1 **Real-time probabilistic seismic hazard assessment based on seismicity anomaly**

3 Yu-Sheng Sun[1*], Hsien-Chi Li[1], Ling-Yun Chang[1], Zheng-Kai Ye[1], and Chien-Chih Chen[1, 2]

4 [1]Department of Earth Sciences, National Central University, Taoyuan City 32001, Taiwan, R.O.C.

5 [2]Earthquake-Disaster and Risk Evaluation and Management Center, National Central University,

6 Taoyuan City 32001, Taiwan, R.O.C.

7 Correspondence: Yu-Sheng Sun (sheng6010@gmail.com)

9 **Abstract**

10 Real-time Probabilistic Seismic Hazard Assessment (PSHA) was developed in this study in

11 consideration of its practicability for daily life and the rate of seismic activity with time. Real-time

12 PSHA follows the traditional PSHA framework, but the statistic occurrence rate is substituted by

13 time-dependent seismic source probability. Over the last decade, the Pattern Informatics method

14 (PI) has been developed as a time-dependent probability model of seismic source. We employed

15 this method as a function of time-dependent seismic source probability, and we selected two major

16 earthquakes in Taiwan as examples to explore real-time PSHA. These are the Meinong earthquake

17 (M_L 6.6) of 5 February, 2016, and the Hualien earthquake (M_L 6.2) of 6 February, 2018. The

18 seismic intensity maps produced by the real-time PSHA method facilitated forecasting the

19   maximum expected seismic intensity for the following 90 days. Compared with real ground motion

20   data from the P-alert network, our seismic intensity forecasting maps showed considerable

21   effectiveness. This result indicated that real-time PSHA is practicable and provides useful

22   information that could be employed in the prevention of earthquake disasters.

24   **1   Introduction**

25   Currently, research on and the application of seismic hazard analyses focus on two major aspects

26   of seismic activity, namely the pre-earthquake and post-earthquake phases. Post-earthquake

27   seismic hazard assessment is employed mainly in the Earthquake Early Warning (EEW) system

28   (Cooper, 1868; Wu et al., 1998; Wu et al., 2013), which provides people with crucial time to seek

29   refuge before the arrival of larger seismic waves. Pre-earthquake seismic hazard assessment

30   conventionally employs Probabilistic Seismic Hazard Analysis (PSHA; Cornell, 1968; SSHAC,

31   1997) mainly for engineering design. PSHA determines the probability of exceeding the ground

32   motion level over a specified time period based on the occurrence rate of earthquakes and ground

33   motion prediction equations (GMPEs). The occurrence rate of earthquakes is generally described

34   by the truncated exponential model (Cosentino et al., 1977) and the characteristic earthquake

35   model (Schwartz and Coppersmith, 1984; Wang et al., 2016). The earthquake occurrence rate

36   computed from these models will not change with time regardless of whether the data being used

37 are from long-term observations or paleoseismic studies. However, seismic activity is a complex

38 dynamic process in time and space, and usually fluctuates greatly over a short time scale (Chen et

39 al., 2006). Furthermore, the assessment is usually computed by using extremely long recurrence

40 intervals, 475 or 2475 years, for the purpose of engineering design (Iervolino et al., 2011).

41 Consequently, it is difficult to verify the accuracy of seismic hazard assessment in relation to the

42 limited lifespan of humans. Although long recurrence intervals are suitable in building

43 construction, the concept of "catastrophic" over such long intervals does not resonate with the

44 general public. In addition, most ordinary people would find it difficult to comprehend an

45 indication such as "10% probability in 50 years". Statistical long-term seismic hazard assessment,

46 therefore, does not have relevance to the daily life of most people.

47

48 However, we believe that short-term, time-dependent, pre-earthquake hazard assessment is

49 necessary for everyone's daily use. Accordingly, we propose a preliminary method to achieve this

50 goal by employing time-dependent seismic source probability instead of the static probability used

51 in long-term assessment. We used the Pattern Informatics (PI) method developed over the past

52 decade (Rundle et al., 2000; Tiampo et al., 2002; Wu et al., 2008a; Chang et al., 2016) as a time-

53 dependent seismic source probability method.

54

55 Anomalous change in seismicity is used widely as precursory indicator for large earthquakes, and

56 is usually classified into seismic activation or seismic quiescence, depending on an ascending or

57 descending number of seismicity occurrences (Chen et al., 2005; Wu et al., 2008b). In the PI

58 method, large earthquakes tend to occur after precursory anomalous seismic changes, and the

59 occurrence probability can be quantified by the magnitude of the spatiotemporal variation in

60 seismicity. In preliminary research, PI performed well in identifying locations in the vicinity of

61 impending large earthquakes. A modified version of PI developed in recent research has apparently

62 improved the accuracy of identifying the occurrence time interval of large earthquakes. After a

63 series of verifications, an occurrence probability of large earthquakes over the following 90 days

64 was found plausible (Chang et al., 2016; Chang, 2018). Accordingly, we used the modified PI

65 method to compute the time-dependent seismic source probability in Taiwan region.

66

67 We illustrate an uncomplicated method to conduct real-time seismic hazard assessment. The

68 crucial difference is to replace statistical seismic probability by the time-dependent probability

69 from the modified PI method. This real-time seismic hazard assessment is able to produce seismic

70 hazard forecasting maps for the following 90 days. Compared with the forecasting time scale and

71 static seismic rate of traditional PSHA, real-time PSHA can be updated by refreshing the

72 earthquake catalog (time-dependent), and can forecast for the near future (short-term). Thus, it can

73   be referred to as "real-time".

74

75   We illustrate the real-time assessment process by two recent large earthquake events in Taiwan,

76   namely the 2016 Meinong earthquake (M_L 6.6) (Lee et al., 2016; Chen et al., 2017; Lee et al., 2017)

77   and the 2018 Hualien earthquake (M_L 6.2) (Hsu et al., 2018). Detailed parameters of the two

78   earthquakes are listed in Table 1. Finally, we verified the reliability of the seismic hazard

79   forecasting maps by comparing them with real ground motion data recorded by the P-alert network.

80

81   **2   Data**

82   **2.1 Central Weather Bureau Seismic Network catalog**

83   We used data from the seismic network catalog maintained by the Central Weather Bureau (CWB)

84   of   Taiwan   (R.O.C.)   (https://www.cwb.gov.tw/V7e/earthquake/seismic.htm   and

85   http://gdms.cwb.gov.tw/index.php, last accessed July 2018). The completeness magnitude (M_c) of

86   this catalog is estimated at approximately 2.0 in local magnitude (M_L) (Wu et al., 2008c). In an

87   analysis of focal depth, Wu et al. (2008b) observed that the focal depth of approximately 80% of

88   earthquakes was shallower than 30 km. Accordingly, we used M_L 2.0 and 30 km as the threshold

89   of magnitude and focal depth, respectively, to select the events to be used in the PI calculation.

90

2.2 P-alert Network

We used the ground motion recordings from the P-alert network to verify the effectiveness of the real-time seismic hazard assessments from our model. The National Taiwan University (NTU) commenced developing the P-alert real-time strong motion network for EEW purposes in 2010 with the support of the Ministry of Science and Technology (MOST) (Wu, 2015). The devices of the P-alert network can record real-time three-component acceleration signals, and publish alerts when the peak initial-displacement amplitude (*Pd*) or the peak ground acceleration (PGA) exceeds predefined thresholds (Wu et al., 2013; Wu, 2015; Wu et al., 2016b). Today, there are more than 600 P-alert stations in Taiwan, most located in elementary schools (Wu et al., 2013; Yang et al., 2018). We mainly adopted the P-alert waveform database maintained by the Taiwan Earthquake Research Center (TEC), and we used the data from the NTU as an auxiliary catalog (data from the P-alert network can be downloaded from the Data Center of the TEC at http://palert.earth.sinica.edu.tw/db/ [last accessed July 2018] or by contacting Prof. Yih-Min Wu at drymwu@ntu.edu.tw for access to the NTU catalog).

