.2 Xingtai earthquake, the 1969 M_b 7.4 Bohai earthquake, and the 1975 M_s 7.3 Haicheng earthquake. Seismicity rate near Tangshan is low. Between January 1976 and January 1986 (Figure 3.8b), Tangshan becomes a center of high seismicity rate because of the intense aftershocks of the Great Tangshan earthquake. The seismicity rates remain high around the epicentral regions of the Xingtai, Bohai, and Haicheng earthquakes, and near the epicenters of other $M \geq 6$ earthquakes during this period. During the next two decades (Figure 3.8c-d), the spatial patterns remain the same but with lower seismicity rates as aftershocks decay. Since 2006, high seismicity rate has faded away around Xingtai but continues around the Tangshan and Haicheng regions (Figure 3.8e-f).

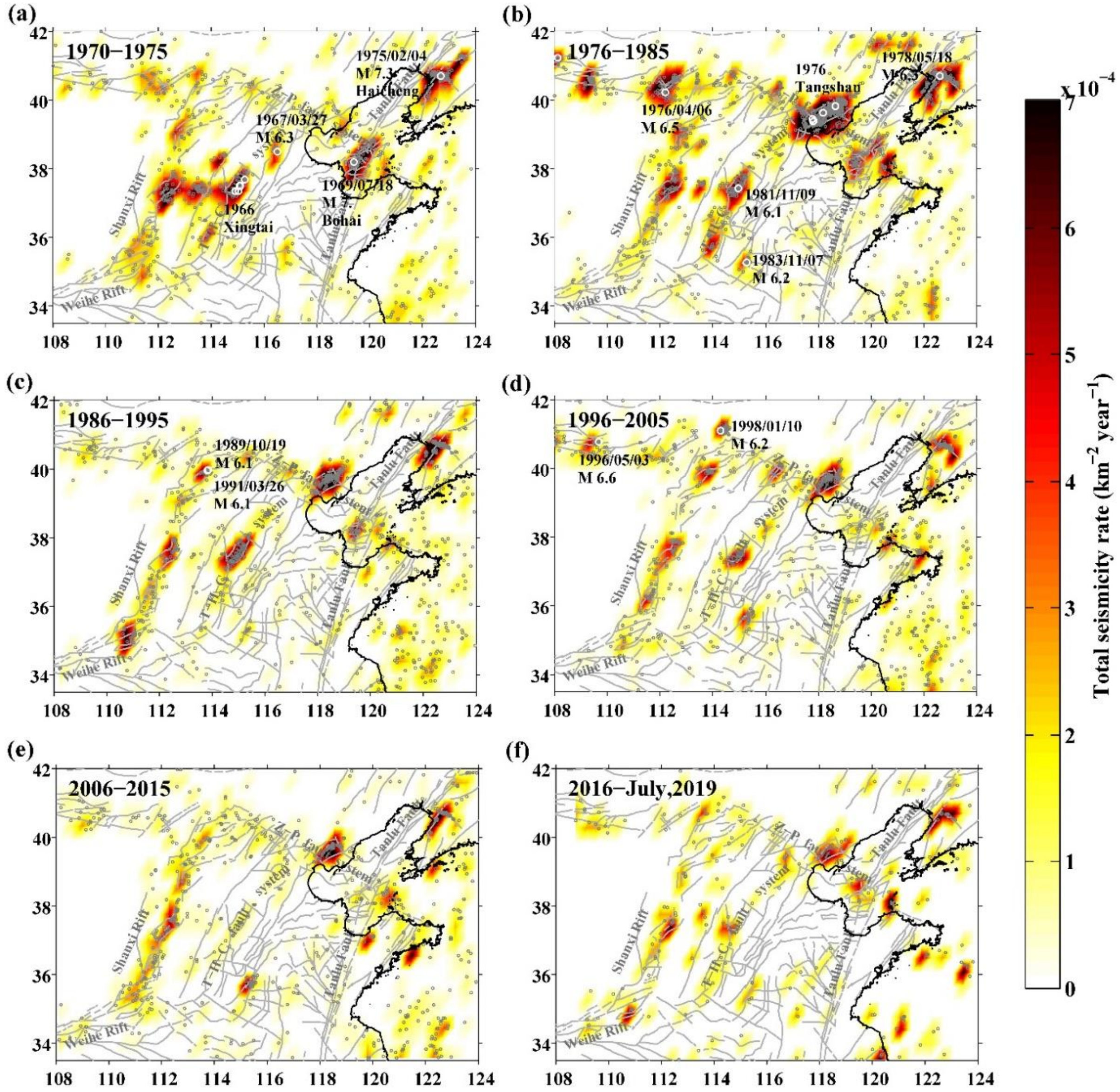


Figure 3.8. Spatiotemporal variation of total seismicity rate ($M \geq 3$) in North China. Small grey circles are all $M \geq 3$ earthquakes during each time period. Large white circles are $M \geq 6$ earthquakes.

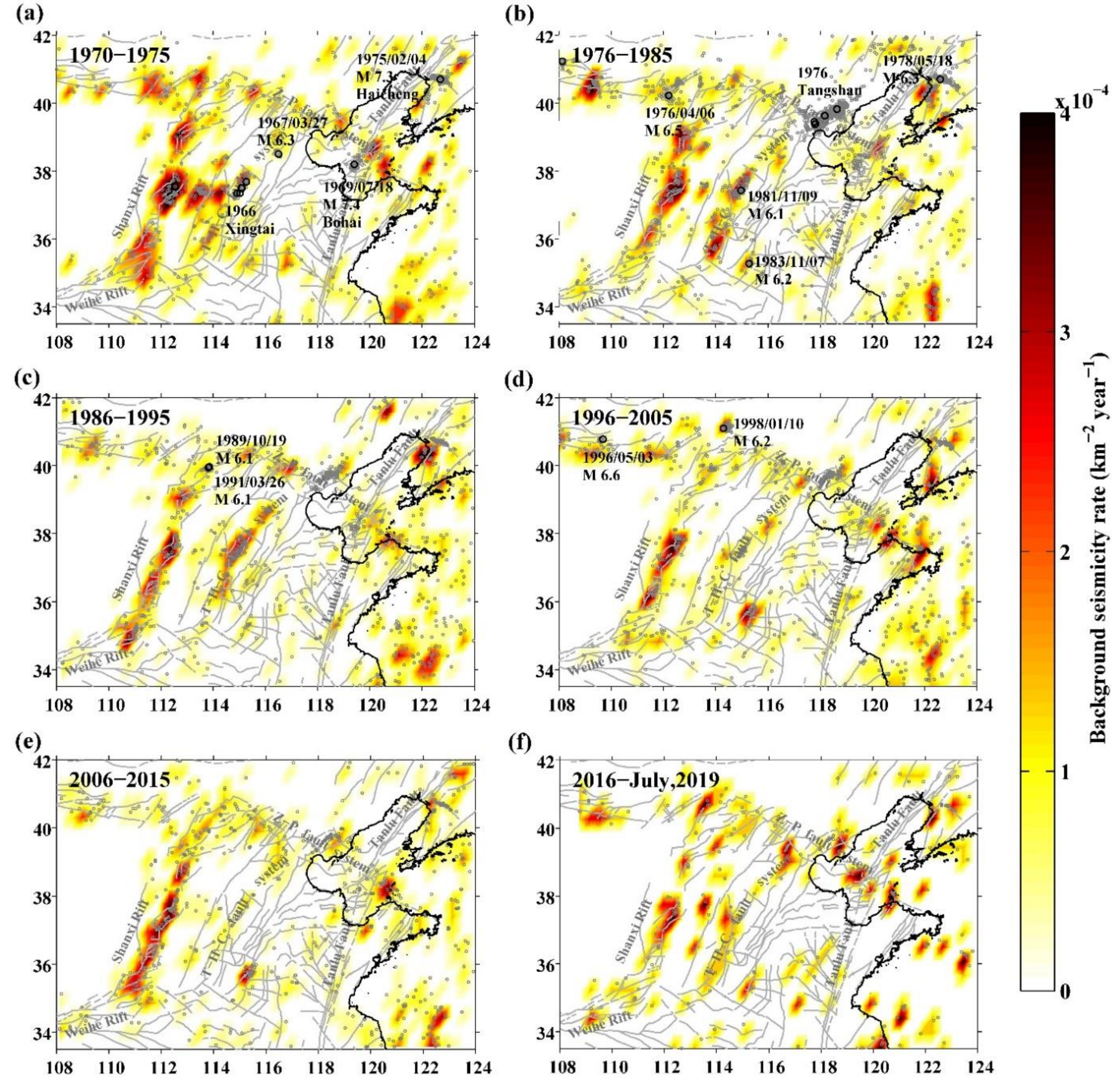


Figure 3.9. Spatiotemporal variation of background seismicity rate ($M \geq 3$) in North China. Small grey circles are all $M \geq 3$ earthquakes during each time period. Large black circles are $M \geq 6$ earthquakes.

Having separated aftershocks from the catalogs, I can examine the background seismicity rate, which forms the foundation for hazard assessment (Cao et al., 1996; Petersen et al., 2014). Unlike the total seismicity rate, background seismicity in North

China is generally stable in space with some variations in time. The regions with high background seismicity rate correlate with major tectonic structures including the Shanxi Rifts, the Zhangjiakou-Penglai (Z-P) fault system, and the Bohai Bay segment of the Tanlu fault (Figure 3.9). After the 1966 Xingtai earthquake and the 1976 Tangshan earthquake, a Tanshan-Hejian-Cixian (T-H-C) active fault system was identified, which experienced thirteen *M*6 and four *M*7 historical earthquakes (J. Xu et al., 1996; Yin et al., 2015). This fault system had slightly higher background seismicity rate than most of North China up to 1996 (Figure 3.9a-c), but was back to "normal" since then (Figure 3.9d-f). Other localities of high seismicity rates in each time period are not clearly associated with mainshocks; they may be related to local stress perturbations. The total background seismicity rate for whole North China stayed nearly the same, with slight decrease in the past decade (Figure 3.10).

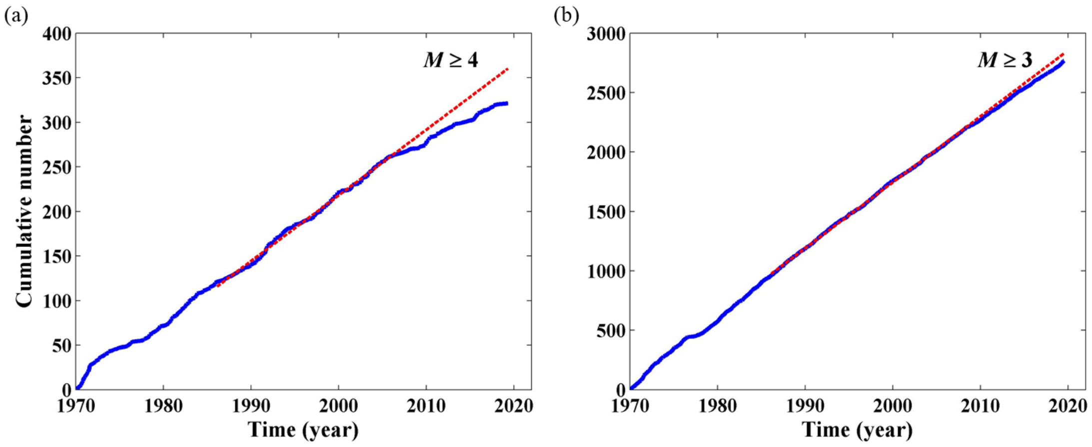


Figure 3.10. The cumulative background earthquakes of $M \geqslant 4$ **(a)** and $M \geqslant 3$ **(b)** in North China. The red lines are linear least-square fits to the average background seismicity rates between 1986 and 2006.

3.5 Correlation between seismicity and strain rate

In the previous sections in this chapter, I mainly focus on the complete instrumental earthquake records after 1970 and the area eastern the Ordos Plateau. In this and later sections, I also use the historical earthquake records and expand the study region to include the Ordos Plateau and its margins (Figure 3.12), so I get more data to describe spatiotemporal patterns as complete as possible and do statistical analysis. The earthquake catalog I used below is a moment-magnitude (M_w) based historical earthquake catalog ($M_w \geq 4$, 780 B.C.-2015) compiled by Cheng et al. (2017). In the broader North China region, the $M \geq 6$ earthquake records are suggested to be complete after 1600 (W.-Q. Huang et al., 1994).

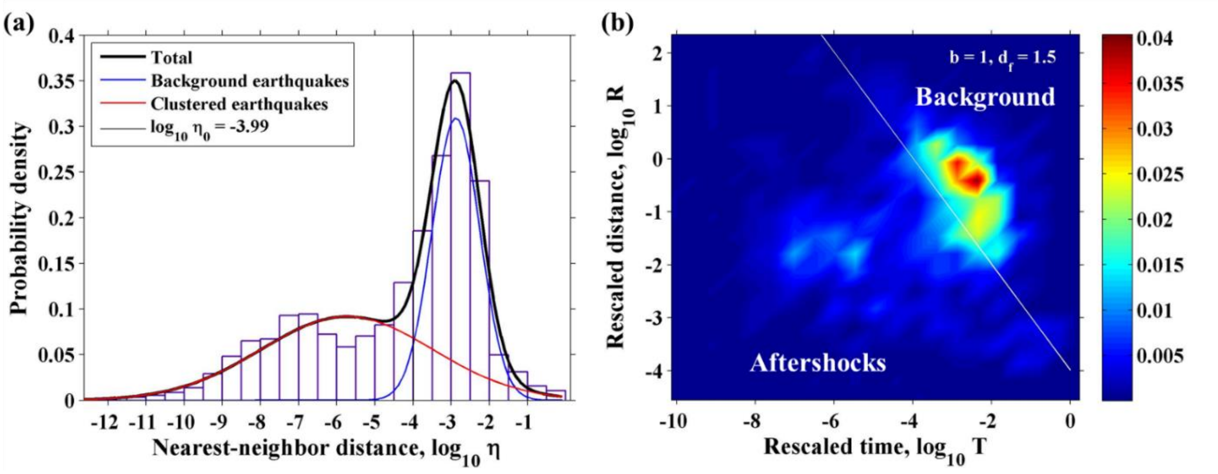


Figure 3.11. Nearest-neighbor method applied to $M_w \geq 4$ earthquakes between 780 B.C. and 2015 in North China: **(a)** 1D density distribution of η, with estimated Gaussian densities for clustered (red) and background (blue) components. The threshold distance η_0 is where the two components are best separated, here at $\log_{10} \eta_0 = -3.99$ (the solid vertical line). **(b)** 2D joint distribution of rescaled distance and time (R, T). The solid line $(\log_{10} T + \log_{10} R = -3.99)$ is the preferred threshold between background and clustered earthquakes. The results are calculated with $b = 1$ and $d_f = 1.5$, and are not sensitive to variations of these values within reasonable ranges.

I declustered the whole historical earthquake catalog of the broader North China region using the NN method; the results are in Figure 3.11. The calculated threshold $\log_{10} \eta_0 = -3.99$ is very close to the threshold $\log_{10} \eta_0 = -4.03$ shown in Figure 3.2, suggesting the NN method is stable with catalog incompleteness. Because most aftershocks of historical earthquakes are missing in the catalog, most events in the catalog are background earthquakes, as indicated by the high event-density cluster in the background domain (above the $\log_{10} \eta_0 = -3.99$ boundary) (Figure 3.11b). The number of aftershocks is relatively small, and hence, the event-density in the aftershock domain is low (Figure 3.11b). Except the horizontal band around $\log_{10} R = -2$, there is an band downward to the right, similar to the result of a single synthetic aftershock series shown in Zaliapin et al. (2008), which suggest that aftershocks was closer to their mainshocks as time went by (the rescaled distance of aftershocks decreases when the rescaled time increases).

After declustering the catalog, I compare the spatial distributions of total seismicity, background seismicity, and aftershocks with the strain rate distribution from Global Strain Rate Model (Figure 3.12) (Kreemer et al., 2014). As shown in Figure 3.12, the central part of the Ordos block has low strain rate, low background seismicity, and is lack of large earthquakes. The high strain rate and concentrated seismicity are found at the west margin of the Ordos block and along the Z-P fault system. Along the Shanxi rift, seismicity is abundant, but strain rate is only medium, mainly in the central and southern parts. In the North China Plain, the strain rate is as low as the one in the central part of the Ordos block, but the background seismicity and large earthquakes are widespread. The aftershock activities concentrate in Xingtai and Tangshan regions because of the 1966 Xingtai and 1976 Tangshan mainshocks. The correlation between strain rate and seismicity in North China is hard to see from the plots, so I do a statistical analysis below.

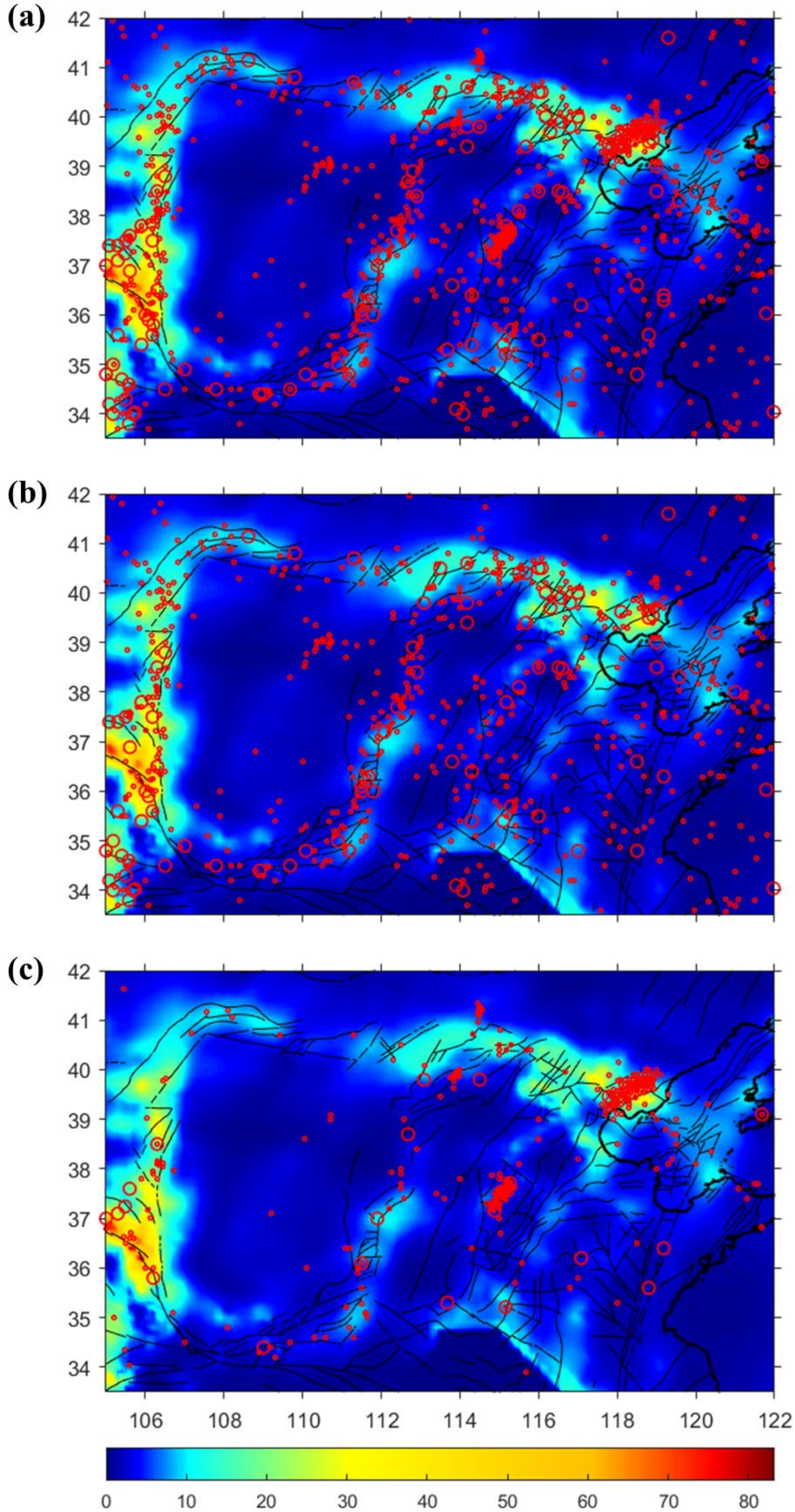

(a)
(b)
(c)
42
41
40
39
38
37
36
35
34
106
108
110
112
114
116
118
120
122
0
10
20
30
40
50
60
70
80
Scalar strain rate, 10^{-9}/year

Figure 3.12. Spatial distribution of scalar strain rate and $M_w \geq 4$ earthquakes between 780 B. C. and 2015. **(a)** Total seismicity. **(b)** Background seismicity. **(c)** Aftershocks. Red circles are earthquake epicenters (Large: $M_w \geq 6$, Small: $4 \leq M_w < 6$). The strain rate data come from Kreemer et al. (2014). The earthquake data come from Cheng et al. (2017).

To quantify this correlation, I follow the approach of Shen et al. (2007) and Zeng et al. (2018) by comparing the cumulative strain rate and earthquakes. I first grid the region according to the resolution of strain rate data into 0.1° by 0.1° cells and then sort the cells by descending strain rate. Strain rate and number of earthquakes in these cells are then separately summed from cells of highest strain rate to cells of lowest strain rate to produce cumulative curves for strain rate and number of earthquakes. For comparing these two curves, the total cumulative strain rate and earthquakes are normalized to unity. The cumulative number of cells is also normalized to unity to express the cumulative proportion of area being scanned. The normalized cumulative strain rates and number of earthquakes are then plotted as a function of the cumulative proportion of scanned area. This kind of plot is actually a flipped version of the Molchan diagram, which is an objective way to test earthquake prediction (Molchan & Kagan, 1992; Zechar et al., 2013) (Figure 3.13). Therefore, the cumulative earthquake curve shows on how well the strain rate predicts previous earthquakes. I use the area skill score introduced by Zechar et al. (2013) to quantify the prediction power. The area skill score is the normalized area below the corresponding cumulative curve. The area skill score = 1 corresponds to a perfect prediction. The area skill score = 0 indicates completely fail of prediction. The area skill score = 0.5 is expected for a random guess (the diagonal line in Figure 3.13).

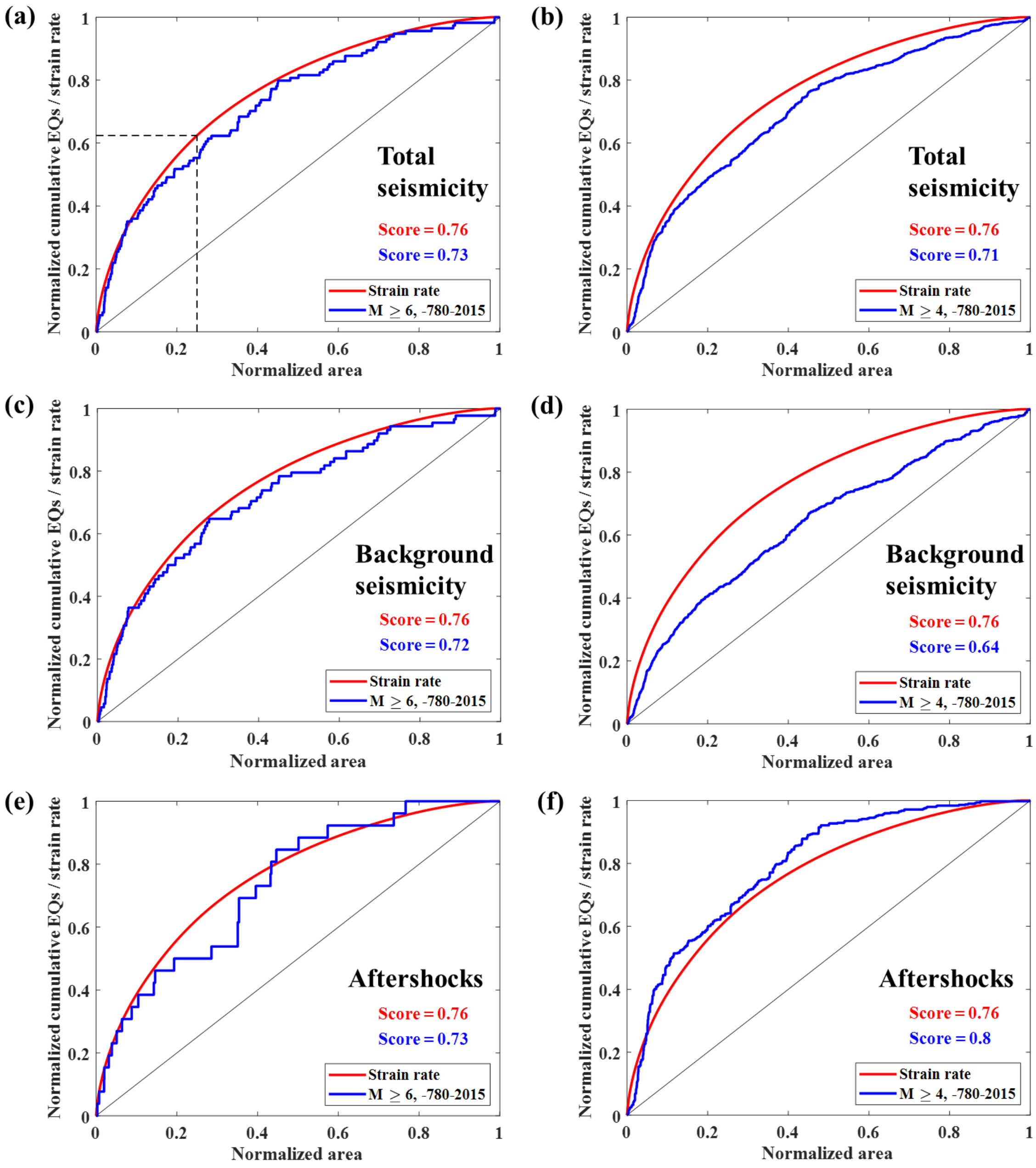


Figure 3.13. Comparison between geodetic strain rate and seismicity in North China. Cumulative strain rate and earthquake count versus fraction of covered area sorted by descending strain rates, with the highest strain rate areas to the left of horizontal axis. The total cumulative strain rate and earthquake count are normalized to unity. **(a), (c), and (e)** are for $M_w \geq 6$ earthquakes. **(b), (d), and (f)** are for $M_w \geq 4$ earthquakes. The black

diagonal line indicates spatial random distribution (area skill score = 0.5). The dash lines are indictors of the strain rate accumulation in 25% area with highest strain rate. The score shown in the figure is the area skill score introduced by Zechar et al. (2013) detailed in main text.

Figure 3.13 shows the cumulative curves for strain rate and different types of seismicity (total seismicity, background seismicity, and aftershocks with $M_w \geq 4$ or $M_w \geq 6$ since 780 B.C. The cumulative curves for strain rate and seismicity are called the strain rate curve and the seismicity curve, respectively. I sort cells according to the strain rate from high to low, so areas with highest strain rate are scanned first (left portion of horizontal axes), then lower strain-rate areas are scanned (right portion of horizontal axes). Therefore, the strain rate curve is concave downward. The earthquakes (e.g. total seismicity, background seismicity, and aftershocks) are also cumulatively counted based on the descending order of strain rate. For total seismicity, either $M_w \geq 4$ or $M_w \geq 6$, the strain rate curves match the seismicity curves well, indicating a strong correlation between the strain rate and total seismicity in North China. The area skill scores of the strain rate and total seismicity are close to each other. For background seismicity, the correlation is still strong for $M_w \geq 6$ earthquakes but weaker for $M_w \geq 4$ earthquakes, suggesting relatively small background earthquakes distributed more broadly than large earthquakes, as shown in Figure 3.12b. For aftershocks, the seismicity correlates well with strain rate for both cut-off magnitudes. For $M_w \geq 4$ aftershocks, the seismicity curve is even above the strain rate curve with a larger area skill score, suggesting aftershocks intensively occurred in high strain rate regions (mainly in the Tangshan region).

