Journal of the Physical Society of Japan **92**, 113001 (2023) **Letters**
https://doi.org/10.7566/JPSJ.92.113001

A Bayesian Inference Method for a Large Magnitude Event in a Spatiotemporal Marked Point Process Representing Seismic Activity

Hiroki Tanaka* and Ken Umeno†

Graduate School of Informatics, Kyoto University, Sakyo, Kyoto 606-8501, Japan

(Received March 18, 2023; accepted September 5, 2023; published online October 3, 2023)

A Bayesian method to forecast the occurrence time of a large-scale earthquake utilizing temporal information on earthquakes with smaller magnitudes was proposed in our recent study for a marked point process that simulates seismic activity. In this paper, we show the extension of this Bayesian approach in the spatiotemporal marked point process, aiming to yield a forecasting method for both the occurrence time and location of the next large earthquake. We particularly discuss the contribution of the correlations between the spatial position and the inter-event time interval at different magnitude scales to probabilistic forecasting.

Probabilistic evaluation of future large-scale earthquake occurrence is one of the important issues for disaster prevention. One approach is to model seismic activity as a nonstationary Poisson process and evaluate the risk using the conditional intensity function.[1–3] The Epidemic Type After-shock Sequence (ETAS) model[2,4–6] gives one such prominent modeling; its model parameter values are estimated from past seismic data including information about small-scale earthquakes.[2,4–7] The conditional intensity function with those determined parameter values allows us to quantitatively evaluate the risk of earthquake occurrence at a certain time, spatial position, and magnitude[1–5,8] and to recognize deviating seismic activity.[6] The conditional intensity function is mostly expressed as the product of two functions; a function that depends only on time and spatial position, and another function only on magnitude.[1–3,8] This is based on the assumption that a seismicity pattern does not affect the spatiotemporal occurrence of its subsequent large-size earthquake. Considering correlations between a significant shock and seismicity ahead of it[6,8–11] is important for improving forecasting.[8]

Another approach without such an assumption and targeting large earthquakes is to evaluate the timing of a major earthquake using the inter-event time distribution at a high magnitude threshold; the risk of large earthquake occurrence timing is evaluated by hazard function based on the inter-event time distribution obtained from seismic history.[1,3,12] Although this approach is a basis for actual forecasting,[13] to the authors' knowledge, this method has not been able to utilize the information on smaller earthquakes than the threshold magnitude, which is used in the first approach.

Thus, we previously suggested a new Bayesian approach for forecasting the occurrence time of a significant earthquake, which improves the second approach by utilizing the information on such small-scale earthquakes.[7] This method focuses on the relationship between the inter-event time intervals (hereafter referred to simply as time intervals) in the point processes at two threshold magnitudes set in a marked point process that represents seismic activity [Fig. 1(a)].[7,14] Bayes' theorem has been considered for such time intervals;[7] its extension, Bayesian updating, has also been considered.[7] Bayesian updating provides a method to forecast the timing of the next event whose magnitude exceeds the upper threshold by utilizing the information of time intervals at the lower magnitude threshold.[7] Thus, this extended method provides a framework for forecasting that can take into account the effect of temporal patterns on the subsequent major event.[7] In this paper, we further extend the scope of this Bayesian approach to the spatiotemporal marked point process [Figs. 1(a) and 1(b)] to forecast not only the occurrence time but the spatial position of a significant future event incorporating the spatial pattern of epicenters. The main objective of this study is to propose a mathematical framework for the Bayesian approach in spatiotemporal marked point processes, which yields an alternative Bayesian method to the preceding study using the Bayesian network.[15] Based on the theory, we discuss possible contributions of spatiotemporal interactions for forecasting.[16]

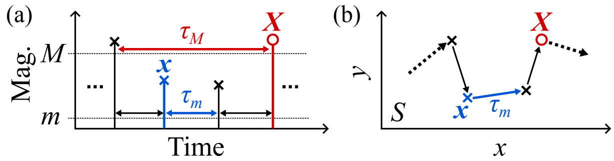

Fig. 1. (Color online) Schematic of a spatiotemporal seismic activity represented by (a) a marked point process with a magnitude as a mark, and (b) corresponding jumps of events in the spatial area S. The upper and lower spatiotemporal pairs $\{\boldsymbol{X}, \tau_M\}$ and $\{\boldsymbol{x}, \tau_m\}$ are shown, and these form the combination $\{\boldsymbol{x}, \tau_m; \boldsymbol{X}, \tau_M\}$.

