From Tornadoes to Earthquakes: Forecast Verification for Binary Events Applied to the 1999 Chi-Chi, Taiwan, Earthquake

Chien-Chih Chen[1],[*], John B. Rundle [2], Hsien-Chi Li[1], James R. Holliday [2],

Kazuyoshi Z. Nanjo [3,2], Donald L. Turcotte [4] and Kristy F. Tiampo [5]

ABSTRACT

Forecast verification procedures for statistical events with binary outcomes typically rely on the use of contingency tables and Relative Operating Characteristic (ROC) diagrams. Originally developed for the statistical evaluation of tornado forecasts on a county-by-county basis, these methods can be adapted to the evaluation of competing earthquake forecasts. Here we apply these methods retrospectively to two forecasts for the M 7.3 1999 Chi-Chi, Taiwan, earthquake. We show that a previously proposed forecast method that is based on evaluating changes in seismic intensity on a regional basis is superior to a forecast based only on the magnitude of seismic intensity in the same region. Our results confirm earlier suggestions that the earthquake preparation process for events such as the Chi-Chi earthquake involves anomalous activation or quiescence, and that signatures of these processes can be detected in seismicity data using appropriate methods.

(Key words: Pattern informatics, Earthquake forecast, Chi-Chi earthquake)

---

[1] Department of Earth Sciences and Institute of Geophysics, National Central University, Chung-Li, Taiwan, ROC

[2] Center for Computational Science and Engineering, University of California, Davis, USA

[3] The Institute of Statistical Mathematics, Minato-ku, Tokyo, Japan

[4] Department of Geology, University of California, Davis, USA

[5] Department of Earth Sciences, University of Western Ontario, London, Canada

[*] *Corresponding author address:* Prof. Chien-Chih Chen, Department of Earth Sciences and Institute of Geophysics, National Central University, Chung-Li, Taiwan, ROC; E-mail: chencc@ncu.edu.tw

1. INTRODUCTION: EARTHQUAKE FORECAST AND ITS VERIFICATION

Earthquake forecasting is an emerging science under active development by a number of groups (Rundle et al. 2000a, 2002, 2003; Tiampo et al. 2002a, b; Huang et al. 2001; Keilis-Borok 2002; Wyss 1997; Wyss and Martirosyan 1998; Wyss and Wiemer 2000; Schorlemmer and Wiemer 2005; Gerstenberger et al. 2005). Researchers have studied a wide variety of proposed earthquake forecast algorithms that were developed with the goal of detecting possible signatures of future earthquakes. In many cases, the proposed anomaly is precursory seismic activation or quiescence. For example, a series of statistical algorithms (*M8*, *MSc*, *CN* etc.) for intermediate-term earthquake predictions, based primarily on seismic activation, have been developed by a Russian group under the direction of V. Keilis-Borok (Keilis-Borok 2002). Since the mid 1980s, this group has regularly issued earthquake alarms for one or more regions with diameters of up to 500 km. The alternative hypothesis, precursory seismic quiescence (Wyss 1997; Wyss and Martirosyan 1998; Wyss and Wiemer 2000), suggests that a decrease in seismicity rate might be a precursor to a forthcoming, major earthquake. An increasingly popular software package *ZMAP* is designed to detect seismic quiescence.

Another recently proposed algorithm (Rundle et al. 2000a, 2002, 2003; Tiampo et al. 2002a, b), the *pattern informatics* (PI) method, detects both seismic activation and quiescence. The physical basis of the PI approach is the hypothesis that earthquakes are the result of self-organizing cooperative behavior and strong space-time correlations arising in an interacting, driven threshold system of faults. Plate tectonic motions supply the driving stresses, the interactions arise primarily from elasticity, and nonlinear damage and friction physics lead to the failure thresholds. Application to Southern California seismicity data suggests that the PI method shows considerable promise as a technique for intermediate-term (3 - 10 years) forecasts of the locations of future large mainshocks and their aftershocks.

Also, another type of probabilistic earthquake forecast has recently been developed by the US Geological Survey (Gerstenberger et al. 2005). This method is based on applying the Omori law of aftershock decay, the Gutenberg-Richter magnitude frequency relation, and observations of peak ground acceleration following large earthquakes. Using a web-based interface, real-time forecasts of ground shaking for the next 24 hours, everywhere in California, are computed on a real-time basis following the occurrence of every earthquake having a magnitude larger than $M = 3$. In this method, the largest ground accelerations will be forecast in the 24 hours after the ground shaking associated with the largest earthquake has subsided.

A critical element of a successful forecast methodology is the establishment of forecast verification procedures. Currently, forecasts of earthquake occurrence are verified using either a statistical maximum likelihood (ML) procedure (Kagan and Jackson 2000), or by means of Molchan's error diagram (Molchan 1997).

