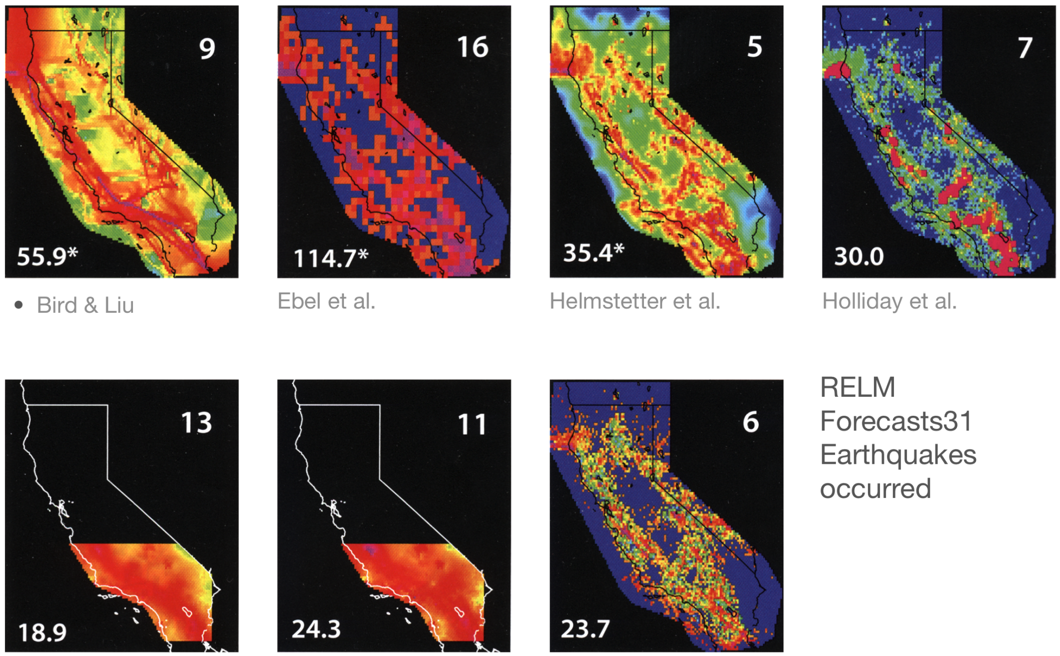
# evaluating earthquake forecasts*


| UC Davis | National Central University | Univ of Western Ontario | Boston University |
|---|---|---|---|
| Donald L. Turcotte | Ya-Ting Lee | Kristy F. Tiampo | W. Klein |
| James R. Holliday | Chien-Chih Chen | | |
| Michael K. Sachs | | | |
| John B. Rundle | | | |



* thanks for generous support from NASA

# RELM forecast comparison test
# January 1, 2006 – December 31, 2010




- Bird & Liu

Ebel et al.

Helmstetter et al.

Holliday et al.

Ward combined

Ward geodedic

Weimer & Schorlemmer

RELM
Forecasts31
Earthquakes
occurred

# RELM test rules

1. Test region: California plus adjacent regions

2. Test region divided into $N_c$ = 7682, 0.1° x 0.1° cells

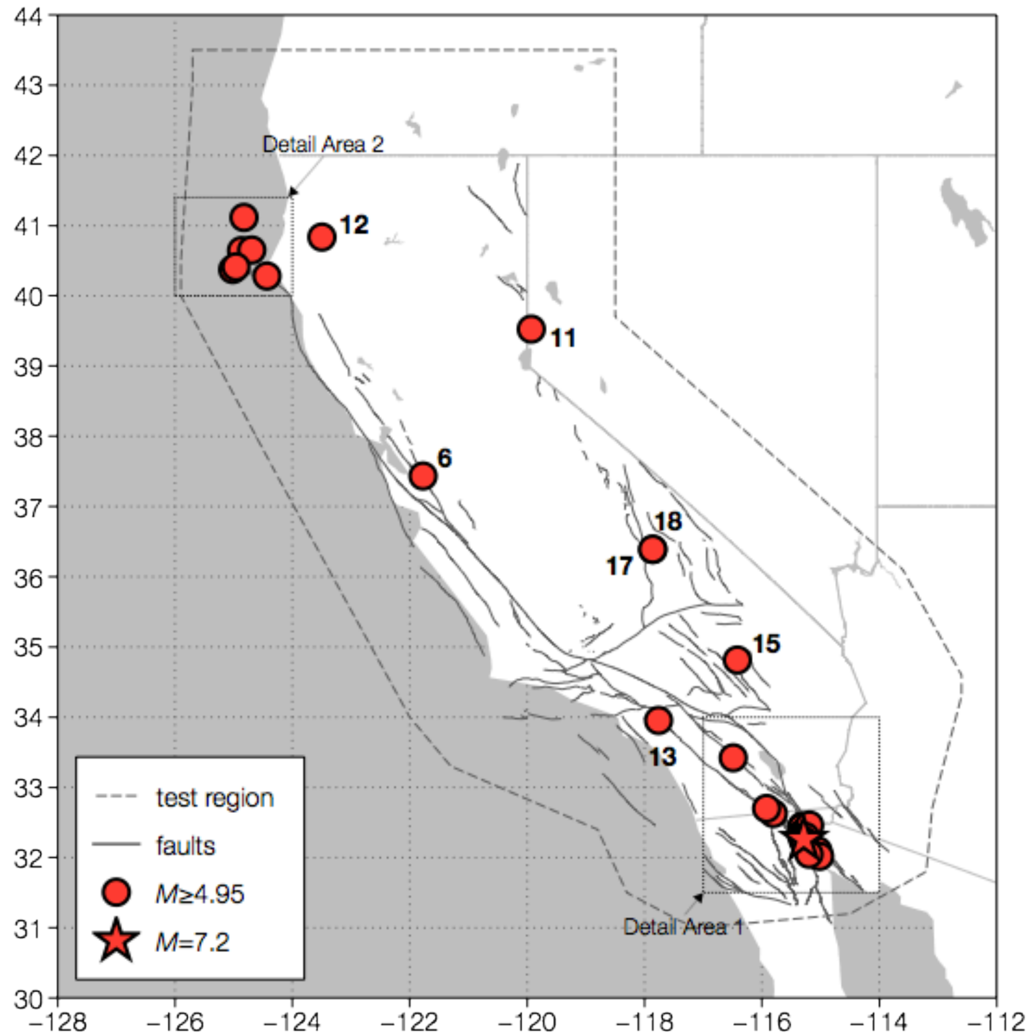
3. Magnitude range $M \geq 5$

4. Time period: 5 years 1/1/06 - 12/31/10

5. Participants submitted the probability $\lambda_{im}$ of the occurrence of an earthquake in cell i, probabilities were submitted for 41 magnitude bins with $\Delta M = 0.1$

# california: the test region & earthquakes

Map of the test region, the coast of California, major faults, and the 31 earthquakes with M ≥ 4.95 that occurred in the test region. Also shown are the square regions where large scale maps are given in subsequent slides.



