

## Retrospective Study on the Predictability of Pattern Informatics to the Wenchuan M8.0 and Yutian M7.3 Earthquakes

YONGXIAN ZHANG,<sup>1,2</sup> XIAOTAO ZHANG,<sup>1</sup> YONGJIA WU,<sup>1</sup> and XIANGCHU YIN<sup>2,3</sup>

**Abstract**—Two large earthquakes occurred in the western part of China in 2008, one of them being the Yutian (35.6°N, 81.6°E) M7.3 earthquake that occurred on March 21 (BJT) and the other the Wenchuan (31.0°N, 103.4°E) M8.0 earthquake that occurred on May 12 (BJT). In this paper, the West Continental China (included in 20.0°–50.0°N, 70.0°–110.0°E region) was the study region for verifying the predictability of the pattern informatics (PI) method using the receiver-operating characteristic curve (ROC) test and *R* score test. Different forecasting maps with different calculating parameters were obtained. The calculating parameters were the grid size  $\Delta x$ , base time  $t_b$ , reference interval  $t_b$  to  $t_1$ , change interval  $t_1$  to  $t_2$ , and forecasting interval  $t_2$  to  $t_3$ . In this paper, the base time  $t_b$  fixed to June 1, 1971, the ending forecast time  $t_3$  fixed to June 1, 2008, and the forecasting interval  $t_2$  to  $t_3$  changed from 1 to 10 years, and the grid sizes were chosen as  $1^\circ \times 1^\circ$  and  $2^\circ \times 2^\circ$ , respectively. The results show that the PI method could forecast the Yutian M7.3 and Wenchuan M8.0 earthquakes only using suitable parameters. Comparing the forecast results of grid sizes  $1^\circ \times 1^\circ$  and  $2^\circ \times 2^\circ$ , the models with  $2^\circ \times 2^\circ$  grids were better. Comparing the forecast results with different forecasting windows from 1 to 10 years, the models with forecasting windows of 4–8 years were better using the ROC test, and the models with forecasting windows of 7–10 years were better using the *R* score test. The forecast efficiency of the model with a grid size of  $2^\circ \times 2^\circ$  and forecast window of 8 years was the best one using either the ROC test or the *R* score test.

**Key words:** PI method, retrospective study, earthquake predictability, Wenchuan M8.0 earthquake, Yutian M7.3 earthquake, ROC test, *R* score test.

### 1. Introduction

After the M8.1 earthquake of West Kunlun Mountain Pass (36.21°N, 90.9°E) on 14 November

2001 (BJT), two large earthquakes occurred 6 years later in West Continental China, one being the Yutian (35.6°N, 81.6°E) M7.3 earthquake (according to CENC) that occurred on 21 March 2008 (BJT) and the other being the Wenchuan (31.0°N, 103.4°E) M8.0 earthquake (according to CENC) that occurred on 12 May 2008 (BJT). As a new approach to earthquake forecasting on a 10-year scale, the pattern informatics (PI) method (RUNDLE *et al.* 2000a, b, 2002, 2003; TIAMPO *et al.* 2002a, b; HOLLIDAY *et al.* 2005, 2006a) has been applied to the forecasting and retrospective verification in the California region (RUNDLE *et al.* 2002; TIAMPO *et al.* 2002b; HOLLIDAY *et al.* 2005), Japan region (NANJO *et al.* 2006a, b), Taiwan region (CHEN *et al.* 2005, 2006; WU *et al.* 2008a, b), China mainland (JIANG and WU 2008; ZHANG *et al.* 2009), and worldwide (HOLLIDAY *et al.* 2005). From the above retrospective studies, most earthquakes did occur in the hotspot boxes or in the Moore neighborhood boxes, and the results of the ROC test show that the PI method outperforms not only random guesses, but also the simple number counting approach based on the clustering hypothesis of earthquakes (the RI forecast).

For the above studies, Holiday's results (HOLLIDAY *et al.* 2005) are involved in verification of the effects of the PI method for the Yutian M7.3 and Wenchuan M8.0 earthquakes. In his study, a retrospective and forecasting map was published, which showed the hotspots for potential  $M \geq 7.0$  earthquakes during the 2000–2010 period (Fig. 1). According to the retrospective study, during the period from 1 January 2000 to 14 December 2004 there were 68 M7 earthquakes worldwide; 57 of these earthquakes occurred within a hotspot or adjoining boxes. Subsequent to his meeting presentation, the  $M = 8.1$  Macquarie Island earthquake occurred on 23 December 2004, and the  $M = 9.0$  Sumatra earthquake

<sup>1</sup> China Earthquake Networks Center, Beijing 100045, China. E-mail: yxzhseis@sina.com; qdxt\_116@163.com; wuyongj@mails.gucas.ac.cn;

<sup>2</sup> LNM, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100080, China. E-mail: xcyin@public.bta.net.cn

<sup>3</sup> Institute of Earthquake Science, China Earthquake Administration, Beijing 100036, China.

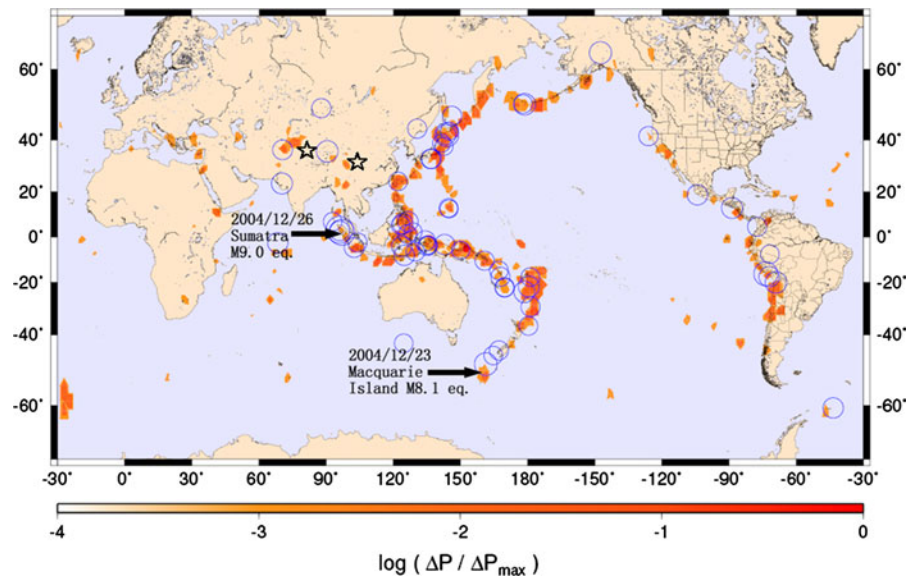


Figure 1

Worldwide application of the PI method (HOLLIDAY *et al.* 2005). Colored areas are the forecast hotspots for the occurrence of  $M \geq 7$  earthquakes during the 2000–2010 period derived using the PI method. The color scale gives values of the  $\log_{10}(PI/P_{\max})$ . Blue circles are the locations of earthquakes with  $M \geq 7$  that have occurred since January 2000 to August 2005. The five-pointed stars denote the Yutian M7.3 and Wenchuan M8.0 earthquakes

occurred on 26 December 2004. The epicenters of both earthquakes were successfully forecasted (Fig. 1). On the other hand, Fig. 1 also allows the chance for verifying for the predictability of the PI method for  $M \geq 7.0$  earthquakes occurring successively, including the Yutian M7.3 and Wenchuan 8.0 earthquakes in 2008 (the five-pointed stars in Fig. 1). From Fig. 1, we can see that the Yutian M7.3 and Wenchuan 8.0 earthquakes did not occur in the hotspot boxes or in the neighborhood boxes.