First, the conditional probability for the spatiotemporal marked point process is defined (Fig. 1). Let τ_M (τ_m) represent the length of an inter-event time interval at upper magnitude threshold M [lower magnitude threshold m ($< M, M := m + \Delta m$)]. Further, let $\boldsymbol{X}(\in S)$ ($\boldsymbol{x}(\in S)$) represent the spatial position of the event at the right (left) end of an upper (lower) inter-event time interval (Fig. 1). Here S is a closed region on the earth's surface. Although there is no restriction on its size, we assume that S encompasses seismogenic zones as centrally as possible; if a seismogenic zone lies on the edge of S, only a limited portion of aftershocks following a mainshock at $\boldsymbol{X}$ can be considered, potentially impeding precise probabilistic evaluation within this theory. Hereafter, the pair of the above-defined spatial position and time interval is referred to as the (spatiotemporal) pair; $\{\boldsymbol{X}, \tau_M\}$ is referred to as the upper (spatiotemporal) pair, and $\{\boldsymbol{x}, \tau_m\}$ the lower (spatiotemporal) pair. When the upper pair $\{\boldsymbol{X}, \tau_M\}$ includes the lower pair $\{\boldsymbol{x}, \tau_m\}$ as shown in Fig. 1(a), these constitute the combination of the upper and lower spatiotemporal pairs $\{\boldsymbol{x}, \tau_m; \boldsymbol{X}, \tau_M\}$. The magnitude of the event at $\boldsymbol{x}$ is exceptionally greater than M when the lower pair is located at the leftmost in an upper pair, though otherwise, it is always less than or equal to M. Thus, $p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)$ represents the spatiotemporal conditional probability density that a lower spatiotemporal pair is $\{\boldsymbol{x},\tau_m\}$ under the condition that it is included in the upper spatiotemporal pair $\{\boldsymbol{X},\tau_M\}$.

We consider Bayes' theorem for the above-defined spatiotemporal conditional probability. Let $dN_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)$ represent the number of the combinations of the upper and lower spatiotemporal pairs such that the upper pair falls within $[\boldsymbol{X},\boldsymbol{X}+d\boldsymbol{X})$ and $[\tau_M,\tau_M+d\tau_M)$, and the lower pair included in it falls within $[\boldsymbol{x},\boldsymbol{x}+d\boldsymbol{x})$ and $[\tau_m,\tau_m+d\tau_m)$, in the time series. For simplicity, $dN_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)$ is referred to as the number of the combination $\{\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M\}$ without mentioning the infinitesimal intervals, and hereafter, other numbers of spatiotemporal pairs or combinations denoted by dN are referred to in the same way. $dN_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)$ can be expressed in two ways:

$$\begin{aligned} dN_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)/d\boldsymbol{x}\,d\tau_m\,d\boldsymbol{X}\,d\tau_M &= N_M p_M(\boldsymbol{X},\tau_M)\frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)\\ &= N_m p_m(\boldsymbol{x},\tau_m)p_{Mm}(\boldsymbol{X},\tau_M|\boldsymbol{x},\tau_m). \end{aligned}\quad (1)$$

Here, N_M (N_m) represents the total number of upper (lower) spatiotemporal pairs in the time series. $p_M(\boldsymbol{X},\tau_M)$ ($p_m(\boldsymbol{x},\tau_m)$) yields the joint probability density that the upper (lower) pair is $\{\boldsymbol{X},\tau_M\}$ ($\{\boldsymbol{x},\tau_m\}$). $p_{Mm}(\boldsymbol{X},\tau_M|\boldsymbol{x},\tau_m)$ yields the inverse probability density that the upper pair is $\{\boldsymbol{X},\tau_M\}$ under the condition that the lower pair $\{\boldsymbol{x},\tau_m\}$ is found within it. $\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}$ represents the average time interval of the spatiotemporal conditional probability:

$$\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M} := \int_S d\boldsymbol{x}\int_0^\infty d\tau_m\tau_m p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M).$$