# our approach

1. Remove magnitude dependence:

$$\lambda_i = \sum_{m=4.95}^{\infty} \lambda_{im}$$

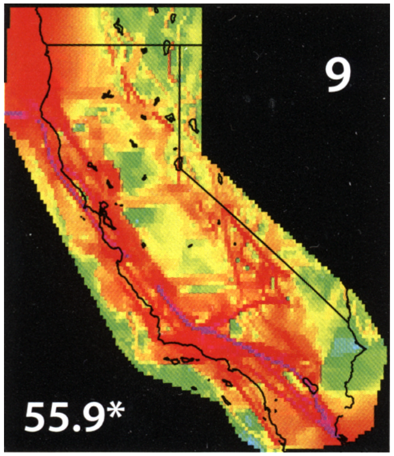
$\lambda_i$ is the probability of occurrence of a M ≥ 4.95 earthquake in cell i

2. Remove dependence on total number of earthquakes forecast (normalize all forecasts over area to the same value:

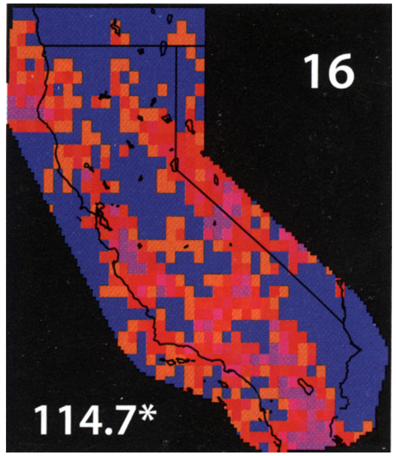
$$N_e = \sum_{i=1}^{7682} \lambda_i \qquad \lambda_{in} = \frac{N_{ce}}{N_e} \lambda_i$$

3. Consider only the 22 cells in which earthquakes occurred,and neglect the multiple occurrence of earthquakes in a cell: $N_{ce}$

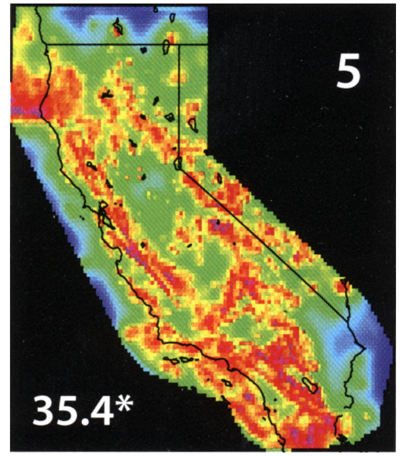
4. Introduce a normalized probability $\lambda_{in}$ (skill score) for the occurrence of an earthquake in cell i



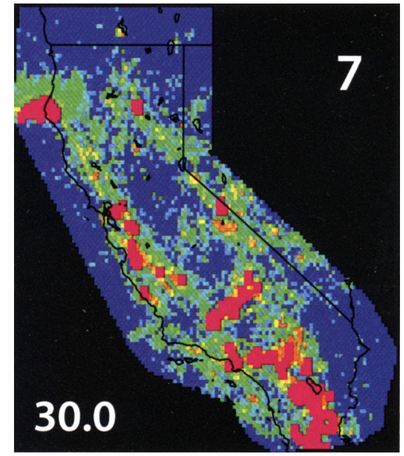

- Bird & Liu



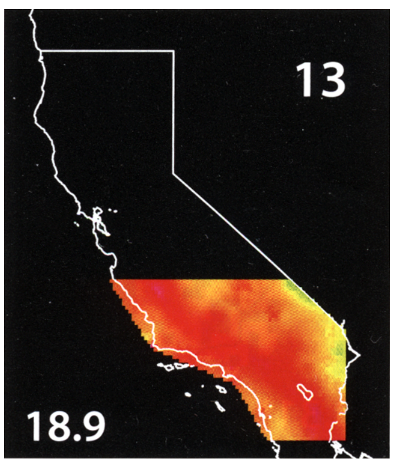

Ebel et al.



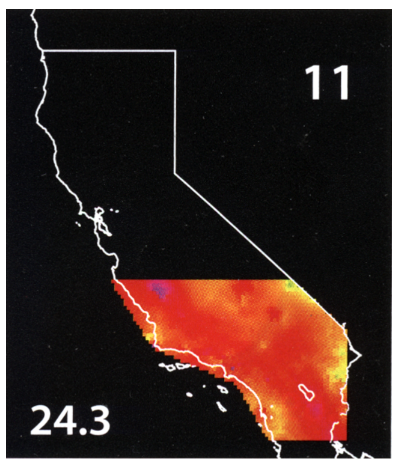

Helmstetter et al.



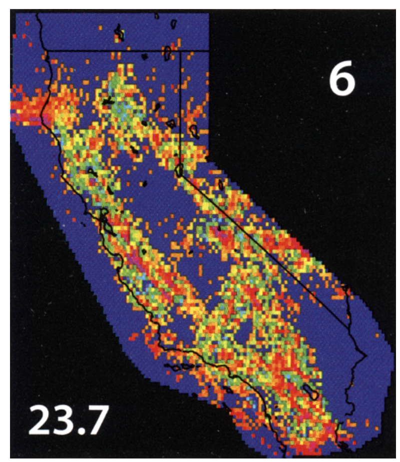

Holliday et al.




Ward combined




Ward geodedic




Weimer & Schorlemmer

RELM Forecasts

31 Earthquakes actually occurred

M>4.95

| Cell ID | EQ ID | B and L | Ebel | Helm. | Holl. | W-C | W-G | W and S |
|---|---|---|---|---|---|---|---|---|
| - A - | 1,7,8,16,24 | 1.99E-02 | 2.20E-02 | 1.17E-01 | 3.32E-02 | 1.87E-02 | 1.28E-02 | 1.24E-01 |
| - B - | 2 | 1.41E-02 | 3.40E-02 | 7.20E-02 | 3.32E-02 | 1.08E-03 | 1.86E-03 | 4.99E-02 |
| - C - | 3 | 7.40E-03 | 6.59E-03 | 7.41E-03 | 3.32E-02 | 8.93E-04 | 1.54E-03 | 7.91E-03 |
| - D - | 4 | 3.54E-02 | 3.29E-02 | 6.97E-02 | 3.32E-02 | 9.50E-04 | 1.64E-03 | 3.59E-02 |
| - E - | 5 | 7.23E-03 | 1.10E-03 | 2.29E-03 | 9.72E-05 | 9.25E-04 | 1.59E-03 | 1.58E-07 |
| - F - | 6 | 9.37E-03 | 2.85E-02 | 3.07E-02 | 3.32E-02 | 5.29E-03 | 8.12E-03 | 4.55E-02 |
| - G - | 9,10 | 9.11E-03 | 5.49E-03 | 2.55E-02 | 3.32E-02 | 2.25E-02 | 1.27E-02 | 2.38E-02 |
| - H - | 11 | 3.42E-04 | 5.49E-03 | 9.15E-04 | 1.62E-04 | 3.77E-04 | 6.49E-04 | 2.06E-04 |
| - I - | 12 | 2.14E-03 | 1.10E-03 | 3.65E-03 | 2.05E-04 | 1.14E-03 | 1.96E-03 | 9.89E-03 |
| - J - | 13 | 1.68E-03 | 8.78E-03 | 1.11E-02 | 3.32E-02 | 8.11E-03 | 5.12E-03 | 1.13E-02 |
| - K - | 14 | 3.12E-02 | 2.20E-02 | 3.30E-02 | 3.32E-02 | 1.93E-02 | 1.17E-02 | 5.90E-02 |
| - L - | 15 | 2.07E-03 | 5.49E-03 | 6.93E-03 | 3.32E-03 | 4.80E-03 | 5.45E-03 | 2.64E-03 |
| - M - | 17,18 | 1.74E-03 | 2.20E-03 | 5.78E-03 | 3.32E-02 | 3.88E-03 | 4.61E-03 | 5.38E-04 |
| - N - | 19 | 5.83E-02 | 6.59E-03 | 1.49E-02 | 3.32E-02 | 1.65E-02 | 1.23E-02 | 7.44E-03 |
| - O - | 20 | 1.25E-02 | 1.43E-02 | 9.45E-03 | 3.32E-02 | 9.30E-04 | 1.60E-03 | 1.62E-02 |
| - P - | 21 | 6.48E-03 | 3.29E-02 | 2.71E-02 | 3.32E-02 | 9.03E-04 | 1.55E-03 | 7.46E-03 |
| - Q - | 22,25,28 | 2.88E-02 | 2.20E-02 | 2.84E-02 | 3.32E-02 | 1.66E-02 | 1.30E-02 | 5.23E-02 |
| - R - | 23,26 | 3.06E-02 | 1.54E-02 | 1.43E-02 | 1.73E-04 | 1.78E-02 | 1.38E-02 | 1.58E-02 |
| - S - | 27 | 2.13E-02 | 5.49E-03 | 1.26E-02 | 3.32E-02 | 9.55E-03 | 7.93E-03 | 1.19E-02 |
| - T - | 29 | 1.83E-02 | 1.32E-02 | 2.43E-02 | 3.32E-02 | 6.35E-03 | 3.90E-03 | 4.99E-02 |
| - U - | 30 | 1.26E-02 | 3.07E-02 | 1.03E-01 | 3.32E-03 | 1.61E-02 | 5.47E-03 | 5.16E-02 |
| - V - | 31 | 6.76E-03 | 1.54E-02 | 5.55E-03 | 3.32E-02 | 1.54E-02 | 1.43E-02 | 2.64E-03 |

