

Detection of a Non-local Seismicity Pattern Prior to Larger-size Earthquakes in Taiwan

Shyh-Yang Sheu^{1,*}, Chiou-Fen Shieh¹, and Ruey-Chyuan Shih¹

(Manuscript received 30 August 2001, in final form 9 April 2002)

ABSTRACT

In this paper, we estimate the varying probability of the occurrence of earthquakes with magnitudes greater than 6 by using a pattern dynamic algorithm. The pattern dynamic algorithm extracts information by constructing a pattern state vector evaluated from past seismic activities. The pattern state vector is expressed as a unit vector in a high-dimensional correlation space. To reduce the bias in constructing a pattern state vector, we only use high quality data provided by the Central Weather Bureau in Taiwan. Earthquakes that occurred from 1973 to 1997 are used as a “learning catalogue” to evaluate the future seismicity patterns. Using a log-likelihood function, we found that the optimized time span of 8 to 10 years is appropriate for a learning catalogue. The spatial distributions of increasing probability for each time span for the earthquakes after 1998 is estimated. Results show that most of the larger-size earthquakes that occurred inland after 1998 were located in the estimated areas. Earthquakes in the eastern, offshore Taiwan area, where the data quality is poor, could not be well located.

(Key words: Seismicity pattern, Pattern dynamic, Correlation space, Learning catalogue)

1. INTRODUCTION

Large earthquakes generally cause death and destruction, such as the 1995 Kobe, Japan earthquake, the 1999 Izmit, Turkey event, which killed more than 17000 people, the 1999 Chi-Chi, Taiwan earthquake, and the 2001 El Salvador earthquake. Seismologists have been seeking a reliable methodology for forecasting earthquakes for at least 30 years. Although many observable quantities and methodologies of earthquake forecasting were explored, very little progress was made (Geller 1991; Geller et al. 1997). Scientists proposed different kinds of

¹Institute of Seismology and Applied Geophysics, National Chung Cheng University, Chia-Yi, Taiwan

* *Corresponding author address:* Prof. Shyh-Yang Sheu, Institute of Seismology and Applied Geophysics, National Chung Cheng University, Min-hsiung, Chia-yi 621, Taiwan, ROC

E-mail: seifent@eq.ccu.edu.tw

phenomenological models to describe seismicity patterns, especially for those in the source region, such as Mogi's donuts (Mogi 1969), seismic gap with seismic quiescence (Haberman 1981; Kisslinger 1988; Wyss and Weimer 1997), and characteristic earthquakes (Schwartz and Coppersmith 1984). In addition, temporal clustering of earthquakes (Frolich 1987; Rundle et al. 1997) and earthquake triggered over large distances (King et al. 1984) were also studied. These models were based on the hypothesis of localized seismicity pattern in the source region. The drawback had been the need of knowing where the event (source region) would occur before the algorithm was applied.

Recently, Mori and Kuramoto (1998) showed that earthquakes could be considered to be an example of a phase dynamical system, in which the evolution of the observables can be associated with the rotation of a state or pattern vector in a high-dimensional correlation space. Some observations also suggested that earthquake phenomena can be described by strongly correlated dynamics (Bufe and Varnes 1993; Press and Allen 1995; Knopoff et al. 1996). Rundle et al. (2000) described in detail the development of the correlation in stress field related to a necessary precondition for the emergence of large earthquakes. Furthermore, numerical experiments showed non-local features of the space-time patterns of earthquakes (Gross and Rundle 1998; Rundle et al. 2000).

Earthquake faults are regarded as a mean-field threshold system, with statistically stationary dynamics over a long time interval (Ferguson et al. 1999). The seismicity patterns clearly reflect a space-time correlation with stress and strain fields. Using the seismicity variation function over a time interval and based on the phase dynamics assumption (i.e., pattern state dynamics are statistically stationary over long time intervals.), we could compute the probability of an anomalous correlation (Tiampo et al. 2000) preceding a large earthquake.

In this study, we estimate the possible future seismicity pattern by assuming the existence of anomalous, non-local space-time patterns and correlations associated with events occurring in Taiwan prior to events with $M \geq 6.0$.