Although the failure forecast to these two earthquakes did not affect the effects of the PI method for most large earthquakes, the reasons causing the failure are the focus of this paper. The predictability of the PI method will be verified by changing the calculated parameters and testing with the receiver-operating characteristic (ROC) diagram and  $R$  score test.

## 2. The PI Method

The pattern informatics method is based on the statistical mechanics of complex systems and can quantify temporal variations in seismicity. The result is a map of areas in a seismogenic region (hotspots) where earthquakes are likely to occur during a

specified period in the future. With the PI method, a forecast map of locations (hotspots) of future earthquakes during a forthcoming period could be worked out. RUNDLE *et al.* (2002) published a forecast map of hotspots of  $M > 5$  earthquakes for California in the 2000–2010 period (<http://www.quakesim.jpl.nasa.gov/scorecard.html>). Subsequently, 19 of the 20 California earthquakes with magnitudes  $M > 5$  occurred in or immediately adjacent to the resulting hotspots up to February 2008, while the areas of the hotspots only covered 4% of the map area (Performance Analysis of Earthquake Forecasts, presentation of RUNDLES *et al.* at the 6th ACES International Workshop in Cairns, Australia, 11–16 May 2008). NANJO *et al.* (2006a, b) modified the PI method for use with the Japanese catalogs and successfully forecast the 23 October 2004  $M = 6.8$  Niigata earthquake. CHEN *et al.* (2005) modified the PI method for use with Chinese Taiwan catalogs and found the Chi-Chi Ms7.6 earthquake located in the hotspot area. The above studies show that the PI method is prospective in medium- to long-term earthquake forecasts, especially for the location.

Following the detail steps of the PI method described by HOLLIDAY *et al.* (2005), we developed codes in the Fortran language that can obtain hotspot

maps with the PI method in the region of interest. The detailed utilization of the PI method for earthquake forecasting is as follows (HOLLIDAY 2005):

- (1) The region of interest is divided into  $N_B$  square boxes with linear dimension  $\Delta x$ . Boxes are identified by a subscript  $i$  and are centered at  $x_i$ . For each box, there is a time series  $N_i(t)$ , which is the number of earthquakes per unit time at time  $t$  larger than the lower cutoff magnitude  $M_c$ . The time series in box  $i$  is defined between a base time  $t_b$  and the present time  $t$ .
- (2) All earthquakes in the region of interest with magnitudes greater than a lower cutoff magnitude  $M_c$  are included. The lower cutoff magnitude  $M_c$  is specified in order to ensure completeness of the data through time, from an initial time  $t_0$  to a final time  $t_2$ .
- (3) Three time intervals are considered:
  - (a) A reference time interval from  $t_b$  to  $t_1$ .
  - (b) A second time interval from  $t_b$  to  $t_2$ ,  $t_2 > t_1$ . The change interval over which seismic activity changes are determined is then  $t_2 - t_1$ . The time  $t_b$  is chosen to lie between  $t_0$  and  $t_1$ . The objective is to quantify anomalous seismic activity in the change interval  $t_1$  to  $t_2$  relative to the reference interval  $t_b$  to  $t_1$ .
  - (c) The forecast time interval  $t_2$  to  $t_3$ , for which the forecast is valid. We take the change and forecast intervals to have the same length.
- (4) The seismic intensity in box  $i$ ,  $I_i(t_b, t)$ , between two times  $t_b < t$ , can then be defined as the average number of earthquakes with magnitudes greater than  $M_c$  that occur in the box per unit time during the specified time interval  $t_b$  to  $t$ . Therefore, using discrete notation, we can write:

$$I_i(t_b, t) = \frac{1}{t - t_b} \sum_{t'=t_b}^t N_i(t') \quad (1)$$

where the sum is performed over increments of the time series, say days.

- (5) In order to compare the intensities from different time intervals, we require that they have the same statistical properties. We therefore normalize the seismic intensities by subtracting the mean seismic activity of all boxes and divided by the

standard deviation of the seismic activity in all boxes. The statistically normalized seismic intensity of box  $i$  during the time interval  $t_b$  to  $t$  is then defined by

$$\hat{I}_i(t_b, t) = \frac{I_i(t_b, t) - \langle I_i(t_b, t) \rangle}{\sigma(t_b, t)} \quad (2)$$

where  $\langle I_i(t_b, t) \rangle$  is the mean intensity averaged over all the boxes, and  $\sigma(t_b, t)$  is the standard deviation of intensity over all the boxes.

- (6) Our measure of anomalous seismicity in box  $i$  is the difference between the two normalized seismic intensities:

$$\Delta I_i(t_b, t_1, t_2) = \hat{I}_i(t_b, t_2) - \hat{I}_i(t_b, t_1) \quad (3)$$

- (7) To reduce the relative importance of random fluctuations (noise) in seismic activity, we compute the average change in intensity,  $\overline{\Delta I_i(t_0, t_1, t_2)}$  over all possible pairs of normalized intensity maps having the same change interval:

$$\overline{\Delta I_i(t_0, t_1, t_2)} = \frac{1}{t_1 - t_0} \sum_{t_b=t_0}^{t_1} \Delta I_i(t_b, t_1, t_2) \quad (4)$$

where the sum is performed over increments of the time series, which here are days.

- (8) We define the probability of a future earthquake in box  $i$ ,  $P_i(t_0, t_1, t_2)$ , as the square of the average intensity change:

$$P_i(t_0, t_1, t_2) = \overline{\Delta I_i(t_b, t_1, t_2)}^2 \quad (5)$$

- (9) To identify anomalous regions, we wish to compute the change in the probability  $P_i(t_0, t_1, t_2)$ , relative to the background so that we subtract the mean probability over all boxes. We denote this change in the probability by

$$\Delta P_i(t_0, t_1, t_2) = P_i(t_0, t_1, t_2) - \langle P_i(t_0, t_1, t_2) \rangle \quad (6)$$

where  $\langle P_i(t_0, t_1, t_2) \rangle$  is the background probability for a large earthquake.

Hotspots are defined to be the regions where  $\Delta P_i(t_0, t_1, t_2)$  is positive. In these regions,  $P_i(t_0, t_1, t_2)$  is larger than the average value for all boxes (the background level).

Note that since the intensities are squared in defining probabilities the hotspots may be due to either increases of seismic activity during the change

time interval (activation) or due to decreases (quiescence). According to BUFE and VARNES (1993), before a major earthquake ( $M_f$ ) occurs, the seismicity of smaller earthquakes with magnitude  $M_f-2$  often show abnormalities like an acceleration or decrease; JAUME and SYKES (1999) concluded that the seismicity of smaller earthquakes with a magnitude of  $M_f-2$  to  $M_f-3$  often shows abnormalities like an acceleration or decrease before a major earthquake ( $M_f$ ). Nevertheless, PAPAACHOS *et al.* (2005) determined that the seismicity of smaller earthquakes with magnitude  $M_f-1.5$  to  $M_f-2$  often shows abnormalities like an acceleration or decrease before a major earthquake ( $M_f$ ). In our study, we hypothesized that earthquakes with magnitudes larger than  $M_c + 2.5$  will occur preferentially in hotspots during the forecast time interval  $t_2$  to  $t_3$ .

