

Testing significance of earthquake precursors: Against complete randomness? Or against earthquake clustering models?

Jiancang Zhuang

Institute of Statistical Mathematics, Tokyo

Timelines with DVJ and statistical seismology

- 1995 First time to know
- 1996 Summer: Working with Ma Li
- 1996.10~1997.10 visiting DVJ in ISOR, VUW, NZ
- 1998.7 1st Statsei Meeting in Hangzhou, China meeting Y. Ogata
- 1999.12-2000.1 Visiting MCS, VUW
- 2000-2003 Studying at ISM
- 2001 Statsei-2, Wellington
- 2002 Workshop in Beijing, DVJ's visiting
- 2003-2010 Visiting each other
- 2008 working on the Bath law, gambling score, Foreshocks

Influential earthquake prediction methods

- ◆ M8
- ◆ RTL (Region-Time-Length)
- ◆ ARM (Accelerating moment release)
- ◆ VAN (Seismo-electric Signals)
- ◆ RTP (Reversed Tracing Precursors)
- ◆ PI (Pattern Informatics)
- ◆ GAP theory
- ◆ Quiescence hypothesis

Institute of Statistical Mathematics, Tokyo

Influential earthquake prediction methods

Seismicity-based precursors:

M8, ARM, RTL (Region-Time-Length), RTP (Reversed Tracing Precursors), PI (Pattern Informatics), GAP theory, Quiescence hypothesis

Non-Seismicity based precursors:

VAN (Seismo-electric Signals), Emission of radon, etc.

Institute of Statistical Mathematics, Tokyo

Influential earthquake prediction methods VAN (Seismo-electric Signals): by Varotsos, Alexopoulos and Nomicos in the 1980s

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand



Influential earthquake prediction methods VAN (Seismo-electric Signals): by Varotsos, Alexopoulos and Nomicos in the 1980s

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand

events. In spite of these differences, both agreed on several points: (1) the VAN method outperforms a Poisson null hypothesis (that is with no clustering) at 95% and higher confidence level, if aftershocks are left in the catalog or only partially removed; and (2) the VAN method does not outperform the Poisson null hypothesis if aftershocks are more fully removed. Kagan [1996] further shows that a null hypothesis which explicitly includes clustering ('Alternative prediction') outperforms the VAN hypothesis. Aceves et al. [1996] did not attempt to include clustering in their null hypothesis. Aceves et al. [1996] appear to favor the VAN hypothesis.

Discussion and conclusions

The results of statistical tests of the VAN earthquake prediction method suggest that although the technique formally is successful, its success might be due either to the posterior adjustment of the prediction rules, or to clustering of shallow earthquakes. A simple prediction algorithm, accounting for non-random seismicity, yields similar forecast results. Since the processing of electric signals is not formalized, it is possible that electric signals were interpreted differently during seismically quiet periods than during periods of seismic activity. This could possibly explain the large difference between the numbers of successful predictions in forward and reverse time [Mulargia and Gasperini, 1992, 1993].

There is still a possibility that precursory electric signals are registered before strong earthquakes as well as during aftershock sequences, and that the connection is thus real. Statistical tests do not usually give a final answer. However, it is clear from the tests that if such a correlation exists, it is weak, and it will be difficult to establish its statistical significance. Jackson [1996] con-

Testing against earthquake clustering

Institute of Statistical Mathematics, Tokyo

Using Hawkes' model to model auxiliary anomalies and earthquake clustering (Zhuang, Vere-Jones, et al. 2005, PAGEOPH)

- Conditional intensity (rate) function (Linlin model, from Ogata)

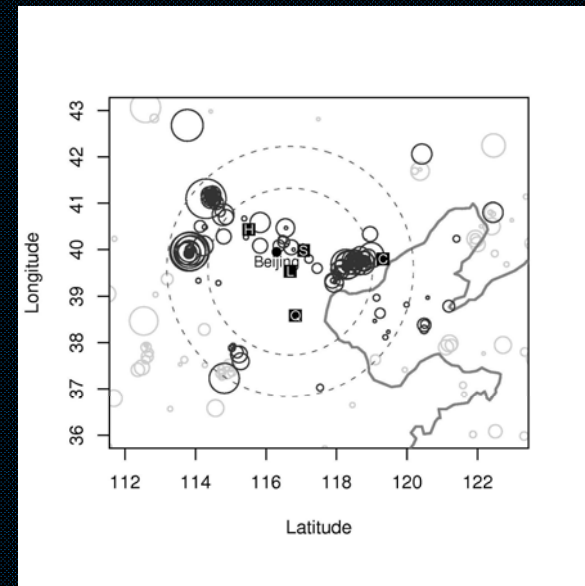
$$\text{EQ rate } \lambda(t) = \Pr\{\text{more than 1 event occurs at } [t, t+dt] \mid \text{observation history}\} \\ = u(t) + \sum_{i: t_i < t} g(t-t_i) + \sum_{j: s_j < t} h(t-s_j)$$

$$\text{Self-exciting term (clustering)} \quad g(t) = e^{-\alpha t} \sum_{k=1}^{K_1} a_k t^{k-1}$$

$$\text{External excitation (precursor input)} \quad h(t) = e^{-\beta t} \sum_{k=1}^{K_2} b_k t^{k-1}$$

$$\text{Background rates (trend)} \quad u(t) = \mu_0 + \mu_1 t + \dots + \mu_{K_0}$$