An essential ingredient in the ML procedure is that the ML test compares the candidate forecast to a forecast from a null hypothesis. The ML test tends to reward successful forecasts and to penalize (strongly) failures-to-predict and (weakly) false alarms. The ML test is now accepted as the standard method of testing earthquake forecasts, which are most often retrospective forecasts ("retro-diction"). However one serious drawback for this procedure is that, if for some reason a large event fails in prediction, the ML test indicates complete failure of the

candidate forecast. As an example, consider a forecast that is perfect for 99 earthquakes, but which completely fails to forecast the 100th event. This forecast will score as a complete failure on the ML test, despite the clear fact that the forecast provides considerable useful information. The typical remedy for this problem is to assume a *background probability*, thus creating opportunities for the inclusion of bias into the statistical test procedures. We believe that the goal of a forecast verification process is to determine the quantity and quality of information provided by a proposed forecast method (Joliffe and Stephenson 2003).

Alternatively, a Relative Operating Characteristic (ROC) diagram (Joliffe and Stephenson 2003; Drton et al. 2003) is used to evaluate the quality of the forecasts for single earthquakes within a given time duration, a given spatial region, and a given magnitude interval. The ROC diagram considers the fraction of failures-to-predict and the fraction of false alarms. This method evaluates the performance of the forecast method relative to random chance by constructing a plot of the fraction of failures-to-predict against the fraction of false alarms (Molchan 1997; Keilis-Borok 2002) for an ensemble of forecasts. Molchan (1997) has used a modification of this method to evaluate the success of intermediate term forecasts.

The ROC diagram has a long history, over 100 years, in the verification of tornado forecasts (Joliffe and Stephenson 2003; Drton et al. 2003). These forecasts take the form of multiple tornado events during the same time interval with each event having a binary set of possible outcomes. For example, during a given time window of several hours duration, a forecast is issued in which a list of counties is given, in each of which it is stated that one or more tornadoes will or will not occur. A 2 by 2 *contingency table* is then constructed whose top row represents the counties in which tornadoes are forecast to occur, and whose bottom row represents counties in which tornadoes are forecast not to occur. Similarly, the left column represents counties in which tornadoes were actually observed to occur, and the right column represents counties in which no tornadoes were observed.

With respect to earthquakes, forecasts are being developed that take exactly this same form (Rundle et al. 2000a, 2002, 2003; Tiampo et al. 2002a, b). A time window is proposed during which the forecast of large earthquakes having a magnitude above a minimum one is considered valid. An example might be a forecast of earthquakes larger than $M = 5$ during a period of 5 or 10 years duration. A map of the seismically active region is then completely covered ("tiled") with "pixels" of two types: pixels in which the epicenters of at least one large earthquake are forecast to occur, and pixels in which large earthquakes are forecast not to occur. In other types of forecasts, large earthquakes are given some continuous probability of occurrence from 0% to 100% in each pixel, but these types of forecasts can also be converted to the binary type by application of a level value or threshold. Pixels having a probability below the threshold are then assigned a forecast rating of *non-occurrence* during the time window, and pixels having a probability above the threshold are assigned a forecast rating of *occurrence*, or *success*. A high threshold value may lead to many *failures-to-predict* (events that occur where no event is forecast), but few *false alarms* (an event is forecast at a location but no event occurs). The level at which the threshold is set is then a matter of public policy specified by emergency planners, representing a balance between the prevalence of failures-to-predict and false alarms.

Our procedure proposed in this paper provides a means of evaluating the effectiveness of

such binary forecasts of multiple earthquake occurrences during the time window. For example, we consider seismicity in the Taiwan region leading to the *M* 7.3, 21 September 1999 Chi-Chi earthquake. This event (Fig. 1) was the largest inland earthquake to occur in the Taiwan region in the 20th century. Chen (2003) investigated the accelerating activity of moderate-sized earthquakes before the Chi-Chi earthquake and pointed out that seismic activation of earthquakes
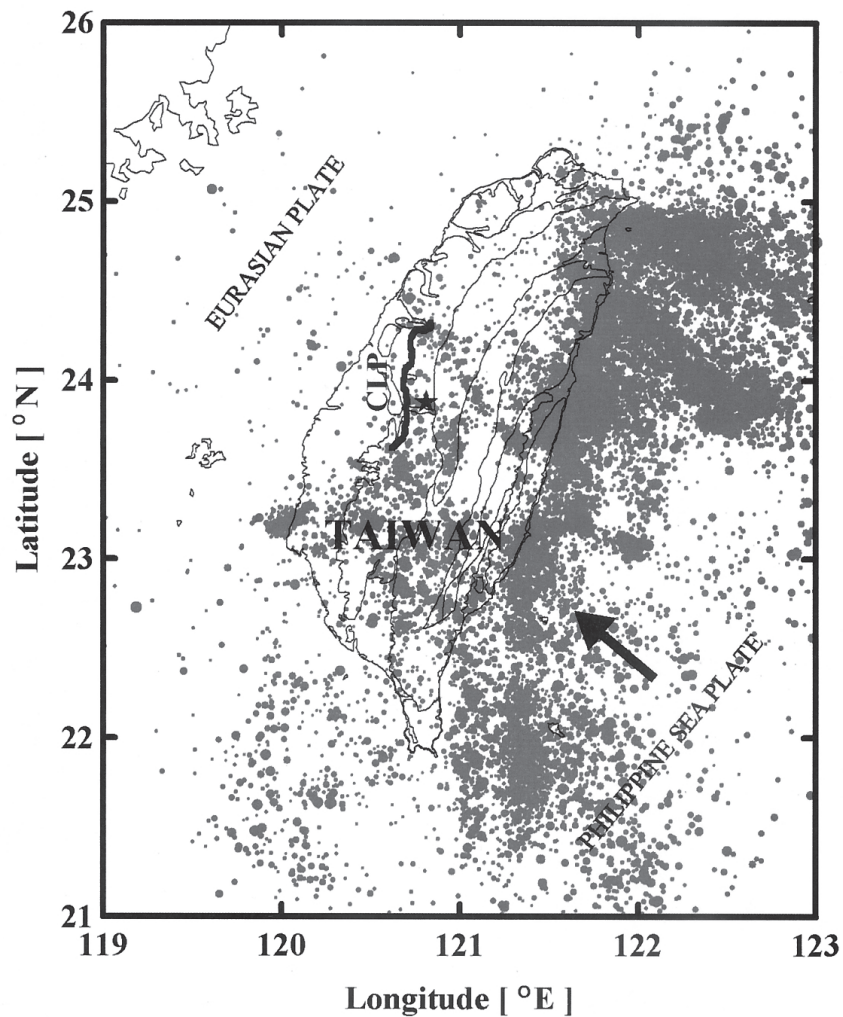