The distribution of the P-alert network is still not uniform (see Fig. 2b and 3b), despite the large number of seismic stations covering Taiwan. Obviously, this could cause problems, which will be discussed later.

109

3 Method

3.1 Pattern Informatics

Phase dynamics is the physical fundamental of the PI method, which describes changes in a system

by rotation of the state vector in the Hilbert space (Rundle et al., 2002; 2003). The evolution of the

state vector in a dynamic fault system is suggested to be related to stress accumulation and release

(Chen et al., 2006). The computational steps we adopted here are a modified version developed by

Chang et al. (2016) and Chang (2018) to improve the spatiotemporal resolution of the PI method.

The research area (119°~123° E, 21°~26° N) was divided into boxes of grid size 0.1°×0.1°, with

each box being indicated by parameter x_i. Because of the M_c and the distribution of the focal

depth (mentioned in Section 2.1), we selected all the events with $M_L \geq 2.0$ and depth ≤ 30 km. In

PI computation, t_1 and t_2 represent the beginning and the end of a change interval, respectively,

with the length of a change interval being 4 years. The start time of calculation, t_0, is defined as

12 years before t_2. Then, t_b is a sampling reference time between t_0 and t_1. The t_b starts

from t_0 and shifts forward 3 days in each calculation until the length of time between t_b and t_1

is a half change interval. The forecasting interval, t_3, starts after t_2 (Chang, 2018). The

seismicity rate in the period t_b to t (t_b to t_1 and t_b to t_2) can be expressed as

126
$$S(x_i, t_b, t) = \frac{1}{t - t_b} \int_{t_b}^{t} n(x_i, t) dt$$

127 (1)

128 We conservatively considered the earthquake number, n, occurring in the x_i and its eight

129 neighboring boxes. The rate change during the change interval can be expressed as

130
$$\Delta S(x_i, t_b, t_1, t_2) = S(x_i, t_b, t_2) - S(x_i, t_b, t_1)$$

131 (2)

132 $S(x_i, t_b, t)$ is a vector in a Hilbert space that records present seismic activity, so that ΔS can be

133 interpreted as an angular drift of S (Rundle et al., 2002; Tiampo 2002). To reduce the time-

134 dependent background seismicity, we used the temporal standard score normalizing ΔS, and

135 obtained $\Delta \tilde{S}$. To compare the high and low levels of seismicity rate change in each grid box at the

136 same t_b, we subsequently used the spatial standard score normalizing $\Delta \tilde{S}$, and obtained $\Delta \hat{S}$. The

137 average of the absolute value at all t_b points in each x_i is

138
$$\Delta s(x_i) = \frac{1}{|\{t_b\}|} \sum_{t_b=t_0}^{t_b} \left| \Delta \hat{S}(x_i, t_b, t_1, t_2) \right|$$

139 (3)

140 Then, the mean squared change in probability

141
$$\Delta P(x_i) = \Delta s^2(x_i)$$

142 was computed (Chen et al., 2005; Chang et al., 2016). We further divided the magnitude range of

143   earthquakes into several segments to separately calculate the relative probabilities $\Delta P(x_i)$. The

144   divided magnitude range is from magnitude 2.0 with window length 0.5, and it shifts forward by

145   0.2 each time. Then, we calculated the relative probability each time, such as $\Delta P(x_i)_{2.0\sim2.5}$,

146     $\Delta P(x_i)_{2.2\sim2.7}$. Finally, we multiplied all the relative probabilities.

147
$$\Delta P_M = \prod \Delta P_{i\sim i+0.5}$$

148   (4)

149   ΔP_M to forecast the occurrence of earthquakes is referred to as the modified pattern informatics

150   method (Chang, 2018). According to Chang et al. (2016), the forecasting interval of the PI method

151   reaches 90 days. Finally, the PI method produced a forecasting probability distribution of seismic

152   sources for $M_L \geq 5.0$ within the forecasting interval.

153

154   **3.2 Real-time PSHA**

155   In the traditional PSHA framework (Cornell, 1968; Wang et al., 2016), the probability of an

156   earthquake occurrence follows the Poisson process and the average recurrence interval for an

157   annual frequency of exceedance can be expressed as

158
$$\upsilon(Z > z) = \sum_{i=1}^{N_S} \dot{N}_i \iint f_{M_i}(m)\, f_{R_i}(r)\, P(Z > z \mid m, r)\, dm\, dr$$

159   (5)

160   where $f_{M_i}(m)$ and $f_{R_i}(r)$ are the probability density functions of magnitude and distance,

161   respectively; $P(Z > z \mid m, r)$ is the conditional probability of ground motion Z exceeding a

162   specified value z for a specific magnitude m and distance r. $\dot{N}_i$ is the annual occurrence rate

163   of earthquakes and is described by the truncated exponential model (Cosentino et al., 1977) and

164   the characteristic earthquake model (Schwartz and Coppersmith, 1984). Finally, to consider all

165   scenarios, the total probability of N_s earthquakes is summarized in a given region.

166

167   In real-time PSHA, the occurrence rate of earthquakes used in the traditional PSHA framework is

168   replaced by seismic forecasting probability to achieve spatiotemporal variability in the hazard

169   assessment. Then, considering the gridded space, real-time PSHA can be expressed as

170
$$v(Z > z) = \sum^{M_s} \sum^{Loc_s} P_{M_i, Loc_i}(m, loc)\, P(Z > z \mid m, loc)$$

171   (6)

172   where $P_{M_i, Loc_i}(m, loc)$, the forecasting probability distribution, is a function of magnitude and

173   location. It specifies an occurrence probability for specific magnitude, M_i, at each spatial location,

174   Loc_i. The summations are to consider the whole of the contribution from any possible magnitude,

175   M_s, and location, Loc_s. We adopted the forecasting probability from the PI method as

176    $P_{M,Loc}(m, loc)$. Loc refers to x_i in the PI method. The forecasting probability of the PI method

177   presents a distribution of cumulative forecasting probability for $M_L \geq 5.0$. We referred to the

178   average character of the Gutenberg-Richter law in Taiwan (Gutenberg and Richter, 1944; Wang

179   et al., 2015) to convert it into the probability density function (PDF). It can be corresponded to the

180   specific magnitude conditions for $P(Z > z \mid m, loc)$. To evaluate the ground motion, we used the

181   GMPE published by Lin et al. (2012),

182     $\ln y = C_1 + F_1 + C_3(8.5 - M_w)^2 + [C_4 + C_5(M_w - 6.3)]\ln\left\{\sqrt{[R^2 + exp(H)^2]}\right\} + C_6 F_{NM}$

183                      $+ C_7 F_{RV} + C_8 \ln\left(\dfrac{V_s 30}{1130}\right)$

184     $F_1 = C_2(M_w - 6.3), \qquad M_w \leq 6.3$

185     $F_1 = -HC_5(M_w - 6.3), \qquad M_w > 6.3$

186   (7)

187   which was also adopted for the Taiwan PSHA in Lee et al. (2017). In Eq. 7, C_1 to C_8 and H are

188   the regression coefficients (Table 2); R is the closest distance (km); F_{NM} and F_{RV} represent the

189   earthquake type, namely $F_{NM} = 1$ and $F_{RV} = 0$ for a normal fault earthquake and $F_{NM} = 0$

190   and $F_{RV} = 1$ for a reverse fault earthquake. In this GMPE, earthquake type is regarded as an

191   important parameter. However, the division of seismic sources in the PI method is no longer based

192   on the geological classification but on the grid box, x_i. Considering that most faults in Taiwan are

193   reverse faults (Shyu et al., 2016), we adopted the reverse fault parameter setting for the entire

194   research area. V_s is eclectically assigned $V_s = 760$. Using the conversion equation from Lin and

195   Lee (2008), which was adopted in Lin (2012), turns M_L into M_w. Finally, the forecasting

196   maximum PGA from real-time PSHA is transferred to seismic intensity according to the seismic

197   intensity scale of the CWB listed in Table 3 (Wu et al., 2003). This implies that the seismic intensity

198   forecasting map presents the maximum seismic intensity that every site will encounter over the

199   following 90 days.

