Geoscience Letters

RESEARCH LETTER

Open Access

Precursory seismic quiescence of major earthquakes along the Sagaing fault zone, central Myanmar: application of the pattern informatics technique

Premwadee Traitangwong[1], Sutthikan Khamsiri[1] and Santi Pailoplee[1*]

Abstract

In this study, the precursory seismic activity before a major earthquake was investigated by using the Pattern Informatics (PI) algorithm along the Sagaing fault zone (SFZ), Central Myanmar. After improving the earthquake catalog, the completeness of seismicity data with $M_w \geq 3.6$ reported during 1980–2020 was used in retrospective testing to find the suitable parameters of the PI algorithm. According to the retrospective test with 6 cases of different forecast period times related to $M_w \geq 5.0$ earthquakes, including verification using the relative operating characteristics (ROC) diagram, the characteristic parameters of both time intervals (change time and forecast time window) = 10 years and target forecast earthquake magnitude $M_w \geq 5.0$ are suitable parameters for PI investigation along the SFZ. Therefore, these parameters were applied with the most up-to-date seismic dataset to evaluate the prospective areas of upcoming major earthquakes. The results reveal that the Myitkyina and the vicinity of Naypyidaw might be at risk of a major earthquake in the future. Therefore, effective earthquake mitigation plans should be urgently arranged.

Keywords Pattern informatics, Seismicity, Seismic quiescence, Sagaing fault zone, Myanmar

Introduction

The Sagaing fault zone (SFZ) is one of the most seismically active faults in Mainland Southeast Asia, trending north–south direction in the central part of Myanmar (Swe 1981). Tectonically, the SFZ is the dextral strike-slip active fault boundary between the Sunda and Burma plates (Curray 2005), which moves at a slip rate of around 23 mm/year. (Bertrand and Rangin 2003) (Fig. 1). Based on the instrumental earthquake records, the SFZ generated at least eight major earthquake events with $M_w \geq 7.0$ between 1906 and 1991 (Kundu and Gahalaut 2012), which caused severe damage around the earthquake sources not only in Myanmar cities such as Myitkyina, Mandalay, Taunggyi, Naypyidaw, Bago, and Rangoon but also in Thailand, especially in the western, northern and central parts, as indicated by the isoseismal maps showing the earthquake intensities level which Myanmar and Thailand experience as level IX and V of the Modified Mercalli intensity scale, respectively (Pailoplee 2012). Therefore, the SFZ should be clarified for its potential future earthquake generation because this area is defined as a hazardous earthquake source.

Previous investigations using statistical seismological techniques in the SFZ evaluated the prospective areas of upcoming major earthquakes by analyzing the precursors of seismic activity. For the b-value of the frequency–magnitude distribution model (FMD; Gutenberg and Richter 1944), Pailoplee (2013) revealed that two prospective areas might be generated by forthcoming earthquakes,

*Correspondence:
Santi Pailoplee
Pailoplee.S@gmail.com
[1] Center of Excellence in Morphology of Earth Surface and Advanced Geohazards in Southeast Asia (MESA CE), Department of Geology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

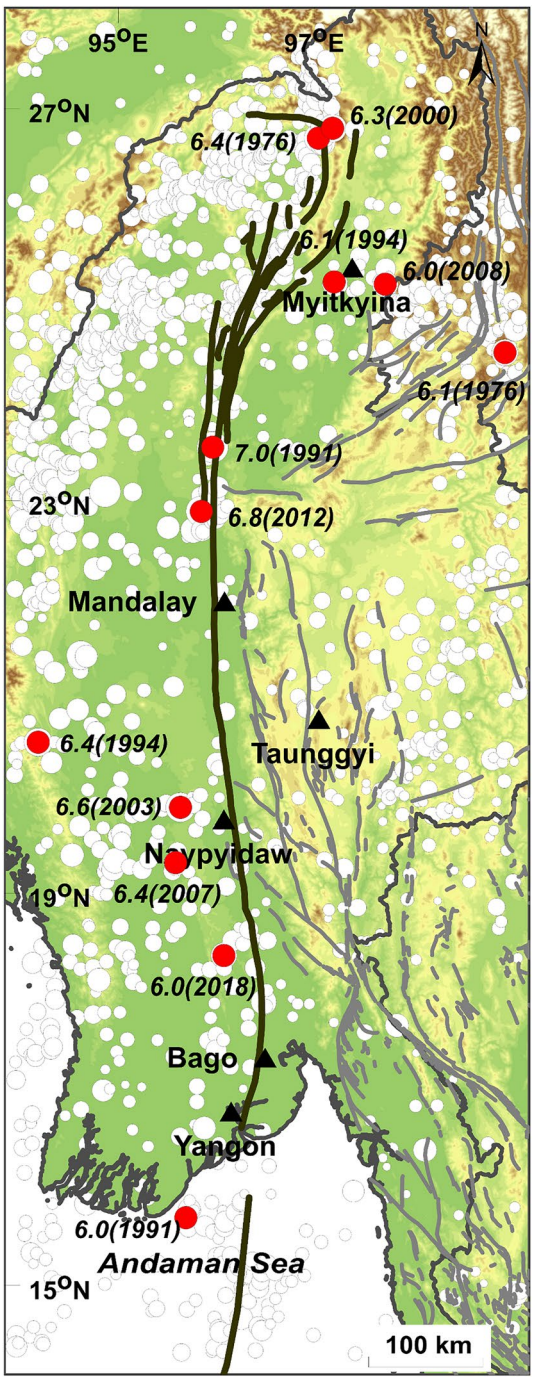

Fig. 1 Map of Myanmar showing the SFZ (dark brown line) trending north–south direction in the central part of Myanmar and the other fault systems (grey lines). Red circles represent the epicenter of earthquakes with $M_w = 6.0$ or larger distributed in the SFZ during 1972–2020, while white circles represent the completeness earthquake catalog utilized in this study. The black triangle indicates the locations of important cities in Myanmar

namely (i) southern Myitkyina, and (ii) Naypyidaw-Mandalay segment. In addition, Traitangwong and Pailoplee (2017) proposed that the western Myitkyina and the area in the vicinity of Naypyidaw are areas at risk from major earthquakes in the future by using region-time-length (RTL) algorithm (Sobolev and Tyupkin 1997; Huang et al. 2001) while using the Z-value (Habermann 1983), there are two prospective areas (Myitkyina and Naypyidaw cities) might be posed by the hazardous earthquake in the future (Pailoplee et al. 2017). The above-mentioned prospective areas are likely to generate the maximum earthquake ranging from 6.0 to 7.2 within the next 50 years (Pailoplee and Nimnate 2022). However, according to the instrument records from 2015 until now (2020), these two risk areas have not generated a magnitude of earthquake greater than 6.0. Therefore, these two still quiet areas need to be monitored and investigated for possible future major earthquakes.

Among the previous technique, Pattern Informatics (PI) algorithm is one of the potential methods that has been successfully used to analyze the precursory seismic quiescence and activation, which is a sign of future major earthquake (e.g., Tiampo et al. 2002; Rundle et al. 2002, 2003; Holiday et al. 2005; Nanjo et al. 2006; Zhang et al. 2009, 2013, 2017; Jiang and Wu 2010; Wu et al. 2011; Mohanty et al. 2014; Kawamura et al. 2014; Chang et al. 2016). Therefore, the main purpose of this study is to investigate the precursory seismic activity and evaluate the prospective areas of the forthcoming major earthquake along the SFZ by the PI method to constrain the prospective areas mentioned previously by the b-value, the RTL algorithm, and the Z-value.

