Regional applicability of quake forecasts using geoelectric statistical moments: Application to Kakioka, Japan

Hong-Jia Chen[1], Katsumi Hattori[2], Chien-Chih Chen[1]

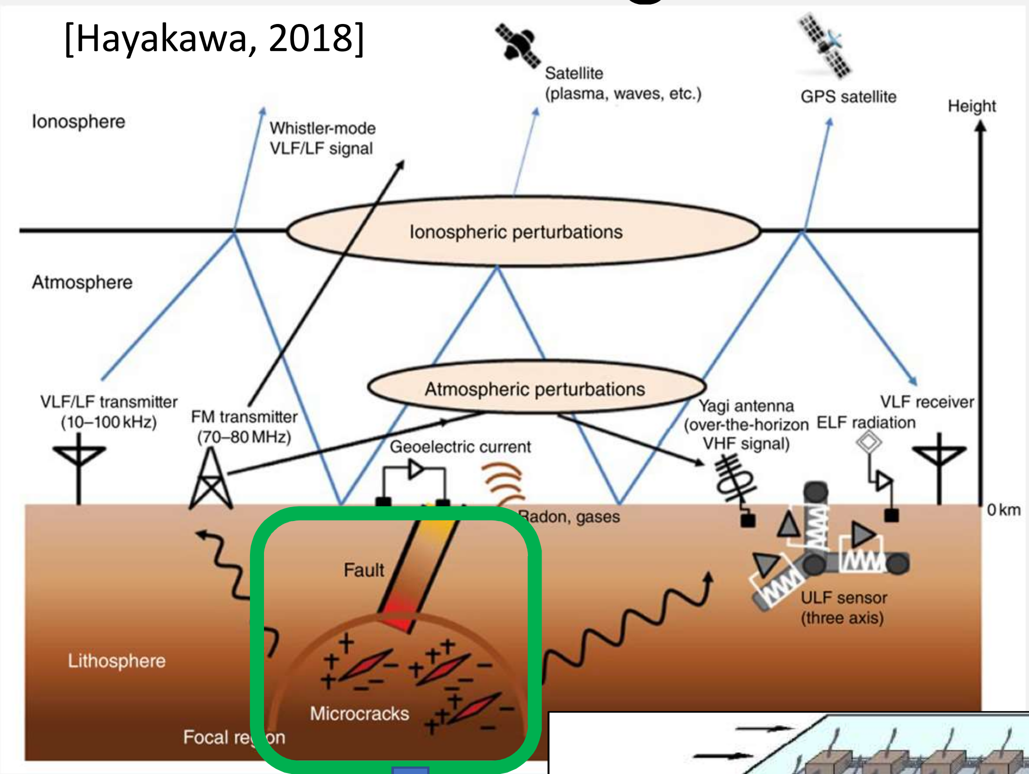
[1]Department of Earth Sciences, National Central University, Taiwan.
[2]Department of Earth Sciences, Graduate School of Science, Chiba University, Japan.

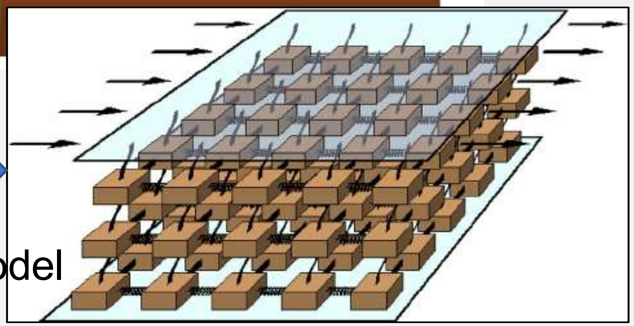
27 May 2022

Takeaway points

1. Quake-related signals exist and affect the distribution of geoelectric time series.
2. The quake-forecasting algorithm based on statistical moments is applicable not only in Taiwan but in Japan, implying its universality.
3. Instead of deterministic quake forecasting, operational forecasting is possible through integrating the multiple measurements and analyzing their early-warning indicators.

Earthquake mechanics and electromagnetics: Self-organized critical phenomena





Mechanics:
Spring-block model

Since the 1980s, studies related to "Earthquake as a critical point": Bak, Ito, Sornette, Turcotte, Keilis-Borok, et al.

Features:

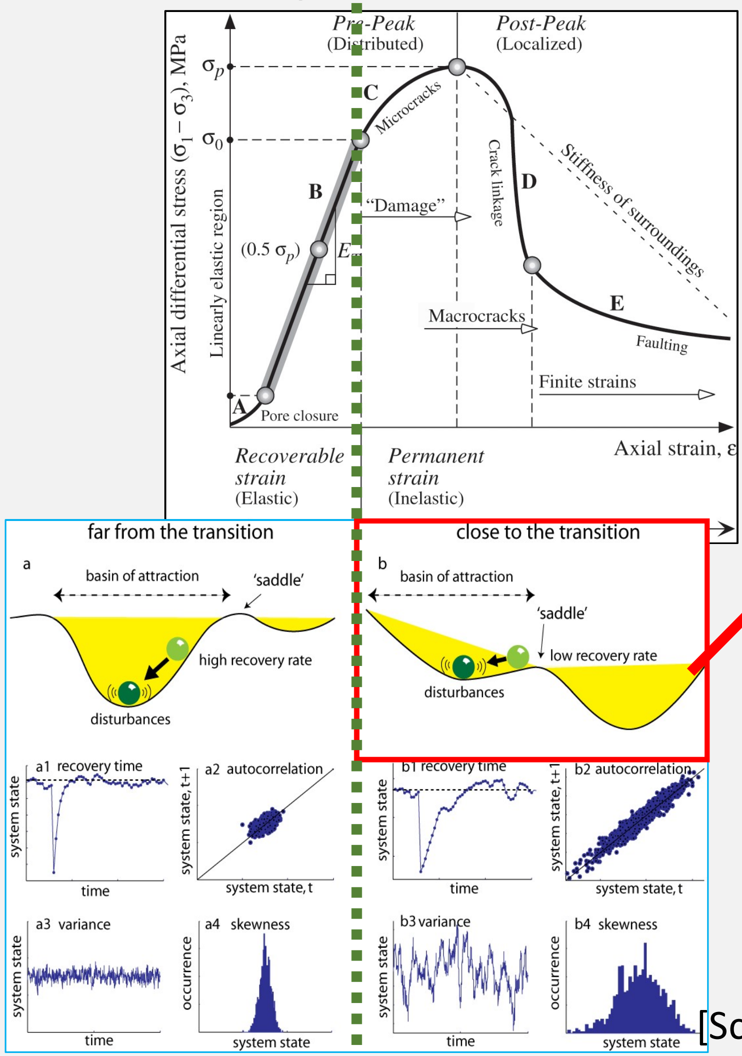
- Scaling law
- Long-range correlation
- Cooperative behavior
- Nonlinear dynamics

To reveal the dynamics, continuous monitoring is needed.
Mechanics: stress, strain, seismicity
Electromagnetics: self-potential, geomagnetic data, ionospheric data (total electron content)