- Normalized probabilities of occurrence $\lambda_{in}$ of an earthquake with $M \geq 4.95$ for the 22 cells in which earthquakes occurred during the test period. Seven submitted forecasts are given: (1) Bird and Liu (B and L), (2) Ebel et al. (Ebel), (3) Helmstetter et al. (Helm.), (4) Holliday et al. (Holl.), (5) Ward combined (W-C), (6) Ward geodetic (W-G), (7) Wiemer and Schorlemmer (W and S). The highest (best) probabilities are in red. Perfect skill $\lambda_{in}$ = 1.00, No skill $\lambda_{in}$ = 22/7682 = $2.86\times10^{-3}$.

| | $\vert N_{\lambda max} \vert$ | $\vert \lambda_{in} \vert$ | $L_{test}$ |
|---|---|---|---|
| Bird and Liu | 3 | $1.53\times10^{-02}$ | -126 |
| Ebel et al. | 1 | $1.51\times10^{-02}$ | -123 |
| Helmstetter et al. | 4 | $2.84\times10^{-02}$ | -114 |
| Holliday et al. | 8 | $2.45\times10^{-02}$ | -123 |
| Ward combined | 0 | $8.55\times10^{-03}$ | -141 |
| Ward geodetic | 0 | $6.53\times10^{-03}$ | -141 |
| Wiemer and Schor. | 6 | $2.66\times10^{-02}$ | -129 |

- Comparisons of the forecasts: Column 1: The number of maximum cell probabilities $N_{\lambda max}$. Column2: The mean cell probabilities forecast $\lambda_{in}$. Column 3: The maximum likelihood scores. The best scores in each category are red.

# comparison of scoring methods

ranking by points:
("match play")

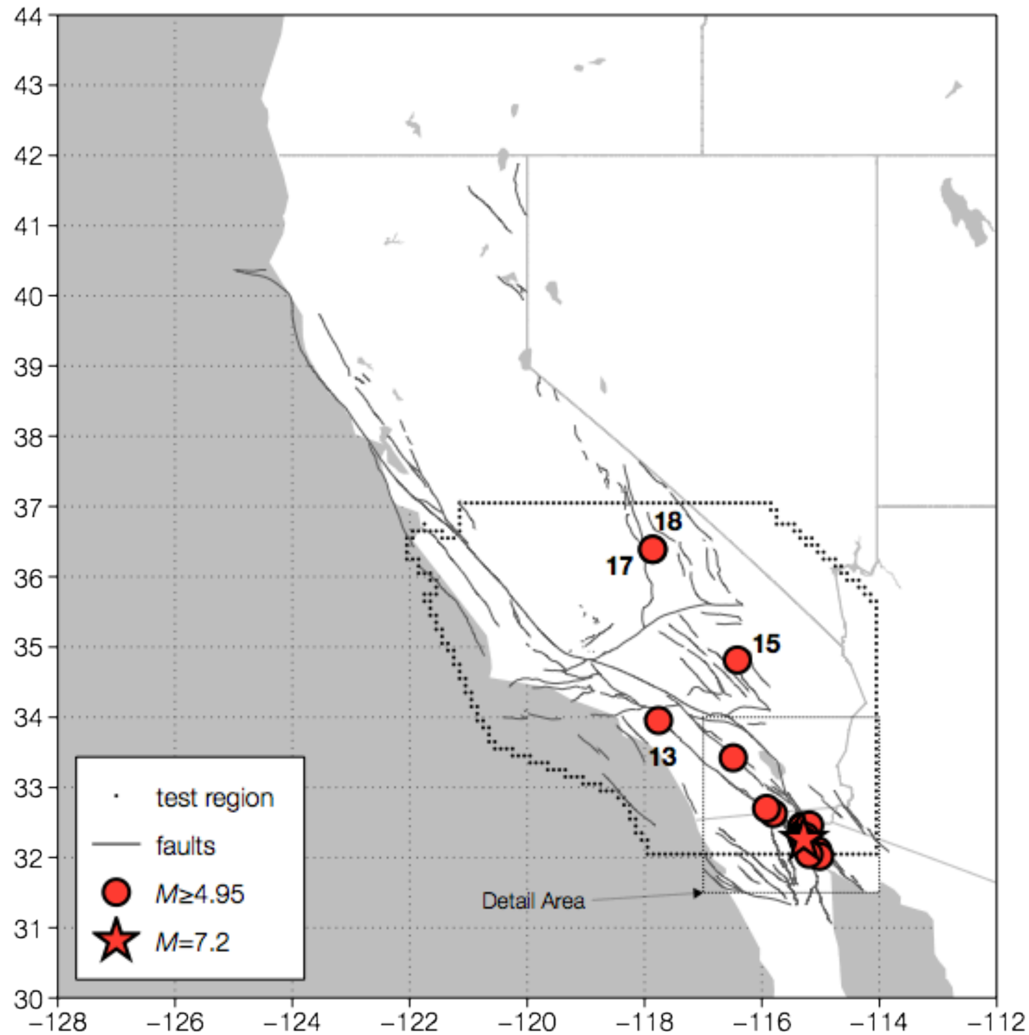
1) Holliday (8 of 22)
2) Weimer-Schorlemmer (6 of 22)
3) Helmstetter (4 of 22)
4) Bird and Liu (3 of 22)
5) Ebel (1 of 22)
6) Ward 1&2 (0 of 22)

ranking by log likelihood:
("stroke play")

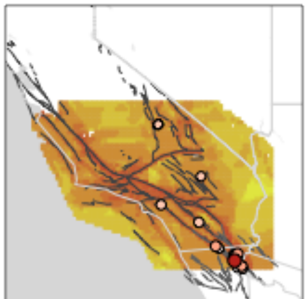
1) Helmstetter (-114)
2) Holliday / Ebel (-123)
3) Bird and Liu (-126)
4) Weimer-Schorlemmer (-129)
5) Ward 1&2 (-141)

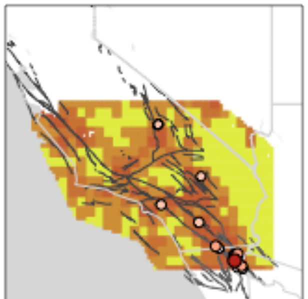
# southern california only: the test region & earthquakes

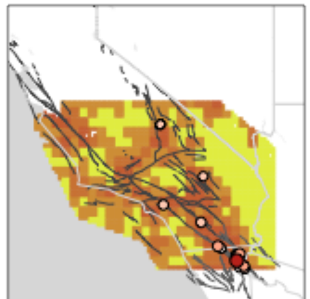
Map of the test region, the coast of California, major faults, and the 22 earthquakes with M ≥ 4.95 that occurred in the test region. Also shown are the square regions where large scale maps are given in subsequent slides.

