

Implications of the Regional Earthquake Likelihood Models test of earthquake forecasts in California

Michael K. Sachs,¹ Ya-Ting Lee,² Donald L. Turcotte,³ James R. Holliday,¹
John B. Rundle^{1,3,4}

¹Physics Department, University of California, Davis, CA, USA; ²Graduate Institute of Geophysics, National Central University, Jhongli, Taiwan, Republic of China; ³Geology Department, University of California, Davis, CA, USA; ⁴Santa Fe Institute, Santa Fe, NM, USA

Abstract

The Regional Earthquake Likelihood Models (RELM) test was the first competitive comparison of prospective earthquake forecasts. The test was carried out over 5 years from 1 January 2006 to 31 December 2010 over a region that included all of California. The test area was divided into 7682 $0.1^\circ \times 0.1^\circ$ spatial cells. Each submitted forecast gave the predicted numbers of earthquakes N_{emi} larger than $M=4.95$ in 0.1 magnitude bins for each cell. In this paper we present a method that separates the forecast of the number of test earthquakes from the forecast of their locations. We first obtain the number N_{em} of forecast earthquakes in magnitude bin m . We then determine the conditional

probability $\lambda_{emi} = \frac{N_{emi}}{N_{em}}$ that an earthquake in magnitude bin m will occur in cell i . The summation of λ_{emi} over all 7682 cells is unity. A random (no skill) forecast gives equal values of λ_{emi} for all spatial cells and magnitude bins. The *skill* of a forecast, in terms of the location of the earthquakes, is measured by the success in assigning large values of λ_{emi} to the cells in which earthquakes occur and low values of λ_{emi} to the cells where earthquakes do not occur. Thirty-one test earthquakes occurred in 27 different combinations of spatial cells i and magnitude bins m , we had the highest value of λ_{emi} for that mi cell. We evaluate the performance of eleven submitted forecasts in two ways. First, we determine the number of mi cells for which the forecast λ_{emi} was the largest, the best forecast is the one with the highest number. Second, we determine the mean value of λ_{emi} for the

27 mi cells for each forecast. The best forecast has the highest mean value of λ_{emi} . The success of a forecast during the test period is dependent on the allocation of the probabilities λ_{emi} between the mi cells, since the sum over the mi cells is unity. We illustrate the forecast distributions of λ_{emi} and discuss their differences. We conclude that the RELM test was successful in illustrating the choices required when a forecast of the location of a future earthquake is made.

Introduction

Prospective forecasts of earthquakes are forecasts of earthquakes that may occur in the future. Retrospective forecasts are forecasts of earthquakes that have already occurred. In principal, a retrospective forecast can be carried out fairly. However, in many cases, the retrospective forecasts are based implicitly or explicitly on the occurrence of the forecast earthquakes.

Examples of successful retrospective forecasts are those based on accelerated moment release (AMR). A systematic increase in Benioff strain is observed prior to an earthquake. Examples of AMR have been given by Bufe and Varnes¹ and Bowman *et al.*² among others. However, the epicenters of the subsequent earthquakes are used to define the regions in which AMR occurs. Hardebeck *et al.*³ have argued that it is not possible to establish AMR without knowing the locations of the subsequent earthquakes, thus AMR cannot be used for prospective forecasting.

A primary example of prospective forecasts has been the sequence of forecasts issued by the International Institute for the Theory of Prediction and Theoretical Geophysics in Moscow. Their forecast algorithms are based on pattern recognition of regional seismicity.^{4,5} Increased rates of occurrence of intermediate magnitude earthquakes are the primary components of their forecasts, thus they have a similarity to AMR. When a threshold of anomalous behavior is reached, a time of increased probability (TIP) is issued. These are alarm-based forecasts. TIPs were released prior to the $m=6.9$ Armenian earthquake on 7 December 1988 and prior to the $m=6.9$ Loma Prieta earthquake on 17 October 1989. Over a period of some 25 years successful TIPs were issued prior to 42 of 47-targeted earthquakes.⁶ However, the results have been subject to criticism. Notable earthquakes were not predicted and there were too many false alarms.⁷

Another example of a prospective forecast of earthquake occurrence was made for $m>5$ earthquakes in California for the period 2000-2010 by Rundle *et al.*⁸ This was also an alarm based forecast. Earthquakes were forecast to occur in specified *hot spots*. Holliday *et al.*⁹ reported that 16 of 18 $m \geq 5$ earthquakes that occurred during the period 2000-2005 occurred in hot spot regions. This forecast was based on the pattern informatics (PI) approach. Precursory seismic activation and quiescence were quantified and when variations

Correspondence: Michael K. Sachs, Physics Department, University of California, Davis, CA 95616, USA. E-mail: mksachs@ucdavis.edu

Key words: earthquake forecasting, RELM testing, forecast verification.

Acknowledgements: YTL is grateful for research support from both the National Science Council (ROC) and the institute of Geophysics (NCU, ROC). JBR and JRH have been supported by NASA Grant NNX08AF69G; MKS has been supported by NASA Earth and Space Science fellowship NNX11AL92H.

Received for publication: 15 November 2011.

Revision received: 20 April 2012.

Accepted for publication: 26 April 2012.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright M.K. Sachs *et al.*, 2012
Licensee PAGEPress, Italy
Research in Geophysics 2012; 2:e10
doi:10.4081/rg.2012.e10

exceeded a prescribed threshold *hot spots* were specified. Forecasts were made for $0.1^\circ \times 0.1^\circ$ cells (about 11×11 km) which corresponded roughly to the rupture length of an $m=6$ earthquake. Precursory seismicity included in the PI approach included $m \geq 3$ earthquakes. The size of the cells and magnitudes of earthquakes scale with AMR studies.¹⁰

A closely related forecasting technique is the relative intensity (RI) approach. The RI forecast is based on the direct extrapolation of the rate of occurrence of small earthquakes using Gutenberg-Richter frequency-magnitude scaling. Shcherbakov *et al.*¹¹ tested the RI method globally. The success of the PI method described above led to a discussion as to whether the PI method is significantly better than the RI method. Comparisons of these approaches have come to different conclusions regarding their validity.^{12,13}

Earthquakes are complex phenomena but they obey several scaling laws. One example is Gutenberg-Richter (GR) frequency-magnitude scaling. The cumulative number of earthquakes $N_{cumulative}$ with magnitudes greater than M in a region over a specified time are well approximated by the relation

$$\log N_{cumulative} = a - bM \quad (1)$$

where b is a near universal constant in the range $0.8 < b < 1.1$ and a is a measure of the level of seismicity. Equation (1) can be used to estimate the risk of large earthquakes based on the rate of occurrence of small earthquakes. This is a primary basis for the time independent seismic risk assessment. It is also the basis for the RI forecasts described above. An essential question is the role of the time dependence of the background seismicity in forecasting future earthquakes.

In order to test earthquake forecasts it is clearly desirable to use prospective forecasts. In order for a prospective forecast test to be useful it should be carried out in a reasonable length of time, say five years, and a reasonable number of earthquakes should be expected to occur. In order to meet these criteria the Regional Earthquake Likelihood Models (RELM) test was developed and carried out. The test region was California and adjacent regions.