In order to make sure our codes work correctly, we calculated the hotspot map of the California region with the same parameters as HOLLIDAY *et al.* (2006b) did. The results produced by our codes were consistent with those by HOLLIDAY *et al.* (2006b), where slight differences between the two maps are likely caused by different earthquake catalogs (ZHANG *et al.* 2009).

### 3. Data and Computing Parameters Chosen

#### 3.1. The Selected Region and Its Seismicity

The west of continental China (included in  $20.0^\circ$ – $50.0^\circ$ N,  $70.0^\circ$ – $110.0^\circ$ E region) was chosen as the study region. This region has frequent large earthquakes. According to the statistical data of the China Earthquake Networks Center, 65 large earthquakes of  $M \geq 7.0$  occurred in the selected region in 1900–2008 (Fig. 2); of these 57 occurred in Chinese territory, 5 in foreign countries very close to the territorial border of China, and 3 more than 300 km away from the territorial border of China. Nine of the 65 earthquakes were tremendous earthquakes of  $M \geq 8.0$  (five-pointed stars in Fig. 2), and the largest one was the Chayu, Tibet ( $28.4^\circ$ N,  $96.7^\circ$ E) M8.6 earthquake that occurred on 15 August 1950 (BJT). Since 2000, there have been three  $M \geq 7.0$  earthquakes in West Continental China. The first was the West Kunlun Pass ( $36.21^\circ$ N,  $96.9^\circ$ E) M8.1 earthquake that occurred on

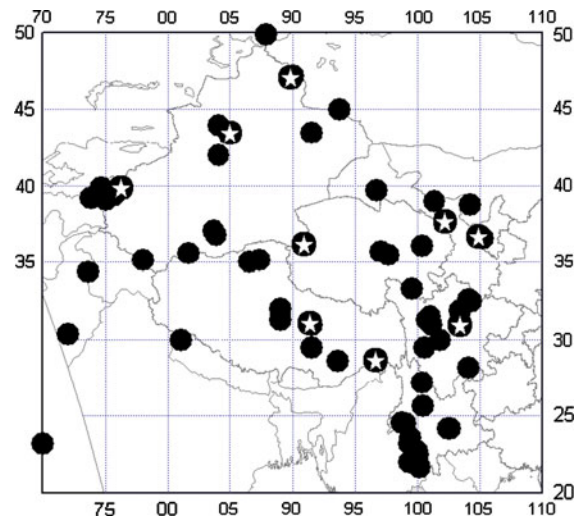


Figure 2

Large earthquakes with  $M \geq 7.0$  occurring in and near West Continental China in 1900–2008 (CENC). Solid circles represent 67 earthquakes larger than 7.0, and five-pointed stars represent 9 earthquakes larger than 8.0

14 November 2001 (BJT). The second was the Yutian ( $35.6^\circ$ N,  $81.6^\circ$ E) M7.3 earthquake that occurred on 21 March 2008 (BJT). The third one is the Wenchuan ( $31.0^\circ$ N,  $103.4^\circ$ E) M8.0 earthquake that occurred on 12 May 2008 (BJT). The West Kunlun Pass M8.1 earthquake and the Yutian M7.3 earthquake did not cause significant casualties because they occurred in depopulated areas, but the Wenchuan M8.0 earthquake resulted in nearly 90,000 dead and missing, as it occurred in a densely populated region.

As mentioned above, the Yutian M7.3 and Wenchuan M8.0 earthquakes did not fall within the forecast hotspots for the 2000–2010 period given by HOLLIDAY *et al.* (2005), as shown in Fig. 1. In order to study the predictability of PI for these two large earthquakes, we needed to carry out a retrospective study with different parameters to calculate the change in probability (Formula 6).

#### 3.2. The Monitoring Ability and Completeness of the Earthquake Catalogue of the Selected Region

The earthquake catalogue employed in this paper is from the CENC (China Earthquake Networks Center) with an initial time of 1970. For the selected

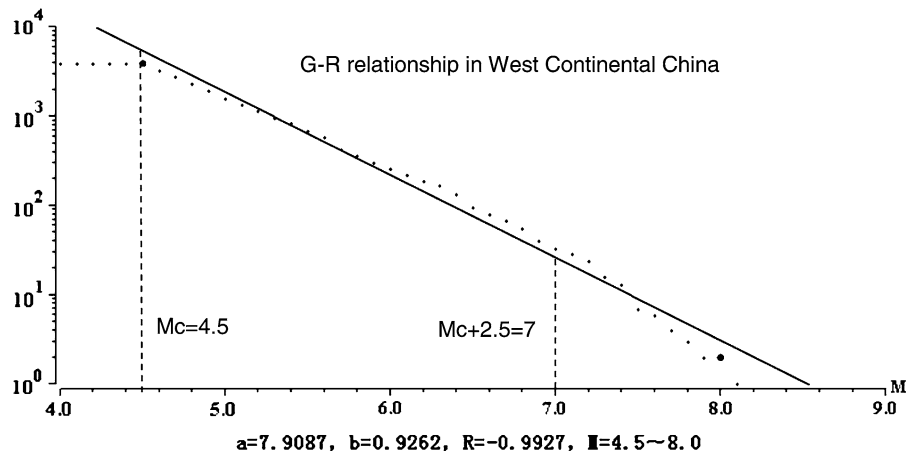


Figure 3  
Completeness of catalogue test by G-R relation in West Continental China

West Continental China region, the China Digital Network has been able to record all earthquakes larger than  $M_L 4.0$  so far. However, during the 1970s–1980s, earthquakes less than  $M_L 4.5$  could not be recorded completely. In order to calculate the forecast hotspots with PI method reasonably, we needed to employ a complete earthquake catalogue. According to the evaluation of seismic monitoring abilities by the CEA (<http://www.csi.ac.cn/yxy/jance/jiancenengli.htm>), the lower cutoff magnitude  $M_c$  is  $M_L 4.5$  for the completeness of the earthquake catalogue in West Continental China since 1970 (Fig. 3).

Figure 4 shows the distribution map of earthquakes larger than  $M_L 4.5$  recorded by the China Digital Seismic Network during the period from 1970 to 2009. From this figure we can see that the recorded earthquakes cover the territory of West Continental China and its neighboring areas.

### 3.3. The Tested Models with Different Computing Parameters

We divided the selected region to be studied into a grid of square boxes with the size of  $1^\circ \times 1^\circ$  and  $2^\circ \times 2^\circ$ . For the  $1^\circ \times 1^\circ$  grid model, the total number of boxes is 1200. For the  $2^\circ \times 2^\circ$  grid model, the total number of boxes is 600.