PI algorithm

Theoretically, Pattern Informatics (PI) has been proposed by Rundle et al. (2002) and Tiampo et al. (2002) based on the statistical dynamic of complex systems (space–time correlations) to analyze the fluctuation of seismicity (quiescence and activation stage) before the earthquake and then forecast the future location of the forthcoming earthquake. According to Holiday et al. (2006), the modified PI method recently proposed that each step is described as follows. First, the study region is binned into N square boxes with linear dimensions Δx. For each box, a time series n (x_i, t) is defined by the number of earthquakes with a magnitude larger than the lower cut-off magnitude M_c during base time t_b and present time t. For the time parameters, there are considered into three intervals: a reference time interval (t_b to t_1), the anomaly training window or change of time interval (t_1 to t_2), and the forecast time interval from t_2 to t_3 for validating the forecast. In each box, the seismic intensity function is

defined to be the average number of earthquakes in box i during the time interval t_b to t as shown in Eq. 1:

$$\mathbf{I}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t})} = \frac{\sum_{\mathbf{t}=\mathbf{t_b}}^{\mathbf{t}} \mathbf{n}_{(\mathbf{x_i},\mathbf{t})}}{\mathbf{t}-\mathbf{t_b}}. \quad (1)$$

The next step is to normalize the seismic intensities by subtracting the mean seismic activity for all boxes and dividing by spatial standard deviation shown in Eq. 2 to compare the activity rates from different times:

$$\widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t})} = \frac{\mathbf{I}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t})} - \frac{1}{\mathbf{N}}\left(\sum_{\mathbf{j}=1}^{\mathbf{N}} \mathbf{I}_{(\mathbf{x_j},\mathbf{t_b},\mathbf{t})}\right)}{\sqrt{\sum_{\mathbf{j}=1}^{\mathbf{N}}\left[\mathbf{I}_{(\mathbf{x_j},\mathbf{t_b},\mathbf{t})} - \frac{1}{\mathbf{N}}\sum_{\mathbf{k}=1}^{\mathbf{N}} \mathbf{I}_{(\mathbf{x_k},\mathbf{t_b},\mathbf{t})}\right]^2}}. \quad (2)$$

And then, this step is measuring the change in the normalized seismic intensities period of t_b to t_2 and t_b to t_1 as shown in Eq. 3:

$$\boldsymbol{\Delta}\widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t_1},\mathbf{t_2})} = \widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t_2})} - \widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t_1})}. \quad (3)$$

The average change in the normalized activity rate function over all possible base-time periods is calculated to reduce the random fluctuation components in seismic activity that are shown in Eq. 4:

$$\boldsymbol{\Delta}\widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})} = \frac{\sum_{\mathbf{t_b}=\mathbf{t_0}}^{\mathbf{t_1}} \boldsymbol{\Delta}\widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_b},\mathbf{t_1},\mathbf{t_2})}}{\mathbf{t_1}-\mathbf{t_0}}. \quad (4)$$

After that, this step defines the probability of change of activity (seismic activation and seismic quiescence) that is related to the square of the average intensity change as shown in Eq. 5:

$$\mathbf{P}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})} = [\boldsymbol{\Delta}\widehat{\mathbf{I}}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})}]^2. \quad (5)$$

Finally, the new probability function is identified as the anomalous region that computes the change in the probability relative to the background by subtracting the mean probability over all boxes as shown in Eq. 6:

$$\widehat{\mathbf{P}}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})} = \mathbf{P}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})} - \frac{1}{\mathbf{N}}\sum_{\mathbf{j}=1}^{\mathbf{N}} \mathbf{P}_{(\mathbf{x_j},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})}. \quad (6)$$

From Eq. 6, hotspots are defined to be the regions where $\widehat{\mathbf{P}}_{(\mathbf{x_i},\mathbf{t_0},\mathbf{t_1},\mathbf{t_2})}$ is positive. $P_{(x_i,t_0,t_1,t_2)}$ is larger than the average value for all boxes (the background level) in these regions. Due to interest in seismic activation and seismic quiescence relative to the background, if boxes have the PI scores less than zero, these scores are replaced by zero. The PI score is zero to one.

This method has been applied to many earthquake sources such as in California, Japan, and worldwide for forecasting future earthquakes. For California, after publishing the forecast map area of M > 5.0 future earthquakes in the 2000–2010 period, it was later found that 19 out of 20 earthquakes (M > 5.0) occurred in the hotspot or anomalous forecast area until February 2008 (Holliday et al. 2005). Moreover, the modified PI method in Japan was also successful, with the Tokyo forecast map published in October 2004, and a few days later, the M_w-6.8 Niigata earthquake occurred very close to the anomalous zone in the forecast map (Nanjo et al. 2006). For this reason, it can be said that the PI method is one of the potential methods and is reliable, which can detect the precursory seismic activity and forecast the forthcoming earthquake in any specific earthquake source including the SFZ.

Dataset and completeness

In this study, the main dataset was the instrumental earthquake data covering a 300-km extension from the SFZ (Fig. 1). The earthquake catalog was compiled by the Incorporated Research Institutions for Seismology (IRIS). These datasets consist of 4562 earthquake events recorded from 1972 to 2022 (shown as line 1 in the cumulative number plot in Fig. 2b which is not close to linearity). For the datasets, the focal depths are less than 50 km. were identified as intra-plate earthquakes of the SFZ. Therefore, the focal depths greater than 50 km, which were defined as intraslab earthquakes of the Sumatra–Andaman subduction, were eliminated. The earthquake magnitude was reported in the different magnitude scales, i.e., moment magnitude (M_w), body-wave magnitude (m_b), and surface-wave magnitude (M_S). Therefore, the magnitude scales in m_b and M_S were converted to M_w according to the empirical relationships, as proposed by Pailoplee et al. (2009) to homogenize seismicity scales of magnitude.

In general, the cluster of earthquakes consists of mainshocks, foreshocks, and aftershocks, with only mainshocks directly reflecting the seismotectonic activities. Therefore, it is necessary to filter out only the mainshocks and eliminate the foreshocks and aftershocks using the seismicity declustering algorithm proposed by Gardner and Knopoff (1974). From Fig. 2a, the red lines in both time and distance distributions in Fig. 2a are used to distinguish between dependent and independent earthquakes with mainshock being above the red lines and foreshock or aftershock being below the red lines (see more detail in the supplementary information). Based on the result, 559 clusters were identified from 4,562 earthquakes with 1,477 events as the mainshocks (Fig. 2a). After declustering, the trend line of the cumulative number seems to be straighter than the previous step seen on line 2 in Fig. 2b.