Forecasts were solicited for $M > 5$ earthquakes during the period 2006 to 2010. It is the purpose of this paper to discuss these forecasts in terms of the earthquakes that actually occurred. The forecasts involve both the number of earthquakes that will occur during the test period and their locations. We present a method that separates these two aspects of the forecasts. Forecasts were required to give the numbers of earthquakes that were expected to occur in 7682 $0.1^\circ \times 0.1^\circ$ spatial cells and 41 magnitude bins. We convert these forecast numbers to a conditioned probability that an earthquake would occur in a specified spatial cell and magnitude bin in the test region during the test period. Forecasts were solicited for main shock only and for main shocks plus aftershocks. We will show that the submitted forecasts for the two cases gave the same conditional probabilities. We will also show that the conditional probabilities have either no dependence or a weak dependence on the magnitude of the forecast earthquakes.

In the next section, we give the details of the RELM test. In the following section, we discuss the earthquakes that occurred in the test region during the test period. The $M=7.2$ El Mayor-Cucapah earthquake on 4 April 2010 was particularly important. Then, we discuss the submitted forecasts. In section 5 we evaluate the performance of the forecasts. We conclude that no single performance measure can be used to measure the success of a submitted forecast. Moreover, we compare the performance of the submitted forecasts on the basis of the submitted probability that an earthquake would occur in a specified cell. The sum of these probabilities is unity so that the allocation of the probabilities between cells distinguishes the good forecasts from the bad forecasts. Finally, we summarize our results.

Regional Earthquake Likelihood Models test

In order to test methods for forecasting future earthquakes the Southern California Earthquake Center formed the working group for RELM in 2000.¹⁴ For the first time a competitive test of prospective earthquake forecasts was to be carried out. Research groups were encouraged to submit forecasts of future earthquakes in California. At the end of the test period, the forecasts would be compared with the actual earthquakes that occurred.

The ground rules for the RELM test were as follows:

- i) The test region to be studied was the state of California, however the selected region extended somewhat beyond the boundaries of the state as shown in Figure 1.
- ii) A five-year time period for the test was selected extending from 1 January 2006 to 31 December 2010. Earthquakes with $M \geq 4.95$ were to be forecast. For California, in years with major aftershock sequences, the level of seismicity is well approximated by GR scaling from Eq. 1 taking $b=1$ and $a=5.4$ per year.¹⁰ The number of $m > 5$ earthquakes expected per year would be 2.5 or 12.5 for 5 years. For $M \geq 6$, only about one earthquake would be expected, so the 5-year period would be much too short. The applicable magnitudes were taken from the Advanced National Seismic System online catalog <http://www.ncedc.org/anss/anss-detail.html>
- iii) Participants were required to submit the number of earthquakes expected to occur in specified spatial cells and magnitude bins during the test period. The test region was subdivided into $N_c=7682$ spatial cells with dimensions $0.1^\circ \times 0.1^\circ$ (approximately 11×11 km). These spatial cells were further divided into 41 magnitude bins: $4.95 \leq M < 5.05$, $5.05 \leq M < 5.15$, $5.15 \leq M < 5.25$, ... , $8.85 \leq M < 8.95$, $8.95 \leq M < \infty$. The participants were required to specify the expected number of earthquakes N_{emi} in magnitude bin m that would occur during the test period in cell i .
- iv) Participants could submit forecasts that included all earthquakes in the test region as well as forecasts that excluded aftershocks.

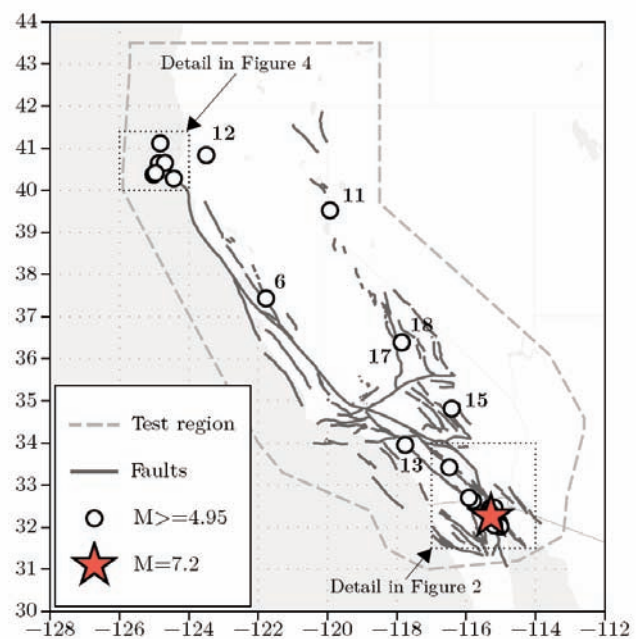


Figure 1. Map of the test region, the coast of California, major faults, and the 31 earthquakes with $M \geq 4.95$ that occurred in the test region. The earthquakes are listed in Table 1. Also shown are the square regions where large-scale maps are given in Figures 2 to 4.

Seventeen forecasts were submitted by eight groups. Before discussing these forecasts in some detail we will discuss the earthquakes that occurred in the test region during the test period with $M \geq 4.95$.

The earthquakes

During the test period 1 January 2006 to 31 December 2010, there were $N_e=31$ earthquakes in the test region with $M \geq 4.95$. The times of occurrence, locations, and magnitudes of these earthquakes are given in Table 1. The locations of the test earthquakes are also shown in Figures 1-4. The earthquakes are identified by the event numbers given in Table 1.

The major earthquake that occurred during the test period was the $M=7.2$ El Mayor-Cucapah earthquake on 4 April 2010 (event #22 in Table 1). This earthquake was on the plate boundary between the North American and Pacific plates. The epicenter was about 50 km

south of the Mexico-United States border, and the aftershocks indicate a rupture zone with a length of about 75 km. Both the epicenter and the aftershock sequence are illustrated in Figure 2.

The earthquakes within a $0.5^\circ \times 0.5^\circ$ region centered on the epicenter are illustrated in Figure 3. The El Mayor earthquake and the test earthquakes that occurred later, 4 April 2010 to 31 December 2010 are given in Figure 3B. Because of their proximity to the El Mayor earthquake in both space and time, events 23, 24, 25, 26, 28, and 31 are almost certainly aftershocks. The El Mayor earthquake and the test earthquakes that occurred earlier, 1 January 2006 to 3 April 2010 are given in Figure 3A. Events 1, 7, 8, 9, 10, 14, 16, and 19 constitute a precursory swarm of eight test earthquakes in this region in the magnitude range 4.97 to 5.80, including four in the 10 day period between 9 February and 19 February 2008 (events 7-10). These events are located some 5 km to 20 km north of the subsequent epicenter of the El Mayor-Cucapah earthquake and lie outside the primary aftershock region of that event, as illustrated in Figure 3A. This swarm of earth-

Table 1. Times of occurrence, locations, and magnitudes of the 31 earthquakes in the test region with $M \geq 4.95$ from 1 January 2006 until 31 December 2010. The $M=7.2$ El Mayor-Cucapah earthquake is in italics.

