

Chapter 3

Multiple Regression Analysis for Estimating Earthquake Magnitude as a Function of Fault Length and Recurrence Interval

Takashi Kumamoto, Kozo Oonishi, Yoko Futagami, and Mark W. Stirling

Abstract Multiple regressions are developed using world earthquake data and active fault data, and the regressions are then evaluated with Akaike's Information Criterion (IEEE Trans Autom Control, 19(6):716–723). The AIC method enables selection of the regression formula with the best fit while taking into consideration the number of parameters. By using parameters relevant to earthquakes and active faults in the regression analyses, we develop a new empirical equation for magnitude estimation as $M_w = 1.13\log L_s + 0.16\log R + 4.62$.

Keywords Multiple regression analysis • Magnitude • Fault length • Recurrence interval

3.1 Introduction

Many empirical equations for estimating earthquake magnitude have been developed in Japan. The most famous of these is the so-called Matsuda's Equation [2], which is widely used for constructing seismic hazard maps of Japan (e.g., The Headquarters for Earthquake Research Promotion [3]). Matsuda's Equation (Eq. 3.1) below is based on 14 earthquakes dated from the 1891 Nobi earthquake (M_{JMA} 8.0: Japan Meteorological Agency magnitude) to the 1970 Southeastern Akita Prefecture earthquake (M_{JMA} 6.2).

T. Kumamoto (✉)

Faculty of Science, Okayama University, Tsushimanaka 3-1-1, Kita-ku, Okayama 700-8530, Japan

e-mail: tkuma@cc.okayama-u.ac.jp

K. Oonishi

Shikoku Electric Power CO., Inc, Marunouchi 2-5, Takamatsu 760-8573, Japan

Y. Futagami

National Research Institute for Cultural Properties, Tokyo, Ueno Park 13-43, Taito-ku, Tokyo 110-8713, Japan

M.W. Stirling

GNS Science, 1 Fairway Drive, Avalon 5010, New Zealand

$$\log L = 0.6M_{JMA} - 2.9 \quad (3.1)$$

In Matsuda's Equation, the fault parameter length L might be regarded as sub-surface. On the contrary, the length of an active fault recognized on the surface from tectonic geomorphology must be used for magnitude estimation prior to the occurrence of the next earthquake. However, a problem occurs in estimating magnitudes of isolated active faults with lengths significantly shorter than the thickness of the seismogenic layer, which is estimated to be roughly 15–20 km in Japan. Thus, another empirical equation for estimating earthquake size is the following empirical equation between seismic moment M_0 and fault area by Irikura and Miyake [4] as:

$$\begin{aligned} S &= 2.23 \times 10^{-15} M_0^{2/3} \text{ for } M_0 < 7.5 \times 10^{25} \text{ dyne-cm} \\ S &= 4.24 \times 10^{-11} M_0^{1/2} \text{ for } M_0 \geq 7.5 \times 10^{25} \text{ dyne-cm} \end{aligned} \quad (3.2)$$

This equation is also widely used especially in strong ground motion prediction in Japan.

However, the question of large uncertainties in the earthquake scaling relation remains in both equations. To resolve the above issues, multiple regressions are developed using world earthquake data and active fault data, and the regressions are then evaluated with AIC (Akaike's Information Criterion, Akaike, 1974). The AIC method enables selection of the regression formula with the best fit while taking into consideration the number of parameters. By using parameters relevant to earthquakes and active faults such as stress drop, average slip rate, and recurrence interval, in the regression analysis, we develop a new empirical equation for magnitude estimation.

3.2 Data

The database used in this paper is compiled from the intraplate earthquake datasets listed below.

1. Earthquake and active fault data of Wells and Coppersmith [5]:

These data were compiled to develop empirical relationships between earthquake magnitude and various fault parameters. The data are for the years 1857–1994, and the number of data items is 244. The data fields include earthquake location, name, date, slip type, magnitude (surface wave magnitude M_s , moment magnitude M_w , seismic moment M_0), subsurface and surface rupture length, fault width, fault area, and maximum/average surface slip amount. Values thought to have low reliability are given in parentheses.

2. Earthquake data of Anderson et al. [6]:

These historical earthquake data were used to estimate earthquake magnitude as a function of the length and slip rate of the causative fault. The earthquake data were collected for the time period 1811–1994, and the number is 43 in total. The data fields include location, Mo, Mw, fault length, and slip rate.