Considering the possible impact of the incomplete earthquake records, I have also checked the results for complete records ($M_w \geq 6$ since 1600 and $M_w \geq 4$ since 1970). The general features shown in Figure 3.13 do not change. Because the portion of complete records is a large part of the total records in the catalog, the general features shown in Figure 3.13 are the features mainly for periods with complete records. Specifically, the features for $M_w \geq 6$ earthquakes are mainly the features in 1600-2015 period. The features for $M_w \geq 4$ earthquakes are mainly the features in 1970-2015 period. The general features shown in Figure 3.8 may vary with time and the length of chosen time periods.

However, the strong correlation between strain rate and large earthquakes ($M_w \geq 6$) (Figure 3.13) does not ensure that the spatial prediction power of strain rate in North China is as good as the one in plate boundary zones. As shown in Figure 3.13a, 55% of the large earthquakes occurred in 25% of the area with the area skill score = 0.73, better than random guess, but the portions of predicted earthquakes are much less than the results for the California-Nevada region, which are 86% of large earthquakes in 25% of the area (Zeng et al., 2018). This large contrast is because areas of relatively high strain rate are much more concentrated in major faults in plate boundary zones (e.g., California) than in diffused fault systems in intraplate regions (e.g., North China). I also counted the number of cells and number of $M_\mathrm{w} \geq 6$ earthquakes that occurred in these cells at different ranges of strain rate (strain rate bins) (Figure 3.14a) to characterize the correlation between strain rate and seismicity. They are normalized to relative frequency so that they can be compared. The results show that, in North China, around 71% of the area have strain rates less than 5 x 10^{-9} but only produced around 38% of $M_\mathrm{w} \geq 6$ earthquakes (Figure 3.14a), so $M_\mathrm{w} \geq 6$ earthquakes tend to occur in regions with relatively high strain rate. To further explore the dependence of $M_w \geq 6$ earthquakes on strain rate, I used the number of $M_w \geq 6$ earthquakes divides the number of cells for each strain-rate bin, and obtained the number of earthquakes per cell for different strain-rate bins (Figure 3.14b). The results show that in general, regions of relatively high strain rate produce more $M_w \geq 6$ earthquake (Figure 3.14b).

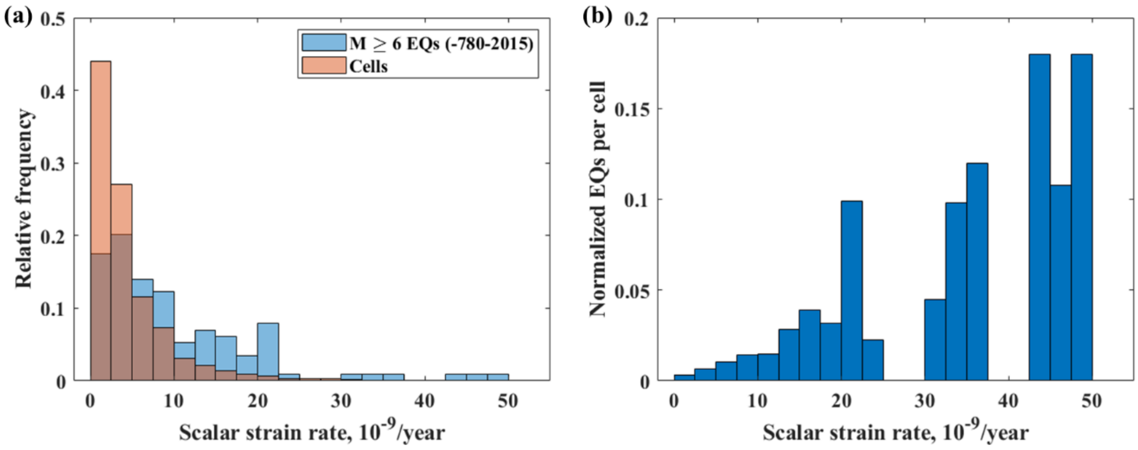


Figure 3.14. (a) Relative frequency distributions of scalar strain rate for the cells and the $M_w \geq 6$ earthquakes occurred within the cells. Shaded area is the overlapping part of these two distributions. If the two distributions are totally overlapped, it means that the earthquakes are not correlated with strain rate. **(b)** Average earthquake occurrence per cell for the given period. The earthquake occurrence per cell for strain rate bines is normalized to unity.

As I have shown in Chapter 2, large earthquakes tend to have a Devil's staircase temporal pattern, with the alternation of active and quiescent periods. The complete records of $M_w \geq 6$ earthquakes in North China show an active period at first between

1600 and 1750, followed by a relatively quiescent period (1750-1900), then another active period since 1900 (Figure 3.15a). I compared the cumulative curves of seismicity in these three periods with the strain rate curve, the results show that the seismicity curves for the two active periods match the strain rate curve well (Figure 3.15b), but the seismicity curve for the quiescent period is significantly below the strain rate curve and close to the spatially random distribution (the diagonal line in Figure 3.15b). Therefore, the spatial patterns of large earthquakes may be different in active periods and quiescent periods.

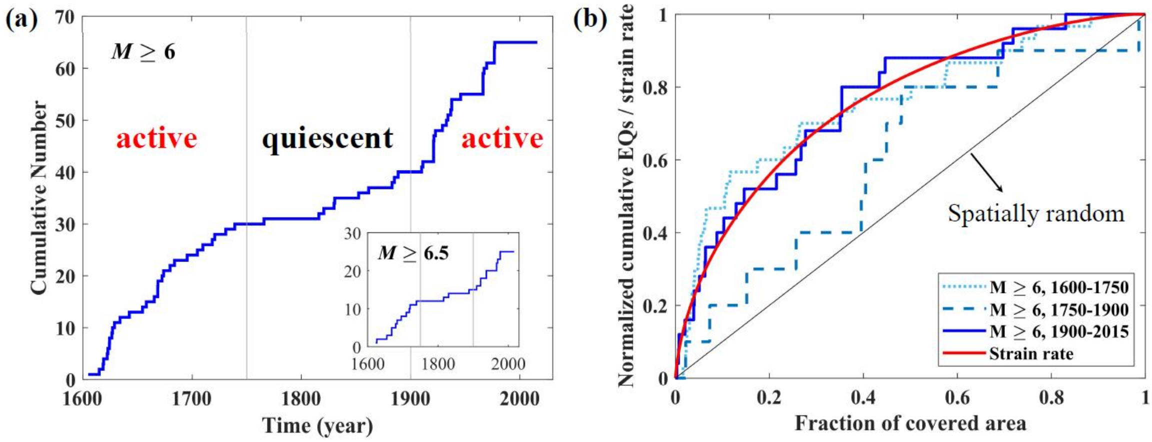


Figure 3.15. (a) Temporal pattern of $M \geq 6$ earthquakes ($M \geq 6.5$ for inset) in North China with two active periods (1600-1750 and 1900-2015) separated by a relatively quiescent period (1750-1900). **(b)** comparison of correlations between strain rate and seismicity in these three periods. The diagonal line suggests a random distribution of earthquakes.

I also explore the temporal variations of the cumulative earthquake curves by using smaller earthquakes. I used both the incomplete records of $M_w \geq 4$ background earthquakes between 1900 and 1969 and the complete records of $M_w \geq 4$ background earthquakes since 1970. For the complete records, I calculate each seismicity curve for a 10-year period window and then move the time window at a 2-year step. For the incomplete records, I calculate in the same way but with a 30-year window to ensure there are enough earthquakes in each window. The results of complete background record since 1970 show that the background seismicity curve matches the strain rate curve at the beginning time periods (red colored curves) and gradually move downward away from the strain rate curve (Figure 3.16a). The blue colored curves with midyear after 2007 are significant below the strain rate curve, suggesting $M_w \geq 4$ earthquakes are more widely distributed in both high and low strain rate regions. Combined with lack of $M_w \geq 6$ earthquakes in the past 40 years (Figure 3.15a), it seems that North China have entered a quiescent period after the 1976 Tangshan earthquake. For the incomplete records between 1900 and 1969, because of limited data and 30-year period window, the results are less reliable with worse time resolution than the ones from complete records, but I found a trend of the background seismicity curve moving from random distribution to strain-rate correlated distribution (Figure 3.16b). However, there is no transition from quiescent period to active period for large earthquakes between 1900 and 1969 (Figure 3.15a). During this period, low strain-rate regions (strain rate $< 5 \times 10^{-9}$) still produced around 38% of $M_w \geq 6$ earthquakes. One possible explanation is that strong correlation between the strain rate and seismicity means large earthquakes are more likely to occur in high strain rate region, but during the quiescent periods, large earthquakes may occur more randomly in space than during clustered periods. The correlations between strain rate and seismicity in different tectonic settings are compared in Chapter 5. The effect of the length of time window is also explored om Chapter 5.

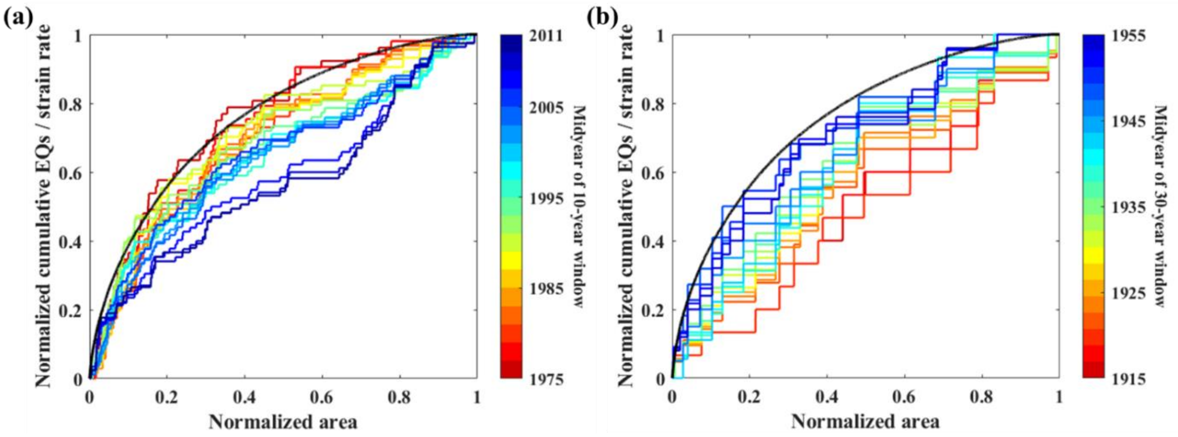


Figure 3.16. (a) The cumulative strain rate (black) and $M_w \geq 4$ earthquake counts within 10-year windows that move in 2-year steps from 1970 to 2015. **(b)** Similar plot for $M_w \geq 4$ earthquakes but with 30-year windows that move in 2-year steps from 1900 to 1970. The colorbar shows the midyear of the moving time windows used to count the cumulative earthquakes.

3.6 Correlation between seismicity and mapped faults

Large earthquakes concentrate and repeat on major faults in plate boundary regions, but in intraplate regions, the fault systems are complex and consist of widespread faults and hidden faults. Large intraplate earthquakes seem to roam between different faults and often occur at unexpected places (M. Liu et al., 2011). The Xingtai and Tangshan earthquakes in North China occurred at the places where no active tectonics or active faults were recognized before the occurrence of these events. Moreover, in intraplate Brazil, most of earthquakes occur near but not directly on the major neotectonic faults (Assumpção et al., 2014). Therefore, I did a statistical analysis on the correlation between seismicity and mapped faults.

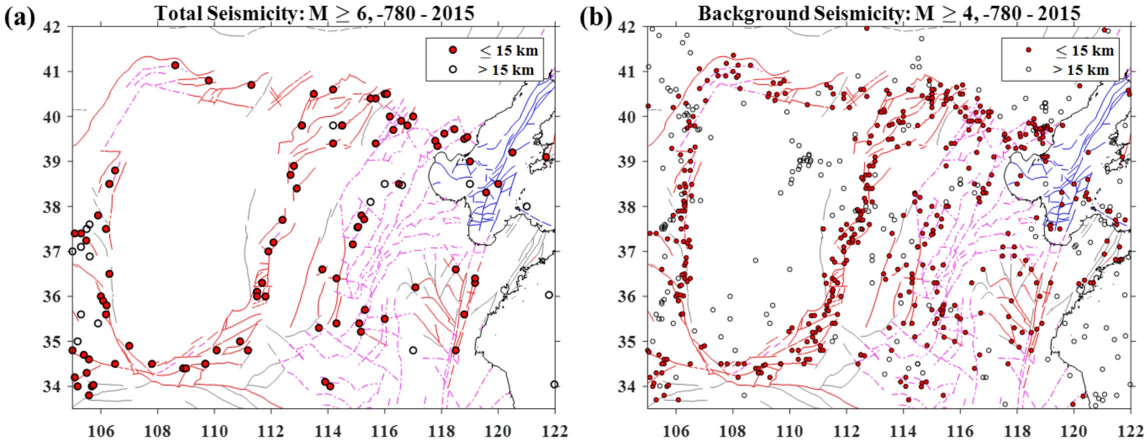


Figure 3.17. Comparison between seismicity and faults. **(a)** Total seismicity with $M_w \geq 6$. **(b)** Background seismicity with $M_w \geq 4$. Red solid lines are Holocene faults. Magenta dashed lines are hidden faults. Grey solid lines are Pleistocene faults. Blue lines are marine faults. Red solid circles are epicenters closer than 15 km from faults. Open circles are epicenters more than 15 km away from any fault.

The fault data I used are from Deng et al. (2007). The mapped faults are divided into four types: Late Pleistocene and Holocene faults, Quaternary faults, hidden faults in plain or basin region, and marine faults. A Late Pleistocene and Holocene fault is a fault that moved in the last 120,000 years. A Quaternary fault is a fault that moved in the

Quaternary but its movement since Late Pleistocene is uncertain. A hidden fault is a fault without evidence of surface ruptures. Faults in the plain or basin region are likely to be hidden faults, because of the covering of loose Quaternary sediment. Most of hidden faults in North China Plain moved in Early and Middle Pleistocene but have no evidence of movement since Late Pleistocene. A marine fault is a fault on sea floor.

The spatial distribution of mapped faults and seismicity in North China is shown in Figure 3.17. The mapped faults are plenty and widespread. The Late Pleistocene and Holocene faults (solid red) are mainly around the Ordos Plateau and near the Tanlu fault. In North China Plain, the hidden faults (dashed magenta) are the major type. The marine faults (blue) distribute in the Bohai Sea. The Quaternary faults are least in number and diffusely distributed. Most earthquakes occurred on or near the mapped faults (Figure 3.17). The Late Pleistocene and Holocene faults generate the largest portion of earthquakes, but many events also occurred on or near the hidden faults (Figure 3.18). Small number of earthquakes occurred on or near the Quaternary and marine faults (Figure 3.18). However, significant number of events (16% for $M_w \geq 6$ events) are observed more than 15 km away from the mapped faults (unmapped) (Figure 3.18). Background earthquakes ($M_w \geq 4$) occurred on the unmapped faults with a larger portion compared with the large earthquakes ($M_w \geq 6$) (Figure 3.18), consistent with the observation that background seismicity is more widespread than large earthquakes (Figure 3.17).

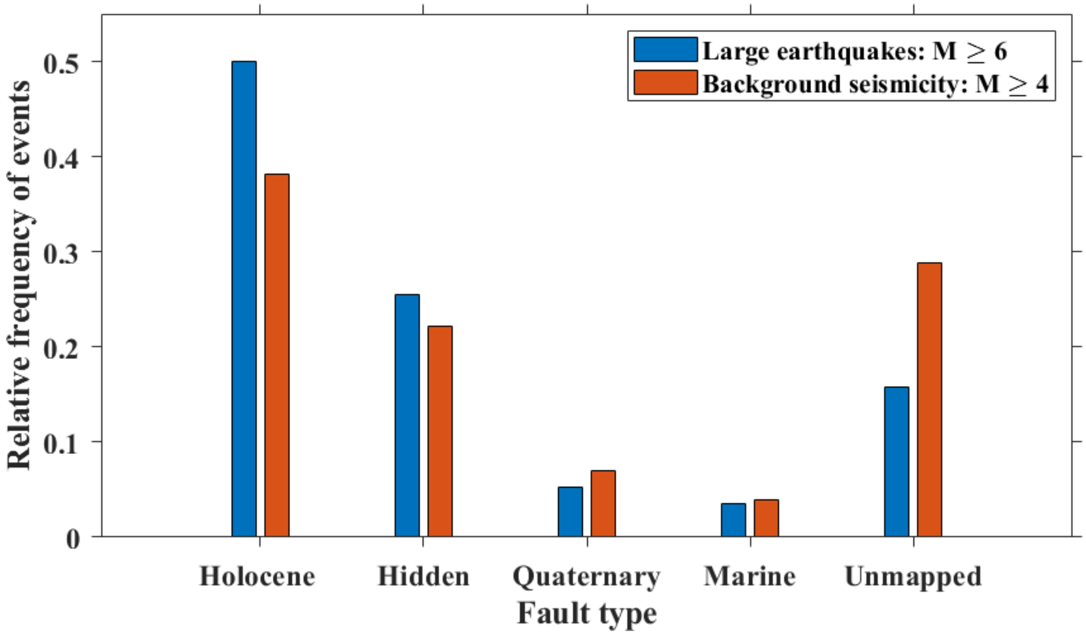


Figure 3.18. Relative frequency distribution of earthquakes on different types of faults.

The results above suggest that the fault distribution is also a good constraint on the spatial distribution of earthquakes. How well the spatial distribution of earthquakes can be indicated by the combined strain rate and the fault data? To answer this question, first, I discretize the fault data into 0.1° by 0.1° cells, same resolution as for the strain rate data, and call them the fault cells. Then the area diagram similar to that for strain rate is constructed. The difference is on how to scan the cells. Here, I divide the 0.1° by 0.1° cells of North China into two categories: the fault cells and ground cells. The fault cells are scanned first by descending strain rate because earthquakes are more likely to occur on faults than outside faults. After all fault cells are scanned, I start to scan the ground cells by descending strain rate (the red curves in Figure 3.19). I also tried to specify the fault cells according to their fault type as shown in Figure 3.18. The way is scanning the cells of the Late Pleistocene and Holocene faults first, then the cells of the hidden faults, then the cells of marine and Quaternary faults, lastly scanning the ground cells (the black curves in Figure 3.19). Each kind of cells are sorted by descending strain rate. Doing so means that I assume that the rifting faults around the Ordos and the Holocene segment of the Tanlu fault are most likely to produce large earthquakes, the hidden faults in North China Plain are the second priority for large earthquake occurrences. The results are shown in Figure 3.19. The area skill score for strain rates increased after combining with the fault data. This is more clear for large earthquakes (Figure 3.19a). Thus, fault data provide additional constraints on the spatial distribution of earthquakes. The earthquake curves based on strain rate and fault data are below the earthquake curves based only on strain rate data in the left part of the plot (cells of higher strain rate). It suggests that some earthquakes occurred on the unmapped faults but in high strain rate regions, and better “prediction” can be made in low strain rate region when scanning areas with mapped faults first. Giving priority to the Late Pleistocene and Holocene faults and hidden faults makes a little improvement on the large earthquake “prediction”, but the general feature does not change. The way that I combine the strain rate and fault data is straightforward but not the best. Possible improvements may be made by using only fault data for low strain rate regions or giving different weight to strain rate and fault data, or assigning wider fault areas.

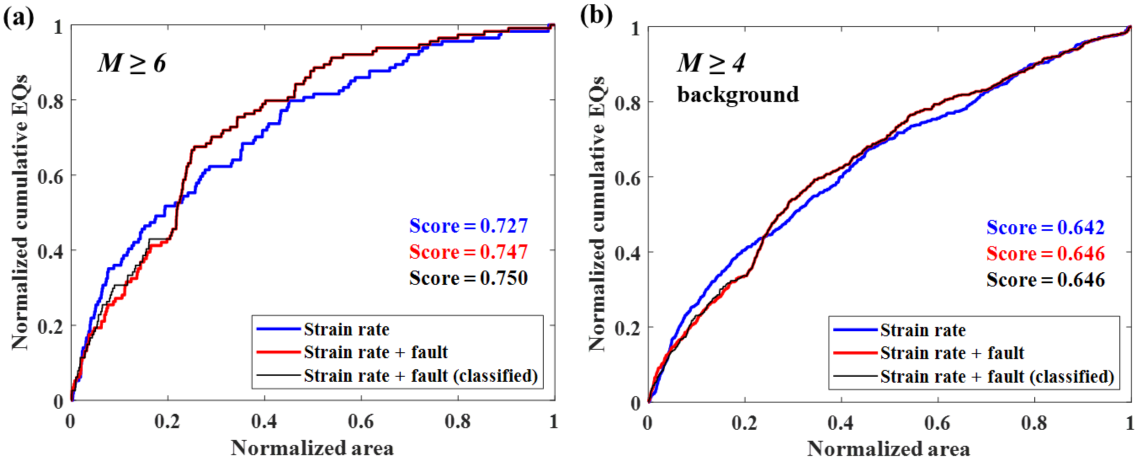


Figure 3.19. Cumulative number of earthquakes versus fraction of covered area sorted based on strain rate or the combination of strain rate and fault data. **(a)** Total seismicity with $M_w \geq 6$. **(b)** Background seismicity with $M_w \geq 4$. The area skill scores are color coded for the different curves.

3.7 Correlation between seismicity and seismic velocity

Mooney et al. (2012) analyzed the correlation between global intraplate seismicity and seismic velocity and found that intraplate seismicity tends to concentrate around edges of cratons, defined by the S-wave anomalies from global seismic tomography at a depth of 175 km. Assumpção et al. (2014) found that the S-wave anomalies at 100 km depth in Brazil show some correlations with the spatial distribution of seismicity; A large percentage of epicenters are found above S-wave anomalies between -1 and -4% and in the range from +3 to +5%. Here I compare the spatial distribution of earthquakes with S-wave anomalies in North China. I used the S-wave anomalies at 100 km depth (Figure 3.20) from Crust and Upper Mantle Velocity Model of North China v2.0 (http://www.craton.cn/data), because the S-wave anomalies at 100 km depth show a good spatial resolution in the checkboard experiments (Zhao et al., 2012) and may reflect the lithospheric thinning and cratonic edge effect suggested by previous studies (Assumpção et al., 2014; Mooney et al., 2012). The S-wave anomalies are obtained by inverting body wave travel-time data (Zhao et al., 2009; Zhao et al., 2012). The data resolution is 0.5° by 0.5°.

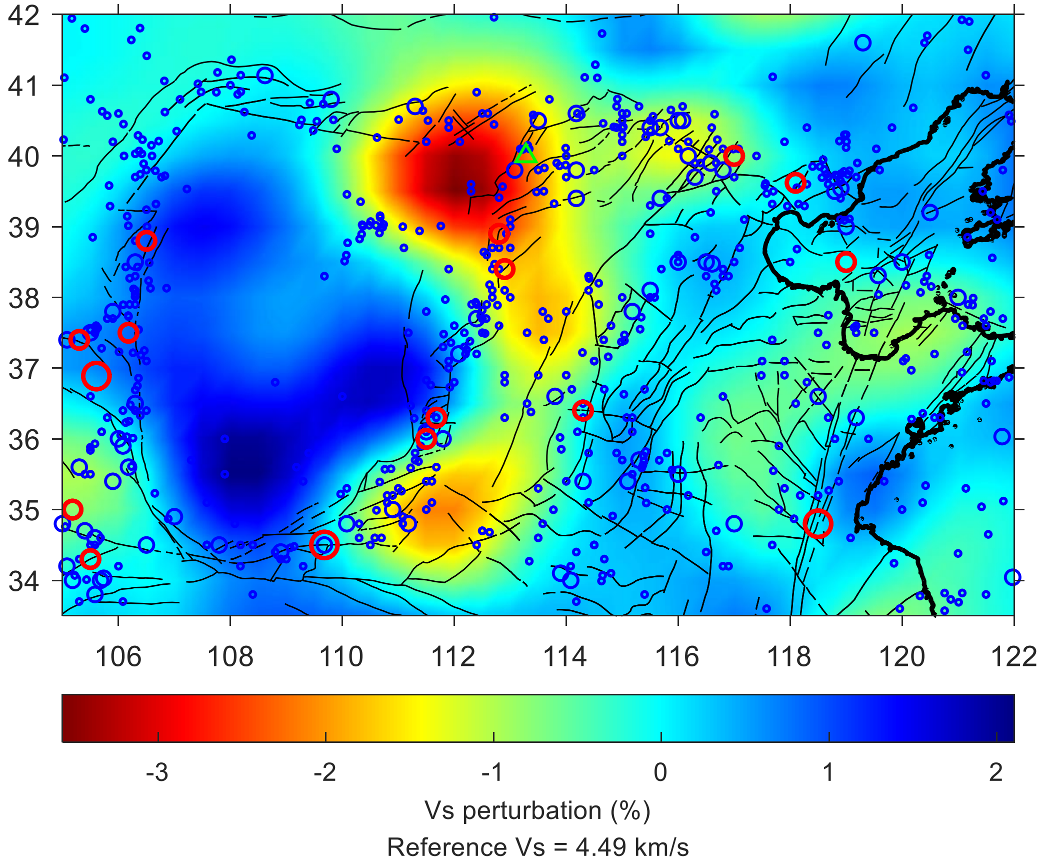


Figure 3.20. Comparison of S-wave anomaly with seismicity. Blue circles are background earthquake epicenters (Large: $6 \leq M_w < 7$, Small: $4 \leq M_w < 6$). Red circles are total earthquake epicenters (Large: $M_w \geq 8$, Medium: $7 \leq M_w < 8$). Green triangle indicates the Datong volcano.

The spatial distribution of the S-wave anomaly generally matches the geological setting in North China, with the highest S-wave velocities in the center of the Ordos Plateau and the lowest S-wave velocities near the Datong volcano (the green triangle in Figure 3.20). Most of the largest earthquakes ($M_w \geq 7$) in North China (red circles in Figure 3.20) occurred around the edges of Ordos block, a relic of the thermally rejuvenated North China craton with a thick lithosphere (Bao et al., 2011), where the S-wave anomalies are between 0 and 0.5% (Figure 3.21a). However, for smaller background earthquakes ($M_w \geq 4$), the distribution of S-wave anomaly largely overlaps with the distribution of S-wave anomaly for the cells (Figure 3.21b), meaning a weak correlation between the S-wave anomaly with background seismicity. To some extent, the areas with low S-wave velocity (anomaly between -2.5% and -0.5%) had more earthquakes than expectation of random distribution, and fewer earthquakes occurred in the areas with high S-wave velocity (anomaly larger than 1%) (Figure 3.21). The lack of seismicity in high S-wave velocity area is due to the cold and strong Ordos block. The relative concentration of seismicity in area of low S-wave velocity may be related with lithospheric thinning in North China. Overall, seismic velocity anomaly alone is not a good spatial predictor of seismicity (close to random guess, shown in Figure 3.22) but may offer some constraints on large earthquakes.

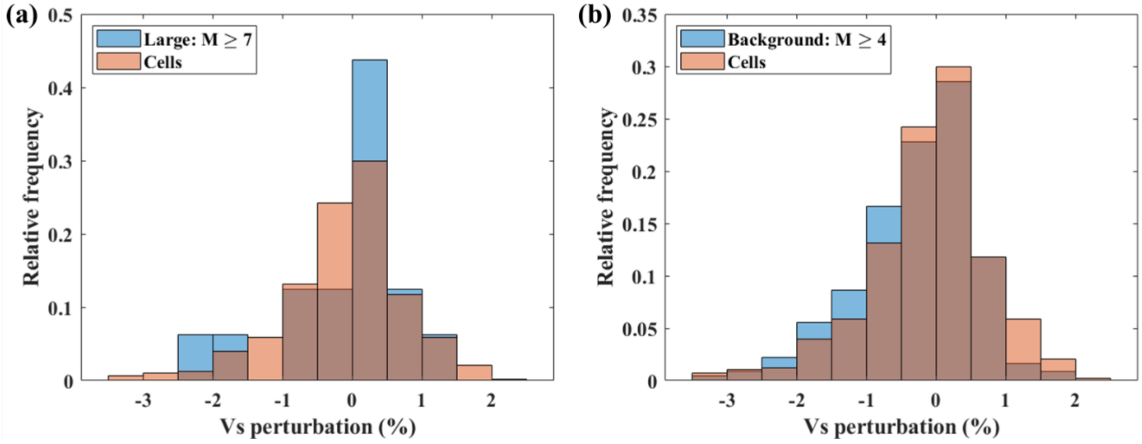


Figure 3.21. Relative frequency distributions of S-wave anomaly for the 0.5° by 0.5° cells and earthquakes occurred in these cells. Shaded area is the overlapping part of these two distributions. If the two distributions are totally overlapped, it means the S-wave anomaly does not correlate with seismicity. **(a)** Large seismicity with $M_w \geq 7$. **(b)** Background seismicity with $M_w \geq 4$.