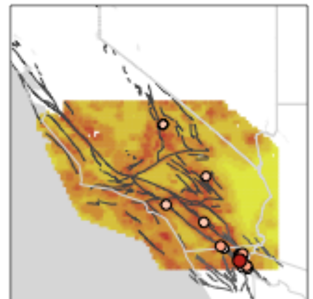


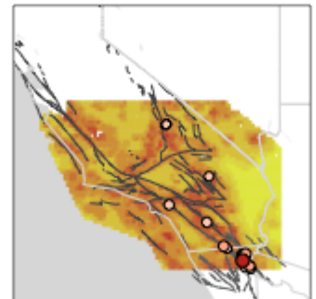
# RELM forecast comparison test
# January 1, 2006 – December 31, 2010

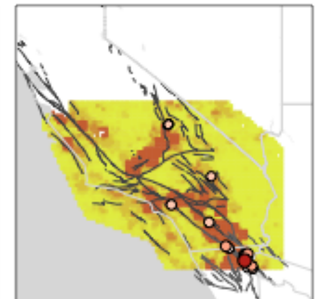

Bird & Liu

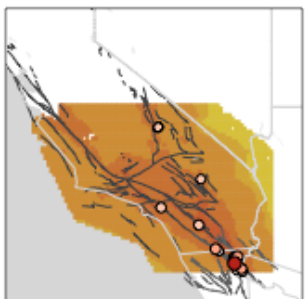

Ebel et al. with aftershocks

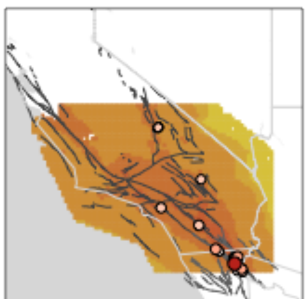

Ebel et al.

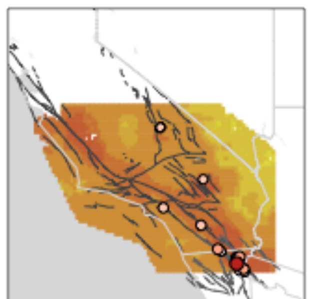

Helmstetter et al. with aftershocks

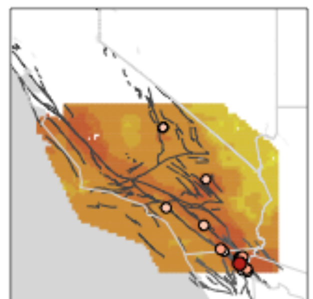

Helmstetter et al.

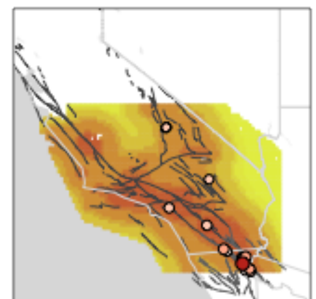

Holliday

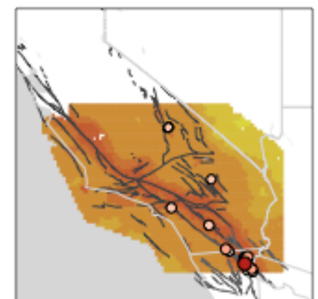

Kagen et al. with aftershocks

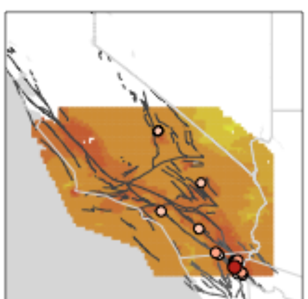

Kagen et al.

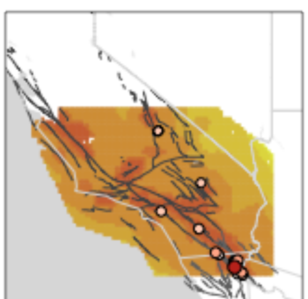

Shen et al. with aftershocks

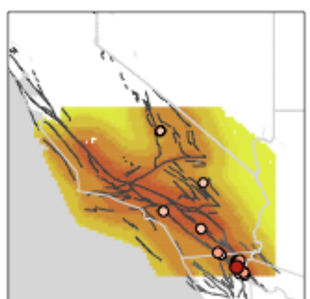

Shen et al.

