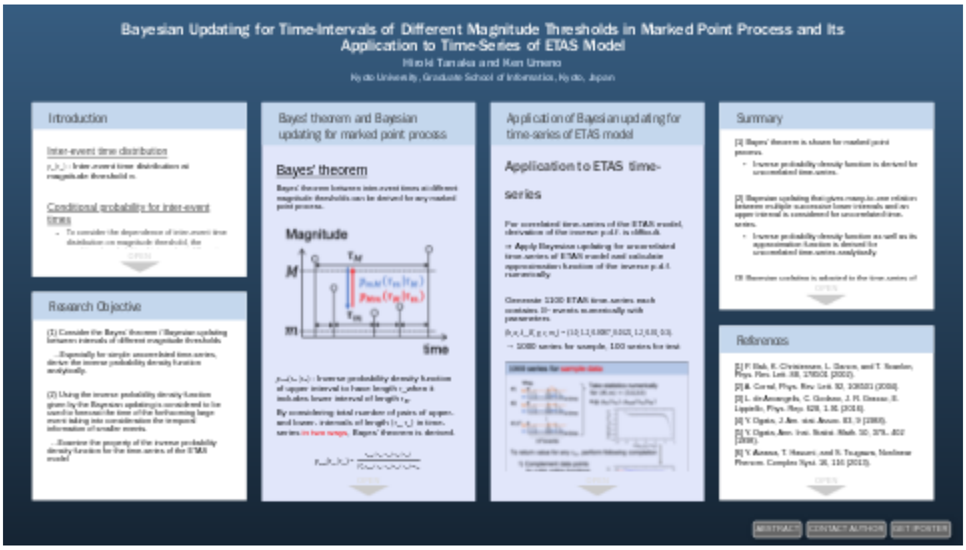
Bayesian Updating for Time-Intervals of Different Magnitude Thresholds in Marked Point Process and Its Application to Time-Series of ETAS Model



Hiroki Tanaka and Ken Umeno

Kyoto University, Graduate School of Informatics, Kyoto, Japan

PRESENTED AT:

INTRODUCTION

Inter-event time distribution

Inter-event time distribution of earthquakes (Fig.1) is important for probabilistic forecasting of forthcoming earthquakes as well as understanding seismic properties [1]. Studies of the properties of inter-event time distributions include,

- Scaling universality [2,3] that is often examined using the ETAS model [4,5]
- Dependence on magnitude threshold [6]

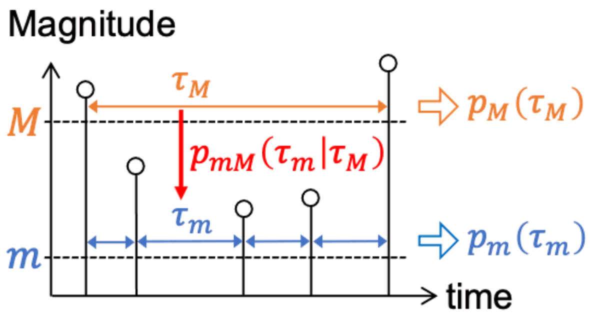


Figure 1. Schematic of M-T diagram. Inter-event time distribution of earthquakes is determined by setting a threshold on this M-T diagram. Denote the inter-event times at magnitude threshold m by τ_m, and inter-event time distribution by $p_m(\tau_m)$.

ETAS model

ETAS model is inhomogeneous Poisson process that combines the GR law and the Omori-Utsu law. Magnitude and the occurrence time of events in this model are generated obeying these laws [4,5] as follows;

$$P(m) \propto 10^{-bm}$$

$$\lambda(t) = \lambda_0 + \sum_{i;t_i<t} \frac{K 10^{\alpha(m_i - m_0)}}{(t-t_i+c)^p}$$

Dependence on magnitude threshold

- To consider the dependence of inter-event time distribution on magnitude threshold, the conditional probability ($p_{mM}(\tau_m|\tau_M)$) between intervals at different magnitude thresholds m and $M\ (= m + \Delta m)$is useful.
- $p_{mM}(\tau_m|\tau_M)$ is the probability density function of lower interval to have length τ_m under the condition that it is included in the upper interval of length τ_M (Fig. 1).
- Inter-event time distributions of different magnitude threshold are connected via the conditional probability in the following way [7].

$$p_m(\tau_m) = 10^{-b\Delta m} \int_{\tau_m}^{\infty} \frac{\tau_M}{\langle\langle \tau_m \rangle\rangle_{\tau_M}} p_{mM}(\tau_m|\tau_M) p_M(\tau_M) d\tau_M,$$

$$\text{where } \langle\langle \tau_m \rangle\rangle_{\tau_M} := \int_0^{\infty} \tau_m p_{mM}(\tau_m|\tau_M) d\tau_m.$$

Utilize conditional probability for forecasting

Conditional probability gives the probability of lower interval given the length of upper interval.

On the contrary, intervals at different magnitude thresholds can be statistically connected by considering the probability of the length of upper interval given the lower intervals.

Such inverse probability is associated with the conditional probability by Bayes' theorem or Bayesian updating.

It is thought that the inverse probability can be used to forecast large event with the information of smaller events. In this study, we derive the inverse probability analytically and numerically, and examine its effectiveness for forecasting.