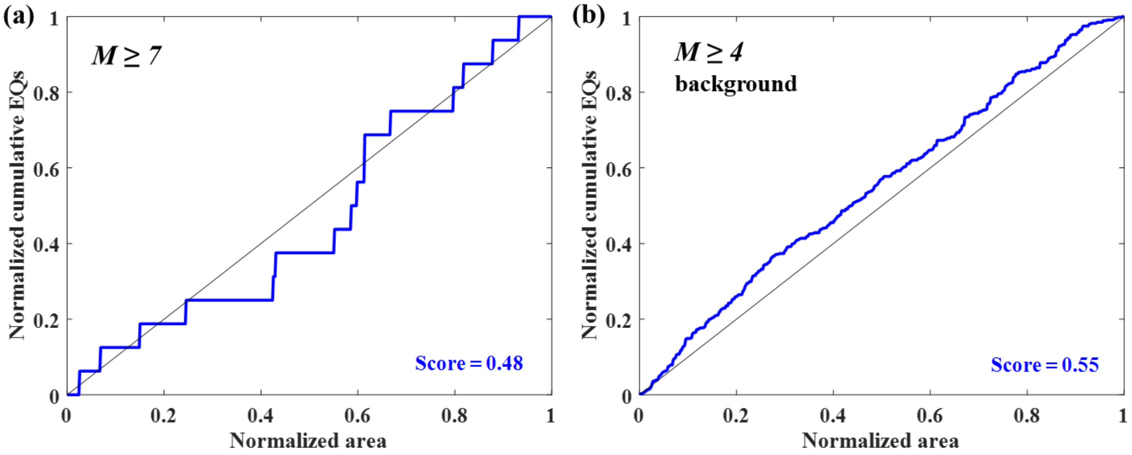


Figure 3.22. Cumulative number of earthquakes versus fraction of covered area sorted by ascending S-wave anomaly. **(a)** Total seismicity with $M_w \geq 7$. **(b)** Background seismicity

with $M_w \geq 4$. The area skill scores are shown in the southeast side of Figure. The diagonal line suggests a random distribution of earthquakes (area skill score = 0.5).

3.8 Discussion

3.8.1 Aftershocks: controversy and uncertainty

One motivation for this chapter is to answer the question regarding recent moderate earthquakes ($M \geq 4.5$) in the Tangshan region: Are they aftershocks of the 1976 Great Tangshan earthquake? Our results, based on the nearest-neighbor method and Omori's law fitting, indicate that the recent moderate earthquakes are likely the aftershocks of the 1976 Great Tangshan earthquake. A supporting indicator is the stress increase in their epicentral regions by the 1976 Tangshan earthquake (Figure 3.7). In other words, they may be triggered by the Great Tangshan earthquake, and therefore are its aftershocks. The same conclusion was reached by Zhong and Shi (2012) and M. Liu and Wang (2012). On the other hand, it is clear from Figure 3.4 that the statistical properties (distance to the main shock in the space-time-magnitude domain) of these recent events are near the boundary separating aftershocks from background earthquakes, so they could also be identified as background earthquakes. This uncertainty of aftershock identification is consistent with the result of Y. Liu et al. (2020) that the probability of the 2019 M_s 4.5 event being triggered by the 1976 Great Tangshan earthquake is 49.4%.

The difficulty of identify or rejecting these events as aftershocks begins with the definition of aftershocks. The term aftershock implies a causative relation with the mainshock and foreshock in an earthquake nucleation process (M. Liu & Stein, 2019;

Scholz, 2002). However, proving this causative relationship has been difficult for natural earthquakes. Some workers think that aftershock, foreshock, and mainshock are events fundamentally the same; they are in a sequence of cascading ruptures triggered by a common mechanism (Felzer et al., 2004; Helmstetter, 2003). In this view, they are just different names for earthquakes in a different part of the earthquake sequence.

In practice, aftershocks usually refer to the smaller earthquakes following a large earthquake (the mainshock) in the same rupture area, and are assumed to be triggered by the mainshock through changes of fault frictional properties and stress (M. Liu & Stein, 2019). They are assumed to occur in the neighborhood of the mainshock with their number decaying with time as described by the Omori's law (Omori, 1894; Utsu, 1961), which is applied in most declustering methods. The main difference of these declustering methods is in choosing the space and time windows for aftershocks, and the reference for the decaying seismicity, i.e., the background seismicity rate. In this chapter, I used the nearest neighbor (NN) method, which statistically identifies aftershocks as clustered events, to decluster the earthquake catalog of North China. Like all statistical methods, the NN method is based only on the statistical property exhibited by events in the catalogs; it has no bearing on the causative physics that is implied in the definition of aftershocks. Thus, the NN may miss some aftershocks that are stress triggered by a mainshock but whose nearest-neighbor distance is beyond the threshold for aftershock identification. The Omori's law is useful to describe the general patterns of decaying seismicity rates in aftershock sequences, but it is not useful for identifying individual aftershock events, because within the duration of aftershocks, an event could be either an aftershock or a background earthquake.

3.8.2 Background seismicity: spatiotemporal variability

One important application of declustering is to obtain background seismicity, which is essential for hazard assessment (van Stiphout et al., 2012). Gardner and Knopoff (1974) and Reasenberg (1985) proposed that background seismicity, obtained by their declustering methods, follow a stationary Poisson process. These two important studies have “established the commonly accepted null model for declustered seismicity - a Poisson point field that is stationary in time and inhomogeneous in space” (Zaliapin & Ben-Zion, 2020). However, Luen and Stark (2012) used the declustering methods of Gardner and Knopoff (1974) and Reasenberg (1985) with different magnitude cut-off catalogs and more powerful statistical tests; they found the declustered catalogs are not consistent with the stationary Poissonian model. Zaliapin and Ben-Zion (2020) developed a more elaborated statistical declustering algorithm based on the NN method and applied it to a global catalog (2000-2015, *M* 5.0-9.1) and earthquake catalogs from California. Their results suggested that the hypothesis of stationary background seismicity cannot be rejected for catalogs with magnitude range (between the maximum and minimum magnitudes) less than 4, but can be rejected for catalogs with magnitude range larger than 4.

Defining background seismicity rates is more difficult in intraplate regions, where earthquake records are often incomplete and not long enough to reflect long term, slow tectonic loading. The background seismicity in North China as identified in this chapter are generally stable in space but with variations in time (Figure 3.9). Regions with high background seismicity rate are found in major tectonic structures, including the Shanxi

Rift and the Z-P fault system, where geodetic strain rates are high and large historic earthquakes were concentrated (Figure 3.12). These results are similar to the spatial patterns of background seismicity rate reported by Xiong et al. (2019). On the other hand, background seismicity shows local temporal variations. The localized highs of background seismicity rate roam with time, with no clear correlations with the epicentral regions of large earthquakes (Figure 3.9). These variations may reflect activity switching among subregions and indicate local stress perturbations. Similar results are found in the background seismicity in California (Zaliapin & Ben-Zion, 2020).

3.8.3 Trade-off between completeness and length of earthquake catalogs

Long and complete records of earthquakes are important for seismic hazard analysis, especially for intraplate regions, where tectonic loading is slow. In this chapter, I used both historical and instrumental earthquake records to explore the spatiotemporal patterns of earthquakes in North China. For the earthquake records in North China, long catalog and complete catalog cannot be satisfied simultaneously, I have to scarify the completeness for the length of records, so I can use as much data as possible. Large earthquakes did occur at places where current seismicity rate is low (Figure 3.1), possibly so do future large earthquakes. Therefore, more historical and paleoseismic earthquake studies are needed to describe long-term behavior of intraplate seismicity and be used in seismic hazard assessment.

3.8.4 Correlation between strain rate and seismicity

The good correlation between seismicity and strain rate is the precondition for using strain rate in earthquake forecast, but a good correlation does not necessarily make strain rate a good earthquake predictor. As I showed in Figure 3.13, the degree of correlation is defined by how well the seismicity curve matches the strain rate curve, but the prediction power is described by the area below the earthquake curve (area skill score). The results show that the strain rate correlates well with large earthquakes in North China, but its prediction power is lower than that in plate boundary regions.

Our results show that large earthquakes tend to concentrate in regions of high strain rate in active periods but diffusely distributed in a quiescent period (Figure 3.15). The rate change of large earthquakes seems to be reflected by the spatiotemporal evolution of background seismicity. During the active period, background earthquakes gradually concentrated within areas of high strain rate and then became more diffusely distributed in both high and low strain rate regions (Figure 3.16). These results are similar to the spatiotemporal patterns of earthquakes in the California-Nevada shown in Zeng et al. (2018). How the correlation between strain rate and seismicity varies in different tectonic settings are explored in Chapter 5.

3.8.5 Implications for intraplate seismic hazard

The declustered earthquake catalogs, assumed containing only independent background events, are often used as input for seismic hazard assessments (Allen et al., 2018; Mueller, 2019). However, as I have shown in this chapter, separating aftershocks from background earthquakes can be challenging. Most declustering methods rely on statistical property variations in the catalogs and are limited by the quality of the catalogs.

The separation of aftershocks from background seismicity in statistical models is not attached to the underlying physics. As Zaliapin and Ben-Zion (2020) stated: “we do not even know if natural seismicity actually operates in these terms (background, clustered, etc.)”.

The challenge is greater in intraplate regions, where slow tectonic loading is shared by complex systems of faults and fault networks. Over the long term, intraplate earthquakes tend to be temporally clustered and spatially roaming (M. Liu & Stein, 2016; M. Liu & Wang, 2012). In North China, my analysis of the catalog since 1970 shows relatively high background seismicity rates in the Shanxi rift and the Z-P fault system (Figure 3.9), where geodetic strain rates are also relatively high and large historic earthquakes were concentrated (Figure 3.12). However, the correlation between background seismicity and strain rates or historic earthquakes is not clear within the North China Plain. Numerous large historic earthquakes, including the 1668 *M* 8.5 Tancheng earthquake on the Tanlu fault, as well as the 1966 Xingtai earthquakes occurred where neither the background seismicity nor geodetic strain rates are significantly above the normal. In the more stable and less active North America craton, Kreemer et al. (2018) found that seismicity has no correlation with strain rate in a plate scale. The overall variation of background seismicity rate near Tangshan and the rest of the North China Plain in recent years (Figure 3.9) provides no clear indication that the region has entered a new period of more active seismic activity (X. Xu & Deng, 1996).

3.9 Conclusions

I used the nearest-neighbor method to separate clustered earthquakes (taken as aftershocks) from background earthquakes in North China and did statistical analysis on the spatial correlation of seismicity with strain rate, fault, and seismic velocity. The results indicate that the aftershock sequence of the 1976 Great Tangshan earthquake is likely continuing. The recent moderate ($M \geq 4.5$) earthquakes, including the 2020/07/12 M_s 5.1 Tangshan earthquake, are statistically identified as aftershocks and may be physically triggered by the 1976 Great Tangshan earthquake. The recent background seismicity rate in the Tangshan region is lower than that before the Great Tangshan earthquake for both $M \geq 4.0$ and $M \geq 3.0$ catalogs. The estimated aftershock duration is around 65-100 years.

The background seismicity in North China is relatively stable in space with variations in time. Relatively high background seismicity is found in the Shanxi Rift and the Z-P fault system, where geodetic strain rates are also relatively high and large historic earthquakes concentrated. However, such correlations are weak in the North China Plain. Numerous large historic earthquakes occurred where neither background seismicity rate nor geodetic strain rate is significantly above the normal, highlighting the complexity of intraplate earthquakes. The temporal variations of background seismicity in North China suggest no clear correlation with large earthquakes in the past century. The variation of background seismicity rate in North China in recent years provides no clear indication that the region has entered a new period of active seismic activity.

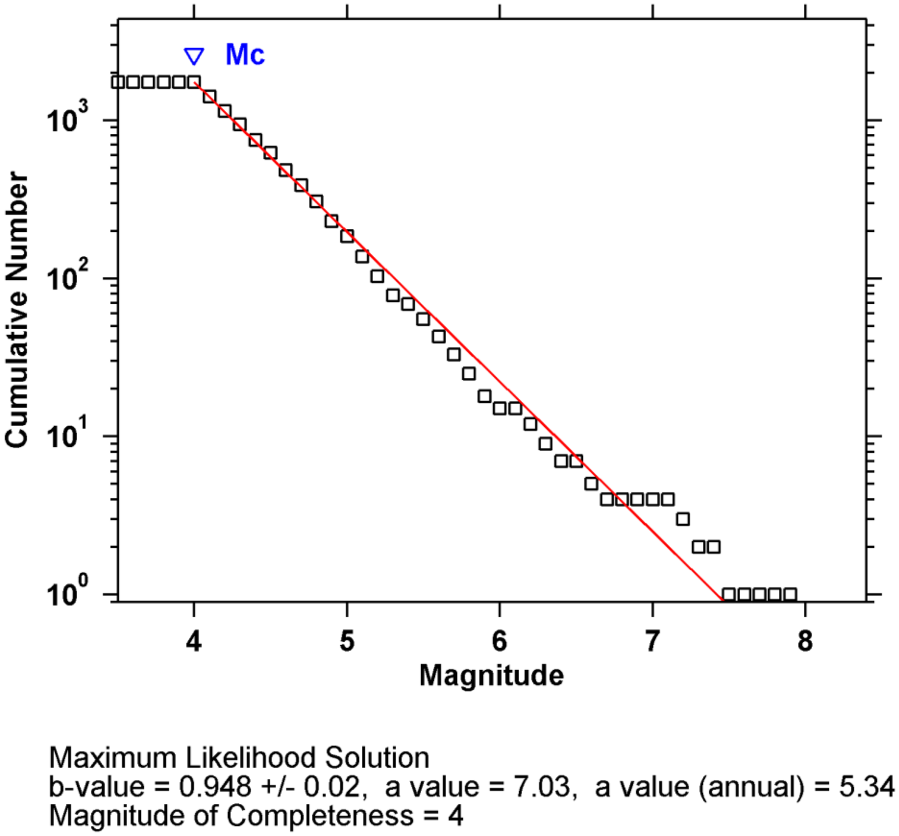
In North China, good correlation is found between large earthquakes and strain rate, but the correlation is worse for smaller earthquakes. The power of strain rate as a spatial predictor of seismicity in North China is significantly better than random guess, but it is not as good as in plate boundary regions, because strain rate is diffuse in North China. The correlation between strain rate and seismicity are strongly time-dependent. Background seismicity gradually concentrated within areas of high strain rate from 1900 to 1970 when the rate of $M \geq 6$ events was high, but low strain-rate regions still produced ~38% of large earthquakes. Since 1976 the rate of $M \geq 6$ earthquakes dropped in North China, so is the correlation between strain rate and earthquakes. These results suggest that the spatial correlation between strain rate and seismicity varies with time, magnitudes, and tectonic settings, all need to be considered when using strain rate for forecasting and hazard assessment.

The fault distribution is also a good constraint on the spatial pattern of seismicity. Many large earthquakes occurred not only on the Late Pleistocene and Holocene faults but also near or on hidden faults. However, some large earthquakes also occurred away from the mapped faults in the high strain rate regions. Background seismicity is more diffusely distributed than large earthquakes with an increased portion on unmapped faults. Combining fault data with strain rate can improve the prediction performance for large earthquakes.

The correlation of seismicity with the S-wave anomaly at 100 km depth is weak, although high velocity regions are aseismic and low velocity regions tend to have little more earthquakes than random distribution. The largest earthquakes ($M \geq 7$) in North China seems to concentrate in the areas with the S-wave anomalies between 0 and 0.5%, roughly corresponding to the edges of the Ordos block.

3.10 Supplementary materials

The b-value of $M \geq 4$ earthquakes in North China is calculated using Equation 1.2. The result is $b = 0.95 \pm 0.02$ (Figure 3.23).



Maximum Likelihood Solution
b-value = 0.948 +/- 0.02, a value = 7.03, a value (annual) = 5.34
Magnitude of Completeness = 4

Figure 3.23. The b-value estimation for $M \geq 4$ earthquakes since 1970 in North China by using the maximum likelihood method.

The d_f of $M \geq 4$ earthquakes in North China is calculated based on Equation 1.6. The result is $d_f = 1.49 \pm 0.02$ (Figure 3.24).

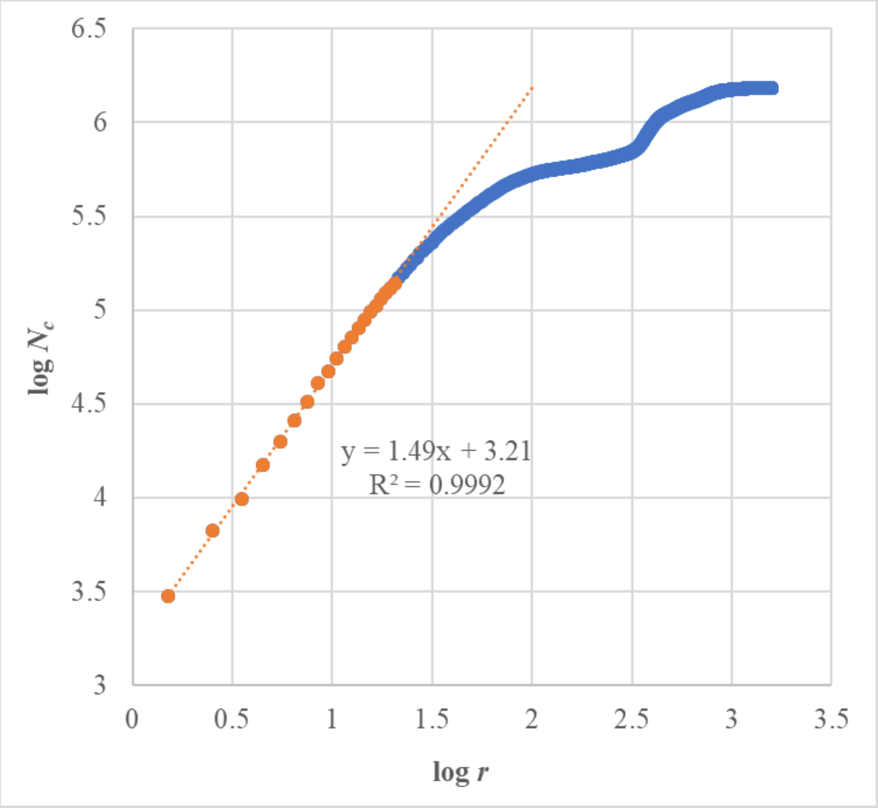


Figure 3.24. The **fractor dimention $\boldsymbol{d_f}$** for $M \geq 4$ earthquakes since 1970 in North China. The value of $\boldsymbol{d_f}$ is estimated by the "correlation integral method" (Grassberger & Procacia, 1983). In practice, all possible spatial distances (r) between two epicenters are considered, and the cumulative number (N_c) of event pairs less than r is computed. The slope of the line of log N vs. log r is the fractal dimension $\boldsymbol{d_f}$.

The separation boundary between background seismicity and aftershocks in the NN method is not sensitive to the variations of the value of b and d_f (Figure 3.25).

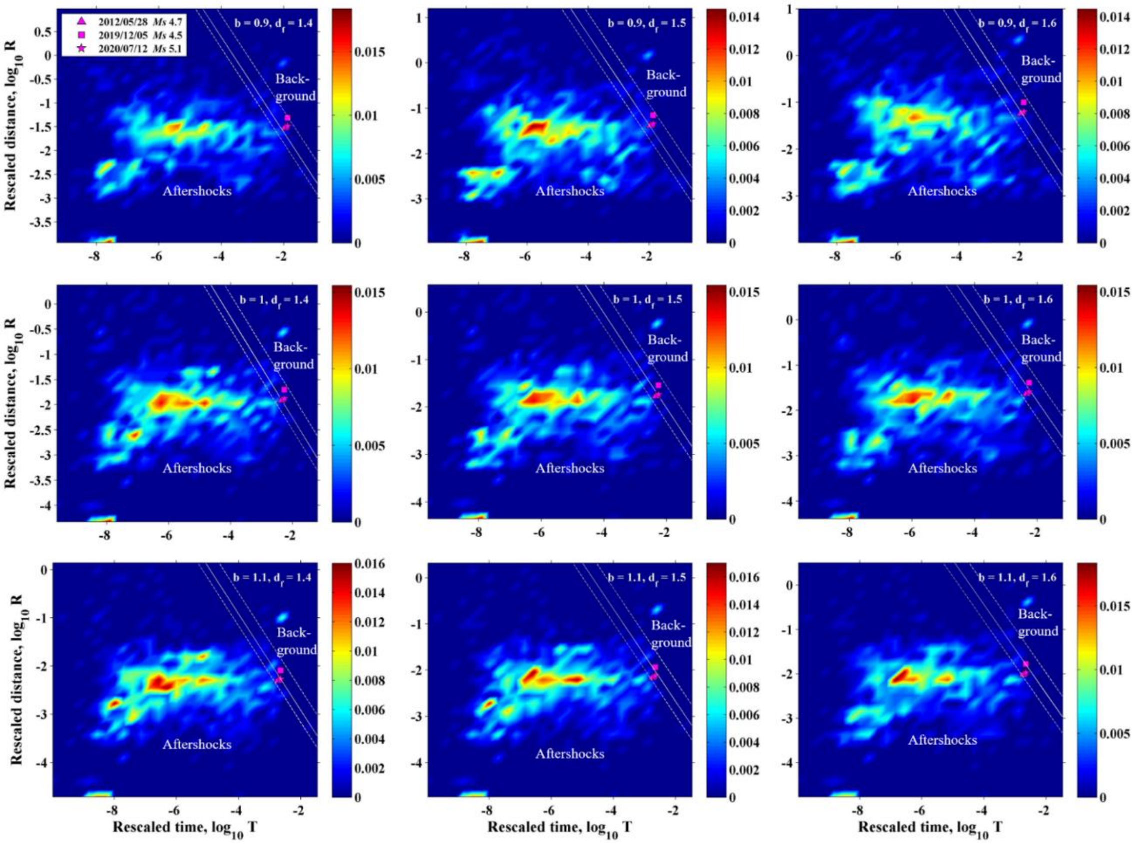


Figure 3.25. Nearest-neighbor method applied to $M \geq 4$ earthquakes since 1970 in the Tangshan region by using different values of *b*-values and d_f. The threshold distance (solid lines) and the range of ≥ 90% confidence level (dashed lines) are based on earthquakes from the entire North China (see Figure 3.2).

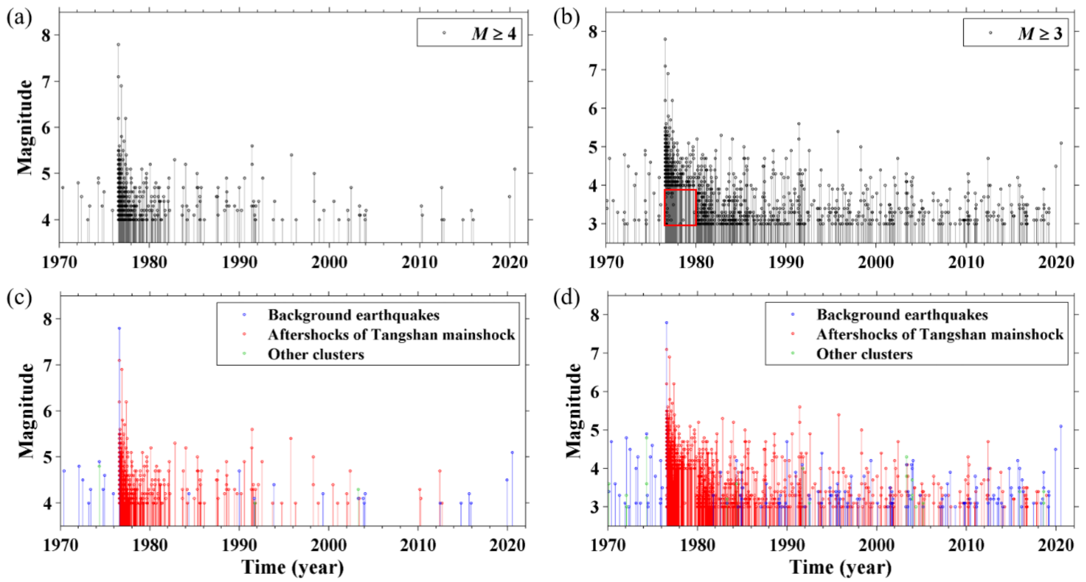


Figure 3.26. (a, b) Magnitude-time plot of earthquakes in the Tangshan region with different cut-off magnitude. The red rectangle in (b) indicates the incomplete records (missing small events shortly after the mainshock). **(c, d)** Same plots with earthquakes identified by the NN method.

References

Aki, K. (1965). Maximum likelihood estimate of b in the formula log N= a-bM and its confidence limits. *Bull. Earthq. Res. Inst., Tokyo Univ., 43*, 237-239.

Allen, T. I., Gibson, G., Ghasemi, H., & Leonard, M. (2018). *The 2018 National Seismic Hazard Assessment for Australia: earthquake epicentre catalogue*: Geoscience Australia.

Assumpção, M., Ferreira, J., Barros, L., Bezerra, F., FranÇa, G. S., Barbosa, J. R., et al. (2014). Intraplate seismicity in Brazil. *Intraplate Earthquakes. Cambridge University Press, New York*, 50-71.

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E, 69*(6), 066106.

Bao, X., Xu, M., Wang, L., Mi, N., Yu, D., & Li, H. (2011). Lithospheric structure of the Ordos Block and its boundary areas inferred from Rayleigh wave dispersion. *499*(1-4), 132-141.

Cao, T., Petersen, M. D., & Reichle, M. S. (1996). Seismic hazard estimate from background seismicity in southern California. *Bulletin of the Seismological Society of America, 86*(5), 1372-1381.

Chen, Y., Liu, M., & Luo, G. (2020). Complex Temporal Patterns of Large Earthquakes: Devil's Staircases. *Bulletin of the Seismological Society of America, XX*, 1–13, doi: 10.1785/0120190148.

Chen, Y., Tsoi, K. L., Chen, F. B., Gao, Z. H., Zou, Q. J., & Chen, Z. L. (1988). *The great Tangshan earthquake of 1976: an anatomy of disaster* (1 ed.). Oxford: Pergamon Press.

Cheng, J., Rong, Y., Magistrale, H., Chen, G., & Xu, X. (2017). An M w-based historical earthquake catalog for Mainland China. *Bulletin of the Seismological Society of America, 107*(5), 2490-2500.

Deng, Q., Ran, Y., Yang, X., Min, W., & Chu, Q. (2007). Active Tectonic Map of China (in Chinese).

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research: Solid Earth, 99*(B2), 2601-2618.

Felzer, K. R., Abercrombie, R. E., & Ekström, G. (2004). A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America, 94*(1), 88-98. http://www.scopus.com/inward/record.url?eid=2-s2.0-1842864389&partnerID=40&md5=671084486bf7e57026c29c0e6b70705f

Frankel, A. (1995). Mapping seismic hazard in the central and eastern United States. *Seismological Research Letters, 66*(4), 8-21.

Gardner, J., & Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America, 64*(5), 1363-1367.

Grassberger, P., & Procacia, I. (1983). Measuring the strangeness of strange attractors. *Physica. D, 9*(1-2), 189-208.

Gu, G., Lin, T., Shi, Z., & Li, Q. (1983). Earthquake catalog of China: 1831 BC–1969 AD.

Guo, H., Jiang, W., & Xie, X. (2017). Multiple faulting events revealed by trench analysis of the seismogenic structure of the 1976 Ms7. 1 Luanxian earthquake, Tangshan Region, China. *Journal of Asian Earth Sciences, 147*, 424-438.

Gutenberg, B., & Richter, C. F. (1954). Seismicity of the earth and related phenomena. *Princeton*.

