地　震　　第 2 輯
第 70 巻（2018）195-213 頁
DOI: 10.4294/zisin.2016-17

論　　説

繰り返し回数の少ない繰り返し地震系列に対する長期的地震発生確率予測の成績と検証

気象研究所地震津波研究部*　田 中 昌 之
岡 田 正 実**
東北大学大学院理学研究科地震・噴火予知研究観測センター***　内 田 直 希

Performance and Validation of Long-Term Predictability for Repeating Earthquake Groups with a Few Recurrences

Masayuki TANAKA
Seismology and Tsunami Research Department, Meteorological Research Institute, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan

Masami OKADA
3-19-6 Minami, Ushiku, Ibaraki 300-1222, Japan

Naoki UCHIDA
Research Center for Prediction of Earthquakes and Volcanic Eruptions, Graduate School of Science, Tohoku University, 6-6 Aramaki, Aoba-ku, Sendai 980-8578, Japan

(Received December 21, 2016; Accepted September 20, 2017; published online on December 18, 2017)

Only a few large or medium-sized repeating earthquakes are known in many sequences due to long recurrence intervals in comparison with observation period. Recurrence intervals are used to estimate long-term earthquake probability; however, the effect of the number of recurrence intervals on prediction performance is unclear. We studied the predictability dependence on the number of recurrences using small interplate repeating earthquakes along the Japan Trench. These earthquakes were extracted from Tohoku University's catalog, and this data was used in the probability forecast experiments from 2006 to 2010. The number of forecasts is 524. Two to five events just prior to the forecasts are picked from each sequence to calculate the probabilities. We calculated the probabilities using the Bayesian statistics log-normal distribution model (LN-Bayes), the log-normal distribution model based on the small sample theory (LN-SST), and the exponential distribution model (Exp-pin). We then evaluated the forecast results using mean log-likelihood and Brier score. The performance of the LN-SST and LN-Bayes models was better than that of the Exp-pin model for almost all cases. In addition, we conducted some statistical tests to measure the consistency of forecast with the observed catalog and confirmed the tendency of underestimation of probabilities and the accuracy dependence of probabilities on the number of recurrences for all models. The LN-Bayes and LN-SST models were examined by random number experiments using a large number of simulated sequences of earthquakes. In statistical tests, we changed the number of repetitions and the elapsed time from the most recent earth-

* 〒305-0052 茨城県つくば市長峰 1-1
** 〒300-1222 茨城県牛久市南 3-19-6
*** 〒980-8578 宮城県仙台市青葉区荒巻字青葉 6-6

quake. As a result, performance improvement, along with the increase of repetition number, was evident when the repetition number was small for both models. The LN-Bayes model is generally better than the LN-SST model if the number of repetitions is very small. However, one of the features of the LN-Bayes model is that probability is saturated when the number of repetitive events is small and longer time has passed since the last event.

Key words: Repeating earthquake, Earthquake forecast, Bayesian approach, Mean log-likelihood, Brier score

§1. は じ め に

ほぼ同じ場所で繰り返す相似地震（小繰り返し地震）は，プレート境界の小さな固着域の破壊で発生すると考えられており，プレート運動を調べる際などに使われている [Igarashi *et al.* (2003), Uchida *et al.* (2004)]. 波形の類似性から，他の地震との客観的識別が比較的に容易であること，規模の大きなものよりも系列数が多く，かつ短期間に繰り返すことなどから，予測研究にも好都合である［例えば，岡田・他 (2008, 2009, 2010, 2011)］. 一方，繰り返し大・中地震の予測は古くから行われており，例えば，Bakun and McEvilly (1984) と Bakun and Lindh (1985) は Parkfield の地震を対象に，Nishenko (1991) は環太平洋で発生する大地震について，プレート境界付近に多数のセグメントを設定して確率予測を行った. Matsuzawa *et al.* (1999, 2002) や長谷川 (2002) は，岩手県釜石沖で繰り返して発生する M4.8 前後の地震に着目した予測を行っている.

地震調査研究推進本部地震調査委員会 (2001) では，国内の主要な活断層及びプレート境界付近の海溝型地震を対象に長期評価を実施し，地震発生確率値を毎年公表している. 確率計算には，更新過程の Brownian Passage Time (BPT) 分布モデルが用いられている. ただし，直近地震（予測時点に最も近い地震）の発生時期が分からないときなどは，指数分布モデルが使われる. BPT 分布はブラウン運動のような揺らぎを伴った定常的なひずみ蓄積と，応力限界での破壊までの時間間隔分布を表すので，地震発生過程との対応はよいが，統計的取り扱いは難しい［Matthews *et al.* (2002), 宇津 (1999, p478)］. BPT 分布の事前分布については，Nomura *et al.* (2011) が繰り返し大地震について調べているが，相似地震に関しては行われていない. このような事情から，本研究では BPT 分布を用いた地震予測は取り上げていない.

一方，対数正規分布を用いた地震予測としては，岡田・他 (2008, 2009, 2010, 2011) により，2006 年 7 月から 2010 年にかけて計 4 回，東北大学作成の相似地震カタログを用いて，更新過程のベイズ統計対数正規分布モデルで予測した実験（以下，「岡田らによる確率予測実験」と呼ぶ）がある. 岡田・他 (2007) によれば，この統計モデルは，事前分布とベイズの定理を用いており，地震の発生間隔が 1 個でも予測確率の計算は可能である.

地震調査研究推進本部地震調査委員会 (2001) が長期評価の対象としている大規模地震や，気象研究所地震火山研究部・他 (2014) などが調査した中規模の繰り返し相似地震は，発生間隔が数年以上と長い. 加えて，昔の地震観測点は今よりもはるかに少なく，地震の検知力は劣り，その後，高感度地震計は整備されたものの，データの保存期間が短いことなどから，地震の繰り返しが 1 回の系列もある. このような繰り返し回数が非常に少ない系列の確率予測にもベイズ統計対数正規分布モデルは適用可能であるが，予測の精度は定かではない. 岡田らによる確率予測実験では，1 系列の地震数が 5 個以上（地震の発生間隔が 4 個以上）の系列を対象にしており，5 個未満の事例はない [Okada *et al.* (2012)]. 地震の発生間隔のデータ数が少なくなれば，使用できる情報が少なくなるので，予測の精度・信頼性の低下は避けられない.

本報告は，予測の精度や成績などがデータ数，予測時期（直近地震からの経過時間）によってどのように変化するかを定量的に調べることを目的とする. 具体的には，岡田らによる確率予測実験で使われた相似地震カタログを用い，地震の発生間隔が 4 個未満の確率予測について，発生間隔のデータ数が変わることに伴う予測成績の変動を調べた. 成績及び成績差については，統計的検定で有意性を調べた. 岡田らによる確率予測実験で事前公開されたのは，ベイズ統計対数正規分布モデルによる結果のみであったが，更新過程の小標本論対数正規分布モデルについても調査した. そして，これら 2 種類の統計モデルによる予測成績との比較用として，更新過程のなかで最も基本的な指数分布モデルを使用した. また，実際の地震活動には，長期的な変動や続発性など，予測モデルが想定していない変動が含まれていることを考慮し，予測モデルの条件をほぼ完全に満たす乱数でシミュレーション（乱数実験）を行った. 乱数実験は，発生間隔のデータ数や予測時期を変えた複数パターンについて行い，相似地震を用いた調査と類似の評価手法で，予測の精度と成績の発生間隔のデータ数依存性を調べ，相似地震の結果との整合性を考察した.

§2.　理　　論

2.1　計算条件

発生確率の計算は，Okada *et al.* (2012) による方法を用いた．データは地震の発生間隔のみを使い，更新過程を用いて数値計算する．その際，予測時点より前に発生した $n+1$ 個の地震の発生時期は，その間に漏れが無く判明しているものとし，n 個の地震の発生間隔 t_i は，互いに独立であるとする．2 種類の統計モデル［ベイズ統計対数正規分布モデル (LN-Bayes)，小標本論対数正規分布モデル (LN-SST)］で発生確率を計算する．また，比較用として，予測時期に依存しない指数分布モデル (Exp-pin) でも発生確率を計算する．

予測時点において，地震数 $n+1$，地震の発生間隔 t_i（日）の自然対数 x_i (＝log t_i) の平均値 $\bar{x}$ と標準偏差 s は，

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \tag{1}$$

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{2}$$

で既知である．直近地震から予測時点までと予測期間終了時点までの経過時間（日）を T_p, T_f とする．T_p から T_f の間に当該地震が 1 個以上発現する確率 P_q は，

$$P_q(T_p,T_f;\theta)=\frac{F(T_f;\theta)-F(T_p;\theta)}{1-F(T_p;\theta)} \tag{3}$$

の条件付き確率で与えられる．ここで，$F(T;\theta)$ は発生間隔 T の分布関数で，θ は分布関数に含まれる全パラメータを意味する．

相似地震カタログを使って発生確率を計算する場合の予測日は，岡田らによる確率予測実験と同じ日（第 1 回：2006 年 7 月 1 日，第 2 回：2008 年 1 月 1 日，第 3 回：2009 年 1 月 1 日，第 4 回：2010 年 1 月 1 日）とし，予測期間 (T_p, T_f) も，岡田らによる確率予測実験と同じく 1 年間（365 日または 366 日）とした．

2.2　統計モデル

2.2.1　ベイズ統計対数正規分布モデル (LN-Bayes)

Okada *et al.* (2012) による方法を用いた．地震の発生間隔 t_i（日）の自然対数 x_i(＝log t_i) が正規分布 $N(\mu,\sigma^2)$ に従うものとする．母集団の平均値 μ の事前分布は，無限区間 $(-\infty,\infty)$ で一様 (＝1) とする．母集団の分散値 σ^2 の事前分布は，半無限区間 $[0,\infty)$ で定義される自然共役な逆ガンマ分布である．予測期間 (T_p, T_f) に当該地震が 1 個以上発現する確率（期待値）$\bar{P}_q$ は，

$$\bar{P}_q=\frac{F_t(Z_f;n+2\phi-1)-F_t(Z_p;n+2\phi-1)}{1-F_t(Z_p;n+2\phi-1)} \tag{4}$$

$$Z_p=\sqrt{n(n+2\phi-1)/(n+1)(ns^2+2\zeta)}\,(x_p-\bar{x})$$

$$Z_f=\sqrt{n(n+2\phi-1)/(n+1)(ns^2+2\zeta)}\,(x_f-\bar{x})$$

で求まる [Okada *et al.* (2012)]．ここで，F_t は t 分布の分布関数を表し，セミコロン“；”以下は自由度である．また，変数 Z は，

$$Z=\sqrt{\frac{n(n+2\phi-1)}{(n+1)(ns^2+2\zeta)}}\,(x_{n+1}-\bar{x}) \tag{5}$$

で自由度 $n+2\phi-1$ の t 分布に従う．x_{n+1} は直近地震から次地震までの発生間隔を表す変数で，予測時点では未知である．ϕ と ζ は，母数 σ^2 の事前分布である逆ガンマ分布 $\Gamma_R(y\,|\,\phi,\zeta)$

$$\Gamma_R(y\,|\,\phi,\zeta)=\frac{1}{\zeta\Gamma(\phi)}\left(\frac{\zeta}{y}\right)^{\phi+1}\exp\left(-\frac{\zeta}{y}\right) \tag{6}$$

の形状パラメータとスケールパラメータである．

今回採用した ϕ と ζ の値は，岡田らによる確率予測実験時の値（2006.7〜2007.6 と 2008.1〜12 の実験：$\phi=2.5$，$\zeta=0.44$，及び 2009.1〜12 と 2010.1〜12 の実験：$\phi=1.5$，$\zeta=0.15$）をそのまま用いた．なお，2009 年以降の実験については，2008 年の実験で使用した系列を使って ϕ と ζ の値を再検討し，ϕ と ζ ともに小さな値 ($\phi=1.5$, $\zeta=0.15$) を用いた場合に成績向上が見込めるとして ϕ と ζ の値を変えている [Okada *et al.* (2012)]．

2.2.2　小標本論対数正規分布モデル (LN-SST)

地震の発生間隔 t_i（日）の自然対数 x_i (＝log t_i) が同一の正規分布 $N(\mu,\sigma^2)$ に従うものとする．正規分布の性質から，$n+1$ 個の発生間隔 x_i で作った変数 Z は，

$$Z=\sqrt{\frac{(\mathrm{n}-1)}{(\mathrm{n}+1)}}\,\frac{(x_{\mathrm{n}+1}-\bar{x})}{s} \tag{7}$$