RESEARCH OBJECTIVE

(1) Consider the Bayes' theorem / Bayesian updating between intervals at different magnitude thresholds. In particular, for simple uncorrelated time-series, derive the inverse probability density function analytically as well as its approximation function.

(2) Apply Bayesian updating to the correlated time-series of the ETAS model to examine the property of the inverse probability density function and possibility to use it to forecast the occurrence time of the forthcoming large event taking into account the temporal information of smaller events.

(1) BAYES' THEOREM AND BAYESIAN UPDATING FOR MARKED POINT PROCESS

Bayes' theorem

Bayes' theorem between inter-event times at different magnitude thresholds can be derived for any marked point process.

Denote $p_{Mm}(\tau_M|\tau_m)$ be the inverse probability density function of upper interval to have length τ_M when it includes lower interval of length τ_m (Fig. 2).

Bayes' theorem is derived by considering total number of pairs of upper- and lower- intervals of length $\{\tau_M, \tau_m\}$ in time-series in two ways.

$$p_{Mm}(\tau_M|\tau_m) = \frac{z_{mM}(\tau_m|\tau_M) z_M(\tau_M)}{\int_{\tau_m}^{\infty} z_{mM}(\tau_m|\tau_M) z_M(\tau_M) d\tau_M}$$

where

$$z_m(\tau_m) = \tau_m p_m(\tau_m) \langle \tau_m \rangle$$

$$z_{mM}(\tau_m|\tau_M) = \tau_m p_{mM}(\tau_m|\tau_M) / \langle\langle \tau_m \rangle\rangle_{\tau_M}$$

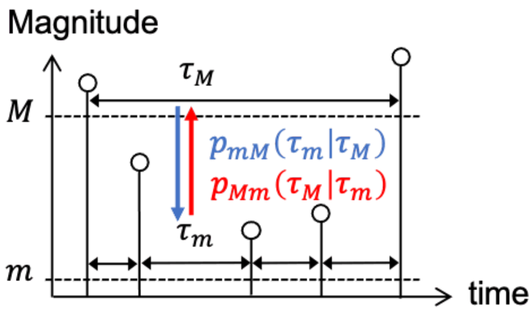


Figure 2. Schematic of the relation between the conditional probability and the inverse probability.

Example : uncorrelated time-series

For uncorrelated time-series generated by the background seismicity ($\lambda(t) \equiv \lambda_0$) in the ETAS model, conditional / inverse probability density functions can be derived analytically as follows.

$$p_{mM}(\tau_m|\tau_M) = \frac{e^{-A_{\Delta m}\frac{\tau_M}{\langle \tau_M \rangle}} \delta(\tau_M - \tau_m) + \frac{A_{\Delta m}}{\langle \tau_M \rangle} e^{-A_{\Delta m}\frac{\tau_m}{\langle \tau_M \rangle}} \left\{ A_{\Delta m}\frac{\tau_M - \tau_m}{\langle \tau_M \rangle} + 2 \right\} \theta(\tau_M - \tau_m)}{A_{\Delta m}\frac{\tau_M}{\langle \tau_M \rangle}}$$

$$p_{Mm}(\tau_M|\tau_m) = \frac{e^{-\frac{\tau_M - \tau_m}{\langle \tau_m \rangle}} \delta(\tau_M - \tau_m) + \frac{A_{\Delta m}}{\langle \tau_M \rangle} e^{-\frac{\tau_M - \tau_m}{\langle \tau_M \rangle}} \left\{ A_{\Delta m}\frac{\tau_M - \tau_m}{\langle \tau_M \rangle} + 2 \right\} \theta(\tau_M - \tau_m)}{(A_{\Delta m} + 1)^2}$$

where $A_{\Delta m} := \langle \tau_M \rangle / \langle \tau_m \rangle - 1 \simeq 10^{b\Delta m} - 1$, $\delta(\cdot)$ is the Dirac's delta function, and $\theta(\cdot)$ is the unit step function. Especially when $\Delta m > b^{-1} \log_{10} 3$, $p_{Mm}(\tau_M|\tau_m)$ has a peak at $\tau_M^{\mathrm{MAX}} = \tau_m + \langle \tau_M \rangle \left(1 - \frac{2}{A_{\Delta m}}\right)$.

Bayesian updating

Bayes' theorem gives one-to-one relationship between upper and lower intervals. This can be extended to Bayesian updating that gives the relation between many lower-intervals and an upper interval.

Denote $p_{Mm}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)})$ be the inverse probability density function of upper interval to have length τ_M when it includes successive lower intervals of length $\{\tau_m^{(1)},\cdots,\tau_m^{(n)}\}$.

Bayesian updating is derived by considering total number of pairs of upper- and successive lower- intervals of length $\{\tau_M,\tau_m^{(1)},\cdots,\tau_m^{(n)}\}$ in time-series in two ways.

For simple uncorrelated time-series, analytic inverse probability density function as well as its approximation function that is applicable to correlated time-series can be derived as follows.