3. Earthquake and active fault data of Mohammadioun and Serva [7]:

The characteristic aspect of this dataset is a list of variations in stress drop for different slip types, including strike-slip, normal, or reverse faults in the 1857–1994 earthquakes worldwide, plus the Umbria earthquake in 1998 (Ms 5.7), the Chi-Chi earthquake in 1999 (Ms 7.6), and the Izmit earthquake in 1999 (Ms 7.4). The number of earthquakes in the dataset is 90, and the data fields include maximum surface slip amount, static stress drop $\Delta\sigma_1$, dynamic stress drop $\Delta\sigma_2$, and the ratio of dynamic stress drop to static stress drop.

4. Earthquake and active fault data of Stirling et al. [8]:

The number of data items is 389, and several scaling laws are developed in this paper to compare instrumental and pre-instrumental data. The data fields include slip type, magnitude (Ms, M_{JMA} , Mw, and Mo), minimum/maximum seismogenic fault length, minimum/maximum surface rupture length, minimum/maximum fault width, and maximum/average surface slip amount.

3.3 Parameters for Analysis

We first examine whether the strength of asperities on the fault plane can be quantified by the stress drop $\Delta\sigma$. Cotton et al. [9] showed the importance of $\Delta\sigma$ for strong ground motion and also showed large variability of $\Delta\sigma$. The 90 data items for static stress drop $\Delta\sigma_1$ and dynamic stress drop $\Delta\sigma_2$ compiled from [7] are used here as stress drop $\Delta\sigma$. The static stress drop $\Delta\sigma_1$ is a value calculated from the average slip amount D_{ave} and coseismic rupture length L from geological/seismological observations with the following equation:

$$\Delta\sigma_1 \propto \frac{D_{ave}}{L} \quad (3.3)$$

The dynamic stress drop $\Delta\sigma_2$ is a value from the spectrum of the seismic wave record.

The recurrence interval R of a fault is estimated directly from the observed displacement of layers and the dating of layers on the historical earthquake record or trench excavation results. However, this estimation is not always possible, and calculated estimation is conducted by dividing the average slip amount on the fault by the average slip rate S_{ave} as the following formula.

$$R = \frac{D_{\text{ave}}}{S_{\text{ave}}} \quad (3.4)$$

The average slip rate S_{ave} is also an indirect value derived from the cumulative slip amount D divided by the dating of layers/tectonic landforms T for which that slip amount was obtained.

$$S_{\text{ave}} = \frac{D}{T} \quad (3.5)$$

The average recurrence interval R is calculated for the average slip rate S_{ave} by using the dataset of 43 earthquakes in this research (Table 3.1). However, it is difficult to estimate the average slip rate S_{ave} in Eq. (3.5) because the cumulative slip amount D is not constant along fault traces. Thus, uncertainties of slip rate might be large, and the simple mean between S_{min} and S_{max} leads to an underestimation or overestimation which is inappropriate for magnitude estimation. Therefore, just as in Anderson et al. (1996), 100 random numbers were generated in the range between S_{min} and S_{max} , and the distribution of the average slip rate S_{ave} was determined in Table 3.1 for calculation of R in Eq. (3.4).

3.4 Results and Discussion

Large variances of earthquake magnitude to the same surface rupture length L_s are observed in Fig. 3.1 from [8]. Then Fig. 3.2 shows the relationship between earthquake magnitude and surface rupture length from our compiled dataset in Table 3.1. In Fig. 3.2, the dynamic stress drop $\Delta\sigma_2$ was divided with four marks: 50 bars or less, 50–70 bars, 70–90 bars, and 90 bars or more. According to Fig. 3.2, the stress drop is almost always 50 bars or less for earthquakes the fault length for which exceeds 100 km, though the data contain many values of 90 bars or more for fault lengths of 20 km or shorter. For fault lengths between 20 km and 100 km, in particular, near the fault length of 40 km which corresponds to an aspect ratio of two seismogenic-layer earthquakes with a large stress drop display large magnitudes, and earthquakes with a small stress drop exhibit a relatively small magnitude, even if the surface rupture length is the same. Therefore, it is possible to infer that the dynamic stress drop $\Delta\sigma_2$ could be an additional parameter for estimating earthquake magnitude.

Then, a total of six variables were set for regression analysis: surface rupture length L_s , seismogenic fault length L_{sub} , maximum slip amount D_{max} , average slip amount D_{ave} , static stress drop $\Delta\sigma_1$, and dynamic stress drop $\Delta\sigma_2$. Single and multiple regression analyses were conducted to estimate moment magnitude M_w , and the goodness of fit arising from varying combinations of one variable, two variables and three variables were evaluated with AIC, the value of the coefficient of correlation for single regression analysis and the values of coefficient of

Table 3.1 Data for multiple regression analysis compiled from [5–8]

