Pure Appl. Geophys. 173 (2016), 153–164

DOI 10.1007/s00024-014-1030-5

Pure and Applied Geophysics

Quantitative Analysis of Seismicity Before Large Taiwanese Earthquakes Using the G-R Law

H.-C. Li,[1] C.-H. Chang,[2] and C.-C. Chen[1]

Abstract—Seismicity has been identified as an example of a natural, nonlinear system for which the distribution of frequency and event size follow a power law called the "Gutenberg–Richter (G-R) law." The parameters of the G-R law, namely *b*- and *a*-values, have been widely used in many studies about seismic hazards, earthquake forecasting models, and other related topics. However, the plausibility of the power law model and applicability of parameters were mainly verified by statistical error σ of the *b*-value, the effectiveness of which is still doubtful. In this research, we used a newly defined *p* value developed by Clauset *et al.* (*Power-Law Distributions in Empirical Data*, SIAM Rev. *51*, 661–703, 2009) instead of the statistical error σ of the *b*-value and verified its effectiveness as a plausibility index of the power-law model. Furthermore, we also verified the effectiveness of K–S statistics as a goodness-of-fit test in estimating the crucial parameter M_c of the power-law model.

1. Introduction

Since seismicity was identified as an example of a natural nonlinear system for which the distribution of frequency and event size in a sufficiently long interval follows the Gutenberg-Richter (G-R) law (i.e. a power law), the parameters of the G-R law, including completeness magnitude $M_c$and *b*- and *a*-values fitted based on M_c, have been used widely in numerous studies focusing on seismic hazards, earthquake forecasting models, and other related topics (Huang *et al.* 2001; Chen 2003; Schorlemmer *et al.* 2005; Huang 2006; Tsai *et al.* 2006; Wu and Chiao 2006; Bhattacharya *et al.* 2011; Rundle *et al.* 2011). The ability to estimate the parameters of the power-law model objectively was definitely crucial to the reliability of these studies; therefore, several techniques were developed for this purpose. They were usually combinations of a fitting method, mainly maximum-likelihood or least-square, with a goodness-of-fit test. Woessner and Wiemer (2005) tested the efficiencies of several techniques systematically using natural seismic catalogs. However, the plausibility of a fitted power-law model to the observed data set was mainly evaluated by the statistical error σ of the *b*-value (Marzocchi and Sandri 2003), which was ruled out as an effective index in our research.

Clauset *et al.* (2009) developed an elaborate method combining the maximum-likelihood method with a goodness-of-fit test based on the Kolmogorov–Smirnov (K–S) statistics to compute the *p* value to evaluate the plausibility of fitted power-law model. In their test, the earthquake catalog of California was suggested to follow a power law while a proper magnitude cut-off was used. A noticeable point that should be emphasized here is that the *p* value defined by Clauset et al. had no relationship with the widely used *p* value in Omori's law. We used their method and another widely accepted maximum-likelihood method developed by Aki (1965) to the Central Weather Bureau (CWB) catalog, which is another important catalog in the world, to investigate precursory anomalies before large earthquakes based on potential power-law behaviors. Besides using K–S statistics as a goodness-of-fit test, we also combined the "Goodness-of-Fit test" (GFT, Wiemer and Wyss 2000) with Aki's method as a comparison to investigate the effectiveness of the *p* value as an index of model plausibility. Through our tests, we verified that the effectiveness of the *p* value was better than the statistical error σ and also concluded that the application of K–S statistics was crucial in correctly

[1] Institute of Geophysics, National Central University, Jhongli, Taiwan 320 ROC. E-mail: klempervinsky@gmail.com
[2] Seismological Center, Central Weather Bureau, Taipei, Taiwan 320 ROC.

identifying power-law behaviors of seismicity rather than the fitted methods used.

2. Seismic Data

The history of instrumental observation for seismicity in Taiwan and nearby islands started in 1897, when the first seismometer was set up in Taipei by the Japanese. Since 1984 the Central Weather Bureau (CWB) has upgraded the instruments to an electromagnetic type and increased the coverage of the network by building more observation stations. After combining the telemetric seismic network of the Institute of Earth Sciences (IES) to the original seismic network of the CWB in 1991, a new real-time digital observation network was formed called the Central Weather Bureau Seismic Network (CWBSN) and is still maintained by the CWB.

In this research, we selected three large ($M \geqq 6.0$) on-land earthquakes in the Taiwanese area, including (1) the 1999 Chichi earthquake in Nantou, (2) the 2003 Chengkung earthquake in Taitung, and (3) the 2010 Jiashan earthquake in Kaohsiung, and summarize their important parameters in Table 1 (Kao and Chen 2000; Kuochen *et al.* 2007; Hsu *et al.* 2011). For each earthquake, we used 12 intervals of interval lengths 30, 60, 90,…, 360 days, respectively, to select seismicity. Each interval started from the previous day before the occurrence of the main-shock and extended back to the past, so the aftershocks were excluded in the data sets used in parameter computation. The purpose of using various interval lengths was to eliminate potential biases caused by using preferential interval lengths subjectively.

3. Methods

3.1. Maximum-Likelihood Estimate by Clauset et al. (2009)

We merely addressed major concepts about the method developed by Clauset *et al.* (2009) here; further details and derivations can be referred to in other articles (Newman 2005) if necessary. A quantity x obeys a power law if it is drawn from a probability distribution

$$p(x) \propto x^{-\alpha} \quad (1)$$

where α is a constant parameter called the "exponent" or "scaling parameter" of the distribution. In usual cases, there must be some lower bound $x_{\min}$, below which the power-law behavior no longer exists and the distribution of x belongs to other types. On the other hand, there are also some statistical fluctuations to larger values of x which are caused by the essential property of enormously low frequency of rare events. Newman (2005) derived that once $x_{\min}$ was known, the probability density of a continuous variable x was

$$p(x) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha} \quad (2)$$

and the exponent of power law distribution could be easily estimated by

$$\alpha = 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}} \right]^{-1}. \quad (3)$$

The quantity x_i, $i = 1, 2, \ldots, n$ gives the measured values of x. The cumulative density function (CDF) was