No.	Origin	Time (UTC)	Latitude	Longitude	M
1	2006/05/24	04:20:26.01	32.3067	-115.2278	5.37
2	2006/07/19	11:41:43.46	40.2807	-124.4332	5.00
3	2007/02/26	12:19:54.48	40.6428	-124.8662	5.40
4	2007/05/09	07:50:03.83	40.3745	-125.0162	5.20
5	2007/06/25	02:32:24.62	41.1155	-124.8245	5.00
6	2007/10/31	03:04:54.81	37.4337	-121.7743	5.45
7	2008/02/09	07:12:04.55	32.3595	-115.2773	5.10
8	2008/02/11	18:29:30.53	32.3272	-115.2568	5.10
9	2008/02/12	04:32:39.24	32.4475	-115.3175	4.97
10	2008/02/19	22:41:29.66	32.4325	-115.3130	5.01
11	2008/04/26	06:40:10.60	39.5253	-119.9289	5.00
12	2008/04/30	03:03:06.90	40.8358	-123.4968	5.40
13	2008/07/29	18:42:15.71	33.9530	-117.7613	5.39
14	2008/11/20	19:23:00.19	32.3288	-115.3318	4.98
15	2008/12/06	04:18:42.85	34.8133	-116.4188	5.06
16	2009/09/19	22:55:17.84	32.3707	-115.2612	5.08
17	2009/10/01	10:01:24.67	36.3878	-117.8587	5.00
18	2009/10/03	01:16:00.31	36.3910	-117.8608	5.19
19	2009/12/30	18:48:57.33	32.4640	-115.1892	5.80
20	2010/01/10	00:27:39.32	40.6520	-124.6925	6.50
21	2010/02/04	20:20:21.97	40.4123	-124.9613	5.88
22	<i>2010/04/04</i>	<i>22:40:42.15</i>	<i>32.2587</i>	<i>-115.2872</i>	<i>7.20</i>
23	2010/04/04	22:50:17.08	32.0972	-115.0467	5.51
24	2010/04/04	23:15:14.24	32.3000	-115.2595	5.43
25	2010/04/04	23:25:06.95	32.2462	-115.2978	5.38
26	2010/04/05	00:07:09.07	32.0180	-115.0172	5.32
27	2010/04/05	03:15:24.46	32.6282	-115.8062	4.97
28	2010/04/08	16:44:25.92	32.2198	-115.2760	5.29
29	2010/06/15	04:26:58.48	32.7002	-115.9213	5.72
30	2010/07/07	23:53:33.53	33.4205	-116.4887	5.43
31	2010/09/14	10:52:18.00	32.0485	-115.1982	4.96

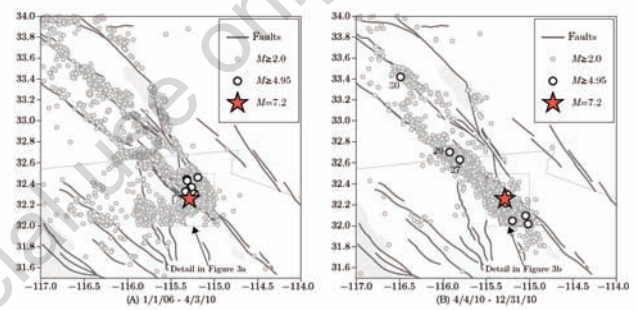


Figure 2. Map of the southeast region around the epicenter of the $M=7.2$ El Mayor-Cucapah earthquake that occurred on 4 April 2010 (event #22 in Table 1, shown as a star). (A) Earthquakes during the period 1 January 2006 through 3 April 2010. (B) Earthquakes during the period 4 April 2010 through 31 December 2010 (includes aftershocks). Included are the test earthquakes given in Table 1 as well as background earthquakes with $M \geq 2.0$. More details in the square region are given in the larger scale maps in Figure 3.

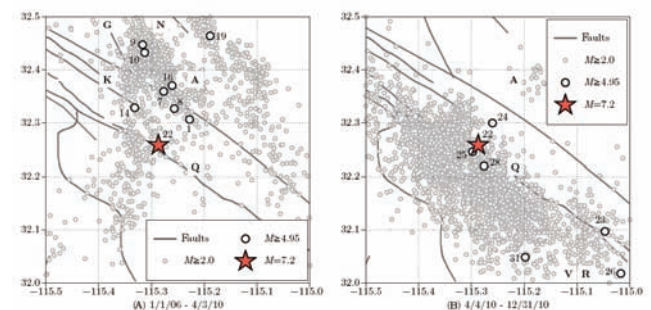


Figure 3. Map of the region in the immediate vicinity of the epicenter of the $M=7.2$ El Mayor-Cucapah earthquake. (A) Earthquakes during the period 1 January 2006 through 3 April 2010. (B) Earthquakes during the period 4 April 2010 through 31 December 2010. Included are the test earthquakes given in Table 1 as well as background earthquakes with $M \geq 2.0$. The association of lettered $0.1^\circ \times 0.1^\circ$ cells in which earthquakes occurred with the numbered earthquakes is illustrated.

quakes certainly cannot be considered foreshocks, due to their relatively small magnitudes and early occurrence, but may represent a seismic activation. Foreshocks are by definition main shocks followed by a larger triggered aftershock.¹⁵ Thus foreshocks have magnitudes that are statistically close to the following main shock.¹⁶

The locations of the earthquakes given in Table 1 identify the $0.1^\circ \times 0.1^\circ$ cells in which the earthquakes occurred. These cells are illustrated in Figure 3. Cells in which earthquakes occurred are identified by capital letters. Earthquakes in Figure 3A occurred in cells A, G, N, K, and Q. Earthquakes in Figure 3B occurred in cells A, Q, R, and V. The association of earthquake event numbers with cell letters is given in Table 2. The occurrence of five test earthquakes in cell A is not surprising since this is the Cerro Prieto geothermal area that is recognized as having a high level of seismic activity.

We next turn to the somewhat larger region ($3.0^\circ \times 2.5^\circ$) illustrated in Figure 2. The El Mayor earthquake and the test earthquakes that occurred later, 4 April 2010 to 31 December 2010 are given in Figure 2B. The aftershock region of the El Mayor-Cucapah earthquake is clearly illustrated, and events 27 and 29 are almost certainly aftershocks. Event 30 may or may not be an aftershock. The El Mayor earthquake and the test earthquakes that occurred earlier, 1 January 2006 to 3 April 2010 are given in Figure 3A. During this period no test earthquakes occurred outside the smaller region considered in Figure 3A.

We next consider the $2^\circ \times 1.4^\circ$ region adjacent to Cape Mendocino, illustrated in Figure 4. Six test earthquakes occurred in this region (events 2, 3, 4, 5, 20, and 21) in the magnitude range 5.0 to 6.5. This is a region of high seismicity, and this concentration of events is expected. Event 21 may or may not be an aftershock of event 20.

There were seven test earthquakes that occurred outside of the regions considered above. These are illustrated in Figure 1, and their magnitudes ranged from 5.0 to 5.45. The pair of earthquakes #17 and #18 are very close in location, magnitude, and time of occurrence. It is very likely that the $M=5.0$ earthquake on 1 October 2009 was a foreshock of the $M=5.19$ earthquake on 3 October 2009.

Submitted forecasts

The submitted forecasts have been discussed in some detail.¹⁴ The nineteen forecasts submitted by eight groups are available on the RELM website (<http://relm.csepesteing.org/>). In order to have a common basis for comparison, we will only consider forecasts that cover the entire test region. Thirteen forecasts were submitted that gave forecast numbers, N_{emi} , of $M \geq 4.95$ earthquakes in 0.1 magnitude bins during the five year test period for all $N_c = 7682$ $0.1^\circ \times 0.1^\circ$ cells.