2. METHODOLOGY

Here we briefly summarize the method used in this study. A more detailed description can be found in Tiampo et al. (2000) and Rundle et al. (2000). The seismicity is first arranged into a time series, $\phi(x_i, t_0, t)$, at a specific locality, x_i , from time t_0 to t , and its amplitude is characterized by the number of earthquakes at unit time.

The method is formulated based on the evolution of seismicity by defining a real-value seismic phase function, $s(x_i, t_0, t)$. The phase function is defined as the time average at x_i over the time interval of $t-t_0$, i.e.,

$$S(x_i, t_0, t) = \frac{1}{t - t_0} \int_{t_0}^t \phi(x_i, t_0, t) dt . \quad (1)$$

Tiampo et al. (2000) showed that $s(x_i, t_0, t)$ provided little evidence to show the forerunning activity of a large earthquake. Following the assumption that seismicity could be described by phase dynamics (Rundle et al. 2000), we redefine the unit vector pointing in the direction of $s(x_i, t_0, t)$ by using an $L2$ norm. Note that seismicity can be considered to be an incoherent super-

position of phase functions, which have a zero-mean. In order to obtain a zero-mean phase function, it is necessary to remove the spatial mean and to create a unit vector by dividing the standard deviation. The spatial mean is

$$\bar{S}(x_i, t_0, t) = \frac{1}{N} \sum_{i=1}^N S(x_i, t_0, t) , \quad (2)$$

and the standard deviation is

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (S(x_i, t_0, t) - \bar{S}(x_i, t_0, t))^2 . \quad (3)$$

The unit vector is therefore defined as

$$\hat{S}(x_i, t_0, t) = \frac{S(x_i, t_0, t) - \bar{S}(x_i, t_0, t)}{\sigma} . \quad (4)$$

Depending on information from the entire study area with N regions, the previous function after normalization is a spatially non-local function. The functions in different regions are correlated with one another; in other words, any two functions might influence each other. Based on the previous derivation, a phase function can be considered a mathematical embodiment that characterizes on earthquake fault system by strongly correlated dynamics. The significant variation in seismicity at x_i between time t_1 and t_2 is

$$\Delta S(x_i, t_1, t_2) = \hat{S}(x_i, t_0, t_2) - \hat{S}(x_i, t_0, t_1) , \quad (5)$$

and its spatial mean is

$$\mu_s = \frac{1}{N} \sum_{i=1}^N |\Delta S(x_i, t_1, t_2)|^2 . \quad (6)$$

Since a correlation function can be interpreted as a probability, Tiampo et al. (2000) showed that while forming a spatial correlation at x_i over a time period, the increase in probability could be statistically obtained based on the principle of conservation of probability, and given as follows

$$\Delta P(x_i, t_1, t_2) = |\Delta S(x_i, t_1, t_2)|^2 - \mu_s . \quad (7)$$

The increase in probability, $\Delta P(x_i, t_1, t_2)$, can be used as a parameter to detect the possible locations of future larger-size earthquakes.

The time span, $t_1 - t_2$, is obtained based on the log-likelihood function (Kagan 1991), L . To compute L , we first calculated the probability, $\Delta P(x_i, t_1, t_2)$, for the i -th $M \geq 6$ earthquake by the following expression:

$$P(x_i, t_1, t_2) = \Delta P(x_i, t_1, t_2) + \mu_s . \quad (8)$$

Then we estimate the likelihood function

$$L = \sum_{i=1}^N \log [P(x_i, t_1, t_2)] , \quad (9)$$

where N is the number of events with $M \geq 6$. By fixing t_2 , t_1 can be determined from the largest value of L .

3. DATA

In this study, we use Taiwan earthquake data during the 1973-2000 period and in the region of from 21.8° N, 120° E to 25.5° N, 122° E, obtained by the Central Weather Bureau Seismic Network (CWBSN). Because of the sparse instrumental coverage before 1991 (Shin 1993), only the data classified as A and B are employed. For the present study, the completeness of earthquake data must be first taken into account. The threshold magnitude (M_c) is the minimum magnitude to show the completeness of earthquakes. The choice of M_c is based on the power-law Gutenberg-Richter relationship, and the maximum value of the first derivative of the frequency-magnitude curve is computed as a first-order approximate of M_c (Wiemer and Katsumate 1999). We estimate the value of M_c of Taiwan earthquakes for each year during 1973-1999. Results are shown in Fig. 1, which showed that the data provided by CWBSN are incomplete when $M < 2.0$. In the following, the events with $M \geq 2.0$ are used.