Hainzl, S., Christophersen, A., Rhoades, D., & Harte, D. (2016). Statistical estimation of the duration of aftershock sequences. *Geophysical Journal International, 205*(2), 1180-1189.

Helmstetter, A. (2003). Is Earthquake Triggering Driven by Small Earthquakes? *Physical Review Letters, 91*(5), 058501. http://link.aps.org/doi/10.1103/PhysRevLett.91.058501

Hicks, A. L. (2011). *Clustering in Multidimensional Spaces with Applications to Statistics Analysis of Earthquake Clustering.*

Huang, B.-S., & Yeh, Y. T. (1997). The fault ruptures of the 1976 Tangshan earthquake sequence inferred from coseismic crustal deformation. *Bulletin of the Seismological Society of America, 87*(4), 1046-1057.

Huang, W.-Q., Li, W.-X., & Cao, X.-F. (1994). Research on completeness of earthquake data in the Chinese mainland (II)—The regional distribution of the beginning years of basically complete earthquake data. *Acta Seismologica Sinica*, *7*(4), 529-538.

Iwata, T. (2008). Low detection capability of global earthquakes after the occurrence of large earthquakes: Investigation of the Harvard CMT catalogue. *Geophysical Journal International*, *174*(3), 849-856.

Jiang, C., Wu, Z., & Zhuang, J. (2013). ETAS model applied to the Earthquake-Sequence Association (ESA) problem: the Tangshan sequence. *CHINESE JOURNAL OF GEOPHYSICS-CHINESE EDITION*, *56*(9), 2971-2981.

Kreemer, C., Blewitt, G., & Klein, E. C. (2014). A geodetic plate motion and Global Strain Rate Model. *Geochemistry, Geophysics, Geosystems, 15*(10), 3849-3889.

Kreemer, C., Hammond, W. C., & Blewitt, G. (2018). A robust estimation of the 3-D intraplate deformation of the North American plate from GPS. *Journal of Geophysical Research: Solid Earth, 123*(5), 4388-4412.

Liu, M., & Stein, S. (2016). Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards. *Earth-Science Reviews, 162*, 364-386.

Liu, M., & Stein, S. (2019). Earthquake, Aftershocks. In H. K. Gupta (Ed.), *Encyclopedia of Solid Earth Geophysics* (pp. https://doi.org/10.1007/1978-1003-1030-10475-10477_10204-10471). The Netherland: Springer International Publishing.

Liu, M., Stein, S., & Wang, H. (2011). 2000 years of migrating earthquakes in North China: How earthquakes in midcontinents differ from those at plate boundaries. *J Lithosphere, 3*(2), 128-132.

Liu, M., & Wang, H. (2012). Roaming earthquakes in China highlight midcontinental hazards. *Eos, Transactions American Geophysical Union, 93*(45), 453-454.

Liu, M., Wang, H., Ye, J., & Cheng, J. (2014). Intraplate earthquakes in North China. In P. Talwani (Ed.), *Intraplate Earthquakes* (pp. 97-125). New York: Cambridge University Press.

Liu, Y., Zhuang, J., & Jiang, C. (2020). Background Seismicity before and after the 1976 Ms 7.8 Tangshan Earthquake: Is Its Aftershock Sequence Still Continuing? *Seismological Research Letters*.

Luen, B., & Stark, P. B. (2012). Poisson tests of declustered catalogues. *Geophysical Journal International*, *189*(1), 691-700.

Marsan, D., & Lengline, O. (2008). Extending earthquakes' reach through cascading. *Science, 319*(5866), 1076-1079.

Mignan, A., Jiang, C., Zechar, J., Wiemer, S., Wu, Z., & Huang, Z. (2013). Completeness of the Mainland China earthquake catalog and implications for the setup of the China Earthquake Forecast Testing Center. *Bulletin of the Seismological Society of America*, *103*(2A), 845-859.

Molchan, G., & Kagan, Y. Y. (1992). Earthquake prediction and its optimization. *Journal of Geophysical Research: Solid Earth, 97*(B4), 4823-4838.

Mooney, W. D., Ritsema, J., & Hwang, Y. K. (2012). Crustal seismicity and the earthquake catalog maximum moment magnitude (Mcmax) in stable continental regions (SCRs): Correlation with the seismic velocity of the lithosphere. *Earth and Planetary Science Letters, 357*, 78-83.

Mueller, C. S. (2019). Earthquake catalogs for the USGS national seismic hazard maps. *Seismological Research Letters, 90*(1), 251-261.

Omori, F. (1894). *On the after-shocks of earthquakes* (Vol. 7): The University.

Peresan, A., & Gentili, S. (2018). Seismic clusters analysis in Northeastern Italy by the nearest-neighbor approach. *Physics of the Earth and Planetary Interiors, 274,* 87-104.

Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D., et al. (2014). *Documentation for the 2014 update of the United States national seismic hazard maps*. Retrieved from

Petersen, M. D., Mueller, C. S., Moschetti, M. P., Hoover, S. M., Rukstales, K. S., McNamara, D. E., et al. (2018). 2018 one-year seismic hazard forecast for the central and eastern United States from induced and natural earthquakes. *Seismological Research Letters, 89*(3), 1049-1061.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982. *Journal of Geophysical Research: Solid Earth, 90*(B7), 5479-5495.

Robinson, R., & Zhou, S. (2005). Stress Interactions within the Tangshan, China, Earthquake Sequence of 1976. *Bulletin of the Seismological Society of America, 95*(6), 2501-2505.

Scholz, C. H. (2002). *The mechanics of earthquakes and faulting* (2nd ed.). Cambridge, UK ; New York: Cambridge University Press.

Shen, Z.-K., Jackson, D. D., & Kagan, Y. Y. (2007). Implications of Geodetic Strain Rate for Future Earthquakes, with a Five-Year Forecast of M5 Earthquakes in Southern California. *Seismological Research Letters, 78*(1), 116-120.

Stein, S., & Liu, M. (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature, 462*(7269), 87-89.

Toda, S., & Stein, R. S. (2018). Why Aftershock Duration Matters for Probabilistic Seismic Hazard Assessment. *Bulletin of the Seismological Society of America, 108*(3A), 1414-1426.

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

van Stiphout, T., Zhuang, J., & Marsan, D. (2012). Theme V-Models and Techniques for Analyzing Seismicity. *Seismicity declustering, community online resource for statistical seismicity analysis*.

Wang, R., Lorenzo-Martin, F., & Roth, F. (2006). PSGRN/PSCMP - a new code for calculating co- and post-seismic deformation, geoid and gravity changes based on the viscoelastic-gravitational dislocation theory. *Computers & Geosciences, 32*(4), 527-541. <Go to ISI>://000236782000010

Xiong, Z., Zhuang, J., & Zhou, S. (2019). Long-Term Earthquake Hazard in North China Estimated from a Modern Catalog. *Bulletin of the Seismological Society of America, 109*(6), 2340-2355.

Xu, H., Sun, Y., & Wu, Z. (2018). The effect of lithospheric structure on the seismic deformation - taking the 1976 Tangshan earthquake and 2001 Kunlunshan earthquake as an exacmole. *CHINESE JOURNAL OF GEOPHYSICS-CHINESE EDITION, 61*(8), 3170-3184.

Xu, J., Niu, L., Wang, C., & Han, Z. (1996). Tangshan-Hejian-Cixian newly-generated seismotectonic zone. *Seismological Geology (in Chinese), 18*(3), 193-198.

Xu, X., & Deng, Q. (1996). Nonlinear characteristics of paleoseismicity in China. *Journal of Geophysical Research: Solid Earth, 101*(B3), 6209-6231. http://dx.doi.org/10.1029/95JB01238

Yin, A., Yu, X., Shen, Z.-K., & Liu-Zeng, J. (2015). A possible seismic gap and high earthquake hazard in the North China Basin. *Geology*.

Zaliapin, I., & Ben-Zion, Y. (2016). A global classification and characterization of earthquake clusters. *Geophysical Journal International, 207*(1), 608-634.

Zaliapin, I., & Ben-Zion, Y. (2013). Earthquake clusters in southern California I: Identification and stability. *Journal of Geophysical Research: Solid Earth, 118*(6), 2847-2864.

Zaliapin, I., & Ben-Zion, Y. (2020). Earthquake declustering using the nearest-neighbor approach in space-time-magnitude domain. *Journal of Geophysical Research: Solid Earth, 125*(4), e2018JB017120.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters, 101*(1), 018501.

Zechar, J. D., Schorlemmer, D., Werner, M. J., Gerstenberger, M. C., Rhoades, D. A., & Jordan, T. H. (2013). Regional earthquake likelihood models I: First-order results. *Bulletin of the Seismological Society of America, 103*(2A), 787-798.

Zeng, Y., Petersen, M. D., & Shen, Z. K. (2018). Earthquake potential in California-Nevada implied by correlation of strain rate and seismicity. *Geophysical Research Letters, 45*(4), 1778-1785.

Zhao, L., Allen, R. M., Zheng, T., & Hung, S. H. (2009). Reactivation of an Archean craton: Constraints from P-and S-wave tomography in North China. *Geophysical Research Letters, 36*(17).

Zhao, L., Allen, R. M., Zheng, T., & Zhu, R. (2012). High-resolution body wave tomography models of the upper mantle beneath eastern China and the adjacent areas. *Geochemistry, Geophysics, Geosystems, 13*(6).

Zhong, Q., & Shi, B. (2012). Aftershock duration of the 1976 Ms 7.8 Tangshan earthquake and implication for seismic hazard estimation. *Acta Seismol Sin, 34*(4), 494-508.

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.

Chapter 4: Possible long-lived aftershocks in the Central and Eastern United State and its implications for seismic hazard assessment

4.1 Introduction

Long-lived aftershock activity in intraplate regions is not a new idea. In 1983, Basham et al. (1983) studied the spatiotemporal patterns of seismicity in the continental margin of eastern Canada and noticed continuing earthquakes near the epicenters of two magnitude 7 earthquakes even though 50 years had passed. Similar continuing seismicity were found at other places without known large events, so they speculated that the long-lived continuing seismicity is aftershocks of previous large events. Dieterich (1994), according to rate-and-state friction of fault properties, proposed that aftershock duration, the time for seismicity rates to return to the background seismicity rate, is directly proportional to the product of a fault constitutive parameter and the normal stress, but inversely proportional to the rate of stressing on the fault. This means that the low rates of shear stressing on intraplate faults would cause relatively long aftershock durations. Ebel et al. (2000), based on the Omori-law decay of aftershocks and historic intensity data, suggest that current clusters of seismicity in the Central and Eastern United State (e.g., New Madrid, Charleston, and Charlevoix) are aftershocks of mainshocks that occurred hundreds of years ago. Stein and Liu (2009) estimated the aftershock durations and fault loading rates for 16 large interplate and intraplate earthquakes and found an inverse correlation between aftershock durations and fault loading rates that is consistent with Dieterich (1994)'s model. Boyd et al. (2015) found that current crustal deformation in the New Madrid seismic zone (NMSZ) can be explained by a dynamic model of postseismic frictional afterslip from the 1450 and February 1812 Reelfoot fault earthquakes, so it is likely that many current earthquakes in the NMSZ are aftershocks of the 1811-1812 events. Chapman et al. (2016) proposed that modern seismicity near Summerville, South Carolina, are the long-lived aftershocks of the 1886 Charleston earthquake, because the seismicity shares common seismic characteristics with the aftershock sequence of the 2011 Virginia earthquake. Toda and Stein (2018) analyzed the aftershock durations of large earthquakes on faults with a wide range of slip rate based on the Omori's law and found that aftershocks can last more than 2000 years on the slowest fault (0.01 mm/yr).

However, the extent of long-lived aftershocks in the Central and Eastern United State (CEUS) is still contentious. Page and Hough (2014) argued that current seismicity in New Madrid, Missouri, is background rather than aftershocks, because it is unlikely to be reproduced by an epidemic-type aftershock sequence (ETAS) model. The ETAS model is a statistical model based on the Omori decay law, which assumes that all earthquakes potentially trigger their own aftershocks (Ogata, 1988). The ETAS model is widely used in the studies about aftershocks (Helmstetter et al., 2006; Zhuang et al., 2019). Fereidoni and Atkinson (2014) also suggests contemporary seismicity in the St. Lawrence Valley is unlikely aftershocks of the 1663 Charlevoix earthquake, because the observed current seismicity rate is significantly higher than the expected aftershock rate of the 1663 Charlevoix earthquake after a long-time decay. For the eastern Tennessee seismic zone (ETSZ), Levandowski and Powell (2019) modeled an extreme case: aftershocks of an *M* 7.9 events ruptured the entire ETSZ 500 years ago, as well as some more reasonable cases. They conclude that at least 90% of modern seismicity is background earthquakes caused by localized release of long-term loading in eastern Tennessee.

Long-lived aftershocks are not considered in current seismic hazard assessment in the CEUS. The earthquake catalog used for the 2018 USGS national hazard map is declustered by the distance-time window method of Gardner and Knopoff (1974), which is based on an empirically analysis of California earthquakes (C. S. Mueller, 2019). For Gardner and Knopoff (1974)'s declustering method, the aftershock duration of a magnitude 8 event is around three years, much shorter than possible hundreds or thousands of years lasting aftershocks in stable continents like the CEUS. The USGS used both fault models and smoothed background seismicity models for the CEUS (Petersen et al., 2014; Petersen et al., 2020). The fault models are only applied to several fault zones (New Madrid, Wabash Valley, Charleston, Cheraw fault, and Meers fault). Other regions used the smoothed background seismicity models based on Frankel (1995), which assume that historical background seismicity is a good predictor for future seismic hazard and tend to diminish hazard for areas lack of historical records. Therefore, as mentioned by Stein and Liu (2009) and Toda and Stein (2018), long-lived aftershocks may be misidentified as background seismicity, thus overestimate the hazard in presently active areas and underestimate it elsewhere. The seismic hazards estimated by the fault models for major fault zones are based on complex logic trees but highly depend on the weight of in-cluster sequence or out of cluster sequence that are manually assigned by experts. The possible long-lived aftershocks may play as an indicator of out of cluster sequence, so may affect the weight assignment in the fault models.

In this chapter, I statistically identified the possible long-lived aftershocks of known large historical earthquakes in the CEUS by using the nearest-neighbor method, focusing on three regions: New Madrid, Charleston, and Charlevoix. I quantified aftershocks by the percentage of current seismicity, estimated the uncertainties caused by mainshock magnitudes and locations, and discussed the implications for hazard assessment.

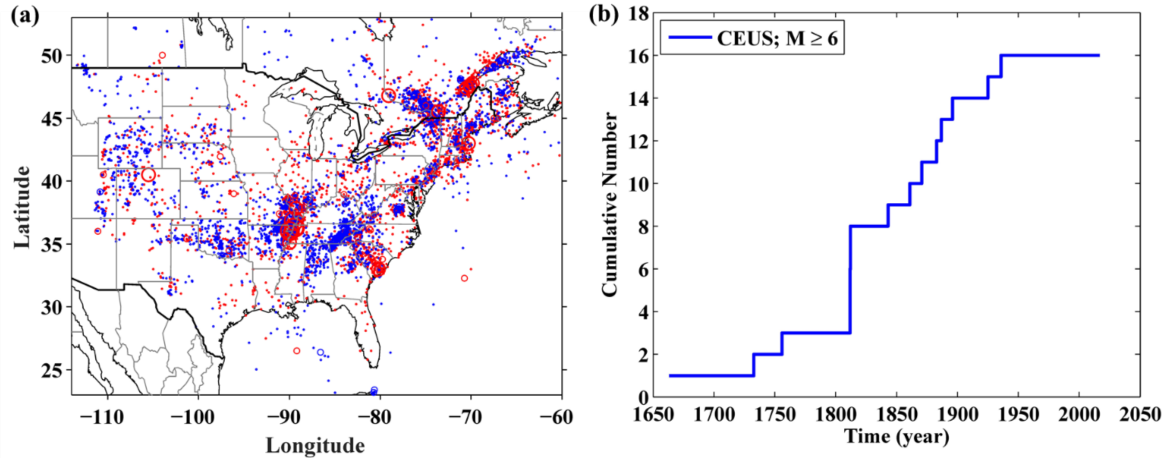


Figure 4.1. (a) Spatial patterns of $M \geq 2.5$ earthquakes in the CEUS. Red circles are events between 1568 and 1979 (incomplete records). Blue circles are events between 1980 and 2016 (complete records). **(b)** Temporal pattern of cumulative $M \geq 6$ earthquakes in the CEUS.

4.2 Data and Method

The earthquake catalog (Figure 4.1a) used in this chapter is a natural earthquake catalog (induced earthquakes are excluded), the raw data for the 2018 USGS national seismic hazard map (https://www.sciencebase.gov/catalog/item/5ad771d1e4b0e2c2dd256fd2) (C. S. Mueller, 2019; Petersen et al., 2020). The catalog is moment magnitude based. The magnitudes of earthquakes used in this chapter are all based on the catalog introduced in C. S. Mueller (2019), may be slightly different from other publications. I use the notation M to represent moment magnitude in the rest of this chapter. The USGS hazard map used the $M \geq 2.7$ catalog and suggests that the catalog is complete for the whole CEUS after 1980 (Petersen et al., 2020). To ensure the reliability of the results, I need to use as many data as possible. I did a simple Gutenberg-Richter frequency-magnitude fit and found the b-value is close to 1 and the magnitude of completeness is 2.5 for events after 1980 (Figure 4.2), so I use the $M \geq 2.5$ and depth ≤ 70 km catalog for analysis in this chapter. I also estimated that the fractal dimension (d_f) of the CEUS epicenters by using the correlation integral method (Grassberger & Procacia, 1983). I found that the value of d_f close to 1 (Figure 4.3), indicating line sources are dominated in the CEUS. Therefore, I use $b = 1, d_f = 1$ to calculate the nearest-neighbor distance in this chapter.

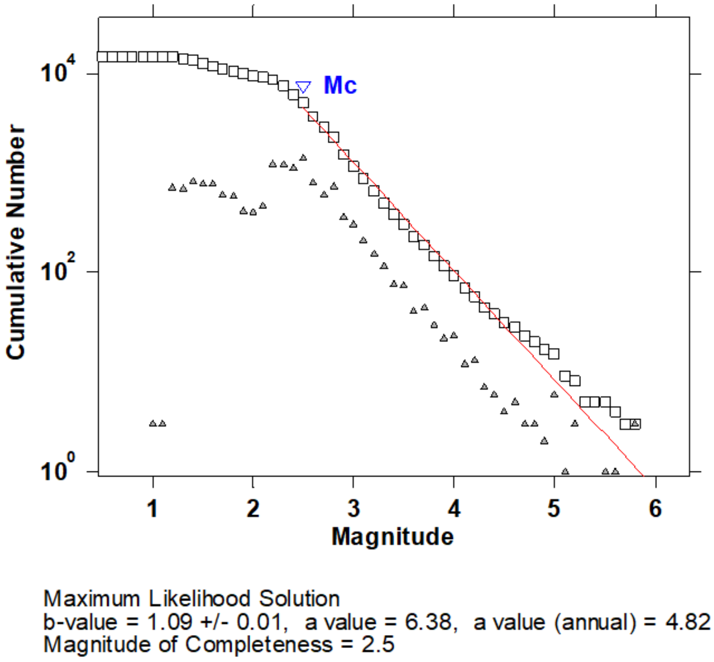


Figure 4.2. The frequency-magnitude distributions of earthquakes since 1980 in CEUS. The squares indicate the cumulative number of evens. The triangles indicate the frequency of events for each bin of 0.1 magnitude increase. The red line is fitted by using the maximum likelihood method. The magnitude of completeness is 2.5.

The method I used to identify the possible aftershocks is the nearest-neighbor method introduced in the Chapter 3. In the Chapter 3 and Chen et al. (2021), I have shown that the nearest-neighbor method, although not assume the Omori's law, has the potential to correctly identify long-lived aftershocks that exhibit the Omori-law type decay in intraplate regions. Incompleteness of records is a major problem for using the Omori's law to identify long-lived aftershocks of large historical earthquakes. However, this problem may be overcome by the nearest-neighbor method, because even earthquake records are incomplete, the strength of the link between a historic large earthquake and a current event can be characterized by their nearest-neighbor distance. As long as the separation boundary η is reasonably chosen, possible long-lived aftershocks can be identified. Furthermore, results of the nearest-neighbor method can be used to quantify the percentage of current seismicity as long-lived aftershocks.

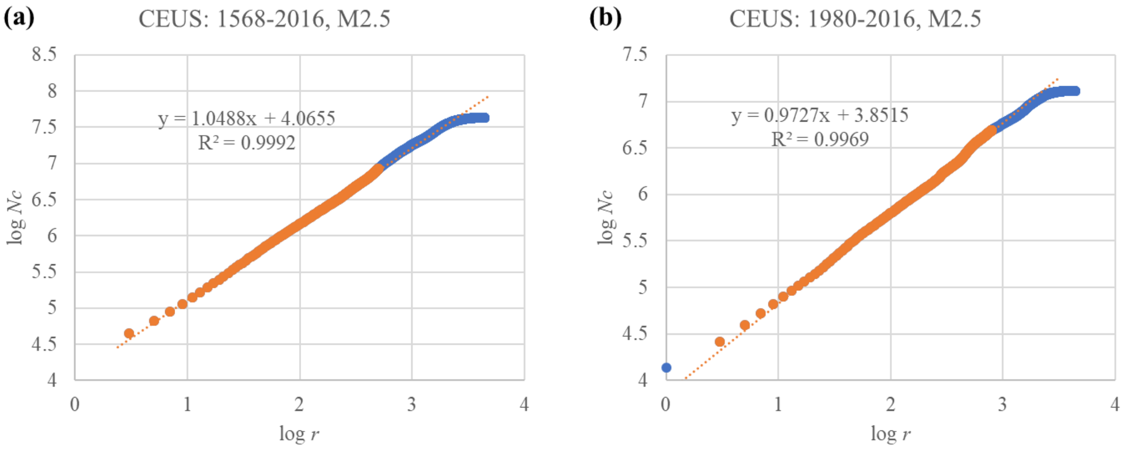


Figure 4.3. The fractal dimension d_f for $M \geq 2.5$ earthquakes **(a)** since 1568 and **(b)** since 1980. The value of d_f is estimated by the correlation integral method (Grassberger & Procacia, 1983). In practice, all possible spatial distances (r) between two epicenters are considered, and the cumulative number (N_c) of event pairs less than r is computed. The slope of the line of log N vs. log r is the fractal dimension d_f.

Because of the lack of large earthquakes ($M \geq 6$) since 1940 in the CEUS (Figure 4.1b), only a small portion of the $M \geq 2.5$ USGS catalog are aftershocks. Therefore, I also use the catalog of aftershocks of the 2011 *M*5.65 Virginia earthquake from Wu et al. (2015) to verify the nearest-neighbor method and show a general distribution of the nearest-neighbor distance between a mainshock and its aftershocks (Figure 4.4). The Wu et al. (2015)'s aftershock catalog contains 1667 events with magnitude range from -0.7 to 3.7 from August 25 to December 31, 2011.

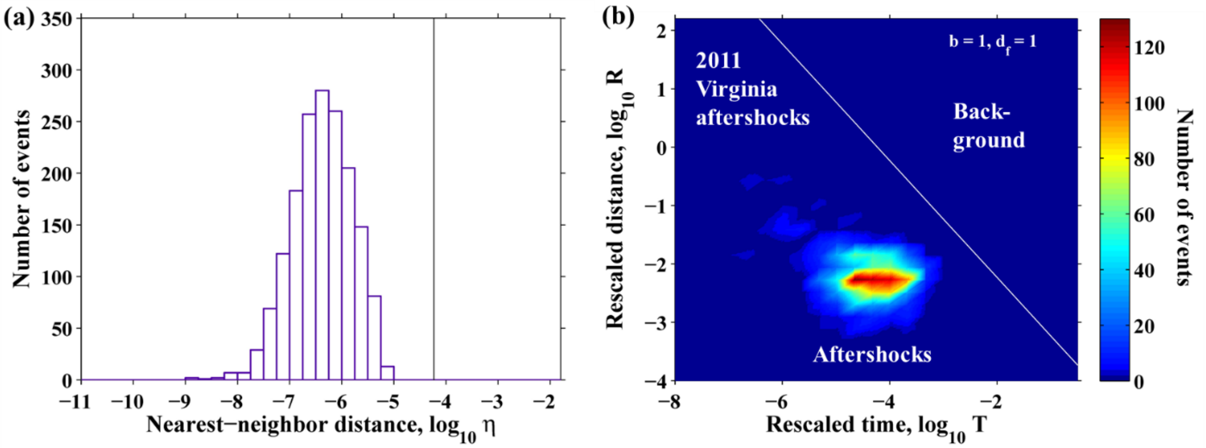


Figure 4.4. **(a)** 1D and **(b)** 2D distributions of the nearest-neighbor distance of the aftershocks ($M \geq$ -0.7) of the 2011 *M*5.65 Virginia earthquake. The solid line ($\log_{10}\eta_0 = -4.23$) is the separation boundary between the aftershocks and background seismicity obtained based on the whole CEUS catalog (See Figure 4.5 and text for details). The aftershock data is from Wu et al. (2015). The results are calculated with $b = 1$ and $d_f = 1$.

4.3 CEUS earthquakes: Aftershocks vs. Background Seismicity

4.3.1 Aftershocks of the 2011 Virginia M5.65 earthquake

The 2011 Virginia M5.65 earthquake was the largest earthquake in the CEUS since 1940. Its aftershocks are well studied by Wu et al. (2015) and Beskardes et al. (2019). Here, I calculated the nearest-neighbor distances between the 2011 Virginia earthquake and its aftershocks within four months by using Wu et al. (2015)'s aftershock catalog. The results show a unimodal distribution of the nearest-neighbor distance (NND) with its peak near $\log_{10}\eta = -6$ (Figure 4.4), similar to the synthetic single aftershock sequence generated by the ETAS model (Zaliapin et al., 2008).

4.3.2 The NND distributions of the CEUS earthquakes

The CEUS catalog, similar to other general earthquake catalogs (e.g., catalogs used in Zaliapin and Ben-Zion (2013) and Chen et al. (2021)), contains both background earthquakes and clustered earthquakes (aftershocks), and shows bimodal distributions of η (Figure 4.5). The aftershocks correspond to the clustered mode on the left of the plot (red curves shown in Figure 4.5) with its peak near $\log_{10}\eta = -6$, similar to the aftershocks of the 2011 Virginia event (Figure 4.4). The background earthquakes correspond to the background mode on the right with its peak near $\log_{10}\eta = -3$ (blue curves shown in Figure 4.5), similar to the synthetic Poisson background earthquakes shown in Zaliapin et al. (2008). Because of the incomplete records of aftershocks of historical large earthquakes, the estimated aftershock modes have low density and may not be reliable. However, enough background events are included in the catalogs, so the estimated background modes are reliable.