で自由度 $n-1$ の t 分布に従う．ここで，$\bar{x}$ と s は，(1) 式と (2) 式の平均値と標準偏差で，$x_{\mathrm{n}+1}$ は直近地震から次地震までの発生間隔を表す変数で，予測時点では未知で $N(\mu,\sigma^2)$ に従う．予測期間 (T_p, T_f) に当該地震が 1 個以上発現する確率 $\bar{P}_q$ は，区間 $(x_\mathrm{p}, x_\mathrm{f})$ における $x_{\mathrm{n}+1}$ の確率と同じであり，

$$\bar{P}_q=\frac{F_t(Z_f;n-1)-F_t(Z_p;n-1)}{1-F_t(Z_p;n-1)} \tag{8}$$

$$Z_p=\sqrt{(n-1)/(n+1)}\,(x_p-\bar{x})/s$$

$$Z_f=\sqrt{(n-1)/(n+1)}\,(x_f-\bar{x})/s$$

で求まる [Okada *et al.* (2012)]．ここで，F_t は t 分布の分布関数を表し，セミコロン“；”以下は自由度を示す．$n=1$ のときは自由度が 0 となり，確率は計算できない．この統計モデルは，発生間隔が対数正規分布に従うことのみを仮定しており，母集団の平均 μ や分散 σ^2 には依存しない．

2.2.3　指数分布モデル (Exp-pin)

地震が時間的に一様かつランダムに発生し，地震の発生間隔が指数分布に従うとする．発生間隔が n 個のとき，発生間隔 t_i（日）の平均値を $\bar{t}$ (＝$1/n\sum_{i=1}^{n}t_i$) とすると，予測期間 (T_p, T_f) に当該地震が 1 個以上発現する確

率 P_q は,

$$P_q(T_p, T_f)=1-\exp\left\{-\frac{(T_f-T_p)}{\bar{t}}\right\} \tag{9}$$

で求まる．発生確率 P_q は，予測期間の長さ T_f-T_P に依存し，直近地震から予測時点までの経過時間 T_p に対して一定である．なお，(9)式は，$T_p, T_f, \bar{t}$ に値を直接代入することで発生確率が求まる．

2.3 確率予測の成績に関する指標

Okada *et al*. (2012) などと同様に，確率予測の成績を表す指標として平均対数尤度とブライアスコアを採用した．

2.3.1 平均対数尤度 (*MLL*)

岡田 (2009) や Okada *et al*. (2012) と同じ方法を用いた．予測確率と予測期間内の当該地震の発現・非発現との適合度を，対数尤度を用いて評価する方法である．予測数を N 個，j 番目の予測確率を P_j とすると，対数尤度 LL は,

$$\begin{aligned} LL &= \sum_{j=1}^{N} c_j \ln(P_j)+\sum_{j=1}^{N}(1-c_j)\ln(1-P_j) \\ &= \sum_{j=1}^{N}\ln(1-P_j)+\sum_{j=1}^{N} c_j\left[\ln\left(\frac{P_j}{1-P_j}\right)\right] \end{aligned} \tag{10}$$

で与えられる．j 番目の観測結果 c_j は,

$$c_j=\begin{cases} 1 & \text{（予測期間内に当該地震が発現したとき）} \\ 0 & \text{（予測期間内に当該地震が発現しなかったとき）} \end{cases} \tag{11}$$

である．

なお，P_j の値が 0 で，c_j が 1 のときは，対数尤度 LL は求まらない．Kagan and Jackson (1995) は，尤度検定（**2.4.2** 参照）でそのような事象が 1 個でもあれば，その予測モデルを直ちに棄却するとしている．また，岡田 (2009) は，そのような場合はモデル不適当（完全不適中）として，その系列を評価対象から除いて評価することを提案している．

平均対数尤度 MLL は,

$$MLL=\frac{LL}{N} \tag{12}$$

で，値は常に負となる．対数尤度 LL は予測数 N が増大すれば単調に減少するが，MLL は比較的安定しており，0 に近いほど予測と観測結果は整合し，良い成績と言える．

2.3.2 ブライアスコア (*BS*)

Brier (1950) によるもので，予測確率と観測結果との整合性を，確率の誤差を 2 乗平均した誤差分散値で評価する方法である．気象分野では広く利用されており，立平 (1999) や Toth *et al*. (2003) によれば，確率値の信頼度と，予測確率で現象発生の有無を分離する分離度の 2 つの評価を兼ね備えている．

予測数を N 個，j 番目の予測確率を P_j とすると，ブライアスコア BS は,

$$BS=\frac{1}{N}\sum_{j=1}^{N}(P_j-c_j)^2 \tag{13}$$

である．j 番目の観測結果 c_j は，(11)式と同じである．BS は 0 から 1 の間の値で，0 に近いほど予測と観測結果は整合し，良い成績と言える．

なお，発生確率が 0 または 1 に近い値であればあるほど，BS の値は小さくなり，変化量はわずかになる特徴がある［山下・他 (2003)，林 (2016)］．

2.4 パラメトリックな検定手法を用いた統計的検定

Okada *et al*. (2012) などと同様に，Kagan and Jackson (1995) が採用した個数検定，尤度検定，尤度比検定を用いた．また，ブライアスコアを用いた検定は，尤度検定の対数尤度をブライアスコア[Brier (1950)] に置き換えたもので，完全不適中の事例が含まれていても不都合は生じない．

2.4.1 個数検定 (N-test)

予測期間内の当該地震の発現数について，予測と観測結果との整合性を統計検定する方法である．予測確率を使い，予測期間内に期待される発現数の確率分布を数値計算する．この理論的な確率分布と予測期間内の当該地震の発現数とを比較する．予測期間内の発現数が両側検定で理論分布の信頼区間（95% または 99%）から外れる予測は棄却できる．発現数が信頼区間の上限値を上回るときは過小予測で，下限値を下回るときは過大予測と見なすことができる．

2.4.2 尤度検定 (L-test)

(10)式の対数尤度 LL を用いて，予測と観測結果とのズレを統計検定する方法である．各予測について，当該地震の発生 ($c_j=1$) と非発生 ($c_j=0$) があり，その確率は P_j と $1-P_j$ で，それぞれの対数尤度は $\log(P_j)$ と $\log(1-P_j)$ である．予測を N 回実施すると，発生・非発生の組み合わせは 2^N 通りになる．予測が互いに独立であるとすれば，各組み合わせの対数尤度と発生確率は，個々の予測の対数尤度の和と発生確率の積である．2^N 個の組み合わせを対数尤度の小さい順に並び替えて，発生確率を順次積算すると，対数尤度 LL の累積確率分布（理論分布）が得られる．

一方で，予測確率と観測結果から対数尤度の合算値 LL_{obs} を (10)式で計算して，理論分布と比較する．理論分布の下側にのみ棄却域をとる片側検定で，LL_{obs} が信頼区間の下限値（理論分布の 5% 点または 1% 点）を下回る場合は棄却できる．

2.4.3 ブライアスコアを用いた検定 (BS-test)

(13)式のブライアスコア BS を用いて，予測と観測結果とのズレを統計検定する手法で，尤度検定の対数尤度 LL をブライアスコア BS に置き換えたものである．予測確率を使い，BS の累積確率分布を，L-test と同様の方法で数値計算する．そして，観測結果からブライアスコア BS_{obs} を(13)式で計算して，理論分布と比較する．理論分布の上側にのみ棄却域をとる片側検定で，BS_{obs} が信頼区間の上限値（理論分布の 95% 点または 99% 点）を上回る場合は棄却できる．

2.4.4 尤度比検定 (R-test)

同じ対象を 2 つの異なる仮説（手法）に基づいて予測したとき，両者の成績の優劣を対数尤度の差（尤度比の対数）を用いて統計検定する方法である．j 番目の予測について，仮説 H0 で予測した確率を P_{0j}，仮説 H1 で予測した確率を P_{1j} とする．予測確率 P_{0j}，P_{1j} を使い，仮説 H0 が正しいと仮定した場合の対数尤度の差 $R\,(=LL_1-LL_0)$ の理論分布を L-test と同じ方法で数値計算する．同様に，仮説 H1 が正しいと仮定した場合の $R\,(=LL_1-LL_0)$ の理論分布を数値計算する．さらに，観測結果から(10)式で求まる仮説 H1 の対数尤度 LL_1 と仮説 H0 の対数尤度 LL_0 の差 $R_{obs}(=LL_1-LL_0)$ を計算して，2 つの理論分布と比較する．R_{obs} の値が仮説 H0 の理論分布の上側に棄却域をとる片側検定で，信頼区間の上限値（理論分布の 95% 点または 99% 点）を上回るときは，仮説 H0 は棄却できる．また，R_{obs} の値が仮説 H1 の理論分布の下側に棄却域をとる片側検定で，信頼区間の下限値（理論分布の 5% 点または 1% 点）を下回るときは，仮説 H1 は棄却できる．R_{obs} の値が仮説 H0 の信頼区間（95% または 99%）にあり，かつ仮説 H1 の棄却域にある（理論分布の 5% 点または 1% 点を下回る）とき，仮説 H0 は仮説 H1 よりも有意に優れていると判定できる．同様に，仮説 H0 が棄却でき，仮説 H1 は棄却とはならないときは，仮説 H1 が仮説 H0 よりも有意に優れていると判定できる．なお，この検定では，仮説 H0 と仮説 H1 がともに観測結果と整合せずに棄却できる場合や，両方とも観測結果と整合する場合があり，予測モデルの優劣が確定できないこともある．

§3. データ

気象研究所と東北大学が共同で，2006 年 7 月から 2010 年にかけて 4 回実施した岡田らによる確率予測実験で採用された 4 つの相似地震カタログを使用した．Okada *et al.* (2012) によれば，東北大学が作成した相似地震カタログから，次の 3 条件を満たす系列が選ばれている．

① 1993 年以降予測日前日までに地震数が 5 個以上

②平均マグニチュードが 2.75 以上

③ 1 系列に含まれる 1994 年三陸はるか沖地震または 2003 年十勝沖地震の余震数が 3 分の 1 未満

なお，③の余震は，本震後 1 か月以内の地震活動図から，地震活動の高い地域を目視で決めて余震域とし，余震域内の地震数の時間変化から余震活動の期間を決めて，選び出している．

本調査では，4 つのカタログから，それぞれ系列毎に予測日に近い順からイベントを 2 個，3 個，4 個，5 個または 5 個以上で切り出し，実験毎と 4 実験を一纏めにした地震数別のデータセット（全 25 ケース）を作成した．以降,「2006 年」とは，2006 年 7 月から 2007 年 6 月を対象にした確率予測実験に関する記述で,「2008 年」,「2009 年」及び「2010 年」も，それぞれの年を対象にした確率予測実験に関する記述である．「2006–10 年」とは，4 実験を一纏めにしたデータセットに関する記述である．

2006–10 年のデータセットで予測した数 N_s は 524 個で，そのうち，予測期間内の当該地震の発現数 N_q は 248 個である．なお，作成したデータセットの中に，完全不適中となる系列が見つかり，その系列を使った予測はすべてのデータセットから除外して作業を進めた．よって，岡田らによる確率予測実験で扱われた予測数の合計よりも 4 個少ない．解析に用いた地震ののべ数は 3871 個である (Table 1)．後の実験の方が，予測数や解析に用いた地震数が多いのは，主に相似地震カタログの期間が延びて，①の条件を満たす系列が増えたためである．また，予測数 524 個の地震数の内訳は，半数以上（283 予測）が 6 個以下である (Table 2)．複数の実験で同じ系列の予測が行われており，同じ系列のものはまとめて 1 個として数えると 166 系列となる．それらの系列は北海道から東北・関東地方の太平洋沖や沿岸に分布している (Fig. 1)．この 166 系列の 1993 年 1 月から 2009 年 12 月までに発生した地震数は 1258 個（発生間隔数にすると 1092 個）で，発生間隔（日）の対数 x_i の頻度分布は Fig. 2 のとおりである．発生間隔の対数の平均値 $\bar{x}$ は 6.430 で，標準偏差 s は 0.676 である．これらの値から求まる発生間隔の期待値 $\overline{T}\,[=\exp(\bar{x}+s^2/2)]$ は約 779 日である．

§4. 結果

4.1 予測成績の比較

各データセット（全 25 ケース）に対し，3 通りの方法（LN-Bayes，LN-SST 及び Exp-pin）で予測した結果について，(12)式の平均対数尤度 (MLL) と，(13)式のブライアスコア (BS) で予測成績を集計したものが Table 3 である．表中の網掛けした数字は，後述の統計検定 (L-

Table 1. Basic information on the catalog used for investigation.