The selected data are divided into two subsets based on time span, the first one is called "learning catalogue" and the second one is "test catalogue". The former is used to estimate the seismicity pattern, which is compared with that of the latter. We use earthquakes during 1973-1997 to form the learning subset, and those during 1998-2000 to form the testing one. Based on the procedure mentioned above, we calculate the increasing probability in the whole region from the learning set and compare the results with earthquakes occurring after 1998.

4. RESULTS

The whole region of Taiwan is divided into 740 boxes, each with an equal area roughly of $11 \times 11 \text{ km}^2$ ($0.1^\circ \times 0.1^\circ$). Aftershocks are included in the analysis procedure for making space-time correlation effective. Unfortunately, high quality data are not enough in the eastern offshore area, so the results there may not be so good. We take one day as a time unit in constructing the seismicity rate, $\phi(x, t_0, t)$, for earthquakes occurring in Taiwan from $t_0=1973$ to $t=1997$. Results are shown in Fig. 2 for all boxes counting from the left top to the right bottom (see Fig. 3). The values of the phase function in Equation (1) for all boxes are displayed in Fig. 3.

For the learning catalogue, the year 1973 is chosen as the starting time, t_0 , in Equation (5), and then the year 1997 is selected as the second end time, i.e., t_2 in Equation (5). Unfortunately, there is no information to determine the first end time, t_1 . The choice of the time span between

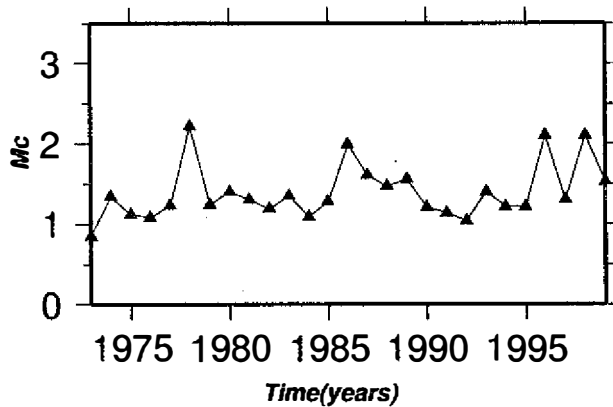


Fig. 1. The minimum magnitude of catalogue completeness (M_c) as a function of time. The aim of M_c estimation is to maximize the number of data available without compromising the accuracy of the results. The results indicate that events with magnitude greater than 2 can be adopted.

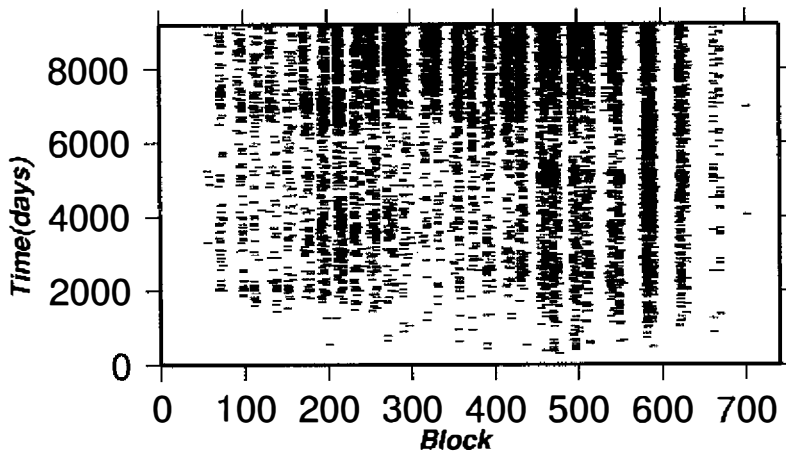


Fig. 2. Time-block plots of seismicity rate, $\phi(x_i, t_0, t)$, in the study region (see Fig. 3) from 1973 (t_0) to 1997(t). There are 740 blocks, each of $\sim 11 \times 11$ km² in size, from the left top corner along S-curve to the right bottom corner in the studied area.