- Analytic form of inverse p.d.f. (same functional form as that of Bayes' theorem).

$$p_{Mm}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)}) = \left(\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right)^2 e^{-\frac{\tau_M-\sum_{i=1}^n\tau_m^{(i)}}{\langle\tau_m\rangle}}\delta(\tau_M-\textstyle\sum_{i=1}^n\tau_m^{(i)})$$
$$+\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle^2}\left(1-\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right)e^{-\frac{\tau_M-\sum_{i=1}^n\tau_m^{(i)}}{\langle\tau_M\rangle}}\left\{A_{\Delta m}\frac{\tau_M-\sum_{i=1}^n\tau_m^{(i)}}{\langle\tau_M\rangle}+2\right\}\theta(\tau_M-\textstyle\sum_{i=1}^n\tau_m^{(i)})$$

- Approximation function

$$p_{Mm}^{\mathrm{approx}}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)}) = p_{Mm}^{\mathrm{kernel}}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)}) - (\mathrm{Correction})$$

where

$$p_{Mm}^{\mathrm{kernel}}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)}) = \frac{\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}}{\left(1-\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right)^{n-1}}\frac{\tau_M}{\langle\tau_m\rangle_{\tau_M}}\left[\prod_{i=1}^n\frac{p_{mM}(\tau_m^{(i)}|\tau_M)}{p_m(\tau_m^{(i)})}\right]p_M(\tau_M),$$
$$(\mathrm{Correction}) = \frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\frac{(n-1)}{\left(1-\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right)^{n-1}}\left[\prod_{i=1}^n\frac{P_i}{p_m(\tau_m^{(i)})}\right]p_M(\tau_M).$$

By numerical calculation, p_{Mm}^{approx} is shown to approximate the inverse probability density function (Fig. 3 and Fig. 4).

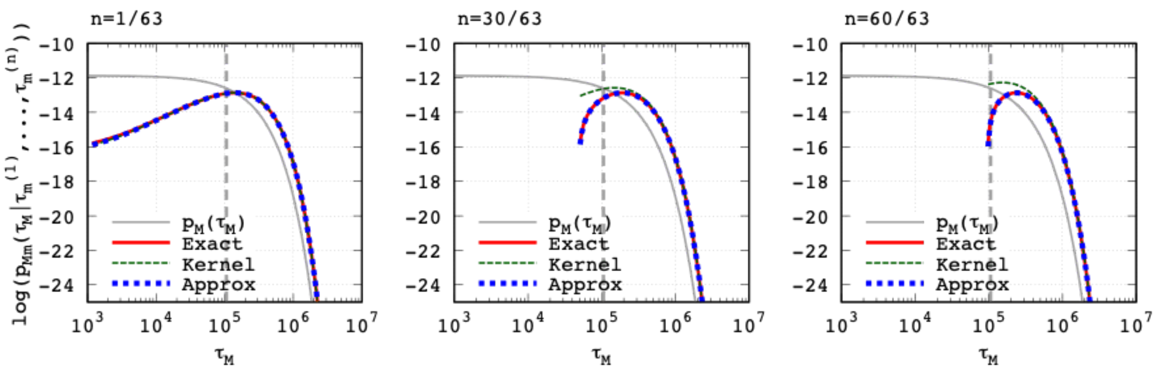


Figure 3. Example of Bayesian updating for uncorrelated time-series.

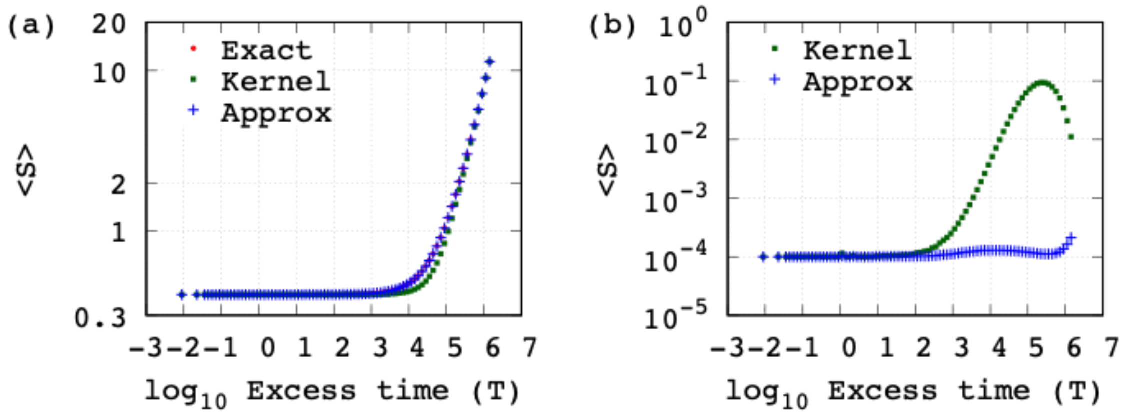


Figure 4. Numerically calculated average relative entropies (a) between inter-event time distribution ($p_M(\tau_M)$) and the inverse probability density functions, and (b) between the analytic inverse probability density function and the approximation functions, for each excess time (T) from former large event.

(2) APPLICATION OF BAYESIAN UPDATING TO TIME-SERIES OF ETAS MODEL

Application to ETAS time-series

For correlated time-series of the ETAS model, derivation of the inverse p.d.f. is difficult.