Forecast period		2006.7–07.6	2008.1–12	2009.1–12	2010.1–12	Total
Forecast sequences (N_s)		92	126	144	162	524
Number of sequences that had earthquakes within the forecast period (N_q)		50	56	70	72	248
Analysis data for forecast	Data period	1993.1–06.6	1993.1–07.12	1993.1–08.12	1993.1–09.12	Total
	Number of earthquakes	666	905	1068	1232	3871

Table 2. Number of earthquakes in one forecast for 524 forecast sequences of the small repeating earthquakes catalog.

Number of earthquakes	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Forecast sequences (N_s)	181	102	77	39	26	21	27	12	8	5	4	8	4	4	4	2	524

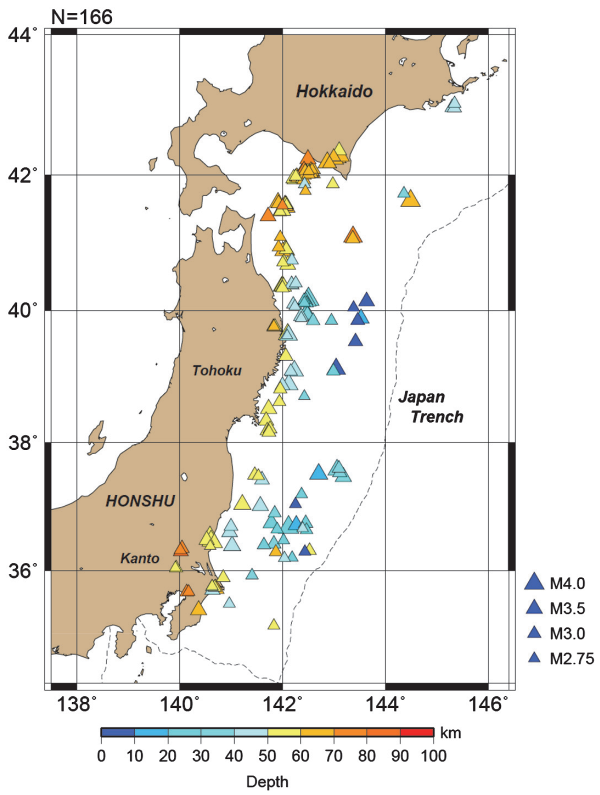

Fig. 1. Distribution of small repeating earthquake clusters of 166 sequences, used in four forecast experiments. The color of the symbol indicates focal depth, and the size of the symbol indicates magnitude. The average magnitude of all the sequences is 2.75 or more. Note that the offshore sequences have large depth uncertainties.

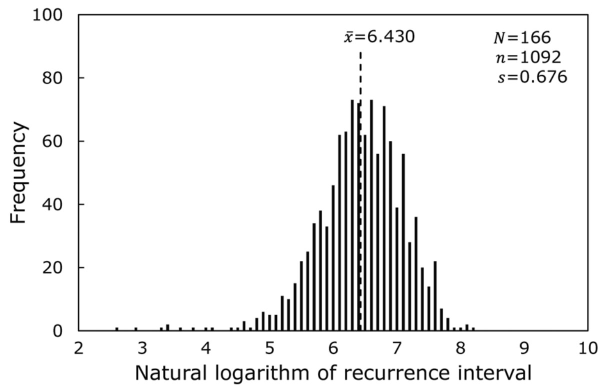

Fig. 2. Frequency distribution of recurrence intervals in the 166 sequences. N denotes the number of all forecast sequences, n denotes the number of recurrence intervals, $\bar{x}$ denotes average of $x_i = \log(t_i)$, and s denotes standard deviation of x_i.

test，BS-test) で棄却できることを示す．なお，LN-SST の場合，地震の発生間隔データが 1 個のときは，原理的に予測確率は計算できないので示していない．

予測成績を比べると，Exp-pin は，すべてのケースで両指標が LN-Bayes より悪く，LN-SST との比較では，2009 年の間隔データが 2 個と 3 個の場合を除き，Exp-pin が悪い．さらに，データが 1 個と 2 個のときは，ほぼすべてのケースで予測確率 0.5 のときの成績 ($MLL = -0.693$, $BS = 0.25$) と同程度かそれ以下であり，非常に悪い結果である．

LN-SST と LN-Bayes の比較では，2008 年は LN-SST が優れており，他の年は同程度か LN-Bayes が優れている．LN-SST は 2009 年の成績が一段と悪く，各年の実験を総合した 2006–10 年の予測成績でも LN-Bayes よりも劣り，成績差は間隔データ数が少ない時に大きい．

間隔データ数を変えた際の *MLL* 及び *BS* の変化例として，2006–10 年の予測成績を Fig. 3 と Fig. 4 に示す．Exp-pin の間隔データが 1 個の場合を除き，両指標ともかなり良く似た傾向であり，データ数の増加に伴う成績向上が認められ，変化は 3 個以下で著しい．2006 年から

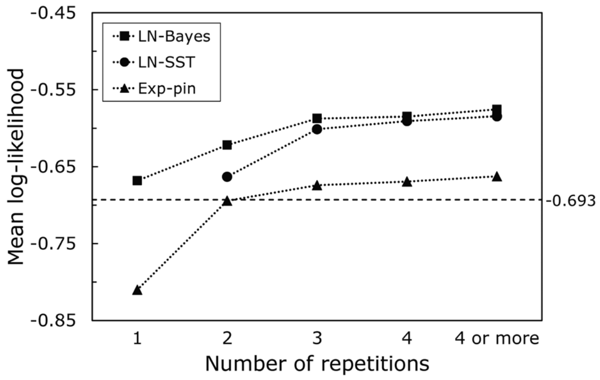

Fig. 3.　Change of the mean log-likelihood for various numbers of repetitions in four forecast experiments. The horizontal dashed line denotes the score of −0.693, corresponding to the forecast probability of 0.5 for all cases. The LN-Bayes model is the log-normal distribution model based on the Bayesian theorem with prior distribution of an inverse gamma. The LN-SST model is the log-normal distribution model based on the small sample theory. The Exp-pin model is the exponential distribution model based on the Poisson process, and the probability of an event is independent of the elapsed time since the last event.

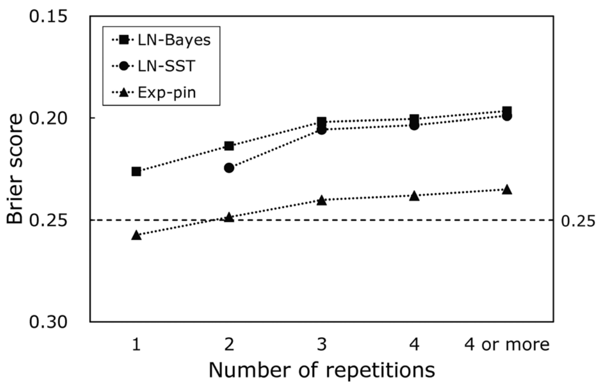

Fig. 4.　Result of the Brier score for each number of repetitions in four forecast experiments. The horizontal dashed line indicates the score of 0.25, corresponding to the forecast probability of 0.5 for all cases. The LN-Bayes, LN-SST, and Exp-pin models are the same as in Fig. 3.

2010 年の各実験の予測成績には，間隔データ数が増えたのにも関わらず，予測成績が悪化した事例がいくつかみられる (Table 3)．しかし，全体的には Fig. 3 及び Fig. 4 と同様にデータ数の増加に伴う成績向上が認められる．

成績の良い LN-Bayes に着目して，予測確率 10% 区切りで集計した予測数，予測確率から推定した発生回数及び当該地震の発現数の関係を Fig. 5 の a，b 及び c に示す．これら 3 図を予測確率 0〜10% で比較すると，間隔データが 1 個のとき (Fig. 5a) は，予測確率から推定した発生回数（灰色棒，3.1 回）に比べて，実際の発現数（黒色棒，15 回）が著しく大きい．間隔データが 3 個 (Fig. 5b) 及び 4 個以上 (Fig. 5c) では，推定した発生回数と実際の発現数の差は小さくなっている．他の予測確率でも，全体として同様な傾向が認められる．すなわち，使用する間隔データ数が増えると，予測確率の「信頼度」は向上する．一方，予測確率 90〜100% に着目すると，これら 3 図とも推定した発生回数と実際の発現数の差は小さく，「信頼度」では大差がない．しかし，間隔データ数の増加に従い，予測回数（白抜き棒）は明瞭に増加し，中間的な予測確率 30〜70% では予測回数の減る傾向が認められる．すなわち，現象の発現・非発現を予測確率で分離する「分離度」がデータ数の増加に伴い，向上している．使用する間隔データ数が増えるにつれて，「信頼度」と「分離度」の両方が向上したことで，全体の予測成績が向上したと言える．

Table 3 をみると，2009 年の成績はすべてのケースで悪く，成績が互いに連動している可能性がある．4 実験について，間隔データが 4 個のときの予測成績の変化を Fig. 6 に示すが，似た傾向を示しており，特に LN-Bayes と LN-SST はよく似た変動である．これは，偶然的な要因で説明するのは困難であり，相似地震活動の変動を反映したためと考えられる．一方で，MLL の標準偏差（Table 3 に示した 4 実験から求めた不偏分散の平方根）は，LN-Bayes が 0.066（BS では 0.027），LN-SST が 0.091 (0.033) であるのに対し，Exp-pin は 0.027 (0.013) と最も小さい．Exp-pin の成績が安定するのは，発生確率を計算する (9) 式で，変数が比較的安定している発生間隔の平均値のみで決まるためである．LN-Bayes と LN-SST は，ともに母数 2 個の対数正規分布を採用しているが，LN-Bayes は，逆ガンマ事前分布で σ^2 の変動性を制約していることから，LN-SST より予測成績が安定していると考えられる．

4.2　統計検定による検証

2.4 節で統計的検定（個数検定，尤度検定，尤度比検定）の方法について紹介したが，ここでは 3 種類の統計モデル（LN-Bayes，LN-SST 及び Exp-pin）の予測について，これらの方法で観測結果との整合性を検証する．以下では，記述を簡潔にするために，95% 信頼区間を採用し，主な結果について紹介する．

4.2.1　N-test の結果

予測期間内の発現数の期待値 $E(N_q)$ と，当該地震の発現数の理論分布の累積確率値 *CDF* は Table 4 の通りである．両側検定で，理論分布の累積確率値 *CDF* が 97.5

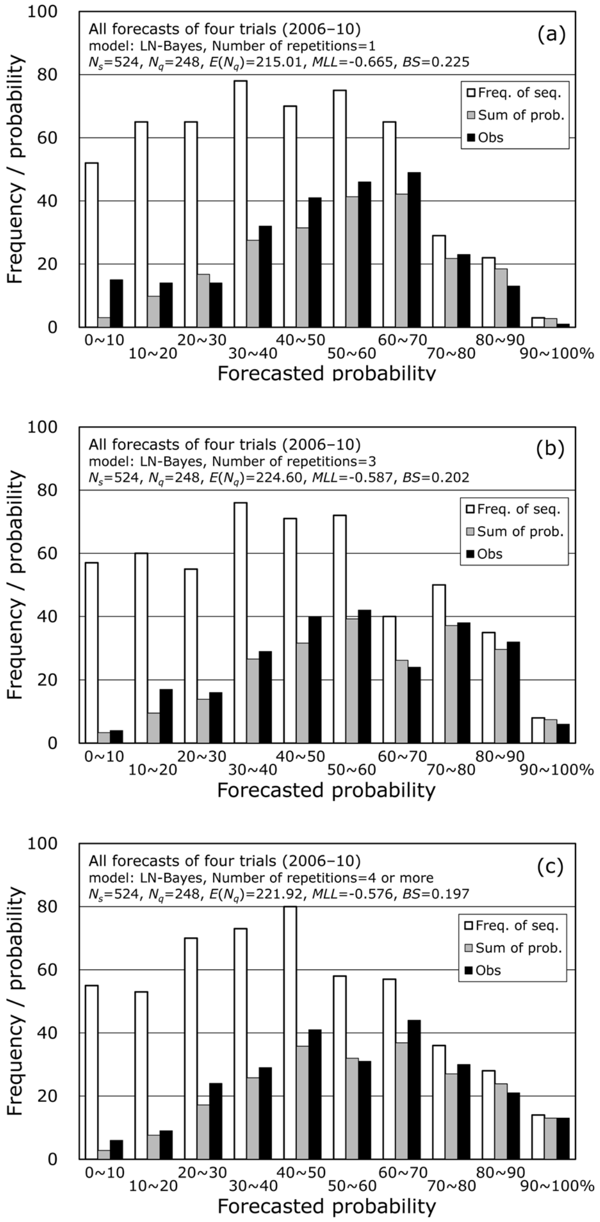

Fig. 5. Frequencies of forecast sequences (left white bars), the expected number of sequences with events that will occur within the forecast period (the sum of forecast probabilities by LN-Bayes, central gray bars), and the actual number of sequences of observed earthquakes within the forecast period (right black bars) for every 10% range of probabilities for four trials. There are (a) 1, (b) 3, and (c) 4 or more repetitions. N_s is the number of all forecast sequences, N_q is the number of sequences that had earthquakes within the forecast period, $E(N_q)$ is the sum of all the probabilities (expected values of N_q) and is the average frequency of the predicted events, MLL is the mean log-likelihood, and BS is the Brier score.