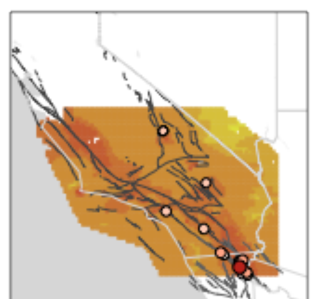

Ward Simulation model

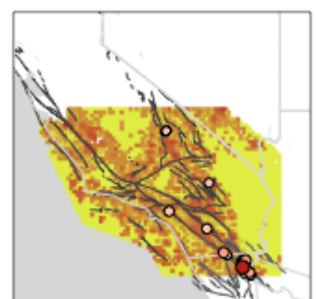

Ward Combined model with Mmax = 8.1


Ward Geodetic model with Mmax = 8.1


Ward Seismic model with Mmax = 8.1


Ward Geologic model with Mmax = 8.1


Ward Geodetic model with Mmax = 8.5


Wiemer & Schorlemmer

| 13 Bins | - A - | - G - | - J - | - K - | - L - | - M - | - N - | - Q - | - R - | - S - | - T - | - U - | - V - | $\bar{\lambda}_{in}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1,7,8,16,24 | 9,10 | 13 | 14 | 15 | 17,18 | 19 | 22,25,28 | 23,26 | 27 | 29 | 30 | 31 | |
| B and L | 2.17E-02 | 9.91E-03 | 1.83E-03 | 3.40E-02 | 2.25E-03 | 1.89E-03 | 6.34E-02 | 3.13E-02 | 3.33E-02 | 2.32E-02 | 1.99E-02 | 1.37E-02 | 7.36E-03 | 2.03E-02 |
| Ebel-AS | 2.60E-02 | 6.50E-03 | 1.04E-02 | 2.60E-02 | 6.50E-03 | 2.60E-03 | 7.79E-03 | 2.60E-02 | 1.82E-02 | 6.50E-03 | 1.56E-02 | 3.64E-02 | 1.82E-02 | 1.59E-02 |
| Ebel | 2.60E-02 | 6.50E-03 | 1.04E-02 | 2.60E-02 | 6.50E-03 | 2.60E-03 | 7.79E-03 | 2.60E-02 | 1.82E-02 | 6.50E-03 | 1.56E-02 | 3.64E-02 | 1.82E-02 | 1.59E-02 |
| Helm.-AS | 1.34E-01 | 2.92E-02 | 1.27E-02 | 3.79E-02 | 7.94E-03 | 6.62E-03 | 1.71E-02 | 3.26E-02 | 1.64E-02 | 1.45E-02 | 2.79E-02 | 1.18E-01 | 6.36E-03 | 3.55E-02 |
| Helm. | 1.34E-01 | 2.92E-02 | 1.27E-02 | 3.79E-02 | 7.94E-03 | 6.62E-03 | 1.71E-02 | 3.26E-02 | 1.64E-02 | 1.45E-02 | 2.79E-02 | 1.18E-01 | 6.36E-03 | 3.55E-02 |
| Holl. | 2.90E-02 | 2.90E-02 | 2.90E-02 | 2.90E-02 | 2.90E-03 | 2.90E-02 | 2.90E-02 | 2.90E-02 | 1.51E-04 | 2.90E-02 | 2.90E-02 | 2.90E-03 | 2.90E-02 | 2.28E-02 |
| Kagen-AS | 9.80E-03 | 1.19E-02 | 7.13E-03 | 1.20E-02 | 4.73E-03 | 4.53E-03 | 9.03E-03 | 1.10E-02 | 1.03E-02 | 8.35E-03 | 8.69E-03 | 1.19E-02 | 7.79E-03 | 9.01E-03 |
| Kagen | 9.80E-03 | 1.19E-02 | 7.13E-03 | 1.20E-02 | 4.73E-03 | 4.53E-03 | 9.03E-03 | 1.10E-02 | 1.03E-02 | 8.35E-03 | 8.69E-03 | 1.19E-02 | 7.79E-03 | 9.01E-03 |
| Shen-AS | 2.29E-02 | 2.35E-02 | 5.20E-03 | 2.38E-02 | 8.75E-03 | 2.76E-03 | 1.59E-02 | 2.34E-02 | 1.90E-02 | 9.20E-03 | 8.80E-03 | 6.96E-03 | 1.76E-02 | 1.44E-02 |
| Shen | 2.29E-02 | 2.35E-02 | 5.21E-03 | 2.38E-02 | 8.76E-03 | 2.76E-03 | 1.59E-02 | 2.34E-02 | 1.90E-02 | 9.21E-03 | 8.82E-03 | 6.96E-03 | 1.76E-02 | 1.44E-02 |
| Ward-SIM | 1.85E-02 | 2.39E-02 | 1.35E-02 | 1.66E-02 | 1.25E-03 | 1.02E-03 | 2.45E-02 | 1.04E-02 | 6.84E-03 | 9.50E-03 | 1.13E-02 | 4.08E-02 | 5.12E-03 | 1.41E-02 |
| Ward -C81 | 1.57E-02 | 1.89E-02 | 6.81E-03 | 1.62E-02 | 4.03E-03 | 3.25E-03 | 1.39E-02 | 1.39E-02 | 1.49E-02 | 8.01E-03 | 5.33E-03 | 1.36E-02 | 1.29E-02 | 1.13E-02 |
| Ward-G81 | 1.38E-02 | 1.37E-02 | 5.50E-03 | 1.26E-02 | 5.85E-03 | 4.95E-03 | 1.32E-02 | 1.40E-02 | 1.48E-02 | 8.52E-03 | 4.19E-03 | 5.87E-03 | 1.53E-02 | 1.02E-02 |
| Ward-S81 | 1.70E-02 | 2.78E-02 | 5.41E-03 | 2.22E-02 | 3.75E-03 | 2.30E-03 | 1.33E-02 | 1.38E-02 | 1.86E-02 | 9.43E-03 | 7.03E-03 | 1.86E-02 | 1.33E-02 | 1.33E-02 |
| Ward-GL81 | 1.76E-02 | 1.83E-02 | 1.06E-02 | 1.59E-02 | 1.17E-03 | 1.34E-03 | 1.57E-02 | 1.41E-02 | 1.10E-02 | 5.58E-03 | 5.46E-03 | 2.14E-02 | 8.33E-03 | 1.13E-02 |
| Ward-G85 | 1.38E-02 | 1.37E-02 | 5.50E-03 | 1.26E-02 | 5.85E-03 | 4.95E-03 | 1.32E-02 | 1.40E-02 | 1.48E-02 | 8.52E-03 | 4.19E-03 | 5.87E-03 | 1.53E-02 | 1.02E-02 |
| W and S | 1.64E-01 | 3.14E-02 | 1.49E-02 | 7.78E-02 | 3.48E-03 | 7.09E-04 | 9.81E-03 | 6.89E-02 | 2.09E-02 | 1.57E-02 | 6.58E-02 | 6.81E-02 | 3.48E-03 | 4.19E-02 |

# comparison of scoring methods - socal

ranking by points:
("match play")

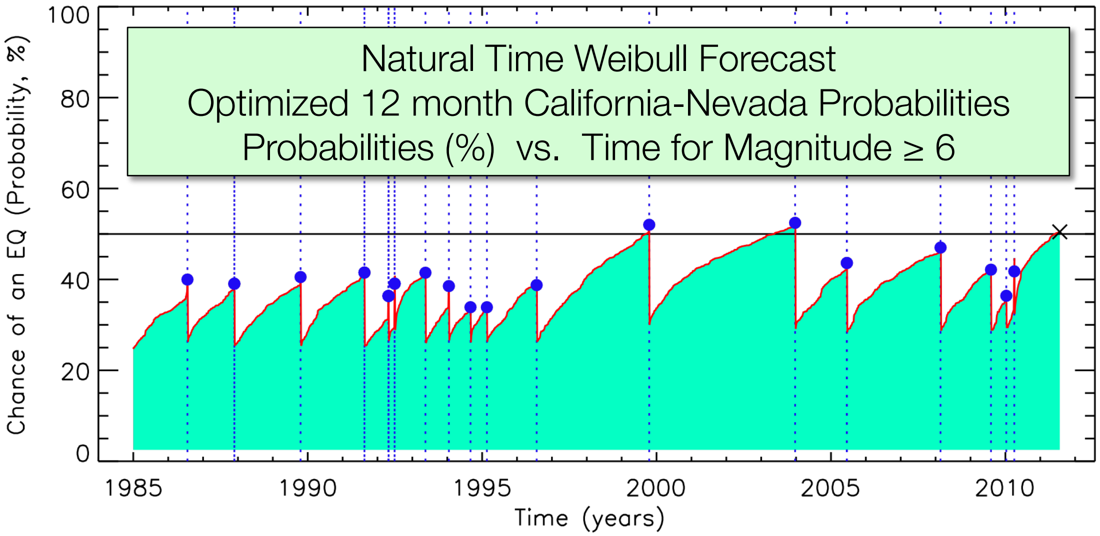
1) Weimer-Schorlemmer (5 of 13)
2) Holliday (4 of 13)
3) Bird-Liu (2 of 13)
4) Shen (1 of 13)
4) Helmstetter (1 of 13)