t_1 and t_2 would be critical for determining the future seismicity pattern. Physically, a wider correlated region should correspond to a longer time span. To determine a proper time span for the studied area, the log-likelihood function, L , must be calculated. We compute L by first calculating the probability, $\Delta p(x_i, t_1, t_2)$, for the boxes, in which with the i -th earthquake of $M \geq 6$ that occurred in the test catalogue [see Equation (8)]. Then we estimate the likelihood function using Equation (9). In other words, we calculate the values of L for several t_1 values,

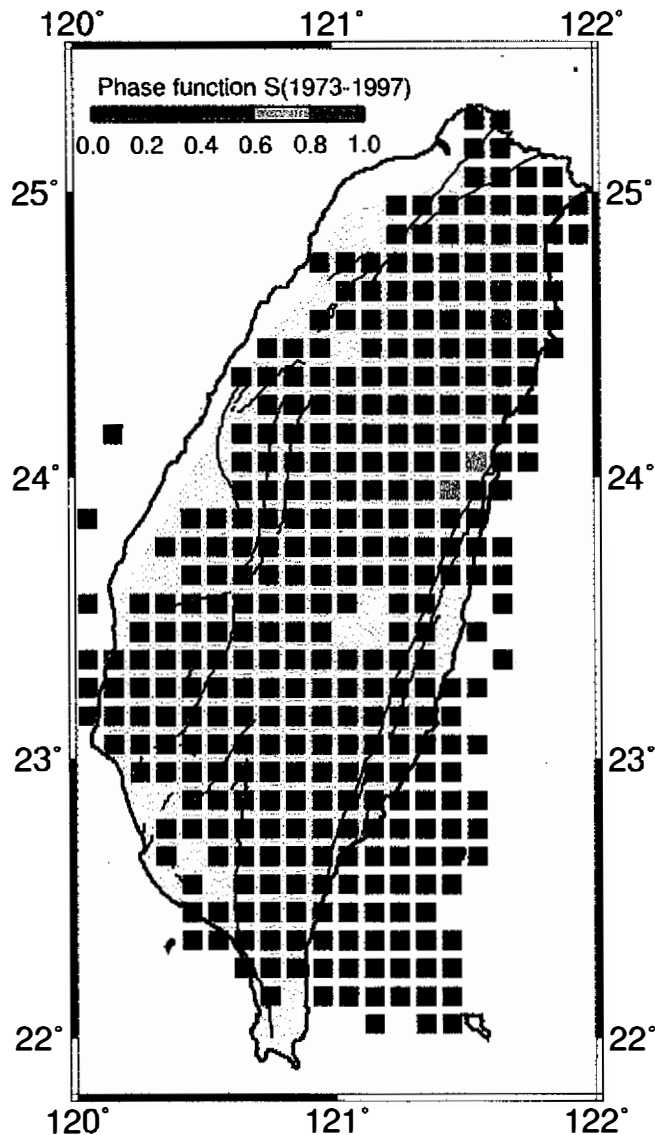


Fig. 3. The phase function, $S(x_i, t_0, t)$, of each divided box (x_i) for the time span at $t_0=1973$, $t=1997$. Not much information can be obtained from this phase function.

and then choose the t_1 value relevant to the largest value of L . Figure 4 shows the variation of L with, $(t_1 - t_2)$. Results show that the local maximum value lies between 8 to 10 years. This value shows the importance of the seismicity pattern of the past 8 to 10 years for the occurrences of future $M \geq 6$ earthquakes. However, a time span is not a constant and must be a function of the expected magnitude of the future earthquake and the box size. Theoretically, a longer expected time span is associated with a greater expected magnitude.