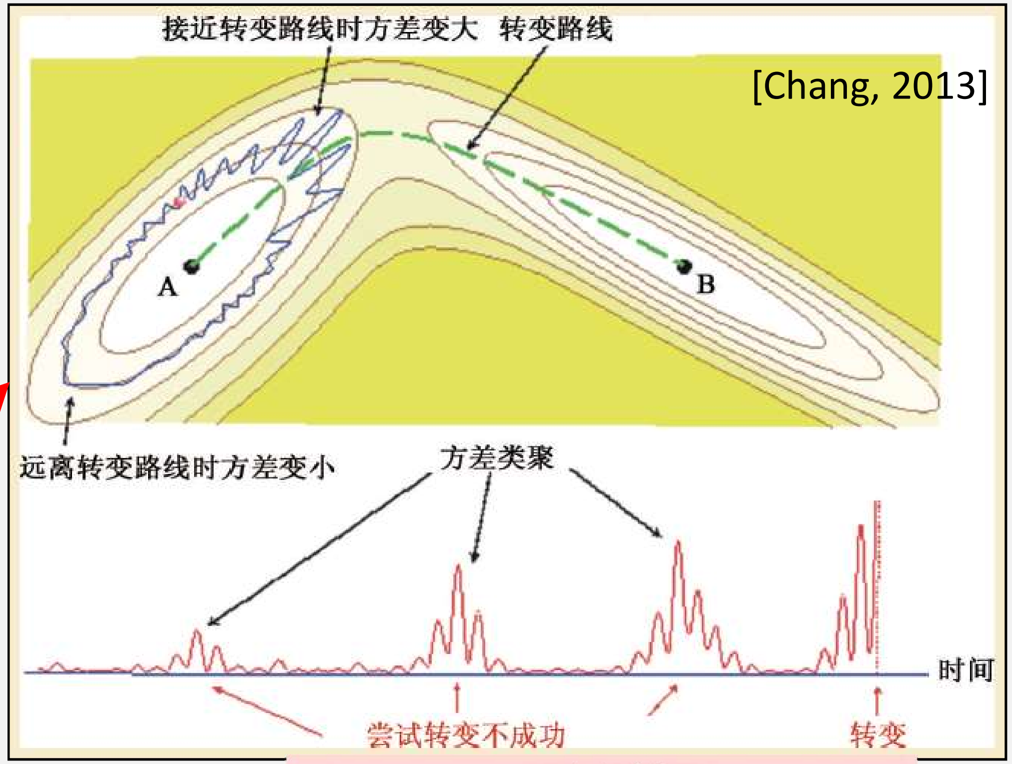
Searching for patterns that relate to the critical transition theory of earthquakes.

Critical transition of a catastrophic event

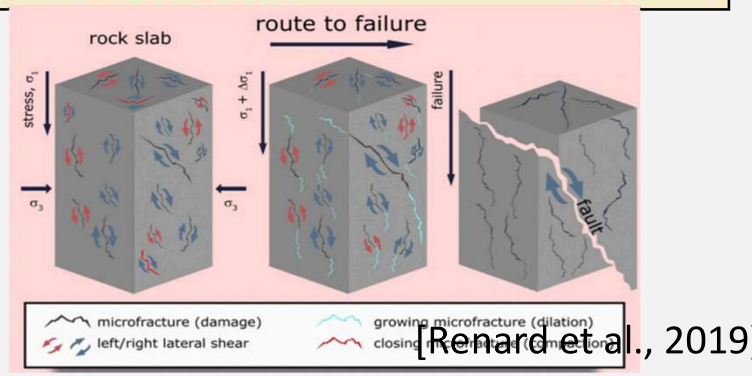
Elastic stage | **After the yielding point**



[Scheffer et. al, 2009]



[Chang, 2013]



[Renard et al., 2019]

Pre-earthquake

"Early-warning" signal for critical transition

Complex dynamical system:

- Crustal system
- Weather system
- Financial system
- Biomass system

Generic leading indicators for critical transition

Table 1. Early warning signals for critical transitions.

	Method/Indicator	Phenomenon: Rising memory	Rising variability	Flickering	Ref.
metrics	Autocorrelation at-lag-1	x			[23]
	Autoregressive coefficient of AR(1) model	x			[19]
	Return rate (inverse of AR(1) coefficient)	x			[23]
	Detrended fluctuation analysis indicator	x			[7]
	Spectral density	x			[20]
	Spectral ratio (of low to high frequencies)	x			[25]
	Spectral exponent	x			[this paper]
	Standard deviation		x	x	[28]
	Coefficient of variation		x	x	[28]
	Skewness		x	x	[29]
	Kurtosis		x	x	[25]
	Conditional heteroskedasticity		x	x	[32]
	BDS test		x	x	[10]
models	Time-varying AR(p) models	x	x		[38]
	Nonparametric drift-diffusion-jump models	x	x	x	[16]
	Threshold AR(p) models			x	[38]
	Potential analysis (potential wells estimator)			x	[43]

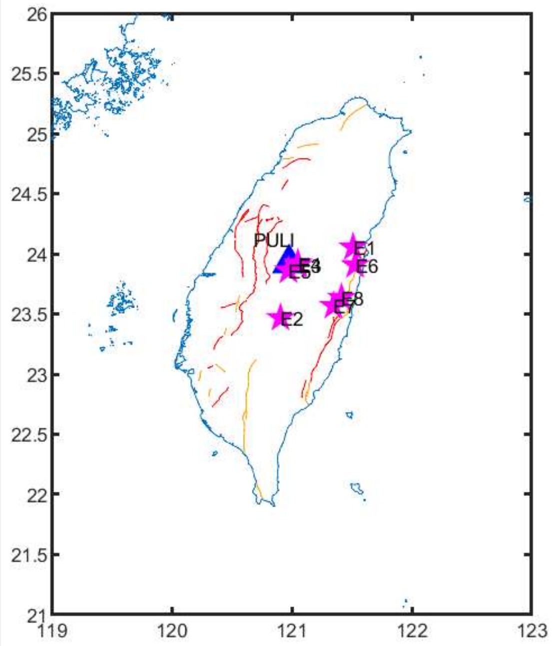
Statistical moments

[Dakos et al., 2012]

Skewness variation of Taiwan

Event #	Time (UTC+8)	Location (N, E)	Depth (km)	Magnitude
E1	2012/8/14 18:55	24.05, 121.51	21.83	5.18
E2	2012/12/31 0:3	23.47, 120.91	4.41	5.28
E3	2013/3/27 10:3	23.90, 121.05	19.43	6.24
E4	2013/3/27 10:4	23.92, 121.05	19.36	5.00
E5	2013/6/2 13:43	23.86, 120.97	14.54	6.48
E6	2013/7/24 23:32	23.91, 121.53	10.59	5.03
E7	2013/10/31 20:2	23.57, 121.35	14.98	6.42
E8	2013/10/31 23:47	23.63, 121.42	9.72	5.08