After that, the Genetic Network Analysis System (GENAS) algorithm (Habermann 1983), operated in the ZMAP software (Wiemer 2001), was used to exclude

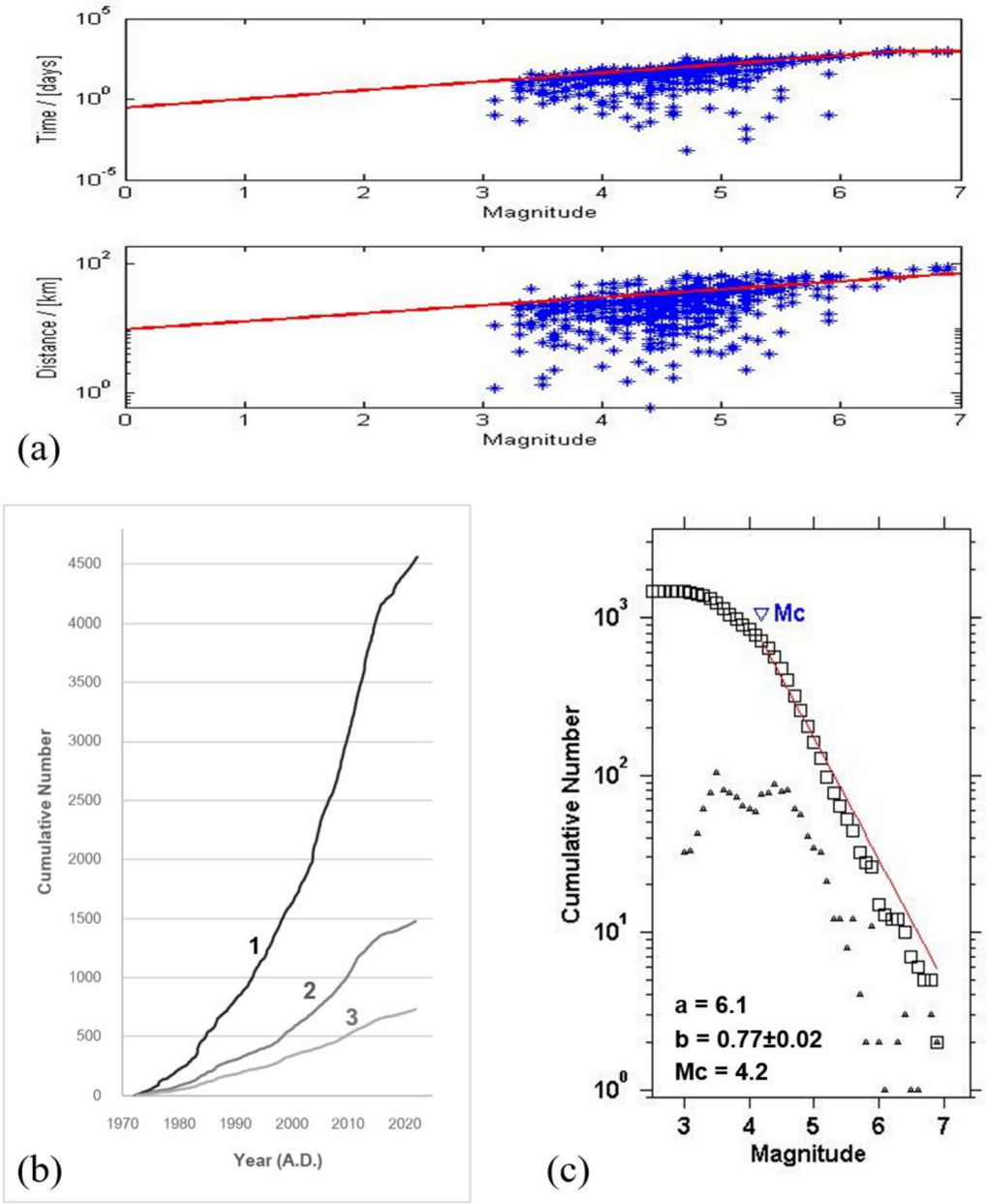

Fig. 2 **a** The declustering algorithm of Gardner and Knopoff (1974) exists within a fixed time and space window. The earthquake data (blue crosses) above the red lines of both time and space windows were identified as mainshocks. **b** Cumulative number of earthquake catalog versus time along the SFZ. The line 1 represents the composite data before declustering, while the line 2 shows the dataset after declustering. Then, after filtering out artificial seismicity based on the GENAS algorithm, the completeness data is represented by the line 3. **c** Frequency–magnitude distribution plots of the completeness seismic data. Triangles and squares represent the number and cumulative number of earthquakes in each magnitude, respectively. The red line is the best fit evaluated by the maximum likelihood method from Woessner and Wiemer (2005) which is used to estimate a and b values. Mc means the magnitude of completeness

man-made changes that could affect the seismicity variation (Wyss 1991; Zuniga and Wiemer 1999). This reveals that the dataset with $M_w \geq 3.6$ events reporting in 1980–2021 did not include significant artifacts, as shown in line 3 in Fig. 2b, which shows that the cumulative number of earthquakes approaches the straight line as the completeness data processes. Finally, the completeness earthquake dataset consists of 736 mainshocks of magnitude range of $3.6 \leq M_w \leq 7.0$ with depths of 0–50 km from 1980 to 2021 (Fig. 1). These datasets were used in the PI investigation.

In addition, the magnitude of completeness (Mc) of the FMD plot was selected at 4.2 (Fig. 2c), following the entire-magnitude range assumption (Woessner and Wiemer 2005). This Mc is necessary to determine the magnitude of the target earthquake for the PI investigation in the next section.

Retrospective investigation

As mentioned above in the PI algorithm section, it is necessary to choose the most suitable parameters for a specific area to achieve efficiency and accuracy in the PI anomalies related to a hazardous earthquake. The parameters for the PI algorithm include: (i) the time parameters (t_0, t_b, t_1, t_2); (ii) space parameters of the scale of boxes in the study area; and (iii) Mc related to the magnitude of the earthquake that will occur during the forecast time interval. For selecting proper parameters, previous studies are used as guidelines for choosing PI parameters as follows. Holliday et al. (2005) used the change time and the forecast time intervals to be both 10 years to forecast target earthquake $M \geq 5.0$ in southern California and central Japan within box size as $0.1^\circ \times 0.1^\circ$ while forecast target earthquake $M \geq 7.0$ for worldwide within box size as $1^\circ \times 1^\circ$ by PI method. Chen et al. (2005) applied the PI method to investigate the variations in seismicity pattern changes before the 1999 Chi-Chi, Taiwan, earthquake by selecting the change time interval around 6 years. For large earthquake forecasting in Japan, Nanjo et al. (2006) used the change time interval to be equal to the forecast time interval (10 years) to forecast earthquake event $M \geq 5.0$ ($Mc = 3.0$) by selecting the box size as $0.1^\circ \times 0.1^\circ$ (11 km.). For the retrospective test for the Wenchuan Earthquake, Sichuan-Yunnan region, Jiang and Wu (2010) adapted the variation parameter and then selected the change time and the forecast time intervals to be 5 years within grid size as 0.2° for forecasting target earthquake $M \geq 5.0$ ($Mc = 3.0$). To study the precursory migration of anomalous seismic activity of the 2011 Tohoku earthquake by the PI method, Kawamura et al. (2013) used a grid size of $0.25^\circ \times 0.25^\circ$ and a minimum Mc of 4.0 related to the forecast target magnitude earthquake (≥ 5.0). For the time intervals, 8 years (1 January 2000–1 January 2008) was chosen for the change time while the forecast time was chosen only to cover the time of the case study (11 March 2011). In addition to previous works that specify the parameters for PI analysis, there are also previous works that vary variables to find the most appropriate parameters. Zhang et al. (2013) studied the predictability of the PI method by testing with the M-8.0 Wenchuan and M-7.3 Yutian, varying both of the time intervals (1–10 years) and the grid size ($1^\circ \times 1^\circ$ and $2^\circ \times 2^\circ$). To obtain the suitable parameters, the relative operating characteristics (ROC) diagram and R score are used to verify. These results from the verification test show that the grid size of $2^\circ \times 2^\circ$ and the forecast time interval of 8 years are the best parameters in this area.