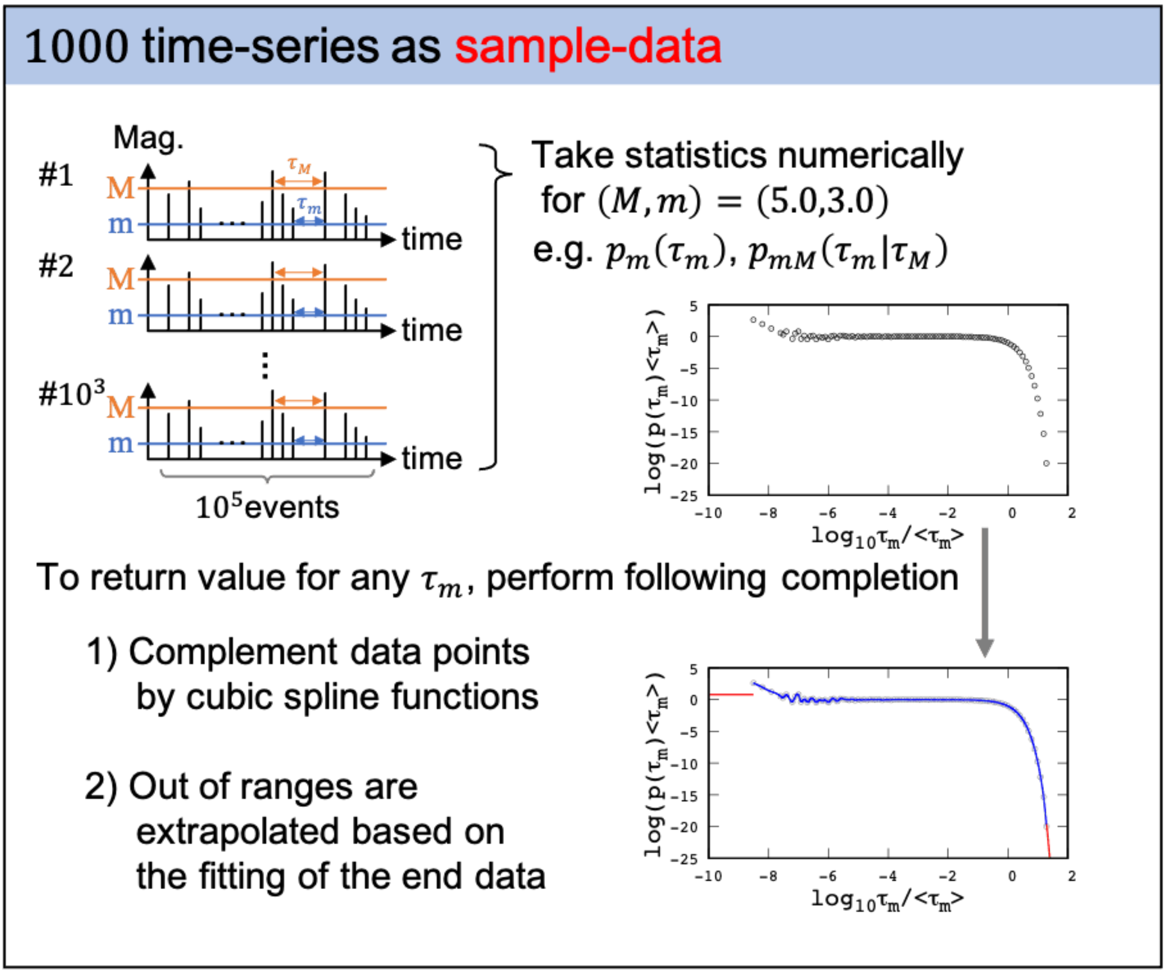
⇒ Apply Bayesian updating for uncorrelated time-series to the ETAS model, and calculate the approximation function of the inverse p.d.f. numerically.

Method

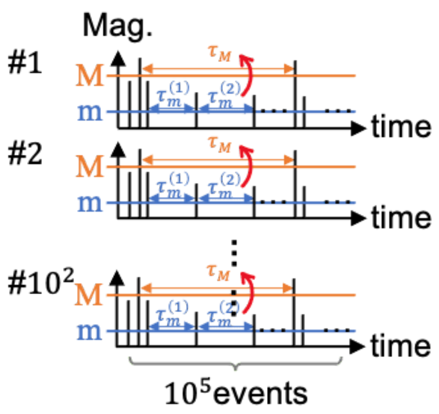
Generate 1100 ETAS time-series each of which contains 10^5 events numerically with parameters

$(b, \alpha, \lambda_0, K, p, c, m_0) = (1.0, 1.2, 0.0007, 0.0125, 1.2, 0.01, 0.3)$.

→ use 1000 as sample-data, 100 as test-data.