% より大きいときは過小な確率予測で，2.5% より小さいときは過大予測である．Table 4 では，どちらの場合も CDF の値に網掛けをしてある．

N-test の一例として，LN-Bayes による 2006–10 年の予測成績について，間隔データが 3 個のときの発現数の理論分布と，予測期間内の当該地震の発現数 N_q の関係を Fig. 7 に示す．この場合，観測結果 N_q＝248 に対応する理論分布の累積確率値 CDF は 99.2% と大きく，過小予測で棄却できることを表している．

Table 4 で示すように，LN-Bayes による予測では，2006–10 年の予測成績は，すべてのケースが過小予測で棄却できる．各年の実験結果は，2006 年はすべてのケースが棄却でき，2008 年，2009 年及び 2010 年は，逆にすべてのケースで観測結果と整合する．各年の 4 実験のうち 3 実験で観測結果と整合しているにも関わらず，2006–10 年の予測成績が棄却できるのは，予測数 N_s が各年の実験の予測数よりも 3 倍以上増えており，有意性の判定が厳しくなったためと考えられる．

LN-SST 及び Exp-pin による予測では，2008 年の予測成績を除き，概して過小予測で，ほとんどのケースが棄却できる．N-test の結果では，Exp-pin と LN-Bayes または LN-SST との差はあまり明瞭でない．これは，Exp-pin による予測は，過去の平均発生間隔が直に反映されており，予測の精度・成績が劣っていても (Table 3)，全体の予測発生数が観測の発現数に比較的近いことが多いためである．

4.2.2　L-test の結果

検定の一例として，LN-Bayes による 2006–10 年の予測成績について，間隔データが 3 個のときの対数尤度の理論分布と，観測結果から求まる対数尤度 LL_{obs} との関係を Fig. 8 に示す．この場合，観測結果から求まる対数尤度 LL_{obs}＝−307.8 に対応する理論分布の累積確率値 CDF は 2.2% と小さく，予測は観測結果と整合せず，棄却できる．

Table 3 の網掛けで示すように，LN-Bayes による予測では，2006–10 年の予測成績は，ほとんどのケースが棄却できる．各年の実験結果は，2006 年と 2008 年は，間隔データが 2 個以上のケースで観測結果と整合するのに対し，2009 年と 2010 年は，ほとんどのケースが棄却でき，実験時期で違いがみられる．

LN-SST による予測では，2006–10 年の予測成績は，すべてのケースが棄却できる．各年の実験結果は，2006 年と 2008 年は，間隔データが 3 個以上のケースで観測結果と整合し，2009 年と 2010 年は，すべてのケースが棄却できる．

Exp-pin による予測では，2006–10 年の予測成績は，ほ

Table 3. Experimental scores of the mean log-likelihood, MLL, and Brier, BS, for various data sets derived from the small repeating earthquakes.

Period	2006.7–07.6		2008.1–12		2009.1–12		2010.1–12		2006–10 (Four trials)	
Number of repetitions	*MLL*	*BS*	*MLL*	*BS*	*MLL*	*BS*	*MLL*	*BS*	*MLL*	*BS*
LN-Bayes										
1	-0.650	0.220	-0.669	0.227	-0.693	0.242	-0.656	0.216	-0.665	0.226
2	-0.574	0.197	-0.612	0.204	-0.674	0.237	-0.610	0.210	-0.622	0.214
3	-0.532	0.182	-0.563	0.188	-0.658	0.232	-0.575	0.197	-0.587	0.202
4	-0.510	0.173	-0.545	0.182	-0.663	0.234	-0.589	0.201	-0.585	0.201
4 or more	-0.509	0.172	-0.531	0.177	-0.650	0.229	-0.583	0.197	-0.576	0.197
LN-SST										
2	-0.568	0.194	-0.557	0.185	-0.825	0.283	-0.656	0.221	-0.663	0.224
3	-0.544	0.187	-0.522	0.177	-0.724	0.251	-0.587	0.199	-0.601	0.206
4	-0.500	0.174	-0.512	0.174	-0.695	0.243	-0.611	0.208	-0.591	0.203
4 or more	-0.509	0.176	-0.497	0.167	-0.681	0.236	-0.610	0.204	-0.585	0.199
Exp-pin										
1	-0.884	0.272	-0.808	0.253	-0.907	0.272	-0.684	0.239	-0.766	0.257
2	-0.705	0.255	-0.675	0.240	-0.713	0.257	-0.686	0.244	-0.694	0.248
3	-0.687	0.246	-0.656	0.232	-0.699	0.252	-0.659	0.233	-0.674	0.240
4	-0.675	0.240	-0.645	0.227	-0.705	0.255	-0.653	0.230	-0.669	0.238
4 or more	-0.665	0.236	-0.639	0.224	-0.698	0.252	-0.648	0.228	-0.663	0.235

Remarks: The shaded numbers indicate that the forecast probabilities were rejected by L-test or BS-test.

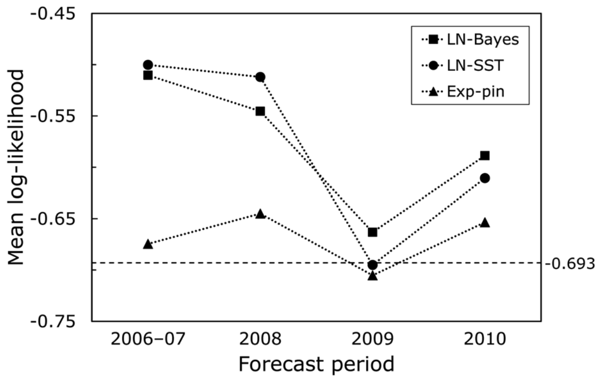

Fig. 6. Change of the mean log-likelihood for four forecast experiments with four repetitions.

とんどのケースが棄却できる．各年の結果では，間隔データが 2 個以下では棄却できるケースが多く，3 個以上のときは観測結果と整合するケースが多い．

4.2.3 BS-test の結果

一例として，LN-Bayes による 2006–10 年の予測成績について，間隔データが 3 個のときのブライアスコアの理論分布と，観測結果から求まるブライアスコア BS_{obs} の関係を Fig. 9 に示す．この場合，観測結果から求まるブライアスコア BS_{obs}=0.202 に対応する理論分布の累積確率値 *CDF* は 97.9% と大きく，予測は観測結果と整合せず，棄却できる．

Table 3 の網掛けで示すように，LN-Bayes による予測では，2006–10 年の予測成績は，間隔データが 4 個以上の場合を除き，棄却できる．各年の実験結果は，2006 年と 2008 年は観測結果と整合するケースが多く，2009 年と 2010 年は棄却できるケースが多い．

LN-SST による予測では，2006–10 年の予測成績は，すべてのケースが棄却できる．各年の実験結果は，2006 年と 2008 年は，間隔データが 3 個以上のケースで観測結果と整合し，2009 年と 2010 年はすべてのケースが棄却できる．

Exp-pin による予測では，2006–10 年の予測成績は，間隔データが 4 個以上の場合以外は棄却できる．各年の実験結果は，2006 年，2008 年及び 2010 年は，間隔データが 2 個以下のケースではほとんど棄却でき，3 個以上はすべてのケースで観測結果と整合する．2009 年はすべてのケースが棄却できる．

4.2.4 R-test の結果

データは同じで統計モデルが異なる場合と，同じ統計モデルでデータ数が異なる場合の比較について，それぞれの結果を紹介する．同じデータで統計モデルの異なる場合の検定の一例として，2006–10 年の間隔データが 3 個のときの Exp-pin と LN-Bayes の比較結果を Fig. 10 に示す．仮説 H0 (Exp-pin) の予測が正しいと仮定したときと，仮説 H1 (LN-Bayes) の予測が正しいと仮定したときの対数尤度の差 R (=LL_1-LL_0) の理論分布 (H0: Exp-pin，H1: LN-Bayes) と，観測結果から求まる対数尤

Table 4. Results of the N-test on probabilities for the small repeating earthquake event within 12 months, by the repetition numbers.

Period	2006.7–07.6		2008.1–12		2009.1–12		2010.1–12		2006–10 (Four trials)	
Number of repetitions	$E(N_q)$	CDF	$E(N_q)$	CDF	$E(N_q)$	CDF	$E(N_q)$	CDF	$E(N_q)$	CDF
LN-Bayes		%		%		%		%		%
1	38.4	99.8	47.7	96.1	61.3	95.9	67.6	81.6	214.9	100.0
2	40.2	99.3	53.4	73.4	62.1	94.7	68.7	76.2	224.3	99.2
3	40.2	99.3	54.3	67.2	63.0	92.8	67.2	84.2	224.6	99.2
4	40.8	99.0	54.3	67.3	60.8	97.0	65.7	89.9	221.7	99.7
4 or more	41.2	98.8	55.6	57.5	60.9	96.9	64.3	93.7	221.9	99.6
LN-SST		%		%		%		%		%
2	34.7	100.0	49.2	95.4	55.5	99.9	60.6	99.0	200.0	100.0
3	36.4	100.0	51.6	86.0	59.4	98.8	62.5	97.4	209.9	100.0
4	38.7	99.9	52.4	81.6	56.9	99.6	62.6	97.2	210.5	100.0
4 or more	39.3	99.8	53.3	75.2	57.7	99.4	61.7	97.9	212.0	100.0
Exp-pin		%		%		%		%		%
1	42.7	96.0	53.4	72.9	63.0	91.1	70.9	60.8	229.9	96.0
2	40.7	98.4	52.3	78.5	58.3	98.4	66.0	86.1	217.4	99.8
3	39.8	99.0	50.7	86.4	57.1	99.0	62.3	95.4	209.9	100.0
4	39.3	99.2	49.4	90.9	55.0	99.7	60.6	97.6	204.4	100.0
4 or more	39.3	99.2	49.5	90.4	55.0	99.7	59.9	98.2	203.7	100.0

Remarks: LN-Bayes LN-SST, and Exp-pin are the Bayesian statistics lognormal distribution model of the recurrence intervals with an inverse gamma prior distribution, the log-normal distribution model on small sample theory, and the exponential distribution model based on the Poisson process, respectively. $E(N)$ and CDF are the expectation number of sequences with qualifying events and the probability of the theoretical cumulative distribution function for the number of sequences with event, calculated from the forecast probabilities, at the observed number of sequences with event, respectively. The shaded numbers indicate that the forecast probabilities were rejected in N-test.