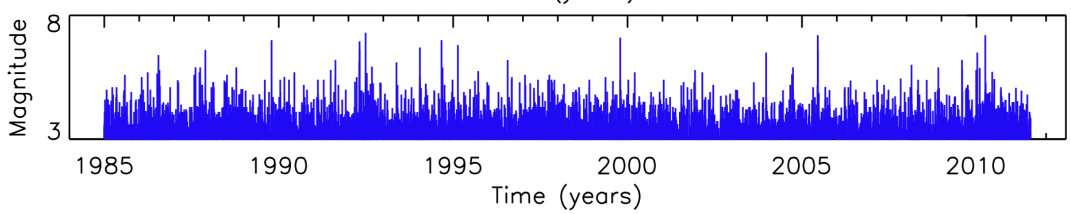
# natural time weibull forecast

- Count number of small eqs since last large eq (natural time) = n
- Use Weibull probability law with N (expected number) as parameter
- Select best parameters based on backtesting
- Use spatial forecast to allocate probability in space
- Account for finite correlation length of earthquake interactions
- Example following

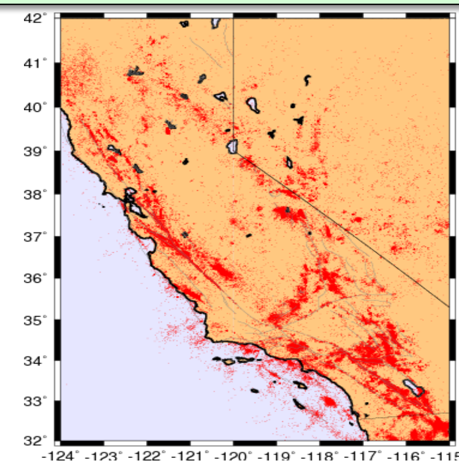
# California-NV Forecast

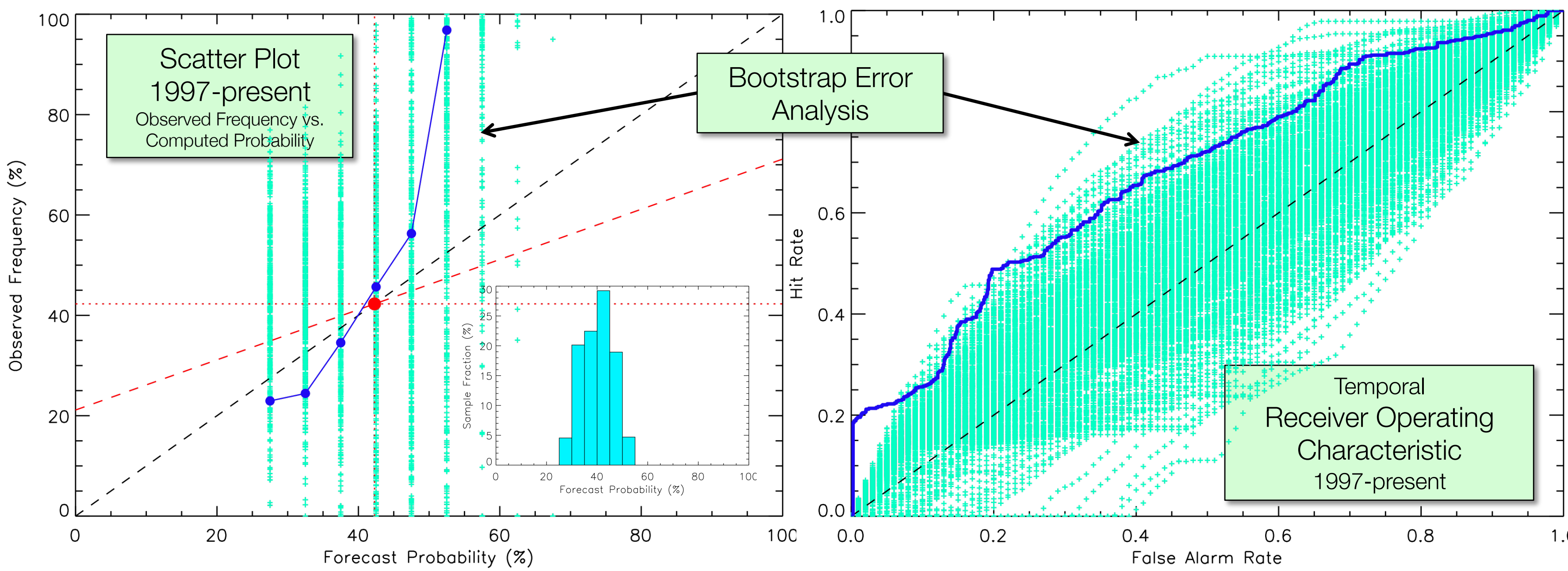
Computed 2011/07/25




Optimal forecasts via backtesting, using common validation and verification testing procedures.



# Forecast Verification – Types of Tests

http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html.

| Method | Answers the Question | Comments |
|---|---|---|
| Bias | How did the frequency of "yes" events compare to the observed frequency of "yes" events? | Indicates tendencies toward underforecasting (Bias<1) or overforecasting (Bias>1). Only measures relative frequencies. |
| Mean Error:<br>Mean Abs. Error<br>Mean Square Error<br>RMS | What is the average forecast error? | Does not measure magnitude of errors. It is possible to get a perfect score for a bad forecast if there are compensating errors. |
| Receiver Operating Characteristic (ROC); Area Skill | How well does the forecast discriminate between events and non-events? | Measures the ability of the forecast to discriminate between two alternative outcomes (resolution). Conditioned on the data. |
| Reliability, Resolution, Sample Skill<br>(Scatter Plots) | How well did the forecast values correspond to the observed values? | Plots the forecast values against the observed values.<br>An accurate forecast will have points near the diagonal.<br>Conditioned on the forecast. |

# scores

| | Reliability | Resolution | Sample Skill | Area Skill |
|---|---|---|---|---|
| **NTW Model** | 0.024 | 0.035 | 0.045 | 0.140 |
| **Random Models** | 0.081±0.031 | 0.072±0.022 | -0.036±0.088 | 0.045±0.125 |

# discussion

1. The general RELM test worked well. The magnitude minimum, cell size and time period were appropriate

2. It would be best to separate the forecasts of number of events from the forecasts of the locations of the events.

3. The relative values of alarm based versus continuous forecasts should be considered. Continuous forecasts are best for insurance purposes.

4. There is no unique scoring that determines the best forecasts.

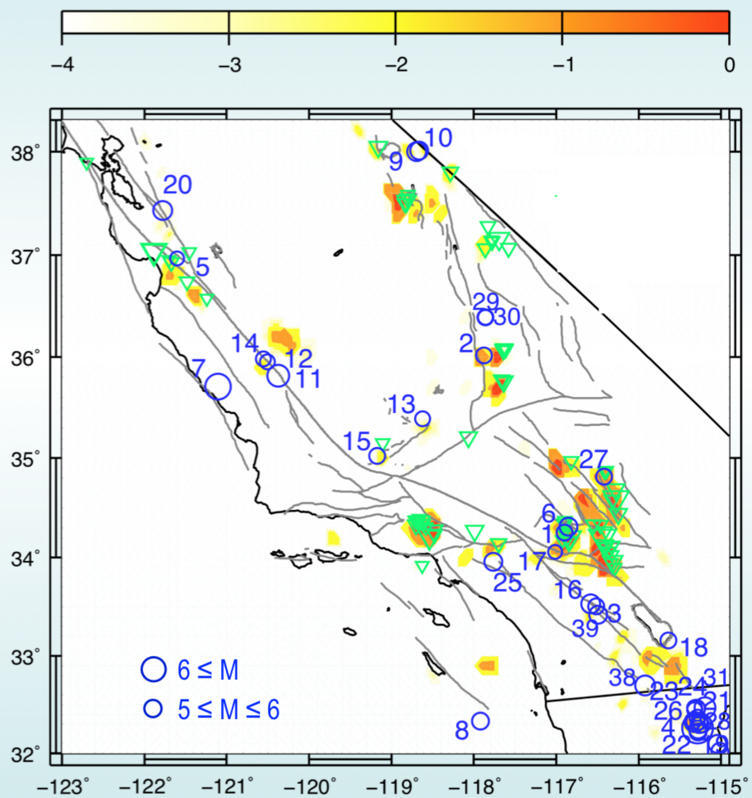
5. Other types of forecasts (e.g., NTW) provide time-dependent view of earthquake hazard