Fig. 1. Map showing the epicenters of earthquakes used in this study (dots) and the Chi-Chi mainshock (star). The legend CLP denotes the Chelungpu fault. Thick arrow indicates the direction of relative motion between the Eurasian and Philippine Sea plates.

with magnitudes larger than 5 began at the end of 1993, lasting about 6 years up to the mainshock (Fig. 3 in Chen 2003). Examination of the frequency-magnitude statistics in the years prior to the earthquake indicates that three distinct stages can be identified. The first stage represents a typical Gutenberg-Richter scaling relation. In the second stage, seismic activation of moderate-sized earthquakes ($M \geq 5$) occurs, while in the third stage, the observations indicate a hybrid of seismic quiescence for small events ($M < 5$) and activation for moderate events. In the context of the self-organizing spinodal model of earthquake fault systems (Rundle et al. 2000b), the time evolution of the frequency-magnitude distributions of earthquakes in Taiwan before the Chi-Chi mainshock represents an example of mixed seismic activation and quiescence.

2. BINARY EARTHQUAKE FORECASTS: HISTORIC INTENSITY AND PATTERN INFORMATICS

To illustrate this approach to earthquake forecast verification, we constructed two types of retrospective binary forecasts for the region of Taiwan in the years leading up to the Chi-Chi event. In both forecasts, the study region is tiled with boxes, or *pixels* of size $0.1° \times 0.1°$. The number of earthquakes with magnitude $M \geq 3.4$, the level at which the earthquake catalogue can be considered complete, in each box down to a depth of 20 km is determined over the time period from January 1987 to June 1999. We have found that the vertical scale of coarse graining must be roughly equal to the horizontal scale, i.e., $0.1°$, within a factor of 2 or so. Otherwise there may be problems with mixing events of different scales. For example, it is not so clear that a magnitude 4 event with a depth of ~100 km should be combined in the same box with other magnitude 4 events from ~10 km. A depth cutoff of 20 km is also comparable to the thickness of the seismogenic zone in the upper crust for the Taiwan region (Wang et al. 1994). The horizontal scale of $0.1°$ is related to the size of target events; here we focus on the forecast of earthquakes with $M \geq 6$. It is true that for an $M \geq 6$ event the fault length might occupy two or more pixels and most fracturing energy could be probably released at other pixels instead of just at the epicenter; for instance, the seismic energy associated with the Chi-Chi earthquake was released mainly in the northern segment of the Chelungpu fault (Wang 2004). However, it is expected that the sign for a larger earthquake should be different from a smaller earthquake. We have further found that some correlated/connective structure from the forecasting algorithm is usually a signature of a forthcoming larger earthquake than magnitude 7 (Chen et al. 2005).

We refer to the first type of forecast as a *relative intensity* (RI) forecast. The RI score for each box-volume is then computed as the total number of earthquakes in the time period divided by the value for the box-volume having the largest value. The RI values are treated as a non-normalized probability density, which is then integrated and normalized to the value 1 over the active area. A threshold value in the interval (0, 1) is then selected. Large earthquakes with $M \geq 6$ are then considered possible only in boxes having an RI value larger than the threshold. The remaining boxes with RI scores smaller than the threshold represent sites at which large earthquakes are forecast not to occur. The physical justification for this type of

forecast is that large earthquakes are considered most likely to occur at sites of high seismic activity.

A second type of forecast is computed by the PI method (Rundle et al. 2000a, 2002, 2003; Tiampo et al. 2002a, b) and its recent modifications (Chen et al. 2005). The seismically active region is again tiled with pixels of size $0.1° \times 0.1°$ and the seismicity to a depth of 20 km is determined. For the basic PI method, the PI score is computed by considering pairs of RI maps, in which the maps have been normalized to have the same statistics. An average over all such pairs of maps is taken in order to suppress the effects of random fluctuations. The result is squared and again treated as a probability density, then integrated and normalized to 1 as before. The physical justification for this procedure is that large earthquakes are considered most likely to occur at sites having both high seismic activity and a high rate of change, either an increase or decrease in activity.