Year 2012 - 2013

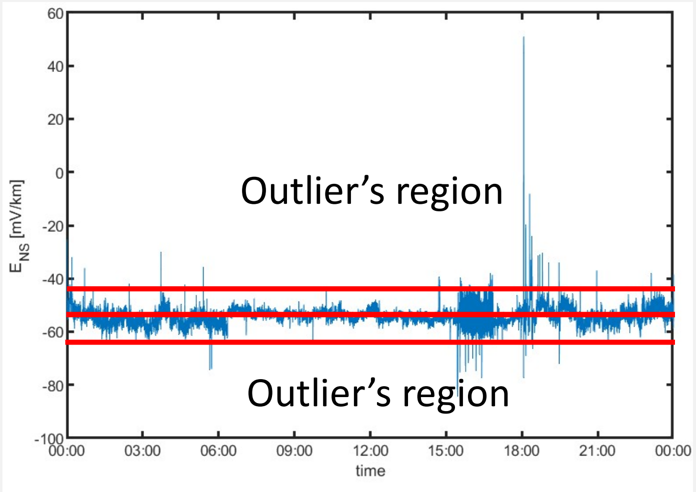


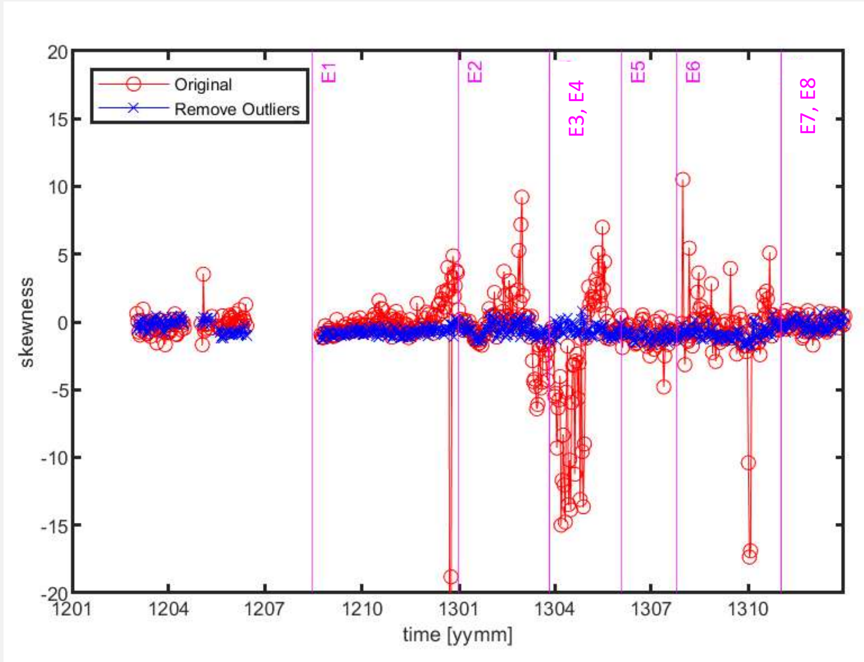
Definition of outliers

Tukey fence (F)

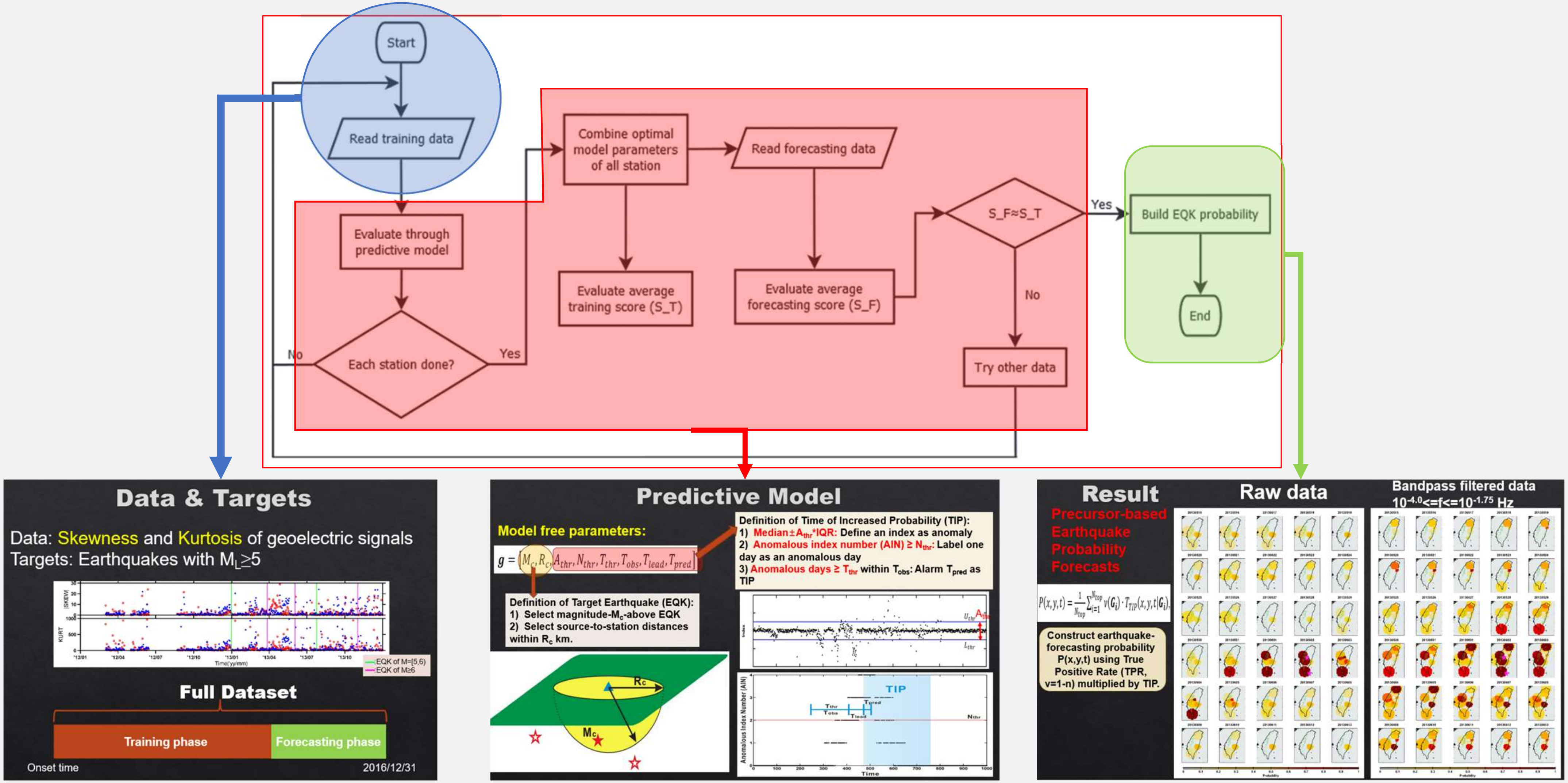
$$F = Median \pm 3 \times IQR$$

IQR means the interquartile range.