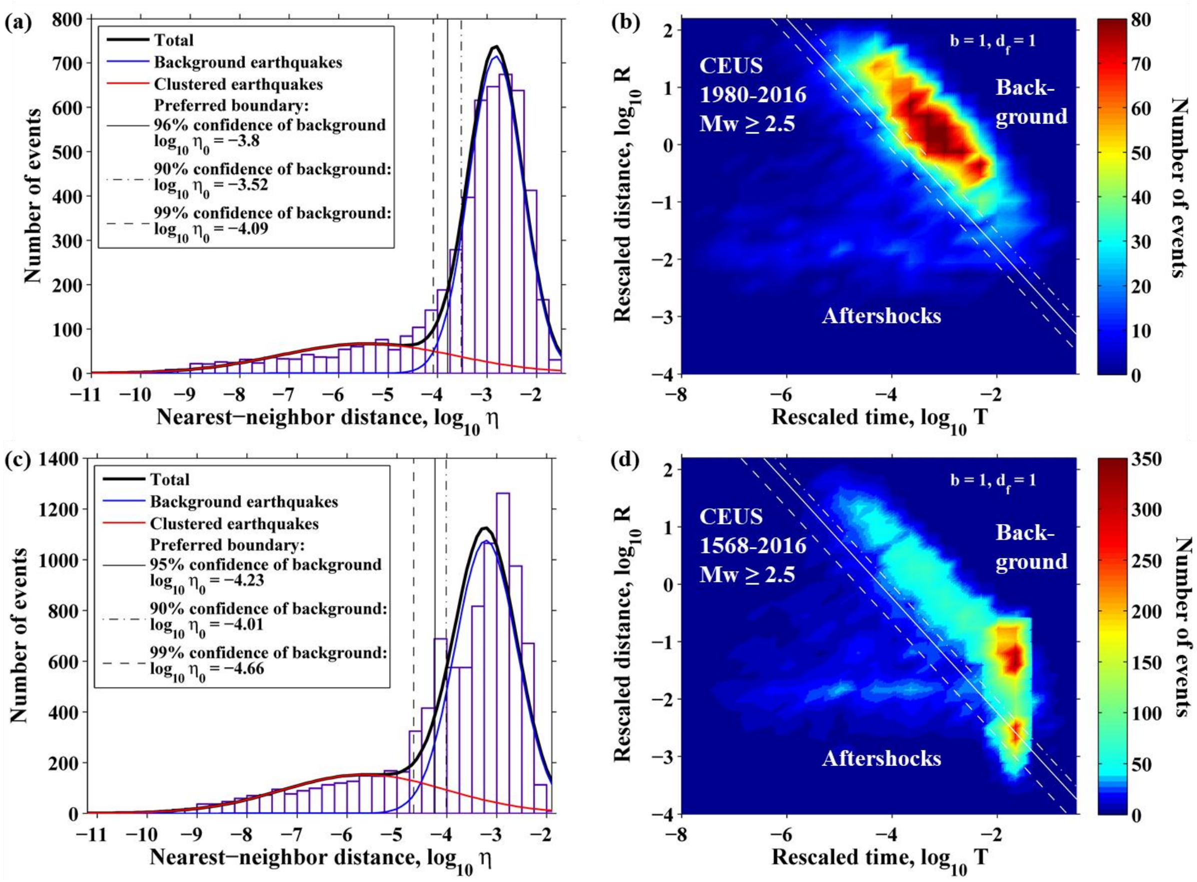


Figure 4.5. The nearest-neighbor method applied to $M \geq 2.5$ earthquakes **(a, b)** since 1980 and **(c, d)** since 1568. **(a, c)** 1D distribution of η, with estimated Gaussian distributions for clustered (red) and background (blue) components. The solid vertical lines are the boundaries that best separate the two components. **(b, d)** 2D joint distribution of rescaled distance and time (R, T). The solid white lines indicate the same boundaries shown in (a, c). The results are calculated with $b = 1$ and $d_f = 1$.

I applied the nearest-neighbor (NN) method to both the complete catalog since 1980 and the whole catalog since 1568 (Figure 4.5). The best separation boundaries are calculated based on a Gaussian mixture model with the background and clustered modes by the Expectation-Maximization algorithm (Zaliapin & Ben-Zion, 2016). The events with η values smaller than the separation boundary value $\log_{10}\eta_0$ are identified as aftershocks. Those with higher values are identified as background earthquakes. Because this chapter focus on possible long-lived aftershocks in the rupture zones of their mainshocks, the identified aftershocks with distance larger than 250 km away from their mainshocks are manually assigned as background earthquakes. For the complete catalog since 1980, the optimal separation boundary is $\log_{10}\eta_0 = -3.8$ (Figure 4.5a). Because no historical large earthquakes are included in this catalog, long-lived aftershocks could be misidentified as background earthquakes even using the nearest-neighbor methods. For the catalog that include events since 1568, the separation boundary at $\log_{10}\eta_0 = -4.23$ has 95% confidence of background identification (Figure 4.5c). This boundary value is smaller than the one for the complete catalog since 1980, meaning a more conservative separation boundary for aftershock identification. I also checked the variation of separation boundary by using $M \geq 3$ and $M \geq 4$ catalog since 1568 and found the value of the separation boundary is quite stable (Figure 4.6), so the incompleteness of records has a minor effect on the separation boundary. This separation boundary ($\log_{10}\eta_0 = -4.23$) can also correctly identify the aftershocks of the 2011 Virginia earthquake (Figure 4.4). After including more events by setting an older starting year in the catalog, two high density patches with large rescaled-time show up in the 2D plot of joint distribution (Figure 4.5d), because many modern events found their potential close relationships to historical large events (examples shown in the New Madrid case below in Figure 4.7). However, only the events with small rescaled-distance and large rescaled-time and below the separation boundary are identified as long-lived aftershocks (Figure 4.5d).

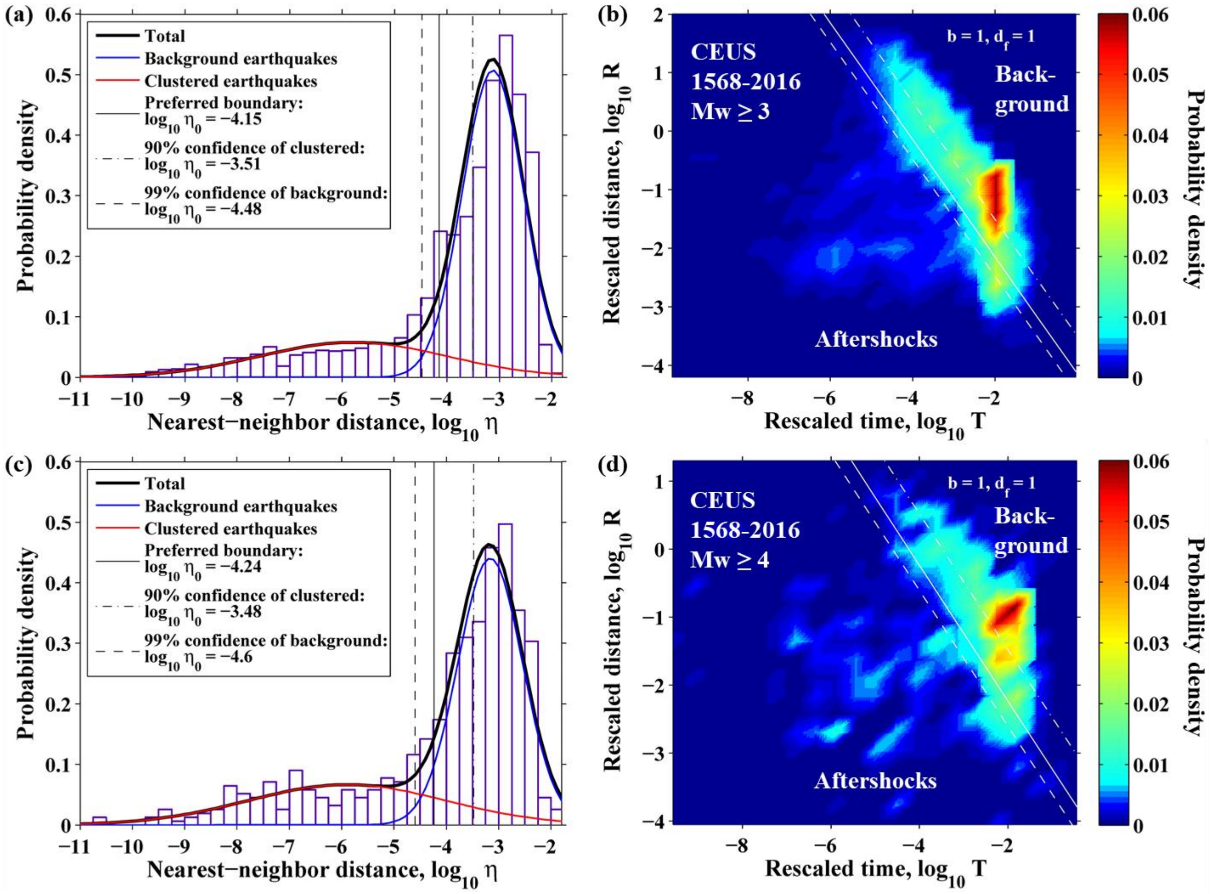


Figure 4.6. The nearest-neighbor method applied to **(a, b)** $M \geq 3$ and **(c, d)** $M \geq 4$ earthquakes since 1568.

4.3.3 Long-lived aftershocks in the New Madrid seismic zone

The New Madrid seismic zone is a currently active seismic zone in the CEUS and the host of several major earthquakes in the past 2300 years (Tuttle et al., 2002; Tuttle et al., 2005). The magnitude and return period of these major earthquakes are still in debate. Magnitudes from 6.8 to 8.4 and return periods from 160 to 10,000 or more years have been proposed (Boyd et al., 2015). Here, I focus on the historical 1811-1812 New Madrid large events and their potential long-lived aftershocks. The magnitudes used are based on the 2018 USGS hazard map catalog (C. S. Mueller, 2019).

The spatial patterns of the nearest-neighbor results and the nearest-neighbor links in the New Madrid seismic zone (NMSZ) are shown in the Figure 4.7. When only events since 1980 are considered, the nearest-neighbor links are disordered and weak for most events (Figure 4.7b), so most events are identified as background earthquakes (Figure 4.7a). After including more past events, many events are found to be correlated with the 1811-1812 mainshocks, so the spatial pattern of the nearest-neighbor links shows several focuses (Figure 4.7d). The events close enough in space or time to the 1811-1812 mainshocks are identified as aftershocks (Figure 4.7c).

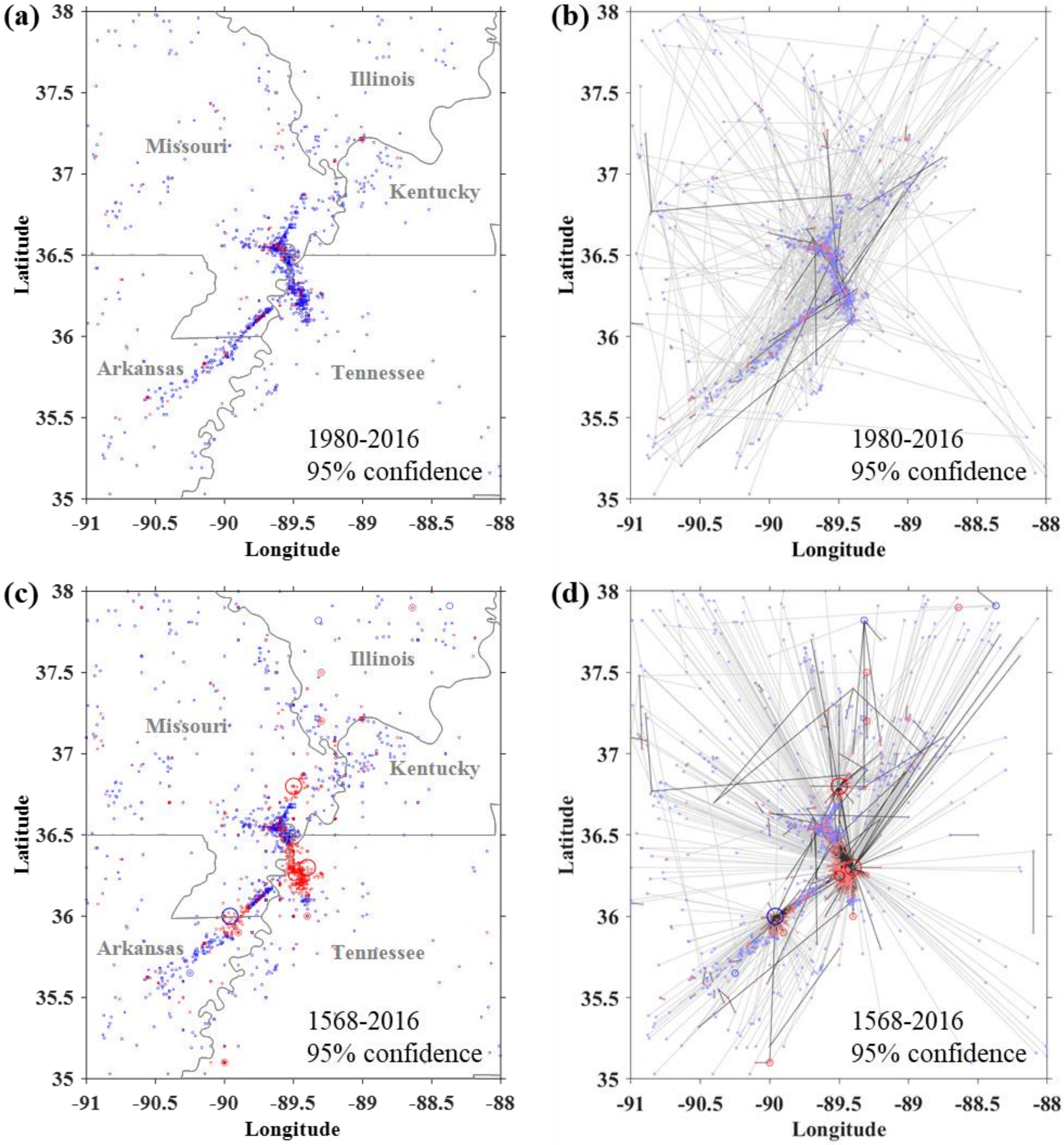


Figure 4.7. Spatial patterns of **(a, c)** the nearest-neighbor results and **(b, d)** the nearest-neighbor links for $M \geq 2.5$ earthquakes in the New Madrid seismic zone. The blue circles indicate the identified background earthquakes. The red circles indicate the identified aftershocks. The black lines indicate strong links between mainshocks and their aftershocks. The grey lines indicate weak lines between the nearest neighbors, so the events are identified as background earthquakes.

The long-lived aftershocks in the NMSZ are only from the four major events: two occurred on December 16, 1811 (NM1 and NM1-A), one occurred on January 23, 1812 (NM2), and the largest one occurred on February 7, 1812 (NM3) (Figure 4.8). The 2018 USGS hazard map catalog is selected and compiled by experts from 20 previous published catalogs (C. S. Mueller, 2019), so I assume the magnitudes and locations offered by this catalog is best based on our current knowledge. Based on this catalog, the nearest-neighbor result show that 23.3% of events in the NMSZ between 1980 and 2016 are aftershocks of the 1811-1812 mainshocks (Table 4.1). All long-lived aftershocks are close to their mainshocks in space (Figure 4.8a). The magnitudes and locations of historical events are estimated based on seismic intensity data and felt area (Bakun & Hopper, 2004; C. S. Mueller, 2019), so they have larger uncertainties than instrumental records. To explore the effect caused by the location uncertainty, I tried to use another epicentetral location (86.4°W, 36.5°N) for the NM3 from Stover and Coffman (1993), which did not consider the NMSZ fault segments suggested by Johnston (1996). The result caused by a different epicenter location shows a different spatial distribution of aftershocks of the NM3 (Figure 4.8b), but aftershocks are still on the central fault segment. For magnitude uncertainties, I calculated two extreme cases by assigning minimum or maximum magnitudes to all four major events. The minimum or maximum magnitudes are based on the 2-sigma uncertainties of magnitude offered by the CEUS Seismic Source Characterization (SSC) project (Coppersmith et al., 2012). The two modified catalogs are called SSC_min and SSC_max separately. Only 10.7% events for SSC_min but 65.0% events for SSC_max is identified as long-lived aftershocks (Figure 4.8c, 4.8d and Table 3.1). Thus, the long-lived aftershocks identified by the nearest-neighbor method are heavily affected by the magnitudes of their mainshocks. For a mainshock, a larger assigned magnitude indicates a longer aftershock activity. The magnitude dependent results can be explained by the Omori's law. At one location, a larger mainshock means a larger aftershock rate immediately after the mainshock and longer decay time to background rate.

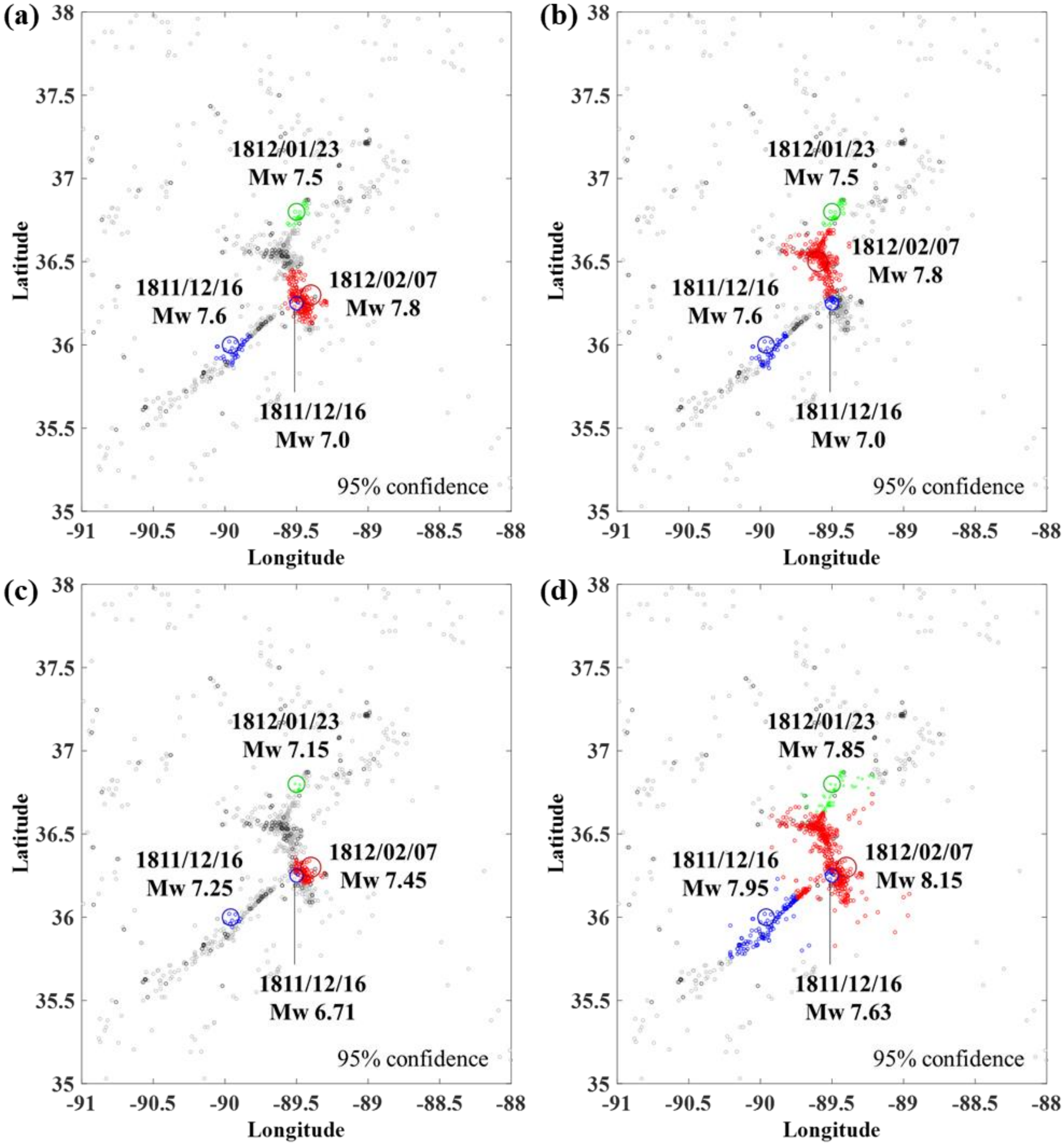


Figure 4.8. The results of possible long-lived aftershocks (1980-2016) of the 1811-1812 mainshocks considering the uncertainties of data. Data: **(a)** the USGS hazard map catalog (C. S. Mueller, 2019), **(b)** same to (a) but the 1812/02/07 event with a more northern epicenter at 86.4°W, 36.5°N (Stover & Coffman, 1993). **(c, d)** same to (a) but the magnitude of the four major events set to minimum and maximum separately based on the 2-sigma uncertainties of magnitude offered by the CEUS Seismic Source Characterization (SSC) project (Coppersmith et al., 2012). Blue circles indicate the aftershocks of the 1811/12/16 events. Green circles indicate the aftershocks of the 1812/01/23 events. Red circles indicate the aftershocks of the 1812/02/07 events. Black circles indicate the aftershocks of other events. Grey circles indicate background earthquakes.

Table 4.1: The results of long-lived aftershocks (1980-2016) of the 1811-1812 mainshocks considering magnitude uncertainties

Catalog	$\log_{10}\eta_0$	Number of aftershocks	Percentage of aftershocks
USGS	-4.23	263	23.3%
SSC_min	-4.15	121	10.7%
SSC_max	-4.14	733	65.0%

4.3.4 Long-lived aftershocks in Charleston and Charlevoix

The Charleston seismic zone is best-known for the 1886 destructive earthquakes occurred near Charleston (Figure 4.9a), South Carolina. Paleoseismic studies indicate that at least two prehistoric earthquakes with comparable magnitude of the 1886 events occurred there in the past 5500 years (Obermeier et al., 1985; Talwani & Cox, 1985).

The Charlevoix seismic zone is an active intraplate seismic zone in St. Lawrence Rift System. Four $M \geq 6$ earthquakes occurred in the past 400 years (Figure 4.9c), and at least three prehistoric large earthquakes occurred in the past 10,000 years (Tuttle & Atkinson, 2010).

The nearest-neighbor results in Charleston and Charlevoix seismic zones show different patterns (Figure 4.9). The 1886 M6.9 Charleston earthquake has not only long-distance aftershocks shortly after its occurrence but also long-lived aftershocks very close to it epicenter (Figure 4.9a, b). Four $M \geq 6$ events occurred in the Charlevoix region, but the earthquakes are more diffusely distributed than in Charleston, so most events are identified as background earthquakes (Figure 4.9c, d).

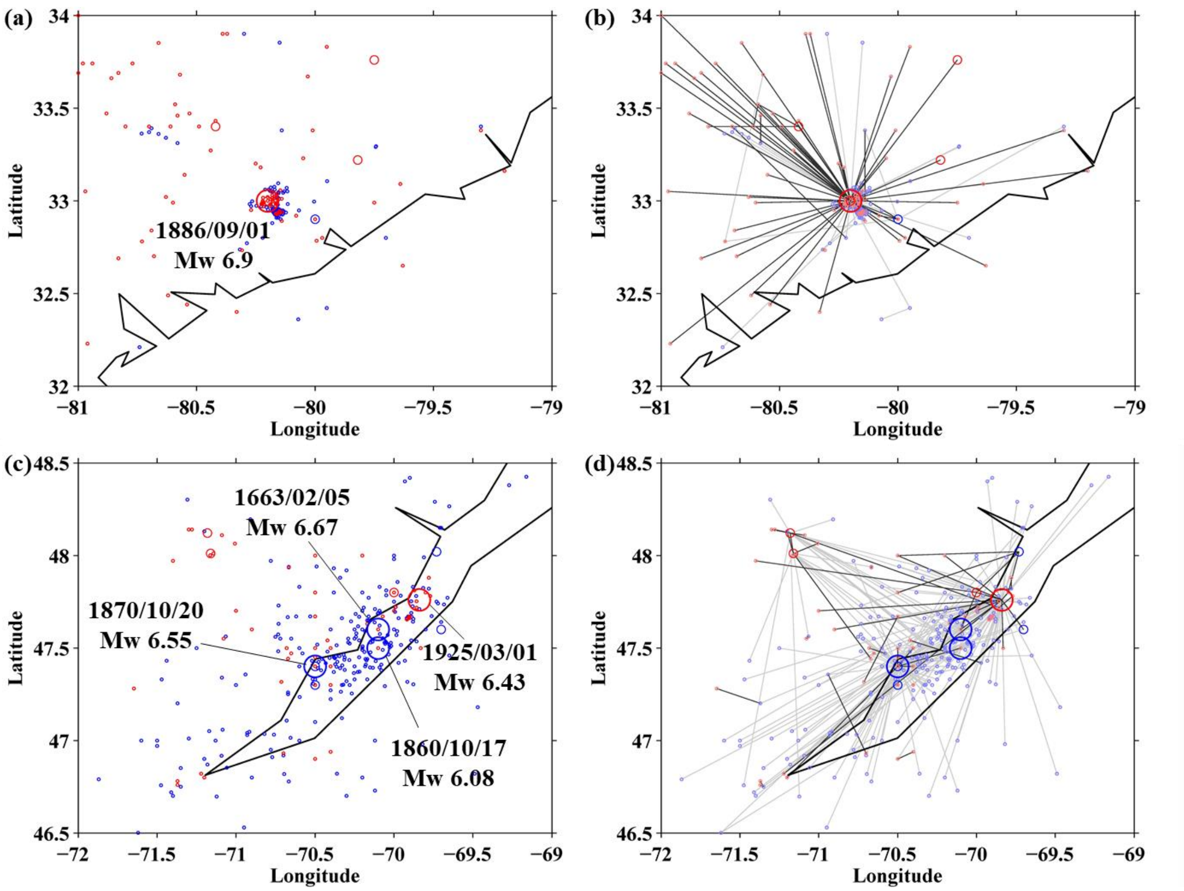


Figure 4.9. Spatial patterns of the nearest-neighbor results and the nearest-neighbor links for $M \geq 2.5$ earthquakes since 1568 in **(a, b)** Charleston and **(c, d)** Charlevoix. Same notations in Figure 4.7.

4.3.5 Inconsistence between current seismicity rate and average strain rate

In previous sections, I showed that long-lived aftershocks are significant in New Madrid and Charleston but lacking in Charlevoix. These results may be explained by the inconsistency of current seismicity rate and average strain rate in these regions (Table 4.2). Background seismicity rate is believed to reflect the average strain rate for a given region, so a high strain rate region should have higher background seismicity rate than a low strain rate region. However, Charlevoix, the region with the highest strain rate of the three regions I studied, has the least seismicity rate (Table 4.2). One possible explanation is that current seismicity rates in New Madrid and Charleston are higher because the catalog contains not only background earthquakes but also long-lived aftershocks.

Table 4.2: Comparison between average strain rate and seismicity rate of $M \geq 2.5$ events (1980-2016) for the three intraplate regions in CEUS

Region	Number of events	Seismicity rate (event yr^{-1} km^{-2})	Average strain rate (10^{-9} yr^{-1})
New Madrid	1128	2.74×10^{-4}	0.52
Charleston	116	0.63×10^{-4}	0.39
Charlevoix	141	0.51×10^{-4}	1.62

* The strain rate data come from Kreemer et al. (2018).

4.4 Discussion

4.4.1 Aftershocks: Controversy and Uncertainty

One motivation of the research presented in this chapter is to answer the question regarding the modern active seismicity in CEUS: Are they aftershocks of previous large historical earthquakes? Our results, based on the nearest-neighbor method, indicate that the modern active seismicity in New Madrid and Charleston likely contain significant amount of the long-lived aftershocks of the 1811-1812 New Madrid earthquakes and the 1886 Charleston earthquake, respectively; but most current seismicity in Charlevoix are likely background earthquakes. The long-lived aftershock activity in New Madrid is consistent with the rate-and-state model of fault friction (Dieterich, 1994; Stein & Liu, 2009), Coulomb failure stress models (K. Mueller et al., 2004), and a postseismic afterslip model (Boyd et al., 2015). Although Page and Hough (2014) proposed an opposite argument because the ETAS model failed to reconstruct long-lived aftershocks that are consistent with observations, Boyd et al. (2015) pointed out that the parameter values used by Page and Hough (2014) in the ETAS model may not be appropriate. The long-lived aftershocks of the 1886 Charleston earthquake is supported by Chapman et al. (2016) finding on common characteristics shared by the modern seismicity in Charleston and the aftershocks of the 2011 Virginia earthquake; earthquakes are concentrated in a shallow and semicircular zone with a substantial diversity of earthquake focal mechanism nodal plan orientation. The lack of long-lived aftershocks in Charlevoix is supported by Fereidoni and Atkinson (2014)'s statistical analyses based on the Omori's law for the 1663 event. However, Fereidoni and Atkinson (2014)'s data and interpretation are contested by Ebel (2016) and defended by Fereidoni and Atkinson (2016). Current studies on the long-lived aftershocks in Charlevoix, including this study, are all based on statistical analysis, so more physics-based studies are needed. Nevertheless, the spatial distributions of seismicity in New Madrid and Charleston are different from that in Charlevoix (Figure 4.7 and 4.9). Modern earthquakes are clustered on the faults that historical large earthquakes ruptured in New Madrid and Charleston, but are more diffusedly distributed in Charlevoix.