100 time-series as test-data



Conduct Bayesian updating using statistical amounts obtained from test-data

In particular, examine the kernel part p^{kernel}_{Mm}

Prior distribution = inter-event time distribution

$$\ln p^{\mathrm{kernel}}_{Mm}\left(\tau^k_M\middle|\tau^{(1)}_m\right) := \ln\left[\frac{\langle\tau_m\rangle\,\tau^k_M}{\langle\tau_M\rangle\langle\langle\tau_m\rangle\rangle_{\tau_M}}\right] + \ln p_{mM}\left(\tau^{(1)}_m\middle|\tau^k_M\right) - \ln p_m\left(\tau^{(1)}_m\right) + \ln p^k_M$$

$$\ln p^{\mathrm{kernel}}_{Mm}\left(\tau^k_M\middle|\tau^{(1)}_m,\tau^{(2)}_m\right) := -\ln\left(1-\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right) + \ln p_{mM}\left(\tau^{(2)}_m\middle|\tau^k_M\right) - \ln p_m\left(\tau^{(2)}_m\right) + \ln p^{\mathrm{kernel}}_{Mm}\left(\tau^k_M\middle|\tau^{(1)}_m\right)$$

$$\ln p^{\mathrm{kernel}}_{Mm}\left(\tau^k_M\middle|\tau^{(1)}_m,\tau^{(2)}_m,\tau^{(3)}_m\right) := -\ln\left(1-\frac{\langle\tau_m\rangle}{\langle\tau_M\rangle}\right) + \ln p_{mM}\left(\tau^{(3)}_m\middle|\tau^k_M\right) - \ln p_m\left(\tau^{(3)}_m\right) + \ln p^{\mathrm{kernel}}_{Mm}\left(\tau^k_M\middle|\tau^{(1)}_m,\tau^{(2)}_m\right)$$

$$\vdots$$

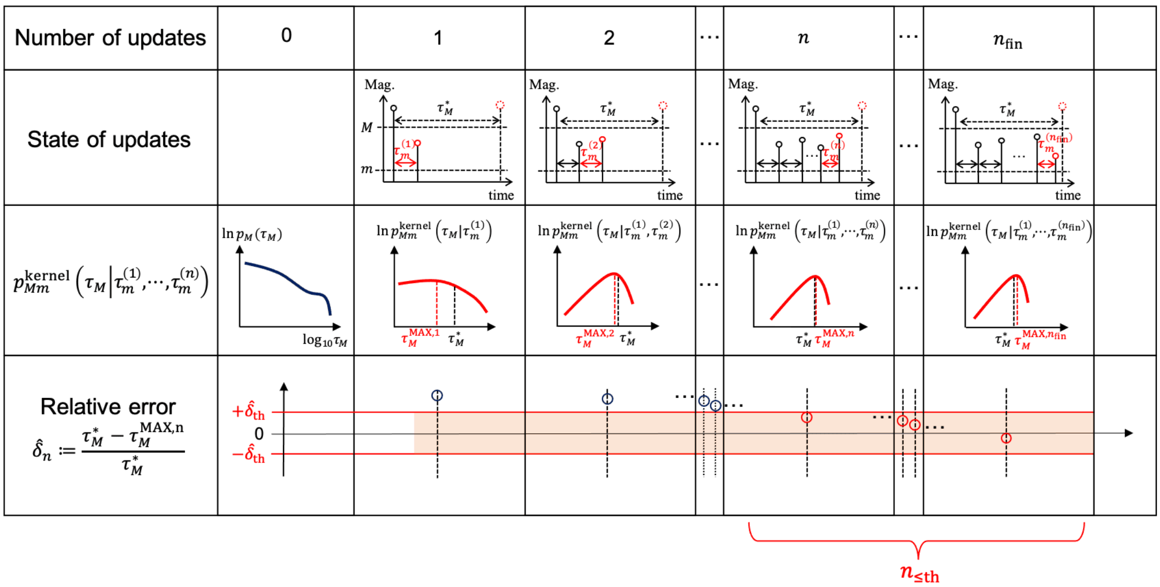
Bayesian inference

- Estimate the occurrence time of forthcoming large event by the mode ($\tau_M^{\mathrm{MAX,n}}$) of $p^{\mathrm{kernel}}_{Mm}(\tau_M|\tau_m^{(1)},\cdots,\tau_m^{(n)})$✻

- Evaluate the quality of estimation at n-th update by relative error $\hat{\delta}_n$ with the actual occurrence time of large event τ_M^*.

⇒ If $|\hat{\delta}_n| \leq \hat{\delta}_{\mathrm{th}} = 0.5$, estimation at n-th update is judged to be good.

Forecasting effectiveness is judged by the number ($n_{\leq \mathrm{th}}$) of times $|\hat{\delta}_n| \leq \hat{\delta}_{\mathrm{th}}$ is continued.



✻ Correction term is shown numerically not to affect the mode of the approximation function significantly, though is effective to the variance.

Examples

(1) Effective forecasting (**Fig. 5**)

- time-series is stationary
- estimate $\tau_M^{\mathrm{MAX,n}}$ stays around τ_M^*
- $|\hat{\delta}_n| \leq \hat{\delta}_{\mathrm{th}}$ holds for long update times

(2) Ineffective forecasting (**Fig.6**)

- time-series is non-stationary
- estimate $\tau_M^{\mathrm{MAX,n}}$ keeps shifting
- $|\hat{\delta}_n| \leq \hat{\delta}_{\mathrm{th}}$ holds only before large event

⇒ These examples suggest that the stationarity of time-series is related to the stability of the estimate, and the effectiveness of the forecasting.