The submitted forecasts are based on a variety of approaches. The Bird and Liu forecast¹⁷ was based on a kinematic model of neotectonics. The Ebel *et al.* forecast¹⁸ was based on the average rate of $M \geq 5$ earthquakes in $3^\circ \times 3^\circ$ cells for the period 1932 to 2004. The Helmstetter *et al.* forecast¹⁹ was based on the extrapolation of past seismicity. The Wiemer and Schorlemmer forecast²² was based on the asperity-based likelihood model. Ward²¹ submitted six forecasts. His seismicity submission was based on the extrapolation of past seismicity (Ward seis.), his geologic submission was based on fault slip data (Ward geol.). Two geodetic based simulations were made, one with a maximum magnitude of 8.1 (Ward geod.) and one with a maximum magnitude of 8.5 (Ward geol. 8.5). His simulation submission was based on a fault-based simulation of earthquakes in California (Ward sim.). His final submission was an average of the first three submissions (Ward combo).

We will now discuss the Holliday *et al.*²⁰ forecast in somewhat greater detail. The basis of this RELM forecast followed the format introduced in the PI forecast methodology.^{8,23} The magnitude range $M \geq 5$ and the cell dimensions $0.1^\circ \times 0.1^\circ$ were the same. However, the PI method was alarm based. Earthquakes were forecast to either occur or not occur in specified regions (hot spots) in a specified time period. In the PI based

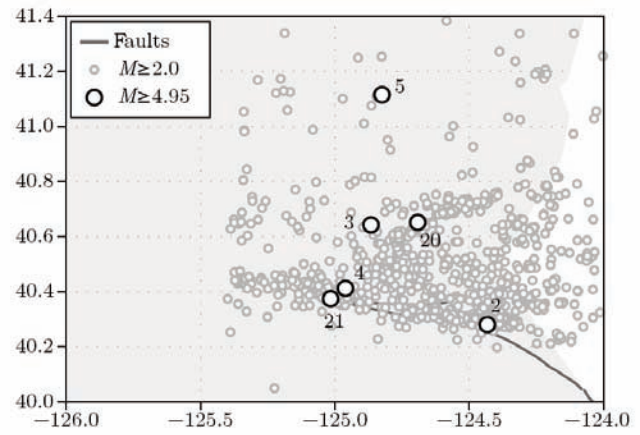


Figure 4. Map of the northwest region near Cape Mendocino. Test earthquakes given in Table 1 are shown as well as background earthquakes with $M \geq 2.0$.

RELM forecast, all hot spot cells are given equal probabilities of an earthquake. Instead of being alarm based, the RELM test was based on numbers of occurrence of earthquakes in each cell in the test region. This required a continuous assessment of earthquake occurrence rather than a binary, alarm-based assessment. To do this, the Holliday *et al.*²⁰ forecast introduced a uniform probability of occurrence for hotspot regions and added smaller probabilities for non-hotspot regions based on the RI of seismicity in the region. As a result the distribution of risk in this forecast was very different from the other forecasts. We will quantify this difference in a later section.

As stated in our description of the RELM test, each participant submitted the forecast for the number of earthquakes N_{emi} in magnitude bin m that would occur in cell i . Thus $41 \times 7682 = 314,962$ values of N_{emi} were submitted in each forecast. In this paper we emphasize that there are two aspects to the RELM forecasts: i) How many earthquakes will occur in the test region during the test period? ii) Where will the forecast earthquakes occur? The number of earthquakes N_e expected is certainly variable but an extrapolation from past seismicity is straightforward. We believe the primary focus of earthquake forecasting is to specify the spatial risk of an earthquake. In this paper we focus our attention on the conditional probability λ_{emi} that a test earthquake will occur in spatial cell i with a magnitude in magnitude bin m .

The submitted forecasts give the number N_{emi} of earthquakes forecast to occur in spatial cell i and in magnitude bin m . The sum of the N_{emi} over all $N_c = 7682$ cells is the total number of earthquakes N_{em} forecast to occur in magnitude bin m during the test period

$$N_{em} = \sum_{i=1}^{N_c} N_{emi}. \quad (2)$$

The total number of earthquakes N_e forecast to occur during the test period is given by

$$N_e = \sum_{m=5}^9 N_{em}. \quad (3)$$

The number of earthquakes N_{em} forecast to occur in magnitude bins

Table 2. Values of the submitted conditional probabilities λ_{zmi} that an earthquake that has a magnitude in spatial cell z . Values are given for the 31 test earthquakes listed in Table 1. The first column is the magnitude bin m , the second column is the spatial cell z , and the third column is the earthquake number. Forecast conditional probabilities λ_{zmi} are given for 13 forecasts. These are Bird and Liu (B-L), Ebel *et al.* main shocks+aftershocks (Ebel) and main shock only (Ebel ms), Helmstetter *et al.* main shocks+aftershocks (Helm) and main shock only (Helm ms), Holliday *et al.* (Holl), Ward seismic model with $M_{max}=8.1$ (Ward seis), Ward geologic model with $M_{max}=8.1$ (Ward geol.), Ward geotectic model with $M_{max}=8.1$ (Ward geot), Ward geotectic model with $M_{max}=8.1$ (Ward geot), Ward geotectic model with $M_{max}=8.1$ (Ward geot), Ward combined model with $M_{max}=8.1$ (Ward combo), Ward geotectic model with $M_{max}=8.5$ (Ward geod. 8.5), Ward simulation model (Ward sim.), Wiemer and Schorlemmer (W-S). The highest (best) forecast probabilities are in *italics*. The number of winning forecasts and the mean forecast probabilities for all test earthquakes λ_{zmi} are given for each forecast. Also given are the total number of earthquakes forecast N_z .