The idea of classifying earthquakes into background earthquakes and aftershocks implies two extreme groups: one only depends on the long-term tectonic loading, and one is triggered and strongly dependent on its mainshock. However, both the long-term tectonic loading and increased stress caused by the mainshock play significant roles for the so-called long-lived aftershocks in the transition period. Therefore, from a physical standpoint, to call a specific event as a long-lived aftershock or background event is meaningless. On the other hand, the percentage of aftershocks is meaningful; it quantifies the influence of a large event for the seismicity of a given region.

The nearest-neighbor method I used is a statistical method based on the long-term Gutenberg-Richter (GR) frequency-magnitude distribution and the fractal distribution of earthquake epicenters (Baiesi & Paczuski, 2004; Zaliapin et al., 2008). Each earthquake is assumed as a point source at its epicenter and its influence depends on its magnitude, location and time, but the point-source assumption may not be appropriate for large historical earthquakes. Therefore, it has no bearing on the causative physics between mainshocks and their aftershocks. Some stress-triggered aftershocks may be missed by the nearest-neighbor method if their nearest-neighbor distances are beyond the threshold for aftershock identification. For the results in this chapter, the number of identified long-lived aftershocks in New Madrid (Figure 4.8a) is more conservative than those suggested by the Coulomb failure stress field in K. Mueller et al. (2004).

4.4.2 Implications for intraplate seismic hazard

The long-lived aftershocks highlight the importance to include historical earthquakes and paleoseismicity in the seismic hazard assessment. The uncertainties of location and magnitude of historic large earthquakes affect the identification of long-lived aftershocks (Figure 4.8 and Ebel (2016)), which need to be seriously considered. Careful historical and paleoseismic studies are needed to reduce data uncertainties. The most recent seismic hazard map for intraplate regions contains two types of source model: one is seismicity-based source model, another is fault-based source model (Allen et al., 2018; Petersen et al., 2014; Petersen et al., 2020). The seismicity-based source model use background seismicity as input and usually yield spatially varying earthquake occurrence rate by smoothing the observed background rate with a given smoothing kernel, assuming past seismicity is a good spatial predictor for future seismic hazard (Frankel, 1995). In this case, if long-lived aftershocks are misinterpreted as background earthquake, the seismic hazard would be overestimated, as pointed out by Stein and Liu (2009) and Toda and Stein (2018). The fault-based model is mainly based on historical and paleoseismic records of large earthquakes on specific faults. Background seismicity usually has little or no weight in this model. For practical purpose, the time-independent Poisson model, rather than the time-dependent quasi-periodic model, is generally used (Petersen et al., 2014). Considering that large intraplate earthquakes usually occur in clusters and separated by irregular but relatively longer quiescence periods (Chen et al., 2020; Clark et al., 2012), the weight for in-cluster sequence and out-of-cluster sequence are subjectively assigned by experts in logical trees in hazard assessments (Petersen et al., 2020). The long-lived aftershocks, if not considered in the hazard assessment, would cause an overestimated weight for in-cluster sequence, so an overestimated hazard. Therefore, the seismic hazards in New Madrid and Charleston may be overestimated. Moreover, the percentage of aftershocks and its uncertainty obtained in this study offers an objective way to quantify the possible long-lived aftershocks and can be used in the seismic hazard assessment.

4.5 Conclusions

In this chapter, I used the nearest-neighbor method to identify the possible long-lived aftershocks in intraplate CEUS, specifically in the New Madrid, Charleston, and Charlevoix seismic zones. I validated the reliability of the nearest-neighbor method by using the aftershocks of the 2011 Virginia earthquake. The results of the CEUS earthquakes indicate that aftershock activities are likely to be continuing in New Madrid and Charleston but unlikely in Charlevoix seismic zones. I also explored how the uncertainties of locations and magnitudes of historic large events affect the results. Changing the location of a mainshock epicenter can change the spatial distribution of aftershocks. Changing the magnitude of a mainshock, on the other hand, can significantly affects the aftershock identification when using the nearest-neighbor method. For a given mainshock, assigning a larger magnitude of a mainshock can significantly increase the number and duration of aftershocks. In New Madrid, around 23.3% of events between 1980 and 2016 are probably aftershocks, but with large uncertainty ranging from 10.7% to 65.0%. Even for the lower estimate, the result indicates that the aftershock sequence of the 1811-1812 mainshocks has continued to the present-day. Ignoring these long-lived aftershocks may cause overestimation of seismic hazard in New Madrid and Charleston seismic zones.

References

Allen, T. I., Gibson, G., Ghasemi, H., & Leonard, M. (2018). *The 2018 National Seismic Hazard Assessment for Australia: earthquake epicentre catalogue*: Geoscience Australia.

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E, 69*(6), 066106.

Bakun, W., & Hopper, M. G. (2004). Magnitudes and locations of the 1811–1812 New Madrid, Missouri, and the 1886 Charleston, South Carolina, earthquakes. *Bulletin of the Seismological Society of America, 94*(1), 64-75.

Basham, P., Adams, J., Hays, W., & Gori, P. (1983). Earthquakes on the continental margin of eastern Canada: Need future large events be confined to the locations of large historical events. *US Geol. Surv. Open-File Rept. 83-843*, 456-467.

Beskardes, G. D., Wu, Q., Hole, J. A., Chapman, M. C., Davenport, K. K., Brown, L. D., & Quiros, D. A. (2019). Aftershock sequence of the 2011 Virginia earthquake derived from the dense AIDA array and backprojection. *Bulletin of the Seismological Society of America, 109*(1), 19-33.

Boyd, O. S., Smalley, R., & Zeng, Y. (2015). Crustal deformation in the New Madrid seismic zone and the role of postseismic processes. *Journal of Geophysical Research: Solid Earth, 120*(8), 5782-5803.

Chapman, M., Beale, J. N., Hardy, A. C., & Wu, Q. (2016). Modern seismicity and the fault responsible for the 1886 Charleston, South Carolina, earthquake. *Bulletin of the Seismological Society of America, 106*(2), 364-372.

Chen, Y., Liu, M., & Luo, G. (2020). Complex Temporal Patterns of Large Earthquakes: Devil’s Staircases. *Bulletin of the Seismological Society of America, 110*(3), 1064-1076.

Chen, Y., Liu, M., & Wang, H. (2021). Aftershocks and background seismicity in Tangshan and the rest of North China. *Journal of Geophysical Research: Solid Earth*, e2020JB021395.

Clark, D., McPherson, A., & Van Dissen, R. (2012). Long-term behaviour of Australian stable continental region (SCR) faults. *Tectonophysics, 566*, 1-30.

Coppersmith, K. J., Salomone, L. A., Fuller, C. W., Glaser, L. L., Hanson, K. L., Hartleb, R. D., et al. (2012). *Central and eastern United States (CEUS) seismic source characterization (SSC) for nuclear facilities*. Retrieved from

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research: Solid Earth, 99*(B2), 2601-2618.

Ebel, J. E. (2016). Comment on “Aftershock Statistics for Earthquakes in the St. Lawrence Valley” by Azadeh Fereidoni and Gail M. Atkinson. *Seismological Research Letters, 87*(1), 149-151. https://doi.org/10.1785/0220150120

Ebel, J. E., Bonjer, K.-P., & Oncescu, M. C. (2000). Paleoseismicity: Seismicity evidence for past large earthquakes. *Seismological Research Letters, 71*(2), 283-294.

Fereidoni, A., & Atkinson, G. M. (2014). Aftershock statistics for earthquakes in the St. Lawrence Valley. *Seismological Research Letters, 85*(5), 1125-1136.

Fereidoni, A., & Atkinson, G. M. (2016). Reply to “Comment on ‘Aftershock Statistics for Earthquakes in the St. Lawrence Valley’ by Azadeh Fereidoni and Gail M. Atkinson” by John Ebel. *Seismological Research Letters, 87*(1), 152-156. https://doi.org/10.1785/0220150170

Frankel, A. (1995). Mapping seismic hazard in the central and eastern United States. *Seismological Research Letters, 66*(4), 8-21.

Gardner, J., & Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America, 64*(5), 1363-1367.

Grassberger, P., & Procacia, I. (1983). Measuring the strangeness of strange attractors. *Physica. D, 9*(1-2), 189-208.

Helmstetter, A., Kagan, Y. Y., & Jackson, D. D. (2006). Comparison of short-term and time-independent earthquake forecast models for southern California. *Bulletin of the Seismological Society of America, 96*(1), 90-106.

Johnston, A. C. (1996). Seismic moment assessment of earthquakes in stable continental regions—I. Instrumental seismicity. *Geophysical Journal International, 124*(2), 381-414.

Kreemer, C., Hammond, W. C., & Blewitt, G. (2018). A robust estimation of the 3-D intraplate deformation of the North American plate from GPS. *Journal of Geophysical Research: Solid Earth, 123*(5), 4388-4412.

Levandowski, W., & Powell, C. A. (2019). Evidence for strain accrual in the Eastern Tennessee Seismic Zone from earthquake statistics. *Seismological Research Letters, 90*(1), 446-451.

Mueller, C. S. (2019). Earthquake catalogs for the USGS national seismic hazard maps. *Seismological Research Letters, 90*(1), 251-261.

Mueller, K., Hough, S. E., & Bilham, R. (2004). Analysing the 1811–1812 New Madrid earthquakes with recent instrumentally recorded aftershocks. *Nature, 429*(6989), 284-288.

Obermeier, S. F., Gohn, G. S., Weems, R. E., Gelinas, R. L., & Rubin, M. (1985). Geologic evidence for recurrent moderate to large earthquakes near Charleston, South Carolina. *Science, 227*(4685), 408-411.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American statistical Association, 83*(401), 9-27.

Page, M. T., & Hough, S. E. (2014). The New Madrid seismic zone: Not dead yet. *Science, 343*(6172), 762-764.

Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D., et al. (2014). *Documentation for the 2014 update of the United States national seismic hazard maps*. Retrieved from

Petersen, M. D., Shumway, A. M., Powers, P. M., Mueller, C. S., Moschetti, M. P., Frankel, A. D., et al. (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. *Earthquake Spectra, 36*(1), 5-41.

Stein, S., & Liu, M. (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature, 462*(7269), 87-89.

Stover, C. W., & Coffman, J. L. (1993). *Seismicity of the United States, 1568-1989 (revised)*: US Government Printing Office.

Talwani, P., & Cox, J. (1985). Paleoseismic evidence for recurrence of earthquakes near Charleston, South Carolina. *Science, 229*(4711), 379-381.

Toda, S., & Stein, R. S. (2018). Why Aftershock Duration Matters for Probabilistic Seismic Hazard Assessment. *Bulletin of the Seismological Society of America, 108*(3A), 1414-1426.

Tuttle, M. P., & Atkinson, G. M. (2010). Localization of large earthquakes in the Charlevoix seismic zone, Quebec, Canada, during the past 10,000 years. *Seismological Research Letters, 81*(1), 140-147.

Tuttle, M. P., Schweig, E. S., Sims, J. D., Lafferty, R. H., Wolf, L. W., & Haynes, M. L. (2002). The earthquake potential of the New Madrid seismic zone. *Bulletin of the Seismological Society of America, 92*(6), 2080-2089.

Tuttle, M. P., Schweig III, E. S., Campbell, J., Thomas, P. M., Sims, J. D., & Lafferty III, R. H. (2005). Evidence for New Madrid earthquakes in AD 300 and 2350 BC. *Seismological Research Letters, 76*(4), 489-501.

Wu, Q., Chapman, M., & Beale, J. (2015). The aftershock sequence of the 2011 Mineral, Virginia, earthquake: Temporal and spatial distribution, focal mechanisms, regional stress, and the role of Coulomb stress transfer. *Bulletin of the Seismological Society of America, 105*(5), 2521-2537.

Zaliapin, I., & Ben-Zion, Y. (2016). A global classification and characterization of earthquake clusters. *Geophysical Journal International, 207*(1), 608-634.

Zaliapin, I., & Ben-Zion, Y. (2013). Earthquake clusters in southern California I: Identification and stability. *Journal of Geophysical Research: Solid Earth, 118*(6), 2847-2864.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters, 101*(1), 018501.

Zhuang, J., Murru, M., Falcone, G., & Guo, Y. (2019). An extensive study of clustering features of seismicity in Italy from 2005 to 2016. *Geophysical Journal International, 216*(1), 302-318.

Chapter 5: Correlations between strain rate and seismicity in different tectonic settings: Complex spectra

5.1 Introduction

Geodetic strain rate characterizes present-day crustal deformation and therefore may be used as a spatial predictor for earthquakes. Good correlations between strain rate and seismicity are found in plate boundary zones. On a global scale, Kreemer et al. (2002) found that seismicity rates of shallow earthquakes are correlated with the strain rates in subduction zones and active continents. In California and Nevada, large earthquakes are found concentrated in the San Andreas Fault system, the Eastern California Shear Zone, and the Walker Lane shear zone, where the strain rate is high (Shen et al., 2007; Zeng et al., 2018). In the Tibetan Plateau, background seismicity rate is correlated with geodetic strain rate: higher strain rate regions have higher background seismicity rates (Stevens & Avouac, 2021). Because of the good correlation, strain rate is used in some probabilistic seismic hazard assessments (Shen et al., 2007; Stevens & Avouac, 2021).

However, in intraplate regions, the correlation between strain rate and seismicity is uncertain. In North China, high strain rates are found along the major active tectonic zones, but some low strain rate regions in North China Plain have significant modern seismicity and large historical earthquakes (Figure 3.1 and 3.7) (Chen et al., 2021; Liu & Wang, 2012). In Fennoscandia, the strain rate field is relatively simple; with NW–SE compression on the Norwegian continental margin and NW–SE extension in the central part of Fennoscandia (Keiding et al., 2015). However, the seismicity in Fennoscandia is more complex and diverse in space (Bungum et al., 2010; Keiding et al., 2015). Especially in the central part of Fennoscandia, focal mechanisms show variations between tension, compression, and strike-slip, and directions also significantly vary between different earthquakes (Keiding et al., 2015). In addition, the geodetic moment rate derived from GPS strain rate is two orders of magnitude larger than the seismic moment rate derived from earthquake catalogs in Fennoscandia (Keiding et al., 2015). In intraplate North America, the plate-scale spatial correlation between seismicity and strain rate is absent: no strain accumulation at the major seismic zones such as the Charleston, South Carolina area, the Eastern Tennessee Seismic Zone, and the New Madrid Seismic Zone (NMSZ) (Calais et al., 2016; Kreemer et al., 2018). In the Saint Lawrence Valley, eastern Canada, Tarayoun et al. (2018) found that high strain rate is concentrated in ancient rift zones where modern seismicity and large historical earthquakes are clustered, but no systematic correlation is found between seismicity and geodetic strain rate patterns in the whole region. Therefore, it is important to explore how well strain rate spatially correlates with seismicity in intraplate regions, therefore whether or not strain rate can be used for seismic hazard assessment in intraplate regions, where seismic records are in general too short to show the complete spatial patterns of seismicity.

The correlation between strain rate and seismicity could also be time-dependent. In California and Nevada, the $M \geq 4$ background earthquakes were gradually changing from diffused distribution in the whole region (1933-1980s) to concentrated distribution in high strain rate areas (1980s-2016), along with an increased rate of $M \geq 6.5$ events (Zeng et al., 2018). It is worthy to check whether such a temporal variation is a general pattern or not, because it may be used to indicate whether a region is in a relatively quiescent period or an active period with clustered of large earthquakes.

In this chapter, I did a systematic analysis about the correlation between strain rate and seismicity. First, I analyzed and compared the spatial correlations between strain rate, seismicity, and seismic moment in different tectonic settings using the approach of Shen et al. (2007) and Zeng et al. (2018). Second, I explored how correlations between strain rate and seismicity in different tectonic settings vary with time. In these two steps, I also explore the effects of seismic catalog completeness, cut-off magnitude, decluttering, and model parameters. At last, I discuss the implication of these results for seismic hazard assessment.

5.2 Data and Method

5.2.1 Earthquake Catalogs

The earthquake catalogs used in this chapter are from four sources: the historical and instrumental earthquake catalog for North China (-780-2015) (Cheng et al., 2017), the earthquake catalogs (1568-2016) used for the 2018 USGS National Seismic Hazard Map (https://www.sciencebase.gov/catalog/item/59e62f3ce4b05fe04cd1cc48) (C. S. Mueller, 2019), the GEM Global Historical Earthquake Catalog (1000-1903) (https://emidius.eu/GEH/) (Albini et al., 2013; Albini et al., 2014), and the ISC-GEM Global Instrumental Earthquake Catalog (1904-2015) (http://www.isc.ac.uk/iscgem/) (Giacomo et al., 2018; Storchak et al., 2013; Storchak et al., 2015). All catalogs use the moment magnitude. The first two catalogs are introduced in Chapters 3 and 4. The Global Historical Earthquake Catalogue (GHEC) is a world catalogue of historical large earthquakes for the period 1000-1903, with magnitude of M_w 7 and above. Because the catalog is based on historical records, the completeness of the catalog significantly varies with regions. The catalog is roughly complete since 1580 for Japan, since 1000 for Anatolia, and since 1780 for Tibetan Plateau (Albini et al., 2013). The ISC-GEM Global Instrumental Earthquake Catalogue (1904-2015) is a world catalogue of instrumental large earthquakes with magnitude of M_w 5.5 and above plus continental events down to M_w 5.0. On a global scale, the catalog is complete for M_w 7.0 after 1918 and complete for M_w 6.0 after 1964 (Michael, 2014).

5.2.2 GPS data and strain rate models

Here, I use the strain rate results from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014) for plate boundary zones and Kreemer et al. (2018) for CEUS.

The Global Strain Rate Model (GSRM v.2.1) is a global model of strain rates in the plate boundary zones constrained by horizontal geodetic site-velocities (Kreemer et al., 2014). The velocities are from both continuous GPS measurements and published studies prior to Kreemer et al. (2014), and are carefully selected to avoid effects of creep and transient phenomena from postseismic deformation and Glacial Isostatic Adjustment (GIA), so the data set consists of velocities best representing the “secular” or interseismic velocities. Based on the velocities, the strain rate field is modeled using the Haines and Holt method (Beavan & Haines, 2001; Haines & Holt, 1993; Holt et al., 2000). In Kreemer et al. (2014)’s model, the plate boundary zones are allowed to deform in grid cells with size 0.1° longitude by 0.1° latitude, while the rest of Earth surface is modeled as rigid spherical caps.

The strain rate model I used for intraplate North America is from Kreemer et al. (2018). Kreemer et al. (2018) applied a new strain rate estimation method, called median estimation of local deformation, which is developed for intraplate regions and is stable against outlier data. In this method, the strain rate for a given point is the median value from a set of N local estimates of strain rate. Each local estimate of strain rate is determined from velocity observations of three distinct and noncollinear locations. The formulation used to link the velocity and strain rate is provided by Ward (1998). The grid size of strain rate results is 0.5° by 0.5° with spatial resolution of ~100 km in the continental CEUS. About the data, Kreemer et al. (2018) use not only the data from continuous GPS networks and data archives but also the data from commercial and state networks and networks that were installed to study the ionosphere, the troposphere, and surface subsidence to increase the number of data needed by the method. Kreemer et al. (2018)'s results show that the main cause of deformation in intraplate North America is glacial statistical adjustment.

5.2.3 Spatial distribution of seismic moment

Besides the spatial distribution of earthquake epicenters, the spatial distribution of seismic moment release may also correlate with the strain rate field. Here, the seismic moment released by each earthquake is converted from its moment magnitude based on the formula of Hanks and Kanamori (1979): $\log_{10} M_0 = 1.5\ M_w + 16.1$, where M_0 is the seismic moment with units of dyne-centimeter. Spatially, the seismic moment released by each earthquake is assumed to be evenly distributed in a circular region centered at the earthquake epicenter with the diameter equal to the empirical rupture length. The empirical rupture length is calculated using the formula from Blaser et al. (2010).

5.2.4 Success diagram, spatial prediction power, and strain rate-seismicity correlation

To quantify the correlation between strain rate and seismicity and test the prediction power of strain rate on earthquake locations, I follow the approach of Shen et al. (2007) and Zeng et al. (2018).

I first grid the region according to the resolution of strain rate data and then sort the grid cells by descending strain rate. Strain rate, number of earthquakes, and seismic moment in these grid cells are then separately summed over the sorted cells to produce their cumulative curves. For comparing strain rate and seismicity, the cumulative strain rate, cumulative earthquake number, and cumulative seismic moment are normalized to unity. The total cumulative number of the sorted cells is also normalized to unity to express the fraction of covered area. The normalized cumulative strain rate, number of earthquakes, and seismic moment are then plotted as a function of the fraction of covered area (Figure 5.1). The fraction of covered area is sorted by descending strain rates, with the highest strain rate areas located to the left of horizontal axis.

For simplicity, I call the cumulative strain rate curve as the strain rate curve, the cumulative earthquake number curve as the seismicity curve, and the cumulative seismic moment curve as the moment curve. Overall, I call this kind of plot “the success diagram" here, because it describes how many earthquakes or seismic moments are "successfully" concentrated in high strain rate regions. This kind of plot is actually a flipped version of the Molchan error diagram, which is an objective way to test earthquake prediction (Molchan & Kagan, 1992; Zechar & Jordan, 2008). Furthermore, to quantify the spatial prediction power of strain rate, I use a parameter called the area skill score. The area skill score, introduced by Zechar and Jordan (2008), is the normalized area below the corresponding cumulative curve of earthquake number or seismic moment. The area skill score = 1 corresponds to a perfect prediction: all earthquakes or seismic moments concentrate on one point. The area skill score = 0 indicates completely fail of prediction: no event or seismic moment releases in the given area. The area skill score = 0.5 is expected for a random guess: earthquakes or seismic moments with no preference on high or low strain rate areas. The random guess appears as the diagonal line in the success diagrams (Figure 5.1).

The correlation between strain rate, earthquake, and seismic moment is indicated in the success diagram. The assumption is that the seismicity (or moment) curve should match with the strain rate curve if the seismicity (or moment) is highly correlated with the strain rate. If the seismicity (or moment) curve is significantly above or below the strain rate curve, the correlation between strain rate and seismicity (or moment) is poor. If the seismicity (or moment) is randomly distributed in space while the strain rate is not, the correlation between strain rate and seismicity (or moment) is absent. I also quantify the degree of such a correlation by calculating the area between the seismicity (or moment) curve and strain rate curve. The larger such an area is, the poorer the correlation is.

5.3 Results

5.3.1 Long-term strain rate-seismicity correlation in different tectonic settings

Here, I studied the spatial correlation between strain rate, earthquake, and seismic moment in California-Nevada, Japan, Anatolian, Tibetan Plateau, North China, and Central and Eastern United State (CEUS), representing the spectrum of tectonic settings ranging from plate boundary zones to stable continental interior. The first four regions are plate boundary zones or regions with averagely high strain rate. The last two regions are intraplate regions with low strain rate. Specifically, California-Nevada region is a transform plate boundary region between the Pacific Plate and North America plate. Japan is a typical subduction zone between Eurasian Plate, Pacific Plate, and Philippine Sea Plate. Anatolia and Tibetan Plateau are continental collision zones between Arabian Plate and Eurasian Plate, and between Indian Plate and Eurasian Plate, respectively. North China is an intraplate region and a reactivated Archaean craton, loaded by far-field plate tectonic process, but it is sometimes also referred to as a diffuse plate boundary region because of its non-rigid deformation (Gordon, 1998; Kreemer et al., 2014). North China is less active and has lower average strain rate than plate boundary regions. CEUS is a stable continental region that is largely unaffected by currently active plate boundary process and mainly affected by glacial isostatic adjustment (Kreemer et al., 2018). In the six studied regions, CEUS is least active with the lowest average strain rate.

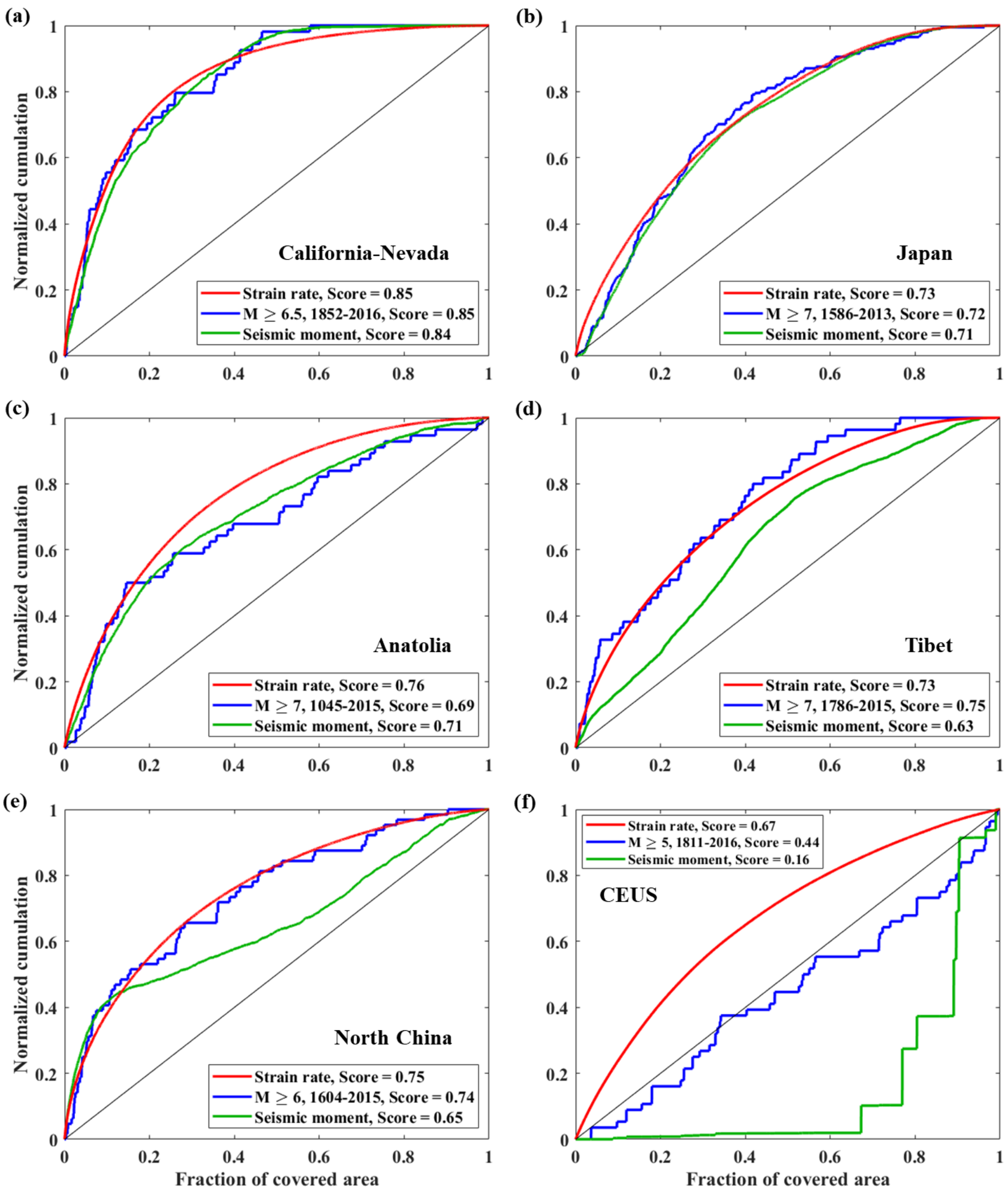


Figure 5.1. Comparison of correlations between strain rate, seismicity, and seismic moment in different tectonic settings. Cumulative strain rate, earthquake count, and seismic moment are plotted against the fraction of covered area sorted by descending strain rates, with the highest strain rate areas located to the left of horizontal axis. The total cumulative strain rate (red), earthquake count (blue), and seismic moment (green)

are normalized to unity. The scores in the legend are the area skill scores. Higher value means better spatial prediction power. The diagonal line indicates random distribution in space (area skill score = 0.5).