In order to study the predictability of the PI method for the Yutian M7.3 and Wenchuan M8.0 earthquakes, we needed to try different calculating parameters of the beginning time  $t_b$ , the reference

time interval  $t_b$  to  $t_1$ , change interval  $t_1$  to  $t_2$  and the forecast time interval  $t_2$  to  $t_3$ . We designed 20 models with different parameters, which are listed in Table 1. For these models, the beginning time  $t_b$  and the ending forecast time  $t_3$  were the same, while the change interval  $t_1$  to  $t_2$  and the forecast time interval  $t_2$  to  $t_3$  ranged from 1 to 10 years, and the reference time interval  $t_b$  to  $t_1$  ranged from 17 to 35 years.

## 4. Results of Retrospective Tests for Different Models

### 4.1. PI Patterns from Different Models

Generally, hotspots are defined to be the regions where  $\Delta P_i(t_0, t_1, t_2)$  is positive. Definitely, there are more hotspots when the threshold of  $\Delta P_i(t_0, t_1, t_2)$  is lower. In order to raise the hit rate and reduce the miss rate, we have to make the decision threshold of possibility gain  $\Delta P_i(t_0, t_1, t_2)$ . After trying different thresholds of  $\Delta P_i(t_0, t_1, t_2)$ , we could find the best fit of  $\Delta P_i(t_0, t_1, t_2)$ , under which the hit rate was relatively higher, while the miss rate was relatively lower.

When we took the threshold possibility as  $\log_{10}(\Delta P_i(t_0, t_1, t_2)/\Delta P_{\max}(t_0, t_1, t_2)) = -0.6$ , the best good fitness hotspot map could be obtained as shown in Fig. 5.

From Fig. 5 we can see that, for the Yutian M7.3 earthquake, it falls in the hotspot box or the neighbor

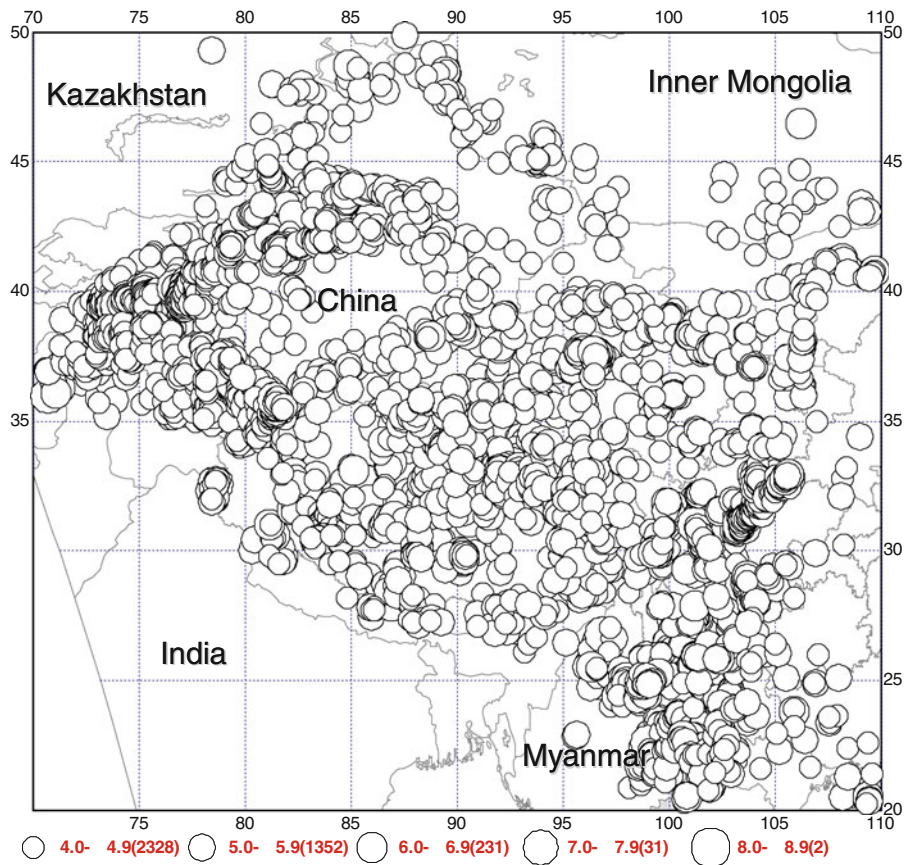


Figure 4

Earthquake distribution map in West Continental China recorded by the China Digital Seismic Network (the earthquake catalogue is from the CENC, with all recorded earthquakes larger than  $M_L 4.5$  for 1970–2009)

box in models 11–19. For the Wenchuan M8.0 earthquake, it drops in the hotspot box or the neighbor box in models 11, 13, 14 and 17–20. From this direct perspective view, models with  $2^\circ \times 2^\circ$  boxes are better than those with  $1^\circ \times 1^\circ$  boxes for predicting the Yutian M7.3 and Wenchuan M8.0 earthquakes.

In order to evaluate the predictive effects of these models quantitatively, we employed the ROC (SWETS 1973; MOLCHAN 1997) test as RUNDLE *et al.* (2000c, 2003), TIAMPO *et al.* (2002c) and HOLLIDAY *et al.* (2005) did. We also employed the  $R$  score evaluation method approved by XU *et al.* (1989) and SHI *et al.* (2000). The evaluation results are as follows.

#### 4.2. ROC Test for Different Models

The receiver operating characteristic test is conducted by systematic changing of the ‘alarm threshold’ of the

‘forecast region’ and counting the ‘hit rate’ and ‘false alarm rate’ compared to real earthquake activity. Following RUNDLE *et al.* (2000c, 2003), TIAMPO *et al.* (2002c) and HOLLIDAY *et al.* (2005), we define that: during  $t_2$  to  $t_3$ , if an earthquake larger than M7.0 occurs in a hotspot box or within the Moore neighborhood of the box, this is a success [the eight boxes surrounding the hotspot box are defined as “the Moore neighborhood” (MOORE 1962)]; if no earthquakes occur in a non-hotspot box, this is also a success; if no earthquake occurs in a hotspot box or within the Moore neighborhood of the hotspot box, this is a false alarm; if an earthquake occurs in a box that is not hotspot box or the Moore neighborhood of the hotspot box, this is a failure to forecast.

According to the above definitions, values  $a$  (Forecast = yes, Observed = yes),  $b$  (Forecast = yes, Observed = no),  $c$  (Forecast = no, Observed = yes) and  $d$  (Forecast = no, Observed = no) are