From Eq. (1), Bayes' theorem is derived as:

$$p_{Mm}(\boldsymbol{X},\tau_M|\boldsymbol{x},\tau_m) = 10^{-b\Delta m}\frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}\frac{p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)}{p_m(\boldsymbol{x},\tau_m)}p_M(\boldsymbol{X},\tau_M),\quad (2)$$

where the GR law[17] ($N_M/N_m = 10^{-b\Delta m}$) is used.

The normalization condition of the inverse probability in Eq. (2) can be checked using the integral equation with the spatiotemporal conditional probability in its kernel, which is derived below. Let $dN_m(\boldsymbol{x},\tau_m)$ represent the number of the spatiotemporal pairs $\{\boldsymbol{x},\tau_m\}$ in the time series, then $dN_m(\boldsymbol{x},\tau_m)$ can be expressed in two ways:

$$\begin{aligned} dN_m(\boldsymbol{x},\tau_m)/d\boldsymbol{x}\,d\tau_m &= N_m p_m(\boldsymbol{x},\tau_m)\\ &= N_M\int_S d\boldsymbol{X}\int_0^\infty d\tau_M\frac{\tau_M p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}p_M(\boldsymbol{X},\tau_M). \end{aligned}$$

Thus, the integral equation that connects the probability density functions of the upper and lower pair is derived as:

$$p_m(\boldsymbol{x},\tau_m) = 10^{-b\Delta m}\times\int_S d\boldsymbol{X}\int_0^\infty d\tau_M\frac{\tau_M p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}p_M(\boldsymbol{X},\tau_M).\quad (3)$$

The inverse probability density function for only upper and lower time intervals or only upper and lower spatial positions can be derived from the joint probability for the combination of an upper and a lower spatiotemporal pair. Let $p_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)$ represent the joint probability density of the combination $\{\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M\}$. The total number of the combinations of an upper and a lower spatiotemporal pair in the time series is N_m, and therefore, $p_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)$ can be expressed in two ways using Eq. (1) as follows:

$$\begin{aligned} p_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M) &= \frac{dN_{mM}(\boldsymbol{x},\tau_m;\boldsymbol{X},\tau_M)}{N_m d\boldsymbol{x}\,d\tau_m\,d\boldsymbol{X}\,d\tau_M}\\ &= 10^{-b\Delta m}\frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M)p_M(\boldsymbol{X},\tau_M)\\ &= p_m(\boldsymbol{x},\tau_m)p_{Mm}(\boldsymbol{X},\tau_M|\boldsymbol{x},\tau_m). \end{aligned}\quad (4)$$

Thus, Bayes' theorem for time intervals already obtained in the previous study[7] can be derived by marginalizing the joint probability in Eq. (4) for $\boldsymbol{x}$ and $\boldsymbol{X}$ as:

$$p_{Mm}(\tau_M|\tau_m) = 10^{-b\Delta m}\frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\tau_M}}\frac{p_{mM}(\tau_m|\tau_M)}{p_m(\tau_m)}p_M(\tau_M).$$

Further, the integral equation for time intervals introduced in the previous study[7,14] can also be derived by marginalizing Eq. (3) for $\boldsymbol{x}$ and $\boldsymbol{X}$, noting that the integrand of the r.h.s. of Eq. (3) is the joint probability in Eq. (4), as:

$$p_m(\tau_m) = 10^{-b\Delta m}\int_0^\infty d\tau_M\frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\tau_M}}p_{mM}(\tau_m|\tau_M)p_M(\tau_M).$$

Bayes' theorem for spatial positions can also be derived. First, we define the following quantity:

$$\langle n(\boldsymbol{x},\tau_m)\rangle_{\boldsymbol{X},\tau_M} := \frac{\tau_M}{\langle\!\langle\tau_m\rangle\!\rangle_{\boldsymbol{X},\tau_M}}p_{mM}(\boldsymbol{x},\tau_m|\boldsymbol{X},\tau_M).\quad (5)$$