Table 1

Parameters of selected large on-land earthquakes in the Taiwan region

	1999 Chichi	2003 Chengkung	2010 Jiashan
Date (UT)	20 Sep. 1999	10 Dec. 2003	4 Mar. 2010
Magnitude	M_w 7.6, M_L 7.3	M_w 6.8, M_L 6.4	M_w 6.3, M_L 6.4
Epicenter	23.85°N, 120.82°E	23.06°N, 121.39°E	22.97°N, 120.71°E
Focal depth (km)	8	17.73	22.64

$$P(x) = \int_{x}^{\infty} p(x')dx' = \left(\frac{x}{x_{min}}\right)^{-\alpha+1} \tag{4}$$

which also followed the power law, but with a different exponent, $\alpha - 1$, which is the b-value of the G-R law in the field of earth science. An estimate of the expected statistical error σ in Eq. (3) is given by

$$\sigma = \sqrt{n}\left[\sum\nolimits_{i=1}^{n} \ln\frac{x_i}{x_{min}}\right]^{-1} = \frac{\alpha - 1}{\sqrt{n}}. \tag{5}$$

The relation between the estimated maximum amplitude of an earthquake on seismograph and magnitude is

$$A(\Delta) \propto 10^{M_L}, \tag{6}$$

where A is the maximum amplitude of an event recorded at an epicentral distance Δ. In order to use Clauset's method for a real earthquake catalog, we had to transform all documented magnitudes in an interval of interest to the analogies of amplitude using Eq. 6 and use the transformed quantity as the variable x in Eq. (2), rather than using the values of magnitude directly.

3.2. Maximum-Likelihood Estimate by Aki (1965)

The widely accepted method in the field of earth science to estimate the parameters in the G-R law is the maximum-likelihood method developed by Aki (1965). In this method, the probability density function of an earthquake with a magnitude greater than M_c is assumed to obey an exponential distribution expressed by

$$f(M, b') = b'e^{-b'(M-M_c)},\ M \geq M_c \tag{7}$$

where $b' = b/log_{10}e$. Suppose that we have a sample of n earthquakes with magnitudes M_1, M_2,..., M_n. The parameter b is estimated by

$$b = \frac{log_{10}e}{\langle M\rangle - M_c + \Delta M/2} \tag{8}$$

where $\langle M\rangle$ is the average magnitude of the used data set and ΔM is the binning width. In Aki's method, the parameter M_c is implicitly assumed to be known. However, we have to determine a suitable M_c value to select a data set and to compute a corresponding b-value in a real application. A statistical error in Eq. (8) is also estimated by Eq. (5) using b instead of $\alpha - 1$.

3.3. Goodness-of-Fit Test

A noticeable point is that the exponent in each method completely depends on an estimated data minimum, i.e. $x_{\min}$ and M_c in Eqs. (3) and (8) respectively. Thus, a good fitness-of-fit test which collaborates with the above-mentioned methods to determine a suitable minimum is critical. Clauset et al. used the Kolmogorov–Smirnov (K–S) statistics as the goodness-of-fit test, which is the maximum distance between the CDFs and a fitted model:

$$D = \max_{x \geq x_{\min}} |S(x) - P(x)|. \tag{9}$$

Here, $S(x)$ is the CDF of the observation data with values greater than or equal to $x_{\min}$, and $P(x)$ is the CDF for the fitted power-law model that best fits the data in the region $x \geq x_{\min}$. The $x_{\min}$ which minimizes D is the estimated lower bound of power-law behavior. We used the same combination in this research and called it C + KS. Furthermore, the K–S statistics was also used in the analysis applying Aki's method and this combination was called A + KS.

Another test used in the analysis applying Aki's method was the "Goodness-of-Fit test (GFT)", which was developed by Wiemer and Wyss (2000) to compute the difference between an observed frequency-magnitude distribution (FMD) and a synthetic distribution of a fitted model. At first we estimated the b- and a-values of the G-R law as a function of cut-off magnitude M_{co} by Eq. (8), then we generated synthetic distributions using the estimated b- and a-values for $M \geq M_{co}$. The residual between an observed and a synthetic distribution is computed by

$$R(a, b, M_{co}) = \frac{\sum_{M_{co}}^{M_{\max}} |B_i - S_i|}{\sum_i B_i} \tag{10}$$

where B_i and S_i are the observed and synthetic cumulative number of earthquakes in each magnitude bin. If the residual value R is smaller, the similarity between an observed and a synthetic distribution is better. Wiemer and Wyss defined the M_c as the first M_{co} at which R was less than a fixed confidence level, 0.1 or 0.05.

We set the precision of each candidate of M_c in 0.1 rather than 0.01 because 0.01 was too small considering the measurement precision of magnitude of a small earthquake; therefore, it did not make sense to use an unrealistic value as small as 0.01.

3.4. *Plausibility of the Hypothesis of the Power Law*

Using the methods addressed in Sects. 3.1 and 3.2, we can estimate the parameters of hypothesized probabilistic models for each observed data set. However, it was critical for us to know whether the hypothetic power-law model was plausible for the observed data set. In this research, we adopted the approach developed by Clauset *et al.* (2009) to generate p values to quantify the plausibility of the hypothesis model.

The first step was to generate a synthetic data set. For the observed data set in an interval, the total number of observations was n, and the number of observations greater than or equal to an estimated $x_{\min}$ (i.e. M_c in this research) was n_{tail}. With the probability n_{tail}/n we generated a random number x_i using the probabilistic model with $x \geq x_{\min}$ and estimated b-value. On the other hand, with the probability $1 - n_{\text{tail}}/n$, we selected x_i at random from the observed data values that conformed the condition $x < x_{\min}$. Repeating the process for all $i = 1,\ldots,$ n, we generated a synthetic data set which followed the hypothesis model above $x_{\min}$, but had the same distribution of the observed data below $x_{\min}$. Based on Clauset's estimation, we generated 2,500 synthetic data sets to allow the p value to be accurate to about two decimal digits. Once the synthetic data sets were generated, we used the methods addressed in Sects. 3.1 or 3.2 to estimate M_c, b-value, and D or R of each synthetic data set.