m	z	No.	B-L ¹⁷	Ebel ¹⁸	Ebel ms ¹⁸	Helm ¹⁹	Helm ms ¹⁹	Holl ²⁰	Ward seis. ²¹	Ward geol. ²¹	Ward combo ²¹	Ward geod. 8.5 ²¹	Ward sim. ²¹	W-S ²²	
5	E	5	3.29e-04	4.99e-05	1.04e-04	1.04e-04	1.04e-04	4.42e-06	5.41e-14	2.94e-11	7.24e-05	4.20e-05	7.24e-05	1.16e-29	7.09e-09
5	B	2	6.45e-04	1.55e-03	3.27e-03	3.27e-03	3.27e-03	1.51e-03	6.93e-11	2.94e-11	8.44e-05	4.90e-05	8.44e-05	1.16e-29	2.49e-03
5	H	11	1.58e-05	2.49e-04	4.16e-05	4.16e-05	4.16e-05	7.36e-06	1.12e-07	3.29e-19	2.95e-05	1.72e-05	2.95e-05	1.16e-29	1.22e-05
5	M	17	7.73e-05	9.98e-05	2.63e-04	2.63e-04	2.63e-04	1.51e-03	1.54e-04	1.00e-04	2.10e-04	1.76e-04	2.10e-04	2.82e-06	3.39e-05
5	S	27	9.77e-04	2.49e-04	5.74e-04	5.74e-04	5.74e-04	1.51e-03	6.32e-04	4.17e-04	3.61e-04	4.34e-04	3.61e-04	1.11e-03	6.44e-04
5	K	14	1.43e-03	9.98e-04	1.50e-03	1.50e-03	1.50e-03	1.51e-03	1.49e-03	1.19e-03	5.32e-04	8.76e-04	5.32e-04	3.06e-04	2.80e-03
5	G	9,10	4.19e-04	2.49e-04	1.16e-03	1.16e-03	1.16e-03	1.51e-03	1.86e-03	1.37e-03	5.78e-04	1.02e-03	5.78e-04	5.13e-04	1.23e-03
5	V	31	3.09e-04	6.99e-04	6.99e-04	2.52e-04	2.52e-04	1.51e-03	8.89e-04	6.23e-04	6.99e-04	6.99e-04	6.99e-04	3.41e-05	1.18e-04
5.1	L	15	9.54e-05	2.49e-04	3.15e-04	3.15e-04	3.15e-04	1.51e-04	2.51e-04	8.76e-05	2.48e-04	2.18e-04	2.48e-04	7.76e-06	1.49e-04
5.1	A	7,8,16	9.08e-04	9.98e-04	5.30e-03	5.30e-03	5.30e-03	1.51e-03	1.14e-03	1.31e-03	5.82e-04	8.50e-04	5.82e-04	8.99e-05	5.92e-03
5.2	D	4	1.60e-03	1.50e-03	3.17e-03	3.17e-03	3.17e-03	1.51e-03	4.82e-12	4.45e-11	7.43e-05	4.32e-05	7.43e-05	1.02e-29	1.77e-03
5.2	M	18	7.82e-05	9.98e-05	2.63e-04	2.63e-04	2.63e-04	1.51e-03	1.54e-04	1.00e-04	1.76e-04	1.76e-04	1.76e-04	3.95e-05	2.89e-05
5.3	Q	28	1.31e-03	9.98e-04	1.29e-03	1.29e-03	1.29e-03	1.51e-03	9.25e-04	1.05e-03	5.91e-04	7.55e-04	5.91e-04	7.89e-05	2.40e-03
5.3	R	26	1.40e-03	6.99e-04	6.99e-04	6.52e-04	6.52e-04	7.85e-06	1.25e-03	8.27e-04	6.27e-04	8.08e-04	6.27e-04	1.85e-05	7.26e-04
5.4	C	3	3.34e-04	2.99e-04	2.99e-04	3.37e-04	3.37e-04	1.51e-03	1.61e-12	6.74e-11	6.99e-05	4.06e-05	6.99e-05	5.69e-30	3.65e-04
5.4	I	12	9.72e-05	4.99e-05	1.66e-04	1.66e-04	1.66e-04	9.33e-06	2.29e-11	6.74e-11	8.91e-05	5.18e-05	8.91e-05	5.69e-30	4.56e-04
5.4	F	6	4.26e-04	1.30e-03	1.40e-03	1.40e-03	1.40e-03	1.51e-03	1.13e-04	4.30e-09	3.69e-04	2.41e-04	3.69e-04	8.27e-09	2.11e-03
5.4	J	13	7.64e-05	3.99e-04	5.06e-04	5.06e-04	5.06e-04	1.51e-03	3.63e-04	7.96e-04	2.33e-04	3.69e-04	2.33e-04	1.08e-03	5.30e-04
5.4	U	30	5.72e-04	1.40e-03	4.68e-03	4.68e-03	4.68e-03	1.51e-04	1.25e-03	1.60e-03	2.48e-04	7.34e-04	2.48e-04	5.39e-03	2.41e-03
5.4	Q	25	1.31e-03	9.98e-04	1.29e-03	1.29e-03	1.29e-03	1.51e-03	9.25e-04	1.05e-03	5.91e-04	7.55e-04	5.91e-04	2.26e-04	2.41e-03
5.4	A	1,24	9.06e-04	9.98e-04	5.30e-03	5.30e-03	5.30e-03	1.51e-03	1.14e-03	1.31e-03	5.82e-04	8.50e-04	5.82e-04	4.24e-04	5.78e-03
5.5	R	23	1.39e-03	6.99e-04	6.99e-04	6.52e-04	6.52e-04	7.85e-06	1.25e-03	8.27e-04	6.27e-04	8.08e-04	6.27e-04	5.64e-04	7.33e-04
5.7	T	29	8.25e-04	5.99e-04	1.11e-03	1.11e-03	1.11e-03	1.51e-03	4.72e-04	4.09e-04	1.77e-04	2.89e-04	1.77e-04	9.06e-04	2.32e-03
5.8	N	19	2.62e-03	2.99e-04	6.78e-04	6.78e-04	6.78e-04	1.51e-03	8.94e-04	1.17e-03	5.59e-04	7.52e-04	5.59e-04	3.05e-03	2.81e-04
5.9	P	21	2.97e-04	1.50e-03	1.23e-03	1.23e-03	1.23e-03	1.51e-03	3.87e-12	1.90e-10	7.06e-05	4.10e-05	7.06e-05	1.34e-29	2.73e-04
6.5	O	20	5.64e-04	6.49e-04	4.30e-04	4.30e-04	4.30e-04	1.51e-03	3.23e-12	6.59e-10	4.23e-05	4.23e-05	4.23e-05	3.52e-29	7.30e-04
7.2	Q	22	1.27e-03	9.98e-04	1.29e-03	1.29e-03	1.29e-03	1.51e-03	9.25e-04	1.05e-03	5.91e-04	7.55e-04	5.91e-04	3.12e-04	2.12e-03
Number of wins			3	1	3	3	3	8	1	0	0	0	0	2	9
λ_{zmi}			7.51e-04	6.99e-04	1.38e-03	1.38e-03	1.38e-03	1.13e-03	5.95e-04	5.67e-04	3.38e-04	4.41e-04	3.38e-04	5.24e-04	1.44e-03
Ne (=31.00 actual)			55.90	114.75	28.19	35.40	21.13	30.00	22.20	18.13	32.08	32.08	32.08	8.20	23.71

during the test period are given in Table 2. The discussion of these values and their relation to the actual number of test earthquakes that occurred $N_c=31$ will be given in the next section.

The forecast conditional probability λ_{emi} that a test earthquake will occur in spatial cell i with a magnitude in magnitude bin m is given by

$$\lambda_{emi} = \frac{N_{emi}}{N_{em}} \quad (4)$$

From Eqs. (2) and (4) we see that

$$\sum_{i=1}^{N_c} \lambda_{emi} = 1. \quad (5)$$

For all submitted forecasts the sum of the conditional probabilities over all cells is one. Thus the forecasts differ only in the allocation of the conditional probability between the $N_c=7682$ spatial cells. This allocation will be discussed in some detail in a later section.

The forecast conditional probabilities λ_{emi} for the spatial cells in which test earthquakes occurred are given in Table 2. For each of the 31 earthquakes (identified by number in Table 1) the appropriate cell i and magnitude bin m are given. The lettering of i for cells has been illustrated in Figure 3. Note that earthquakes 9 and 10 occurred in magnitude bin $m=5$ and cell $i=G$, earthquakes 7, 8 and 16 occurred in magnitude bin $m=5.1$ and cell $i=A$. Thus we consider 27 cells i and magnitude bins m which we refer to as mi cells. Ebel *et al.*¹⁸ and Helmstetter *et al.*¹⁹ submitted separate forecasts for all earthquakes and for only main shocks. From Table 2 we see that the forecast numbers N_{em} are substantially higher when aftershocks are considered. However, from Table 2 we see that the conditional probabilities λ_{emi} are identical with and without aftershocks. Because of our rescaling approach, we eliminate the difference between these two types of forecasts in terms of forecast locations.