This is the average number of specific lower spatiotemporal pairs $\{\boldsymbol{x},\tau_m\}$ included in the upper pair $\{\boldsymbol{X},\tau_M\}$. Bayes' theorem for the spatial position is obtained by marginalizing the joint probability in Eq. (4) for τ_m and τ_M while using Eq. (5).

$$p_{Mm}(\boldsymbol{X}|\boldsymbol{x}) = 10^{-b\Delta m}\frac{\langle n(\boldsymbol{x})\rangle_{\boldsymbol{X}}}{p_m(\boldsymbol{x})}p_M(\boldsymbol{X}).\quad (6)$$

Here, $\langle n(\boldsymbol{x})\rangle_{\boldsymbol{X}}$ is the average number of such events with magnitude $\in(m,M]$ that occur at $\boldsymbol{x}$ and are in between two consecutive large events with magnitudes $> M$, the latter of which occurs at $\boldsymbol{X}$. The integral equation for the spatial position can be derived by marginalizing Eq. (3) for τ_m and τ_M using Eq. (5) and noting again that the integrand in Eq. (3) is the joint probability in Eq. (4), as follows.

$$p_m(\boldsymbol{x}) = 10^{-b\Delta m}\int_S d\boldsymbol{X}\langle n(\boldsymbol{x})\rangle_{\boldsymbol{X}}p_M(\boldsymbol{X}).\quad (7)$$

The normalization condition for the inverse probability in Eq. (6) can be checked using Eq. (7). In particular, consider the case where spatial positions and time intervals are independent. In this case, the average number of lower intervals in an upper interval is $10^{b\Delta m}$ according to the GR law, and therefore, $\langle n(\boldsymbol{x})\rangle_{\boldsymbol{X}} = 10^{b\Delta m}p_{mM}(\boldsymbol{x}|\boldsymbol{X})$. Thus, Eqs. (6) and (7) are simplified as:

$$p_{Mm}(\boldsymbol{X}|\boldsymbol{x}) = \frac{p_{mM}(\boldsymbol{x}|\boldsymbol{X})}{p_m(\boldsymbol{x})}p_M(\boldsymbol{X}),$$

$$p_m(\boldsymbol{x}) = \int_S d\boldsymbol{X}p_{mM}(\boldsymbol{x}|\boldsymbol{X})p_M(\boldsymbol{X}).$$

We extend Bayes' theorem to Bayesian updating. Hereafter, for simplicity, the sequence of consecutive lower spatiotemporal pairs $\{\boldsymbol{x}_1,\tau_m^{(1)},\dots,\boldsymbol{x}_n,\tau_m^{(n)}\}$, such that all the pairs are included in the same upper pair, is denoted by $\{\boldsymbol{x}_{1:n},\tau_m^{(1:n)}\}$ (Fig. 2). When this sequence is included in

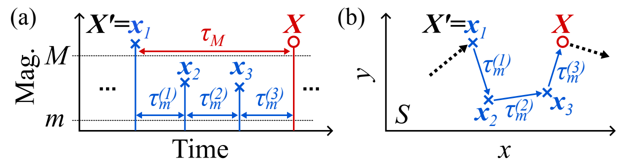

Fig. 2. (Color online) Schematic of spatiotemporal seismic activity showing time intervals and spatial positions of events considered in the Bayesian updating; the upper pair $\{X, \tau_M\}$ and the consecutive lower pairs $\{x_{1:n}, \tau_m^{(1:n)}\}$ ($n = 3$ in the figure) included in it are indicated. These form the combination of upper and lower pairs $\{x_{1:n}, \tau_m^{(1:n)}; X, \tau_M\}$.