In this paper, we introduce three simple modifications of the PI map, called here the PI-XM map (please refer to Appendix for details). The first modification is that only the most-active X% = 30% of the pixels having events with $M \geq 3.4$ are used in the PI analysis. The second modification is that in computing the PI-XM map, the seismicity is averaged over the pixel and the 8 surrounding pixels comprising its Moore neighborhood (Wolfram 2002). Under this more general nomenclature, the forecast map shown below would be termed a *PI-30M* map, and the original PI method would be termed a *PI-100* map. The third modification is that for de-emphasizing the influence of the very active regions, a temporal normalization of seismicity rate changes is implemented in the PI analysis.

In Fig. 2, we show forecast maps obtained from the RI score and PI-30M score for the Taiwan region during the interval from January 1987 to June 1999. For the PI-30M map, average change in seismicity rate is computed over the time period from 1 November 1993 to 30 June 1999, representing forecasts of likely locations of multiple large events over the time period from 1 July 1999 to 2005 (Rundle et al. 2000a, 2002, 2003; Tiampo et al. 2002a, b). Threshold values of 0.292 in the RI map and 0.492 in the PI-30M map are chosen, which generate the same numbers of *hotspot* pixels in both maps. With these values, it can be seen from the PI-30M map that the epicentral area of the Chi-Chi mainshock is located in a connected cluster of hotspot pixels, while no such signature appears on the RI map. The occurrence of the Chi-Chi mainshock, and the location of its epicenter, was unanticipated by seismologists in Taiwan, since large earthquakes had occurred for the most part in eastern Taiwan and offshore. Central Taiwan had been relatively inactive with respect to large earthquakes, inasmuch as the area had not experienced a large destructive event since 1935, when the *M* 7.1 Hsinchu-Taichung earthquake (24.4°N, 120.8°E) occurred.

Also, one interesting feature on the PI-30M map (Fig. 2B) is the appearance of numerous hotspots just around the southeastern coast. An *M* 6.4 earthquake occurred exactly in this cluster of hotspots on 1 April 2006 (http://www.cwb.gov.tw). It thus could be scored as one successful prediction in the earthquake location since this paper was submitted in March 2006. On the other hand, although the RI map seems able to successfully forecast some events occurring in the offshore area in eastern Taiwan, at this moment, we hold a conservative viewpoint for forecasting earthquakes in the offshore area. This is due to a higher error of locating an earthquake offshore than inland.
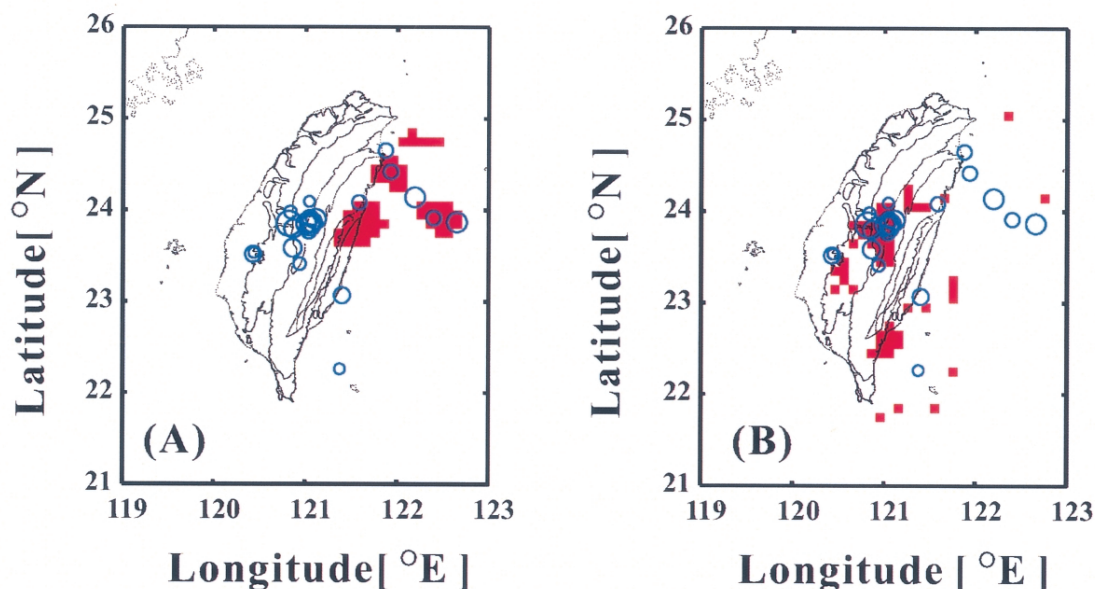
Chen et al.