It would be desirable to identify whether the single forecasts were for all earthquakes or for only main shocks. Unfortunately the submissions were ambiguous on this subject. Based on the equality of the values of λ_{emi} for the Ebel *et al.*¹⁸ and Helmstetter *et al.*¹⁹ forecasts we will hypothesize that the results for the other

forecasts are not significantly dependent on whether they were for all earthquakes or for main shocks only.

Evaluation of Results

During the formulation of the RELM project a comprehensive testing strategy was also developed.²⁴ A suite of likelihood tests were proposed which would be implemented through a testing center.²⁵ The approach utilized an L-test, N-test, and R-test. These tests were applied to the raw submitted data. This approach was applied to the first 2½ years of RELM results by Schorlemmer *et al.*²⁶ Zechar *et al.*²⁷ recognized a problem with the original proposed likelihood tests and proposed a modification. We note that in the testing approach given by Schorlemmer *et al.*²⁴ it was suggested that the declustering algorithm given by Reasenberg²⁸ be used to separate aftershocks from main shocks. Unfortunately the single submission forecasts were not clearly defined to include all earthquakes or just main shocks.

The primary purpose of this paper is to present a complementary approach. Our approach has the advantage that the evaluation of the

numbers of earthquakes forecast can be separated from the forecast of their locations. A preliminary version of our approach has been given by Lee *et al.*²⁹

We first consider the forecasts of the number of test earthquakes that would occur during the test period given in Table 2. The total number of test earthquakes was 31. Based on the discussion given in section 4 we concluded that 9-11 earthquakes were aftershocks, which means there were 20-23 main shocks. Thus there were 29-35% aftershocks and 65-71% main shocks. For the two sets of forecasts that distinguished aftershocks, Ebel *et al.*¹⁸ had 86 (75%) aftershocks and 29 (25%) main shocks and Helmstetter *et al.*¹⁹ had 14 (40%) aftershocks and 21 (60%) main shocks. The latter values were reasonably close to the actual values. For the single submissions Holliday *et al.*²⁰ forecast 30 earthquakes and Wiemer and Schorlemmer²² forecast 24 earthquakes, both reasonable values. The 55 earthquakes forecast by Bird and Liu¹⁷ was high and Ward²¹ had values from 8 to 56 earthquakes for his 6 forecasts.

The probability λ_{emi} is the forecast conditional probability that a test earthquake will occur in spatial cell i and magnitude bin m . These probabilities, given in Table 2, can be used to compare the spatial aspect of RELM forecasts. We first discuss two aspects of the forecast values of λ_{emi} :

i) Values of λ_{emi} have been given for both main shock and main shock plus aftershock forecasts by Ebel *et al.*¹⁸ and Helmstetter *et al.*¹⁹ In both cases the values of λ_{emi} are identical with and without aftershocks. Thus 13 submitted forecasts are reduced to 11 when considering forecast locations.

ii) It is of interest to compare the forecast probabilities for two earthquakes that occur in a cell with different magnitudes. As a specific example we consider two earthquakes that occurred in cell Q, #22 with $m=7.2$ and #28 with $m=5.3$. For seven of the 11 forecasts the values of λ_{emi} are identical. The others have relatively close values. By normalizing the conditional probabilities with the number of earthquakes N_{em} in magnitude bin m we have isolated the frequency-magnitude (b value) statistics.

We now address the question, which forecast is best at specifying the location of future earthquakes. As a specific example, we consider the $M=7.2$ El Mayor-Cuicapa earthquake (event #22). The 13 submitted values of λ_{emi} are given in Table 2. They range from $\lambda_{emi}=2.12 \times 10^{-3}$ for the Wiemer and Schorlemmer²² forecast to $\lambda_{emi}=3.12 \times 10^{-4}$ for the Ward²¹ simulation forecast. The Wiemer and Schorlemmer²² forecast was the best forecast in that it gave the highest probability of occurrence in this spatial cell and magnitude bin. Using these criteria the best forecasts are highlighted in bold in Table 2.

There were three spatial-magnitude bins that had multiple earthquakes. These were G-5 with two (earthquakes 9 and 10), A-5.1 with three (earthquakes 7, 8, and 16) and A-5.4 with two (earthquakes 1 and 24). Earthquakes occurred in 27 spatial magnitude bins. Wiemer and Schorlemmer²² forecast the largest λ_{emi} for 9 bins, Holliday *et al.*²⁰ for 8 bins, Helmstetter *et al.*¹⁹ and Bird and Liu¹⁷ for 3 bins, Ward²¹ simulation model 2 bins, Ward²¹ seismic model 1 bin, and Ebel *et al.*¹⁸ 1 bin.

There are other ways to evaluate the results of the forecasts. One forecast might do very well (high λ_{emi}) for some test earthquakes and do poorly (low λ_{emi}) for other test earthquakes. The overall validity of a forecast can be quantified using the mean forecast probability $\bar{\lambda}_{emi}$ for the 27 spatial-magnitude bins. These values are also given in Table 2. The best overall forecast by this measure was Wiemer and Schorlemmer²² with $\bar{\lambda}_{emi} = 1.44 \times 10^{-3}$ and Helmstetter *et al.*¹⁹ with $\bar{\lambda}_{emi} = 1.38 \times 10^{-3}$. It is also of interest to compare the submitted forecasts to a random forecast. Consider an earthquake in magnitude bin m , the sum of all conditional probabilities λ_{emi} is unity as given by Eq.

(5). A random forecast would give equal probabilities to all cells λ_{emi}^{rand} . From Eq. (5) this value is given by

$$\lambda_{emi}^{rand} = \frac{1}{N_c} = \frac{1}{7682} = 1.30 \times 10^{-4}. \quad (6)$$

The best mean forecasts exceed this value by about a factor of 50. All the winning forecasts given in Table 2 exceed this random forecast.

Distribution of cell probabilities

The basic purpose of this paper is to better understand the physics and statistics of earthquake forecasts. We have concentrated our discussion on the relative probabilities of where earthquakes will occur. In order to do this we have introduced the conditional probability λ_{emi} that an earthquake in magnitude bin m will lie in spatial cell i . The sum of λ_{emi} over all bins i is unity. Thus the allocation of conditional probabilities λ_{emi} between cells is the essential feature of a successful forecast. To explore this we will consider the forecast probability λ_{ei} that an earthquake with magnitude $m \geq 4.95$ will occur in cell i . The definition of this probability is

$$\lambda_{ei} = \frac{\sum_m N_{emi}}{\sum_i \sum_m N_{emi}}. \quad (7)$$

Once again we have

$$\sum_i \lambda_{ei} = 1. \quad (8)$$

The values of the probability λ_{ei} are similar to the values of the conditional probability λ_{emi} , but the sum over magnitude bins eliminates the weak dependence of λ_{emi} on m due to different values of the b -value used in the forecasts. We rank these probabilities λ_{ei} in a forecast from the highest to the lowest. The highest forecast probability is N_1 and the lowest forecast probability is N_{7682} .