The p value was defined as the fraction of synthetic data sets with D or R greater than the value calculated by real observed data. If most of the synthetic data sets had values of D or R greater than the observed data sets and, therefore, made the p value exceed the threshold value 0.1 (based on Clauset's definition), the hypothesis model was more suitable to describe the observed data set than most of the synthetic data sets. We concluded the hypothesis model to be plausible for the observed data set. Otherwise, if $p < 0.1$, the hypothesis model was ruled out as a plausible one for the observed data set. However, we should emphasize again that a high p value merely verified the plausibility of the hypothesis model to the observed data. It did not necessarily mean that the hypothesis model was the "correct" distribution for the data. It is still possible to find other hypothesis models that are plausible to the same data set.

4. *Results*

Tables 2, 3, and 4 list all estimated parameters of optimal fitted models, including M_c, σ, a-, b-, and p value, of the data set in each interval. To compare the a-value of each interval, we normalized all rates of earthquake to "number per 30 days" and calculated the a-value at magnitude 2.0.

4.1. *Chichi earthquake, Nantou*

By Table 2, we found that the methods C + KS and A + KS estimated the same values of M_c except the cases using 30-, 270-, and 300-day intervals. In the cases with the same M_c, the b-values estimated by C + KS and A + KS for the same data set had tiny differences smaller than 0.01. A similar situation also existed for several cases of the Chengkung earthquake in 2003 and the Jiashan earthquake in 2010. These tiny differences revealed an essential difference while using Eqs. 3 or 8 to compute the b-value. We illustrate the case using a 90-day interval in Fig. 1a–c, in which the inverted triangles indicate the observed cumulative FMD and the open circles indicate the predicted cumulative FMD of fitted power-law model using the method C + KS, A + KS, and A + G, respectively. The patterns of predicted FMD in Fig. 1a by C + KS and 1b by A + KS were highly similar by visual inspection and both were verified as plausible power-law models to the selected data set by their p values exceeding 0.1. Actually, the power-law model was shown to be plausible for all cases using the method C + KS or A + KS because their p values were all greater than 0.1.

Table 2

Parameters estimated using the data prior to the 1999 Chichi earthquake

Interval (day)	a-value			b-value			M_c			p Value			σ		
	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G
30	2.7949	2.7270	2.6857	0.9872	0.9178	0.8587	2.7	2.4	2.1	0.81	0.59	0.45	0.0876	0.0607	0.0430
60	2.7216	2.7157	2.6421	0.9338	0.9238	0.8338	2.6	2.6	2.1	0.98	0.90	0.34	0.0548	0.0542	0.0310
90	2.6884	2.6840	2.6288	0.8760	0.8672	0.7942	2.5	2.5	2.1	0.64	0.77	0.07	0.0379	0.0375	0.0244
120	2.7229	2.7176	2.6476	0.8877	0.8787	0.7990	2.6	2.6	2.2	0.47	0.68	0.08	0.0357	0.0353	0.0228
150	2.7132	2.7089	2.6434	0.8699	0.8613	0.7807	2.5	2.5	2.1	0.26	0.34	0.03	0.0282	0.0280	0.0182
180	2.7091	2.7047	2.6483	0.8808	0.8719	0.8005	2.5	2.5	2.1	0.65	0.85	0.08	0.0264	0.0261	0.0170
210	2.7149	2.7104	2.6428	0.8955	0.8864	0.8019	2.5	2.5	2.1	0.83	0.53	0.04	0.0249	0.0246	0.0159
240	2.7233	2.7186	2.6597	0.9123	0.9028	0.8322	2.5	2.5	2.2	0.72	0.48	0.05	0.0237	0.0235	0.0167
270	2.7383	2.7297	2.6740	0.9211	0.9085	0.8406	2.6	2.5	2.2	0.54	0.27	0.03	0.0248	0.0221	0.0157
300	2.7557	2.7531	2.6977	0.9260	0.9187	0.8504	2.5	2.6	2.2	0.17	0.22	0.02	0.0209	0.0230	0.0146
330	2.7759	2.7699	2.7123	0.9387	0.9287	0.8580	2.6	2.6	2.2	0.31	0.33	0.02	0.0222	0.0219	0.0139
360	2.7894	2.7833	2.7242	0.9390	0.9290	0.8571	2.6	2.6	2.2	0.31	0.31	0.02	0.0209	0.0207	0.0131

Each interval started from the day before the main shock occurred and extended back to the past. The a-values in Tables 2, 3, and 4 were estimated at magnitude 2.0

C Clauset, *A* Aki, *KS* Kolmogorov–Smirnov statistics, *G* goodness-of-fit test by Woessner and Wiemer

Table 3

Parameters estimated using the data prior to the 2003 Chengkung earthquake

Interval (day)	a-Value			b-Value			M_c			p Value			σ		
	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G
30	2.8014	2.8014	2.7871	0.9115	0.9021	0.8678	2.0	2.0	1.9	0.15	0.28	0.21	0.0362	0.0359	0.0317
60	2.8523	2.8513	2.8228	0.9380	0.9279	0.8646	2.1	2.1	1.9	0.70	0.41	0	0.0277	0.0274	0.0215
90	2.8559	2.8550	2.8449	0.9327	0.9228	0.9038	2.1	2.1	2.0	0.23	0.59	0.38	0.0224	0.0221	0.0197
120	2.8942	2.8522	2.8395	1.0055	0.9401	0.9162	2.7	2.1	2.0	0.86	0.02	0.06	0.0404	0.0196	0.0174
150	2.8949	2.8871	2.8421	0.9915	0.9803	0.9018	2.7	2.7	2.0	0.57	0.37	0.01	0.0352	0.0348	0.0153
180	2.8916	2.8784	2.8702	0.9831	0.9721	0.9116	3.2	3.2	2.0	0.74	0.66	0	0.0559	0.0553	0.0137
210	2.9182	2.9058	2.9085	0.9526	0.9423	0.8885	3.2	3.2	2.0	0.94	0.92	0	0.0466	0.0461	0.0118
240	2.9155	2.9027	2.9186	0.9688	0.9581	0.9273	3.2	3.2	2.1	1	0.99	0	0.0455	0.0450	0.0127
270	2.8897	2.8773	2.9145	0.9498	0.9395	0.9269	3.2	3.2	2.1	0.97	0.93	0	0.0422	0.0418	0.0120
300	2.9102	2.8974	2.8879	0.9651	0.9545	0.8896	3.2	3.2	2.0	0.91	0.84	0	0.0406	0.0402	0.0101
330	2.9132	2.9004	2.8861	0.9703	0.9596	0.8951	3.2	3.2	2.0	0.95	0.89	0	0.0391	0.0386	0.0097
360	2.9077	2.8949	2.8847	0.9701	0.9594	0.8981	3.2	3.2	2.0	0.82	0.70	0	0.0376	0.0372	0.0094