The results of success diagrams are shown in Figure 5.1. The complete earthquake catalogs in each region are used. In California- Nevada and Japan, good spatial correlations between strain rate, seismicity, and seismic moment are found (Figure 5.1 a, b). The spatial correlations can be clearly seen in the maps (Figure 5.2, 5.3). In the Anatolia, the seismicity and moment curves match with each other, but they are slightly below the strain rate curve (Figure 5.1 c). The deviation may be caused by the lack of $M \geq 7$ earthquake in the regions of medium strain rates in central and western Anatolian peninsula (Figure 5.3). However, these regions have lower strain rate in the updated high-resolution results (Weiss et al., 2020), which would improve the correlation between strain rate and seismicity in the Anatolia. In Tibetan Plateau, strain rate has a good spatial correlation with large earthquakes ($M \geq 7$), but is poorly correlated with seismic moment (Figure 5.1d, 5.4). The cause of the deviation between strain rate and seismic moment may be that the seismic records in Tibetan Plateau is too short to include several earthquake cycles.

In intraplate North China, the plot is similar to the one in Tibetan Plateau, with good correlation between strain rate and seismicity but poor for seismic moment (Figure 5.1e). North China has lower average strain rate than plate boundary regions, so the recurrence interval in North China is even longer, and the seismic moment curve can be changed by a few large earthquakes. For example, the 1668 $M8.4$ Tancheng earthquake

(Figure 5.6b) is one of the largest earthquakes in North China; it occurred in an area of low strain rate (Figure 5.6). In the CEUS, I used a longer seismic catalog than that in Kreemer et al. (2018) and get similar results: the correlation between strain rate, seismicity, and seismic moment is poor or absent (Figure 5.1f, 5.7). The seismicity curve is close to the diagonal line with area skill score = 0.44, so the $M \geq 5$ seismicity in the CEUS is close to be randomly distributed in space and not correlated with strain rate. I also checked the $M \geq 5$ background seismicity and small modern seismicity ($M \geq 2.5$); the results are similar (Figure 5.8). The highest strain rate regions near the Great lakes are lack of $M \geq 5$ earthquakes (Figure 5.7). Most seismic moments are released in New Madrid and Charlevoix seismic zones where strain rates are relatively low (Figure 5.1f, 5.7). Many factors may cause the significant different between strain rate, seismicity, and seismic moment in the CEUS and will be discussed in the Discussion section. Overall, good correlations between strain rate and seismicity are found in all studied regions except the CEUS. The correlation between strain rate and seismic moment varies with regions.

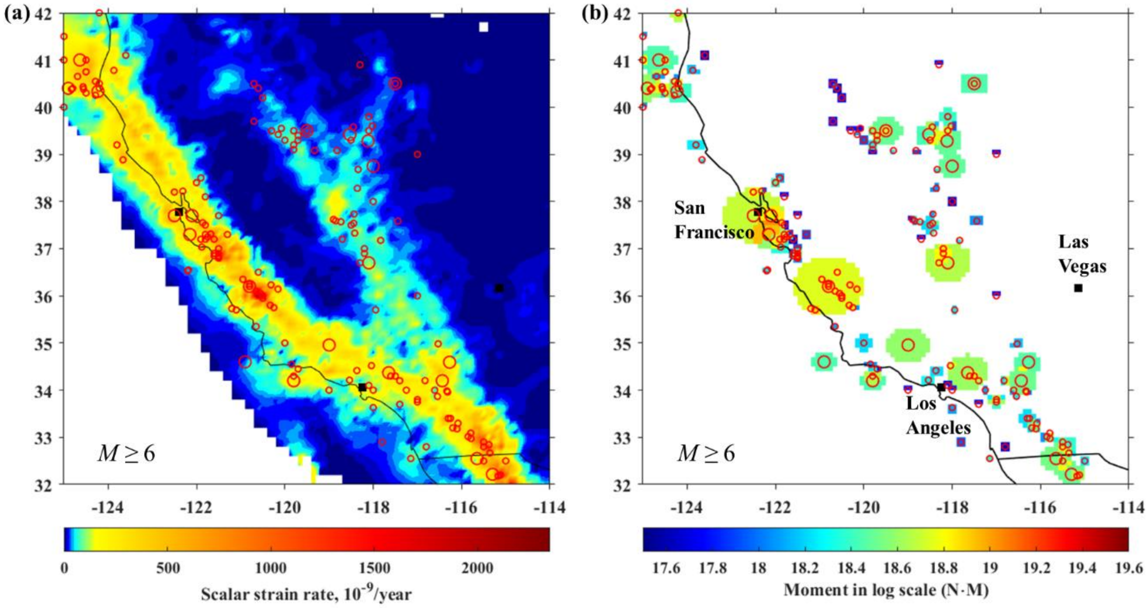


Figure 5.2. Comparison between strain rate, seismicity, and seismic moment in California and Nevada. **(a)** spatial distribution of $M \geq 6$ earthquakes (circles) between 1769 and 2016 and scalar strain rate (color contours). **(b)** spatial distribution of $M \geq 6$ earthquakes between 1769 and 2016 and their seismic moment. The earthquake data are from K. Mueller et al. (2004) used for the 2018 USGS National Seismic Hazard Map. The strain rate data are from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014).

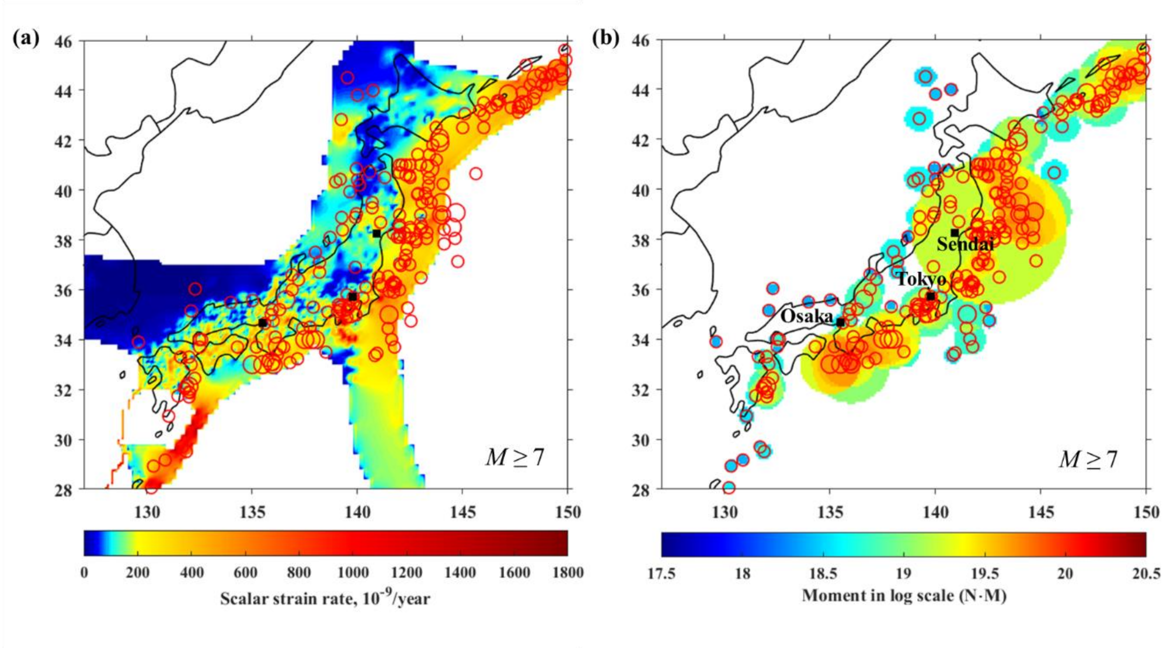


Figure 5.3. Comparison between strain rate, seismicity, and seismic moment in Japan. **(a)** spatial distribution of $M \geq 7$ earthquakes between 1096 and 2015 and scalar strain rate. **(b)** spatial distribution of $M \geq 7$ earthquakes between 1096 and 2015 and their seismic moment. The earthquake data is from the GEM Global Historical Earthquake Catalogue (1000-1903) and the ISC-GEM Global Instrumental Earthquake Catalogue (1904-2015). The strain rate data are from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014).

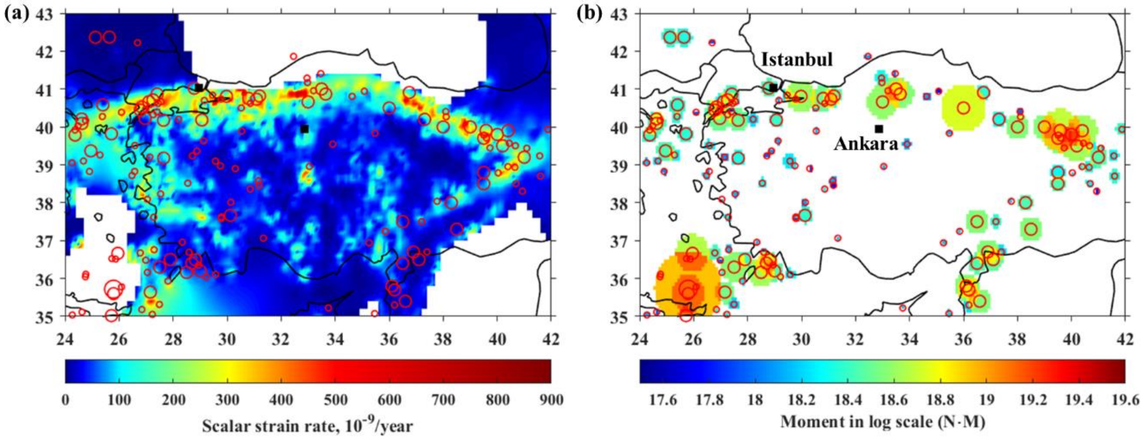


Figure 5.4. Comparison between strain rate, seismicity, and seismic moment in Anatolia. **(a)** spatial distribution of $M \geq 6$ earthquakes between 1045 and 2015 and scalar strain rate. **(b)** spatial distribution of $M \geq 6$ earthquakes between 1045 and 2015 and their seismic moment. The earthquake data is from the GEM Global Historical Earthquake Catalogue (1000-1903) and the ISC-GEM Global Instrumental Earthquake Catalogue (1904-2015). The strain rate data are from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014).

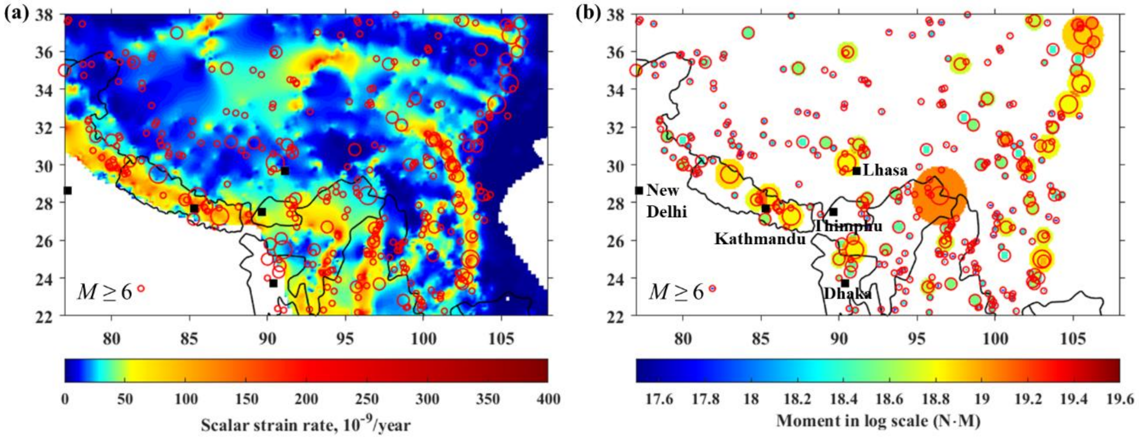


Figure 5.5. Comparison between strain rate, seismicity, and seismic moment in Tibet. **(a)** spatial distribution of $M \geq 6$ earthquakes between 1117 and 2015 and scalar strain rate.

(b) spatial distribution of $M \geq 6$ earthquakes between 1117 and 2015 and their seismic moment. The earthquake data is from the GEM Global Historical Earthquake Catalogue (1000-1903) and the ISC-GEM Global Instrumental Earthquake Catalogue (1904-2015). The strain rate data are from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014).

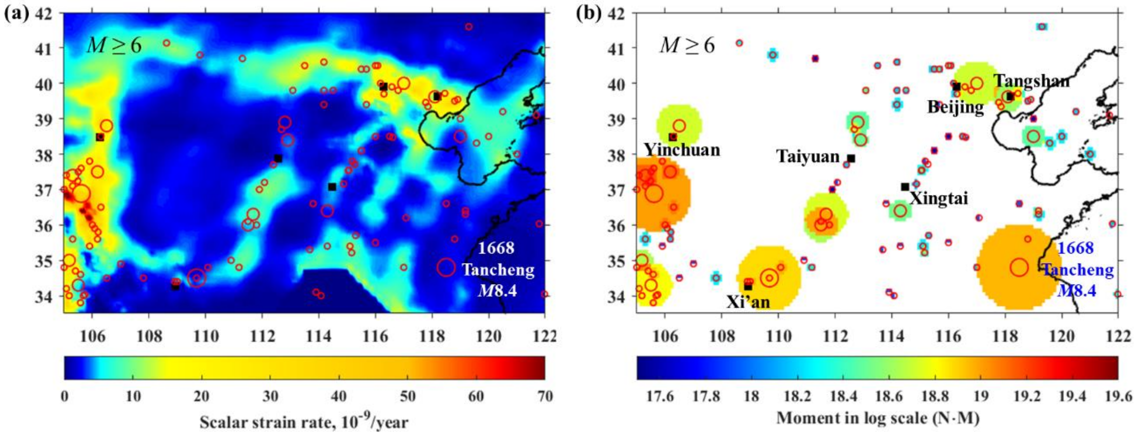


Figure 5.6. Comparison between strain rate, seismicity, and seismic moment in North China. **(a)** spatial distribution of $M \geq 6$ earthquakes between -780 and 2015 and scalar strain rate. **(b)** spatial distribution of $M \geq 6$ earthquakes between -780 and 2015 and their seismic moment. The earthquake data is from Cheng et al. (2017). The strain rate data are from the Global Strain Rate Model (GSRM v.2.1) (Kreemer et al., 2014).

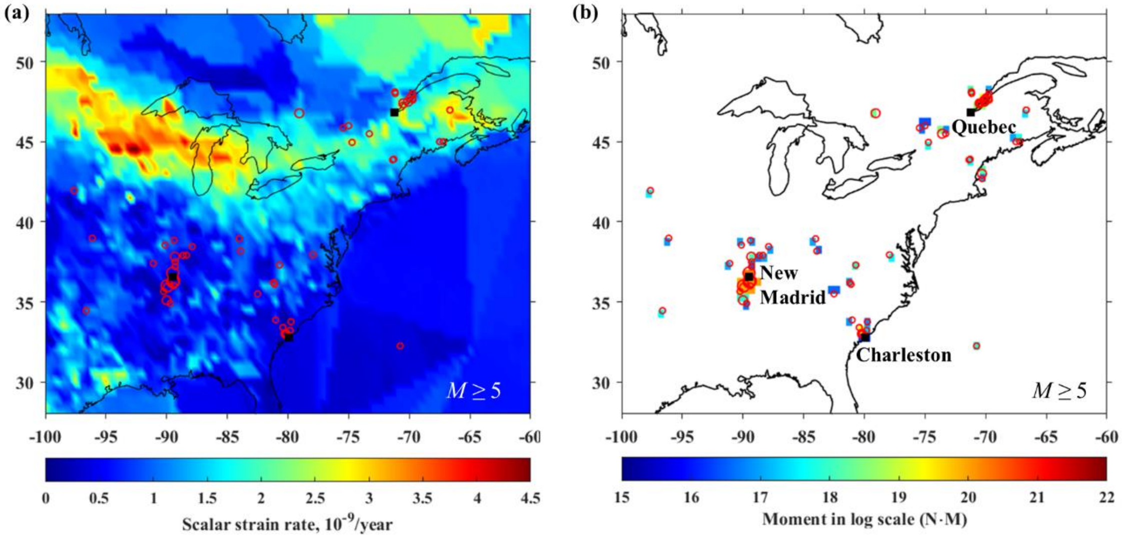


Figure 5.7. Comparison between strain rate, seismicity, and seismic moment in the CEUS. **(a)** spatial distribution of $M \geq 5$ earthquakes between 1568 and 2016 and scalar strain rate. **(b)** spatial distribution of $M \geq 5$ earthquakes between 1568 and 2016 and their seismic moment. The earthquake data are from K. Mueller et al. (2004) used for the 2018 USGS National Seismic Hazard Map. The strain rate data are from Kreemer et al. (2018).

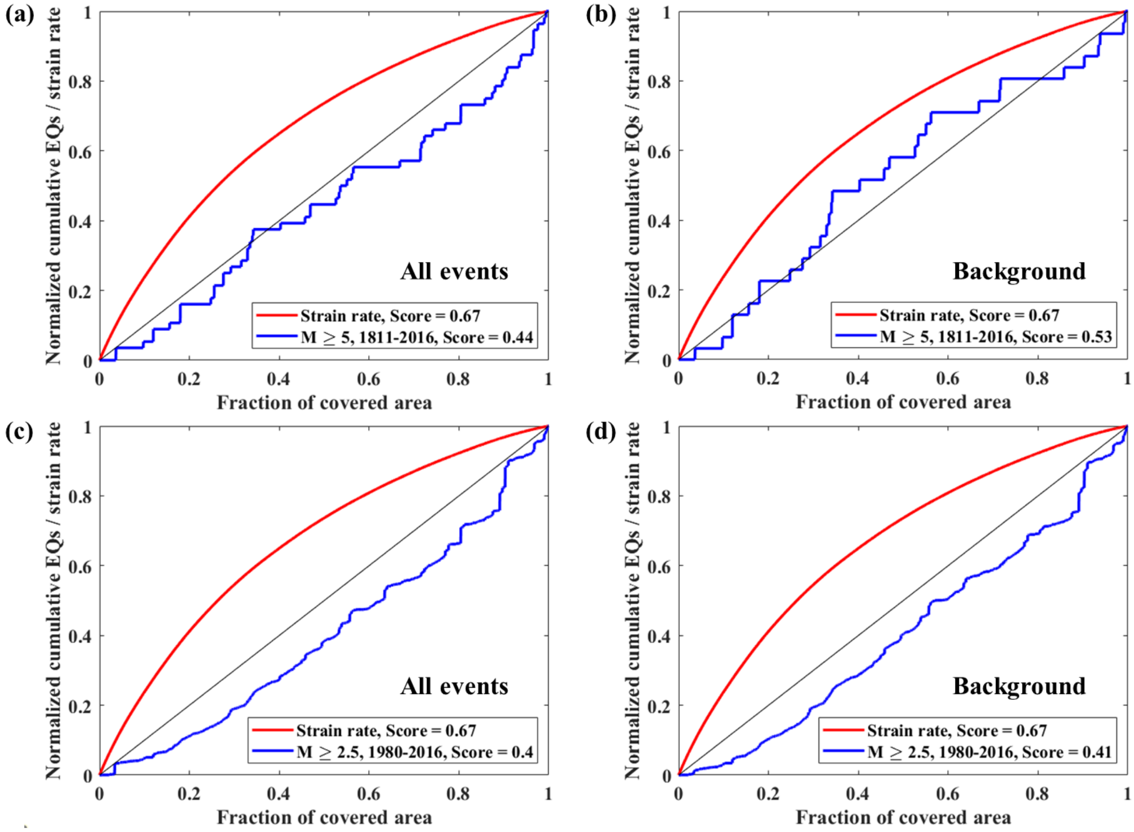


Figure 5.8. Comparison of the correlations between strain rate and seismicity in the CEUS for **(a)** all event or **(b)** background events of $M \geq 5$ between 1811 and 2016, and **(c)** all event or **(d)** background events of $M \geq 2.5$ between 1980 and 2016.

I also quantified the spatial prediction power of strain rate by using the area skill score, the area below the corresponding curve in the Figure 5.1. The area skill score is a parameter characterizing how concentrated one factor (e.g., earthquakes) is in a given area with a certain scanning order (e.g., based on strain rate). In this chapter, each tectonic region is scanned from cells of high strain rate to cells of low strain rate (the slope of the strain rate curve decreases as the fraction of covered area increases). Therefore, the area skill score of a seismicity (or seismic moment) curve characterize how concentrated seismicity (or seismic moment) is in high strain rate areas of a tectonic region. When the area skill score of a seismicity curve is larger than 0.5, it means more earthquakes occurred in high strain areas than in low strain rate areas, so strain rate as a spatial predictor is better than random guess. The results of different tectonic settings are shown in the legends in Figure 5.1 and in Table 5.1. The best spatial prediction of seismicity is found in California with the area skill score equals to 0.85, which is significantly higher than other regions (Figure 5.1). Japan, Anatolia, Tibetan Plateau, and North China have close area skill score of seismicity around 0.73. In Tibetan Plateau and North China, the area skill score of seismic moment (~0.65) is significantly lower than the area skill score of seismicity, but higher than the score of a random guess (0.5). For the CEUS, the prediction power of strain rate for seismicity (area skill score = 0.44, Figure 5.1) is significantly poorer than other studied regions. The differences between these three groups are significant and are shown in Figure 5.9.

Table 5.1. Comparison between different tectonic setting on spatial prediction power and strain rate-seismicity correlation based on Figure 5.1

Region	Time	Magnitude	Area skill score of strain rate	Difference between strain rate and EQ	Difference between strain rate and moment
California	1852-2016	M ≥ 6.5	0.85	0.025	0.028
Japan	1586-2015	M ≥ 7	0.73	0.023	0.022
Anatolia	1045-2015	M ≥ 7	0.76	0.077	0.057

Tibet	1786-2015	M ≥ 7	0.73	0.027	0.094
North China	1604-2015	M ≥ 6	0.75	0.021	0.111
CEUS	1568-2016	M ≥ 5	0.67	0.225	0.510

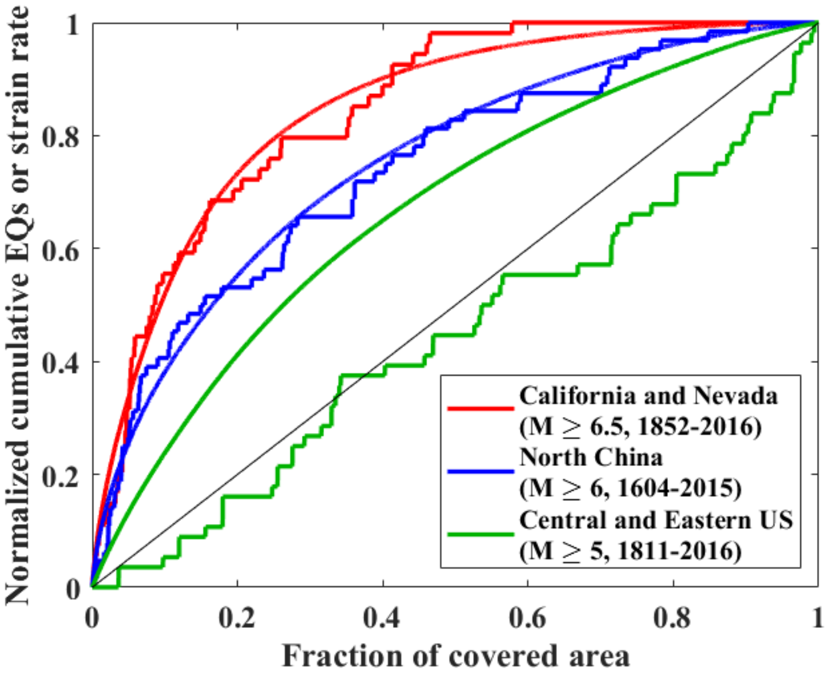


Figure 5.9. Comparison between strain rate and seismicity in different tectonic settings.

5.3.2 Temporal variations of strain rate-seismicity correlation

As I introduced in Chapter 2 and 3, temporal patterns of large earthquakes are complex and often show the pattern of Devil's staircses, with the alternation of active and quiescent periods (Figure 2.1, 3.15a). In North China, good correlation between strain rate and seismicity is found in the active periods but poor correlation is found in the relatively quiescent period (Figure 3.15b). Similar results are found in California-Nevada (Zeng et al., 2018).

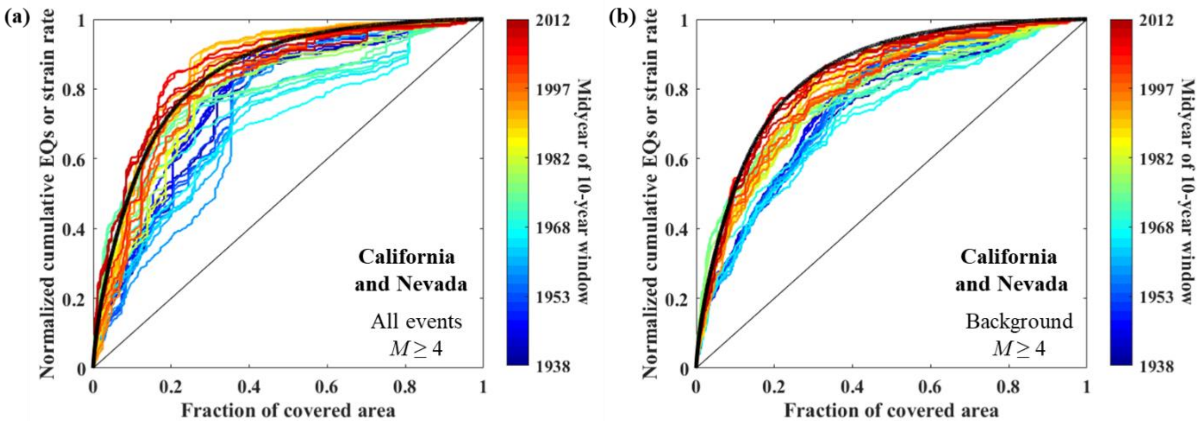


Figure 5.10. Temporal variations of strain rate-seismicity correlation in California and Nevada for **(a)** all events or **(b)** background earthquakes. The cumulative earthquakes are counted within a 10-year window that moves in 2-year steps from 1933 to 2016. The colorbar shows the midyear of the moving 10-year time windows used to calculate the cumulative earthquake counts. The thick black curve is the cumulative strain rate. The diagonal line indicates spatially random distribution.

I also explored the temporal variations of correlation between strain rate and relatively small earthquakes in California, North China, and CEUS. Both the original catalogs and declustered catalogs are studied to check if the temporal trends are consistent. The declustered catalogs are obtained using the nearest-neighbor method introduced in Chapter 2. In California and Nevada, the correlation between strain rate and seismicity also varies with time (Figure 5.10). Both background earthquakes and all events follow the same trend: with poor correlations from 1933 to the 1980s and better correlations from the 1980s to 2016, similar to the results for background earthquakes of Zeng et al. (2018). A reverse trend is observed in North China from 1970 to 2015 (Figure 5.11). For all events (Figure 5.11a), the seismicity curves are above the strain rate curve at the early periods because most events are aftershocks of the 1976 Tangshan earthquake, which caused the concentration of seismicity in high strain rate areas. As time passed, aftershock activity decayed and background earthquakes that occurred in the low strain rate regions become relatively dominant, therefore the correlation between strain rate and seismicity gets worse. Similar trend is also found for background seismicity (Figure 5.11b). After 2000, almost all events in North China are background earthquakes and diffusely distributed in both high and low strain rate regions. Together with the lack of large earthquakes in North China after the 1976 Tangshan earthquake (Figure 3.15a), North China may have entered a relatively quiescent period. In the CEUS, the temporal variation of the correlation exists but is smaller than that in California-Nevada and North China (Figure 5.12). The seismicity curves are below the spatial random line (the diagonal line) and moves downward as time passed, indicating even more small events occurred in low strain rate areas.

Comparing the temporal variations of the seismicity curves between California-Nevada and North China (Figure 5.10, 5.11), I found that spatial distribution of seismicity in North China is more time-dependent than that in California-Nevada. In California-Nevada, even for the period with the worst correlation between strain rate and seismicity, the seismicity curve is still significantly above the diagonal line (Figure 5.10), so seismicity is still concentrated in high strain rate regions, although not as concentrated as the strain rate itself. However, in North China, during the most uncorrelated period, the spatial distribution of seismicity is close to random (Figure 5.11).