In Figure 5 we give the distribution of forecast probabilities λ_{ei} for the forecasts of Bird and Liu,¹⁷ Ebel *et al.*,¹⁸ Helmstetter *et al.*,¹⁹ Holliday *et al.*,²⁰ and Wiemer and Schorlemmer *et al.*²² are given in Figure 5. The areas under the curves are equal to 1. For the highest probabilities $0 < N_c < 100$ the forecasts of Helmstetter *et al.* and Wiemer and Schorlemmer forecast the highest probabilities λ_{ei} , however the forecasts of Wiemer and Schorlemmer are slightly higher because they gave reduced probabilities for large N_c . In the range of $100 < N_c < 637$ the forecast by Holliday *et al.* gave the highest probabilities of occurrence. This behavior can be attributed to the alarm basis of the PI forecast. The highest probability cells, $0 < N_c < 637$ were given the same probability values. For the range $637 < N_c < 7682$ the forecast probabilities were much lower. In the range $637 < N_c < 4000$ the forecast probabilities given by Bird and Liu and Ebel *et al.* were the highest and were approximately equal. For the range $400 < N_c < 7682$ the forecast probabilities given by Bird and Liu were the highest. The forecast by Wiemer and Schorlemmer gave the largest range of values and the forecast by Bird and Liu gave the smallest range of values. If a range of low probabilities are given in order to enhance the values of the highest probabilities the risk is that an unexpected earthquake will occur in the cells with a low forecast probability. An example is the forecast (Table 2) $\lambda_{emi} = 7 \times 10^{-9}$ given by Wiemer and Schorlemmer²² for test earthquake #1. The highest forecast for this earthquake was $\lambda_{emi} = 7.29 \times 10^{-4}$ by Bird and Liu.¹⁷

A no skill forecast would assign a probability $\lambda_{emi}^{rand} = 7.29 \times 10^{-4}$ to all cells (Eq. 6). Out of the 7682 cells the Helmstetter *et al.*¹⁹ and the Wiemer and Schorlemmer²² forecasts had 1400 cells with higher than no skill probabilities, Bird and Liu¹⁷ had 1900 and Ebel *et al.*¹⁸ 2100.

As we have previously discussed earthquake forecasts can be either probabilistic or alarm based. The submission rules for RELM were probabilistic. The only forecast that had an alarm based distribution of forecasts was Holliday *et al.*²⁰ In the probabilities λ_{emi} listed in Table 2, the hot spot (alarm) cells had values $\lambda_{emi} = 1.15 \times 10^{-3}$. Of the 27 cells in which earthquakes occurred, 20 occurred in hot spot cells. In 8 of the 20 cells, the hot spot forecasts had the highest probabilities of occurrence. The hot spot cells comprised 8.3% of the test region (637 of the 7682 cells). This alarm-based behavior is clearly illustrated in Figure 5.

Discussion

The RELM test provided a well-defined set of prospective earthquake forecasts and a well-defined set of test earthquakes. In this paper we present a method for evaluating the RELM forecasts. We believe our approach has significant advantages but look forward to comparing our results with those obtained by other authors.

RELM forecasts provide the numbers N_{emi} of earthquakes expected to occur in magnitude bins m and spatial cells i . The basis of our approach is:

- i) To use Eq. (2) to determine the forecast number N_{em} of earthquakes expected to occur in magnitude bin m .
- ii) To use Eq. (4) to determine the conditional probability λ_{emi} that an earthquake with magnitude in magnitude bin m will occur in cell i . In addition Eq. (4) is used to determine the total number N_e of forecast earthquakes.

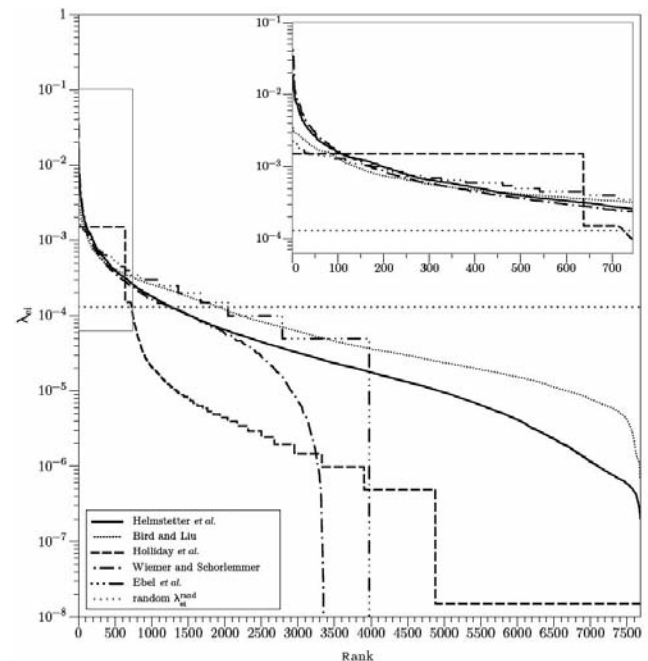


Figure 5. Distribution of forecast probabilities λ_{ei} that an earthquake with $m \geq 4.95$ will occur in cell i . The 7682 forecast cell probabilities are ranked from highest to lowest. In each forecast the sum of the probabilities is unity. The no skill forecast is also included.

The conditional probability λ_{emi} is the forecast probability that an earthquake with magnitude m will occur in cell i . The sum of the λ_{emi} over all cells is unity. The allocation of the λ_{emi} to cells is the forecast of where earthquakes are expected to occur. When separate forecasts were submitted for all earthquakes and for only main shocks the values of λ_{emi} were identical. In addition the values of λ_{emi} were either independent of m or only weakly dependent on m .

A random (no skill) forecast would have assigned equal values of λ_{emi} to all cells. The success (skill) of a forecast is measured by the forecasters ability to assign large values of λ_{emi} to the cells where test earthquakes occurred and small values where they did not occur.

During the test period 31 earthquakes occurred with $M > 4.95$. These earthquakes occurred in 27 different combinations of spatial cell i and magnitude bin m as shown in Table 2. The largest forecast values of λ_{emi} for each of these 27 values of m and i were the best forecasts, these are highlighted in Table 2. The largest values ranged from $\lambda_{emi} = 5.92 \times 10^{-3}$ for earthquakes 7, 8 and 16 in magnitude bin $m = 5.1$ and in cell A to $\lambda_{emi} = 2.49 \times 10^{-4}$ for earthquake 11 in magnitude bin $m = 5$ and cell H. All the largest values exceeded the random (no skill) value $= 1.30 \times 10^{-4}$ given in Eq. 6. One measure of the *best* forecast is the forecast with the largest number of highest λ_{emi} . From Table 2 the winner by this measure is the Wiemer and Schorlemmer²² forecast with 9 highest values followed by Holliday *et al.*²⁰ with 8 highest values. An alternative measure of the *best* forecast is the highest mean value of λ_{emi} for the 27 values of m and i . The values of for the forecasts are given in Table 2. The winner by this measure is again the forecast by Wiemer and Schorlemmer²² with $\bar{\lambda}_{emi} = 1.44 \times 10^{-3}$ followed by Helmstetter *et al.*¹⁹ with $\bar{\lambda}_{emi} = 1.38 \times 10^{-3}$. These values are about one order of magnitude better than the random (no skill) value $\lambda_{emi}^{rand} = 1.30 \times 10^{-4}$.