Table 4

Parameters estimated using the data prior to the 2010 Jiashan earthquake

Interval (day)	a-Value			b-Value			M_c			p Value			σ		
	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G	C + KS	A + KS	A + G
30	2.5208	2.5126	2.7520	0.8487	0.8405	0.9418	3.0	3.0	2.0	0.93	0.82	0.03	0.1238	0.1226	0.0396
60	2.7675	2.7675	2.3036	0.9914	0.9802	0.7796	2.0	2.0	3.3	0	0	0.06	0.0290	0.0286	0.1248
90	2.9594	2.9544	2.9137	1.0519	1.0393	0.9817	2.4	2.4	2.2	0	0	0	0.0327	0.0323	0.0248
120	2.7360	2.7253	2.8988	0.8866	0.8776	0.9720	3.2	3.2	2.2	0.68	0.81	0	0.0647	0.0640	0.0216
150	2.8492	2.8401	2.8812	0.9445	0.9344	0.9193	2.9	2.9	2.1	0.62	0.41	0	0.0423	0.0418	0.0166
180	2.9166	2.9119	2.8652	1.0134	1.0017	0.9279	2.4	2.4	2.1	0	0	0	0.0230	0.0227	0.0156
210	2.8581	2.7516	2.8879	0.9818	0.9155	0.9753	2.9	3.3	2.2	0.28	0.58	0	0.0382	0.0573	0.0166
240	2.7878	2.9285	2.9059	0.8879	0.9631	0.9258	3.3	2.2	2.1	0.74	0	0	0.0479	0.0146	0.0128
270	2.9558	2.9538	2.9316	0.9479	0.9377	0.9018	2.2	2.2	2.1	0	0	0	0.0131	0.0129	0.0114
300	2.8306	2.8189	2.9155	0.8912	0.8821	0.9063	3.3	3.3	2.1	0.62	0.60	0	0.0411	0.0407	0.0111
330	2.9553	2.9263	2.9034	0.9867	0.9453	0.9081	2.6	2.2	2.1	0.05	0	0	0.0196	0.0122	0.0107
360	2.9171	2.9088	2.8948	0.9537	0.9434	0.8997	2.8	2.8	2.1	0.03	0.05	0	0.0231	0.0228	0.0103

On the other hand, the pattern of the model shown in Fig. 1c is quite different from Fig. 1a, b. The observed data began to deviate from the open circles slightly at magnitude 3.3 and more apparently at magnitude 3.8. A similar pattern, in which there were apparent gaps between the observed FMD and predicted FMD by model, existed generally in the cases using A + G for estimating parameters. The values of M_c estimated by A + G were much smaller than the values estimated by other methods. Almost all p values estimated by A + G were smaller than 0.1, and thus revealed that the statistical distribution of corresponding data sets did not obey the power law except two cases using 30- and 90-day intervals. In contrast to the performances of other methods, we suggested that the method A + G failed to determine appropriate data sets of which the statistical distributions followed the hypothetic G-R law before the Chichi earthquake.

4.2. Chengkung Earthquake, Taitung

The methods C + KS and A + KS estimated the same M_c values and, therefore, very close b-values in all cases except the one using the 120-day interval. But the values of M_c estimated by A + G which were in the range 1.9–2.1 were much smaller than the values by other methods for the cases using interval length longer than 120 days. The unique difference of M_c between the method C + KS and A + KS in the case using the 120-day interval was noticeable, and the results are shown in Fig. 2a–c. By visual inspection, the open circles in Fig. 2a are very close to the inverted triangles until magnitude 4.9, when the observed FMD fluctuates enormously and, therefore, clearly deviates from the predicted FMD. On the other hand, the open circles in Fig. 2b began to deviate from the observed FMD at magnitude 3.2, and the open circles in Fig. 2c also clearly deviated from the observed FMD at magnitude 3.1. The p values revealed that the power-law model was plausible for the data set used in Fig. 2a, but implausible for the data sets in Fig. 2b, c with smaller M_c. The method C + KS was the only one which could uncover potential power-law behaviors in the case using the 120-day interval.