Fig. 2. (A) RI and (B) PI-30M maps for the Taiwan region during the interval January 1987 to June 1999. For the PI-30M map, t_0 = 1 January 1987, t_1 = 1 November 1993, t_2 = 30 June 1999 (please refer to Appendix for details). Red pixels represent 0.1° × 0.1° boxes with RI values larger than 0.292, and with PI-30M values larger than 0.492, respectively. Thus, the numbers of red pixels in both RI and PI-30M maps are 54. Large earthquakes with $M \geq 6.0$ since 30 June 1999 are denoted by blue circles.

3. FORECAST VERIFICATION: CONTINGENCY TABLE AND ROC DIAGRAM

The first step in our analysis is the construction of 2 by 2 contingency tables for the PI-30M and RI forecast maps. The hotspot pixels in each map represent the forecast locations. A hotspot pixel upon which *at least* one large future earthquake during the forecast period occurs is counted as a *successful forecast*. A hotspot pixel upon which *no* large future earthquake occurs during the forecast period is counted as an *unsuccessful forecast*, or alternately, a *false alarm*. A white pixel upon which *at least* one large future earthquake during the forecast period occurs is counted as a *failure-to-forecast*. A white pixel upon which no large future earthquake occurs during the forecast period is counted as a *successful forecast of non-occurrence*.

Verification of the PI-30M and RI forecasts proceeds in exactly the same way as for tornado forecasts. For a given number of hotspot pixels, which is controlled by the value of probability threshold in each map, a contingency table (Table 1) is constructed for both PI-30M and RI maps. Values for the table elements *a* (Forecast = yes, Observed = yes),

Table 1. Contingency table corresponding to the RI and PI-30M forecast maps shown in Fig. 2. Each pixel is categorized as to whether it represents a hotspot pixel ($a + b$) or a white pixel ($c + d$). An actual earthquake that occurs on a hotspot pixel is a successful forecast of occurrence (a). No earthquake occurring on a white pixel represents a successful forecast of non-occurrence (d). An actual earthquake that occurs on a white pixel is a failure-to-predict (c). A hotspot pixel on which no earthquake occurs is a false alarm (b).

Observed ⟍ Forecast	Yes		No	
Yes	(a)	4 (RI); 7 (PI-30M)	(b)	50 (RI); 47 (PI-30M)
No	(c)	15 (RI); 12 (PI-30M)	(d)	1755 (RI); 1758 (PI-30M)

b (Forecast = yes, Observed = no), c (Forecast = no, Observed = yes), d (Forecast = no, Observed = no) are obtained for each map. The *hit rate* is $H = a / (a + c)$, which is the fraction of large earthquakes that occur at a hotspot. The *false alarm rate* is $F = b / (b + d)$, which is the fraction of non-observed earthquakes that are incorrectly forecast.

To analyze the information in the PI-30M and RI maps relative to the benchmark forecast, which is typically considered to be a random forecast RAN, standard procedure (Joliffe and Stephenson 2003) is to consider all possible forecasts together, which are obtained by increasing F from 0 (corresponding to no hotspots on the map) to 1 (all active pixels on the map identified as hotspots). Results are displayed in Fig. 3, which shows an ROC diagram, a plot of H against F for the PI-30M map (red line) and RI map (blue line). The diagonal black line, i.e., $H = F$, is the corresponding random forecast RAN. In the ROC diagram, higher skill of the forecast algorithm is indicated by a larger area between the ROC curve of the forecast algorithm and the diagonal line of the random forecast. For a perfect forecast this area is equal to 0.5, whereas for a forecast with no skill this area approaches to 0 (Joliffe and Stephenson 2003). It can be seen that the PI-30M map outperforms the RI map at all values of F.

A means of quantifying the relative performance of the test forecast $S_T(F)$ relative to a benchmark forecast $S_B(F)$ as the false alarm rate F is varied is the *skill score* $S_{T,B}(F)$ (Joliffe and Stephenson 2003):

$$S_{T,B}(F) \equiv \frac{H_T(F) - H_B(F)}{1 - H_B(F)} \quad . \tag{1}$$

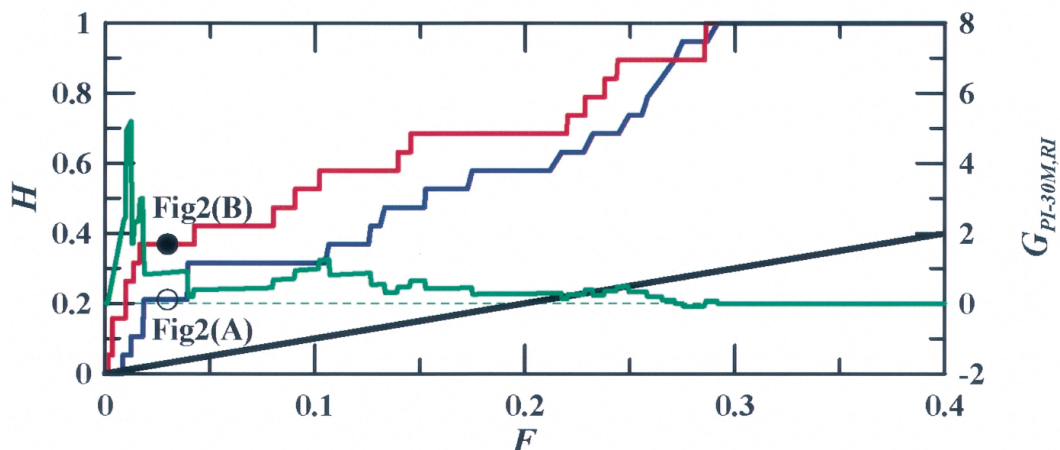Skill scores for the PI-30M map relative to the RAN map, $S_{PI-30M,RAN}(F)$, and for the RI