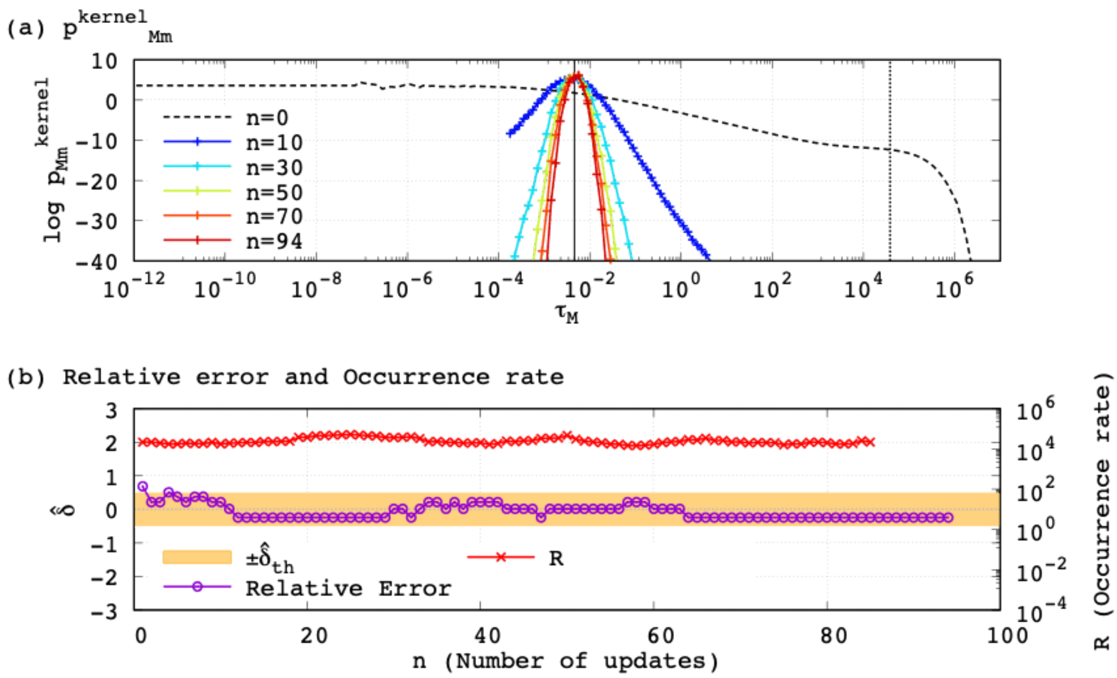


Figure 5. (a) An example of time evolution of p_{Mm}^{kernel}. Vertical line is the actual occurrence time (τ_M^*) of large event, and vertical dotted line is the average interval $\langle \tau M \rangle$. (b) Corresponding time-evolution of relative error $\hat{\delta}_n$ and occurrence rate R.

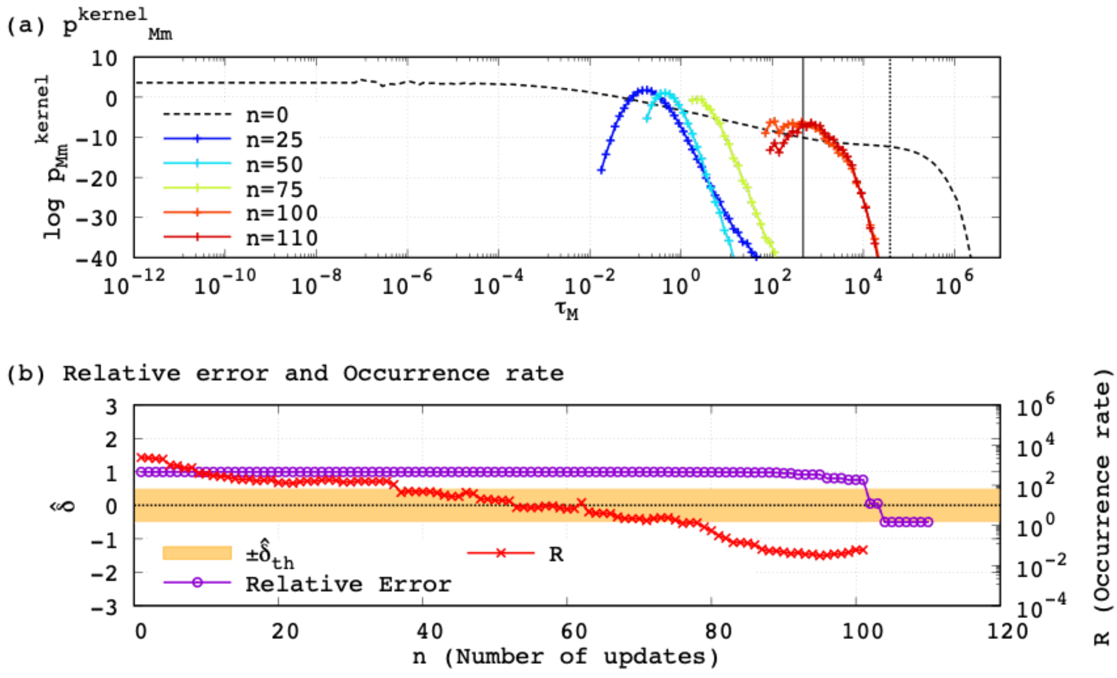


Figure 6. Another example of (a) p^{kernel}_{Mm}, (b) $\hat{\delta}_n$ and R.