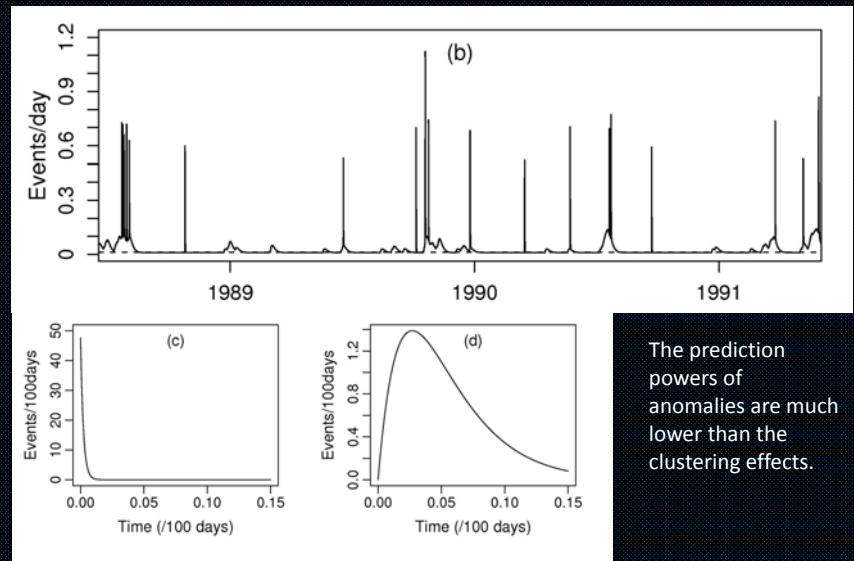
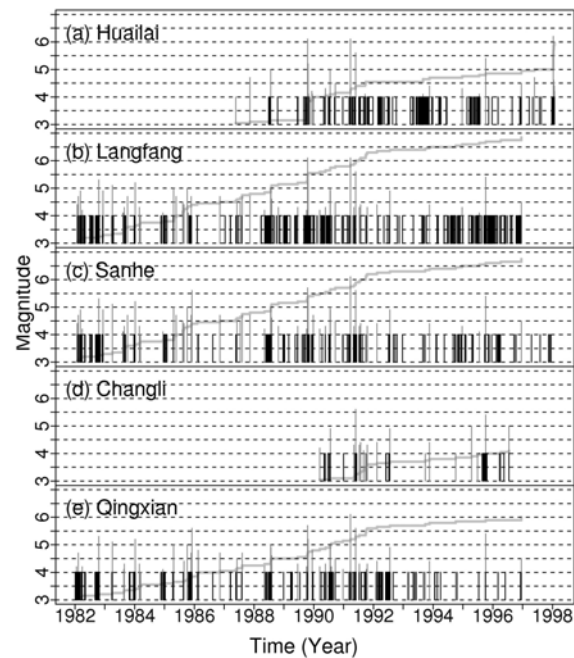
$$\text{Target process } \{t_i\} \quad \text{External input process } \{s_j\}$$



Station names

H: Huailai
L: Langfang
S: Sanhe
Q: Qingxian
C: Changli

Earthquake data:
1982.01.01 – 1998.01.31
4.0 ≤ M ≤ 6.3



The prediction powers of anomalies are much lower than the clustering effects.

Sanhe station: (b) Hazard rate (c) self-excitation (d) external excitation

Influential earthquake prediction methods

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand

M8 (Keilis-Borok and Kossobokov, 1984): designed by retroactive analysis of the seismicity preceding the greatest ($M8+$) earthquakes worldwide. Prediction is aimed at earthquakes of magnitude $M0$ and above. Overlapping circles with the diameter $D(M_0)$ scan the seismic territory. Within each circle the sequence of earthquakes is considered with aftershocks removed $\{t_i, m_i, h_i, b_i(e)\}$, $i = 1, 2 \dots$. Here t_i is the origin time, $t_i \leq t + 1$; m_i is the magnitude, h_i is focal depth, and $b_i(e)$ is the number of aftershocks during the first e days.

$N(t)$, the number of main shocks;

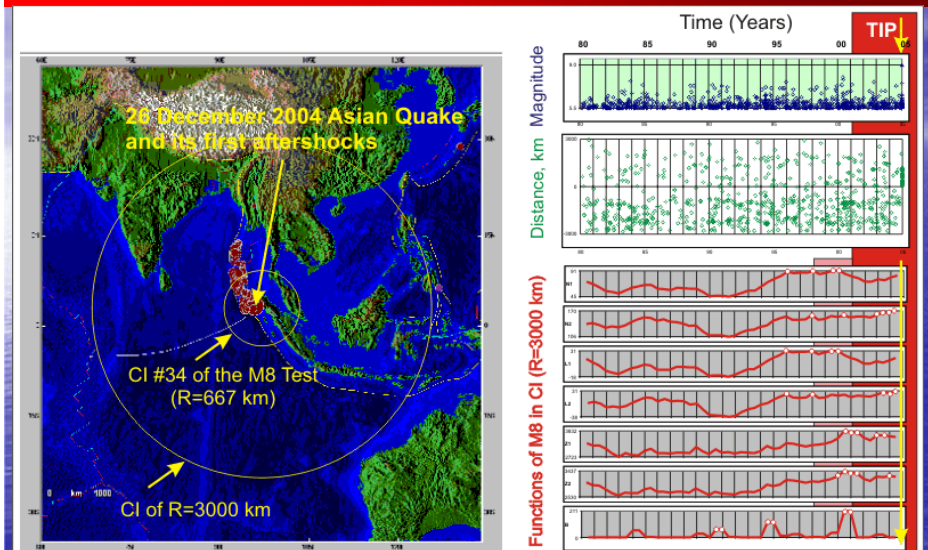
$L(t)$, the deviation of $N(t)$ from the long-term trend,

$Z(t)$, linear concentration of the main shocks estimated as the ratio of the average diameter of the source, l , to the average distance, r , between them; and $B(t) = \max\{b_i\}$, the maximal number of aftershocks.

Each of the functions N , L , Z is calculated for $C = 20$ and $C = 10$.

An alarm or a TIP, "time of increased probability", is declared for five years, when at least six out of seven functions, including B , become "very large" within a narrow time window $(t - u, t)$. To stabilize prediction, this condition is required for two consecutive moments, t and $t + 0.5$ years.