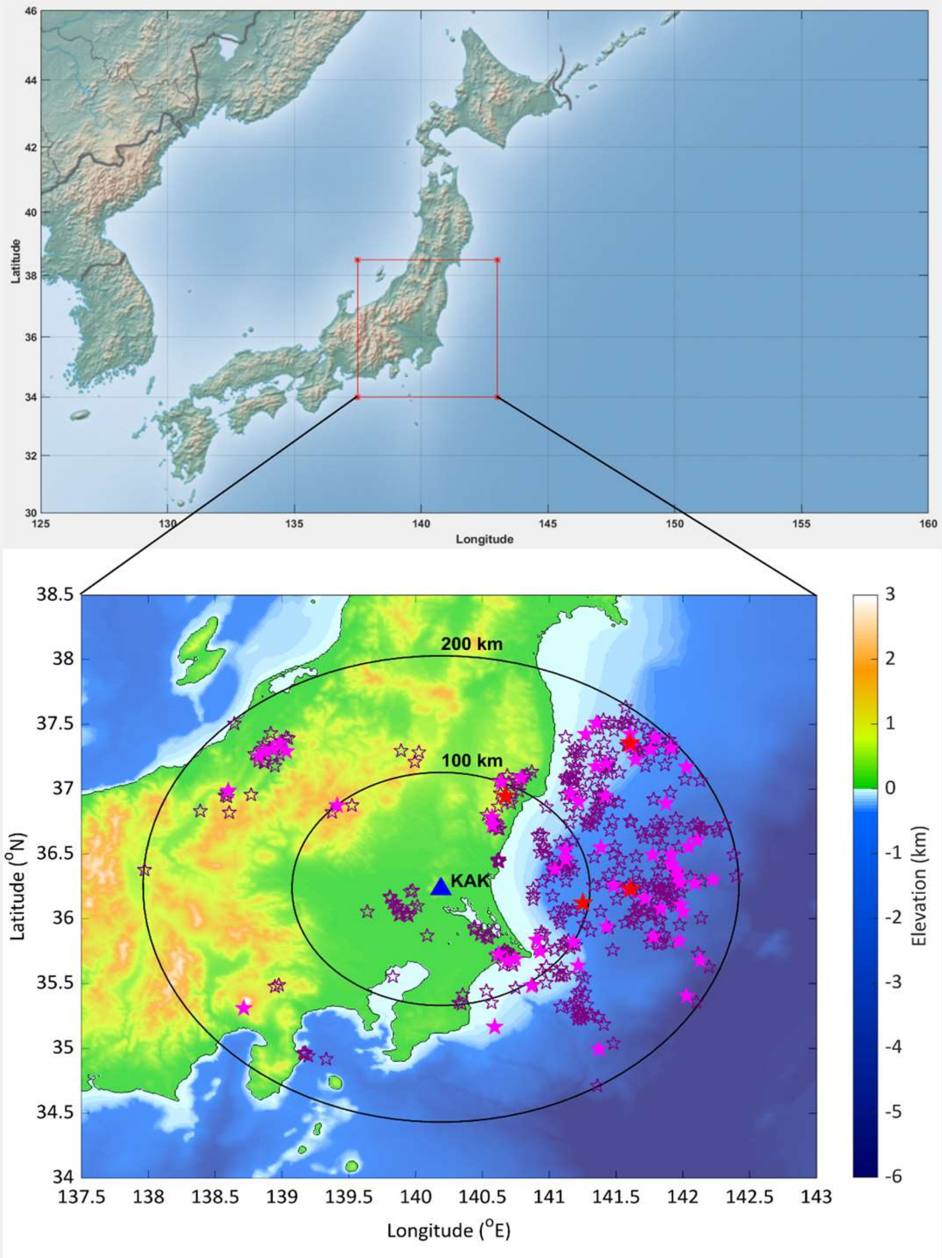
Development of quake-forecasting algorithm



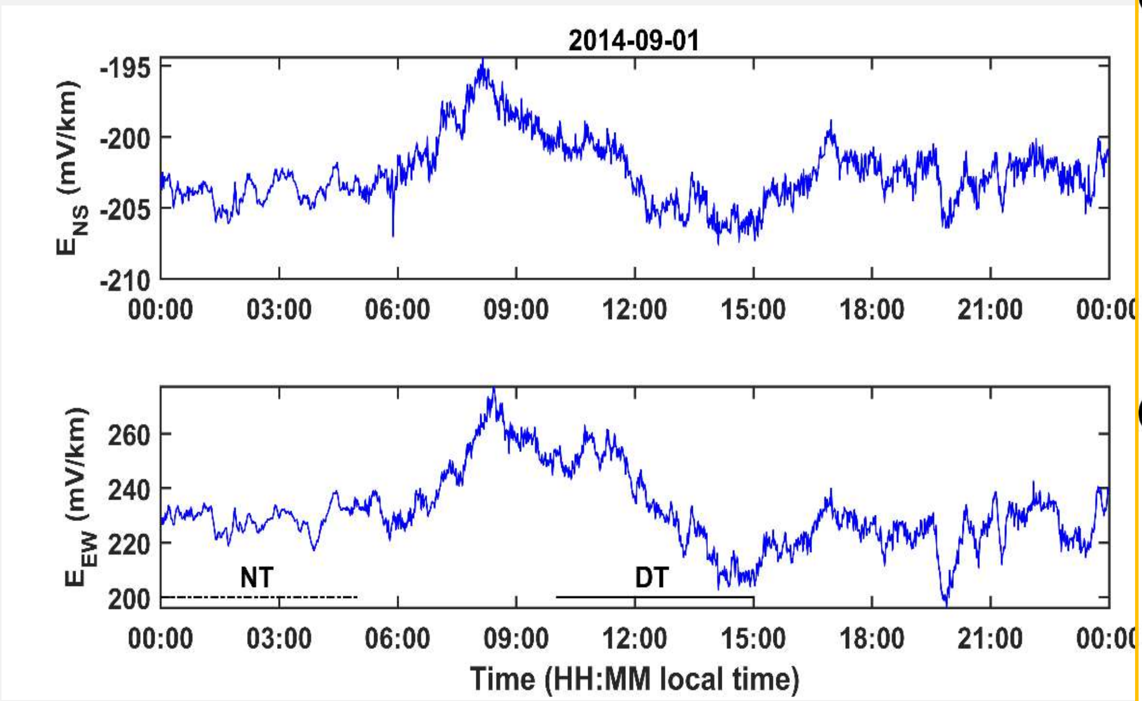
Testing the algorithm: Application to Kakioka, Japan

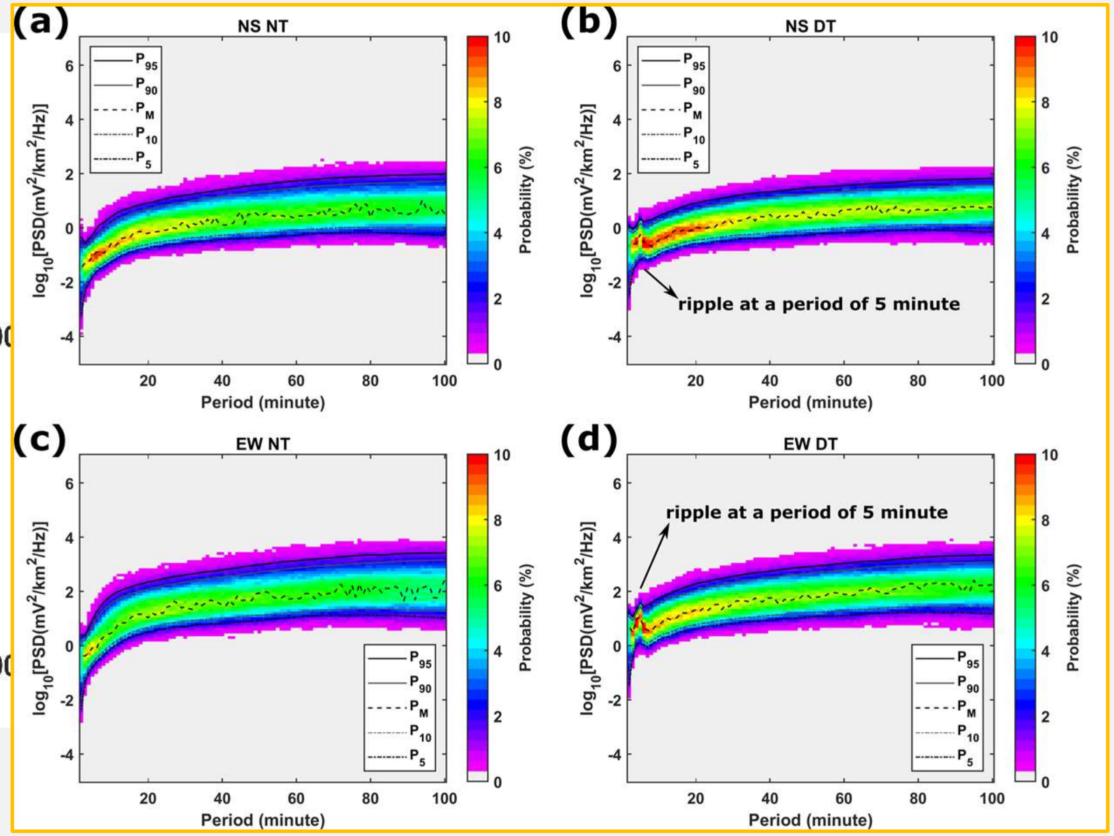
Station	Value
Name	KAK
Place	Kakioka
Longitude	140.1 E
Latitude	36.23 N
Year	**1993-2018 (26 years)**