The success of a RELM forecast is dependent on the allocation of the probabilities λ_{emi} between the 7682 cells. In Figure 5 we give the distribution of forecast probabilities λ_{ei} for five forecasts. The variability is clearly illustrated. It is interesting to compare the forecast by Wiemer and Schorlemmer²² to the forecast by Helmstetter *et al.*¹⁹ Wiemer and Schorlemmer²² had slightly higher values of λ_{ei} in the high probability regions. This is the reason that Wiemer and Schorlemmer²² had the best forecasts for earthquakes 1, 6, 7, 8, 14, 16, 22, 24, 25 and 29. In order to have these higher values of λ_{ei} , Wiemer and Schorlemmer²² had very low values of λ_{ei} in low probability regions. The consequence of the balance was the very poor forecast ($\lambda_{ei} = 7.09 \times 10^{-9}$) for earthquake number 5. It is a matter of choice how this poor forecast should be penalized.

The forecast by Holliday *et al.*²⁰ differed from the other forecasts since it was alarm based. This is clearly seen in Figure 5 where the 637 high probability cells had equal forecast probabilities, $\lambda_{ei} = 1.51 \times 10^{-3}$. The consequence was that this forecast was not highest in high probability regions but was the highest in moderate probability regions. Thus this forecast had the highest probabilities ($\lambda_{ei} = 1.51 \times 10^{-3}$) for earthquakes 3, 13, 17, 18, 20, 21, 27 and 31. Overall 23 of the 31 earthquakes occurred in hot spot regions that included 637 out of 7682 (8.3%) of the cells. However, the Holliday *et al.*²⁰ forecast had very low forecast probabilities (2.9×10^{-6}) for four of the earthquakes.

Conclusions

In summary we conclude that the RELM test was extremely useful in providing an understanding of the trade offs in forecasting the locations of future earthquakes. The RELM forecasts were primarily based on the extrapolation of the rates of earthquake activity to forecast

where future earthquakes occur. The results quantified the statistical validity of this approach. It should be noted that there were two important limitations to the RELM test approach. The first is that prospective test earthquakes have considerable statistical variability. The RELM test earthquakes were dominated by events associated with the $M = 7.2$ El Mayor-Cucapah earthquake. The second is the relevance of the test earthquakes to the occurrence of larger earthquakes. The minimum magnitude earthquake in the RELM test was $M = 4.95$. From Table 1 we see that 29 of the 31 earthquakes have magnitudes $M < 6$. This activity certainly correlated quite well with pre-existing background seismicity. Earthquakes with $M > 7.0$ are likely to occur on mapped faults. Many of these faults (*i.e.* the San Andreas) have low levels of seismicity on them. The implications of this low level of seismicity are not clear.

References

1. Bufe CG, Varnes DJ. Predictive modeling of the seismic cycle of the greater San Francisco Bay region. *J Geophys Res* 1993;98:9871-83.
2. Bowman DD, Ouillon G, Sammis CG, et al. An observational test of the critical earthquake concept. *J Geophys Res* 1998;103:24359-72.
3. Hardebeck JL, Fetzner KR, Michael AJ. Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. *J Geophys Res* 2008;113:B08310.
4. Keilis-Borok VI. The lithosphere of the earth as a nonlinear system with implications for earthquake prediction. *Rev Geophys* 1990;28:19-34.
5. Keilis-Borok VI, Kossobokov VG. Premonitory activation of earthquake flow-algorithm m8. *Phys Earth Planet Interiors* 1990;61:73-83.
6. Kossobokov VG, Shebalin P. Earthquake prediction. In: Keilis-Borok VI, Soloviev AA, (eds). *Nonlinear dynamics of the lithosphere and earthquake prediction*. Berlin: Springer-Verlag; 2003. pp. 141-207.
7. Minster JB, Williams N. Systematic global testing of intermediate-term earthquake prediction algorithm. In: Mora P, (ed). *1st ACES Workshop Proceedings*. Brisbane: GOPRINT; 1999.
8. Rundle JB, Tiampo KF, Klein W, Martins JSS. 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* 2002;99 Suppl 1:2514-21.
9. Holliday JR, Rundle JB, Tiampo KF, Turcotte DL. Using earthquake intensities to forecast earthquake occurrence times. *Nonlinear Proc Geoph* 2006;13:585-93.
10. Rundle JB, Turcotte DL, Shcherbakov R, et al. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev Geophys* 2003;41:1019.
11. Shcherbakov R, Turcotte DL, Rundle JB, et al. Forecasting the locations of future large earthquakes: an analysis and verification. *Pure Appl Geophys* 2010;167:743-9.
12. Zechar JD, Jordan TH. Testing alarm-based earthquake predictions. *Geophys J Int* 2008;172:715-24.
13. Holliday JR, Nanjo KZ, Tiampo KF, et al. Earthquake forecasting and its verification. *Nonlinear Proc Geoph* 2005;12:965-77.
14. Field EH. Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seismol Res Lett* 2007;78:7-16.
15. Savage MK, dePolo DM. Foreshock probabilities in the western great-basin eastern Sierra Nevada. *B Seismol Soc Am* 1993;83:1910-38.
16. Helmstetter A, Sornette D. Foreshocks explained by cascades of triggered seismicity. *J Geophys Res* 2003;108:2457.
17. Bird P, Liu Z. Seismic hazard inferred from tectonics: California. *Seismol Res Lett* 2007;78:37-48.

18. Ebel JE, Chambers DW, Kafka AL, Baglivo JA. Non-poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California. *Seismol Res Lett* 2007;78:57-65.
19. Helmstetter A, Kagan YY, Jackson DD. High-resolution time-independent grid-based forecast for $m \geq 5$ earthquakes in California. *Seismol Res Lett* 2007;78:78-86.
20. Holliday JR, Chen CC, Tiampo KF, et al. A RELM earthquake forecast based on pattern informatics. *Seismol Res Lett* 2007;78:87-93.
21. Ward SN. Methods for evaluating earthquake potential and likelihood in and around California. *Seismol Res Lett* 2007;78:121.
22. Wiemer S, Schorlemmer D. ALM: an asperity-based likelihood model for California. *Seismol Res Lett* 2007;78:134-40.
23. Tiampo KF, Rundle JB, McGinnis S, et al. Eigen patterns in southern California seismicity. *J Geophys Res* 2002;107:2354.
24. Schorlemmer D, Gerstenberger MC, Wiemer S, et al. Earthquake likelihood model testing. *Seismol Res Lett* 2007;78:17-29.
25. Schorlemmer D, Gerstenberger MC. RELM testing center. *Seismol Res Lett* 2007;78:31-6.
26. Schorlemmer D, Zechar JD, Werner MJ, et al. First results of the regional earthquake likelihood models experiment. *Pure Appl Geophys* 2010;167:859-76.
27. Zechar JD, Gerstenberger MC, Rhoades DA. Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *B Seismol Soc Am* 2010;100:1184-95.
28. Reasenber PA. Second-order moments of central California seismicity 1969-1982. *J Geophys Res* 1985;90:479-95.
29. Lee YT, Turcotte DL, Holliday JR, et al. Results of the regional earthquake likelihood models (RELM) test of earthquake forecasts in California. *Proc Natl Acad Sci USA* 2011;108:16533-8.

Non-commercial use only