Event	Date	Lat	Lon	Type	Mw	Ls	D _{ave}	S _{min}	S _{max}	S _{ave}	Δσ ₁	Δσ ₂
New Madrid	1811/12/16	36.6	-89.6		8.2			0.01	2	1.005		
Marlborough	1848/10/15	-40.6	174.5	S	7.5	95	6	4	10	7		
Fort Tejon	1857/1/09	34.9	-119.1	S	7.7	297	5.53	16	43	29.5	9.5	18.8
Hayward	1868/10/21	37.5	-122	S	6.8	52		8	10	9	5.63	21.7
Owens Valley	1872/3/26	36.7	-118.1	SN	7.6	108	6	1	3	2	30.6	51.2
Canterbury	1888/9/1	-42.5	172.6	S	7.2	35	2.5	11	25	18		
Nobi	1891/10/27	35.6	136.6	S	7.4	80	4.2	1	10	5.5	30	80.4
Rikuu/Senya	1896/8/31	39.5	140.7	R	7.2		2.59	0.1	1	0.55	33	57
San Francisco	1906/4/18	37.8	-122.6	S	7.9	432	3.9	15	28	21.5	4.24	1.13
Pleasant Valley	1915/10/3	40.3	-117.6	N	7.2	62	3.5	0.3	1	0.65	28.1	59
Tango	1927/3/7	35.5	135.2	SR	6.8	14	2.9	0.01	1	0.505	64.3	657
North Izu	1930/11/25	35.1	139.1	SR	6.9	35	2.95	1	10	5.5	32.6	83
Long Beach	1933/3/11	33.6	-118	S	6.4		0.2	0.1	6	3.05		
Parkfield	1934/6/7	35.8	-120.4	S	6.1	20		29	39	34		
Erzincan	1939/12/25	38	40.3	S	7.7	360	2.43	5	25	15	6.25	5.92
Imperial Valley	1940/5/19	32.8	-115.5	S	6.9	60	1.83	18	23	20.5	29.5	31
Erbaa	1942/12/20	40.7	36.5	SN	6.7	47	0.59	5	25	15	12.8	45
Bolu	1944/2/1	40.9	32.6	S	7.5	180	2.28	5	25	15	6	10
Kem Country	1952/7/21	35.3	-118.7	RS	7.4	57	2.6	3	8.5	5.75	15.8	80
Canakkale	1953/3/18	39.9	27.4	S	7.2	58	2.45	5	25	15	22.5	32.6
Fairview Peak	1954/12/16	39.2	-118.2	SN	7.2	57	2.45	0.01	1	0.505	21.6	33.5
Dixie Valley	1954/12/16	39.6	-118.2	SN	6.9	45	1.8	0.3	1	0.65	25.3	23.9
San Miguel	1956/2/9	31.7	-115.9	SR	6.6	22	0.5	0.1	0.5	0.3	12.3	83
Hebgen Lake	1959/5/15	44.8	-111.2	N	7.3	26.5	2.47	0.8	2.5	1.65	69	212
Niigata	1964/6/18	38.4	139.2	R	7.6		3.3	0.01	1	0.505		

(continued)

Table 3.1 (continued)

Event	Date	Lat	Lon	Type	Mw	Ls	Dave	S _{min}	S _{max}	S _{ave}	$\Delta\sigma_1$	$\Delta\sigma_2$
Parkfield	1966/6/28	35.8	-120.4	S	6.3	38.5	0.38	29	39	34	1.56	15.1
Mudurna Valley	1967/7/22	40.7	31.2	S	7.3	80	1.35	5	25	15	9.75	28.4
Borrego Mountain	1968/4/9	33.2	-116.1	S	6.8	31	0.54	1.4	5	3.2	3.68	41.9
San Fernando	1971/2/9	34.3	-118.4	FG	5.6	16	1.48	2	7.5	4.75	47	67.4
Luhuo	1973/2/8	31.3	100.7	S	7.5	89	1.3	5	10	7.5	12.1	20.4
Coyote Lake	1979/8/6	37	-121.5	S	5.8	14.4		15	19	17	3.13	19.8
El Centro	1979/10/15	32.8	-115.4	S	6.5	30.5	0.8	18	23	20.5	7.87	36.1
Daolu	1981/1/23	31	101.2	S	6.6	44		5	10	7.5	10.2	24.7
Coalinga	1983/5/2	36.3	-120.5	RS	6.4			1	7	4		
Borah Peak	1983/10/28	44.2	-113.8	NS	6.9	34	1.15	0.07	0.3	0.185	23.8	87
Morgan Hill	1984/4/24	37.32	-121.7	S	6.3			3	6.4	4.7		
No Palm Springs	1986/7/8	33.9	-116.6	SR	6.1	9		14	25	19.5		
Edgecumbe	1987/3/2	-38	176.5	N	6.5	14	1.7	1.3	2.8	2.05	48.3	67
Superstition Hills	1987/11/24	33	-115.8	S	6.6	27	0.54	2	6	4	10.2	36.5
Loma Prieta	1989/10/18	37.2	-121.9	SR	6.9			12	28	20		
Luzon	1990/7/16	15.7	121.1	S	7.7	110		10	20	15	15.5	31
Landers	1992/6/28	34.2	-116.4	S	7.3	71	2.95	0.08	2	1.04	25.3	48.1
Northridge	1994/1/17	34.2	-118.5	R	6.7			1.4	1.7	1.55		