Earthquake catalog	Value
Year	1993-2018
Magnitude	$M_L \geq 5$
Location	Around KAK within a 200-km radius
Symbol Color	Purple: M_L=5-6 Magenta: M_L=6-7 Red: $M_L \geq 7$



Geoelectric time series & probabilistic power spectral density



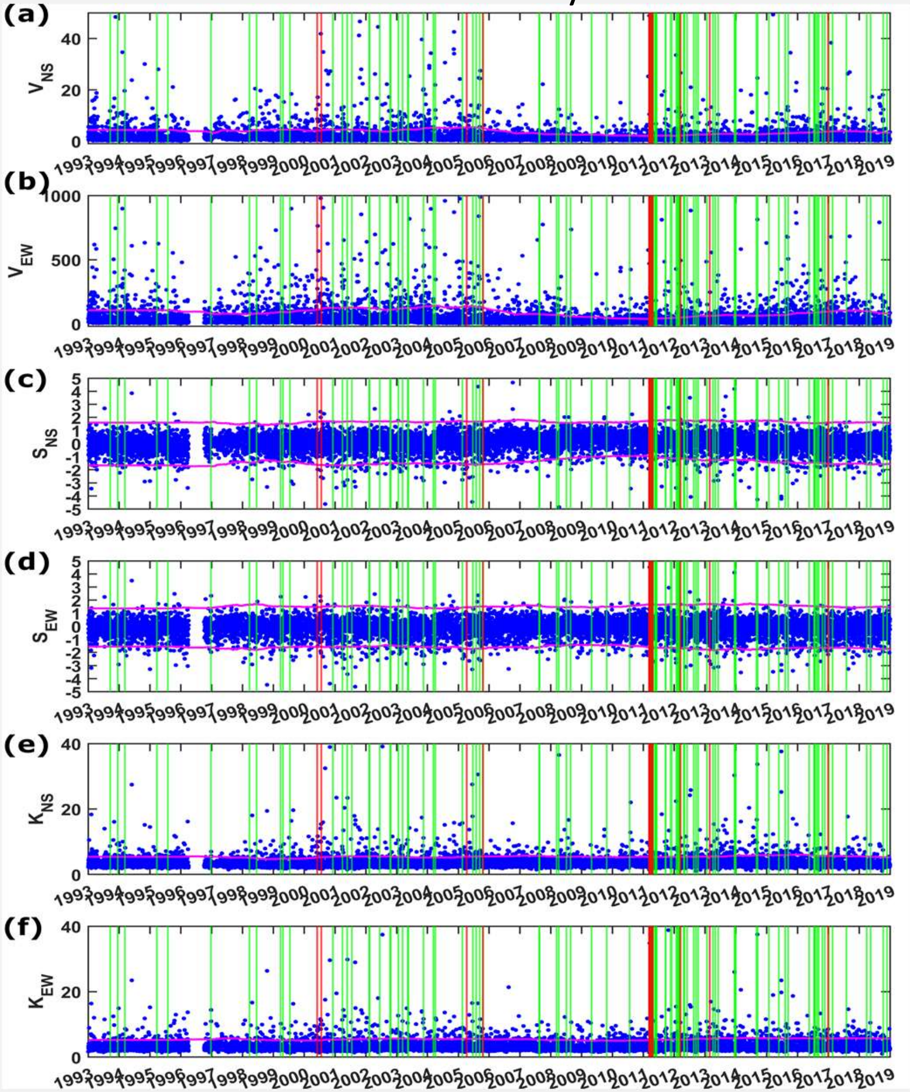


Sampling rate: 1 min for N- and E-components.
Nighttime (NT) data: smoothness
Daytime (DT) data: roughness

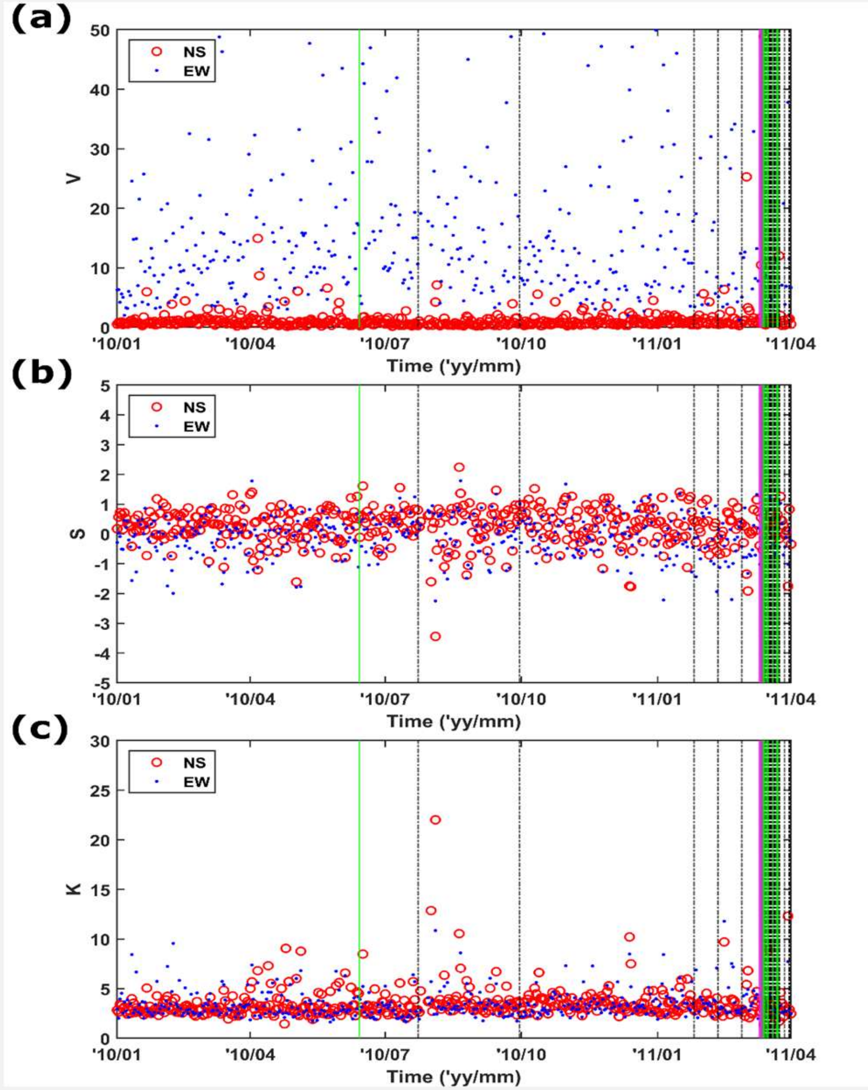
Comparing the PPSD of NT and DT data,
a clear ripple shows at a period of 5 min in the DT data.
We infer that such a ripple is generated by human activity.