Table 1  
Twenty PI models with different parameters

Model no.	Grid size	$t_b$	$t_1$	$t_2$	$t_3$	Forecast period/year
1	1° × 1°	1971-6-1	2006-6-1	2007-6-1	2008-6-1	1
2	1° × 1°	1971-6-1	2004-6-1	2006-6-1	2008-6-1	2
3	1° × 1°	1971-6-1	2002-6-1	2005-6-1	2008-6-1	3
4	1° × 1°	1971-6-1	2000-6-1	2004-6-1	2008-6-1	4
5	1° × 1°	1971-6-1	1998-6-1	2003-6-1	2008-6-1	5
6	1° × 1°	1971-6-1	1996-6-1	2002-6-1	2008-6-1	6
7	1° × 1°	1971-6-1	1994-6-1	2001-6-1	2008-6-1	7
8	1° × 1°	1971-6-1	1992-6-1	2000-6-1	2008-6-1	8
9	1° × 1°	1971-6-1	1990-6-1	1999-6-1	2008-6-1	9
10	1° × 1°	1971-6-1	1988-6-1	1998-6-1	2008-6-1	10
11	2° × 2°	1971-6-1	2006-6-1	2007-6-1	2008-6-1	1
12	2° × 2°	1971-6-1	2004-6-1	2006-6-1	2008-6-1	2
13	2° × 2°	1971-6-1	2002-6-1	2005-6-1	2008-6-1	3
14	2° × 2°	1971-6-1	2000-6-1	2004-6-1	2008-6-1	4
15	2° × 2°	1971-6-1	1998-6-1	2003-6-1	2008-6-1	5
16	2° × 2°	1971-6-1	1996-6-1	2002-6-1	2008-6-1	6
17	2° × 2°	1971-6-1	1994-6-1	2001-6-1	2008-6-1	7
18	2° × 2°	1971-6-1	1992-6-1	2000-6-1	2008-6-1	8
19	2° × 2°	1971-6-1	1990-6-1	1999-6-1	2008-6-1	9
20	2° × 2°	1971-6-1	1988-6-1	1998-6-1	2008-6-1	10

obtained for the hotspot map. The fraction of colored boxes, also called the probability of forecast of occurrence, is  $r = (a + b)/N$ , where the total number of boxes is  $N = a + b + c + d$ . The hit rate is  $H = a/(a + c)$  and is the fraction of large earthquakes that occur on a hotspot. The false alarm rate is  $F = b/(b + d)$  and is the fraction of non-observed earthquakes that are incorrectly forecast.

For each model, we obtain the PI hotspot maps under different thresholds of  $\Delta P_i(t_0, t_1, t_2)$  from 0 to  $\Delta P_{\max}(t_0, t_1, t_2)$  firstly; and then calculate the hit rate and false rate of according to above mentioned method.

Figure 6 shows the diagrams of the ROC test for different models. From these results we can see that, compared to the random forecast ( $H = F$ ), all 20 models have higher forecast efficacy since  $H > F$  in most cases.

In order to evaluate the degree of forecast efficiency of the models quantitatively, we define a parameter  $E_f$  as the following,

$$E_f = \sum_i H_i \cdot \Delta F_i \quad (7)$$

where  $H_i$  denotes the hit rate associated with the false rate of  $F_i$ . The essence of  $E_f$  is the area surrounded by

the curve of  $H(F)$ , the line of  $H = 0$  and the line of  $F = F_{\max}$ . Here  $F_{\max}$  is the false rate under the threshold of  $\Delta P_i(t_0, t_1, t_2) = 0$ . For the 20 models in this paper,  $F_{\max} = 0.07$ , as shown in Fig. 6. If the hit rate is bigger under the same false rate, the parameter  $E_f$  is bigger; hence, the forecast efficiency could be determined by  $E_f$ . Bigger  $E_f$  means higher forecast efficiency.

The values of the  $E_f$  corresponding to the 20 models are 0.0275, 0.0175, 0.017, 0.014, 0.02, 0.017, 0.0035, 0.025, 0.0305, 0.033, 0.023, 0.0115, 0.035, 0.0485, 0.043, 0.048, 0.056, 0.059, 0.03 and 0.028, respectively.

Figure 7 shows the pictorial diagram of  $E_f$  varied with different models. From this figure we can see that model 18 (with a 2° × 2° box and 8-year forecast window) outperforms all other models, because when the hit rate reaches 1, the false alarm rate is the lowest (square dashed line in Fig. 6d). The next best model after model 18 is model 17, as shown in Fig. 6 (diamond solid line in Fig. 6d). The third best models are model 16 (dotted solid line in Fig. 6d) and model 14 (asterisk dashed line in Fig. 6c). Generally, the forecast efficiency of the models with 2° × 2° boxes are higher than those with 1° × 1° boxes.

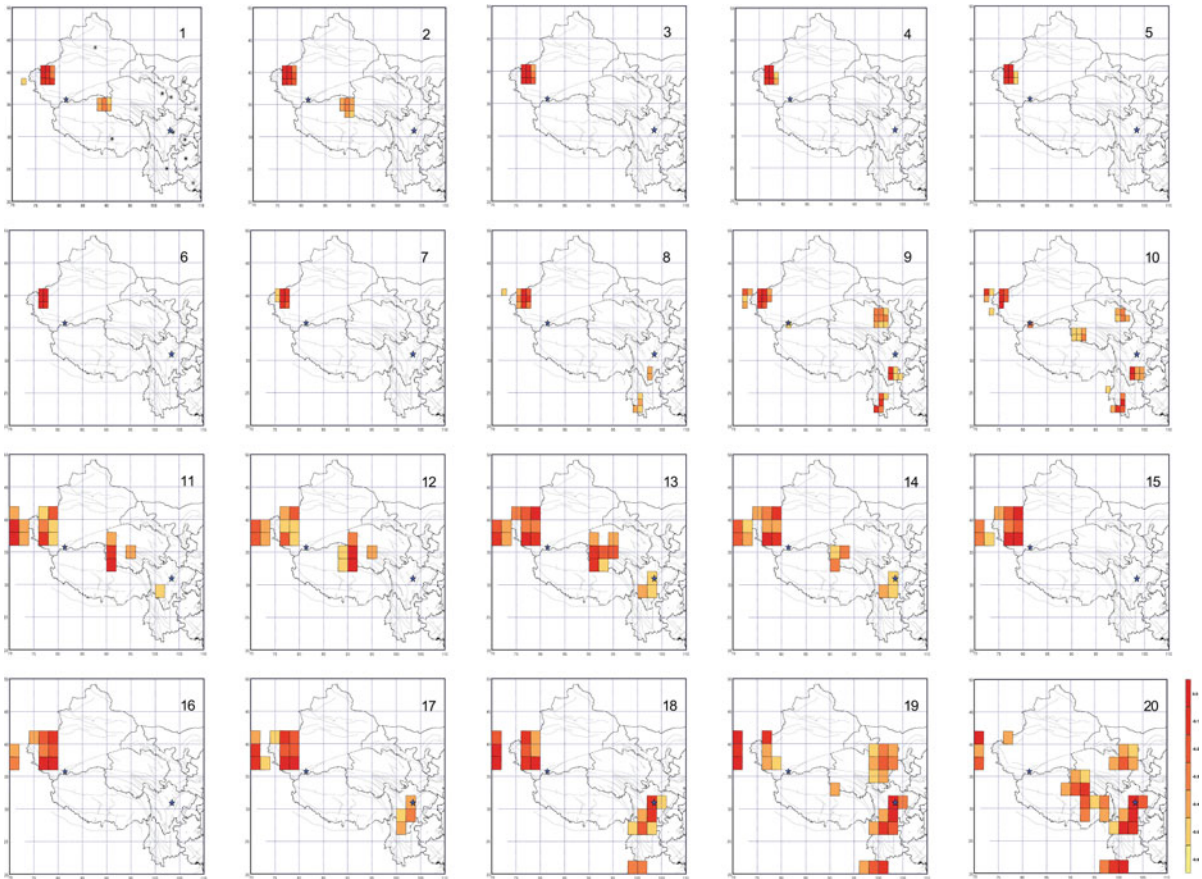


Figure 5

Hotspot maps for the 20 models under the threshold possibility  $\log_{10}(\Delta P_i(t_0, t_1, t_2)/\Delta P_{\max}(t_0, t_1, t_2)) = -0.6$ . The five-point stars denote the positions of Yutian M7.3 and Wenchuan M8.0 earthquakes in 2008, respectively. The numbers in the figures represent the models, corresponding to the model no. in Table 1