Considering these previous studies, the study region is divided into a grid of square boxes with a size of $0.25^\circ \times 0.25^\circ$, which has a total of 7,506 boxes. The lower magnitude cutoff is set as M_w 4.2. Refer to the assumption of Holiday et al. (2005), the target earthquake in the forecast time is set to be greater than $Mc + 2$. In this study area, it is set at a magnitude larger than 5.0, which is different from Holiday's assumption because the number of earthquakes larger than 6.2 in this area is not many and not enough to be chosen as the case studies. For the time parameters, the initial time was $t_0 = 1972$, the beginning time was starting at 1980 (start of the completeness dataset), and the time step was every 5 years until 2015, while the change time and forecast time intervals were set to be the same interval that varied between 5 or 10 years for the retrospective forecasting of the PI algorithm. Thus, each PI value investigation that depended on time interval conditions was tested iteratively.

For finding the best condition of time interval parameters in this study area, the ROC diagram was chosen to verify the retrospective test with different time interval conditions between 5 and 10 years. The ROC diagram (Swets 1973; Jolliffe and Stephenson 2003; Mason 2003) is one of the well-known methods for validating the binary forecast, like the PI method. This technique is used to count the hit rate and false alarm rate from real earthquake activity compared with the alarm boxes of the forecast region. The forecast is considered successful when determined by maximizing the proportion of earthquakes that occur in alarm boxes while decreasing the fraction of alarm boxes that occur outside of boxes. As a consequence of this test, there are four categories for each box: (a) successful forecast: earthquake occurs in the hotspot boxes or within its 8 Moore neighborhood (alarm and earthquake); (b) false alarm: no earthquake occurs in a hotspot box or within its 8 Moore neighborhood (alarm but not earthquake); (c) failure to predict: an earthquake occurs in non-hotspot (no alarm but an earthquake); (d) successful forecast of non-occurrence: no earthquake occurs in a white non-hotspot box (no alarm and no earthquake), where the Moore neighborhood is composed of nine boxes—the forecast hotspot box and the eight boxes that surround it (Moore 1962). From the studies of Holliday et al. (2005) and Nanjo et al. (2006), it was concluded that if future earthquakes occur in the forecasted anomalous area or within the margin of error of ± 1 box dimension, it indicates that the forecast was successful. So, a margin of error of 28 km was set in this study. In addition, values for the hot spot map consist of values

a (forecast = yes, observed = yes), b (forecast = yes, observe = no), c (forecast = no, observed = yes), and d (forecast = no, observed = no), which will be explained in the contingency table (see Table 1 in the supplementary information). The fraction of colored boxes also called the probability of forecast of occurrence, is r = (a + b)/N, where the total number of boxes (N) is N = a + b + c + d. The hit rate (H) and a false alarm rate (F) are used to indicate the success of a forecast as follows: the hit rate (H) is H = a / (a + c), is the fraction of large earthquakes that occur on a hot spot and the false alarm rate (F) is F = b / (b + d), is the proportion of non-observed earthquakes that are incorrectly forecast as Table 1 in the supplementary information.. From the ROC diagram of both conditions as 5 and 10 years (Fig. 3), it can be seen that no matter how the threshold changes, the hit rate is always greater than the false alarm rate, and any value of PI in both conditions exceeding the random value (represented in the horizontal blue dashed line). These mean the PI method definitely outperforms the random forecast in this region. When comparing the two conditions, this plot shows that 3 out of 6 case studies of 5 years condition (represented in black color) have significantly lower the hit rate than the 10 years condition (represented in grey color), with these hit rates lower than 0.1 as shown in Fig. 3. Additionally, the hit rate of three cases in the 10 years ranged from 0.288 to 0.317 was significantly higher than other cases in the 5 years. Overall, it can be concluded that the hit rate in 10 years condition is greater than in 5 years condition. For these reasons, the 10-year condition is the suitable time interval parameter for the PI investigation.

For the retrospective test with the PI method, the spatial distribution map of PI value was evaluated and mapped into the hotspot areas (color codes) that represent areas where ΔP is positive and is associated with a relative probability increase for strong earthquakes. For the color code that represents the seismicity changes region, it is explained as follows. Warm tones represent the areas with large changes in seismicity (seismic quiescence and activation) during the change time interval, which indicates that these areas have a high chance of earthquake occurrence after the change interval, with the red color indicating the highest probability of earthquake occurrence. The cold tones show the areas with small changes in seismicity, which indicates that these areas have a low chance of earthquake occurrence after the change interval (Fig. 4).

From the PI investigation in six case studies, these forecast results of earthquakes with magnitude ≥ 5.0 (shown by stars; red stars represent the epicenter of earthquakes that occur in or within 28 km of the anomalous area, and brown stars represent those that occur more than 28 km from the anomalous area.) with different time intervals from 1990 until 2020, respectively, are shown in Fig. 4a–f and explained as follows. From Fig. 4c, the forecast hotspots for the occurrence of $M_w \geq 5.0$ from 2000 to 2010, defined as $t_1 = 1$ January 1990 and $t_2 = 1$ January 2000 show that there are 15 earthquakes in total, of which

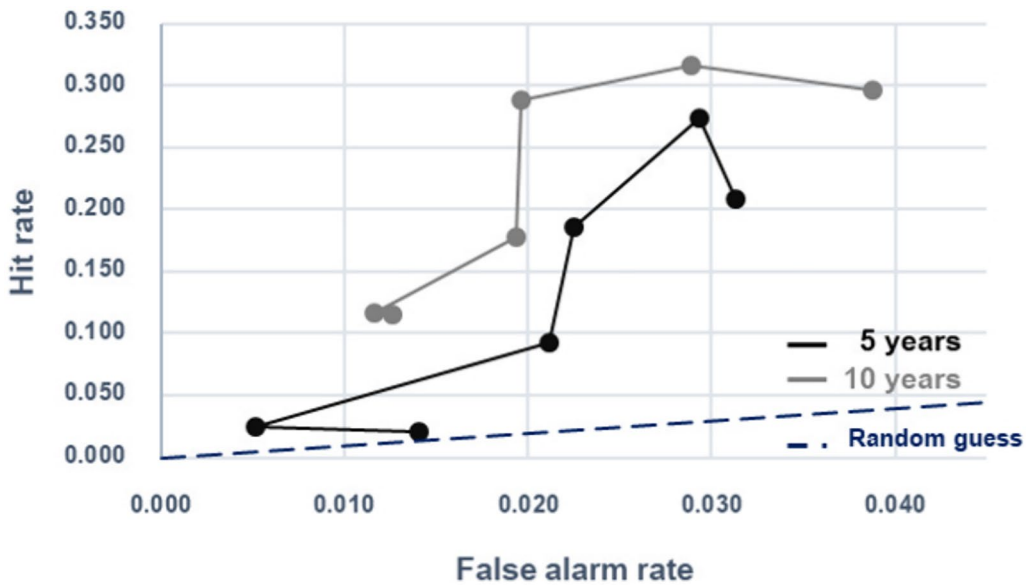

Fig. 3 ROC diagram which plots of hit rate and false alarm rate for earthquake forecasts by PI method using different time intervals between 5 and 10 years (represented by black and grey lines, respectively)