200

201   **3.3 Performance verification**

202   **3.3.1 Receiver Operating Characteristic curve**

203   The Receiver Operating Characteristic (ROC) diagram is a binary classification model used widely

204   as a tool to quantify the performance of earthquake prediction (Holliday et al., 2006; Nanjo et al.,

205   2006; Wu et al., 2016a). We used the ROC diagram as an objective quantitative indicator to

206   evaluate the performance of the seismic forecasting probability computed by the PI method. For

207   each box, x_i, there are four situations (parameters) when comparing forecasting hotspots and

208   target earthquakes. Namely, a means any target earthquake in a hotspot, b means no target

209   earthquake in a hotspot, c means no hotspot but at least one target earthquake, d means no target

210   earthquake and no hotspot. The true positive rate (TPR) is defined as $a/(a + c)$ and the false

211   positive rate (FPR) is defined as $b/(b + d)$. The values of a, b, c, and d change with the

212   threshold of forecasting probability and, therefore, TPR and FPR change as well. The value of the

213   area under the ROC curve (AUC) varies between 0 and 1. AUC=1 is a perfect prediction and

214   AUC=0.5 is a random guess. For each PI forecasting map, we generated 1000 random test maps

215   by re-distributing the hotspots randomly over the research area to examine the possibility that a

216   specific distribution of hotspots could be generated by chance. In Fig. 1c and 1d, the blue line is

217   the 95% confidence interval based on two standard deviations. The standard deviation is calculated

218   by the random test results in each bin of the x-axis. The 95% confidence interval helps to

219   differentiate the distributing range of random tests and the significance of the forecasting

220   probability.

221

222   **3.3.2   Average Percent Hit Rate**

223   The success rate of forecasting seismic intensity is a predictive accuracy of classification problems

224   for which the average percent hit rate (APHR) is arguably the most intuitive discrimination

225   measure. The APHR is a rate at which the forecasting data are classified into the correct classes

226   (Sharda and Delen, 2006). We used the APHR to quantify the forecasting performance of real-time

227   seismic hazard assessments. In the APHR, the exact hit rate, which only counts the correct

228   classifications to the exact same class, can be expressed as:

229
$$\text{APHR}_{\text{exact}} = \frac{1}{N} \sum_{i=1}^{g} p_i$$

230   (8)

231   where, in this study, N is the total number of the P-alert stations or the boxes on the forecasting

232   hazard map, g is the total number of seismic intensity classes (=8, according to the CWB seismic

233   intensity scale), and p_i is the total number of samples classified as class i. In the random test, we

234   generated 1000 random tests by randomly re-distributing the forecasting maximum seismic

235   intensity over the research area and the stations to examine the possibility that a specific

236   distribution of the forecast could be generated by chance.

237

238   **4   Results**

239   **4.1 Forecasting earthquake occurrences**

240   Figure 1a and 1b show the forecasting probability maps computed with the PI method, and Fig. 1c

241   and 1d show the corresponding forecasting performance verified by the ROC tests. In the case of

242   the 2016 Meinong earthquake, t_0, t_1, and t_2 are 2004/01/31, 2012/01/31, and 2016/01/31. In

243   the case of the 2018 Hualien earthquake, t_0, t_1, and t_2 are 2006/01/31, 2014/01/31, and

244   2018/01/31. The forecasting intervals of both cases are 90 days after t_2. The cyan stars in Fig. 1a

245   and 1b indicate the main shock of the 2016 Meinong and 2018 Hualian earthquakes, and the largest

246   earthquake in the forecasting interval. The gray circles in Fig. 1a and 1b are the earthquakes with

247   magnitude $M_L \geq 5.0$ in the forecasting interval, with more-detailed information about these

248   earthquakes presented in Table 1. Notably, both main shocks and most large earthquakes are

249   located in or in close proximity to hotspots. Overall, simply from visual inspection, the

250   performance of the PI forecasting probabilities appeared satisfactory.

251

252   In Fig. 1c and 1d, the red curves are located far above the blue curves (95% confidence interval).

253   The AUCs of the red curves are 0.91 and 0.94, and are apparently larger than the AUCs of the blue

254   curves, which are 0.73 and 0.70. The ROC tests quantitatively verified that the performance of the

255   PI forecasting probability was significant, and that these patterns were not generated by chance by

256   the random distribution of hotspots. Both distributions of hotspots were found to be physically

257   meaningful. In view of the above, we were able to use these probability maps as the function of

258   earthquake occurrence rate in subsequent calculations for real-time PSHA.

259

260   **4.2 Real-time PSHA**

261   In Figs 2 and 3, panel (a) shows the map forecasting the maximum seismic intensity estimated by

262   real-time PSHA for the forecasting interval, and panel (b) shows the map indicating the maximum

263   seismic intensity recorded by the P-alert network during the forecasting interval. To ensure that it

264   was the absolute maximum intensity during the forecasting interval, we used only the stations that

265   had recorded all the target events ($M_L \geq 5.0$) in the forecasting interval. Although there are over

266   600 P-alert stations distributed widely in Taiwan, some boxes do not contain any station, e.g., the

267   Central Mountain Range (see Fig. 5a and 5b). Therefore, we estimated the intensities in these

268   boxes by interpolating. However, clearly, our strategy generated an artificial effect, which will be

269   shown later.

270

271   Comparing Fig. 2a and 2b, we suggest that both seismic intensity distributions are remarkably

272   similar. An apparent deviation between the forecasted seismic intensities and the recorded values

273   occurs in southwestern Taiwan, particularly the area closer to the 2016 Meinong main shock.

274   Figure 2c shows the difference in seismic intensity between Fig. 2a and 2b, with the blue and red

275   colors indicating that the forecasting value in a box was underestimated or overestimated,

276   respectively. Most boxes have an intensity difference in the range -1 to 1, but some boxes in

277   southwestern Taiwan are underestimated, with the differences being mostly 2 or even up to 3.

278

279   Comparing Fig. 3a and 3b, we suggest that both seismic intensity distributions are still extremely

280   similar. In this instance, an apparent deviation between the forecasted seismic intensities and the

281   recorded values occurs in southern Taiwan and a part of the southwestern area. Figure 3c shows

282   that most boxes in southern Taiwan have a smaller recorded intensity, with the recorded intensities

283   in a part of southwestern Taiwan being larger than the forecasting values.

284

285   Figure 4 shows the verifications generated by the APHR to quantitatively evaluate the performance

286   of forecasting the seismic intensity. We considered the denominator of two classifications in Eq.

287   8, i.e., the total number of P-alert stations and the total number of boxes in the research area. The

288   results are indicated by "P-alert" and "Map", respectively, in Fig. 4. When comparing forecasting

289   intensity with recorded value, both cases "forecasting = recorded" and "forecasting = recorded +1"

290   indicate "successful forecasting". However, defining the tolerance range, which depends on the

291   perspectives and allowances of different users, is debatable (Hsu et al., 2018). In this study, we

292   tolerated an overestimation of 1 intensity rather than underestimation, as, in relation to the

293   prevention or mitigation of earthquake disasters, "overestimation" was considered preferable to

294   "underestimation".