#### 4.3. *R* Score Test for Different Models

From 1989, an evaluation method called the *R* score was developed by Xu *et al.* (1989) for the efficacy of earthquake forecast methods in China. Shi *et al.* (2000) developed the algorithm of the *R* score test for the annual potential seismic hazard regions forecast by the China Earthquake Administration (CEA). Compared with the ROC test, “the Moore neighborhood” is not considered in the *R* score test, i.e. during  $t_2$  to  $t_3$ , if an earthquake larger than M7.0 occurs in a hotspot box, this is a success; if no earthquake occurs in a non-hotspot box, this is also a success; if no earthquake occurs in a hotspot box, this is a false alarm; if an earthquake occurs in a box that is not hotspot box, this is a failure to forecast. So the *R* score test is more rigorous than the ROC test.

According to the above definitions, values *a* (Forecast = yes, Observed = yes), *b* (Forecast = yes, Observed = no), *c* (Forecast = no, Observed = yes) and *d* (Forecast = no, Observed = no) are obtained for the hotspot map. The hit rate is  $H = a / (a + c)$ . The false alarm rate is  $F = b / (b + d)$ . *R* score =  $H - F$ .

Figure 8 shows the relationships between the hit rate and false alarm rate of the 20 models by *R* score test. From these results we can see that, compared to the random forecast ( $H = F$ ), all 20 models have higher forecast efficacy since  $H > F$  in most cases.

We can also calculate the values of the  $E_f$  corresponding to the 20 models by Formulae (7). They are 0.0195, 0.0175, 0.015, 0.003, 0.0095, 0.002, 0.0035, 0.016, 0.0185, 0.0245, 0.0045, 0, 0.0015, 0.008, 0.007, 0.0065, 0.017, 0.025, 0.03 and 0.0285, respectively.



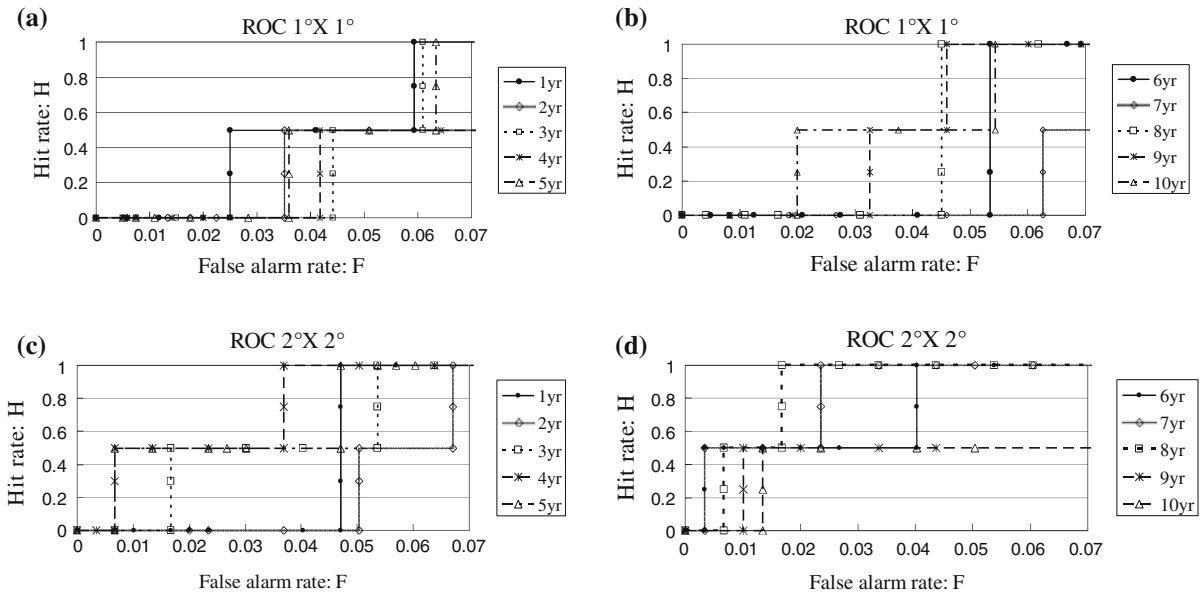


Figure 6

ROC test for different models (hit rate and false rate are calculated by changing the threshold possibility  $\Delta P_i(t_0, t_1, t_2)$  from 0 to  $\Delta P_{\max}(t_0, t_1, t_2)$ )

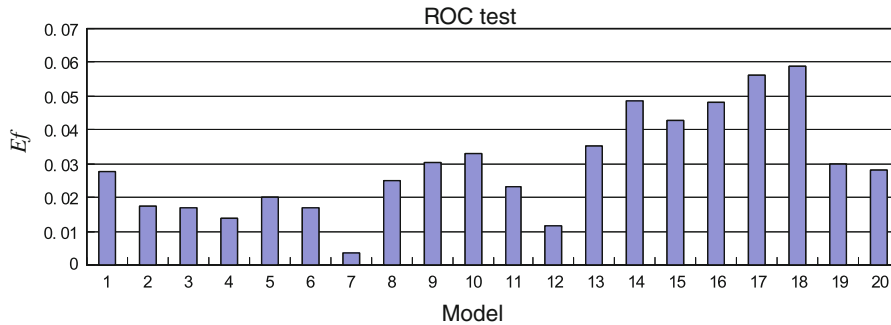


Figure 7

Pictorial diagram of  $E_f$  varied with different models using the ROC test

Figure 9 shows the pictorial diagram of  $E_f$  varied with different models using the  $R$  score test. From this figure we can see that model 19 (with a  $2^\circ \times 2^\circ$  box and 9-year forecast window) outperforms all other models because when the hit rate reaches to 0.5 the false alarm rate is the lowest (asterisk dashed line in Fig. 8d). The next best model after model 19 is model 20, as shown in Fig. 8 (triangle dashed line in Fig. 8d). The third best model is model 18 (square dashed line in Fig. 7d).

We can also draw the conclusion that generally the forecast efficiency of those models with longer

forecast windows and larger scale boxes is higher than that of those with shorter forecast windows and smaller scale boxes. Compared with the ROC test, the hit rate universally decreases; hence the  $E_f$  value decreases.