12 are red stars and 3 are grey stars. The 12 red stars occurred in or near the anomalous activity area, which is the Myitkyina and surrounding areas, Naypyidaw, and the Andaman Sea. In particular, the Myitkyina and Naypyidaw are high PI value zones, which are in the range of 0.8–0.9 and 0.4–0.6, respectively, and the epicenter of more than 6 events was located in a cluster of hotspots in these areas with 5 events (composed of M_w 6.0 (2008), M_w 5.4 (2000), M_w 5.2 (2007), M_w 5.1 (2005), and M_w 5.0 (2005)) in the Myitkyina and 3 events (composed of M_w 6.6 (2003), M_w 6.4 (2007), and M_w 5.0 (2009) in the Naypyidaw, as shown in red stars. Likewise, for the forecast interval from 2005 to 2015 (Fig. 4d), defined as $t_1 = 1$ January 1995 and $t_2 = 1$ January 2005, it is clear that the anomalous zones with a high PI value remain located in the same areas as Myitkyina (PI = 0.7–0.8), Naypyidaw (PI = 0.3–0.5). In addition to the earthquakes that occurred in 2000–2010, it was also found that in these anomalous zones, there were three new earthquakes during the periods 2010–2015, which were the magnitude of 5.5, 5.1, and 5.9 in 2011, 2012, and 2014, respectively. While the anomaly in the Andaman Sea remains in the same zones with a low PI value of approximately 0.1–0.2, it is considered that there is a low probability of an earthquake occurring in this area and is in accordance with a forecast that earthquakes will occur far away from the surrounding of the anomalous zones, which is M_w 5.2 in 2008 and M_w 5.3 in 2009. Another anomalous zone that has increased is Bago. The earthquakes fall into the surrounding anomalous area, even though it does not occur exactly at the epicenter or at the highest anomalous PI, as well.

Thereafter, select $t_1 = 1$ January 2000 and $t_2 = 1$ January 2010 to forecast the possible locations of the occurrence of earthquakes during the period 2010–2020, the anomalous zone around Bago still shows high PI values (0.2–0.5), and M_w 5.1 earthquake fell into this area in 2017 as seen in the red star in Fig. 4e. Moreover, it can be seen that earthquakes (M_w 5.0 in 2016 and M_w 5.1 in 2018) will occur in the Andaman Sea, which is close to the anomalously high PI zones, with a value of approximately 0.7–0.8. The vicinity of Myitkyina is still active. When moving the period to 5 years ($t_1 = 1$ January 2005 and $t_2 = 1$ January 2015), the forecast result for anomalous zones in 2015–2025 is still exposed near the same areas as in the period 2010–2020, namely the Myitkyina, Naypyidaw, Bago and Andaman Sea. Moreover, it can be seen that the anomalous zones around Bago and Andaman Sea have expanded and have high values of 0.4–0.6 and 0.7–0.9, respectively, which makes the forecast more accurate, as indicated by the M_w 5.1 Bago earthquake in 2017 and the M_w 5.0 Andaman Sea earthquake in 2016 as seen in Fig. 4f.

According to the spatial distribution of anomalous PI value during the forecast period 1990–2000, defined as $t_1 = 1$ January 1980 and $t_2 = 1$ January 1990, the anomalous high PI values were evident in the Andaman Sea, the northwest of the Myitkyina, and the Naypyidaw. Figure 4a shows that the epicenters of the 1991 earthquake located in an anomalous area with a high PI value in the Andaman Sea (PI = 0.6–0.7) that is fairly conforming to the high probability of occurrence. Meanwhile, the M_w 5.2 (1999) and M_w 5.1 (1995) earthquakes are located around the anomalous areas of Naypyidaw and the northwest of Myitkyina, with low PI values of 0.2–0.3, which is also fairly conforming to the low probability of occurrence. In addition to the case of forecast period 1990–2000, the case of forecast period 1995–2005, defined as $t_1 = 1$ January 1985 and $t_2 = 1$ January 1995, it was shown that although the epicenter of all earthquakes (represented by the stars) was not located in the anomalous areas, 4 earthquakes occurred in the surrounding of areas with anomalous PI values. Figure 4b shows that the magnitudes of 6.6 earthquakes in 2003 and 5.2 earthquakes in 1999 occurred close to the anomalous areas in Naypyidaw, while the magnitudes of 5.1 earthquakes in 1995 and 5.0 earthquakes in 2003 occurred close to the anomalous areas in the northwest of Myitkyina.

According to the retrospective test and the ROC diagram verification technique, selecting the characteristic parameters that include the time intervals equal to 10 years, a spatial grid size of 0.25°, and a target earthquake magnitude ≥ 5.0 (Mc = 4.2) will result in high PI values which indicates the seismogenic region reliably with reasonable accuracy to evaluate the prospective areas of upcoming major earthquake along the SFZ in the next session.

(See figure on next page.)

Fig. 4 Map of the SFZ demonstrating the spatial distributions of the PI values evaluated and mapped as the hotspot area in this study with different periods. The intervals of change time and forecast time are taken both as 10 years, starting from 1990 to 2015, with the sliding step being 5 years, which is shown in six case studies: **a** period 1990–2000, **b** period 1995–2005, **c** period 2000–2010, **d** period 2005–2015, **e** period 2010–2020 and **f** period 2015–present. The color code in the PI hazard map (red to blue colors) represents areas where ΔP is positive which implies the seismic activity rate. Red stars indicate the epicenters of the magnitude ≥ 5.0 earthquakes occurring in or within 28 km of the anomalous zones in each case of the forecast period considered in this study