Statistical analysis

Statistical analysis on forecasting effectivity (Figs. 7 and 8) suggests that

Stationarity of time-series

→ Stability of estimate

→ Longer $n_{\le \mathrm{th}}$ ⇔ Effective forecasting

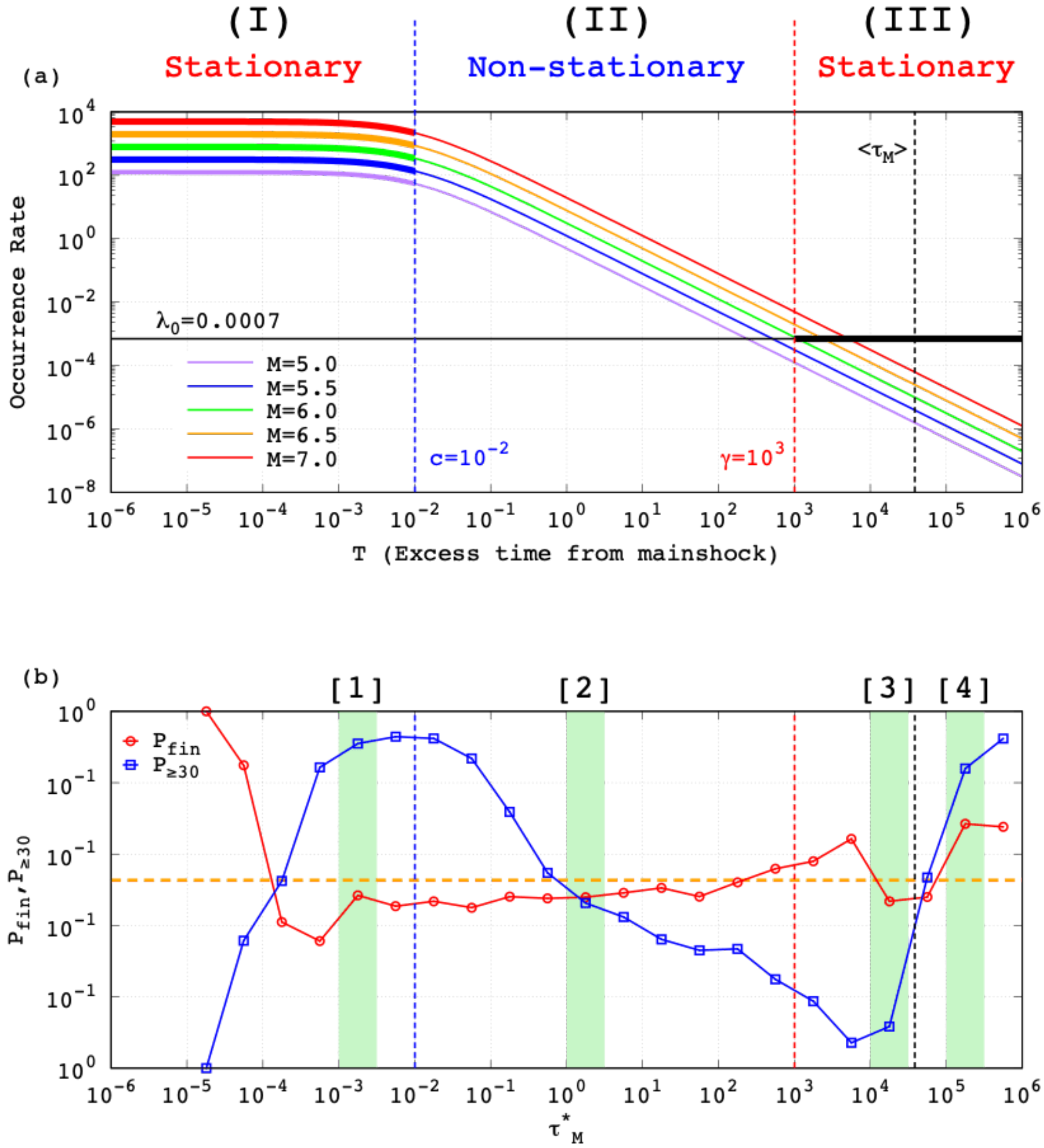


Figure 7. (a) Omori-Utsu law in this ETAS model. Time from the mainshock (T) can be categorized into three regions by stationarity. (b) Probability (P_{fin}) that $|\hat{\delta}_{\text{nfin}}| \le \hat{\delta}_{\text{th}}$ ($n_{\le\text{th}} > 0$) is around 0.53, especially takes around 0.7 for $\tau_M^* > \langle\tau_M\rangle$. Also, percentage ($P_{\ge 30}$) of those satisfies $n_{\le\text{th}} > 0$ that specifically $n_{\le\text{th}} \ge 30$ is high when stationarity is dominant in time-series.