Early-warning signal: Variance, skewness, kurtosis (VSK)

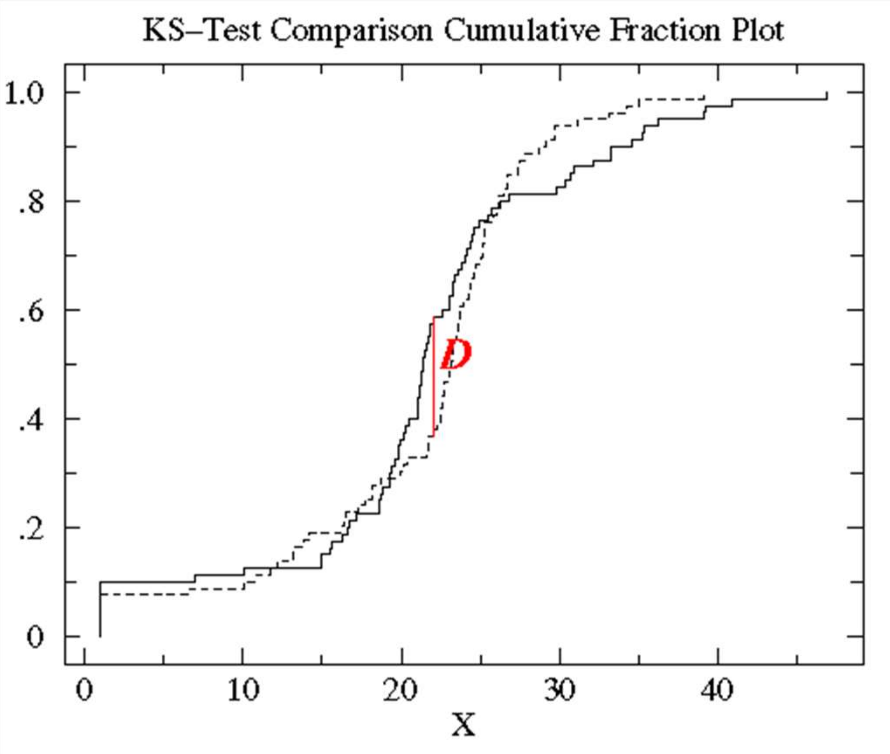
1993-2018: 26 years



Before 11 March 2011 (Tohoku EQ)



VSK distribution of Normal day vs DST<-50 day



Kolmogorov-Smirnov test returns a test decision for the null hypothesis that the data in sample 1 and sample 2 are from the same continuous distribution.

Data selection:
Normal day: the day with -50≤DST≤50
Magnetic disturbance day: DST<-50 and within the following 3 days

H0: Null hypothesis is that the data (VSK) for the normal days and magnetic disturbance days are from the same continuous distribution.

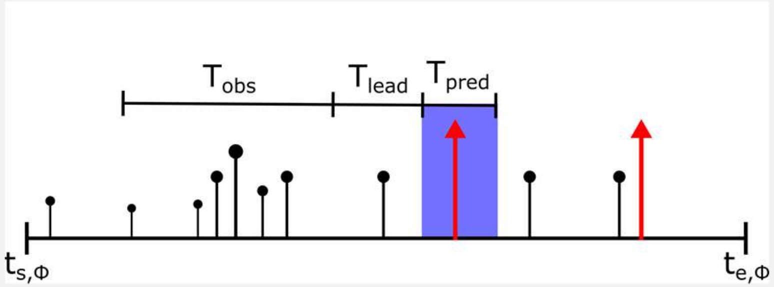
Alpha=0.05	H0	P
V_{NS}	1	0.0037
V_{EW}	1	0.0356
S_{NS}	1	5e-8
S_{EW}	1	2e-7
K_{NS}	1	6e-4
K_{EW}	1	0.0056

This explains that such an analysis algorithm using VSK indices can be less affected by the magnetic distances.

Model parameter description

Parameter	Description	Values
M_{thr}	Magnitude threshold for selecting target earthquakes	5:0.5:7
R_{thr}	Epicenter distance threshold for selecting target earthquakes	50:25:200
A_{thr}	A threshold for defining anomalous indices	1:1:10
P_{thr}	A threshold for issuing times of increased probability	0.1:0.1:0.7
T_{obs}	Observation time window	5:5:60
T_{lead}	Leading time window	0:5:60
T_{pred}	Prediction time window	1, 5:5:30
ts and te	Start date and end date of a time period	
tau	Ratio of alarmed cells	
n	Ratio of missed events	
d	fitness function: d = 1-tau-n	

Model formulation



$$\theta_u(t;y) = Median(y_i) + A_{thr} * IQR(y_i), i \in [t - \Delta t, t]$$
$$\theta_l(t;y) = Median(y_i) - A_{thr} * IQR(y_i), i \in [t - \Delta t, t]$$

$$F_{obs}(t|\boldsymbol{g}) = B[(r_i \le R_{thr}) \cdot (m_i \ge M_{thr}) \cdot (t = t_i)], t \in [t_s, t_e]$$

$$N_{AIN}(t|\boldsymbol{g}) = \sum_y B\left[\left(y(t) > \theta_u(t;y)\right) + \left(y(t) < \theta_l(t;y)\right)\right]$$

$$\bar{N}_{AIN,T_{obs}}(t|\boldsymbol{g}) = \frac{\sum_{i=t-T_{obs}+1}^{t} N_{AIN}(i|\boldsymbol{g})}{T_{obs}}$$

$$R_{AIN,T_{obs}}(t|\boldsymbol{g}) = \frac{\bar{N}_{AIN,T_{obs}}(t|\boldsymbol{g})}{\max_{t \in [t_s,t_e]} \bar{N}_{AIN,T_{obs}}(t|\boldsymbol{g})}$$

$$F_{at}(t|\boldsymbol{g}) = B\left[R_{AIN,T_{obs}}(t|\boldsymbol{g}) \ge P_{thr}\right]$$

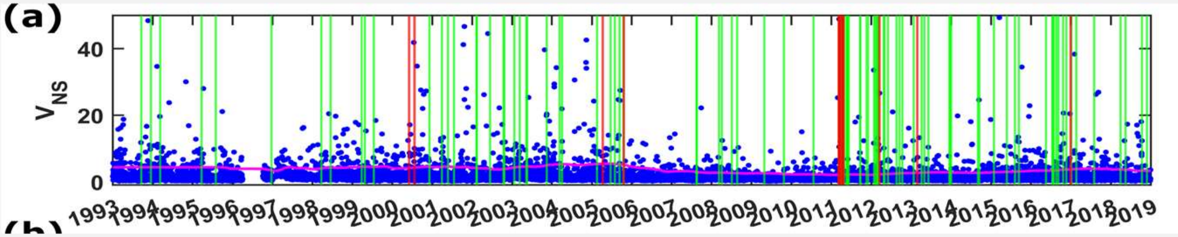
$$F_{tip}(t|\boldsymbol{g}) = B\left[\left(t = t_i|_{F_{at}=1} + T_{lead} + 1\right) + \left(t = t_i|_{F_{at}=1} + T_{lead} + 2\right) + \cdots + \left(t = t_i|_{F_{at}=1} + T_{lead} + T_{pred}\right)\right], t \in [t_s, t_e] \quad (11)$$

$$\tau(\boldsymbol{g}) = \frac{\sum_t F_{tip}(t|\boldsymbol{g})}{\sum_t B[F_{tip}(t|\boldsymbol{g}) \ge 0]}$$