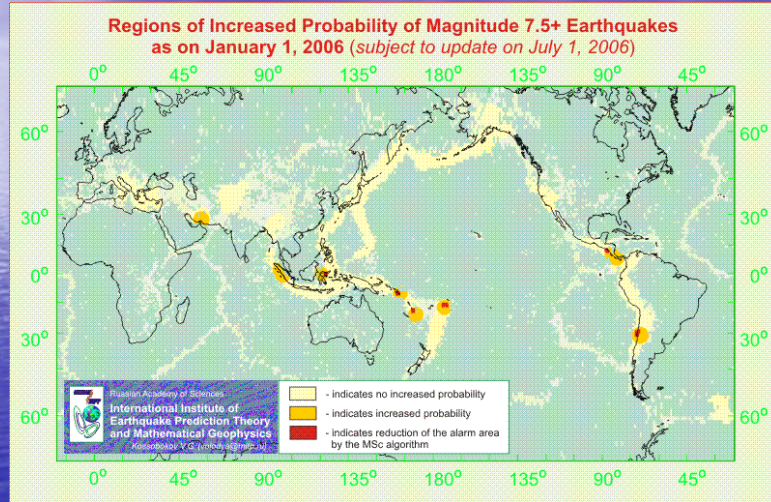
26/12/2004 Mw9.0 Great Asian mega-thrust earthquake



(A slide from Kossobokov's seminar)

Real-time prediction of the world largest earthquakes

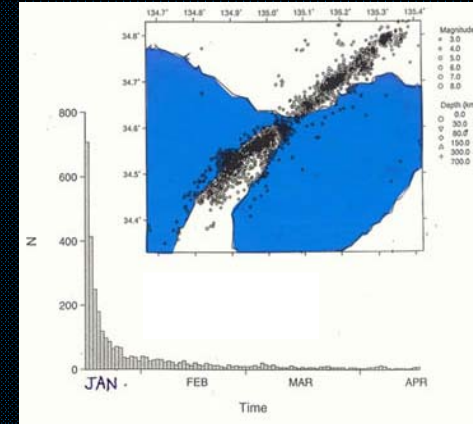
(<http://www.mitp.ru> or <http://www.phys.ualberta.ca/mirrors/mitp>)



(A slide from Kossobokov's seminar)

1.2 Introduction to the ETAS model

- Epidemic-type aftershock sequence (ETAS) model



$$n(t) = \frac{K}{(t+c)}$$

(Omori, 1894)

$$n(t) = \frac{K}{(t+c)^p}$$

(Utsu, 1961)

14

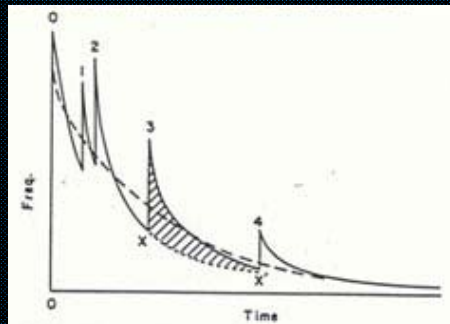
$$n(t) = K(t+c)^{-p} + H(t-T_1)K_1(t-T_1+c_1)^{-p_1} + \dots + H(t-T_n)K_n(t-T_n+c_n)^{-p_n}$$

(Utsu, 1970; Ogata, 1983)

$$\lambda(t) = \mu + \sum_{t_i < t} \frac{K_i}{(t-t_i+c)^p} = \mu + K_0 \sum_{t_i < t} \frac{e^{\alpha(M_i-M_0)}}{(t-t_i+c)^p}$$

(Ogata, 1988; 1989)

$\lambda(t)$: Conditional intensity, hazard function conditioning on the past history



15

Temporal ETAS model

- Conditional intensity

$$\lambda(t) = \mu + \sum_{t_i < t} \kappa(m_i)g(t-t_i)$$

- Direct productivity:

$$\kappa(m) = Ae^{\alpha(m-m_c)}, \quad m \geq m_c$$

- Time p.d.f (Omori-Utsu):

$$g(t) = (p-1)(1+t/c)^{-p}/c, \quad t > 0$$

- Likelihood function

$$\log L = \sum_{t_i \in [0, T]} \log \lambda(t_i) - \int_0^T \lambda(t) dt$$

(Ogata, 1988)

16

Space-time ETAS model

$$\lambda(t, x, y, m) = s(m) \left[\mu(x, y) + \sum_{i: t_i < t} \kappa(m_i) g(t - t_i) f(x - x_i, y - y_i) \right]$$

1. Magnitude p.d.f (G-R): $s(m) = \beta e^{-\beta(m-m_c)}, \quad m \geq m_c$
2. Direct productivity: $\kappa(m) = A e^{\alpha(m-m_c)}, \quad m \geq m_c$
3. Time p.d.f (Omori-Utsu): $g(t) = (p-1)(1+t/c)^{-p} / c, \quad t > 0$
4. Location p.d.f: $f(x, y | m) = \frac{q-1}{\pi D e^{\gamma(m-m_c)}} \left(1 + \frac{x^2 + y^2}{D e^{\gamma(m-m_c)}} \right)^{-q}$

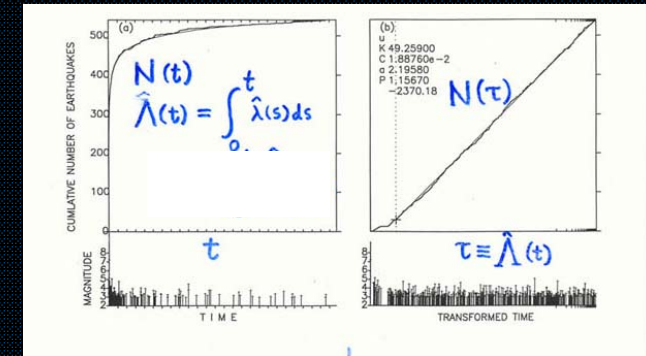
(Ogata, 1988, 2006; zhuang et al. 2005)

17

Combining quiescence with clustering

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand

1. Transformed Time sequence (Ogata, 1992, JGR)



$$t_i \rightarrow \tau_i = \int_0^{t_i} \lambda(u) du$$

If $\{t_i\}$ is the observation of a process with $\lambda(t)$, the $\{\tau_i\}$ is a standard Poisson process.