295

296   All the red lines are above the maximum hit rate of the random tests and higher than 0.5, not to

297   mention the random guesses of the eight choices of the seismic intensity scale on each station or

298   box. This implies that the forecasting ability of the generated seismic intensity maps is significantly

299   effective, and that this satisfactory performance could not be ascribed to chance. Furthermore, both

300   hit rates of the "P-alert" cases are higher than the rates of the "map" cases. However, this result

301   could be attributed to the influence of the artificial effect generated by the interpolation of seismic

302   intensity from the P-alert data of nonuniform distribution. Finally, it should be emphasized that we

303   focused only on earthquakes with $M_L \geq 5$ and we cannot deny the possibility that a $M_L < 5$

304   earthquake could cause large seismic intensity in the near field.

305

306   **5   Discussion**

307   The results of the APHR performance test indicated that the maps and stations employed to

308   forecast the maximum seismic intensity by real-time PSHA were significant and effective. Figure

309   5 is a concretization of the APHR verification and provides more detail. It clearly shows the P-

310   alert station distributions of the "hit" and "not hit", considering only the station-to-station

311   prediction relationship between the forecasts and records. In both instances, most of the P-alert

312   stations are hit (Fig. 5a and 5b), and the hit percentages are distributed along the diagonal and

313   tolerant ranges (Fig. 5c and 5d). However, some locations or stations produced incorrect forecasts.

314   In the case of the 2016 Meinong earthquake, the stations located in southwestern Taiwan do not

315   match the real records and, at high seismic intensities (>3), the forecasting results at some stations

316   are underestimated (Fig. 5c), particularly in the southwestern area. In the case of the 2018 Hualien

317   earthquake, the result from the P-alert APHR appears superior to the former (Meinong), and the

318   distribution of the hit percentage is more concentrated along the diagonal and tolerant ranges (Fig.

319   5d). Nevertheless, the forecasts in southern and part of southwestern Taiwan were not hit.

320

321 In both instances, the differences between the forecasting results and the recorded seismic

322 intensities could be ascribed mainly to three aspects. First, the forecasting model that determines

323 the probability distributions of earthquake occurrences is critical in real-time PSHA. If the

324 probability distribution misses or is a false alarm in somewhere, it directly leads to inaccurate

325 forecasts in real-time PSHA. In the PI results, some differences were located on the hotspots with

326 relatively higher probability, e.g., the area in 22.6º to 23ºN and 120.9º to 121.3ºE in Fig. 1a, and

327 22.7º to 23.1ºN and 120.4º to 120.8ºE in Fig. 1b. Compared with the locations of the earthquakes,

328 these hotspots shifted slightly and it appeared acceptable. However, in the results of real-time

329 PSHA, this led to underestimation of the maximum seismic intensity in the area close to the

330 epicenters and overestimation in the area without any earthquake events but with high probability

331 of earthquake occurrence. For instance, in the case of the 2018 Hualien earthquake, the maximum

332 seismic intensity in the southwestern area was underestimated and that in the southern area

333 overestimated (see Figs 3 and 5b). Therefore, in real-time PSHA, a more accurate and precise

334 forecasting model would facilitate obtaining results that are more positive. Furthermore, even if

335 the PI results performed well in the ROC test, the PI method still needed improvement.

336

337 Second, the evaluation of earthquake ground motion is subject to the limitations of GMPEs. We

338   adopted the GMPE produced by Lin et al. (2012), whose data ($M_L \geq 5.0$) within 50 km represent

339   less than 14% of all the data for regression of GMPE. Therefore, when there is shortage of data in

340   the near field and, for larger events, in the regression of GMPEs, the applicability of GMPEs

341   become limited (Edwards and Fäh, 2014). Accordingly, the limited applicability of GMPEs

342   probably caused the deviation in evaluation of the seismic intensity forecasting maps, e.g., the

343   underestimation of the areas around the two main shocks (Figs 2c and 3c). Furthermore, it is

344   difficult to properly and comprehensively evaluate the site effect in GMPEs, but it dramatically

345   affects the behavior of seismic waves. For example, the amplitudes in the Meinong earthquake

346   were amplified extending along the northwest (in Fig. 2b) because the Western Plain is composed

347   of thick and low-velocity sedimentary deposits (see Fig. 4 in Lee et al., 2016). Consequently, the

348   site effect leads to underestimation in the seismic intensity forecast (Figs 2c and 5a).

349

350   In addition, the directivity effect plays a vital role in the distribution of ground motion. As regards

351   the main shocks in the two study cases, the rupture characteristic had a strong directivity effect

352   that caused significant amplification of the ground motion along the rupture direction (Lee et al.,

353   2016; Hsu et al., 2018). However, basically, GMPEs indicate the statistical distribution of PGA

354   generated by all the data at the same radical distance without considering the possible effect of

355   rupture directivity. As a result, GMPEs are only able to provide the ground motion estimation of

356 radial extension. Furthermore, the forecasting model does not include information on the rupture

357 direction. Therefore, we suggest that some differences along the rupture direction could be

358 ascribed to this effect.

359

360 **6   Conclusion**

361 This study presents a method to achieve real-time seismic hazard assessment by replacing the static

362 seismic rate, i.e., the truncated and characteristic earthquake models, with the time-dependent

363 seismic source probability of the PI method. With regard to this time-dependent seismic source

364 probability, ROC tests verified quantitatively that the performance of the PI forecasting

365 probabilities in the forecasting intervals was quite effective. Therefore, we were able to use the

366 significant probability distributions as the function of the earthquake occurrence rate, $P(m, loc)$,

367 in real-time PSHA. The hit rates of our seismic intensity forecasting maps generated with real-

368 time PSHA outperformed the random guesses and were higher than 0.5 for both the Meinong and

369 the Hualien earthquakes. Therefore, we suggest that real-time PSHA maps are effective forecasting

370 tools, and their satisfactory performance cannot be attributed to coincidence. We demonstrated that

371 real-time seismic hazard assessment was attainable and could be realized and updated by time-

372 dependent seismic source probability.

373

374 In future, different time-dependent seismic source probability models of earthquake occurrences

375 could be introduced to provide estimation that is more accurate and robust. In addition, a possible

376 improvement to our results could be from estimated PGA distribution, not only by means of state-

377 of-the-art machine learning tools for an extensive databank of the P-alert network but also by

378 physics-based numerical simulations (PBS) of seismic ground motion, instead of empirical

379 GMPEs. Presumably, a real-time forecasting map of seismic intensity would enable governments

380 or businesses to prepare efficiently for earthquake disasters. Furthermore, the seismicity intensity

381 scale based on PGA is related to the vulnerability level of buildings, which will also change over

382 time because of degradation and upgrades (e.g., obsolescence, retrofitting actions, and climate

383 events). Therefore, real-time PSHA and change in vulnerability should be considered when

384 assessing seismic risk fluctuation with time.

385

386 ***Code and data availability***. Earthquake catalog can be found on the Central Weather Bureau

387 website   (https://www.cwb.gov.tw/V7e/earthquake/seismic.htm  and

388 http://gdms.cwb.gov.tw/index.php). Data from the P-alert network can be downloaded from the

389 Data Center of the TEC at http://palert.earth.sinica.edu.tw/db/ or by contacting Prof. Yih-Min Wu

390 at drymwu@ntu.edu.tw for access to the NTU catalog.

391

392   *Author contributions*. YSS conducted this study and wrote the main program of real-time PSHA

393   and the article. HCL and CCC provided the main idea and helpful feedback, and revised the paper.

394   LYC and CCC provided expertise in the Pattern Informatics model and wrote the modified PI code.

395   ZKY helped simulate the ground motion.

396

397   *Competing interests*. The authors declare that they have no conflict of interest.