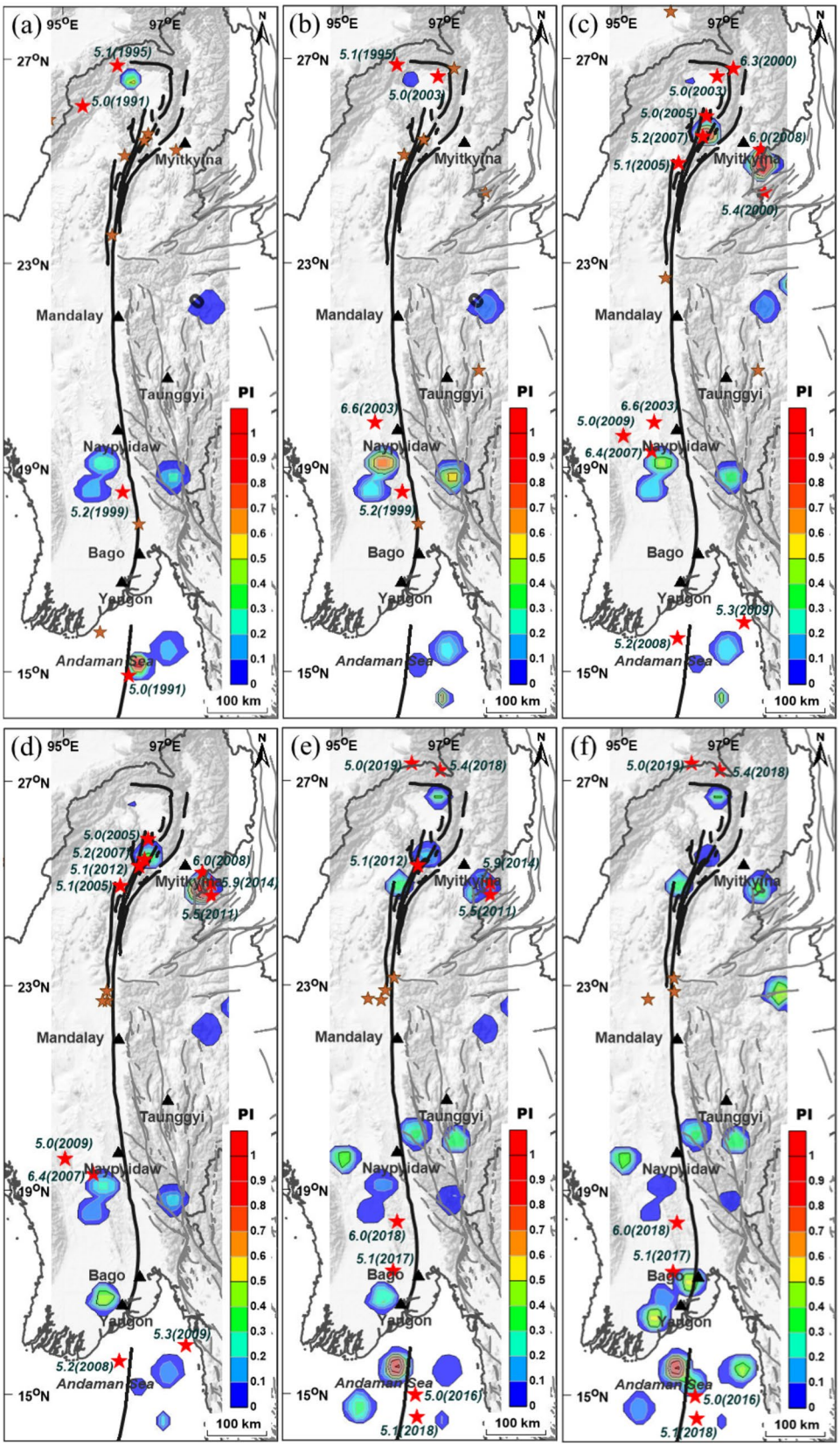

Fig. 4 (See legend on previous page.)

Present-day investigation

Using the suitable parameters obtained from the previous section, the present-day situation was evaluated by using the PL algorithm with the most up-to-date earthquake dataset. After investigating the spatial distribution of PI value during 2020–2030, it was discovered that the regions two conspicuous regions along the SFZ show high PI anomalies, namely the Myitkyina and the area in the vicinity of Naypyidaw (Fig. 5a).

To confirm that the anomalous areas obtained by PI algorithm investigation are not involved with the random phenomena of earthquake occurrence, a statistical method called a stochastic process (Huang 2005) was chosen. This is a new validation application implemented in the PI method. After generating a random seismicity dataset, it was calculated by the PI method, setting the same suitable parameters as used on real data. For the synthetic seismicity dataset, there will be a total of 10,000 events with the range of magnitude 3.0–8.0. As a result of random seismicity dataset, the PI value in plot (Fig. 5b.) are set a spacing every 0.1. According to the stochastic result, it is revealed that the relationship between the probability of occurrence and the PI value fluctuates gradually stabilized when the PI value increased to 0.7 and remained constant until the PI value equal to 1.0 as shown in Fig. 5b. In other words, the higher the PI value,

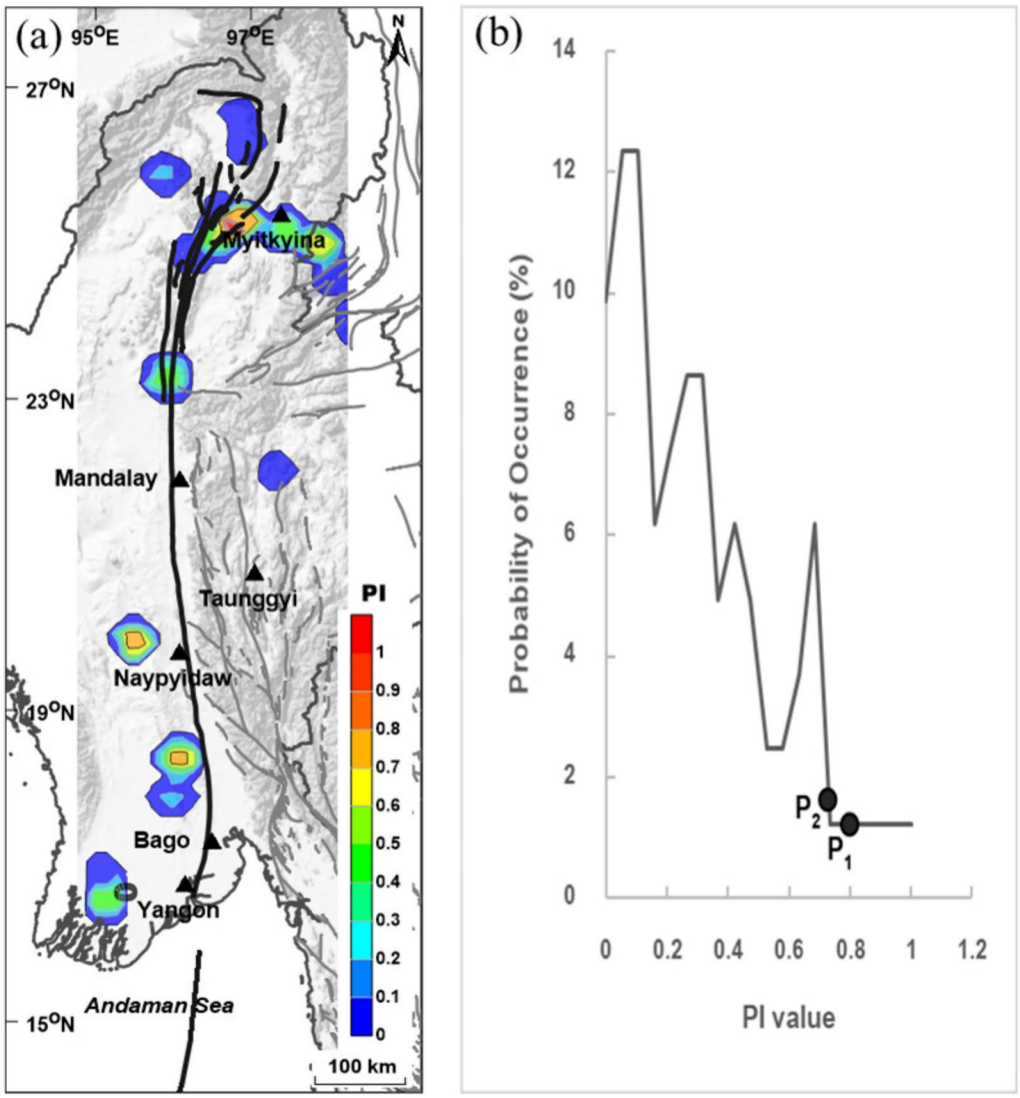

Fig. 5 **a** Map of the SFZ illustrating the spatial distribution of the present-day PI value, which forecasts future major earthquakes during the period 2020–2030. **b** The stochastic test of the PI value estimated from random seismic data shows the probability of occurrence at each PI value level, where P_1 and P_2 points represent the high PI values found in the anomalous zones of the Myitkyina and the Naypyidaw, respectively

the lower the chance of random occurrences. For the two prominent anomalous areas along the SFZ, the obtained PI values in the Myitkyina and the Naypyidaw are 0.8 and 0.7 (point as P_1 and P_2), respectively. When comparing with the probability of occurrence, it is implied that the anomalies observed in this study have only a 1% probability of being random.

