

12-2 予測の評価

Evaluation of earthquake forecast

東京大学地震研究所

Earthquake Research Institute, The University of Tokyo

1. はじめに

現在日本においては、地震活動に基づいた地震活動評価による地震発生予測検証実験を Collaboratory for the Study of Earthquake Predictability (CSEP) プロトコルに基づき実施されている。CSEP とは、グローバルな地震予測の検証環境を厳密に規定し、地震予測を客観的に評価する枠組みのことである。すべての地震予測モデルは、同一の地震カタログを用いて予測パラメータが最適化され、予測がなされる。その予測は、prospective に実施されるので、モデラーによる予測の変更が原理的にできない仕組みであり、予測結果は共通の手法により評価がなされる。モデラーとは異なる第三者により地震研究所に設置されたテストセンターにおいて、その予測検証実験が遂行されている。日本においては、AllJapan, Mainland, Kanto の予測領域、1 日、3 ヶ月、1 年、3 年の予測期間が設定され、合計 12(=3x4)のテストが行われている（第 1 図）。

2. CSEP のテスト

地震の予測が観測結果を満たしているかどうかの評価に、CSEP においては、地震数の予測が観測を満たしているかどうかをチェックする N-テスト、地震活動の分布が観測を満たしているかどうかの S-テスト、予測のマグニチュードの規模別頻度分布が観測を満たしているかどうかの M-テスト、これら 3 つのテストの総合テストである L-テストが統計的に厳密に行われる。これらのテストは consistency テストとよばれるが、観測結果は一つであるため、その評価の手法に特徴がある。これらのテストの手順は以下のようなものとなる。①期待値マップを元に数値シミュレーションにより地震を発生させる。② ①の疑似地震により対数尤度を計算する。③ ①、②を繰り返して（通常は 1000 回）対数尤度の分布を得る。④ 実際に観測された地震による対数尤度が③の分布のどこにくるかを見る（第 2 図）。この分布は、第 3 図をみてわかるように地震予測モデル毎に異なり、このテストから得られる評価値（スコア）はモデル間の絶対的な評価ではなく相対的なものであることが特徴である。

3. 予測結果

表 1 から 3 に、AllJapan テスト領域の 3 ヶ月テストクラスの結果を示す。東北沖地震直前のラウンドにおいては、すべての予測モデルは観測数を課題評価しており、N-テストをパスしていない。東北沖地震を含むラウンドにおいては、すべての予測モデルがテストをパスしていないことがわかる。東北沖地震後は予測モデルがテストをパスしないラウンドが続くが、2012/8/1 からの 12 ラウンドから、予測モデルがテストをパスする状況に回復している。

4. CSEP テストの問題点

CSEP のテストは、(1) L-テストは、地震予測数を少し多めであるとパスしやすい。(2) N-テスト結果と L-テスト結果は相関する。(3) N-テスト結果と S-テスト結果は、相関がない。(4) M-テストはほぼパスする。(予測モデルはほぼ GR 則を満たす。(5) N-テストは地震観測数が増えるとパスしに

くなる．などの特徴があるが，スコアがモデル毎の相対的な指標であるということと，日本においては S-テストをパスしないモデルが多いことがある．また，このテストの評価は，「予測が観測を満たしている」ということではなく，テストをパスの意味するところは「予測モデルは観測を満たしていないとはいえない」が正しい解釈である．また，M-テストは下限マグニチュードを変えるとスコアが変わる．テスト領域の解像度を変えて S-テストを行うとスコアが変わるなど，評価テストの改善・改良が今後が必要になってくると考えている．

5. 終わりに

2009 年 11 月 1 日から，日本における地震発生予測検証実験がスタートし，これまでに数多くの検証結果が得られている．予測の評価手法である CSEP の consistency テストの特性も徐々に明らかとなり，テスト自体の改善・改良が必要であるが，この検証実験を継続して実施することが予測モデルのさらなる向上につながると考えている．

(鶴岡 弘)