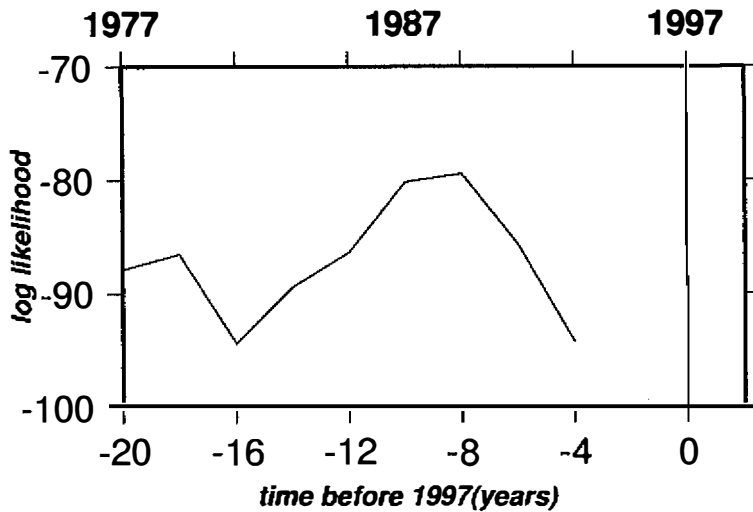


Fig. 4. A plot of log-likelihood function (L) versus time span, t_1-t_2 , where t_2 is fixed at year 1997. The optimal choice of the time span estimated from the value of L is 8 to 10 years.

We calculate the increase in probability for several seismicity patterns in different time intervals. The results for the seismicity patterns with four time spans, i.e., 5, 10, 15, 20 years, are illustrated in Fig. 5. In Fig. 5, t_2 is 1997 for four different t_1 values, i.e., 1974, 1978, 1988, and 1993. A positive value of $\Delta p(x_i, t_1, t_2)$ represents an increase in probability for future earthquakes and, thus, a box with a positive value is marked with a color. A negative value denotes a decrease in probability and is not marked.

Because there are several moderate earthquakes occurring after 1998, the data from 1998 to 2000 provide a good test catalogue. The colored boxes in Fig. 5 are scaled by a largest positive value of $\Delta p(x_i, t_1, t_2)$. The colored circles represent the earthquakes with $M \geq 6$, occurring after 1998, in the test catalogue, the blue circles denote the aftershocks that occurred in the week after the Chi-Chi earthquake, and the green, solid inverted triangles represent the events that occurred during the time span (t_1 to t_2) of calculation.

5. DISCUSSIONS

As shown in Fig. 5, in the areas with colored boxes with a positive value of $\Delta p(x_i, t_1, t_2)$, these moderate or large earthquakes occurred (including the Chi-Chi earthquake and two Chia-Yi earthquakes). This indicates that in inland Taiwan, there is a high correlation between the occurrences of larger-sized earthquakes and the areas with colored boxes. However, the correlation is low in offshore Taiwan. Of course, the probability of estimated seismicity pattern in offshore Taiwan is low, because the instrumental coverage there is poor as mentioned above. Comparing our results to those obtained in California by Tiampo et al. (2000), misfits are lower in this study. Although the time span of 8-10 years is selected from the previous analysis, comparison of the results for four time spans (as shown in Fig. 5) reveals that the