398

399   **Acknowledgments**

400   The authors are grateful for supports from the Ministry of Science and Technology (ROC, Taiwan),

401   the Department of Earth Sciences at National Central University (ROC, Taiwan) and the

402   Earthquake-Disaster & Risk Evaluation and Management Center (E-DREaM) at National Central

403   University.

404

405   **References**

406   Chang, L.-Y., Chen, C.-c., Wu, Y.-H., Lin, T.-W., Chang, C.-H., and Kan, C.-W.: A Strategy for

407   a Routine Pattern Informatics Operation Applied to Taiwan, Pure Appl. Geophys., 173, 235-

408   244, doi:10.1007/s00024-015-1079-9, 2016.

409   Chang, L.-Y.: A study on an improved pattern informatics method and the soup-of-group model

410   for earthquakes. Doctoral dissertation, Department of Earth Sciences, National Central

411   University, Taiwan, R. O. C., 2018.

412 Chen, C.-c., Rundle, J. B., Holliday, J. R., Nanjo, K. Z., Turcotte, D. L., Li, S.-C., and Tiampo, K.

413   F.: The 1999 Chi-Chi, Taiwan, earthquake as a typical example of seismic activation and

414   quiescence, Geophys. Res. Lett., 32, L22315, doi:10.1029/2005GL023991, 2005.

415 Chen, C.-c., Rundle, J. B., Li, H.-C., Holliday, J. R., Turcotte, D. L., and Tiampo, K. F.: Critical

416   point theory of earthquakes: Observation of correlated and cooperative behavior on

417   earthquake fault systems, Geophys. Res. Lett., 33, L18302, doi:10.1029/2006GL027323,

418   2006.

419 Chen, H.-J., Chen, C.-c., Ouillon, G., and Sornette, D.: Using geoelectric field skewness and

420   kurtosis to forecast the 2016/2/6, M L 6.6 Meinong, Taiwan Earthquake, Terr. Atmos. Ocean.

421   Sci., 28, 745-761, doi:10.3319/TAO.2016.11.01.01, 2017.

422 Cooper, J. D.: Letter to editor, San Francisco Daily Evening Bulletin, Nov. 3, 1868.

423 Cornell, C. A.: Engineering seismic risk analysis, Bull. Seismol. Soc. Am., 58, 1583-1606, 1968.

424 Cosentino, P., Ficarra, V., and Luzio, D.: Truncated exponential frequency-magnitude relationship

425   in earthquake statistics, Bull. Seismol. Soc. Am., 67, 1615-1623, 1977.

426 Edwards, B., and Fäh, D.: Ground motion prediction equations, ETH-Zürich, 2014.

427 Holliday, J. R., Rundle, J. B., Turcotte, D. L., Klein, W., Tiampo, K. F., and Donnellan, A.: Space-

428   Time Clustering and Correlations of Major Earthquakes, Phys. Rev. Lett., 97, 238501,

429   doi:10.1103/PhysRevLett.97.238501, 2006.

430 Gutenberg, B. and Richter, C. F.: Frequency of earthquakes in California, Bull. Seism. Soc. Am.,

431   34, 185–188, 1944.

432 Hsu, T. Y., Lin, P. Y., Wang, H. H., Chiang, H. W., Chang, Y. W., Kuo, C. H., Lin, C. M., and

433   Wen, K. L.: Comparing the Performance of the NEEWS Earthquake Early Warning System

434   Against the CWB System During the 6 February 2018 Mw 6.2 Hualien Earthquake, Geophys.

435   Res. Lett., 45, 6001-6007, doi:10.1029/2018GL078079, 2018.

436 Iervolino, I., Chioccarelli, E., and Convertito, V.: Engineering design earthquakes from

437   multimodal hazard disaggregation, Soil Dynam. Earthquake Eng., 31, 1212-1231,

438   https://doi.org/10.1016/j.soildyn.2011.05.001, 2011.

439 Lee, S.-J., Yeh, T.-Y., and Lin, Y.-Y.: Anomalously Large Ground Motion in the 2016 ML 6.6

440   Meinong, Taiwan, Earthquake: A Synergy Effect of Source Rupture and Site Amplification,

441   Seismol. Res. Lett., 87, 1319-1326, doi:10.1785/0220160082, 2016.

442 Lee, Y.-T., Wang, Y.-J., Chan, C.-H., and Ma, K.-F.: The 2016 Meinong earthquake to TEM

443   PSHA2015, Terr. Atmos. Ocean. Sci., 28, 703-713, doi:10.3319/TAO.2016.12.28.02, 2017.

444 Lin, P.-S., Lee, C.-T.: Ground-Motion Attenuation Relationships for Subduction-Zone

445   Earthquakes in Northeastern Taiwan, Bull. Seism. Soc. Am., 98 (1): 220–240.

446   doi: https://doi.org/10.1785/0120060002, 2008.

447 Lin, P.-S., Hsie, P.-S., Lee, Y.-R., Cheng, C.-T., and Shao, K.-S.: The research of probabilistic

448   seismic hazard analysis and geological survey of nuclear power plant: Construction of ground

449   motion prediction equation for response spectra., Commission Report of the Institute of

450   Nuclear Energy Research, Atomic Energy Council, Executive Yuan. (in Chinese), 2012.

451 Nanjo, K. Z., Holliday, J. R., Chen, C.-c., Rundle, J. B., and Turcotte, D. L.: Application of a

452   modified pattern informatics method to forecasting the locations of future large earthquakes

453   in   the   central   Japan,   Tectonophysics,   424,   351-366,

454   https://doi.org/10.1016/j.tecto.2006.03.043, 2006.

455 Rundle, J. B., Klein, W., Tiampo, K., and Gross, S.: Linear pattern dynamics in nonlinear threshold

456   systems, Phys. Rev. E, 61, 2418-2431, doi:10.1103/PhysRevE.61.2418, 2000.

457 Rundle, J. B., Tiampo, K. F., Klein, W., and Sá Martins, J. S.: Self-organization in leaky threshold

458   systems: The influence of near-mean field dynamics and its implications for earthquakes,

459   neurobiology,   and   forecasting,   Proc.   Nat.   Acad.   Sci.,   99,   2514-2521,

460   doi:10.1073/pnas.012581899, 2002.

461 Rundle, J. B., Turcotte, D. L., Shcherbakov, R., Klein, W., and Sammis, C.: Statistical physics

462   approach to understanding the multiscale dynamics of earthquake fault systems, Rev.

463   Geophys., 41, 1019, doi:10.1029/2003RG000135, 2003.

464   Senior Seismic Hazard Analysis Committee (SSHAC): Recommendations for probabilistic

465   seismic hazard analysis: guidance on uncertainty and use of experts, US Nuclear Regulatory

466   Commission Washington, DC, 1997.

467   Schwartz, D. P., and Coppersmith, K. J.: Fault behavior and characteristic earthquakes: Examples

468   from the Wasatch and San Andreas Fault Zones, J. Geophys. Res., 89, 5681-5698,

469   doi:10.1029/JB089iB07p05681, 1984.

470   Sharda, R., and Delen, D.: Predicting box-office success of motion pictures with neural networks,

471   Expert Syst. Appl., 30, 243-254, https://doi.org/10.1016/j.eswa.2005.07.018, 2006.

472   Shyu, J. B. H., Chuang, Y.-R., Chen, Y.-L., Lee, Y.-R., and Cheng, C.-T.: A New On-Land

473   Seismogenic Structure Source Database from the Taiwan Earthquake Model (TEM) Project

474   for Seismic Hazard Analysis of Taiwan, Terr. Atmos. Ocean. Sci., 27, 311-323,

475   doi:10.3319/TAO.2015.11.27.02(TEM), 2016.

476   Tiampo, K. F., Rundle, J. B., McGinnis, S., Gross, S. J., and Klein, W.: Mean-field threshold

477   systems and phase dynamics: An application to earthquake fault systems, Europhys. Lett., 60,

478   481, doi:10.1209/epl/i2002-00289-y, 2002.