Rules of CSEP-Japan

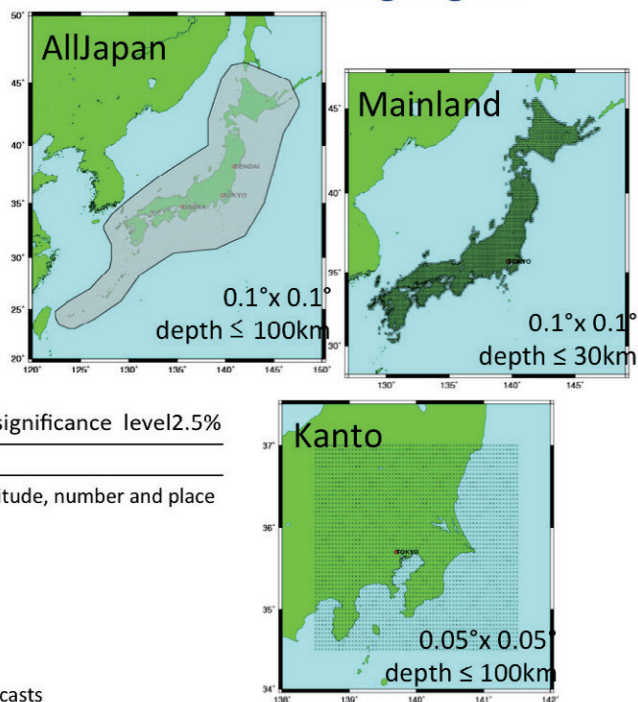
Test class

1-day	$4 \leq M \leq 9$
3-month	$4 \leq M \leq 9$
1-year	$5 \leq M \leq 9$
3-year	$5 \leq M \leq 9$

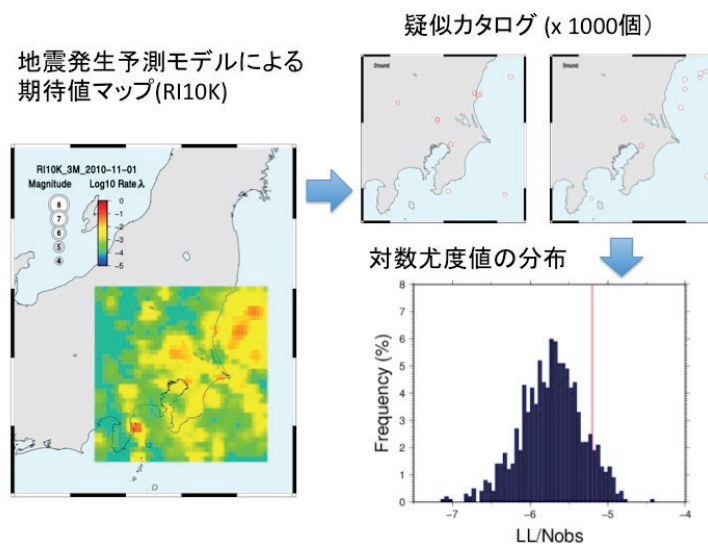
Evaluation

test		significance level 2.5%
L	Likelihood of distributions of magnitude, number and place	
CL	Conditional L-test	
N	Event number	
M	Magnitude distribution	
S	Spatial distribution	
W/T	Pair-wise comparison between forecasts	

Testing region

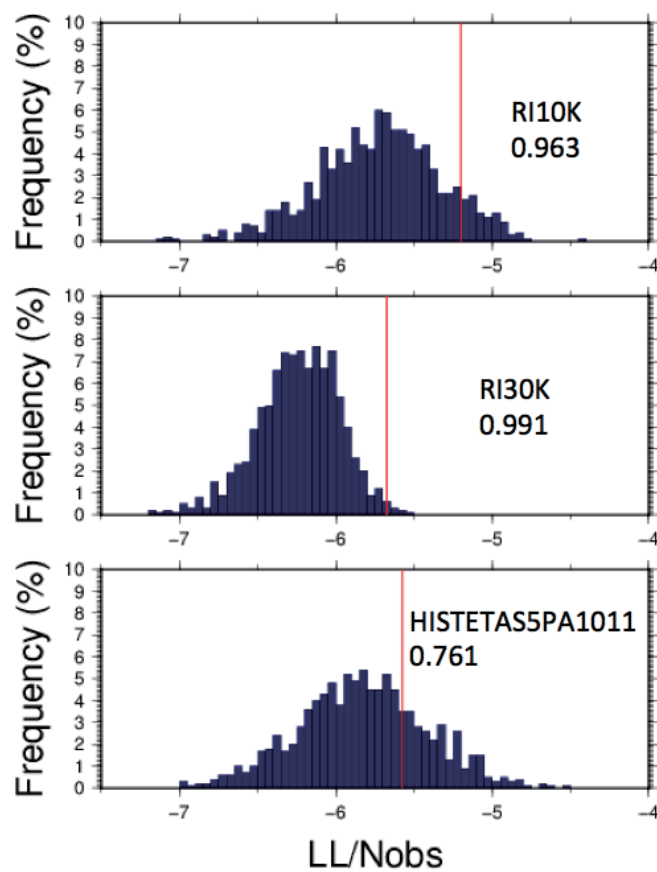


第 1 図．日本の地震予測検証実験のルール．
Fig.1. Rules of CSEP-Japan.