In addition, to constrain the prospective areas that might generate hazardous earthquakes in the future along the SFZ, comparing the resulting map with the previous works by the FMD b-value (Fig. 6a; Pailoplee 2013), the RTL algorithm (Fig. 6b; Traitangwong and Pailoplee 2017) and the Z-value (Fig. 6c; Pailoplee et al. 2017). The prospective areas for forthcoming earthquake sources evaluated by these methods indicated the same areas as the PI method studies (Fig. 5a), but it can be seen that the PI method can identify more specifically the anomalous areas that are smaller and clearer than other methods. Therefore, this PI study supports a high possibility of a major earthquake being generated soon at Myitkyina and in the vicinity of Naypyidaw.

Furthermore, Pailoplee (2022) studied seismic activity in terms of the possible maximum magnitude that might be generated in the next 5–100 years. From the result of Pailoplee (2022), it can be evaluating the maximum magnitude of earthquake expected to generate in the prospective areas proposed in this study (Fig. 5a), Pailoplee (2013), Traitangwong and Pailoplee (2017), and Pailoplee et al (2017). This study focuses on the human life span, which corresponds to the next 50 years, so the M_w in the next 50 years was used. From the potential hazard areas mentioned, there are two areas: Myitkyina and Naypyidaw. When assessed by Pailoplee and Nimnate (2022), it was revealed that there is a probability for the maximum earthquakes occurring in the next 50 years with M_w of 6.4–7.2 and M_w of 5.6–6.2 in the Myitkyina and the Naypyidaw, respectively (Fig. 6d).

Conclusion remarks

In this study, the PI algorithm was used to investigate the precursory seismic quiescence before the occurrence of major earthquakes and then to generate the map of spatial distributions along the SFZ. After improving the earthquake dataset, the completeness earthquake catalog ($M_w \geq 3.6$ reported during 1980–2020) was used to investigate. To find out the suitable parameters, the retrospective tests were performed from 1990 to 2015, moving every 5 years, and verified

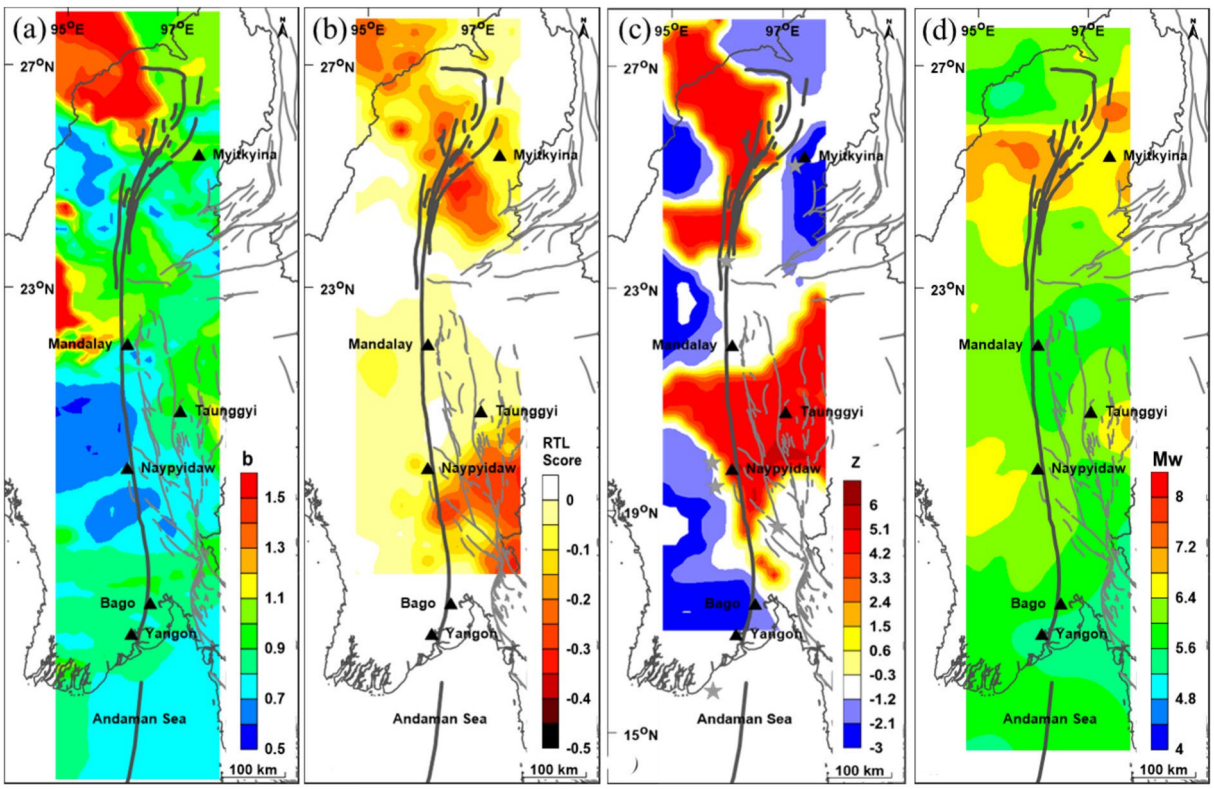

Fig. 6 **a** Map of the SFZ showing the spatial distribution of **a** the precursory b-value evaluated by the power-law FMD (Pailoplee 2013), **b** the present-day RTL score evaluated during 2009.84–2013.56 (Traitangwong and Pailoplee 2017), **c** the present-day Z-value determined in the time slices at 2012.83., and **d** the maximum earthquake magnitude (M_w) capable of being generated in the next 50 years

by the ROC Diagram Technique. These tests found that utilizing the suitable parameters of grid size as 0.25°, time intervals = 10 years., Mc = 4.2, and the target magnitude of the earthquake ≥ 5.0 can clearly show the consistency of the anomalous areas with high PI values and the earthquakes that occur in or near these areas. Although there were some cases, such as from 1990 to 2000, in anomalous areas (high PI values), there were no earthquakes due to limitations in the amount of seismic data used for analysis.

For the PI investigation of the present-day seismicity data of 2010–2020, this result revealed two prominent areas of high PI value anomalies along the SFZ at the Myitkyina and in the vicinity of the Naypyidaw, which agree with the anomalous maps proposed by the FMD b-value (Pailoplee 2013), the RTL algorithm (Traitangwong and Pailoplee 2017) and the Z-value (Pailoplee et al. 2017). Furthermore, to confirm that the obtained anomalous PI before the major earthquake was likely caused by tectonic activity directly (not random), the stochastic process was tested. Various statistical seismology studies mentioned above suggest that the areas at risk for future major earthquakes are still the same. So, it is implied that these areas are still active. These make sense because there has never been a large earthquake in Myanmar. Therefore, plans for preventing and mitigating potential future disasters in Myanmar should be contributed urgently.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40562-024-00351-9.