#### 4.4. Better Models for Large Earthquake Forecasts

From the above results of the ROC test and  $R$  score test, we can see that models 18, 17, 16 and 14 are the top four models in the ROC test, while models 19, 20, 18 and 10 are the top four models using the  $R$  score test. For a combination test of the ROC

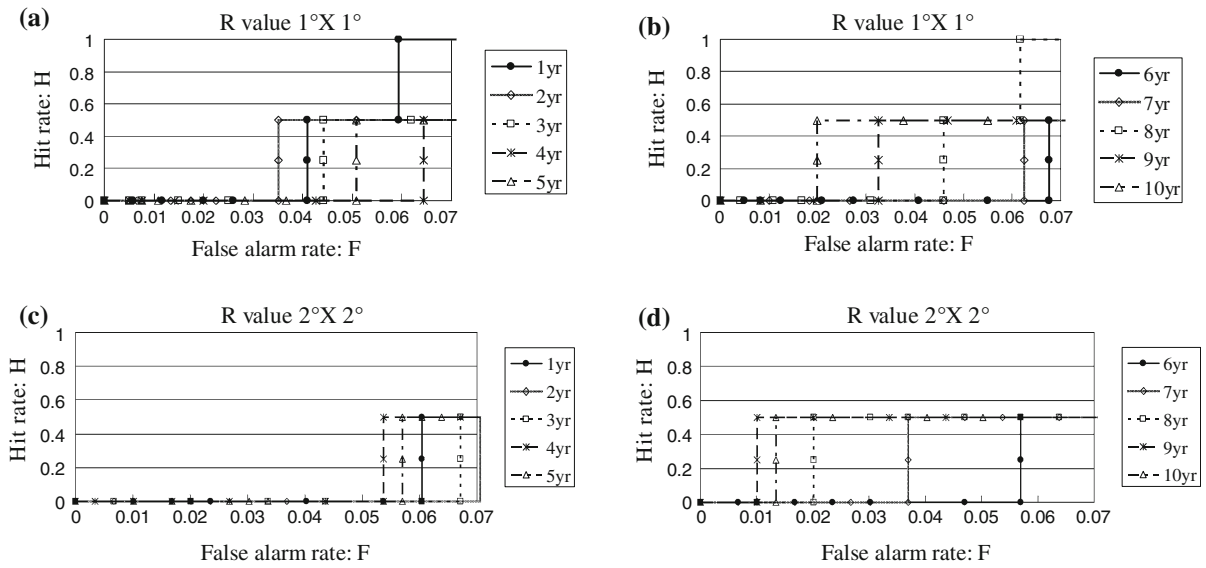


Figure 8

$R$  score test for different models (hit rate and false rate are calculated by changing the threshold possibility  $\Delta P_i(t_0, t_1, t_2)$  from 0 to  $\Delta P_{\max}(t_0, t_1, t_2)$ )

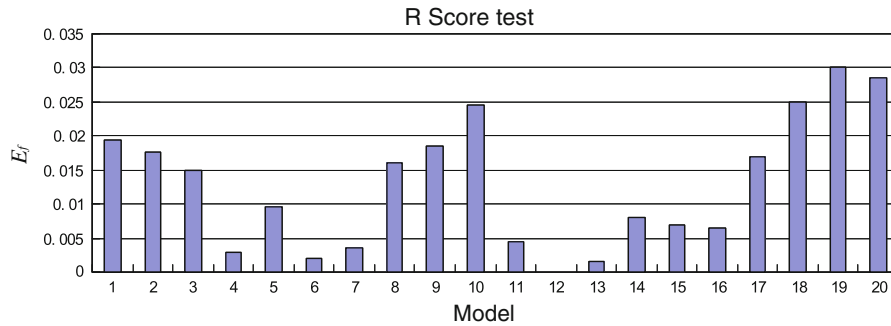


Figure 9

Pictorial diagram of  $E_f$  varied with different models using the  $R$  score test

and  $R$  scores, model 18 is the best model with higher forecast efficacy. The quantitative evaluation results are coincident with the direct perspective view in Fig. 5.

### 5. Conclusions and Discussions

In this paper, West Continental China (included in the region of  $20.0^{\circ}$ – $50.0^{\circ}$ N,  $70.0^{\circ}$ – $110.0^{\circ}$ E) was taken as the study region to verify the predictability of the PI method by the ROC test and  $R$  score test. From the results, we can draw the following conclusions:

1. The PI method has much higher forecast efficacy than the random forecast using the quantitative ROC test and  $R$  score test. It is really an optimal method for long-term earthquake forecasts.
2. For large earthquakes like the Yutian M7.3 and Wenchuan M8.0 earthquakes, taking a grid size of  $2^{\circ} \times 2^{\circ}$  and forecast window of 8 years could raise the forecast efficacy most, and Yutian M7.3 and Wenchuan M8.0 earthquakes could drop in the hotspots. The essence of this conclusion may be that larger earthquakes have bigger critical seismogenic size and longer seismogenic time. For example, BUFE and VARNES (1993) give the

relationship between the critical seismogenic size  $R$  and the magnitude  $M$  as  $\log R = -0.2 + 0.36M$ . By this formula, when  $M \geq 5$ ,  $R \geq 40$  km, and when  $M \geq 7$ ,  $R \geq 209$  km. So models with  $2^\circ \times 2^\circ$  grid size are better in forecasting larger earthquakes than those with  $1^\circ \times 1^\circ$  grid size.

In Holliday's model (HOLLIDAY *et al.* 2005) with a grid size of  $1^\circ \times 1^\circ$  and forecast window of 10 years, the Yutian M7.3 and Wenchuan 8.0 earthquakes did not drop in the hotspot map (Fig. 1). In our retrospective study, the Yutian M7.3 and Wenchuan 8.0 earthquakes did not drop in the hotspots or their Moore neighborhood in any of the models with a grid size of  $1^\circ \times 1^\circ$  (Model 1–10 in Fig. 5), but both of the two large earthquakes dropped in the hotspots or their Moore neighborhood of models 11, 13, 14, 17, 18 and 19 (Fig. 5). In the rest of the models with a grid size of  $2^\circ \times 2^\circ$  (models 12, 15, 16 and 20 in Fig. 5), both Yutian M7.3 and Wenchuan M8.0 earthquake could drop in the hotspot or its Moore neighborhood. The different results between Holliday's model (2005) and ours may be caused by two factors. The first is that different catalogues were employed. Holliday's is from the global seismic network with a lower magnitude cutoff of  $M_c = 5$  and initial time  $t_0 = 1,965$ , while ours is from the China Earthquake Networks Center (CENC) with a lower magnitude cutoff of  $M_c = 4.0$  and initial time  $t_0 = 1,970$ . The second factor may be that the grid sizes are different in Holliday's model and ours. Which factor is key can be investigated in further study.

Further study is valuable to verify whether Rundle's model with a  $0.1^\circ \times 0.1^\circ$  box for forecasting  $M \geq 5.0$  earthquakes in California region (RUNDLE *et al.* 2002; TIAMPO *et al.* 2002b; HOLLIDAY *et al.* 2005) could be improved by modulating the grid size from  $0.1^\circ \times 0.1^\circ$  to  $0.4^\circ \times 0.4^\circ$ .

In summary, before applying the PI model for earthquake forecasting in a region, the suitable time points  $t_1$ ,  $t_2$  and  $t_3$ , suitable thresholds of lower cutoff of magnitude, suitable thresholds of  $\log_{10}(\Delta P / \Delta P_{\max})$ , etc., should be selected after systematic retrospective study by ROC test or  $R$  score test. Higher forecasting efficacy can be reached with the suitable models.