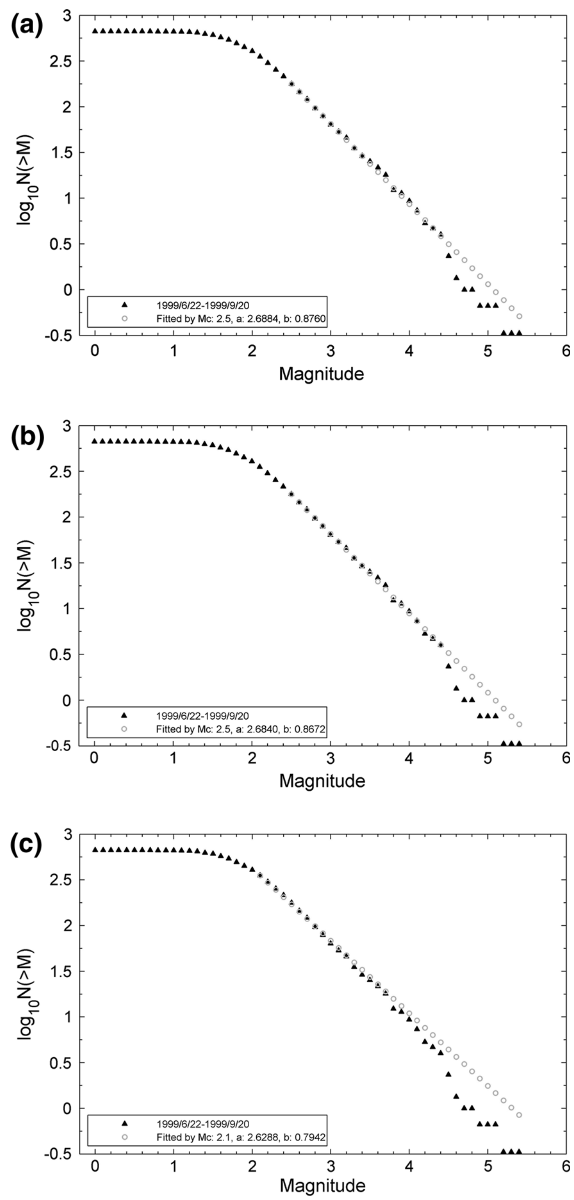

Figure 1
Observed cumulative FMD (*inverted triangle*) and corresponding fitted power-law models (*open circle*) before the 1999 Chichi earthquake. The interval length is 30 days and the method used to fit model is C + KS for (**a**), A + KS for (**b**), and A + G for (**c**)

4.3. Jiashan Earthquake, Kaohsiung

A noticeable property of all the observed FMDs before the 2010 Jiashan earthquake was the apparent discontinuities of the FMD curves. We illustrated this property by the case using a 150-day interval in Fig. 3a–c. The method C + KS and A + KS estimated the same M_c value, 2.9, in contrast to a smaller value, 2.1, by A + G. In Fig. 3a and b, the open circles were located very close to the inverted triangles in the magnitude ranging from 2.9 to 4.5. However, the open circles in Fig. 3c deviated clearly from the observed FMD at magnitude greater than 2.9. Based on inspection of the p value, we suggest that the power-law model is implausible for the data set used in Fig. 3c, while a smaller M_c value was estimated and, therefore, included cumbersome smaller earthquakes.

Although the method C + KS and A + KS identified appropriate data sets for which the power-law model was plausible, both methods also estimated smaller M_c values in several cases, including 60-, 90-, 180-, 270-, 330-, and 360-day intervals. Unlike the cases with larger M_c, all p values of the cases with smaller M_c were below the threshold 0.1 and, therefore, rejected the hypothesis of a power-law model. We illustrated the cases using a 240-day interval in Fig. 3d–f, in which the discontinuity of observed FMD was not so clear by visual inspection. The method C + KS estimated a larger M_c value 3.3 and p value 0.74 in contrast to the smaller M_c values and zero p values by other methods. The inverted triangles deviated from the open circles in magnitudes ranging from 3 to 4.5 in Fig. 3f and merely ranging from 3.2 to 3.4 in Fig. 3e. In fact, the deviations of the observed FMD from predicted values were not very apparent in Fig. 3e and f, but the power-law model was still rejected as plausible for the data sets by the zero p values. We suggest that these cases sufficiently illustrate the advantage in identifying the behaviors of seismicity with an objective quantitative index such as the p value rather than by visual inspection.

5. Discussions

We had three topics to discuss in this research, which were (1) the reliability of statistical error σ as an indicator of model plausibility, (2) the effectiveness of three fitted methods to identify potential power-law behaviors of seismicity, and (3) the implications of precursory seismic anomalies identified by appropriate fitted methods. We discussed the reliability of p value and statistical error σ first.