第 2 図．観測分布の導出．

Fig. 2. Generation of the observation distribution.



第 3 図．モデル毎の対数尤度分布 (モデル : RI10K, RI30K, HISTETAS5PA1011)

Fig. 3. Log likelihood distribution of 3 models (Model: RI10, RI30K, HISTETAS5PA1011)

TEST		L	N(under)	N(over)	M	S	forecast(obs)	fore/obs
MODEL	LL(L)	gamma	delta1	delta2	kappa	zeta	forecast(obs)	fore/obs
HISTETAS5PA	-553.96521	1.000	1.00000	0.00000	0.262	0.155	132.21126(68.0)	1.94
HISTETAS7PA	-554.48529	1.000	1.00000	0.00000	0.257	0.479	132.20472(68.0)	1.94
RI10K	-555.94433	1.000	1.00000	0.00000	0.284	0.065	124.21431(68.0)	1.83
MARFS	-574.73038	1.000	1.00000	0.00000	0.349	0.001	130.57214(68.0)	1.92
RI30K	-576.21008	1.000	1.00000	0.00000	0.334	0.037	124.21430(68.0)	1.83
RI50K	-587.82008	1.000	1.00000	0.00000	0.297	0.082	124.21430(68.0)	1.83
RI100K	-599.47050	1.000	1.00000	0.00000	0.332	0.291	124.21430(68.0)	1.83
TRIPLES1011	-616.66693	1.000	1.00000	0.00000	0.561	0.647	122.89603(68.0)	1.81

表 1. AllJapan の第 5 ラウンド(2010/11/1-2011/2/1)の評価結果.

Table 1. Evaluation results for AllJapan 5round (2010/11/1-2011/2/1)

TEST		L	N(under)	N(over)	M	S	forecast(obs)	fore/obs
MODEL	LL(L)	gamma	delta1	delta2	kappa	zeta	forecast(obs)	fore/obs
HISTETAS7PA	-26311.15385	0.000	0.00000	1.00000	0.000	0.000	126.75051(3702.0)	0.03
HISTETAS5PA	-26439.18496	0.000	0.00000	1.00000	0.000	0.000	126.75947(3702.0)	0.03
RI100K	-26482.17245	0.000	0.00000	1.00000	0.000	0.000	119.41082(3702.0)	0.03
RI50K	-26766.78596	0.000	0.00000	1.00000	0.000	0.000	119.41082(3702.0)	0.03
RI30K	-26964.84176	0.000	0.00000	1.00000	0.000	0.000	119.41082(3702.0)	0.03
TRIPLES1011	-27729.81732	0.000	0.00000	1.00000	0.000	0.000	118.55716(3702.0)	0.03
RI10K	-27800.54045	0.000	0.00000	1.00000	0.000	0.000	119.41082(3702.0)	0.03
MARFS	-27890.97891	0.000	0.00000	1.00000	0.000	0.000	114.74027(3702.0)	0.03

表 2. AllJapan の第 6 ラウンド(2011/2/1-2011/5/1)の評価結果.

Table 1. Evaluation results for AllJapan 6round (2011/2/1-2011/5/1)

TEST		L	N(under)	N(over)	M	S	forecast(obs)	fore/obs
MODEL	LL(L)	gamma	delta1	delta2	kappa	zeta	forecast(obs)	fore/obs
HISTETAS5PA	-1185.07647	0.969	0.94834	0.05970	0.768	0.878	209.81296(187.0)	1.12
RI10K	-1195.06581	0.577	0.34128	0.68527	0.831	0.863	181.13562(187.0)	0.97
MARFS	-1195.17379	0.658	0.42438	0.60411	0.875	0.839	184.07388(187.0)	0.98
RI30K	-1229.50003	0.689	0.34128	0.68527	0.830	0.988	181.13562(187.0)	0.97
HISTETAS7PA	-1231.85582	0.994	0.94713	0.06106	0.722	0.996	209.64589(187.0)	1.12
RI50K	-1257.38876	0.685	0.34128	0.68527	0.841	1.000	181.13562(187.0)	0.97
TRIPLES1011	-1274.27537	0.116	0.00183	0.99856	0.582	0.997	149.73916(187.0)	0.80
RI100K	-1302.52153	0.700	0.34128	0.68527	0.846	1.000	181.13562(187.0)	0.97

表 3. AllJapan の第 12 ラウンド(2012/8/1-2012/11/1)の評価結果.

Table 1. Evaluation results for AllJapan 12round (2012/8/1-2012/11/1)