# Pattern Informatics ("PI") Forecast

## ***Status of the Real Time Earthquake Forecast Experiment (Original Version)***

*( JB Rundle et al., PNAS, v99, Supl 1, 2514-2521, Feb 19, 2002; KF Tiampo et al., Europhys. Lett., 60, 481-487, 2002; JB Rundle et al.,Rev. Geophys. Space Phys., 41(4), DOI 10.1029/2003RG000135 ,2003. http://quakesim.jpl.nasa.gov )*

### *How are We Doing?*
### *(Composite N-S Catalog)*




***Plot of $Log_{10}$ (Seismic Potential)***

Increase in Potential for significant earthquakes, ~ 2000 to 2010

Thirty-nine significant earthquakes (blue circles) have occurred in Central or Southern California. Margin of error of the anomalies is +/- 11 km; Data from S. CA. and N. CA catalogs:

<u>After the work was completed</u>

1. Big Bear I, M = 5.1, Feb 10, 2001
2. Coso I, M = 5.1, July 17, 2001

<u>After the paper was in press ( September 1, 2001 )</u>

3. Anza I, M = 5.1, Oct 31, 2001

<u>After the paper was published ( February 19, 2002 )</u>

4. Baja I, M = 5.7, Feb 22, 2002
5. Gilroy, M=4.9 - 5.1, May 13, 2002
6. Big Bear II, M=5.4, Feb 22, 2003
7. San Simeon, M = 6.5, Dec 22, 2003
8. San Clemente Island, M = 5.2, June 15, 2004
9. Bodie I, M=5.5, Sept. 18, 2004
10. Bodie II, M=5.4, Sept. 18, 2004
11. Parkfield I, M = 6.0, Sept. 28, 2004
12. Parkfield II, M = 5.2, Sept. 29, 2004
13. Arvin, M = 5.0, Sept. 29, 2004
14. Parkfield III, M = 5.0, Sept. 30, 2004
15. Wheeler Ridge, M = 5.2, April 16, 2005
16. Anza II, M = 5.2, June 12, 2005
17. Yucaipa, M = 4.9 - 5.2, June 16, 2005
18. Obsidian Butte, M = 5.1, Sept. 2, 2005
19. Baja II, M = 5.4, May 23, 2006
20. Alum Rock, M=5.6, Oct. 30, 2007
21. Baja III, M = 5.4, Feb 9, 2008
22. Baja IV, M = 5.1, Feb 11, 2008
23. Baja V, M = 5.0, Feb 12, 2008
24. Baja VI, M = 5.0, Feb 19, 2008
25. Chino Hills, M = 5.4, July 29, 2008
26. Baja VII, M = 5.0, Nov 20, 2008
27. Mojave, M = 5.1, Dec 6, 2008
28. Baja VIII, M=5.1, Sept 19, 2009
29. Coso II, M=5.0, Oct 1, 2009
30. Coso III, M=5.2, Oct 3, 2009

31.-37. Baja VIII-XV, M=5.4 to M=7.2: Dec 4, 2009 – April 8, 2010

38. Mexicali, M=5.7, June 15, 2010
39. Anza III, M=5.4, July 7, 2010

Note: This original forecast was made using both the full Southern California catalog plus the full Northern California catalog. The S. Calif catalog was used south of lattitude 36º, and the N. Calif. catalog was used north of 36º . No corrections were applied for the different event statistics in the two catalogs. Green triangles mark locations of large earthquakes (M ≥ 5.0) between Jan 1, 1990 – Dec 31, 1999.

CL#03-2015

# forecasts

## Consider 3 Forecast Hypotheses

- Pattern Informatics ("PI"): Future large earthquakes where largest recent changes in seismicity have occurred

- Relative Intensity ("RI"): Future large earthquakes where most recent seismicity has occurred

- National Seismic Hazard Map ("NSHM"): Future large earthquakes based on smoothed seismicity map used in the official NSHM

# Forecast:

## Pattern Informatics (PI)

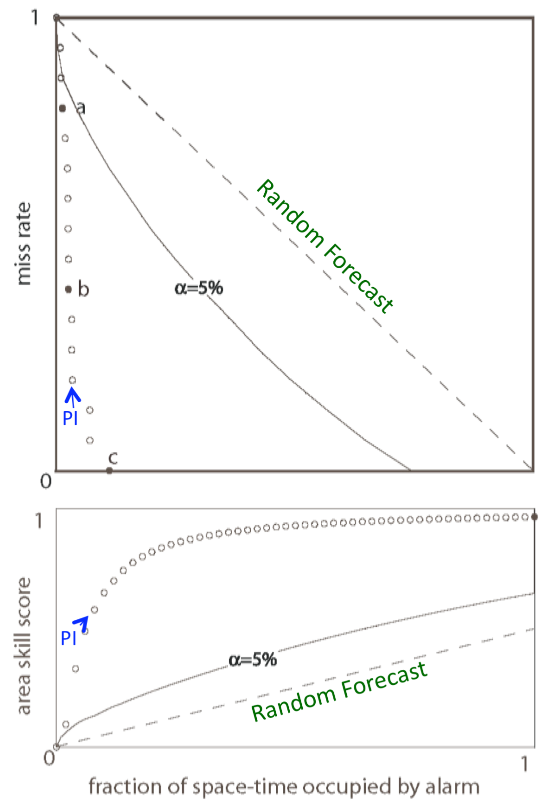
Using Molchan diagrams

Dashed line = Random Forecast. (1 -$\alpha$) = confidence level computed from binomial statistics

Forecast maps cover the period January 1, 2000 - present

15 M > 5 test earthquakes cover the period January 1, 2000 – May 24, 2006

Plots are Figure 6 from:
J. Zechar and T Jordan,
Geophys. J. Int., 172, 715-724 (2008).




PI only

# Forecast:

## Relative Intensity (RI)

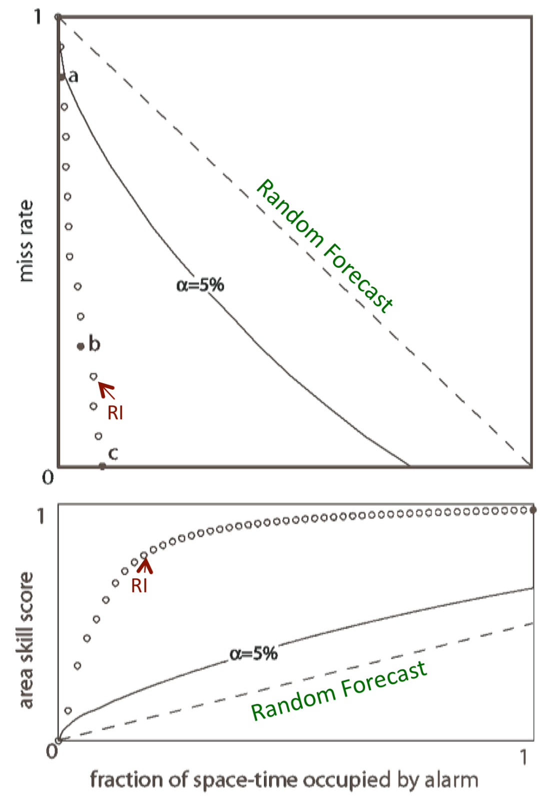
Using Molchan diagrams

Dashed line = Random Forecast. (1 -$\alpha$) = confidence level computed from binomial statistics

Forecast maps cover the period January 1, 2000 - present

15 M > 5 test earthquakes cover the period January 1, 2000 – May 24, 2006

Plots are Figure 7 from:
J. Zechar and T Jordan,
Geophys. J. Int., 172, 715-724 (2008).