度の差 $R_{obs}(=LL_1-LL_0=45.4)$ がプロットされている．片側検定であり，仮説 H0 の棄却域は対数尤度の差 R_{obs} について仮説 H0 の理論分布の累積確率値 CDF(H0) が 95% より大きい範囲であり，仮説 H1 の棄却域は，仮説 H1 の理論分布の累積確率値 CDF(H1) が 5% より小さい範囲である．この図の場合，仮説 H1 は観測結果と整合し，仮説 H0 は観測結果と整合せず，棄却できるので，LN-Bayes が Exp-pin より有意に優れていると判定できる．

① 異なる統計モデルで比較した場合

Exp-pin と LN-Bayes の比較では，2006–10 年の予測成績は，Exp-pin がすべてのケースで棄却でき，かつ LN-Bayes はすべてのケースで観測結果と整合する．よって，Exp-pin より LN-Bayes が優れていると判定できる．各年の実験結果でも，ほとんどのケースで LN-Bayes が Exp-pin より優れている．しかし，2009 年の間隔データが 3 個と 4 個のときは，LN-Bayes も観測結果と整合せず，優劣が確定できない場合もある．

Exp-pin と LN-SST の比較では，2006–10 年の予測成績は，両モデルともすべてのケースが棄却でき，優劣は確定できない．各年の実験結果は，2006 年と 2008 年はすべてのケースで Exp-pin より LN-SST は優れているが，2009 年と 2010 年は両モデルともすべてのケースが棄却でき，優劣は確定できない．

LN-SST と LN-Bayes の比較では，年によって優劣が異なる．2006–10 年の予測成績は，間隔データが 4 個以上のケースで LN-Bayes は優れているが，他のケースでは，両モデルとも棄却でき，優劣は確定できない．各年の実験結果は，2006 年は，間隔データが 3 個のときを除き，優劣は確定できない．2008 年はすべてのケースで LN-SST が優れており，2009 年と 2010 年では逆にすべてのケースで LN-Bayes が優れている．

② 同じ統計モデルを異なるデータ数で比較した場合

LN-Bayes による予測では，2006–10 年の予測成績は，間隔データが 1 個と 2 個，及び 2 個と 3 個のときは，データ数の多い方が優れており，3 個と 4 個では，両予測とも観測結果と整合せず，優劣は確定できない．各年の実験結果は，ほとんどのケースでデータ数の多い方が優れている．しかし，2009 年の 3 個と 4 個では，両予測とも観測結果と整合し，優劣は確定できない．また，

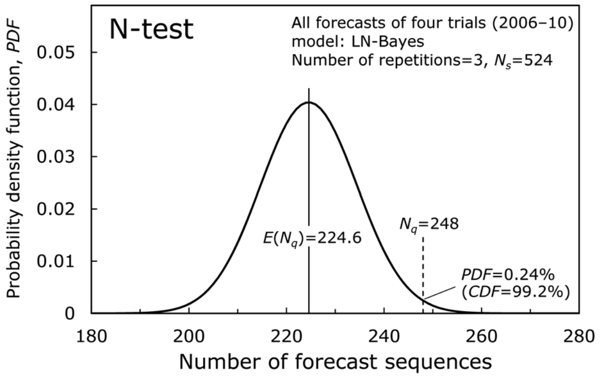

Fig. 7. Result of N-test for 524 forecasts of four trials from 2006 to 2010 using the LN-Bayes model. The solid curved line indicates the theoretical distribution of the frequency of events calculated from forecast probabilities, the vertical dashed line indicates the observed frequency of events, and the solid vertical line indicates the expected value calculated from the forecast probability. N_s, N_q, $E(N_q)$ are the same as in Fig. 5. CDF is the cumulative distribution function at the score of observation.

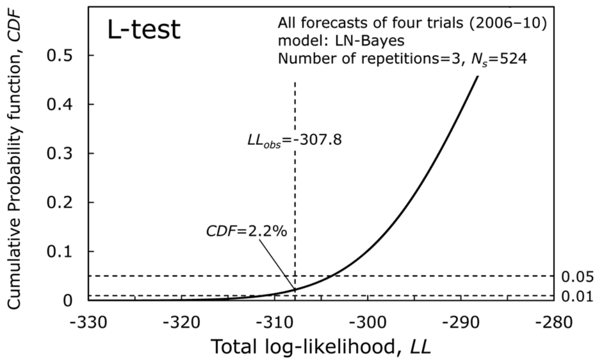

Fig. 8. Result of L-test for 524 forecasts of four trials from 2006 to 2010 using the LN-Bayes model. The solid curved line indicates the theoretical distribution of total log-likelihood (LL), and the vertical dashed line indicates its observed score. The dashed horizontal line indicates 0.05 and 0.01 of probability. N_s is the same as in Fig. 5.

2010 年の 3 個と 4 個では，データ数の少ない方が優れている．

LN-SST による予測では，2006–10 年の予測成績は，LN-Bayes と同じく，間隔データが 2 個と 3 個では，データ数の多い方が優れており，3 個と 4 個では，優劣は確定できない．各年の実験結果は，ほとんどのケースでデータ数の多い方が優れている．しかし，2010 年の 3 個と 4 個のときは，データ数の少ない方が優れている．

Exp-pin による予測では，2006–10 年の予測成績は，

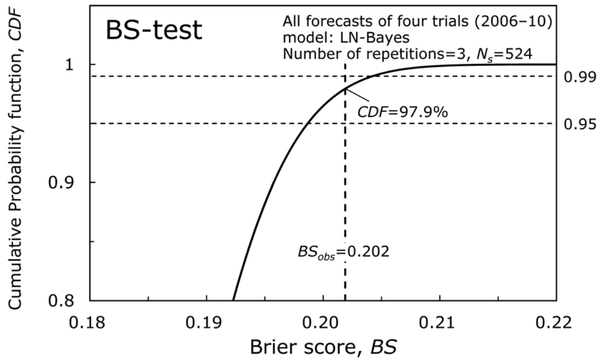

Fig. 9. Result of BS-test for 524 forecasts of four trials from 2006 to 2010 using the LN-Bayes model. The solid curved line indicates the theoretical distribution of the Brier score (BS), and the vertical dashed line indicates its observed score. N_s is the same as in Fig. 5.

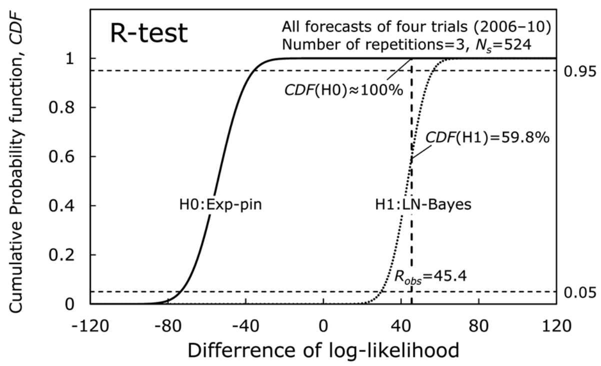

Fig. 10. Result of R-test for the Exp-pin model and the LN-Bayes model. The solid curved line (H0) indicates the theoretical distribution of the log-likelihood difference when the Exp-pin model is correct, and the dashed curved line (H1) indicates it when the LN-Bayes model is correct. The vertical dashed line indicates the observed log likelihood difference. N_s is the same as in Fig. 5.

データ数の多い方がすべて優れている．各年の実験結果は，ほとんどのケースでデータ数の多い方が優れており，優劣の確定できないケースはあるが，データ数の少ない方が優れているケースはない．

2010 年の 3 個と 4 個の比較で，LN-Bayes と LN-SST の結果は，ともにデータ数の少ない方が優れている．原因としては，地震活動の異常かもしれないが，データ数の少ない方が優れるケースは 40 回実施した R-test のうちの 2 回のみであり，有意水準 5% と同程度の割合であることから，偶然的な要因による揺らぎの結果と見なすこともできる．

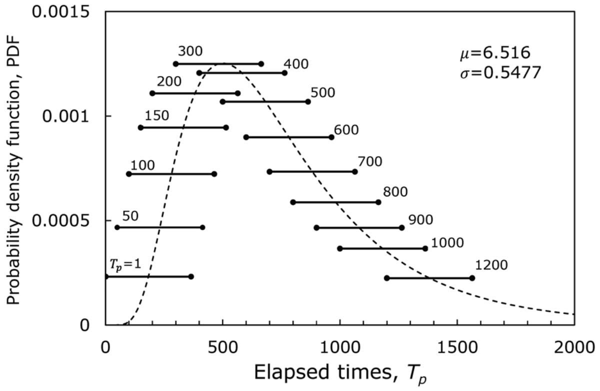

Fig. 11.　Forecast intervals to be used for a random number of experiments. The horizontal line indicates the range of the forecast period, which is fixed to 365 days. T_p is the elapsed time from the last event to the forecast period. The dashed curved line exhibits a log-normal distribution, $LN(\mu, \sigma^2)$, used for simulation by the LN-SST model.

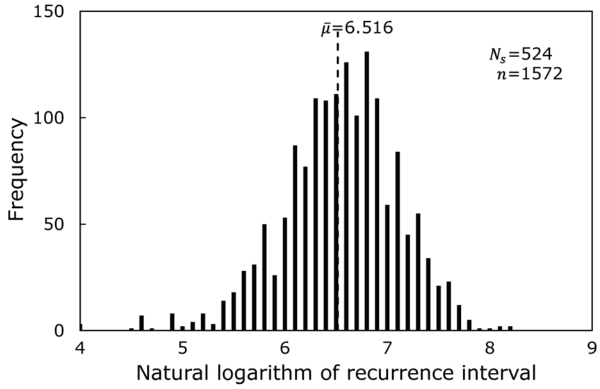

Fig. 12.　Distribution of recurrence intervals using the LN-Bayes model. $\bar{\mu}$ is the average logarithm of the recurrence intervals. N_s is the number of all forecast sequences. n is the number of recurrence intervals.

§5.　考　　察

LN-SST と LN-Bayes の 2 つ統計モデルについて，全 524 予測の相似地震を使い，比較用として指数分布モデルを用いて，繰り返し回数が非常に少ないときの予測成績や予測と観測結果の整合性について調査した．現実の地震活動では，様々なスケールで時間・空間的に変動し，群発性などがしばしば見られるので，対数正規分布が正しいと仮定しても，更新過程（発生間隔の相互独立性）など，予測モデルの前提条件が十分に成り立っているかどうかは疑問である．予測モデルを評価するためには，実際の地震に対する予測実験とともに，統計モデルの条件をほぼ完全に満たす疑似データを作成し，発生確率予測をシミュレーションすることが望ましい．そこで，シミュレーションの結果と相似地震の予測実験結果を比較・検討し，予測モデルの信頼性を確認し，相似地震のデータでは判明できなかった点について調べた．

5.1　乱数実験用の疑似データ

シミュレーションに使用する疑似地震カタログと，予測後に発現する疑似発現データは，相似地震の統計から導いた対数正規分布のパラメータを使い，一様乱数から作成した．発生間隔のデータ数は，LN-Bayes が 1, 2, 3, 4, 5, 7, 10, 15, 20, 30, 60 及び 100 個の計 12 ケースで，LN-SST が，1 個の場合を除いた 11 ケースである．予測日（直近地震からの経過時間）T_p は，地震 1 日後から 1200 日後まで計 25 パターン設定した (Fig. 11)．統計モデル，発生間隔のデータ数及び予測日の組み合わせ毎に 50,000 系列の疑似データセットを用意した．

LN-Bayes で使用する事前分布は，μ について一様で，σ^2 については (6) 式の逆ガンマ分布である．乱数実験に用いる発生間隔の対数 x_B は，

$$x_B=\mu+\sigma_B F^{-1}(P) \tag{14}$$

で作成した．ここで，添え字 B は系列の順番を表し，B＝1，2，…，50,000 であり，x_B は系列 B の疑似データ群である．μ は全 524 予測の相似地震から各系列直近の発生間隔 3 個の平均の平均値 $\mu=6.516$ を用いた (Fig. 12)．$F^{-1}(P)$ は標準正規分布の分布関数 $F(x)$ の逆関数で，

$$P=F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt \tag{15}$$

である．P は区間 (0, 1) の一様乱数である．σ_B は系列毎に異なり，一様乱数 P_2 を用いて，

$$\sigma_B=\sqrt{\frac{\zeta^2}{G^{-1}(P_2|\phi,\zeta)}} \tag{16}$$

で与える．$G^{-1}(x)$ はガンマ分布の累積分布関数 $G(x)$ の逆関数である．ϕ, ζ は形状パラメータとスケールパラメータで，2009 年と 2010 年の予測実験で使われた値 ($\phi=1.5, \zeta=0.15$) を用いた．P_2 は区間 (0, 1) の一様乱数で，P_1 とは別々に発生させた一様乱数である．

LN-SST の乱数実験に用いる発生間隔の対数 x_s は，

$$x_s=\mu+\sigma F^{-1}(P) \tag{17}$$

で作成した．ここで，添え字 s は系列の順番を表し，s＝1，2，・・・，50,000 で，x_s は系列 s の疑似データ群である．μ は LN-Bayes と同じく 6.516 の値を用いた．σ は，LN-Bayes で用いる逆ガンマ分布の平均の平方根値 $\sigma=0.5477(=\sqrt{\zeta/(\varphi-1)})$ を用いた．μ, σ の値から求まる平均発生間隔は約 785 日である．

なお，一様乱数 は，Mersenne Twister の 2002 年版 (mt19937ar.c) [Matsumoto and Nishimura (1998)] で発生させたものを用いた．Matsumoto and Nishimura (1998)

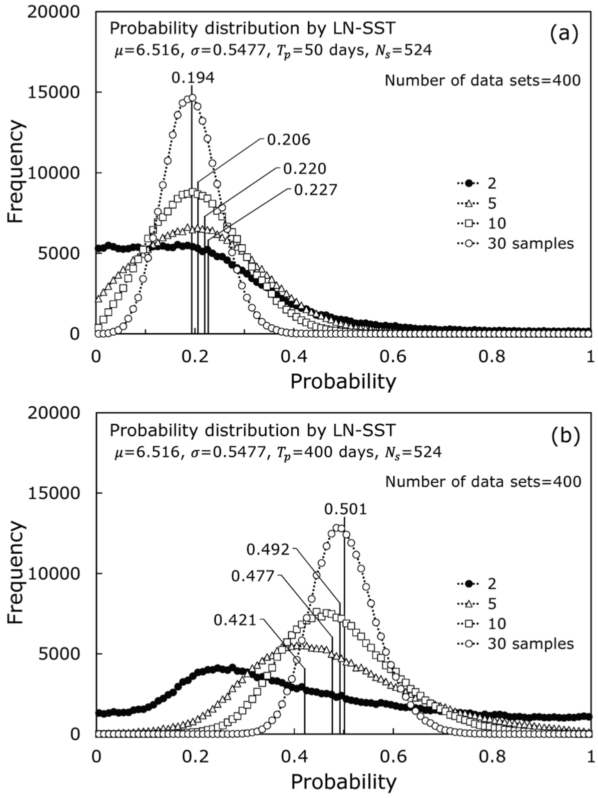

Fig. 13.　Frequency distribution of simulated forecast probabilities using the LN-SST model. T_p is (a) 50 days, and (b) 400 days. μ is the average of logarithm of the recurrence intervals. σ is the standard deviation. T_p is the same as in Fig. 11. N_s is the number of all forecast sequences. The four vertical lines indicate the average of the simulated probabilities.