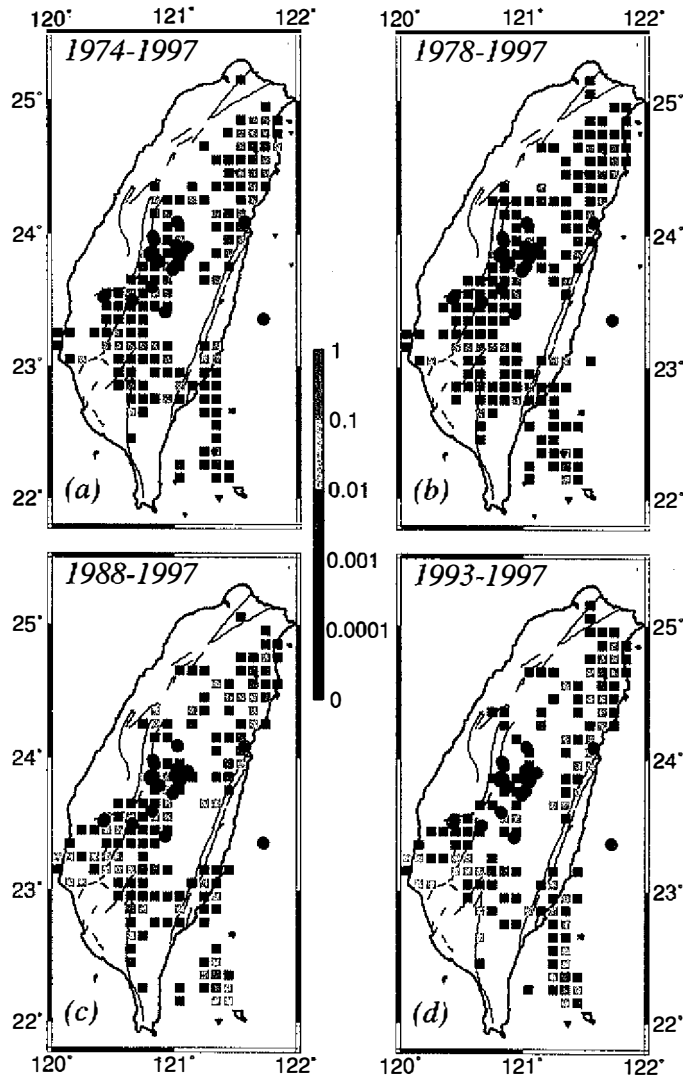


Fig. 5. Comparisons of the increase in probability, $\Delta p(x_i, t_1, t_2)$, for four time spans (i.e., t_1 to t_2) (a) 1974 - 1997 (b) 1978 - 1997 (c) 1988 - 1997 (d) 1993 - 1997. The results were calculated for data from $t_0=1973$ to $t_2=1997$ with four different t_1 values, i.e., 1974, 1978, 1988, and 1993. The probability is scaled by the largest positive value of $\Delta p(x_i, t_1, t_2)$. The increase in probability is marked with a color, while a negative value denotes a decrease in probability and the related box is not marked. The green circles represent events that occurred after 1998 with magnitude greater than 6, and the blue circle stands for the 1999 Chi-Chi event and its aftershocks in one week. The inverted triangles represent the past events ($M \geq 6$) that occurred before 1997.

seismicity patterns for the four time spans are similar, and, thus our estimation is reliable. The results shown in Fig. 5 clearly illustrate that the estimated seismicity patterns fit well with the observed data. The damaging 1999 Chi-Chi earthquake and the two Chia-Yi earthquakes are all located in the areas with positive values of increase in probability.

Earthquake occurrence is highly correlated to stress-strain accumulation. The reason why we didn't analyze this correlation in this work is that it is a nonlinear effect and could not be detected directly. The complexities of nonlinear dynamics in the earthquake fault system are generally not clear, and cannot be understood just from a deterministic point of view. The pattern dynamic analysis use a linear operator extracted from real data (seismicity). In the analysis procedure, we avoided analyzing the complicated stress-strain field. Nevertheless, the stress-strain field is included in the linear operator implicitly. Such an analysis has been widely adopted in a branch of quantum physics, so-called phenomenology.

6. CONCLUSIONS

Results indicate that the seismicity prior to a main shock dictates a non-local space-time pattern. Our results could provide a seismicity pattern for forecasting an impending earthquake, but not for predicting it. Seismicity may reflect the structures, dynamics, and properties of the underlying high-dimensional nonlinear system associated with the fault system. By performing the pattern dynamic analysis, we are capable of detecting the non-local seismicity patterns. These patterns match well with the larger-sized earthquakes shown in the test catalogue. This implies that the existence of correlated structures in the stress and strain fields can provide a very large effect on the occurrence of future larger-sized earthquakes. In this study, we only focus on the so-called class A and class B earthquakes. Most events in offshore Taiwan are classified to be class C and class D ones because of poor instrumental coverage in the area, so that the estimated seismicity pattern estimated for offshore Taiwan can not be reliable. The estimated seismicity pattern cannot contain information of the waiting time for a future large-sized earthquake and the order of future large-sized earthquakes. Relocating the recorded earthquakes and further modifying the theory will be helpful for future investigations.