Combining quiescence with clustering

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand

2. Quiescence in background seismicity (Zhuang et al., 2005)

- Time varying seismicity rate (conditional intensity or stochastic intensity) **at event j**

$$\lambda(t_j, x_j, y_j, m_j) = s(m_j) \left[\mu(x_j, y_j) + \sum_{i: t_i < t} \kappa(m_i) g(t_j - t_i) f(x_j - x_i, y_j - y_i) \right]$$

Contribution from background seismicity

Contribution from the i -th event

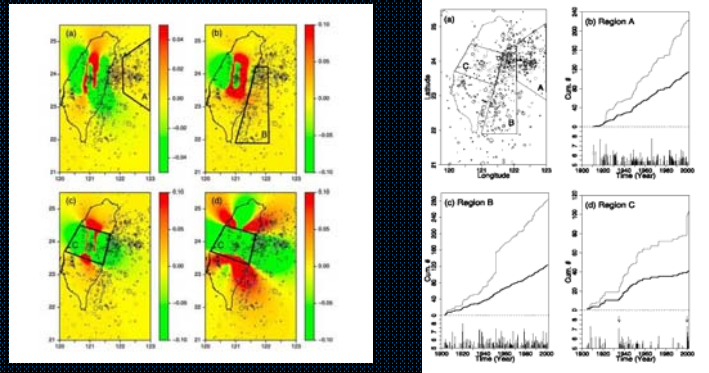
$$\phi_j = \frac{s(m_j) \mu(x_j, y_j)}{\lambda(t_j, x_j, y_j, m_j)}$$

Combining quiescence with clustering

2. Quiescence in background seismicity (Zhuang et al., 2005)

Background seismicity

$$S(t) = \sum_{t_i < t} \varphi_i$$



Combining the ETAS model with PI

(Jiang & Wu, 2010, NHESS)

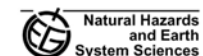
Pattern Informatics (PI)

- PI is a technique developed at University of California, Davis for analyzing earthquake seismic records to forecast regions with high future seismic activity.
 - They have correctly forecasted the locations of 15 of last 16 earthquakes with magnitude > 5.0 in California.
- See Tiampo, K. F., Rundle, J. B., McGinnis, S. A., & Klein, W. Pattern dynamics and forecast methods in seismically active regions. *Pure Ap. Geophys.* 159, 2429-2467 (2002).
 - <http://citebase.eprints.org/cgi-bin/fulltext?format=application/pdf&identifier=oai%3AarXiv.org%3Acond-mat%2F0102032>
- PI is being applied other regions of the world, and John has gotten a lot of press.
 - Google "John Rundle UC Davis Pattern Informatics"

Combining the ETAS model with PI

(Jiang & Wu, 2010, NHESS)

Nat. Hazards Earth Syst. Sci., 11, 697–706, 2011
www.nat-hazards-earth-syst-sci.net/11/697/2011/
doi:10.5194/nhe-11-697-2011
© Author(s) 2011. CC Attribution 3.0 License



PI forecast with or without de-clustering: an experiment for the Sichuan-Yunnan region

C. S. Jiang and Z. L. Wu

Institute of Geophysics, China Earthquake Administration, Beijing 100081, China

Received: 27 July 2010 – Revised: 27 December 2010 – Accepted: 4 February 2011 – Published: 7 March 2011

Abstract. Pattern Informatics (PI) algorithm uses earthquake catalogues to estimate the increase of the probability of strong earthquakes. The main measure in the algorithm is the number of earthquakes above a threshold magnitude. Since aftershocks occupy a significant proportion of the total number of earthquakes, whether de-clustering affects the performance of the forecast is one of the concerns in the application of this algorithm. This problem is of special interest after a great earthquake, when aftershocks become predominant in regional seismic activity. To investigate this problem, the PI forecasts are systematically analyzed for the Sichuan-Yunnan region of southwest China. In this region there have occurred some earthquakes larger than M_s 7.0, including the 2008 Wenchuan earthquake. In the analysis, the epidemic-type aftershock sequences (ETAS) model was used for de-clustering. The PI algorithm was revised to consider de-clustering, by replacing the number of earthquakes by the sum of the ETAS-assessed probabilities for an event to

uses earthquakes catalogues to identify the increase of probability of strong earthquakes. The main measure of this algorithm is the number of earthquakes above a threshold magnitude. In the perspective of nonlinear dynamics, the PI algorithm tests whether the seismicity before a strong earthquake is different from what it "should be" in the usual time. The null hypothesis for this test is that the seismicity is stationary, or more strongly, being Poissonian. For counts of mainshocks, generally this assumption has few problems, but for an earthquake sequence which mixes mainshocks and aftershocks, this may not be the case. It is well-known that the occurrence of aftershocks does not observe a Poissonian process, rather it observes the Omori-Utsu law. As a matter of fact, in some of the analysis of aftershocks, to identify the "abnormal change" of seismicity above the "normal" background, transformation is used so that the aftershock sequence becomes Poissonian along the "transformed time" axis (Omori, 2007). In other studies, de-clustering, that is

Combining the ETAS model with PI

(Jiang & Wu, 2010, NHESS)

Events in Box i , $N_i(t)$, replaced by

$$\sum_{\text{Event } k \in \text{Box } i} \varphi_k$$

Conclusion: using background events in the PI forecasts seems to improve performance slightly, for the region that they investigate.