the upper pair $\{X, \tau_M\}$, these constitute the combination of the upper and lower pairs $\{x_{1:n}, \tau_m^{(1:n)}; X, \tau_M\}$. Let $dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)$ represent the total number of such combination in the time series. Thus, $dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)$ can be expressed in two ways:

$$\begin{aligned} &dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)/dX\, d\tau_M\, dx^n\, d\tau_m^n \\ &= N_M p_M(X, \tau_M)\langle n(x_{1:n}, \tau_m^{(1:n)})\rangle_{X,\tau_M} \\ &= N_m R_n p_m(x_{1:n}, \tau_m^{(1:n)}) p_{Mm}(X, \tau_M | x_{1:n}, \tau_m^{(1:n)}). \end{aligned} \tag{8}$$

Here, R_n represents the proportion of the number of sequences of consecutive n lower pairs such that all these pairs are included in the same upper pair in the time series to the total number of sequences of consecutive n lower pairs in the time series. Thus, $N_m R_n$ represents the total number of consecutive n lower pairs belonging to the same upper pair and the total number of combinations of an upper and consecutive n lower pairs in a time series. For example, in the time series of the background seismicity generated with a constant occurrence rate in the ETAS model, $R_n = (1 - \langle\tau_m\rangle/\langle\tau_M\rangle)^{n-1}$ by the GR law.[7] Further, $p_m(x_{1:n}, \tau_m^{(1:n)})$ represents the joint probability density that a sequence of consecutive n lower pairs belonging to the same upper interval takes $\{x_{1:n}, \tau_m^{(1:n)}\}$. $p_{Mm}(X, \tau_M | x_{1:n}, \tau_m^{(1:n)})$ represents the inverse probability density that the upper pair is $\{X, \tau_M\}$ when $\{x_{1:n}, \tau_m^{(1:n)}\}$ is found within it. $\langle n(x_{1:n}, \tau_m^{(1:n)})\rangle_{X,\tau_M}$ represents the average number of the sequence $\{x_{1:n}, \tau_m^{(1:n)}\}$ in the upper pair $\{X, \tau_M\}$. From Eq. (8), the inverse probability density is:

$$\begin{aligned} &p_{Mm}(X, \tau_M | x_{1:n}, \tau_m^{(1:n)}) \\ &= 10^{-b\Delta m} R_n^{-1} \frac{\langle n(x_{1:n}, \tau_m^{(1:n)})\rangle_{X,\tau_M}}{p_m(x_{1:n}, \tau_m^{(1:n)})} p_M(X, \tau_M). \end{aligned} \tag{9}$$

We derive the inverse probability density for only time intervals or only spatial positions. Let $p_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)$ represent the joint probability density that the combination of an upper and consecutive lower spatiotemporal pairs is $\{x_{1:n}, \tau_m^{(1:n)}; X, \tau_M\}$, which can be expressed as:

$$\begin{aligned} p_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M) &= \frac{dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)}{N_m R_n dX\, d\tau_M\, dx^n\, d\tau_m^n} \\ &= 10^{-b\Delta m} R_n^{-1} \langle n(x_{1:n}, \tau_m^{(1:n)})\rangle_{X,\tau_M} p_M(X, \tau_M) \\ &= p_m(x_{1:n}, \tau_m^{(1:n)}) p_{Mm}(X, \tau_M | x_{1:n}, \tau_m^{(1:n)}). \end{aligned} \tag{10}$$

Thus, the inverse probability density function for time intervals is derived by marginalizing Eq. (10) for $X, x_1, \ldots, x_n$,

$$p_{Mm}(\tau_M | \tau_m^{(1:n)}) = 10^{-b\Delta m} R_n^{-1} \frac{\langle n(\tau_m^{(1:n)})\rangle_{\tau_M}}{p_m(\tau_m^{(1:n)})} p_M(\tau_M). \tag{11}$$

This is consistent with the result obtained in the previous study.[7] On the other hand, the inverse probability density for spatial positions can be derived by marginalizing Eq. (10) for $\tau_M, \tau_m^{(1)}, \ldots, \tau_m^{(n)}$,

$$p_{Mm}(X | x_{1:n}) = 10^{-b\Delta m} R_n^{-1} \frac{\langle n(x_{1:n})\rangle_X}{p_m(x_{1:n})} p_M(X). \tag{12}$$

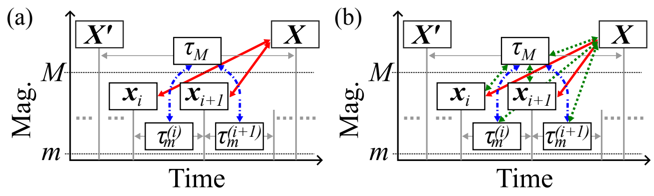

Fig. 3. (Color online) Two cases of correlations assumed between time intervals and spatial positions. Colored arrows represent correlations between (red, solid) spatial positions, (blue, dot-dash) time intervals, and (green, dotted) a time interval and a spatial position, respectively.