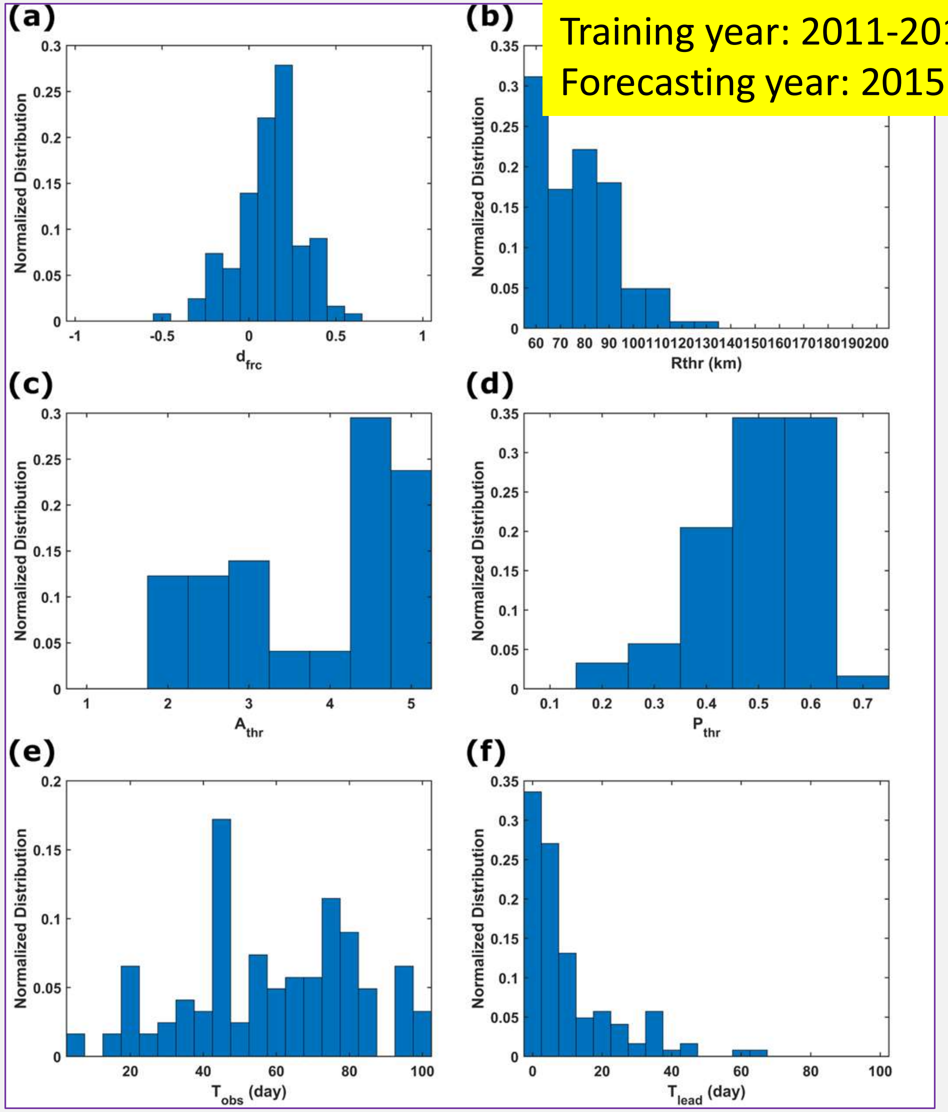
$$n(\boldsymbol{g}) = \frac{\sum_t B[(F_{tip}(t|\boldsymbol{g})=0) \cdot (F_{obs}(t|\boldsymbol{g})=1)]}{\sum_t B[(F_{tip}(t|\boldsymbol{g}) \ge 0) \cdot (F_{obs}(t|\boldsymbol{g})=1)]}$$

$$d(\boldsymbol{g}) = 1 - \tau(\boldsymbol{g}) - n(\boldsymbol{g})$$

4 years training + 1 year forecasting



Optimal model



Training year: 2011-2014
Forecasting year: 2015

Molchan Diagram

(1) Fraction of missing EQKs

$$\tau(\boldsymbol{g}) = \frac{\sum_t I(TIP(t|\boldsymbol{g})=1)}{\sum_t I(TIP(t|\boldsymbol{g})\geq 0)}$$

(2) Fraction of alarmed cells

$$n(\boldsymbol{g}) = \frac{\sum_t I(TIP(t|\boldsymbol{g})=0 \ \& \ EQK(t|\boldsymbol{g})=1)}{\sum_t I(TIP(t|\boldsymbol{g})\geq 0 \ \& \ EQK(t|\boldsymbol{g})=1)}$$

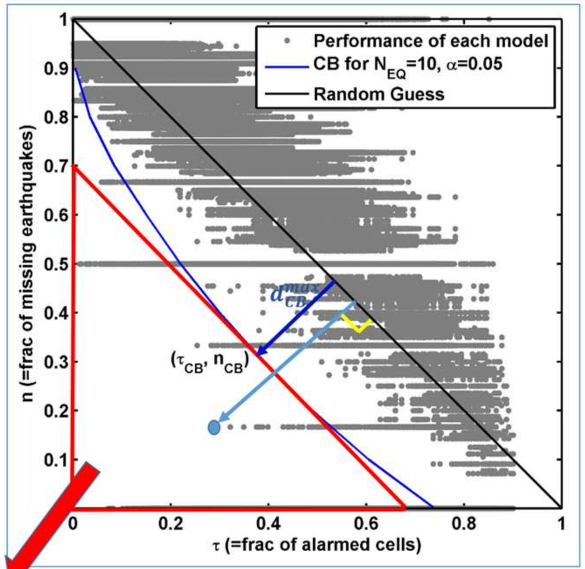
(3) Scoring of a parameter set

$$d(\boldsymbol{g}) = 1 - \tau(\boldsymbol{g}) - n(\boldsymbol{g})$$

The quantity d is the distance from a point to the random guess line.

1) d>0 means the performance is better than random.
2) d=0 means the performance is random.
3) d<0 means the performance is worse than random.

The model is significant when $d > d_{CB}^{max}$.