によれば，この乱数は，乱数列の周期が $2^{19937}-1$（10^{6000} 以上）で，623 次元超立方体の中に均等分布するという長所を持っている．なお，本調査で使用した乱数は 10^8 個程度である．

5.2　シミュレーション結果

5.2.1　予測確率の分布と精度

実際の地震活動では，当該地震の正確な発生確率を知ることはできないので，相似地震の解析では，観測値（1: 当該地震発生，0: 非発生）を真の値とし，予測確率と観測値の差を誤差と見なして，予測成績を調査した．本来ならば，予測確率の精度としては，真の発生確率と直接比較するのが合理的である．与えられた条件から正しい発生確率が確定しても，予測確率値の分布を理論的に求めることは困難なので，シミュレーションで予測確率の分布と精度を推定する．

LN-SST による予測シミュレーションでは，疑似データを作成する際に使用した対数正規母集団のパラメータ μ, σ^2 が，すべての予測に対して，$\mu=6.516$，$\sigma=0.5477$ で，予測期間 ΔT は 365 日である．予測日 T_p を指定すれば，(3) 式または (8) 式で容易に発生確率が計算できる．

予測日 $T_p=50$ 日の場合は，$x_p=\log T_p=3.9120$，$x_f=\log(T_p+\Delta T)=6.0283$ で，正確な発生確率は $P_0=0.1866$ である．間隔データが 2, 5, 10 及び 30 個の場合について，それぞれ約 21 万個の予測確率を求め，確率 1% ごとに区分・集計した結果が Fig. 13a である．データが 2 個の場合は，予測確率が広範囲に分布し，逆 J 字型に近い分布になるので，極端に小さい予測値がかなりの割合を占めることがわかる．間隔データ数が増えるにつれて分布が凸型で，平均値付近に集中した分布になる．図中に予測確率の平均値を間隔データ数毎に示してあるが，いずれも正確な発生確率 P_0 (=0.1866) より過大である．したがって，予測確率は不偏推定値ではない．一方で，間隔データ数が増えるにつれて予測確率の平均は大きい方から正しい値 P_0 に近づく．予測確率の標準偏差が 0.175（データ 2 個），0.123（5 個），0.093（10 個），0.056（30 個）と小さくなるので，正しい確率値に対する二乗平均平方根誤差も小さくなり，予測確率の精度が向上することがわかる．

一方，$T_p=400$ 日の場合は，$x_p=5.9915$, $x_f=6.6399$ で，正確な発生確率は $P_0=0.5059$ である．予測確率の分布はいずれも凸型で，間隔データ数が増えるにつれて，平均値付近に集中し，予測確率の平均値は，小さい側から正しい値 P_0 に近づくとともに，標準偏差も小さくなるので，データ数の増加による予測精度の向上が確認できる (Fig. 13b).

なお，$T_p=200$ 日の場合は，$P_0=0.3634$ で，$T_p=400$ と同様，データ数の増加に伴い，予測確率の平均値は小さい側から P_0 に近づき，正しい確率値と予測確率の平均値との差は $T_p=400$ 日の場合の半分弱であった．

LN-SST では，(7) 式の確率変数 Z $[=\sqrt{(n-1)/(n+1)}\cdot(x_{i+1}-\bar{x})/s]$ が自由度 $n-1$ の t 分布に従うことから，この分布を用いた (8) 式の条件付き確率を予測確率としている．間隔データ数 n が増大すると，(7) 式の平方根の部分は 1 に収束し，$\bar{x}, s$ は母数 μ, σ に，Z は $(x_{i+1}-\mu)/\sigma$ に確率収束する．さらに，自由度 $n-1$ の t 分布は正規分布 $N(0, 1)$ に近づくので，LN-SST による予測確率は，正しい発生確率 P_0 に確率収束する一致性推定量である．

一方，LN-Bayes による予測シミュレーションでは，(14) 式で疑似データ作成に使用する σ^2 は，(6) 式の逆ガンマ分布に従う乱数 σ_j^2 であり，事前分布のパラメータ値を固定しても，正しい発生確率は一義的に決まらない．そのため，LN-Bayes の精度をシミュレーションで

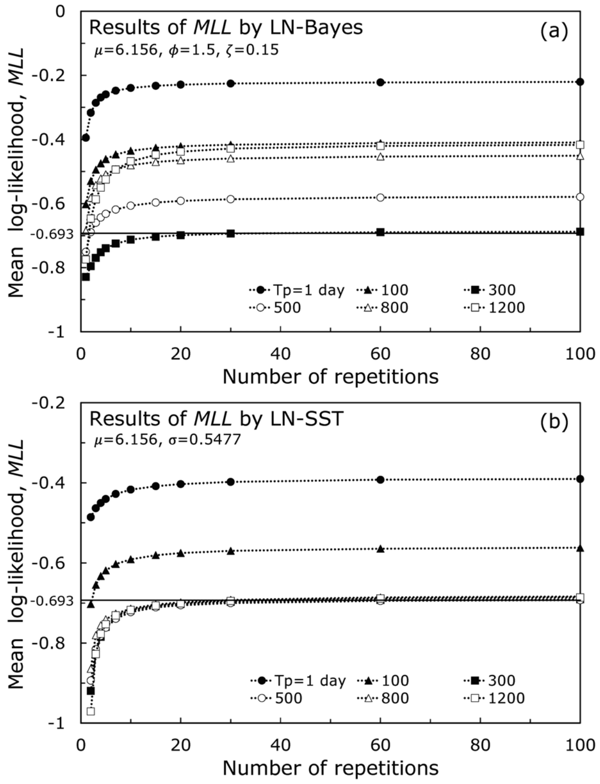

Fig. 14.　Result of mean log-likelihood using (a) the LN-Bayes model, and (b) the LN-SST model. T_p is the same as in Fig. 11, and the forecast period is fixed to 365 days. The horizontal dashed line indicates the score of -0.693, corresponding to the forecast probability of 0.5 for all cases.

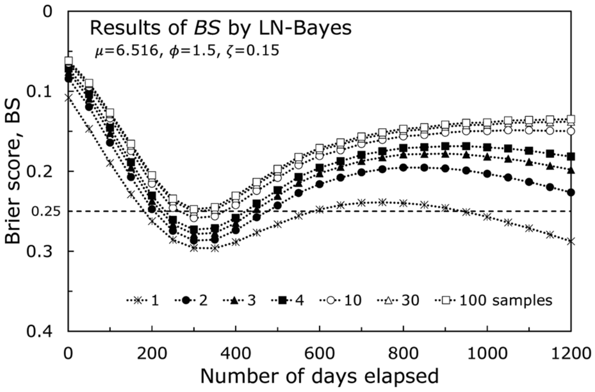

Fig. 15.　Result of the Brier score using the LN-Bayes model. The horizontal dashed line indicates the score of 0.25, corresponding to a forecast probability of 0.5 for all cases.

確認することは難しい．以下では定性的な考察を紹介する．

LN-Bayes では，(5)式の確率変数 Z が自由度 $n+2\phi-1$ の t 分布に従うので，(4)式の条件付き確率を予測確率としている．n が増大すると，LN-SST の場合と同様に，個々の予測確率の精度が向上し，正しい発生確率に確率収束する．

5.2.2　指標による予測成績の比較

LN-Bayes と LN-SST の予測成績について，間隔データ数の違いによる *MLL* と *BS* の変化を調査した．両モデルによる予測成績 *MLL* を Fig. 14 の a と b に示すが，すべての予測日でデータ数が増加するにつれて成績は向上する．成績変化はデータが少ない場合に顕著で，30 個以上はほぼ横ばいである．*BS* の場合も同様な結果である．

相似地震の場合は，Table 2 にあるように，各系列の地震数が 5〜7 個（発生間隔が 4〜6 個）の場合が多いので，予測成績はデータ数にも依存することになる．Fig. 3 と Fig. 4 に示した LN-Bayes 及び LN-SST の予測成績の変化は，Fig. 14 (a) と Fig. 14 (b) で繰り返し回数が少ない場合の変化とかなりよく似ており，データ数増加に伴う予測確率の精度向上が反映したものと考えられる．

シミュレーションの結果 (Figs. 14 (a) and 14 (b)) によって，予測日によっても成績（*MLL* と *BS*）は大きく異なることがわかる．LN-Bayes の予測成績 *BS* の変化を Fig. 15 に示す．地震直後の予測成績が最良で，予測日が遅くなるにつれて成績が悪くなり，T_p＝300 日頃が最も悪い．その後，T_p＝800 日頃まで少しずつ成績は良くなるが，データが少ないときは，それ以後再び悪くなる．*MLL* も同様な変動を示す．

予測確率の精度と予測成績との関係を考察する．真の発生確率 P_0 の現象に対し，予測確率を q とした場合の *MLL* と *BS* の期待値は，

$$\begin{aligned} E(MLL) &= P_0 \log q + (1-P_0)\log(1-q) \\ &= P_0 \log P_0 + (1-P_0)\log(1-P_0) \\ &\quad - \left\{ \frac{(P_0-q)^2}{2P_0(1-P_0)} + \frac{(P_0-q)^3(1-2P_0)}{3P_0{}^2(1-P_0)^2} + \cdots \right\} \end{aligned} \tag{18}$$

$$\begin{aligned} E(BS) &= P_0(1-q)^2 + (1-P_0)q^2 \\ &= P_0(1-P_0) + (P_0-q)^2 \end{aligned} \tag{19}$$

である．P_0 が一定ならば，$q=P_0$ で *MLL* は最大に，*BS* は最小になり，最良のスコア

$$MLL : P_0 \log P_0 + (1-P_0)\log(1-P_0) \tag{20}$$

$$BS : P_0(1-P_0) \tag{21}$$

が期待できる．この最良期待値は，P_0＝0.5 で最も悪く，P_0 が 0 または 1 に近いほど良いスコアになる．

(18), (19)式から，予測確率 q が P_0 の近傍にあれば，最良のスコアとの差は，*MLL* が $(P_0-q)^2/2P_0(1-P_0)$ 程度で，*BS* は $(P_0-q)^2$ であり，成績の変動は小さい．逆に，予測確率 q が P_0 から大きく離れていると，q の変化でスコアは大きく変動する．

Table 5. Results of N-test on probabilities of simulation using random numbers.