Finally, we examine how spatiotemporal correlations appear in Bayesian updating by comparing two cases of correlations among spatial positions and time intervals as shown in Fig. 3. Variables connected by an arrow in Fig. 3 are assumed to be correlated (meaning neither independent nor conditionally independent given other variables), and variables not connected are assumed to be not only independent, but conditionally independent given other variables. However, the conditional independence between the lower time intervals ($\tau_m^{(i)}$'s) in the same upper time interval is not assumed given the length of their upper time interval τ_M. In the first case, space-time correlations are not assumed [Fig. 3(a)], whereas in the second case, correlations between spatial position and time interval indicated with the green dotted arrows in Fig. 3(b) are added. For the first case, Eq. (8) is rewritten as:[18]

$$\begin{aligned} &dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)/dX\, d\tau_M\, dx^n\, d\tau_m^n \\ &= N_M p_M(X) p_M(\tau_M) \langle n(\tau_m^{(1:n)})\rangle_{\tau_M} \prod_{i=1}^{n} p_{mM}(x_i | X) \\ &= N_m R_n p_{Mm}(X | x_{1:n}) p_{Mm}(\tau_M | \tau_m^{(1:n)}) \prod_{i=1}^{n} p_m(x_i) p_m(\tau_m^{(i)}). \end{aligned} \tag{13}$$

Therefore, the Bayesian updating is expressed as the product of the following two respective updates for the time interval and the spatial position.

$$p_{Mm}(\tau_M | \tau_m^{(1:n)}) = 10^{-b\Delta m} R_n^{-1} \frac{\langle n(\tau_m^{(1:n)})\rangle_{\tau_M}}{\prod_{i=1}^{n} p_m(\tau_m^{(i)})} p_M(\tau_M), \tag{14}$$

$$p_{Mm}(X | x_{1:n}) = \frac{\prod_{i=1}^{n} p_{mM}(x_i | X)}{\prod_{i=1}^{n} p_m(x_i)} p_M(X). \tag{15}$$

For the second case, Eq. (8) is rewritten as:[18]

$$\begin{aligned} &dN_{mM}(x_{1:n}, \tau_m^{(1:n)}; X, \tau_M)/dX\, d\tau_M\, dx^n\, d\tau_m^n \\ &= N_M p_M(X, \tau_M) \langle n(\tau_m^{(1:n)})\rangle_{X,\tau_M} \prod_{i=1}^{n} p_{mM}(x_i | X, \tau_M) \\ &= N_m R_n p_{Mm}(X, \tau_M | x_{1:n}, \tau_m^{(1:n)}) \prod_{i=1}^{n} p_m(x_i) p_m(\tau_m^{(i)}). \end{aligned} \tag{16}$$

Thus, the inverse probability is:

$$p_{Mm}(\boldsymbol{X}, \tau_M | \boldsymbol{x}_{1:n}, \tau_m^{(1:n)}) = 10^{-b\Delta m} R_n^{-1} \times \frac{\langle n(\tau_m^{(1:n)}) \rangle_{X,\tau_M}}{\prod_{i=1}^{n} p_m(\tau_m^{(i)})} \frac{\prod_{i=1}^{n} p_{mM}(\boldsymbol{x}_i | \boldsymbol{X}, \tau_M)}{\prod_{i=1}^{n} p_m(\boldsymbol{x}_i)} p_M(\boldsymbol{X}, \tau_M).$$

The joint probability is obtained by taking the ratio of Eq. (16) to $N_m R_n$, in the same way as Eq. (10).