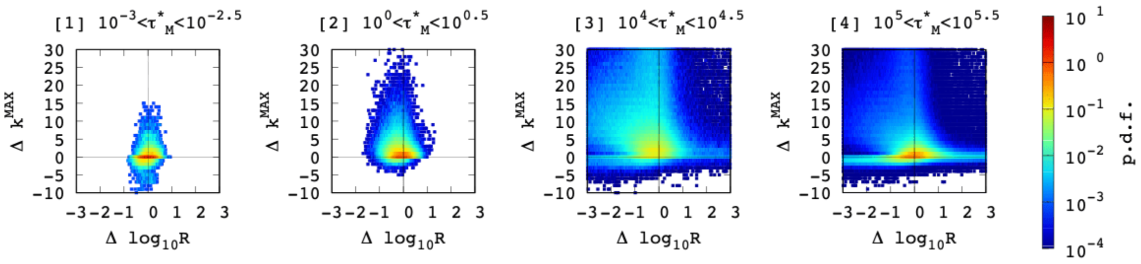


Figure 8. Joint probability density function of log-Occurrence rate ($\Delta \log_{10} R$) and log-Estimate (Δk^{MAX} where $\tau_M^{\text{MAX}} = 10^{(k^{\text{MAX}}+0.5)0.1}$), for some τ_M^*.

SUMMARY

(1) Bayes' theorem is considered for marked point process, and the inverse probability density function is derived for uncorrelated time-series.

(2) Bayesian updating that gives the relation between multiple successive lower intervals and an upper interval is considered for uncorrelated time-series. Inverse probability density function and its approximation function are derived.

(3) Bayesian updating is adopted to the time-series of the ETAS model. Mode of the kernel-part of the approximation function is used as the estimate of forthcoming large event. Goodness of estimation is evaluated by relative error with the actual occurrence time of the large event, and effectiveness of forecasting is examined by the length of successive such good estimations. As a result, it is suggested that the stability of the estimate (=effectiveness of forecasting) is related to the stationarity of the time-series. In the ETAS model we used, the forecasting is comparably effective immediately or long after the former large event, because in these two cases the time-series is dominated by stationarity. This suggests the possibility of using the method in the long-term or immediately after the mainshock.

REFERENCES

[1] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).

[2] A. Corral, Phys. Rev. Lett. **92**, 108501 (2004).

[3] L. de Arcangelis, C. Godano, J. R. Grasso, E. Lippiello, Phys. Rep. **628**, 1-91 (2016).

[4] Y. Ogata, J. Am. stat. Assoc. **83**, 9 (1988).

[5] Y. Ogata, Ann. Inst. Statist. Math. **50**, 379– 402 (1998).

[6] Y. Aizawa, T. Hasumi, and S. Tsugawa, Nonlinear Phenom. Complex Syst. **16**, 116 (2013).

[7] H. Tanaka and Y. Aizawa, J. Phys. Soc. Jpn **86**, 024004 (2017).

ABSTRACT

In this presentation, we introduce a Bayesian updating method for inter-event times of different magnitude thresholds in marked point process, apply it to time-series of the ETAS model [1], and discuss the effectiveness in probabilistic forecasting of forthcoming large event considering the information on smaller events.

To investigate magnitude threshold dependence of the inter-event time distribution of earthquakes, the conditional probability between inter-event times of different magnitude thresholds is proposed [2]. This gives the one-to-one statistical relationship between inter-event times of different magnitude thresholds.

Firstly, we show the Bayes' theorem on this conditional probability and derive the representation of the inverse probability density function.

Secondly, we extend it to the Bayesian updating that gives the relationship between multiple intervals for lower threshold and an interval for upper threshold. We show the derivation of the inverse probability density function and its approximation function for uncorrelated marked point process (background seismicity in the ETAS model). The condition for the inverse probability density function to have a peak is also shown. The approximation function consists of two parts, a kernel-part that determines its outline and a correction term. The former has an easy form to handle numerically and is applicable to the time-series with correlations among events.

Thirdly, based on the results for uncorrelated time-series, we apply the Bayesian updating method to time-series of the ETAS model. The mode of the approximation function is numerically shown to be nearly the same as that of the kernel-part. Therefore, the mode of the kernel part is used as the estimate of the occurrence time of forthcoming large event. By using the relative error between the estimate and the actual occurrence time of large event, effectiveness of the estimation with the approximation function is statistically evaluated. As a result, it is shown that if the time-series is dominated by stationary part, immediately or long after the large event, the forecasting is effectively conducted.

[1] Y. Ogata, Ann. Inst. Statist. Math. 50(2), 379 (1998).

[2] H. Tanaka and Y. Aizawa, J. Phys. Soc. Jpn. 86, 024004 (2017).

REFERENCES

[1] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. 88, 178501 (2002).

[2] A. Corral, Phys. Rev. Lett. 92, 108501 (2004).

[3] L. de Arcangelis, C. Godano, J. R. Grasso, E. Lippiello, Phys. Rep. 628, 1-91 (2016).

[4] Y. Ogata, J. Am. stat. Assoc. 83, 9 (1988).

[5] Y. Ogata, Ann. Inst. Statist. Math. 50, 379– 402 (1998).

[6] Y. Aizawa, T. Hasumi, and S. Tsugawa, Nonlinear Phenom. Complex Syst. 16, 116 (2013).

[7] H. Tanaka and Y. Aizawa, J. Phys. Soc. Jpn 86, 024004 (2017).