3.1 The PI algorithm

In the PI algorithm, the whole region under study is binned into boxes or "pixels" with size $D \times D$ centered at a point x_i . Each point x_i is associated with a time series $N_i(t)$, where $N_i(t)$ is the time-dependent average rate of earthquakes with magnitude greater than the cutoff magnitude M_c in box i and its Moore neighborhood. $N_i(t)$ is calculated for box i within a period starting from time t_0 to time t ($t > t_0$). The "seismic activity intensity" function of box i is defined as the average rate of occurrence of earthquakes:

$$I_i(t_0, t) = \frac{1}{t - t_0} \sum_{t'=t_0}^t N_i(t') \quad (1)$$

The probability of a future strong earthquake in box i , $P_i(t_0, t_1, t_2)$, is defined as the square of the average intensity fluctuation:

$$P_i(t_0, t_1, t_2) = \Delta I_i(t_0, t_1, t_2)^2 \quad (2)$$

in which t_1 is the starting time of the "anomaly identification window", t_2 the ending time of the "anomaly identification window" and the starting time of the "forecast window". The "sliding window" for PI calculation is selected to start from t_0 . Subtracting the mean probability over all boxes and denoting this change as the probability-increase of future earthquakes via

$$\Delta P_i(t_0, t_1, t_2) = P_i(t_0, t_1, t_2) - \langle P_i(t_0, t_1, t_2) \rangle \quad (3)$$

the "hotspots" are defined to be the boxes where $\Delta P_i(t_0, t_1, t_2)$ is positive, or the probability function $P_i(t_0, t_1, t_2)$ is larger than the background level. Physically, since the probability function has quadratic form, either "seismic activation" or seismic quiescence can be reflected by the PI "hotspot" map, which is one of the reasons why PI forecasts outperform the "relative intensity" (RI) forecasts.

Space-time ETAS model

- Time varying seismicity rate (conditional intensity or stochastic intensity) **at event j**

$$\lambda(t_j, x_j, y_j, m_j) = s(m_j) \left[\mu(x_j, y_j) + \sum_{i: t_i < t} \kappa(m_i) g(t_j - t_i) f(x_j - x_i, y_j - y_i) \right]$$

Contribution from background seismicity

Contribution from the i -th event

$$\varphi_j = \frac{s(m_j) \mu(x_j, y_j)}{\lambda(t_j, x_j, y_j, m_j)}$$

Combining the ETAS model and LURR

(Zhang & Zhuang 2011, Tectophysics)

Outlines

1. What is LURR (Load/Unload response ratio)?
2. The ETAS model and a new model
3. Estimating LURR through residuals
4. Application to the Wenchuan earthquake

David Vere-Jones Seminar Series
2011-7-27
Waikanae, New Zealand

Introduction to the LURR (load/unload response ration)

Basic theory

- Three phases of solid materials when being loaded: Elastic, Plastic, & fracture.

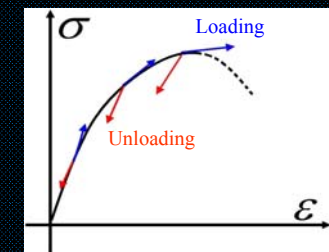
- LURR: useful for describing evolution of damaging process

Response rate $X = \lim_{\Delta\sigma \rightarrow 0} \frac{\Delta\epsilon}{\Delta\sigma}$

LURR $Y = \frac{X_+}{X_-}$

X_+ , Loading response rate

X_- , Unloading response rate



Typical stress strain curve for solid materials

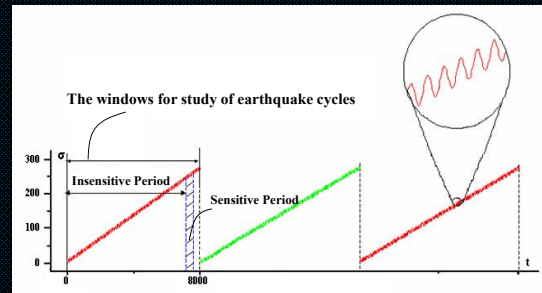
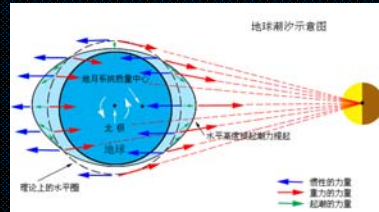
How to load/unload the crust

$$\sigma_{re} = \sigma_{te} + \sigma_{ti} \approx kt + A \sin(\omega t + \varphi)$$

re: resultant stress

te: tectonic stress

ti: tidal stress



29

Using tidal stresses

Choose the response variable

Energies released by earthquakes

$$Y = \frac{\left(\sum_{i=1}^{N^+} E_i^\kappa \right)_+}{\left(\sum_{i=1}^{N^-} E_i^\kappa \right)_-}$$