479   Wang, J.-H., Chen, K.-C., Leu, P.-L., and Chang, J.-H.: b-Values Observations in Taiwan: A

480   Review, Terr. Atmos. Ocean. Sci., 26, 475-492, doi:10.3319/TAO.2015.04.28.01(T), 2015.

481   Wang, Y.-J., Chan, C.-H., Lee, Y.-T., Ma, K.-F., Shyu, J. B. H., Rau, R.-J., and Cheng, C.-T.:

482   Probabilistic Seismic Hazard Assessment for Taiwan, Terr. Atmos. Ocean. Sci., 27, 325-340,

483   doi:10.3319/TAO.2016.05.03.01, 2016.

484   Wu, Y.-M.: Progress on development of an earthquake early warning system using low cost

485   sensors, Pure Appl. Geophys. 172, 2343–2351, doi: 10.1007/s00024-014-0933-5, 2015.

486   Wu, Y.-M., Shin, T.-C., and Tsai, Y.-B.: Quick and reliable determination of magnitude for

487   seismic early warning, Bull. Seismol. Soc. Am., 88, 1254-1259, 1998.

488   Wu, Y.-M., Teng, T.-l., Shin, T.-C., and Hsiao, N.-C.: Relationship between Peak Ground

489   Acceleration, Peak Ground Velocity, and Intensity in Taiwan, Bull. Seismol. Soc. Am., 93,

490   386-396, doi:10.1785/0120020097, 2003.

491   Wu, Y.-H., Chen, C.-c., and Rundle, J. B.: Detecting precursory earthquake migration patterns

492   using the pattern informatics method, Geophys. Res. Lett., 35, L19304,

493   doi:10.1029/2008GL035215, 2008a.

494   Wu, Y.-H., Chen, C.-c., and Rundle, J. B.: Precursory seismic activation of the Pingtung (Taiwan)

495   offshore doublet earthquakes on 26 December 2006: A pattern informatics analysis, Terr.

496   Atmos. Ocean. Sci., 19, 743-749, doi:10.3319/TAO.2008.19.6.743(PT), 2008b.

497   Wu, Y.-M., Chang, C.-H., Zhao, L., Teng, T.-L., and Nakamura, M.: A Comprehensive Relocation

498   of Earthquakes in Taiwan from 1991 to 2005, Bull. Seismol. Soc. Am., 98, 1471-1481,

499   doi:10.1785/0120070166, 2008c.

500   Wu, Y.-M., Chen, D.-Y., Lin, T.-L., Hsieh, C.-Y., Chin, T.-L., Chang, W.-Y., and Li, W.-S.: A

501   High-Density Seismic Network for Earthquake Early Warning in Taiwan Based on Low Cost

502   Sensors, Seismol. Res. Lett., 84, 1048-1054, doi:10.1785/0220130085, 2013.

503   Wu, Y.-H., Chen, C.-c., and Li, H.-C.: Conditional Probabilities for Large Events Estimated by

504   Small Earthquake Rate, Pure Appl. Geophys., 173, 183-196, doi:10.1007/s00024-014-1019-

505   0, 2016a.

506   Wu, Y.-M., Liang, W.-T., Mittal, H., Chao, W.-A., Lin, C.-H., Huang, B.-S., and Lin, C.-M.:

507   Performance of a Low-Cost Earthquake Early Warning System (P-Alert) during the 2016 ML

508   6.4 Meinong (Taiwan) Earthquake, Seismol. Res. Lett., 87, 1050-1059,

509   doi:10.1785/0220160058, 2016b.

510   Yang, B. M., Huang, T.-C., and Wu, Y.-M.: ShakingAlarm: A Nontraditional Regional Earthquake

511   Early Warning System Based on Time-Dependent Anisotropic Peak Ground-Motion

512   Attenuation Relationships, Bull. Seismol. Soc. Am., 108, 1219-1230,

513   doi:10.1785/0120170105, 2018.

514

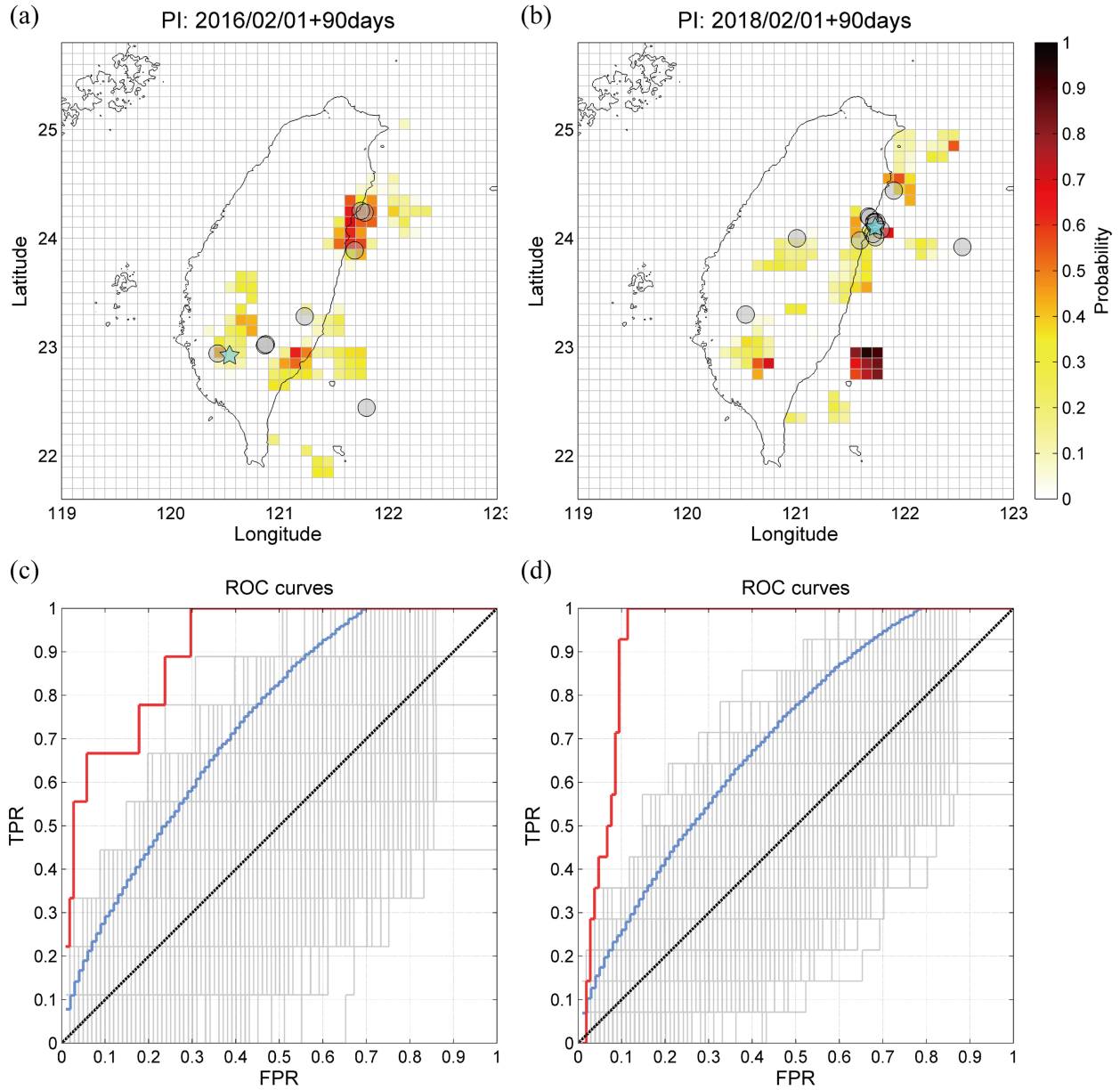



Figure 1. Panels (a) and (b) show the forecasting probability maps of the Meinong earthquake and

the Hualien earthquake, respectively. Panels (c) and (d) are the ROC curves of (a) and (b),

respectively. Red, gray, blue, and black curves represent the forecasting probability map, random

tests, 95% confidence interval, and the average of random tests, respectively.