	Number of repetitions								
LN-Bayes	1	2	3	4	5	7	10	30	100
Number of days elapsed	%	%	%	%	%	%	%	%	%
1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	15.0	36.6
50	0.0	0.0	0.0	0.0	0.0	0.0	0.4	18.1	38.3
100	0.0	0.0	0.0	0.0	0.0	0.1	1.7	26.8	49.9
150	0.0	0.0	0.0	0.0	0.0	0.0	0.9	16.7	33.5
200	0.0	0.0	0.0	0.0	0.1	1.0	5.0	27.8	41.2
250	0.0	0.1	0.6	2.2	5.1	11.4	19.6	37.9	45.7
300	93.2	72.6	64.0	60.0	57.1	54.4	53.6	51.0	50.4
350	100.0	100.0	99.8	98.4	95.7	89.1	79.5	57.2	48.4
400	100.0	100.0	100.0	100.0	99.9	98.9	96.4	79.7	66.9
500	100.0	100.0	100.0	100.0	100.0	99.9	98.4	72.3	52.5
600	100.0	100.0	100.0	100.0	100.0	100.0	99.9	86.6	69.0
700	100.0	100.0	100.0	100.0	100.0	100.0	100.0	83.7	58.5
800	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.3	78.4
900	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.5	68.4
1000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.4	73.0
1100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.9	82.2
1200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	91.8
LN-SST		2	3	4	5	7	10	30	100
Number of days elapsed		%	%	%	%	%	%	%	%
1		0.0	0.0	0.0	0.0	0.0	0.0	6.6	29.5
50		0.0	0.0	0.0	0.0	0.0	0.2	13.4	34.5
100		13.3	1.6	1.4	1.8	4.1	9.0	29.0	41.1
150		96.5	72.1	56.7	50.6	46.7	46.0	48.8	51.3
200		100.0	98.8	93.2	87.4	78.5	71.5	58.1	53.0
250		100.0	100.0	99.9	99.2	96.5	90.4	70.8	61.4
300		100.0	100.0	100.0	99.8	98.0	93.5	73.9	61.1
350		100.0	100.0	100.0	99.7	96.5	88.8	66.1	55.1
400		100.0	100.0	100.0	99.9	98.0	92.2	68.3	56.1
500		100.0	100.0	100.0	100.0	98.7	93.3	71.4	60.0
600		100.0	100.0	100.0	100.0	99.6	96.8	74.4	60.9
700		100.0	100.0	100.0	100.0	99.9	97.1	76.6	60.4
800		100.0	100.0	100.0	100.0	100.0	100.0	88.1	71.8
900		100.0	100.0	100.0	100.0	100.0	100.0	84.5	59.8
1000		100.0	100.0	100.0	100.0	100.0	100.0	90.2	58.9
1100		100.0	100.0	100.0	100.0	100.0	100.0	92.9	60.9
1200		100.0	100.0	100.0	100.0	100.0	100.0	98.8	74.4

Remarks: LN-Bayes and LN-SST are the same as in Table 4. *BS* is the Brier score of the real observation. Shading means the same as Table 4.

予測確率の分布 (Figs. 13a and 13b) をみると，発生間隔のデータが2個から5個に増えると，P_0から大きく外れた予測確率はかなり大幅に減るので，予測成績は大きく向上する．一方，データが10個の場合は，予測確率がすでにP_0近傍に多く分布しており，30個に増えても*MLL*と*BS*の変化は比較的小さい．間隔データ数が大きい場合は，予測精度がよいので，*MLL*と*BS*はデータ数にほとんど依存せず，(20)，(21)式に示すように，P_0によってスコアはほぼ決まることから，予測成績はT_pに大きく依存することになる．

LN-SSTのシミュレーションの場合では，$T_p=1$日で最小の発生確率値$P_0=0.131$となり，$T_p=400$日でほぼ$P_0=0.5$になり，$T_p=800$日を超えると$P_0=0.58$程度となり，その後はあまり変化しない．Fig. 14bで，データ数nが大きい場合をみると，*MLL*の最大は$T_p=1$日のときで，T_pが100日，300日と増大するにつれて*MLL*の値は小さくなり，$T_p=300$日を超えると*MLL*の値はほとんど変化しなくなる．これは上記のT_p，P_0及び*MLL*との関係を反映したものである．LN-Bayesの場合は，P_0が確率分布するので，T_pと*MLL*との関係は複雑であるが，T_pが大きくなるにつれて，*MLL*は一度小さくなり，その後増大する変化がみられる (Fig. 14a).

5.2.3　予測確率の偏りと個数検定 (N-test)

5.2.1で予測確率の不偏性を調べた際には，計算条件を固定して，真の発生確率1個に対し，シミュレーションで予測確率の分布を求め，平均と真の発生確率との相

違を調べた．一方，個数検定 (N-test) では，予測確率から計算したイベント発生回数の理論分布と，観測された現象の発生回数を比較し，多数の予測確率が全体として過大または過小であるかを調べる．シミュレーションで得られた結果についても参考までに N-test を試みた (Table 5).

LN-SST による予測の N-test 結果は，間隔データが 10 個以下で，かつ T_p=50 日以内のときは，すべてのケースが下側 2.5% の棄却域にあり，予測が過大である．一方，T_p=200 日の間隔データが 2 個と 3 個のときと，T_p=250 日以上の間隔データが 7 個以下のときは，ほとんどのケースが上側 2.5% の棄却域にあり，予測は過小である．これらの結果は，先に得られた予測確率の偏りの結果と矛盾していない．なお，T_p=150 日のときは，すべてのケースが棄却とはならず，予測確率全体の偏りが非常に小さいことを示唆している．また，間隔データが 30 個以上では，棄却できるケースはほとんどない.

LN-Bayes による予測の N-test 結果は，間隔データが 10 個以下で，かつ T_p=200 日以内のときは，ほとんどのケースが下側 2.5% の棄却域にあり，予測は過大である．一方，T_p=400 日以上では，間隔データが 10 個以下のほとんどのケースが上側 2.5% の棄却域にあり，予測は過小である．すなわち，予測確率の過大・過小は予測日（直近地震からの経過時間）に強く依存することを示している.

5.2.4　LN-Bayes と LN-SST の成績比較

LN-Bayes と LN-SST は，使用した疑似データの作成方法が異なるので，R-test で比較することはできない．そこで，両方のデータセットに対し，両予測モデルを適用し，予測成績を比較した．LN-Bayes 用に作成した疑似データセットを使って，LN-Bayes と LN-SST による予測成績 *MLL* の差を Fig. 16 に示す．一部の例外を除き，LN-Bayes が LN-SST より良い成績である．特に，間隔データが数個以下のときに事前分布の効果が見られ，ほとんどのケースで明瞭な成績差が認められる．しかし，間隔データが増えるにつれて成績の差は小さくなり，間隔データが 10 個以上のケースでの成績差は非常に小さい．間隔データ数の増加に伴う成績差の縮小は，相似地震の場合 (Figs. 3 and 4) と整合している.

LN-SST 用の疑似データセットによる予測も，LN-Bayes の予測成績が LN-SST より優れている．間隔データ数に対する依存性も，LN-Bayes 用の疑似データセットを使用した場合 (Fig. 16) とよく似ている．これは，LN-Bayes で使用する σ^2 の逆ガンマ事前分布の平均 (=0.3) を用いて，LN-SST 用の疑似データセット ($\sigma=\sqrt{0.3}$ =0.5477) を作成したためと考えられる．当該系列の母

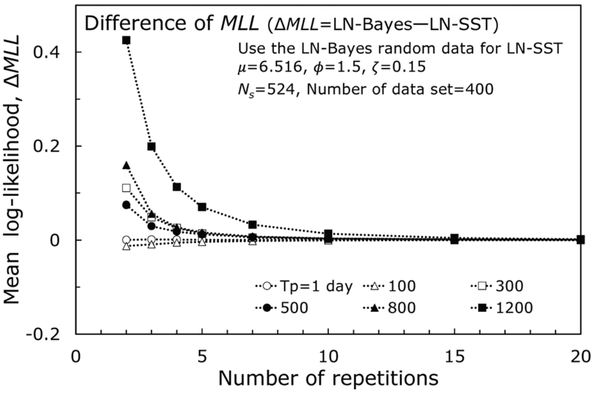

Fig. 16.　Difference in mean log-likelihood between LN-Bayes and LN-SST models. T_p is the same as in Fig. 11, and the forecast period is fixed to 365 days.

数 σ^2 と逆ガンマ事前分布がどの程度整合していれば，LN-Bayes の予測成績が LN-SST よりよくなるのか興味深い問題である.

5.3　パラメータ値の違いによる成績への影響確認

今回の LN-Bayes によるシミュレーションでは，逆ガンマ事前分布のパラメータ値に，相似地震を使った 2009 年と 2010 年の確率予測実験時の値 (ϕ=1.5, ζ=0.15) を採用した．しかし，2006 年と 2008 年の解析では，2006 年と 2008 年の確率予測実験時の値 (ϕ=2.5, ζ=0.44) を用いた（**2.2.1** 参照)．後者の値 (ϕ=2.5, ζ=0.44) を使ってシミュレーションを行い，パラメータ値の違いによる成績への影響を確認・評価した．結果は，平均対数尤度 *MLL* とブライアスコア *BS* ともに，前者のパラメータ値 (ϕ=1.5, ζ=0.15) を使った方が成績は良く，その差は，*MLL* が 0.005〜0.011，*BS* が 0.002〜0.005 であった．そして，間隔データ数への依存性にはあまり影響しないことも確認できた.

5.4　予測確率の上限

間隔データが 2 個で，予測日 T_p が 1 日，150 日そして 800 日のときに LN-Bayes で予測した予測確率の頻度分布を Fig. 17 に示す．直近地震から予測時点までの経過時間 T_p の対数が地震の発生間隔の対数の平均 μ を超えると，予測確率の頻度分布が最大確率値の近傍に集中し，分布に頭打ちが見られるようになる.

予測日 T_p を一定にすると，予測期間 (x_p=log T_p, x_f=log(T_p+ΔT)) が決まり，n 個の疑似データ x_i から得られる平均 $\bar{x}$ と標準偏差 s に対する予測確率 $P_q(\bar{x}, s)$ は，(4) 式から計算できる．ここで，$\bar{x}$ と s (>0) を $\bar{x}_t = x_p - \sqrt{2\zeta/(ns^2+2\zeta)}\cdot(x_p-\bar{x})$ と s=0 に置き換えて，予測確率 $P_q(\bar{x}, s)$ と $P_q(\bar{x}_t, 0)$ を比較すると，Z_p は $Z_p(\bar{x}, s)=Z_p(\bar{x}_t, 0)$ で，Z_f は $Z_f(\bar{x}, s)<Z_f(\bar{x}_t, 0)$ となる．したがって，予測確

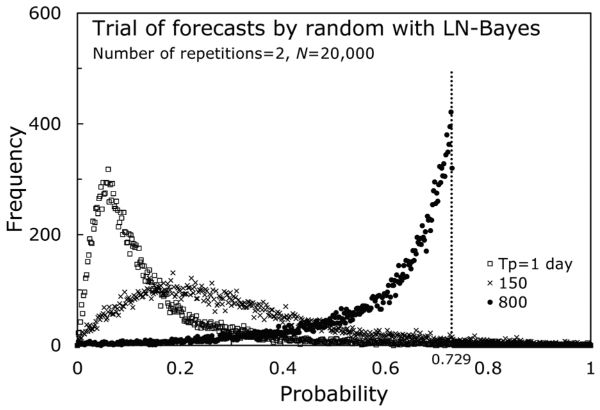

Fig. 17.　Frequency distribution of forecast probabilities determined by random experiments with the LN-Bayes model. N is the number of forecast sequences. The number of repetitions is the number of recurrence intervals. T_p is the same as in Fig. 11. The vertical dotted line is the maximum value of the forecast probability when T_p is 800 days.