Acknowledgements The constructive and thoughtful comments by two anonymous reviewers were very helpful in improving the article, particularly regarding problems of writing. We are grateful for the effort of an anonymous reviewer in modifying the English and clarifying some parts of this article. The authors also would like to express their gratitude to the Central Weather Bureau for providing available data. This research was part of the integrated project "Earthquake and Active Fault" formed by J. H. Wang and supported by the National Science Council of Taiwan under grant NSC 89-2921-M-194-001-EAF.

REFERENCES

- Bufe, C. G., and D. J., Varnes, 1993: Predictive modeling of the seismic cycle of the greater San Francisco Bay Region. *J. Geophys. Res.*, **98**, 9871-9883.
- Ferguson, C. D., W. Klein, and J. B. Rundle, 1999: Spinodals, scaling and ergodicity in a

- model of an earthquake fault with long-range stress transfer. *Phys. Rev E.*, **60**, 1359-1373.
- Frohlich, C., 1987: Aftershocks and temporal clustering of deep earthquakes. *J. Geophys. Res.*, **92**, 13944-13956.
- Geller, R. J., 1991: Shake-up for earthquake prediction. *Nature*, **352**, 275-276.
- Geller, R. J., D. D. Jackson, Y. Y. Kagan, and F. Mulargia, 1997: Earthquakes cannot be predicted. *Science*, **275**, 1616-1617.
- Gross, S., and J. B. Rundle, 1998: A systematic test of time-to-failure analysis. *Geophys. J. Int.*, **133**, 57-64.
- Kagan, Y. Y., 1991: Likelihood analysis of earthquake catalogues. *Geophys. J. Int.*, **106**, 135-148.
- Haberman, R. E., 1981: In earthquake prediction: an international review, edited by D. W. Simpson and P. G. Richards, pp. 29-42, AGU, Washington, D. C.
- King, G. C. P., R. S. Stein, and J. Lin, 1984: Static stress changes and the triggering of earthquakes. *Bull. Seis. Soc. Am.*, **84**, 935-953.
- Kisslinger, C., 1988: An experiment in earthquake prediction and the 7 May 1986 Andrean of Island earthquake. *Bull. Seism. Soc. Am.*, **78**, 218-229.
- Knopoff, L., T. Levshina, V. I. Keilis-Borok, and C. Mattoni, 1996: Increased long-term intermediate-magnitude earthquake activity prior to strong earthquakes in California. *J. Geophys. Res.*, **101**, 5779-5796.
- Mogi, K., 1969: Some features of recent seismicity in and near Japan (2). Activity before and after great earthquakes. *Bull. Earthquake Res. Inst. Univ. Tokyo*, **47**, 395-417.
- Mori, H., and Y. Kuramoto, 1998: Dissipative structures and chaos, Springer-Berlin, New York, 299pp.
- Press, F., and C. R. Allen, 1995: Patterns of seismic release in southern California region. *J. Geophys. Res.*, **100**, 6421-6430.
- Rundle, J. B., S. Gross, W. Klein, C. Ferguson, and D. L. Turcotte, 1997: The statistical mechanics of earthquakes. *Tectonophysics*, **277**, 147-164.
- Rundle, J. B., W. Klein, K. F. Tiampo, and S. Gross, 2000: Dynamics of seismicity patterns in systems of earthquake faults. *Phys. Rev. E.*, **61**, 2418-2431.
- Schwartz, D. P., and K. J. Coppersmith, 1984: Fault behavior and characteristic earthquakes: example from the Wasatch and San Andreas faults. *J. Geophys. Res.*, **89**, 5681-5698.
- Shin, T. C., 1993: The calculation of local magnitude from the simulated Wood-Anderson seismograms of the short-period seismograms in the Taiwan area. *TAO*, **4**, 155-170.
- Tiampo, K., J. B. Rundle, S. McGinnis, W. Klein, and S. J. Gross, 2000: post at <http://cires.colorado.edu/C4/news.html>
- Wiemer, S., and K. Katsumate, 1999: Spatial variability of seismicity parameters in after-shock zones. *J. Geophys. Res.*, **104**, 13135-13151.
- Wyss, M., and S. Wiemer, 1997: Two current seismic quiescences within 40 km of Tokyo. *Geophys. J. Int.*, **128**, 459-473.