$\kappa=1$, E^κ energy

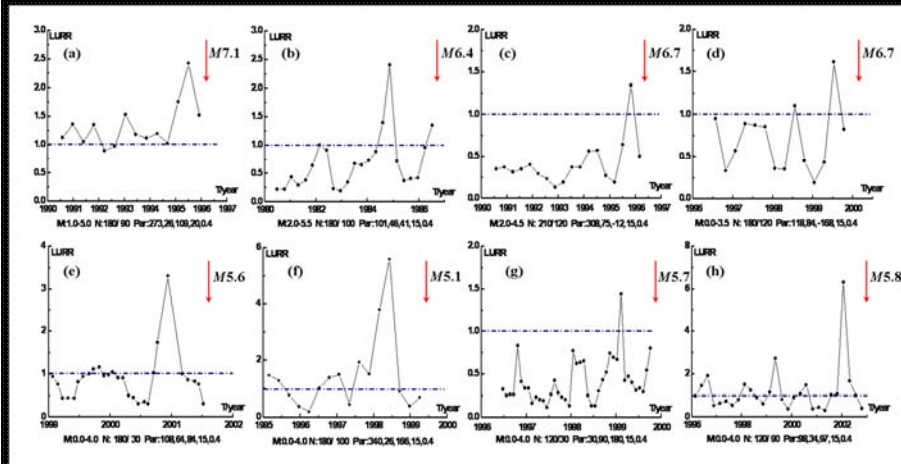
$\kappa=1/2$, E^κ Benioff strain

$\kappa=0$, $\sum_{i=1}^N E^\kappa$: number of events

small earthquakes \leftrightarrow micro-fractures

30

Examples for the LURR research groups



LURR before several earthquake
in China

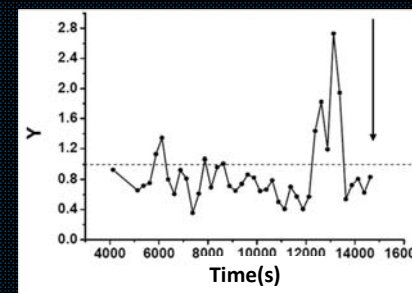
Yin et al., 2000

31

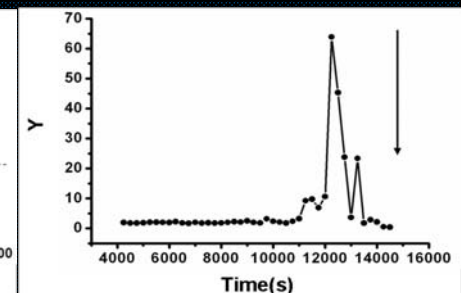
Laboratory experiment on rock fractures

$$Y = \frac{\left(\sum_{i=1}^{N^+} E_i^\kappa \right)_+}{\left(\sum_{i=1}^{N^-} E_i^\kappa \right)_-}$$

$$Y = \frac{(\Delta \varepsilon_1)_+ / (\Delta P_1)_+}{(\Delta \varepsilon_1)_- / (\Delta P_1)_-}$$



For Granite G2



Zhang et al., 2006

32

Comibine LURR and ETAS model

Baseline model (ETAS model) $\lambda_0(t, m) = \lambda_0(t)s(m)$

New model with response rate $\lambda(t, m) = X(t)\lambda_0(t, m)$

$$X(t) = \begin{cases} X_+(t), & \text{loading period} \\ X_-(t), & \text{unloading period} \end{cases}$$

X_+ : loading response

X_- : unloading response

LURR

$$Y(t) = \frac{X_+(t)}{X_-(t)}$$

33

Comibine LURR and ETAS model

Baseline model (ETAS model) $\lambda_0(t, m) = \lambda_0(t)s(m)$

New model with response rate $\lambda(t, m) = X(t)\lambda_0(t, m)$

$$X(t) = \begin{cases} X_+(t), & \text{loading period} \\ X_-(t), & \text{unloading period} \end{cases}$$

X_+ : loading response

X_- : unloading response

LURR

$$Y(t) = \frac{X_+(t)}{X_-(t)}$$

How to estimate?!

34

A property of the conditional intensity

Given a marked point process N equipped with a conditional intensity $\lambda(t, m)$, if $h(t, m)$ is a predictable marked process, then for any fixed interval S

$$E \left[\sum_{i: (t_i, m_i) \in N \cap S} h(t_i, m_i) \right] = E \left[\int_M \int_S h(t, m) \lambda(t, m) dm dt \right]$$

35

Combine LURR and ETAS model

Baseline model (ETAS model) $\lambda_0(t, m) = \lambda_0(t)s(m)$

New model with response rate $\lambda(t, m) = X(t)\lambda_0(t, m)$

$$X(t) = \begin{cases} X_+(t), & \text{loading period} \\ X_-(t), & \text{unloading period} \end{cases}$$

X_+ : loading response

X_- : unloading response

LURR

$$Y(t) = \frac{X_+(t)}{X_-(t)}$$

Choose

$$h(t, m) = \frac{f(m)}{\lambda_0(t, m)}$$

Some known function of m

36

A property of the conditional intensity

Given a marked point process N equipped with a conditional intensity $\lambda(t, m)$, if $h(t, m)$ is a predictable marked process, then for any fixed interval S

$$E \left[\sum_{i: (t_i, m_i) \in N \cap S} h(t_i, m_i) \right] = E \left[\int_M \int_S h(t, m) \lambda(t, m) dm dt \right]$$

37

Combining LURR and ETAS model

For a loading period, S_+

$$\begin{aligned} E \left[\sum_{i: t_i \in S_+} f(m_i) / \lambda_0(t_i) \right] &= E \left[\int_M \int_{S_+} \frac{f(m) \lambda(t, m)}{\lambda_0(t)} dm dt \right] \\ &= E \left[\int_M \int_{S_+} f(m) s(m) X(t) dm dt \right] \\ &= X_+ | S_+ | E \left[\int_M f(m) s(m) dm \right] \\ &= X_+ | S_+ | E[f(m)] \end{aligned}$$

Similarly, for the unloading period, S_-

$$E \left[\sum_{i: t_i \in S_-} f(m_i) / \lambda_0(t_i) \right] = X_- | S_- | E[f(m)]$$

LURR

$$Y = \frac{X_+}{X_-} = \frac{|S_-| E \left[\sum_{i: t_i \in S_+} f(m_i) / \lambda_0(t_i) \right]}{|S_+| E \left[\sum_{i: t_i \in S_-} f(m_i) / \lambda_0(t_i) \right]} \approx \frac{|S_-| \sum_{i: t_i \in S_+} f(m_i) / \lambda_0(t_i)}{|S_+| \sum_{i: t_i \in S_-} f(m_i) / \lambda_0(t_i)}$$