率 $P_q(\bar{x}, s)$ と $P_q(\bar{x}_t, 0)$ は，$P_q(\bar{x}, s) < P_q(\bar{x}_t, 0)$ の関係であり，P_q の最大値は $s=0$ の場合に限られる．正規分布 $N(\mu, \sigma^2)$ の性質から，確率変数 s^2/σ^2 は自由度 $n-1$ の χ^2 分布に従うので，$s=0$ 近傍の確率密度は $n=2$ のときに最大で，n が大きくなると急速に小さくなる．このため，$n=2$ のときに予測確率の頭打ちが生じても，n が大きくなると明瞭な頭打ちは起きなくなる．

一方，$n=2$ かつ $s=0$ の場合には，$\bar{x}$ が x_p にかなり近いとき（数値実験結果から求めた回帰直線式では $\bar{x}=0.864x_p+0.610$）に予測確率の最大値が出現する．$\bar{x}$ の分布は正規分布 $N(\mu, \sigma^2/2)$ であり，$\bar{x}$ が μ から離れると出現する割合は小さくなる．したがって，x_p が μ より小さいときは，予測確率が最大値となる近傍に出現する割合は小さい．さらに，Fig. 17 に示すように，P_q の変動幅が大きいので，頭打ちは目立たない．逆に，x_p が μ より大きいときは，P_q の変動幅は小さいので，予測確率の頭打ちが明瞭になる．

5.5　予測モデルの選択と検証

本報告では，発生間隔の分布として対数正規分布を採用し，LN-Bayes と LN-SST による予測の精度や成績を調べた．一方，地震調査研究推進本部地震調査委員会 (2001) では，大地震の長期評価を開始する際に，5 種類の分布（BPT 分布，対数正規分布，ガンマ分布，ワイブル分布，2 重指数分布）を比較検討し，総合的判断で BPT 分布を採用した．ただし，直近地震の発生時期が判らない場合などでは，指数分布を使用している．データ処理については，パラメータ推定値または標準値 $\alpha=0.24$ を (3) 式の条件付き確率の式に直接代入する Plug-in 方式を採用している．

過去のデータが多数あり，パラメータが十分な精度で求まるときは，物理的解釈が容易な分布型を優先的に採用することは順当である．しかし，データが少ないときは，データ処理の方法によって予測の精度・成績はかなり異なるので，発生間隔分布の型とデータ処理の方法を同時に検討する必要がある．

Plug-in 方式は多様な分布に適用できるが，母数の推定誤差が予測確率に直に反映されるので，データが少ないときは，予測確率は不安定で，予測成績も劣ることが多い．

ベイズ統計は，多数の事例から得られた経験則をパラメータの事前分布として使用し，経験則と当該データを使用することから，少ないデータでも予測成績が比較的良く，色々な分布型に適用できるので，好都合である．しかし，事前分布の型とそのパラメータ値を決めるには，データ収集とその解析に相当な労力を要する．

小標本論（精密標本論）は，事前分布を必要とせず，比較的安定した結果が得られるが，適用できるものが正規分布と対数正規分布に事実上限られる．

予測モデルを評価する重要な要素は，対象地震に対する予測の精度・成績と信頼性である．事前に予測した結果の検証は予測期間終了後であり，予測期間が長期間であれば，検証は必然的に遅くなる．実用的な予測モデルを選定する際は，様々な視点からシミュレーションを実施し，モデル間の成績比較を十分に行うことが望ましい．

§6.　ま と め

本報告では，相似地震（小繰り返し地震）カタログと疑似データセットを使って，ベイズ統計対数正規分布モデル (LN-Bayes)，小標本論対数正規分布モデル (LN-SST)，および Plug-in 方式の指数分布モデル (Exp-pin) による発生確率予測の検証実験を行い，繰り返し回数と予測成績との関係など，少数データに由来する基本的な特性を調査した．

相似地震カタログを用いた検証実験では，予測に使用する地震の発生間隔データを少なくすると，3 つの統計モデルとも予測成績（平均対数尤度 *MLL* とブライアスコア *BS*）は悪くなる．成績変化は，予測確率の信頼度と分離度が低下することによるものであり，間隔データが 3 個以下と非常に少ない場合に顕著である．

各統計モデルの予測成績を比較すると，定常ポアソン過程に基づく Exp-pin は，全体として LN-Bayes および LN-SST より有意に劣り，相似地震の発生率（単位時間

あたりの発生回数）が非定常であること，すなわち直近地震からの経過時間に依存することを確認した．

LN-Bayes と LN-SST の比較では，LN-SST の方が良い場合もあるが，全体としては LN-Bayes の方が優れており，間隔データが 2 個および 3 個のときの成績差はかなり大きい．4 回実施した岡田らによる確率予測実験での成績変動は，LN-SST より LN-Bayes が小さく，比較的安定している．すなわち，間隔データ数が非常に少ないときは，事前分布の情報が予測精度の向上に役立っている．

発生予測のシミュレーションでは，LN-SST 用の疑似データ（母数 μ, σ^2 が共通一定の対数正規乱数）と LN-Bayes 用の疑似データ（母数 μ が共通一定で，σ^2 が逆ガンマ分布に従う乱数で，系列毎に異なるもの）を作成した．

LN-SST による予測確率は不偏推定量ではなく，直近地震からの経過時間が短い予測日の場合は，予測確率の平均は正しい発生確率より大きく，予測日が平均発生間隔より後の場合は，正しい発生確率よりも予測確率は小さくなる．このことは，予測結果を個数検定 (N-test) した結果と整合する．間隔データが増加するにつれて，予測確率の平均は正しい発生確率に近づき，分散は小さくなることから，予測確率の精度が向上することを確認した．理論的考察から，LN-Bayes と LN-SST の予測確率は一致性推定量であることを示した．

間隔データの増加に伴う成績向上は，LN-Bayes と LN-SST とも設定したすべての予測日で見られた．成績変化は，間隔データが少ないときほど顕著であり，データが 30 個以上になるとほぼ一定である．データが少ないときの成績向上は，相似地震の場合と類似しており，間隔データが非常に少ないときは，データを 1 個でも増やせれば，予測成績の向上が期待できる．

一方，間隔データ数を固定すると，成績のスコアは，予測日が遅くなるにつれて，最良の値から急速に悪化し，最悪の値に達したのち回復する．平均発生間隔よりも予測日が遅くなるとあまり変化しなくなるが，間隔データが少ないときは下降傾向が見られた．このような成績変化は，発生確率が 0.5 のときに悪く，0 または 1 に近いときは良くなりやすいこと，および予測日によって真の発生確率が変化することで説明できる．

LN-Bayes と LN-SST の比較のために，2 種類の疑似データセットで，シミュレーションを行った．両データセットとも，LN-Bayes の成績が LN-SST より基本的に優れており，成績差は間隔データが少ないときに大きく，データ数が 10 個程度以上になると小さい．LN-SST 用のデータセットでも LN-Bayes の成績が良かったのは，データセットに使用した母数 σ^2 が，LN-Bayes で使われている σ^2 の事前分布の平均と一致していたためと考えられる．データ数が少ないときに見られる両モデルの成績差の変化は，相似地震の場合と似ている．

なお，LN-Bayes の事前分布のパラメータ値の違いは，予測成績にあまり影響しないことをシミュレーションで確認した．間隔データが 2 個の場合には，予測日が平均発生間隔よりも後になると，予測確率の頭打ちが見られることについて考察した．

謝　　辞

本調査には，東北大学で作成した相似地震カタログを使わせていただきました．また，国立研究開発法人防災科学技術研究所,北海道大学，弘前大学，東北大学，東京大学，名古屋大学，京都大学，高知大学，九州大学，鹿児島大学，国立研究開発法人産業技術総合研究所，国土地理院，国立研究開発法人海洋研究開発機構，青森県，東京都，静岡県，神奈川県温泉地学研究所及び気象庁のデータを用いて気象庁・文部科学省が協力してデータ処理した結果を使用させていただきました．乱数作成には，広島大学大学院理学研究科松本眞教授のホームページ (http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/MT2002/mt19937ar.html，参照 2017-02-25) からダウンロードした Mersenne Twister の 2002 年版 (mt19937ar.c) を使用させて頂きました．Fig. 1 の作成には，GMT [Wessel and Smith (1991)] を使用させて頂きました．また，気象庁気象大学校の吉田康宏教授と匿名の査読者 1 名からは貴重なコメントを頂き，論文の改訂に大変役に立ちました．ここに記して感謝いたします．

文　　献

Bakun, W.H. and T.V. McEvilly, 1984, Recurrence models and Parkfield, California, Earthquakes, J. Geophys. Res., **89**, 3051–3058, doi: 10.1029/JB089iB05p03051.

Bakun, W.H. and A. G. Lindh, 1985, The Parkfield, California, Earthquake Prediction Experiment, Science, v. **229**, no. 4714, 619–624, doi: 10.1126/science.229.4714.619.

Brier, G.W., 1950, Verification of forecasts expressed in terms of probability, Mon. Weather Rev., **78**, 1–3.

長谷川　昭，2002，的中した地震発生予測—プレート境界の活動の様子が見えてきた，科学，**72**，581–583.

林　豊，2016，確率予測の採点方式「拡張ブライアスコア」とその適用例，地震予知連絡会会報，**96**，486–489.

Igarashi, T., T. Matsuzawa, and A. Hasegawa, 2003, Repeating earthquakes and interplate aseismic slip in the northeastern Japan subduction zone, J. Geophys. Res., **108**, 2249, doi: 10.1029/2002JB001920.

地震調査研究推進本部地震調査委員会，2001，長期的な

地震発生確率の評価手法について，99 pp.

Kagan, Y.Y. and D.D. Jackson, 1995, New seismic gap hypothesis: Five years after, J. Geophys. Res., **100**, 3943-3959.

気象研究所地震火山研究部・地震火山部・気象大学校・札幌管区気象台・仙台管区気象台・大阪管区気象台・福岡管区気象台・沖縄気象台，2014，日本各地域の繰り返し相似地震発生状況に関する研究，気象研究所技術報告，72.

Matsumoto, M., T. Nishimura, 1998, Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Transactions on Modeling and Computer Simulation (TOMACS), Vol. 8 (1), 3-30, doi: 10.1145/272991.272995.

Matsuzawa, T., T. Igarashi, and A. Hasegawa, 1999, Characristic small earthquake sequence off Sanriku, Japan, EOS Trans. Am. Geophys. Union, **80**, Fall Meet. Suppl., Abstract S41B-07, F724.

Matsuzawa, T., T. Igarashi, and A. Hasegawa, 2002, Characteristic small-earthquake sequence off Sanriku, northeastern Honshu, Japan, Geophys. Res. Lett., **29**, 1543, doi: 10.1029/2001GL014632.

Matthews, M.V., W.L. Ellsworth, and P.A. Reasenberg, 2002, A Brownian model for recurrent earthquakes, Bull. Seismol. Soc. Am., **92**, 2233-2250, 2002.

Nishenko, S.P., 1991, Circum-Pacific seismic potential: 1989-1999, Pure Appl. Geophys., **135**, 169-259, doi: 10.1007/BF00880240.

Nomura, S., Y. Ogata, F. Komaki, and S. Toda, 2011, Bayesian forecasting of recurrent earthquakes and predictive performance for a small sample size, J. Geophys. Res. **116**, doi: 10.1029/2010JB007917.

岡田正実，2009，繰り返し地震および余震の確率予測，地震 2，**61**，S143-S153.

岡田正実・高山博之・弘瀬冬樹・内田直希，2007，地震長期確率予測に使用する更新過程対数正規分布モデルのパラメータ事前分布，地震 2，**60**，85-100.

岡田正実・内田直希・高山博之，2008，相似地震の確率予測実験，<http://www.aob.gp.tohoku.ac.jp/~uchida/kenkyuu/souji-yosoku/souji-kakuritsu2008.html>，(参照 2015-01-21).

岡田正実・内田直希・青木重樹，2009，相似地震の確率予測実験，<http://www.aob.gp.tohoku.ac.jp/~uchida/kenkyuu/souji-yosoku/souji-kakuritsu2009.html>，(参照 2015-01-21).

岡田正実・内田直希・青木重樹，2010，相似地震の確率予測実験，<http://www.aob.gp.tohoku.ac.jp/~uchida/kenkyuu/souji-yosoku/souji-kakuritsu2010.html>，(参照 2015-01-21).

岡田正実・内田直希・青木重樹，2011，相似地震の確率予測実験，<http://www.aob.gp.tohoku.ac.jp/~uchida/kenkyuu/souji-yosoku/souji-kakuritsu2011.html>，(参照 2015-01-21).

Okada, M., N. Uchida, and S. Aoki, 2012, Statistical forecasts and tests for small interplate repeating earthquakes along the Japan Trench, Earth Planets Space, **64**, 703-715, doi: 10.5047/eps.2011.02.008.

立平良三，1999，気象予報による意思決定―不確実情報の経済価値―，東京堂出版.

Toth, Z., O. Talagrand, G. Candille, and Y. Zhu, 2003, Probability and ensemble forecasts, Forecast Verification: A Practitioner's Guide in Atmospheric Science. Edited by I.T. Jolliffe and D.B. Stephenson, John Wiley & Sons, Ltd., 137-163.

Uchida, N., A. Hasegawa, T. Matsuzawa, and T. Igarashi, 2004, Pre- and post-seismic slow slip on the plate boundary off Sanriku, NE Japan associated with three interplate earthquakes as estimated from small repeating earthquake data, Tectonophysics, **385**, 1-15, doi: 10.1016/j.tecto.2004.04.015.

宇津徳治，1999，地震活動総説，東京大学出版会，896pp.

Wessel, P. and W.H.F. Smith, 1988, New, improved version of generic mapping tools released, EOS Trans. Am. Geophys. Union, vol. 79 (47), 579, doi: 10.1029/98EO00426.

山下智志・川口　昇・敦賀智裕，2003，信用リスクモデルの評価方法に関する考察と比較，金融庁金融研究研修センターディスカッションペーパー，<http://www.fsa.go.jp/frtc/seika/discussion/2003/20031031.pdf>，(参照 2015-12-25).