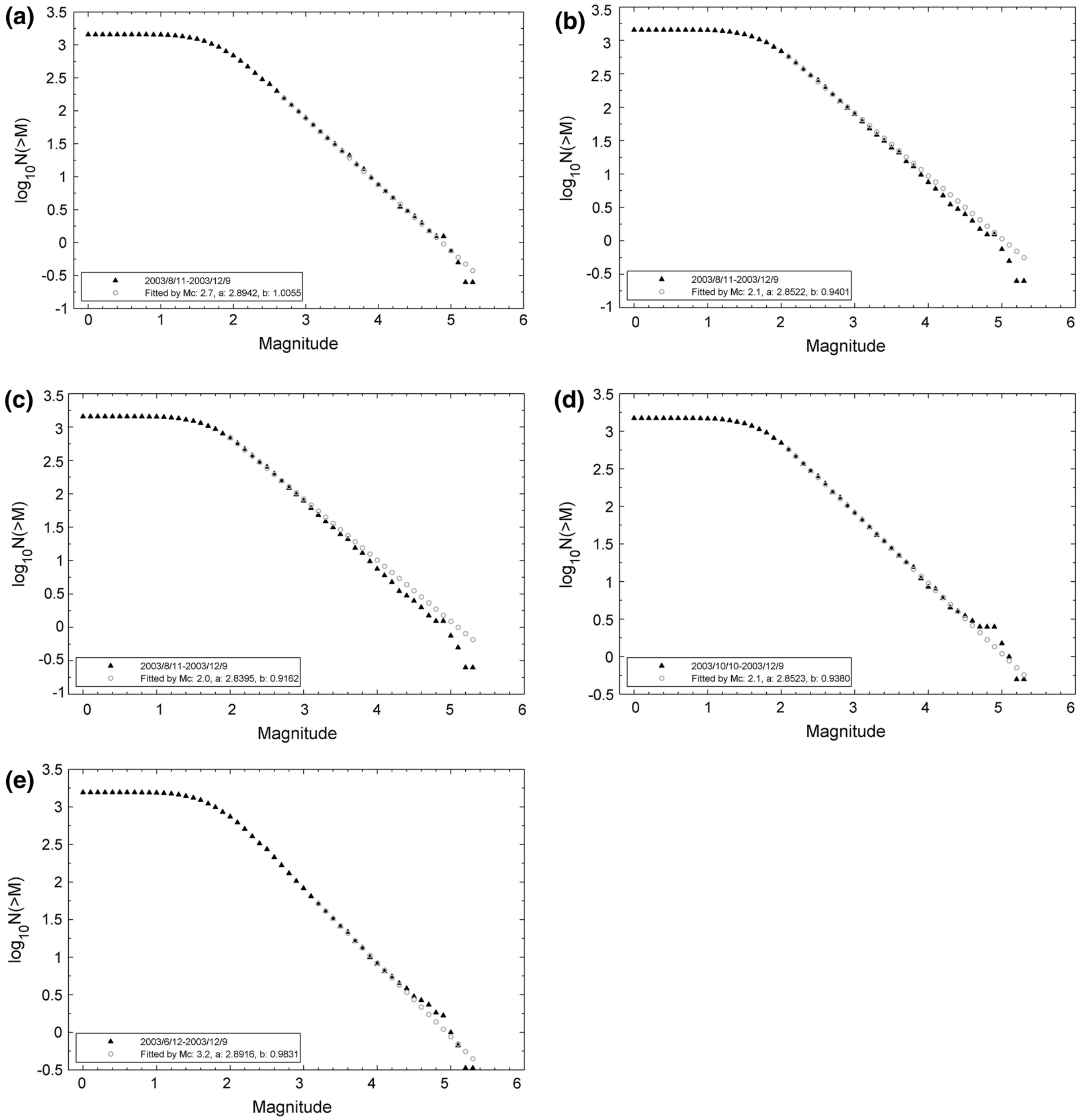

Figure 2

Observed cumulative FMD and corresponding fitted power-law models before the 2003 Chengkung earthquake. The interval length is 120 days for (**a**) to (**c**), 60 days for (**d**), and 180 days for (**e**). The fitting method is C + KS for (**a**), (**d**), and (**e**), A + KS for (**b**), and A + G for (**c**)

5.1. *Reliability of σ as Indicator of Model Plausibility*

An important parameter in Eq. 5 to estimate the statistical error σ is "number of earthquakes, n" in the denominator. This parameter definitely depends on the length of interval and the value of M_c which were used to select data set. If the values of $\alpha - 1$ in the nominator, i.e. b-value in the G-R law, merely fluctuated slightly between different data sets, the data sets with more data points would

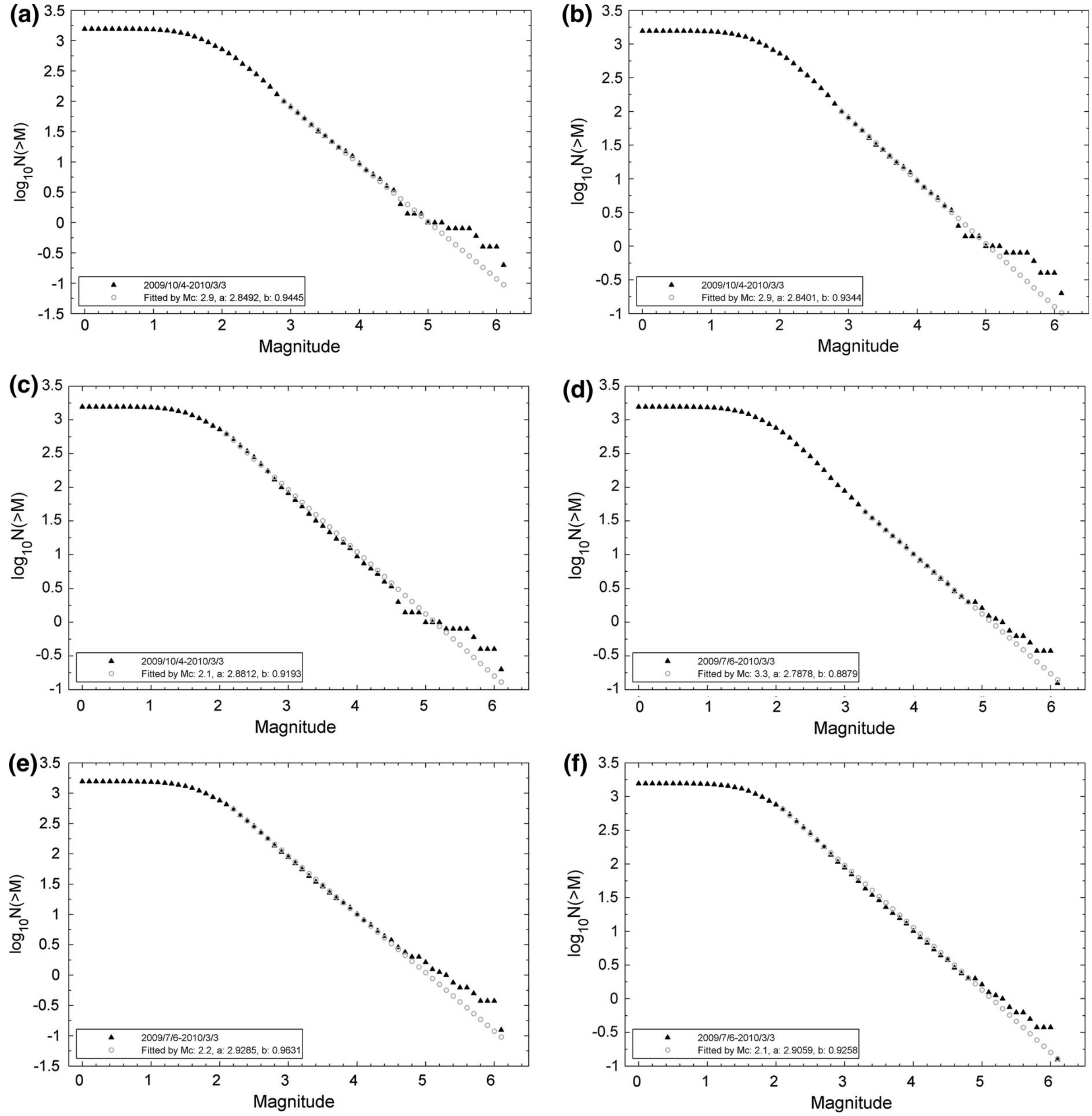

Figure 3

Observed cumulative FMD and corresponding fitted power-law models before the 2010 Jiashan earthquake. The interval length is 150 days for (**a**) to (**c**), and 240 days for (**d**) to (**f**). The fitting method is C + KS for (**a**) and (**d**), A + KS for (**b**), and (**e**), and A + G for (**c**) and (**f**)

generate smaller values of σ. This property could be easily observed in each column of σ using the method A + G in Tables 2, 3, and 4. For example, all M_c values merely fluctuated between 1.9 and 2.1, and the differences of b-value between different time intervals also merely fluctuated smaller than 0.1 in Table 3. The dominant factor affecting the number of data points is the length of time interval used, and we indeed observed that the values of σ decreased regularly with the elongation of time intervals, which also corresponded to the increase of n.