38

Combining LURR and ETAS model

For a loading period, S_+

$$\begin{aligned} E \left[\sum_{i: t_i \in S_+} f(m_i) / \lambda_0(t_i) \right] &= E \left[\int_M \int_{S_+} \frac{f(m) \lambda(t, m)}{\lambda_0(t)} dm dt \right] \\ &= E \left[\int_M \int_{S_+} f(m) s(m) X(t) dm dt \right] \\ &= X_+ | S_+ | E \left[\int_M f(m) s(m) dm \right] \\ &= X_+ | S_+ | E[f(m)] \end{aligned}$$

Similarly, for the unloading period, S_-

$$E \left[\sum_{i: t_i \in S_-} f(m_i) / \lambda_0(t_i) \right] = X_- | S_- | E[f(m)]$$

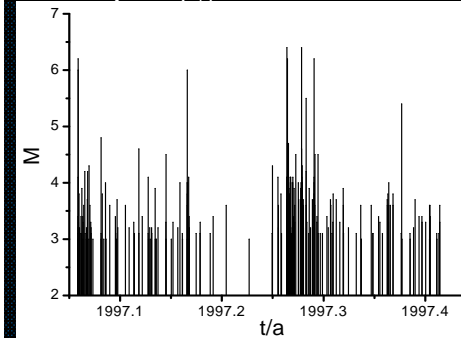
LURR

$$Y_2 \approx \frac{|S_-| \sum_{i: t_i \in S_+} E_i^{1/2} / \lambda_0(t_i)}{|S_+| \sum_{i: t_i \in S_-} E_i^{1/2} / \lambda_0(t_i)}$$

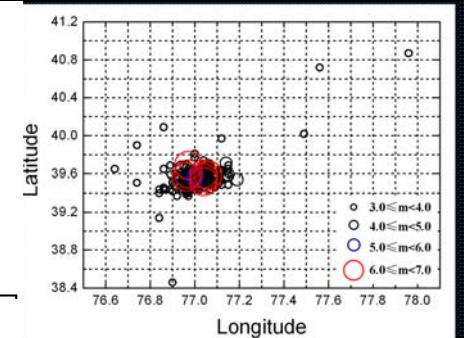
Ratio-unbiased estimate

39

3.1 1997 Jiashi earthquake swarm, Xinjiang, China



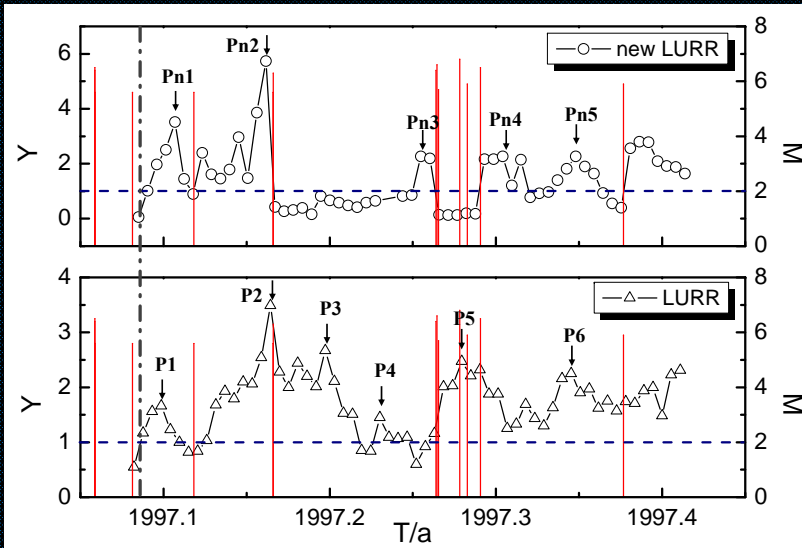
M-t plot



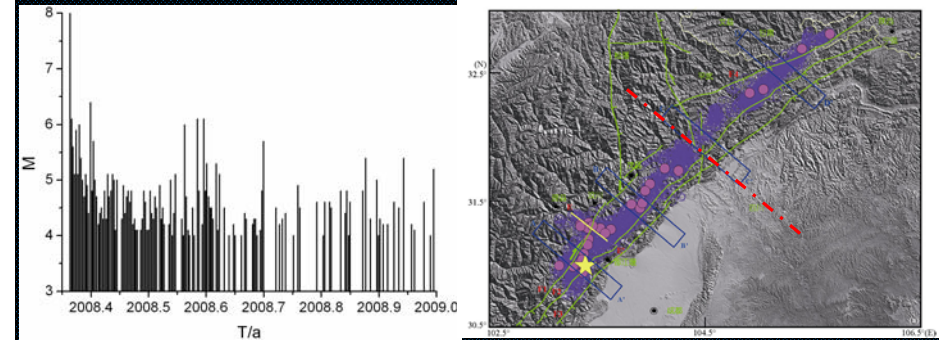
Epicenter locations

40

LURR Curves



Aftershock sequence of the 2008 Wenchuan Earthquake, Sichuan, China



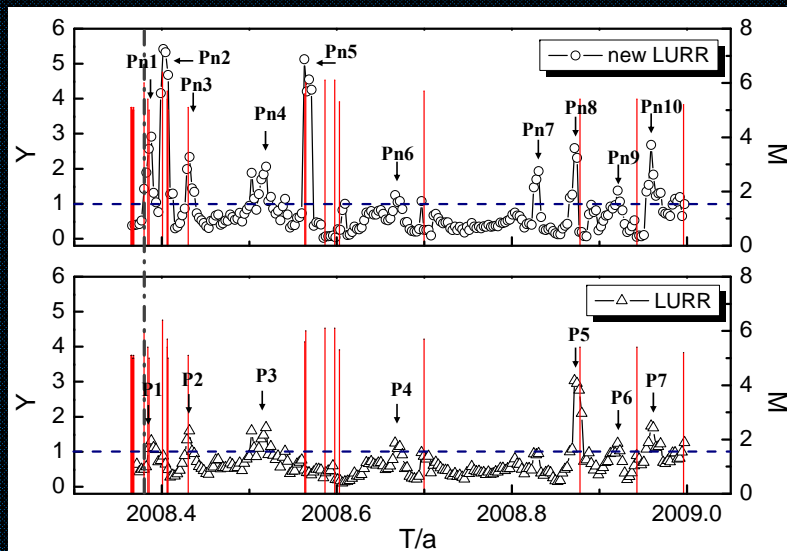
M-t plot

Locations of earthquake epicenters

(Huang et al., 2008; Burchfiel et al., 2008)