519

520

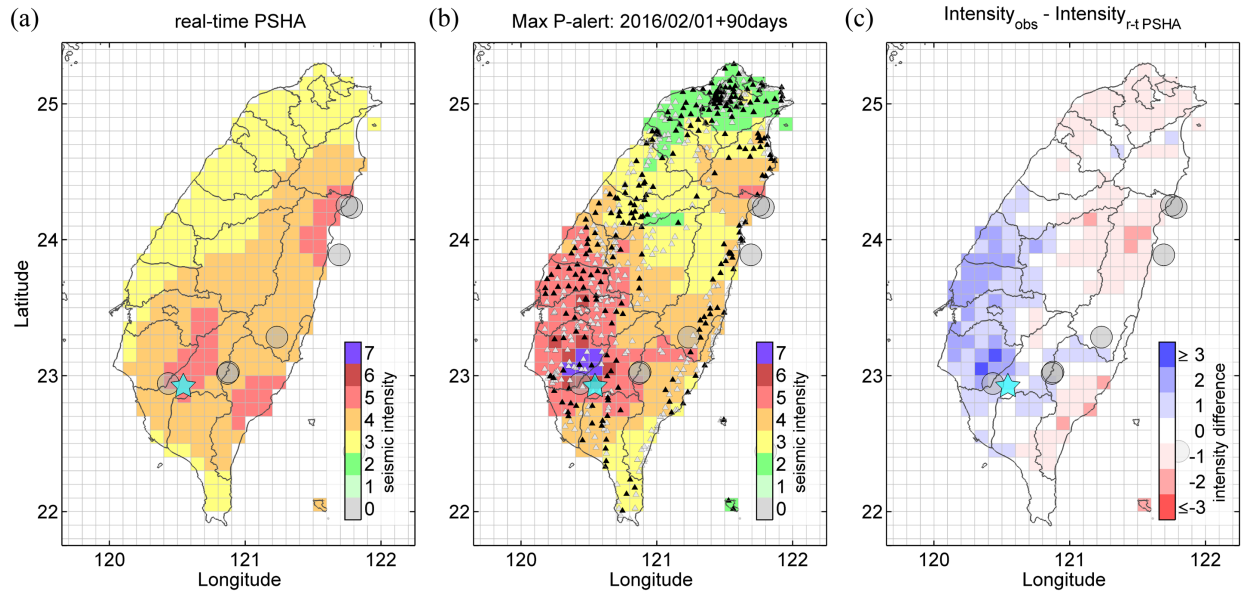



Figure 2. The 2016 Meinong earthquake: (a) Map of forecasted maximum seismic intensity by

real-time PSHA. The forecasting interval of seismic intensity is 90 days. (b) Map of maximum

seismic intensity recorded by the P-alert network. Black and white triangles indicate the P-alert

stations that we used and did not use, respectively, in the verification. (c) Difference in seismic

intensity between the forecast and the record. The cyan star represents the Meinong earthquake,

and the gray circles represent the earthquakes with magnitude $M_L \geq 5$ in this forecasting interval.

527

528

529

530

531

532

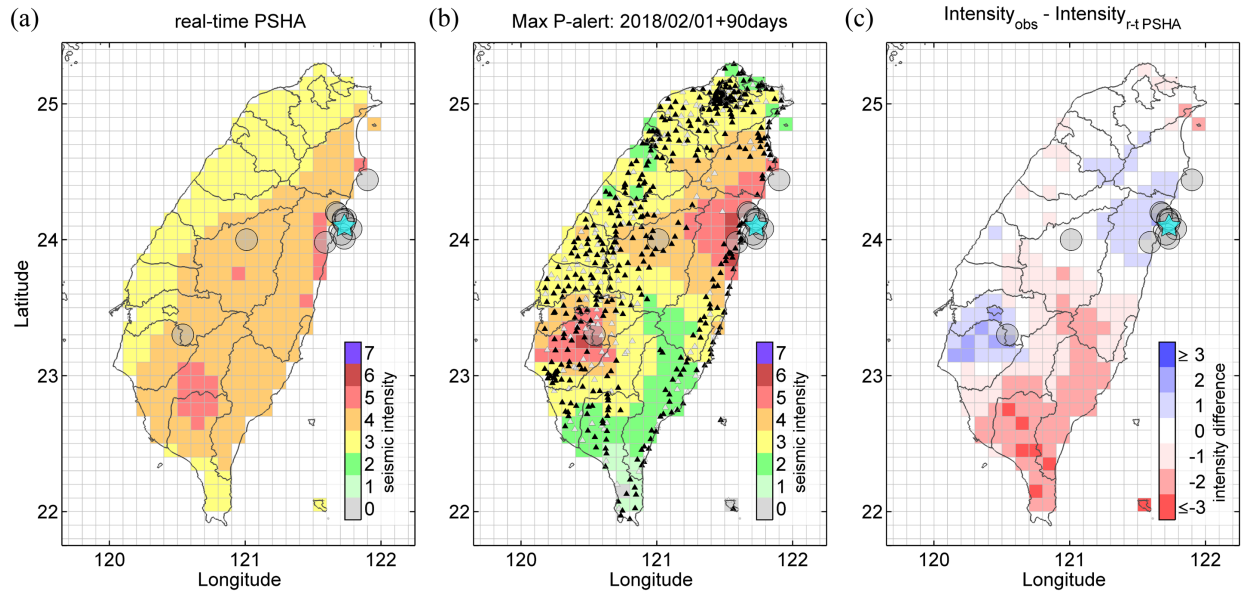



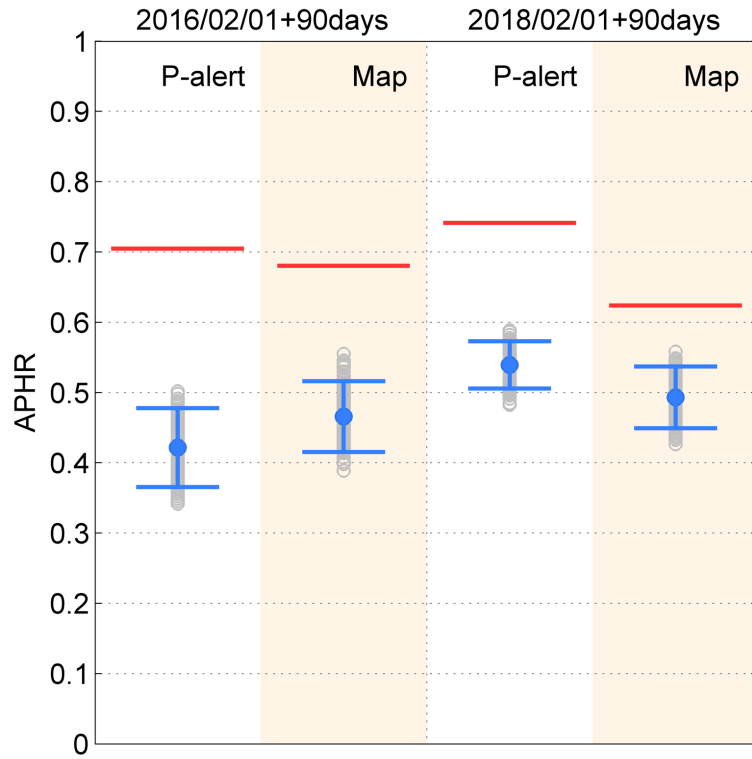
Figure 3. The 2018 Hualian earthquake: (a) Map of forecasting maximum seismic intensity. (b)

Map of maximum seismic intensity recorded by the P-alert network. (c) Difference in seismic

intensity between the forecast and the record. The cyan star represents the Hualian earthquake.