[Zechar & Jordan, 2008, GJI]

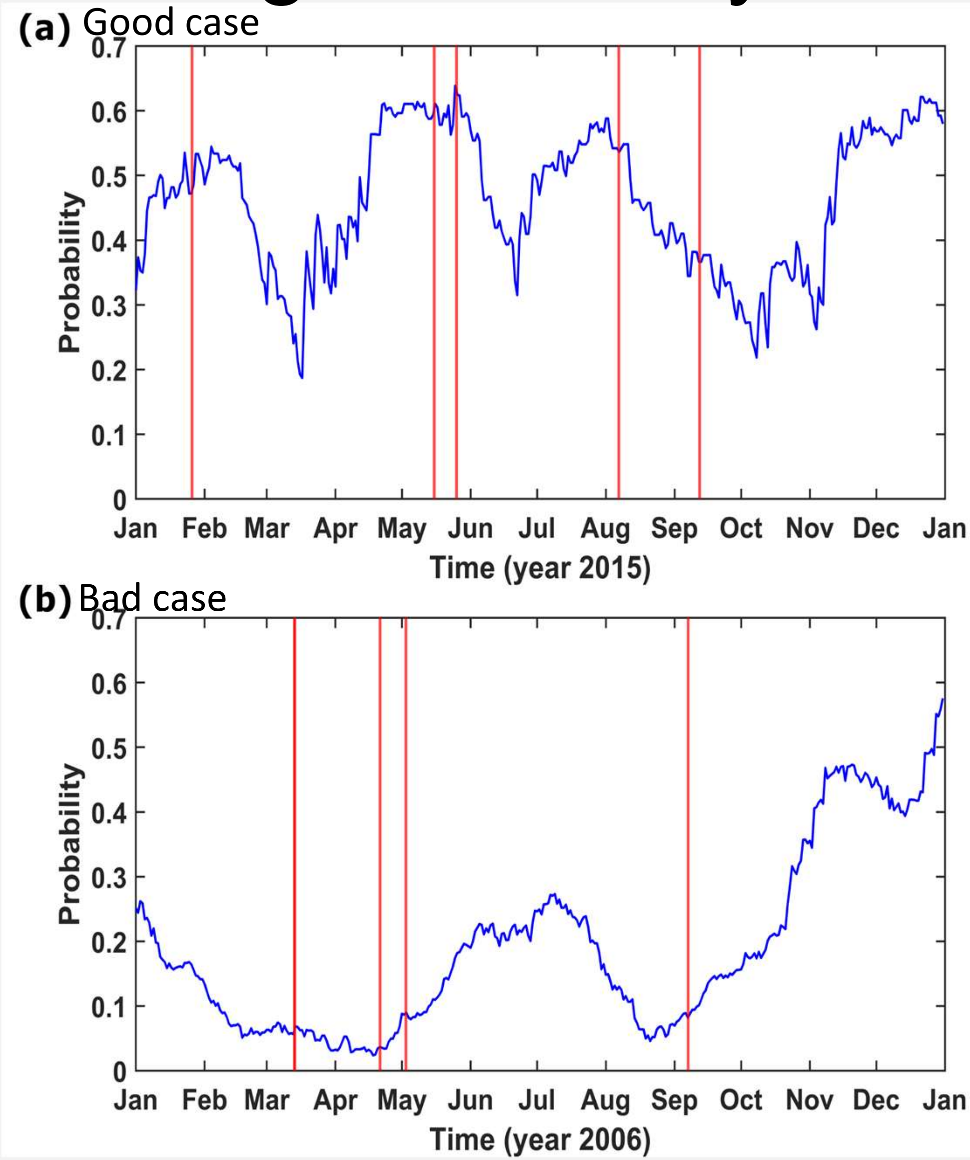
1) **Under the confidence bound:** Parameters mean informative
2) **Molchan diagram:** Evaluate whether a prediction strategy is good

Temporal Forecasting Probability

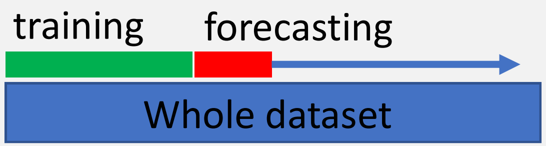
Definition of precursor-based earthquake probability

Constructing earthquake-forecasting probability P(x, y, t) using True Positive Rate multiplied by TIP.

$$P(x,y,t) = \frac{1}{N_{top}} \sum_{i=1}^{N_{top}} v(\boldsymbol{G}_i) \cdot T_{TIP}(x,y,t|\boldsymbol{G}_i),$$



Forecasting performance



Performance
D_{avg} = 0.49
PG_{avg} = 3.68
14 positives out of 22

Case	Training phase 1	Training phase 2	Forecasting phase	PG	D	ρ
1	1993/1/1-1994/12/31	1995/1/1-1996/12/31	1997/1/1-1997/12/31	2.94	0.66	1.3
2	1994/1/1-1995/12/31	1996/1/1-1997/12/31	1998/1/1-1998/12/31	1.96	0.39	1.48
3	1995/1/1-1996/12/31	1997/1/1-1998/12/31	1999/1/1-1999/12/31	1.72	0.42	1.14
4	1996/1/1-1997/12/31	1998/1/1-1999/12/31	2000/1/1-2000/12/31	1.69	0.41	0.72
5	1997/1/1-1998/12/31	1999/1/1-2000/12/31	2001/1/1-2001/12/31	11.41	0.85	1.91
6	1998/1/1-1999/12/31	2000/1/1-2001/12/31	2002/1/1-2002/12/31	2.06	0.48	0.83
7	1999/1/1-2000/12/31	2001/1/1-2002/12/31	2003/1/1-2003/12/31	1.69	0.34	1.84
8	2000/1/1-2001/12/31	2002/1/1-2003/12/31	2004/1/1-2004/12/31	3.08	0.59	1.66
9	2001/1/1-2002/12/31	2003/1/1-2004/12/31	2005/1/1-2005/12/31	3.43	0.49	1.93
10	2002/1/1-2003/12/31	2004/1/1-2005/12/31	2006/1/1-2006/12/31	3.61	0.72	0.1
11	2003/1/1-2004/12/31	2005/1/1-2006/12/31	2007/1/1-2007/12/31	3.07	0.67	2
12	2004/1/1-2005/12/31	2006/1/1-2007/12/31	2008/1/1-2008/12/31	4.52	0.52	1.57
13	2005/1/1-2006/12/31	2007/1/1-2008/12/31	2009/1/1-2009/12/31	2.5	0.35	1.28
14	2006/1/1-2007/12/31	2008/1/1-2009/12/31	2010/1/1-2010/12/31	15.21	0.93	0.83
15	2007/1/1-2008/12/31	2009/1/1-2010/12/31	2011/1/1-2011/12/31	1.71	0.41	1.49
16	2008/1/1-2009/12/31	2010/1/1-2011/12/31	2012/1/1-2012/12/31	1.45	0.28	0.88
17	2009/1/1-2010/12/31	2011/1/1-2012/12/31	2013/1/1-2013/12/31	1.26	0.14	0.08
18	2010/1/1-2011/12/31	2012/1/1-2013/12/31	2014/1/1-2014/12/31	6.76	0.85	1.42
19	2011/1/1-2012/12/31	2013/1/1-2014/12/31	2015/1/1-2015/12/31	3.38	0.55	4.08
20	2012/1/1-2013/12/31	2014/1/1-2015/12/31	2016/1/1-2016/12/31	1.51	0.22	0.12
21	2013/1/1-2014/12/31	2015/1/1-2016/12/31	2017/1/1-2017/12/31	1.4	0.14	0.14
22	2014/1/1-2015/12/31	2016/1/1-2017/12/31	2018/1/1-2018/12/31	4.56	0.39	1.1

Conclusion

Approaching the operational forecasting, we will need:

- Multi-stations
- Continuous data with long-term monitoring
- Other precursory indicators
- Other physical or chemical quantities

- **Contact: Hong-Jia Chen (redhouse6341@gmail.com)**

Thank you for listening!