### Acknowledgments

The authors gratefully acknowledge support from the Ministry of Science and Technology under grant nos. 2010DFB20190 and 2008BAC35B05. We also thank Prof. J.B. Rundle for giving us some comments on PI methods during the 6th ACES International Workshop in Cairns, Australia, 11–16 May 2008. We also thank the China Earthquake Networks Center (CENC) for earthquake data.

### REFERENCES

- BUFE, C.G., VARNES, D.J. (1993) *Predictive modeling of the seismic cycle of the greater San Francisco Bay region*. J.Geophys. Res.98, 9871–9883.
- CHEN, C.C., RUNDLE, J.B., HOLLIDAY, J.R., NANJO, K.Z., TURCOTTE, D.L., LI, S.C., TIAMPO, K.F. (2005) *The 1999 Chi-chi, Taiwan, earthquake as a typical example of seismic activation and quiescence*. Geophysical Research Letters, vol. 32. L22315.
- CHEN, C.C., RUNDLE, J.B., LI, H.C., *et al.* (2006) *From tornadoes to earthquakes: Forecast verification for binary events applied to the 1999 Chi-Chi, Taiwan, earthquake*, TERRESTRIAL ATMOSPHERIC AND OCEANIC SCIENCES, 17 (3): 503–516.
- MOORE E.F. (1962) Machine models of self reproduction, in proceedings of the fourteenth symposium on applied mathematics, American Mathematical Society, pp. 17–33.
- HOLLIDAY, J.R., NANJO, K.Z., TIAMPO, K.F., RUNDLE, J.B. and TURCOTTE, D.L., 2005. *Earthquake forecasting and its verification*, Nonlinear Processes in Geophysics, 12: 965–977.
- HOLLIDAY, J.R., RUNDLE, J.B., TIAMPO, K.F., KLEIN, W. and DONNELLAN, A., 2006a. *Modification of the pattern informatics method for forecasting large earthquake events using complex eigenfactors*. Tectonophysics, 413: 87–91.
- HOLLIDAY, J.R., RUNDLE, J.B., TIAMPO, K.F., KLEIN, W., and DONNELLAN, A. (2006b) *Systematic procedural and sensitivity analysis of the pattern informatics method for forecasting large ( $M \geq 5$ ) earthquake events in southern California*, Pure Appl. Geophys., 163:2433–2454.
- JAUME, S.C., SYKES, L.R. (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.
- JIANG, C.S. and WU, Z.L., 2008. *Retrospective forecasting test of a statistical physics model for earthquakes in Sichuan-Yunnan region*, Science in China Series D: Earth Sciences, 51(10): 1401–1410, doi: 10.1007/s11430-008-0112-6.
- MOLCHAN G.M. (1997) *Earthquake prediction as a decision-making problem*. Pure Appl Geophys, 1997, 149: 233–247.
- NANJO, K.Z.; RUNDLE, J.B.; HOLLIDAY, J.R.; *et al.* (2006a) *Pattern informatics and its application for optimal forecasting of large earthquakes in Japan*, PURE AND APPLIED GEOPHYSICS, 163 (11–12): 2417–2432.
- NANJO, K.Z., HOLLIDAY J.R., CHEN C.C., RUNDLE J.B., TURCOTTE D.L. (2006b) *Application of a modified pattern informatics method to forecasting the locations of future large earthquakes in the central Japan*, Tectonophysics, 424: 351–366.

- PAPAZACHOS, C.B., KARAKAISIS, G.F., SCORDILIS, E.M., PAPAZACHOS, B.C. (2005) *Global observational properties of the critical earthquake model*. Bull. Seismol. Soc. Am. 95, 1841–1855.
- RUNDLE, J.B., KLEIN, W., GROSS, S.J., and TIAMPO, K.F. (2000a) Dynamics of seismicity patterns in systems of earthquake faults, in *Geo-complexity and the Physics of Earthquakes*, edited by J.B. Rundle, D.L. Turcotte, and W. Klein, vol. 120 of Geophys. Monogr. Ser., pp. 127–146, AGU, Washington, D. C.
- RUNDLE, J.B., KLEIN, W., TIAMPO, K.F., and GROSS, S.J. (2000b) *Linear Pattern Dynamics in Nonlinear Threshold Systems*, Phys. Rev. E., 61, 2418–2432.
- RUNDLE, J.B., KLEIN, W., TURCOTTE, D.L., et al. (2000c) *Precursory seismic activation and critical-point phenomena*. Pure Appl Geophys, 2000c, 157:2165–2182.
- RUNDLE, J.B., TIAMPO, K.F., KLEIN, W., and MARTINS, J.S.S. (2002) *Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neuro-biology, and forecasting*, Proc. Natl. Acad. Sci. USA, 99, 2514–2521, Suppl. 1.
- RUNDLE, J.B., TURCOTTE, D. L., SHCHERBAKOV, R., KLEIN, W., and SAMMIS, C. (2003) *Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems*, Rev. Geophys., 41, 1019–1038.
- SHI Y.L., LIU J., ZHANG G.M. (2000) *The evaluation of Chinese annual earthquake prediction in the 90s*. J Graduate School Academia Sin (in Chinese with English abstract), 2000, 17: 63–69.
- SWETS J.A. (1973) *The relative operating characteristic in psychology*. Science, 1973, 182: 990–1000.
- TIAMPO, K.F., RUNDLE, J.B., MCGINNIS, S., GROSS, S.J., and KLEIN, W. (2002a) *Eigenpatterns in southern California seismicity*, J. Geophys. Res., 107, 2354.
- TIAMPO, K.F., RUNDLE, J.B., MCGINNIS, S., and KLEIN, W. (2002b) *Pattern dynamics and forecast methods in seismically active regions*, Pure App. Geophys, 159, 2429–2467.
- TIAMPO K.F., RUNDLE J.B., MCGINNIS S., et al. (2002c) *Mean-field threshold systems and phase dynamics: An application to earthquake fault systems*. Europhys Lett, 2002c, 60: 481–487.
- WU, Y.H., CHEN, C.C., RUNDLE, J.B. (2008a) *Detecting precursory earthquake migration patterns using the pattern informatics method*, GEOPHYSICAL RESEARCH LETTERS, 35 (19): Art. No. L19304.
- WU, Y.H., CHEN, C.C., RUNDLE, J.B. (2008b) *Precursory Seismic Activation of the Pingtung (Taiwan) Offshore Doublet Earthquakes on 26 December 2006: A Pattern Informatics Analysis*, TERRESTRIAL ATMOSPHERIC AND OCEANIC SCIENCES, 19 (6): 743–749 DEC 2008b.
- XU S.X. (1989) *Mark evaluation for earthquake prediction efficacy*. In: Department of Science and Technology Monitoring, State Seismological Bureau, ed. *Collected Papers of Research on Practical Methods of Earthquake Prediction (Volume of Seismology)* (in Chinese). Beijing: Academic Books and Periodical Press, 1989. 586–590.
- ZHANG, Y.X., ZHANG, X.T., YIN, X.C. and WU, Y.J. (2009) *Study on the Forecast Effects of PI Method to the North and Southwest China, Currency and Computation: Practice and Experience*. Published online in Wiley InterScience (<http://www.interscience.wiley.com>). doi:10.1002/cpe.1515.

(Received January 29, 2011, revised October 13, 2011, accepted November 30, 2011, Published online January 13, 2012)