$$\begin{aligned} p_{mM}(\boldsymbol{x}_{1:n}, \tau_m^{(1:n)}; \boldsymbol{X}, \tau_M) &= 10^{-b\Delta m} R_n^{-1} \\ &\times \langle n(\tau_m^{(1:n)}) \rangle_{X,\tau_M} p_M(\boldsymbol{X}, \tau_M) \prod_{i=1}^{n} p_{mM}(\boldsymbol{x}_i | \boldsymbol{X}, \tau_M) \\ &= p_{Mm}(\boldsymbol{X}, \tau_M | \boldsymbol{x}_{1:n}, \tau_m^{(1:n)}) \prod_{i=1}^{n} p_m(\boldsymbol{x}_i) p_m(\tau_m^{(i)}). \end{aligned} \tag{17}$$

Marginalizing Eq. (17) with respect to $\boldsymbol{X}$ or τ_M, the inverse probability density functions for τ_M and $\boldsymbol{X}$ are obtained as follows.

$$p_{Mm}(\tau_M | \boldsymbol{x}_{1:n}, \tau_m^{(1:n)}) = 10^{-b\Delta m} R_n^{-1} \times \frac{\langle n(\tau_m^{(1:n)}) \rangle_{\tau_M}}{\prod_{i=1}^{n} p_m(\tau_m^{(i)})} \frac{\prod_{i=1}^{n} p_{mM}(\boldsymbol{x}_i | \tau_M)}{\prod_{i=1}^{n} p_m(\boldsymbol{x}_i)} p_M(\tau_M), \tag{18}$$

$$p_{Mm}(\boldsymbol{X} | \boldsymbol{x}_{1:n}, \tau_m^{(1:n)}) = 10^{-b\Delta m} R_n^{-1} \times \frac{\langle n(\tau_m^{(1:n)}) \rangle_{X}}{\prod_{i=1}^{n} p_m(\tau_m^{(i)})} \frac{\prod_{i=1}^{n} p_{mM}(\boldsymbol{x}_i | \boldsymbol{X})}{\prod_{i=1}^{n} p_m(\boldsymbol{x}_i)} p_M(\boldsymbol{X}). \tag{19}$$

Equations (18) and (19) show that the terms representing the correlations between spatial position and time interval are multiplied explicitly to the closed updates for time and space in Eqs. (14) and (15), respectively.

In summary, we proposed a Bayesian inference method of the occurrence time and location of the next significant event in a spatiotemporal marked point process using occurrence patterns of smaller events. In seismic activity, temporal quiescence, spatial gap, or activation (foreshocks) is sometimes recognized to precede major earthquakes.[6,19] The question of whether the conditional probability can quantitatively treat such qualitative spatiotemporal characteristics of seismic activity and whether the Bayesian updating method can be used for better probabilistic forecasting in actual seismic activity is one for the future.

This study only presents a theoretical framework. To verify the effectiveness of this framework for forecasting, numerical examination, as performed in the previous study,[7] with a sufficient number of synthetic seismic data that stochastic models (e.g., the hierarchical space-time ETAS model[5]) or physical models (e.g., the Olami–Feder–Christensen model[20]) can generate is necessary. Furthermore, for forecasting actual earthquakes, it is required to examine the framework with seismic catalog data. The preliminary analysis[16] using a seismic catalog in Southern California[21] does not show apparent improvement in forecasting by adding spatial information ($\boldsymbol{x}$) to the inverse probability. The cause of this seems to be in setting the spatial domain S and the way to subdivide it. Improving this point is important in future detailed analysis.

In this paper, the spatiotemporal pair is defined in a way that builds upon previous work,[7] and Bayesian updating is considered based on it. However, this definition includes exceptional events whose magnitude exceeds M in the lower pair. Further, the information of the lower pair immediately before the next large event is not used in real-time Bayesian updating.[7] Therefore, in the practical use of the Bayesian approach, it may be better to define the lower pair as a lower time interval and the spatial position of its subsequent event whose magnitude $\leq M$, in the same way as the upper pair. Bayesian updating based on this definition should be considered excluding the rightmost lower pair in an upper pair.