RI only

# Forecast:

## Nat. Seism. Hazard Map (NSHM)

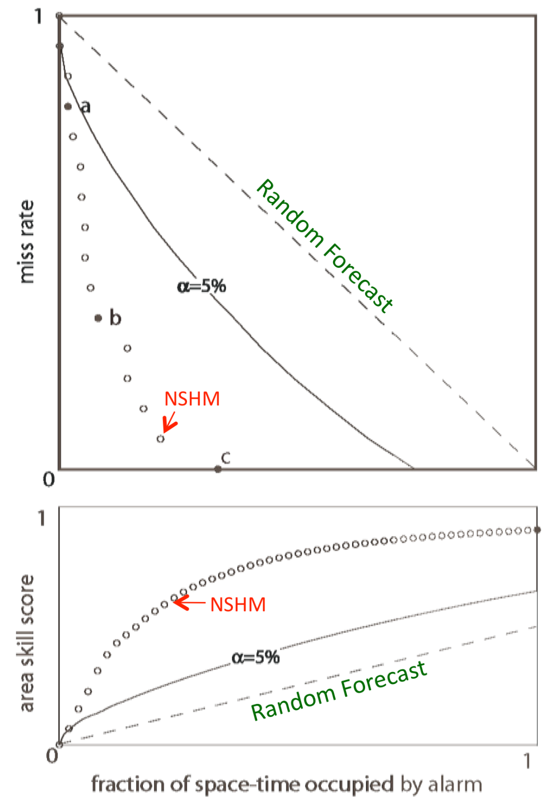
Using Molchan diagrams

Dashed line = Random Forecast. (1 -$\alpha$) = confidence level computed from binomial statistics

Forecast maps cover the period January 1, 2000 - present

15 M > 5 test earthquakes cover the period January 1, 2000 – May 24, 2006

Plots are Figure 8 from:
J. Zechar and T Jordan,
Geophys. J. Int., 172, 715-724 (2008).




NSHM only

# Comparison of Forecasts for:

Pattern Informatics (PI),
Relative Intensity (RI),
National Seismic Hazard Map (NSHM)

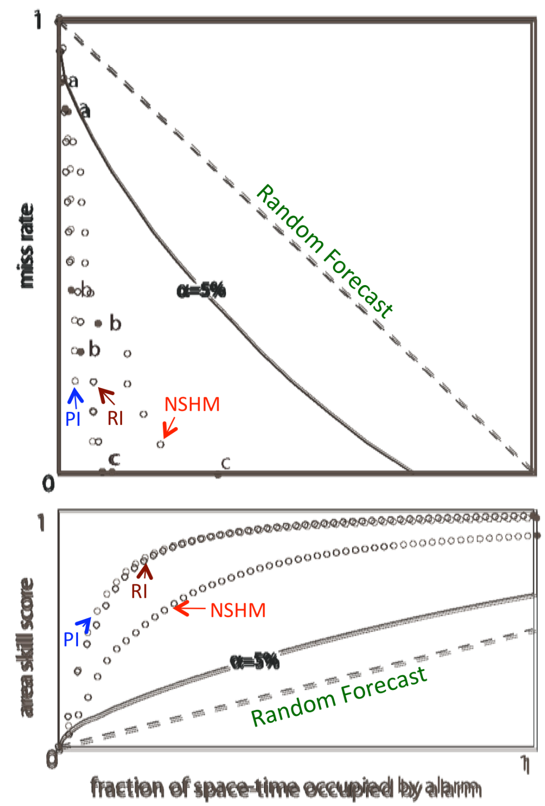
## Using Molchan diagrams

Dashed line = Random Forecast. (1 -$\alpha$) = confidence level computed from binomial statistics

Forecast maps cover the period January 1, 2000 - present

15 M > 5 test earthquakes cover the period January 1, 2000 – May 24, 2006

Plots are superposition of Figures 6-8 from:
J. Zechar and T Jordan,
Geophys. J. Int., 172, 715-724 (2008).




NSHM + RI + PI superposed

# Nanjo, 2010: Molchan Diagrams

RI-BASED EARTHQUAKE FORECAST MODELS

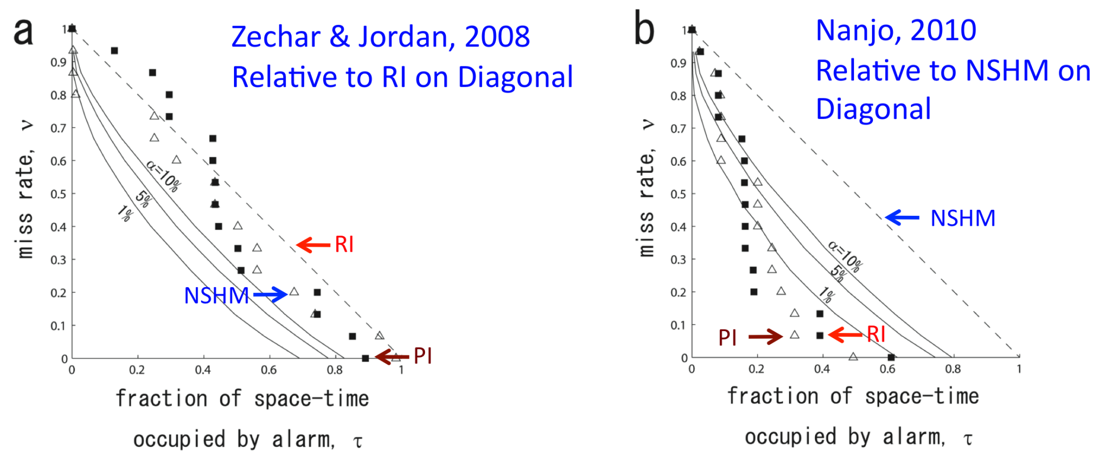



**Figure 1.** Molchan trajectory analysis. (a) PI (filled squares) and NSHM (open triangles) relative to the RI reference model. (b) RI (filled squares) and PI (open triangles) relative to the NSHM reference model. The Molchan trajectories are shown by the plots of the miss rate $\nu$ as functions of fraction of space-time occupied by alarm $\tau$. Each plot also shows the $\alpha = 1\%$, 5% and 10% significance boundaries. In the Molchan trajectory plots, the points below these boundaries reject the null hypothesis. In (a), neither PI nor NSHM provide significant gain relative to RI. In (b), both PI and RI can be significantly better at forecasting future earthquakes than the NSHM for $\nu$~0.0-0.5 at the 95% confidence level (or $\alpha < 5\%$). (a) is a reproduction using the data and program code supplemented in the online version of Zechar and Jordan [2008], while the same data and program code were used to create (b).

# summary – part I

## Molchan Testing

- 3 Forecasts A,B,C
- Choose one forecast as reference, e.g., C
- Compute relative scores of A,B with respect to C
- Results depend on which forecast is chosen as reference