Fig. 3. Relative Operating Characteristic (ROC) curves for PI-30M (red line, left scale), RI (blue line, left scale) and RAN (black line, left scale). Also shown in this plot is the gain function $G_{PI-30M,RI}$ in forecast skill (green line, right scale) of PI-30M map relative to the RI map. The ROC plot for the RI and PI-30M maps is the fraction H of large earthquakes successfully forecast as a function of false alarm rate F. $G_{PI-30M,RI}(F)$ is computed by equations (1) and (2). The solid and open circles correspond to the maps shown in Fig. 2, as summarized in Table 1.

map relative to the RAN map, $S_{RI,RAN}(F)$, were computed. We then defined the gain function in forecast skill $G_{PI-30M,RI}(F)$ of the PI-30M map relative to the RI map as:

$$G_{PI-30M,RI}(F) \equiv \frac{S_{PI-30M,RAN}(F)}{S_{RI,RAN}(F)} - 1 \quad . \tag{2}$$

In Fig. 3, the green line represents $G_{PI-30M,RI}(F)$. It can be seen that $G_{PI-30M,RI}(F)$ can be larger than 5. The average gain $\langle G_{PI-30M,RI} \rangle$ over the interval from $F = 0$ to $F = 0.29$ is 53%. The analysis thus indicates that the PI-30M map is substantially better at locating the sites of future large earthquakes than the RI map.

4. DISCUSSION: THE ROLE OF FLUCTUATIONS IN SEISMICITY RATES

It is known from many laboratory observations that earthquake occurrence is associated with a high level of stress on the fault (Wiemer and Wyss 1994; Hainzl et al. 2000; Rundle et al. 2000b; Huang et al. 2001; Chen 2003; Keilis-Borok 2002; Wyss 1997; Wyss and Martirosyan

1998; Wyss and Wiemer 2000). Ideas relating to anomalous changes in seismicity rates prior to a major earthquake have been frequently proposed and widely discussed in the literature, taking the form of either precursory quiescence or precursory activation (Wiemer and Wyss 1994; Hainzl et al. 2000; Rundle et al. 2000b; Huang et al. 2001; Chen 2003; Wyss 1997; Wyss and Martirosyan 1998; Wyss and Wiemer 2000; Wyss and Habermann 1988; Jaume and Sykes 1999). In addition to this study, the phenomenon of precursory seismic activation and/ or quiescence before the Chi-Chi earthquake was also examined from the twin perspectives of both time scales by Chen (2003) and spatial scales by Chen et al. (2005).

These observations have led to the question of whether increasing stress levels prior to major earthquakes are associated more with the intensity of seismic activity, or alternatively with changes in seismicity rate. Changes in seismicity rate are essential to the earthquake population, expressed by the well-known Omori's law for the decay of aftershock rate. More recently, it was also found that changes in the rate of foreshock activity may be fitted with a function of that form, the *inverse* Omori's law (Jones and Molnar 1979; Shaw 1993; Ziv 2003).

It has been proposed that a correlation exists between stress level in the Earth's crust and time rate of change of seismic activity (Fig. 3 in Stein 1999; Fig. 5 in Toda et al. 2002), rather than between stress level and the intensity of seismic activity. Our results, that the PI-30M map is a substantially better indicator of future earthquake locations than is the RI map, provide strong support for the hypothesis that earthquake occurrence is associated with changes in activity rather than with the intensity of activity.

REFERENCES

Chen, C. C., 2003: Accelerating seismicity of moderate-size earthquakes before the 1999 Chi-Chi, Taiwan, earthquake: Testing time-prediction of the self-organizing spinodal model of earthquake. *Geophys. J. Int.*, **155**, F1-F5.

Chen, C. C., J. B. Rundle, J. R. Holliday, K. Z. Nanjo, D. L.Turcotte, S. C. Li, and K. F. Tiampo, 2005: The 1999 Chi-Chi, Taiwan, earthquake as a typical example of seismic activation and quiescence. *Geophys. Res. Letts.*, **32**, L22315, doi: 10.1029/2005GL023991.

Drton, M., C. Marzban, P. Guttorp, and J. T. Schaefer, 2003: A Markov chain model of tornadic activity. *Mon. Wea. Rev.*, **131**, 2941-2953.

Gerstenberger, M. C., S. Wiemer, L. M. Jones, and P. A. Reasenberg, 2005: Real-time forecasts of tomorrow's earthquakes in California. *Nature*, **435**, 328-331.