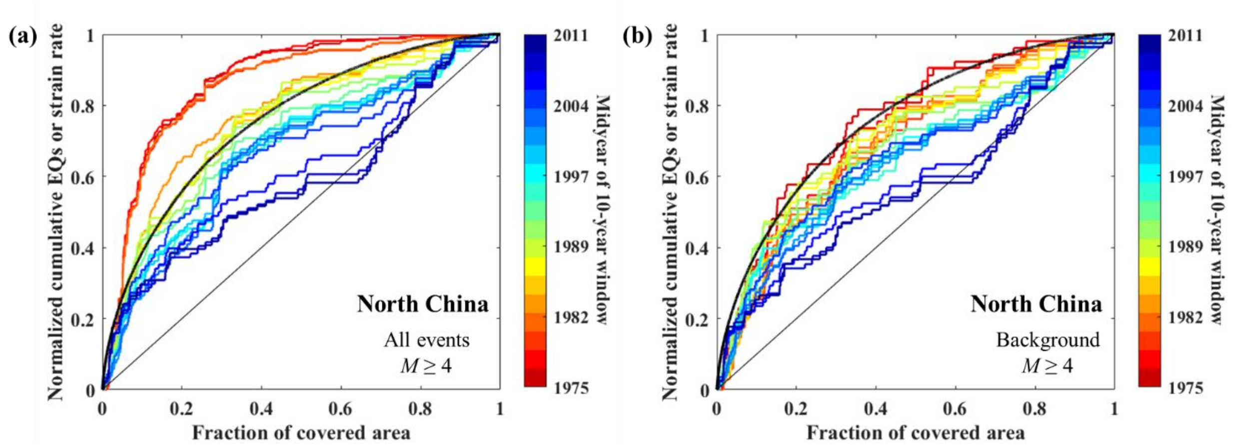


Figure 5.11. Temporal variations of strain rate-seismicity correlation in North China for **(a)** all events or **(b)** background earthquakes. The cumulative earthquakes are counted within a 10-year window that moves in 2-year steps from 1970 to 2015.

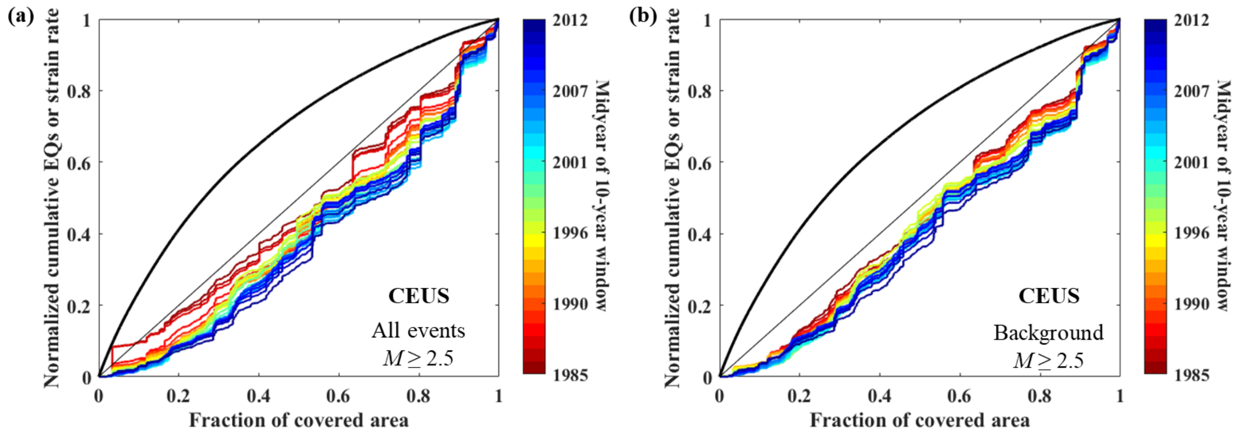


Figure 5.12. Temporal variations of strain rate-seismicity correlation in CEUS for **(a)** all events or **(b)** background earthquakes. The cumulative earthquakes are counted within a 10-year window that moves in 1-year steps from 1980 to 2016.

I further checked to see if the length of the time window affects the observed temporal trend. The answer is not. In addition to the 10-year window, I also used a 4-year window and a 20-year window for North China earthquakes (Figure 5.13). Although the results based on the 4-year window have larger uncertainties because of fewer events, the temporal trend is still clear (Figure 5.13a, b). The results based on the 20-year window, because of more events included, show a more stable temporal trend of seismicity from concentration on high strain rate regions to more random and diffused distribution (Figure 5.13c, d).

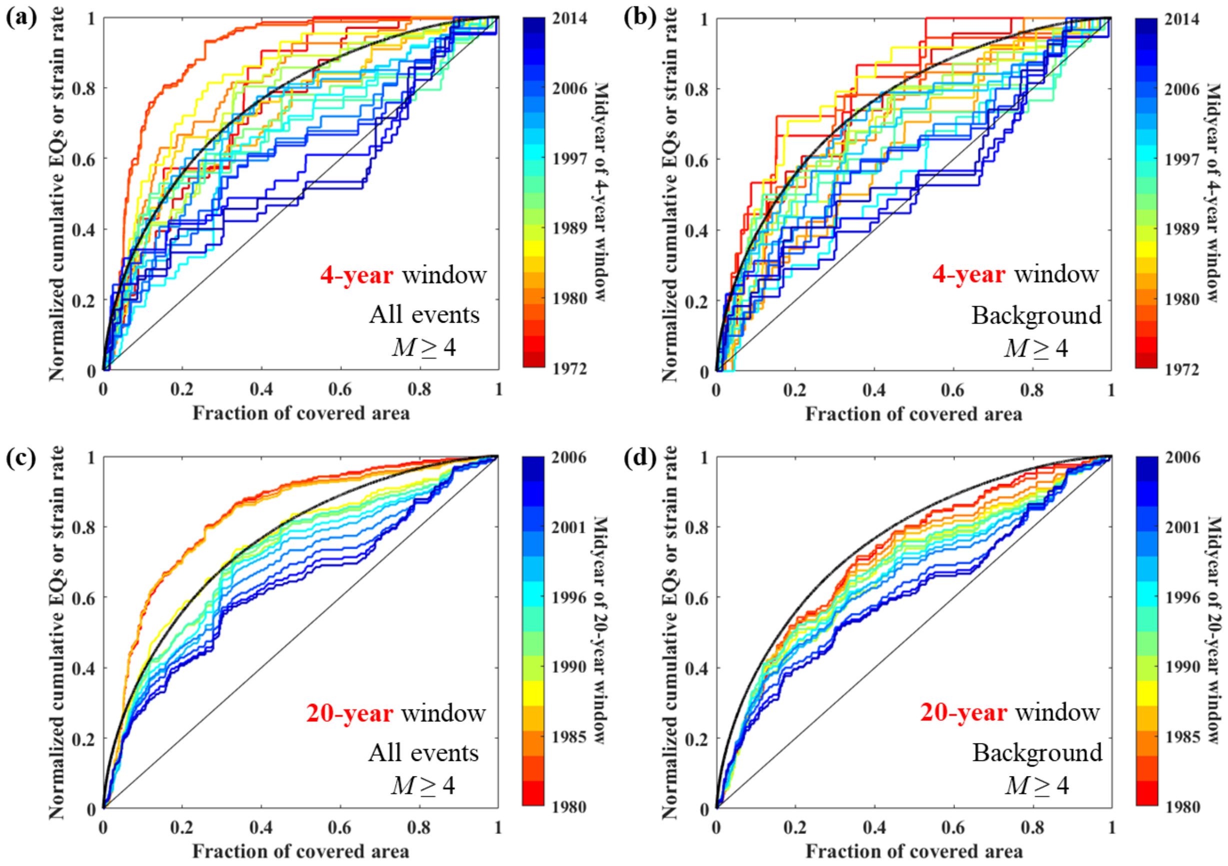


Figure 5.13. Temporal variations of strain rate-seismicity correlation in North China by using different length of time window.

5.4 Discussion

5.4.1 Deviations from the dichotomy between plate boundaries and plate interiors: strain rate-seismicity correlations in different tectonic settings

In this chapter, I characterized the correlation between strain rate, seismicity, and seismic moment by using the success diagram following the approach of Shen et al. (2007) and Zeng et al. (2018). The spatial prediction power of strain rate is evaluated by using the area score (Zechar & Jordan, 2008) of cumulative number of earthquakes (or cumulative seismic moment) counting from high strain rate areas to low strain rate areas. I found that the strain rate-seismicity correlation is complex and varies in different tectonic settings (Figure 5.1). In relatively simple plate boundary zones (e.g., California and Japan), high strain rates are highly concentrate on major faults where large earthquakes occur and records are complete, so good correlations between strain rate, seismicity, and seismic moment are found (Figure 5.1a, b) (Shen et al., 2007; Zeng et al., 2018). Therefore, strain rate has a good spatial prediction power of seismicity. In broad plate boundary zones (e.g., Tibetan Plateau) and boundaries of micro plates (e.g., Anatolia), correlations between strain rate and seismicity are still good, but the spatial prediction power of strain rate gets worse (Figure 5.1c, d), because the strain rate distributions are not as concentrated as that in relatively simple plate boundary zones. In reactivated cratons (e.g., North China), while the distributions of fault systems, strain rate, and seismicity are complex (Chen et al., 2021; Liu & Wang, 2012), a reasonably good correlation between strain rate and seismicity is found (Figure 5.1e). The spatial prediction power of strain rate is comparable to complex plate boundary zones and better than random guess. However, the correlation between strain rate and seismic moment in reactivated cratons may be not good, even in a record of hundreds of years (Figure 5.1e). Because of long recurrence time of large earthquakes in intraplate regions, spatial patterns of seismicity depend on the length of catalogs, and large earthquakes in low strain rate areas may dominate the moment release in the earthquake catalogs. In stable plate interiors (e.g., the CEUS), correlations between strain rate are found to be poor or absent for both large and small earthquakes (Figure 5.1f, 5.8), similar to previous studies (Calais et al., 2016; Kreemer et al., 2018). In such settings, earthquakes are nearly randomly distributed in space, so strain rate cannot be used as a useful spatial predictor of seismicity in stable intraplate continents.

5.4.2 The complexity of intraplate seismicity

Although both North China and the CEUS are intraplate regions, the correlations between strain rate and seismicity in these two regions are significantly different (Figure 5.1e, f), which highlights the complexity of intraplate seismicity. The lengths of earthquake records in intraplate regions are often too short compared to their average recurrence intervals, so spatiotemporal patterns of intraplate earthquakes may not be complete. The relatively good strain rate-seismicity correlation in North China may indicate that the observed geodetic strain rate reflects the long-term interseismic loading. However, the long-lasting aftershocks and postseismic deformation of the 1966 Xingtai earthquake and 1976 Tangshan earthquake in North China (Chen et al., 2021; Liu &

Wang, 2012) may cause an overestimation about the correlation between strain rate and seismicity in North China. Unlike North China, in CEUS, glacial isostatic adjustment (GIA) is the main cause of the observed strain rate field, and poor or no correlation between strain rate and seismicity is observed (Figure 5.1f) (Calais et al., 2006; Kreemer et al., 2018). Based on the lack of strain rate-seismicity correlation in CEUS, Kreemer et al. (2018) argue that “intraplate seismicity does not reflect the release of geodetic strain, and the largest, GIA-controlled, strain rate does not load faults, except perhaps in zones of weakness such as continental margins”. Nevertheless, because of limitations and uncertainties of observations, intraplate seismicity is still enigmatic and need further studies.

5.4.3 Implications to seismic hazard

It is important to know how good strain rate is for predicting locations of earthquakes. In this chapter, I found that the spatial prediction power of strain rate varies with different tectonic regions, which should be considered in seismic hazard assessment. Except the stable continents, using strain rate to predicts earthquake locations (or seismic moment distributions) is better than random guess in plate boundary zones and active intraplate regions. Strain rate-based models can be useful for intraplate regions that current strain rate fields reflect long-term interseismic loading patterns. Large intraplate earthquakes often occur on unmapped faults because of short records, but observed high strain rate areas may offer information about the possible high risk regions. If the strain rate spatially correlates with the seismic moment in a long term, the deviations between the strain rate curve and the seismic moment curve in a short-term record may offer information about where strain is not sufficiently released. For example, in North China (Figure 5.1e), future large earthquakes are more likely to occur in the medium strain rate regions because the stored energy there has not been sufficiently released in the past 400 years.

The temporal variations of the spatial distribution of seismicity show dynamic patterns of seismicity and need to be considered in the seismic hazard assessment. The standard seismic hazard maps often estimate seismic hazard in next 50 years (Petersen et al., 2014; Petersen et al., 2020). However, in a 50-year time window, the spatial distributions of seismicity can vary significantly (Figure 5.10, 5.11). This effect is smaller in plate boundary zones, because recurrence intervals there are shorter than that in intraplate regions and a large portion of events is still concentrated in high strain areas even in the period of the worst strain rate-seismicity correlation (Figure 5.10). In intraplate regions like North China, because of long recurrence intervals, spatial distributions of seismicity can be significantly different in different time periods (Figure 3.15, 5.11). In an active period, earthquakes occur in high strain rate areas, but in a relatively quiescent period, earthquakes are closer to be randomly distributed. The temporal variations of spatial distributions of small seismicity show clear trend from one period to another (Figure 5.10, 5.11), so may be used to tell if a region is entering a more active period or a relative quiescent period, as suggested by Zeng et al. (2018).

5.5 Conclusions

In this chapter, I systematical studied the correlation between strain rate and seismicity and evaluated how good strain rate is as a spatial indictor of earthquakes in different tectonic settings. The strain rate-seismicity correlation is good not only in plate boundary zones but also in some intraplate regions like North China. However, the correlation between strain rate and seismicity is poor or absent in stable plate interiors like the CEUS. The spatial prediction power of strain rate varies with different tectonic settings. In relatively narrow plate boundary zones, the spatial prediction power of strain rate is the best, with most events concentrated in high strain rate areas. The spatial prediction power of strain rate become worse in broad plate boundary zones and active continents, but is still better than random guess. In stable continents, spatial patterns of seismicity are close to a random distribution: lack of preference on high strain rate areas, so strain rate cannot be used as a spatial indictor of seismicity there.

The strain rate-seismicity correlation is time-dependent, not only for all events but also for background earthquakes. Good correlations are found in seismically active periods and poor in relatively quiescent periods, so it may offer information about current seismicity rate level for a given region. The strain rate-seismicity correlation has larger temporal variations in intraplate regions than that in plate boundary zones, so spatial patterns of seismicity in intraplate regions are less stable in time and should be considered in seismic hazard assessments.

References

Albini, P., Musson, R., Gomez Capera, A., Locati, M., Rovida, A., Stucchi, M., & Viganò, D. (2013). Global historical earthquake archive and catalogue (1000-1903).

Albini, P., Musson, R. M., Rovida, A., Locati, M., Gomez Capera, A. A., & Viganò, D. (2014). The global earthquake history. *Earthquake Spectra, 30*(2), 607-624.

Beavan, J., & Haines, J. (2001). Contemporary horizontal velocity and strain rate fields of the Pacific-Australian plate boundary zone through New Zealand. *Journal of Geophysical Research: Solid Earth*, *106*(B1), 741-770.

Blaser, L., Krüger, F., Ohrnberger, M., & Scherbaum, F. (2010). Scaling relations of earthquake source parameter estimates with special focus on subduction environment. *Bulletin of the Seismological Society of America, 100*(6), 2914-2926.

Bungum, H., Olesen, O., Pascal, C., Gibbons, S., Lindholm, C., & Vestøl, O. (2010). To what extent is the present seismicity of Norway driven by post-glacial rebound? *Journal of the geological society, 167*(2), 373-384.

Calais, E., Camelbeeck, T., Stein, S., Liu, M., & Craig, T. (2016). A new paradigm for large earthquakes in stable continental plate interiors. *Geophysical Research Letters, 43*(20).

Calais, E., Han, J., DeMets, C., & Nocquet, J. (2006). Deformation of the North American plate interior from a decade of continuous GPS measurements. *Journal of Geophysical Research: Solid Earth, 111*(B6).

Chen, Y., Liu, M., & Wang, H. (2021). Aftershocks and background seismicity in Tangshan and the rest of North China. *Journal of Geophysical Research: Solid Earth*, e2020JB021395.

Cheng, J., Rong, Y., Magistrale, H., Chen, G., & Xu, X. (2017). An M w-based historical earthquake catalog for Mainland China. *Bulletin of the Seismological Society of America, 107*(5), 2490-2500.

Giacomo, D. D., Engdahl, E. R., & Storchak, D. A. (2018). The ISC-GEM Earthquake Catalogue (1904–2014): status after the Extension Project. *Earth System Science Data, 10*(4), 1877-1899.

Gordon, R. G. (1998). The plate tectonic approximation: Plate nonrigidity, diffuse plate boundaries, and global plate reconstructions. *Annual Review of Earth and Planetary Sciences, 26*(1), 615-642.

Haines, A., & Holt, W. (1993). A procedure for obtaining the complete horizontal motions within zones of distributed deformation from the inversion of strain rate data. *Journal of Geophysical Research: Solid Earth, 98*(B7), 12057-12082.

Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth, 84*(B5), 2348-2350.

Holt, W. E., Shen-Tu, B., Haines, J., & Jackson, J. (2000). On the determination of self-consistent strain rate fields within zones of distributed continental deformation. *Washington DC American Geophysical Union Geophysical Monograph Series, 121*, 113-141.

Keiding, M., Kreemer, C., Lindholm, C., Gradmann, S., Olesen, O., & Kierulf, H. (2015). A comparison of strain rates and seismicity for Fennoscandia: depth dependency of deformation from glacial isostatic adjustment. *Geophysical Journal International, 202*(2), 1021-1028.

Kreemer, C., Blewitt, G., & Klein, E. C. (2014). A geodetic plate motion and Global Strain Rate Model. *Geochemistry, Geophysics, Geosystems, 15*(10), 3849-3889.

Kreemer, C., Hammond, W. C., & Blewitt, G. (2018). A robust estimation of the 3-D intraplate deformation of the North American plate from GPS. *Journal of Geophysical Research: Solid Earth, 123*(5), 4388-4412.

Kreemer, C., Holt, W. E., & Haines, A. J. (2002). The global moment rate distribution within plate boundary zones. *Plate boundary zones, 30*, 173-190.

Liu, M., & Wang, H. (2012). Roaming earthquakes in China highlight midcontinental hazards. *Eos, Transactions American Geophysical Union, 93*(45), 453-454.

Michael, A. J. (2014). How complete is the ISC-GEM global earthquake catalog? *Bulletin of the Seismological Society of America, 104*(4), 1829-1837.

Molchan, G., & Kagan, Y. Y. (1992). Earthquake prediction and its optimization. *Journal of Geophysical Research: Solid Earth, 97*(B4), 4823-4838.

Mueller, C. S. (2019). Earthquake catalogs for the USGS national seismic hazard maps. *Seismological Research Letters, 90*(1), 251-261.

Mueller, K., Hough, S. E., & Bilham, R. (2004). Analysing the 1811–1812 New Madrid earthquakes with recent instrumentally recorded aftershocks. *Nature, 429*(6989), 284-288.

Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D., et al. (2014). *Documentation for the 2014 update of the United States national seismic hazard maps*. Retrieved from

Petersen, M. D., Shumway, A. M., Powers, P. M., Mueller, C. S., Moschetti, M. P., Frankel, A. D., et al. (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. *Earthquake Spectra, 36*(1), 5-41.

Shen, Z.-K., Jackson, D. D., & Kagan, Y. Y. (2007). Implications of Geodetic Strain Rate for Future Earthquakes, with a Five-Year Forecast of M5 Earthquakes in Southern California. *Seismological Research Letters, 78*(1), 116-120.

Stevens, V., & Avouac, J. (2021). On the relationship between strain rate and seismicity in the India–Asia collision zone: implications for probabilistic seismic hazard. *Geophysical Journal International, 226*(1), 220-245.

Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H., et al. (2013). Public release of the ISC–GEM global instrumental earthquake catalogue (1900–2009). *Seismological Research Letters, 84*(5), 810-815.

Storchak, D. A., Di Giacomo, D., Engdahl, E., Harris, J., Bondár, I., Lee, W. H., et al. (2015). The ISC-GEM global instrumental earthquake catalogue (1900–2009): introduction. *Physics of the Earth and Planetary Interiors, 239*, 48-63.

Tarayoun, A., Mazzotti, S., Craymer, M., & Henton, J. (2018). Structural inheritance control on intraplate present-day deformation: GPS strain rate variations in the Saint Lawrence Valley, eastern Canada. *Journal of Geophysical Research: Solid Earth, 123*(8), 7004-7020.

Ward, S. N. (1998). On the consistency of earthquake moment rates, geological fault data, and space geodetic strain: the United States. *Geophysical Journal International, 134*(1), 172-186.

Weiss, J. R., Walters, R. J., Morishita, Y., Wright, T. J., Lazecky, M., Wang, H., et al. (2020). High-resolution surface velocities and strain for Anatolia from Sentinel-1 InSAR and GNSS data. *Geophysical Research Letters, 47*(17), e2020GL087376.

Zechar, J. D., & Jordan, T. H. (2008). Testing alarm-based earthquake predictions. *Geophysical Journal International, 172*(2), 715-724.

Zeng, Y., Petersen, M. D., & Shen, Z. K. (2018). Earthquake potential in California-Nevada implied by correlation of strain rate and seismicity. *Geophysical Research Letters, 45*(4), 1778-1785.

Chapter 6: Conclusions and suggestions for future research

Chapter 6.1: Conclusions

My studies presented in this dissertation have improved the understanding about spatiotemporal patterns of earthquakes. For all fault systems ranging from single faults to different tectonic regions and to the whole world, I found that the periodic or quasiperiodic occurrence of earthquakes predicted by the elastic rebound model is unusual. Clusters of earthquakes separated by irregular and relatively longer quiescent periods are generally observed. Such temporal patterns of earthquakes can be mathematically described by the Devil's Staircase, a fractal property of complex systems commonly seen in nature. For the possible tectonic factors and causes for the Devil's Staircase patterns of large earthquakes, I found that the lengths of quiescent periods between clusters are inversely related to tectonic loading rate. Earthquake clustering can be attributed to viscoelastic relaxation and fault interaction. The Devil's Staircase patterns of large earthquakes have important implications for earthquake hazard assessment. Because catalogs for large earthquakes often too short to reflect their complete patterns, the mean recurrence interval, being estimated using the catalogs and used in seismic hazard analyses, can vary significantly depending on whether the catalogs include only the most recent active period or cover a spectrum of clusters and quiescent intervals. The probability distribution of inter-event time is found burstier than the Poisson model commonly used in probabilistic seismic hazard assessments, implying a higher probability of repeating events soon after a large earthquake. Large fault ruptures, which often involve multiple fault segments or faults, behave as nonlinear complex fault systems and show the Devil's Staircase patterns, requiring earthquake studies using a system approach.

Compared with seismicity at plate boundary regions, seismicity in plate interiors is more complex but less studied. I studied the intraplate seismicity in North China and the CEUS in this dissertation, and focused on the spatial temporal patterns of background seismicity, possible long-lived aftershocks of historical large earthquakes, and correlations of seismicity with different factors (e.g., strain rate, fault, seismic velocity). For North China, the recent moderate earthquakes ($M \geq 4.5$) in the Tangshan region, China are statistically identified aftershocks and may be physically trigged by the 1976 Great Tangshan earthquake. The aftershock activity of the 1976 Great Tangshan earthquake is likely continuing, and its duration is around 65-100 years. The background seismicity in North China is relatively stable in space with variations in time. The active tectonic structures, the Shanxi Rift and the Z-P fault system, are found concentrations of large historical earthquakes, background seismicity, and relatively high strain rate. However, in the North China Plain, the correlation between large historical earthquakes, background seismicity, and strain rate is weak. A significant number of large earthquakes occurred in places where geodetic strain rate or background seismicity is relatively low, highlighting the complexity of intraplate seismicity. In the whole North China, the correlation between strain rate and large earthquakes is generally good in a 2000-year record, but the correlation is worse for smaller background seismicity. The spatial prediction power of strain rate in North China is better than random guess but not as good as in plate boundary regions. The fault map is also a good constraint on the spatial pattern of seismicity in North China. However, it is worthy of noticing that some large earthquakes occurred away from mapped faults and on faults that were formerly unknown. Combining fault data with strain rate data can improve the prediction performance for large earthquakes. The correlation between seismicity and the S-wave anomaly at 100km depth is weak. I also noticed that North China is lack of large earthquakes in the past 40 years and current background seismicity is close to random distribution in space, so North China may have entered a relatively quiescent period.

For the CEUS, the hypothesis that current earthquake clusters are aftershocks of large historical earthquakes are tested by using a statistical method called the nearest-neighbor method. The results show that modern seismicity in the New Madrid seismic zone and the Charleston seismic zone include long-lived aftershocks of the 1811-1812 New Madrid earthquakes and the 1886 Charleston earthquake, while contemporary seismicity in the Charlevoix seismic zone are likely to be background seismicity. I found that the magnitude of a given historical earthquake significantly affects the number of its aftershocks and the duration of its aftershock activity; the larger the historical earthquake is, the more its aftershocks are and the longer its aftershock activity lasts. Even assigning the smallest magnitudes to the 1811-1812 New Madrid earthquakes, still 10.7% of events in the New Madrid seismic zone between 1980 and 2016 are identified as the aftershocks of the 1811-1812 events. Misidentifying long-lived aftershocks as background earthquakes may cause overestimation of seismic hazard in their rupture zones in intraplate regions.

The spatial correlation between strain rate and seismicity and the prediction power of strain rate are also studied in this dissertation. I found the correlation between strain rate and seismicity, rather than being consistently good, varies in different tectonic settings and different time periods. The good correlation between strain rate and seismicity is found in plate boundary regions and during seismic active periods. In contrast, in stable plate interiors and during relatively quiescent periods, the correlation between strain rate and seismicity are found poor or absent. The temporal variation of the strain rate-seismicity correlation is larger in intraplate regions than that in plate boundary regions. In intraplate regions, seismicity may be highly concentrated in regions of high strain rate in one period, but become nearly randomly distributed in space in the next period. The strong time-dependence of intraplate seismicity should be considered in seismic hazard assessment. The spatial prediction power of strain rate for seismicity varies in different tectonic settings, because it is decided not only by how good strain rate-seismicity correlation is but also by how concentrated strain rate is for a give region. Strain rate has best prediction power in plate boundary zones, has worse but still good prediction power in broad plate boundary regions and some intraplate regions. However, in stable continents, strain rate cannot be used as a spatial predictor of earthquakes, because its effect is close to random guess.

Chapter 6.2: Suggestions for future researches

The inadequacy of the elastic rebound model and the complex Devil’s staircase pattern of earthquakes require a system approach. Multiple fault segments of a mature fault or multiple faults in a region forms a nonlinear complex system. The behaviors of interactive fault segments or faults as a whole cannot be fully understood by only studying individual fault segments or faults. Therefore, rather than focusing only on stress accumulation and release on individual faults, the stress revolution of multiple faults needs to be explored. The fault interaction and viscoelastic relaxation, mentioned in Chapter 2, need to be further explored to better understand the physics of earthquakes.

Another important thing is to completely map faults in intraplate regions. I pointed out in Chapter 3 that some large intraplate earthquakes occurred on faults that were previously unknown. In addition, many induced earthquakes have also occurred on unmapped faults. Good correlation between strain rate and seismicity in intraplate region like North China may help to identify some unmapped faults.

I did a simple exploration on how to combin fault data and strain rate data to improve the spatial prediction of earthquakes in Chapter 3, but it is too simple and is not the optimal solution. Future researches can dig deeper into this multifactor problem. For example, exploring the permutation and combination of strain rate, fault, heat flow, and seismic velocity. Then the methods that are built from the researches of the multifactor problem can be applied to different regions.

VITA

Yuxuan Chen was born in Xiangtan, Hunan Province, China. He received his Bachelor of Science in Geophysics from the University of Science and Technology of China, Anhui, China in Spring 2017. He began his doctoral studies at the Department of Geological Science, University of Missouri-Columbia in Fall 2017. His research interests are in intraplate seismicity, earthquake clustering, and seismic hazard.