Acknowledgment This work was supported by JST SPRING, Grant Number JPMJSP2110. Data sources for the seismic catalog were provided by the Caltech/USGS Southern California Seismic Network (SCSN) DOI: 10.7914/SN/CI, and the Southern California Earthquake Data Center (SCEDC) DOI: 10.7909/C3WD3xH1.

*tanaka.hiroki.43s@st.kyoto-u.ac.jp
†umeno.ken.8z@kyoto-u.ac.jp

1) D. Vere-Jones, Int. J. Forecasting **11**, 503 (1995).
2) Y. Ogata, Tokei Suri **63**, 3 (2015) [in Japanese].
3) T. Omi and S. Nomura, *Tenkatei no Jikeiretsukaiseki* (Time Series Analysis for Point Processes) (Kyoritsu, Tokyo, 2021) [in Japanese].
4) Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988); Y. Ogata, Ann. Inst. Stat. Math. **50**, 379 (1998); Y. Ogata and J. Zhuang, Tectonophysics **413**, 13 (2006).
5) Y. Ogata, K. Katsura, and M. Tanemura, J. R. Stat. Soc. Ser. C **52**, 499 (2003); Y. Ogata, Earth Planets Space **74**, 110 (2022).
6) Y. Ogata, Tectonophysics **169**, 159 (1989).
7) H. Tanaka and K. Umeno, submitted to J. Phys. Soc. Jpn.; ESS Open Archive, December 16 (2023) [DOI: 10.22541/essoar.167340710.08314293/v1].
8) Y. Ogata, K. Katsura, H. Tsuruoka, and N. Hirata, Seismol. Res. Lett. **89**, 1298 (2018).
9) D. A. Rhoades and F. F. Evison, Pure Appl. Geophys. **161**, 47 (2004).
10) C. Smyth and J. Mori, Earth Planets Space **63**, 231 (2011).
11) E. Lippiello, L. de Arcangelis, and C. Godano, Phys. Rev. Lett. **100**, 038501 (2008); E. Lippiello, C. Godano, and L. de Arcangelis, Geophys. Res. Lett. **39**, L05309 (2012).
12) C. H. Scholz, *The Mechanics of Earthquakes and Faulting* (Cambridge University Press, Cambridge, U.K., 2002) 2nd ed., pp. 408–412.
13) Web page of the Headquarters for Earthquake Research Promotion [https://www.jishin.go.jp/reports/research_report/choukihyoka_01b/] [in Japanese], last access: August 2023.
14) H. Tanaka and Y. Aizawa, J. Phys. Soc. Jpn. **86**, 024004 (2017).
15) Y. Zhang, H. Zhao, X. He, F.-D. Pei, and G.-G. Li, Physica A **445**, 138 (2016).
16) (Supplemental Material) A preliminary analysis for Southern California showing how the mathematical framework of this study can be analyzed in actual seismic catalogs is provided online.
17) B. Gutenberg and C. F. Richter, Bull. Seismol. Soc. Am. **34**, 185 (1944).
18) (Supplemental Material) Derivation of Eqs. (13) and (16) is provided online.
19) T. Utsu, *Jishingaku* (Seismology) (Kyoritsu, Tokyo, 2012) pp. 199–202, pp. 217–218, pp. 322–325 [in Japanese]; C. H. Scholz, *The Mechanics of Earthquakes and Faulting* (Cambridge University Press, Cambridge, U.K., 2019) 3rd ed., pp. 343–350.
20) Z. Olami, H. J. S. Feder, and K. Christensen, Phys. Rev. Lett. **68**, 1244 (1992).
21) E. Hauksson, W. Yang, and P. M. Shearer, Bull. Seismol. Soc. Am. **102**, 2239 (2012); Caltech/USGS Southern California Seismic Network (SCSN) [DOI: 10.7914/SN/CI]; Southern California Earthquake Data Center (SCEDC) [DOI: 10.7909/C3WD3xH1]; [http://scedc.caltech.edu/research-tools/alt-2011-dd-hauksson-yang-shearer.html], last access: June 2023.