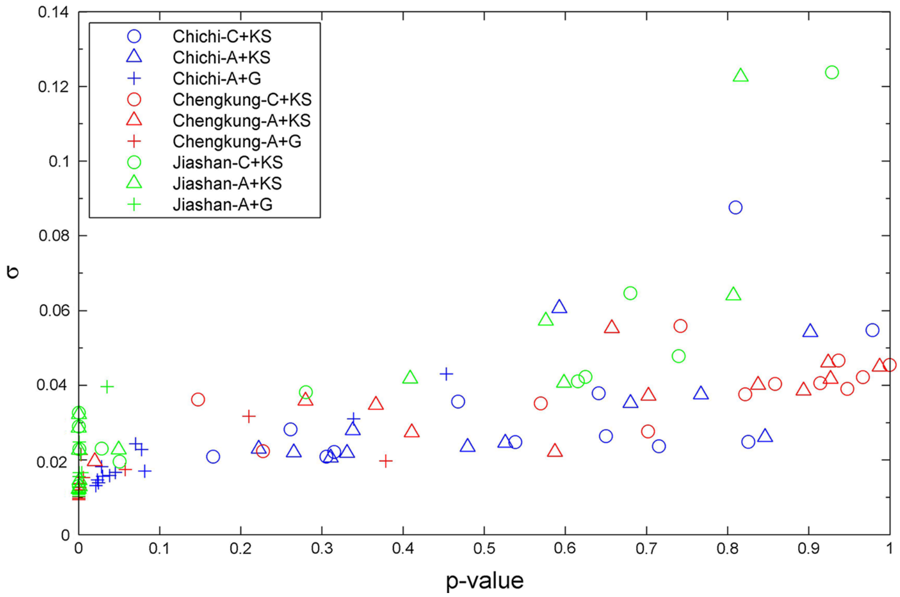

Figure 4

Plot of σ and p value. *Blue*, *red*, and *green* indexes indicate the cases of 1999 Chichi, 2003 Chengkung, and 2010 Jiashan earthquakes, respectively. Indexes marked by *circle*, *triangle*, and *cross* indicate the cases using method C + KS, A + KS, and A + G, respectively

However, smaller σ values did not guarantee the plausibility of the power-law model to a data set. We illustrated this property by considering the cases using the method A + G in Table 3. The data set selected by a 90-day interval generated a value of σ larger than other cases using longer time intervals, but its p value proved that the power-law model was plausible for this data set. Although the data sets using intervals longer 90 days had smaller σ values, the power-law models were ruled out because their p values were below the predefined threshold. A similar situation also could be observed in Tables 2 and 4.

Furthermore, we summarized the distribution between p value and σ in Fig. 4 and observed a very diverse pattern. For example, the values of σ between 0.04 and 0.05 corresponded with the p values ranging from 0.03 to 1. The value of Pearson correlation coefficient between p value and σ was 0.5833, which verified that no apparent linear correlation existed between these two parameters. An immediate consequence was that the statistical error σ could not be a reliable index to verify the plausibility of fitted models. Therefore, we strongly recommended using the p value instead of σ for the purpose of model plausibility.

5.2. *Effectiveness of Fitted Methods for Potential Power-Law Behaviors of Seismicity*

The second topic that we would like to discuss is the effectiveness of the three methods addressed in Sect. 3 in identifying potential power-law behaviors of seismicity. Our definition of effectiveness was very simple, for a case using a specific time interval, if the power-law model was proved to be plausible to either data set selected by two different methods but was ruled out in the data set selected by the third method; we would suggest that the effectiveness of the third method was not comparable to other methods. We had to emphasize that either Clauset's or Aki's maximum-likelihood method could generate a set of G-R law parameters for any input data set. The effectiveness of each method actually depended on the goodness-of-fit method which collaborated with the maximum-likelihood methods to determine the crucial parameter M_c. In this research, both the methods C + KS and A + KS excavated potential power-law behaviors in most cases. There were merely two cases, i.e. the 120-day interval of the 2003 Chengkung earthquake and the 240-day interval of the 2010 Jiashan earthquake, in which the method A + KS failed while C + KS worked. However, the hypothesis of the power-law model was ruled out in 89 % of the data sets selected using A + G and almost all of the M_c values estimated by the GFT were much smaller than the values by the K–S statistics. We suggested that this property corresponded with the observation by Woessner and Wiemer (2005) that the GFT might underestimate the value of M_c. Moreover, we found that the subjectively defined threshold value of residual R in the GFT severely affected the estimation of M_c. Any researcher who attempts to apply this method should be very careful of this affect. Therefore, we suggested that the K–S statistics was a more effective goodness-of-fit method rather than the GFT method for our purpose.

5.3. Precursory Seismic Anomalies

The final topic which we would discuss is actually based on the discussions addressed above. An objective and reliable index to evaluate the plausibility of the power-law model is the most important premise to assess precursory anomalies of seismicity based on the G-R law. With a reliable index such as the p value, it will be helpful in preventing to evaluate seismic anomalies mistakenly using improper parameters such as the b-value of G-R law. We observed both the characteristics of seismic activation and quiescence before three selected large earthquakes in this research with the aid of the p value.

In Fig. 1a and b, the observed FMD (inverted triangles) were located below the predicted FMD (open circles) at magnitude greater than 4.4. In fact, a similar property existed in all of the cases using various interval length except 30 days. We suggested that this deficiency of earthquakes with moderate magnitude lasted at least 1 year before the Chichi main-shock. Similar precursory seismic quiescence had also been reported in previous research by Wu and Chiao (2006) based on the continuously decreasing b-value using monthly seismicity. They attributed this seismic quiescence to the reduction of smaller earthquakes because the monthly numbers of larger earthquakes ($M_L > 4.0$) before the Chichi main-shock were suggested to fluctuate in a normal range consistently. An opposite observation about precursory anomalies before the 1999 Chichi earthquake was made by Chen (2003), who identified a seismic activation of earthquake with magnitude greater than 5 in the interval 1998–September 20, 1999. Chen fitted the cumulative FMD using earthquakes with moderate magnitude, but we had a lot of smaller earthquakes in the optimal data set because of M_c 2.5.