536

537

538

539

540   **Figure 4.** Performance test of APHR. The red line indicates the forecasts of real-time PSHA, the

541   gray circle indicates the result of a random test by randomly re-distributing seismic intensities, the

542   blue error bar indicates the interval with two standard deviations over all random tests.

543

544

545

546

547

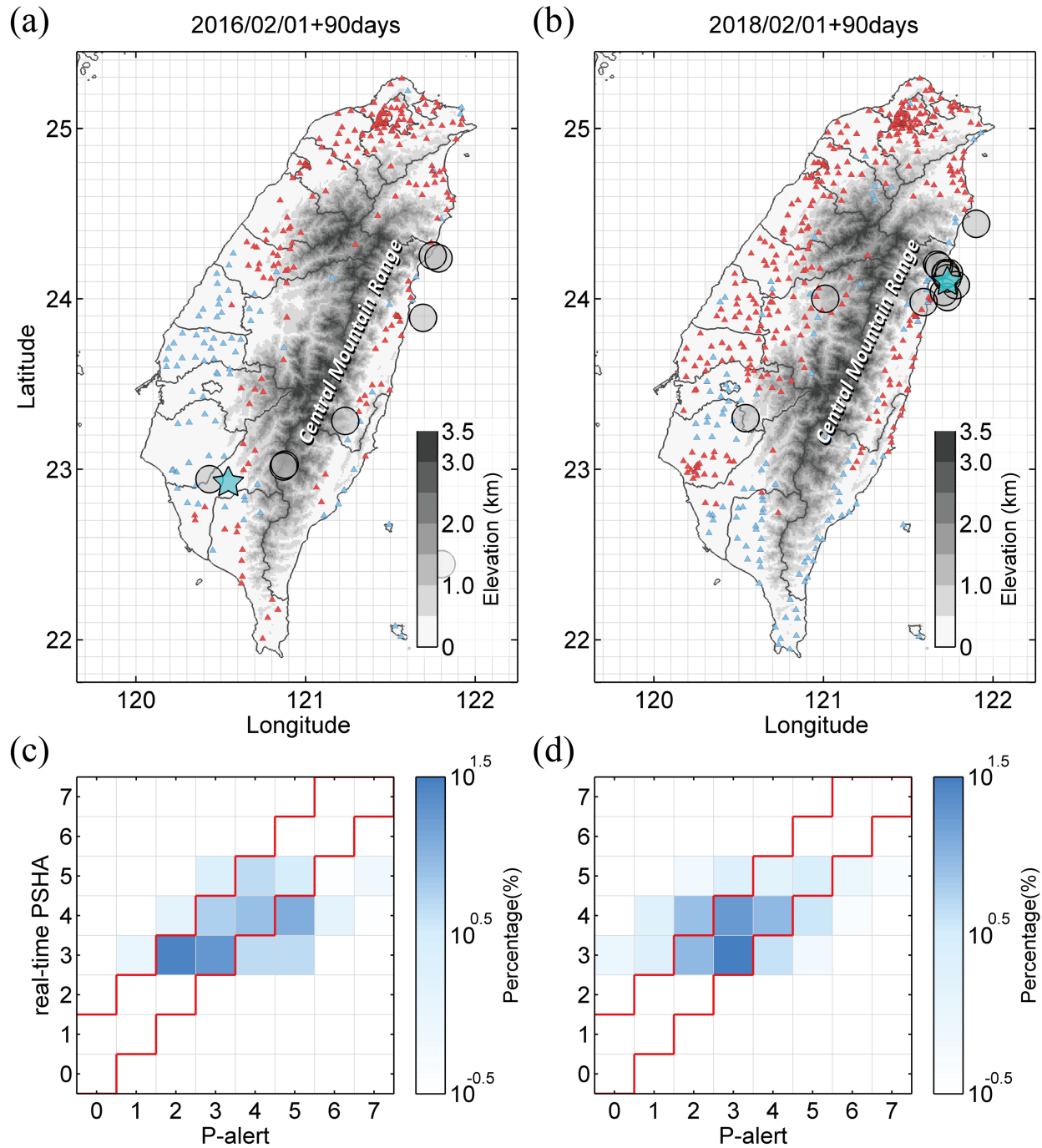



Figure 5. Panels (a) and (b) are the P-alert station distributions indicating "hit" and "not hit". The

red and blue triangles represent "hit" and "not hit", respectively. Panels (c) and (d) are the

distributions of the hit percentage for the 2016 Meinong and 2018 Hualian earthquakes,

551   respectively. The red line area represents the acceptable prediction range.

552

Table 1. Earthquakes occurring in the forecast interval.

(a) Meinong case: 2016/02/01~2016/05/01

Date	Hour	Min.	Lon.	Lat.	Depth	M_L	P-alert	Num.
02/05	**19**	**57**	**120.54**	**22.92**	**14.64**	**6.60**	**TEC**	**338**
02/05	19	58	120.43	22.94	18.10	5.26	Nan	Nan
02/09	00	47	121.69	23.89	5.69	5.12	TEC	341
02/18	01	09	120.87	23.02	5.44	5.27	TEC	357
02/18	01	18	120.88	23.03	4.26	5.13	TEC	357
04/16	10	55	121.80	22.44	11.83	5.22	TEC	436
04/27	15	17	121.78	24.24	11.94	5.67	NTU	424
04/27	15	27	121.75	24.25	12.99	5.13	NTU	425
04/27	18	19	121.23	23.28	15.21	5.52	NTU	423

(b) Hualian case: 2018/02/01~2018/05/02

Date	Hour	Min.	Lon.	Lat.	Depth	M_L	P-alert	Num.
02/04	13	12	121.67	24.20	15.10	5.10	TEC	543
02/04	13	56	121.74	24.15	10.60	5.80	TEC	519
02/04	13	57	121.68	24.19	11.10	5.10	Nan	Nan
02/04	14	13	121.72	24.15	10.30	5.50	TEC	517
02/05	15	58	121.72	24.14	10.00	5.00	TEC	522
02/06	**15**	**50**	**121.73**	**24.10**	**6.30**	**6.20**	**TEC**	**520**
02/06	15	53	121.59	23.98	5.10	5.00	TEC	520
02/06	18	00	121.73	24.12	6.70	5.30	TEC	516
02/06	18	07	121.71	24.04	4.20	5.30	TEC	516
02/06	19	15	121.73	24.01	5.70	5.40	TEC	516
02/07	15	21	121.78	24.08	7.80	5.80	TEC	523
02/25	18	28	121.90	24.44	17.70	5.20	TEC	533
03/20	09	22	120.54	23.30	11.20	5.30	TEC	539
03/29	00	17	121.01	24.00	11.10	5.00	NTU	388

553

554

| 04/23 | 17 | 10 | 122.53 | 23.92 | 19.30 | 5.10 | NTU | 381 |

555 "P-alert" indicates that the P-alert recording was obtained from the Taiwan Earthquake Research

556 Center (TEC) or the National Taiwan University (NTU). "Num." indicates the number of recording

557 stations. "Nan" indicates no P-alert data were recorded from either TEC or NTU, even if the event

558 was recorded by CWB. Bold font represents the Meinong and Hualian earthquakes.

559

560 **Table 2.** Coefficients in the GMPE.

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	H
1.3979	0.3700	0.0000	-1.2273	0.2086	-0.1934	0.1122	-0.4359	1.4877

561

562 **Table 3.** Seismic intensity scale of CWB.

Intensity Scale		Ground Acceleration 563 (cm/s², gal)
Micro	0	<0.8
Very minor	1	0.8~2.5
Minor	2	2.5~8.0
Light	3	8~25
Moderate	4	25~80
Strong	5	80~250
Very Strong	6	250~400
Great	7	≥400