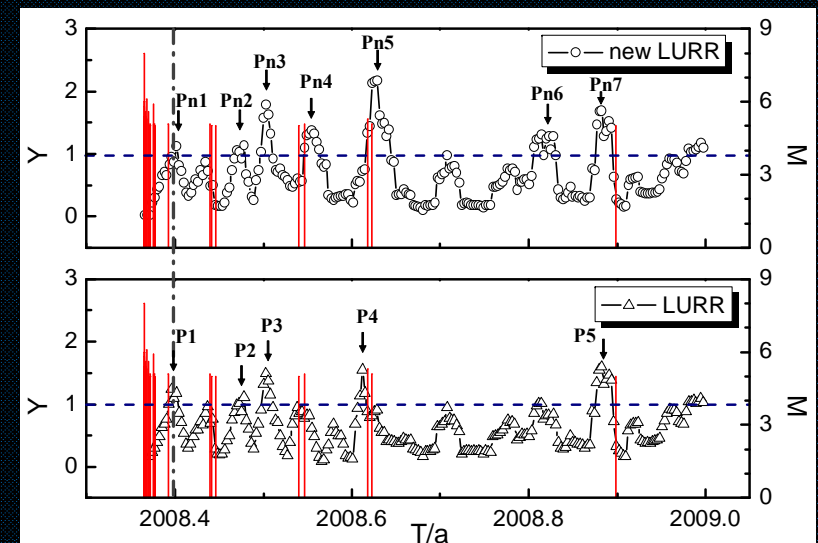
42

Northeast part



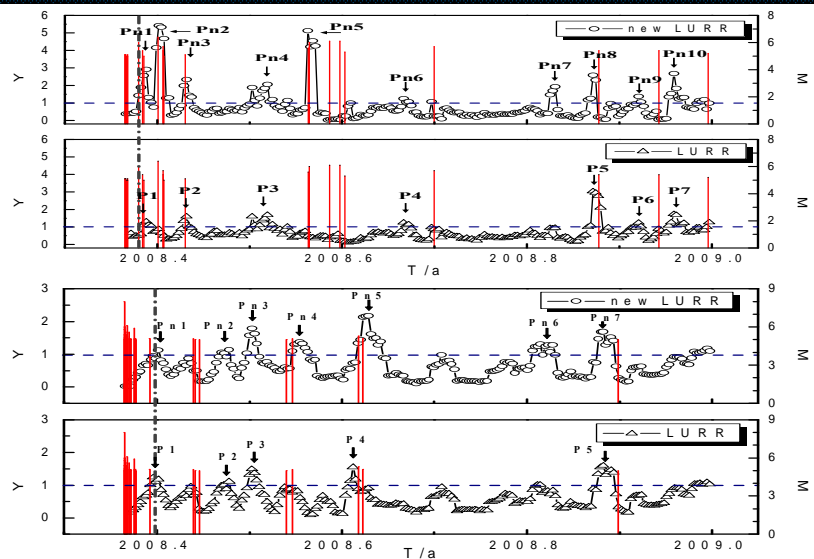
43

Southwest part



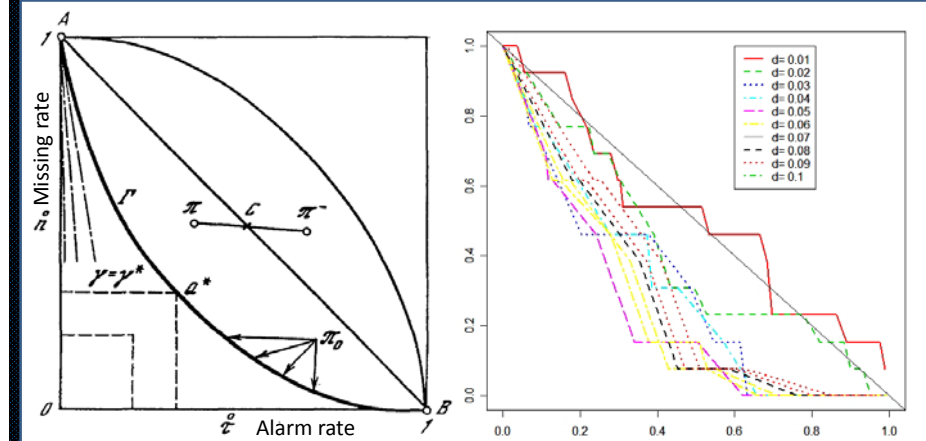
44

NE



45

Molchan's error diagram



Molchan's error diagram

(Molchan 1997)

Results for the
Northeast region.

Conclusions

- ❑ Up to now, biggest known non-randomness is the clustering effect of earthquakes, which out-performs the other available earthquake prediction methods.
- ❑ Making use of the ETAS model and the techniques of residual analysis, almost all the available prediction algorithms can be revised into new versions with the elimination of the earthquake clustering effect.
- ❑ Even though the clustering effect of earthquakes can be eliminated with the techniques introduced in this presentation, the final prediction performance depends on potential of the prediction powers of the algorithm.
- ❑ Further studies are necessary to apply the above techniques to other algorithms.