On the other hand, we observed the phenomena of seismic activation before the 2003 Chengkung and 2010 Jiashan earthquake. In Fig. 2a for the 2003 Chengkung earthquake, the observed FMD were slightly larger than the predicted FMD at the magnitude greater than 4.8. We also illustrate two cases using 60- and 180-day intervals in Fig. 2d and e of which both were estimated by the method C + KS. The observed FMD exceeded the predicted values at the magnitude greater than 4.6 in Fig. 2d and 4.3 in Fig. 2e. For the 2003 Chengkung earthquake, it was also reported that there was local activation of a moderate-size earthquake before the main-shock by Wu *et al.* (2008) based on a positive Z value surrounding the rupture region of the Chengkung earthquake. For the 2010 Jiashan earthquake, the observed FMD were obviously larger than the predicted FMD at the magnitude greater than 5.0 in Fig. 3a and 4.8 in Fig. 3d. We also observed similar characteristics in other cases that the power-law model verified to be plausible. We suggest that the 2010 Jiashan earthquakes might be another example of precursory seismic activation based on our observations, but further analyses are still necessary for more related details.

In this research, we revealed the effectiveness of the K–S statistics as a goodness-of-test method in identifying potential power-law behaviors of seismicity through the test using p value. Furthermore, either maximum-likelihood method developed by Clauset or Aki would generate almost the same values of G-R law parameters if a data set was already verified by the K–S statistics that obeyed the power-law model. The plausibility of the hypothesis model and applicability of parameters such as the b-value should always be verified first to reduce the risk of mistakenly evaluating the seismicity.

References

AKI, K. (1965), *Maximum Likelihood Estimates of b in the Formula* log $n = a\text{-}bm$ *and Its Confidence Limits*, Bull. Earthq. Res. Inst. Univ. Tokyo *43*, 237–239.

BHATTACHARYA, P., CHAKRABARTI, B.K., and KAMAL (2011), *A Fractal Model of Earthquake Occurrence: Theory, Simulations and Comparisons with the Aftershock Data*, J. Phys. Conf. Ser. *319*, 012004.

CHEN, C.-C. (2003), *Accelerating Seismicity of Moderate-Size Earthquakes before the 1999 Chi-Chi, Taiwan, Earthquake: Testing Time-Prediction of the Self-Organizing Spinodal Model of Earthquakes*, Geophys. J. Int. *155*, F1–F5.

CLAUSET, A., SHALIZI, C.R. and NEWMAN, M.E. (2009), *Power-Law Distributions in Empirical Data*, SIAM Rev. *51*, 661–703.

HSU, Y.-J., YU, S.-B., KUO, L.-C., TSAI, Y.-C., and CHEN, H.-Y. (2011), *Coseismic Deformation of the 2010 Jiashian, Taiwan Earthquake and Implications for Fault Activities in Southwestern Taiwan*, Tectonophysics *502*, 328–335.

HUANG, Q.-H., SOBOLEV, G.A., and 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.

Huang, Q.-H. (2006), *Search for reliable precursors: A case study of the seismic quiescence of the 2000 western Tottori prefecture earthquake*, J. Geophys. Res *111*, B04301.

KAO, H., and CHEN, W.-P. (2000), *The Chi-Chi Earthquake Sequence: Active, Out-of-Sequence Thrust Faulting in Taiwan*, Science *288*, 2346–2349.

KUOCHEN, H., WU, Y.-M., CHEN, Y.-G., and CHEN, R.-Y. (2007), *2003 Mw6.8 Chengkung Earthquake and Its Related Seismogenic Structures*, J. Asian Earth Sci. *31*, 332–339.

MARZOCCHI, W., and SANDRI, L. (2003), *A Review and New Insights on the Estimation of the b-value and Its Uncertainty*, Ann. Geophys. *46*, 1271–1282.

NEWMAN, M.E. (2005), *Power Laws, Pareto Distributions and Zipf's Law*, Contemp. Phys. *46*, 323–351.

RUNDLE, J.B., HOLLIDAY, J.R., YODER, M., SACHS, M.K., DONNELLAN, A., TURCOTTE, D.L., TIAMPO, K.F., KLEIN, W., and KELLOGG, L.H. (2011), *Earthquake Precursors: Activation or Quiescence?* Geophys. J. Int. *187*, 225–236.

SCHORLEMMER, D., WIEMMER, S., and WYSS, M. (2005), *Variations in Earthquake-Size Distribution across Different Stress Regimes*, Nature *437*, 539–542.

TSAI, Y.-B., LIU, J.-Y., MA, K.-F., YEN, H.-Y., CHEN, K.-S., CHEN, Y.-I., and LEE, C.-P. (2006), *Precursory Phenomena Associated with the 1999 Chi-Chi Earthquake in Taiwan as Identified under iSTEP Program*, Phys. Chem. Earth *31*, 365–377.

WIEMER, S., and WYSS, M. (2000), *Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan*, Bull. Seismol. Soc. Am. *90*, 859–869.

WOESSNER, J., and WIEMER, S. (2005), *Assessing the Quality of Earthquake Catalogs: Estimating the Magnitude of Completeness and Its Uncertainty*, Bull. Seismol. Soc. Am. *95*, 684–698.

WU, Y.-M., and CHIAO, L.-Y. (2006), *Seismic Quiescence before the 1999 Chi-Chi, Taiwan, Mw 7.6 Earthquake*, Bull. Seismol. Soc. Am. *96*, 321–327.

WU Y.-M., CHEN, C.-C., ZHAO, L., and CHANG, C.-H. (2008), *Seismicity Characteristics before the 2003 Chengkung, Taiwan, earthquake*, Tectonophysics *45*, 177–182.

(Received August 31, 2014, revised December 24, 2014, accepted December 29, 2014, Published online January 24, 2015)