Hainzl, S., G. Zoller, J. Kurths, and J. Zschau, 2000: Seismic quiescence as an indicator for large earthquakes in a system of self-organized criticality. *Geophys. Res. Lett.*, **27**, 597-600.

Huang, Q., G. A. Sobolev, and T. Nagao, 2001: Characteristics of the seismic quiescence and activation patterns before the M = 7.2 Kobe earthquake, January 17, 1995. *Tectonophys.*, **337**, 99-116.

Jaume, S. C., and L. R. Sykes, 1999: Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes. *Pure Appl. Geophys.*, **155**, 279-306.

Joliffe, I. T., and D. B. Stephenson, 2003: Forecast Verification: A Practitioner's Guide in Atmospheric Science. John Wiley & Sons, England.

Jones, L. M., and P. Molnar, 1979: Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults. *J. Geophys. Res.*, **72**, 3596-3608.

Kagan, Y. Y., and D. D. Jackson, 2000: Probabilistic forecasting of earthquakes. *Geophys. J. Int.*, **143**, 438-453.

Keilis-Borok, V., 2002: Earthquake prediction State-of-the-art and emerging possibilities. *Annu. Rev. Earth Planet. Sci.*, **30**, 1-33.

Molchan, G. M., 1997: Earthquake prediction as a decision-making problem. *Pure Appl. Geophys.*, **149**, 233-247.

Rundle, J. B., W. Klein, K. F. Tiampo, and S. J. Gross, 2000a: Linear pattern dynamics in nonlinear threshold systems. *Phys. Rev. E*, **61**, 2418-2431.

Rundle, J. B., W. Klein, D. L. Turcotte, and B. D. Malamud, 2000b: Precursory seismic activation and critical-point phenomena. *Pure Appl. Geophys.*, **157**, 2165-2182.

Rundle, J. B., K. F. Tiampo, W. Klein, and J. S. S. Martins, 2002: Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 2514-2521.

Rundle, J. B., D. L. Turcotte, R. Shcherbakov, W. Klein, and C. Sammis, 2003: Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev. Geophys.*, **41**, 1019, doi:10.1029/2003RG000135.

Schorlemmer, D., and S. Wiemer, 2005: Microseismicity data forecast rupture area. *Nature*, **434**, 1086.

Shaw, B. E., 1993: Generalized Omori law for foreshocks and aftershocks from a simple dynamics. *Geophys. Res. Lett.*, **20**, 907-910.

Stein, R. S., 1999: The role of stress transfer in earthquake occurrence. *Nature*, **402**, 605-609.

Tiampo, K. F., J. B. Rundle, S. McGinnis, S. J. Gross, and W. Klein, 2002a: Mean-field threshold systems and phase dynamics: An application to earthquake fault systems. *Europhys. Lett.*, **60**, 481-487.

Tiampo, K. F., J. B. Rundle, S. McGinnis, and W. Klein, 2002b: Pattern dynamics and forecast methods in seismically active regions. *Pure Appl. Geophys.*, **159**, 2429-2467.

Toda, S., R. S. Stein, and T. Sagiya, 2002: Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nat*ure, **419**, 58-61.

Wang, J. H., K. C. Chen, and T. Q. Lee, 1994. Depth distribution of shallow earthquakes in Taiwan. *J. Geol. Soc. China*, **37**, 125-142.

Wang, J. H., 2004: The seismic efficiency of the 1999 Chi-Chi, Taiwan, earthquake. *Geophys. Res. Letts.*, **31**, L10613, doi: 10.1029/2004GL019417.

Wiemer, S., and M. Wyss, 1994: Seismic quiescence before the Landers (M = 7.5) and Big Bear (M = 6.5) 1992 earthquakes. *Bull. Seism. Soc. Am.*, **84**, 900-916.

Wolfram, S., 2002: A New Kind of Science. Wolfram Media Inc., Champaign, IL.

Wyss, M., 1997: Nomination of precursory seismic quiescence as a significant precursor. *Pure Appl. Geophys.*, **149**, 79-114.

Wyss, M., and R. E. Habermann, 1988: Precursory seismic quiescence. *Pure Appl. Geophys.*, **126**, 319-332.

Wyss, M., and A. H. Martirosyan, 1998: Seismic quiescence before the M 7, 1988, Spitak earthquake, Armenia. *Geophys. J. Int.*, **134**, 329-340.

Wyss, M., and S. Wiemer, 2000: Change in the probability for earthquakes in southern California due to the Landers magnitude 7.3 Earthquake. *Science*, **290**, 1334-1338.

Ziv, A., 2003: Foreshocks, aftershocks, and remote triggering in quasi-static fault models. *J. Geophys. Res.*, **108**, doi:10.1029/2002JB002318.

APPENDIX: NEW MODIFIED PI METHOD

The detailed procedures of a new modified PI method are described in the following:

1. The region of interest is divided into a grid of N_B square boxes with linear dimension Δx.

2. All earthquakes in the catalog to be used are included, consistent with the requirement of catalog completeness. The latter implies that a lower cutoff magnitude M_c must be used to ensure statistical uniformity of the data.