Additional file 1.

Acknowledgements
The Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72nd anniversary of his Majesty King Bhumibol Aduladej is gratefully acknowledged. Thanks are extended to this research fund supported by Thailand Science research and Innovation Fund Chulalongkorn University. The thoughtful comments and suggestions by the editors and anonymous reviewers that enhanced the quality of this manuscript significantly are acknowledged.

Author contributions
All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Premwadee Traitangwong. The first draft of the manuscript was written by Premwadee Traitangwong and Santi Pailoplee and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by The Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72nd anniversary of his Majesty King Bhumibol Aduladej (Premwadee Traitangwong). Santi Pailoplee has received the research support funded by Thailand Science research and Innovation Fund Chulalongkorn University.

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 25 December 2023 Accepted: 29 July 2024
Published online: 07 August 2024

References

Bertrand G, Rangin C (2003) Tectonics of the western margin of the Shan plateau (central Myanmar): implications for the India-Indochina oblique convergence since the Oligocene. J Asian Earth Sci 21:1139–1157
Chang LY, Chen CC, Wu YH, Lin TW, Chang CH, Kan CW (2016) A strategy for a routine pattern informatics operation applied to Taiwan. Pure Appl Geophys 173:235–244
Chen CC, Rundle JB, Holliday JR, Nanjo KZ, Turcotte DL, Li SC, Tiampo KF (2005) The 1999 Chi-Chi, Taiwan, earthquake as a typical example of seismic activation and quiescence. Geophys Res Lett 32:L22315
Curray JR (2005) Tectonics and history of the Andaman Sea region. J Asian Earth Sci 25:187–232
Gardner JK, Knopoff L (1974) Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bull Seismol Soc Am 64(1):363–367
Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34:185–188
Habermann RE (1983) Teleseismic detection in the Aleutian Island Arc. J Geophys Res 88:5056–5064
Holliday JR, Nanjo KZ, Tiampo KF, Rundle JB, Turcotte DL (2005) Earthquake forecasting and its verification. Nonlinear Process Geophys 12:965–977
Holliday JR, Rundle JB, Tiampo KF, Klein W, Donnellan A (2006) Modification of the pattern informatics method for forecasting large earthquake events using complex eigenfactors. Tectonophysics 413:87–91
Huang Q (2005) A method of evaluating reliability of earthquake precursors. Chin J Geophys 48:701–707
Huang Q, Sobolev GA, Nagao T (2001) Characteristics of the seismic quiescence and activation patterns before the M = 7.2 Kobe earthquake, January 17, 1995. Tectonophysics 337:99–116
Jiang C, Wu Z (2010) PI forecast for the sichuan-yunnan region: retrospective test after the May 12, 2008, Wenchuan earthquake. Pure Appl Geophys 167:751–761
Jolliffe IT, Stephenson DB (2003) Forecast Verification. John Wiley, Chichester
Kawamura M, Wu YH, Kudo T, Chen CC (2013) Precursory migration of anomalous seismic activity revealed by the pattern informatics method: a case study of the 2011 Tohoku earthquake. Japan Bull Seismol Soc Am 103(2B):1171–1180
Kawamura M, Wu YH, Kudo T, Chen CC (2014) A statistical feature of anomalous seismic activity prior to large shallow earthquakes in Japan revealed by the pattern informatics method. Nat Hazards Earth Syst Sci 14(4):849–859
Kundu B, Gahalaut VK (2012) Earthquake occurrence processes in the Indo-Burmese wedge and Sagaing fault region. Tectonophysics 525:135–146
Mason IB (2003) Binary events. In: Joliffe IT, Stephenson DB (eds) Forecast verification. John Wiley, Chichester, pp 37–76
Mohanty WK, Mohapatra AK, Verma AK, Tiampo KF, Kislay K (2014) Earthquake forecasting and its verification in northeast India. Geomat Nat Hazards Risk 7:194–214
Moore EF (1962) Machine models of self reproduction. In: Belman R (eds). Proceedings of the Fourteenth symposium on applied mathematics. American Mathematical Society: Providence. pp 17-33.
Nanjo KZ, Rundle JB, Holliday JR, Turcotte DL (2006) Pattern informatics and its application for optimal forecasting of large earthquakes in Japan. Pure Appl Geophys 163:2417–2432

Pailoplee S (2012) Relationship between modified Mercalli intensity and peak ground acceleration in Myanmar. Nat Sci 4:624–630
Pailoplee S (2013) Mapping asperities along the Sagaing fault zone, Myanmar using b-value anomalies. J Earthq Tsunami 7(5):1371001
Pailoplee S, Nimnate P (2022) Earthquake activities along the Sagaing fault zone, central Myanmar: implications for fault segmentation. J Earthq Tsunami 16(5):2250013
Pailoplee S, Sugiyama Y, Charusiri P (2009) Deterministic and probabilistic seismic hazard analyses in Thailand and adjacent areas using active fault data. Earth Planets Space 61:1313–1325
Pailoplee S, Panyatip S, Charusiri P (2017) Precursory seismicity rate changes prior to the large and major earthquakes along the Sagaing fault zone. Central Myanmar Arab J Geosci 10:444
Rundle JB, Tiampo KF, Klein W, Martins JSA (2002) Self-organization in leaky threshold systems: the influence of near mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. Proc Natl Acad Sci USA 99(1):2514–2521
Rundle JB, Turcotte DL, Shcherbakov R, Klein W, Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. Rev Geophys 41:1019–1038
Sobolev GA, Tyupkin YS (1997) Low-seismicity precursors of large earthquakes in Kamchatka. J Volcanol Seismol 18:433–446
Swe W (1981) A major strike-slip fault in Burma. Contrib Burmese Geol 1(1):63–67
Swets JA (1973) The relative operating characteristic in psychology. Science 182:990–1000
Tiampo KF, Rundle JB, McGinnis S, Gross SJ, Klein W (2002) Eigenpatterns in southern California seismicity. J Geophys Res. https://doi.org/10.1029/2001JB000562
Traitangwong P, Pailoplee S (2017) Precursory seismic quiescence along the Sagaing fault zone, central Myanmar-application of the region-time-length algorithm. Geosci J 21(4):543–552
Wiemer S (2001) A software package to analyse seismicity: ZMAP. Seismol Res Lett 72:373–382
Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 95(2):684–698
Wu YH, Chen CC, Rundle JB (2011) Precursory small earthquake migration patterns. Terra Nova 23:369–374
Wyss M (1991) Evaluation of proposed earthquake precursors. evaluation of proposed earthquake precursors. EOS Trans Am Geophys Union 72(38):411–411
Zhang YX, Zhang XT, Yin XC, Wu YJ (2009) Study on the forecast effects of PI method to the North and Southwest China. Concurr Comput 22(12):1559–1568
Zhang YX, Zhang XT, Wu YJ, Yin XC (2013) Retrospective study on the predictability of pattern informatics to the Wenchuan M8.0 and Yutian M7.3 earthquakes. Pure Appl Geophys 170:197–208
Zhang YX, Xia CY, Song C, Zhang XT, Wu YJ, Xue Y (2017) Test of the predictability of the PI method for recent large earthquakes in and near Tibetan plateau. Pure Appl Geophys 174:2411–2426
Zuniga FR, Wiemer S (1999) Seismicity patterns: are they always related to natural causes? Pageoph 155:713–726

Publisher's Note