3. Four times are defined: i) t_0: the time at which the calculation begins, data before t_0 would not be used in PI calculation. ii) t_1: the beginning of the time interval over which changes will be computed. iii) t_2: the end of the time interval over which changes will be computed. iv) t_b: a time shifted between t_0 and t_1 that will be used for normalizing to eliminate the strong effect due to a large number of bursting aftershocks occurring at some sites during t_1 to t_2 and averaging to reduce random fluctuations due to noise as well.

4. In any given region of N_B total boxes, a subset $N_A \leq N_B$ of the boxes will typically contain at least one earthquake with $M \geq M_c$ during the time interval t_0 to t_2. These N_A boxes comprise the "active boxes". It is also possible that for de-noising the PI map, only the most-active X%, say 30%, of the boxes having events with $M \geq M_c$ are considered in the next steps.

Chen et al.

5. The seismic rate function $I(t_b, t)$ of a box is defined as the average number of earthquakes with magnitudes greater than M_c that occur in the box centered at x_i and its Moore neighbor boxes per unit time during a specified time interval $[t_b, t]$, $t = t_1$ or t_2. Consequently, a seismic rate change in $I(t_b, t)$ at x_i over the change interval $[t_1, t_2]$ is computed, $\Delta I(x_i, t_b, t_1, t_2) = I(x_i, t_b, t_2) - I(x_i, t_b, t_1)$. Simply $\Delta I(x_i, t_b)$ may denote $\Delta I(x_i, t_b, t_1, t_2)$, since t_1 and t_2 are fixed throughout the PI calculation.

6. At a given site x_i, the temporal average $\overline{\Delta I(x_i)}$ of function $\Delta I(x_i, t_b)$ and its standard deviation $\sigma_{\Delta I}(x_i)$ are taken over all t_b:

$$\overline{\Delta I(x_i)} \equiv \frac{1}{n(t_b)} \sum_{k=1}^{n(t_b)} \Delta I(x_i, t_{b_k}) \quad , \tag{A1}$$

$$\sigma_{\Delta I}(x_i) \equiv \frac{\sqrt{\sum_{k=1}^{n(t_b)} [\Delta I(x_i, t_{b_k}) - \overline{\Delta I(x_i)}]^2}}{\sqrt{n(t_b)}} \quad , \tag{A2}$$

where $n(t_b)$ is the number of t_b. A "temporally" normalized rate change function $\Delta I_T(x_i, t_b)$ at a given site x_i is then the difference between the seismic rate change $\Delta I(x_i, t_b)$ at x_i and the temporal mean $\overline{\Delta I(x_i)}$, divided by its temporal standard deviation $\sigma_{\Delta I}(x_i)$. This temporal normalization is found to be very useful for eliminating the strong effect due to a large number of bursting aftershocks occurring at some sites during t_1 to t_2.

7. Similarly, at a given t_b, the spatial average $\overline{\Delta I_T(t_b)}$ of function $\Delta I_T(x_i, t_b)$ and its standard deviation $\sigma_{\Delta I_T}(t_b)$ are taken over all N_B boxes:

$$\overline{\Delta I_T(t)} \equiv \frac{1}{N_A} \sum_{i=1}^{N_A} \Delta I_T(x_i, t_b), \tag{A3}$$

$$\sigma_{\Delta I_T}(t) \equiv \frac{\sqrt{\sum_{i=1}^{N_A} [\Delta I_T(x_i, t_b) - \overline{\Delta I_T(t_b)}]^2}}{\sqrt{N_A}} \quad . \tag{A4}$$

A "spatially" normalized function $\Delta I_{TS}(x_i, t_b)$ is the difference between $\Delta I_T(x_i, t_b)$ and the spatial mean $\overline{\Delta I_T(t)}$, divided by its spatial standard deviation $\sigma_{\Delta I_T}(t)$.

8. By definition, a precursor is a change in a time dependent quantity, an increase (positive $\Delta I_{TS}(x_i, t_b)$) or a decrease (negative $\Delta I_{TS}(x_i, t_b)$) from a baseline value. Since a preparing process of a great earthquake may experience seismic quiescence and seismic activation, the absolute value of $\Delta I_{TS}(x_i, t_b)$ is therefore taken by calculating the average of the spatiotemporally normalized rate change function $\Delta I_{TS}(x_i, t_b)$, in order to reduce fluctuations

due to random noise. $<\left|\Delta I_{TS}(x_i)\right|>$ is computed by averaging $\left|\Delta I_{TS}(x_i, t_b)\right|$ over all possible values of t_b from t_0 to t_1.

9. Finally, the probability of an event occurring within a box $P(x_i)$ is defined as the square of $<\left|\Delta I_{TS}(x_i)\right|>$, i.e., $P(x_i) \equiv <\left|\Delta I_{TS}(x_i)\right|>^2$. They also defined the mean μ_P of $P(x_i)$ over all boxes. μ_P can then be considered to be the average background probability for a large event. In the PI maps, the change in probability $\Delta P(x_i) \equiv <\left|\Delta I_{TS}(x_i)\right|>^2 - \mu_P$ was then color coded and plotted.