Fig. 3.1 Relationship between earthquake magnitude and surface rupture length in [8]

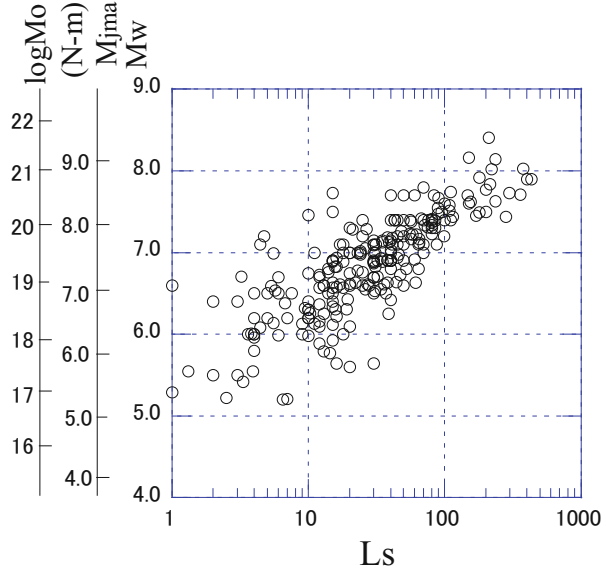
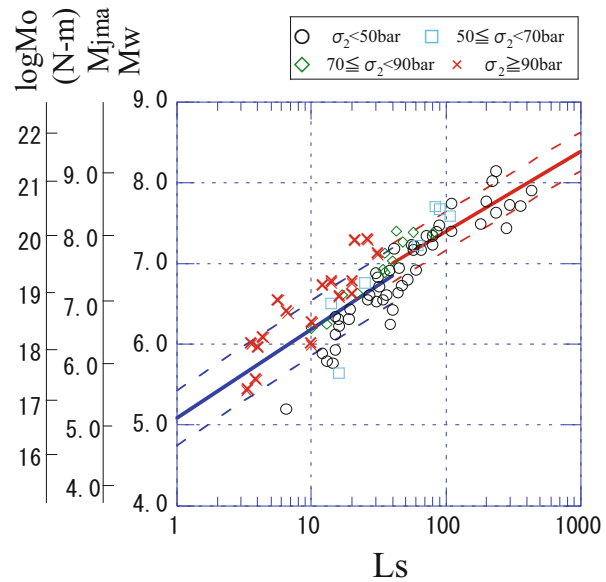


Fig. 3.2 Relationship between earthquake magnitude and surface rupture length with different marks according to stress drop in Table 3.1



determination (Table 3.2). Among various combinations, the Akaike’s Information Criterion [1] value was minimized by taking account of the two variables, surface rupture length L_s , and dynamic stress drop $\Delta\sigma_2$ as Eq. (3.6), though the coefficient of correlation between L_s and $\Delta\sigma_2$ shows 0.69, which means that some multicollinearity effects are assumed.

Table 3.2 Comparison of goodness of fit from single and multiple regression analyses

Response variable	Explanatory variable	AIC	Coef.	
Mw	$\log\Delta\sigma_1$	168.00	0.12	
	$\log\Delta\sigma_2$	160.00	0.04	
Response variable	Explanatory variable	AIC	Coef.	Coef. in variables
Mw	$\log L_s - \log\Delta\sigma_2$	-26.16	0.90	0.69
	$\log L_{\text{sub}} - \log\Delta\sigma_2$	29.65	0.78	0.36
	$D_{\text{max}} - \log\Delta\sigma_2$	96.00	0.59	0.17
	$D_{\text{ave}} - \log\Delta\sigma_2$	61.32	0.64	0.18
	$\log L_s - \log\Delta\sigma_1$	28.34	0.81	0.05
	$\log L_{\text{sub}} - \log\Delta\sigma_1$	29.05	0.78	0.18
	$D_{\text{max}} - \log\Delta\sigma_1$	102.53	0.56	0.43
	$D_{\text{ave}} - \log\Delta\sigma_1$	68.36	0.60	0.23
	$\log L_s - D_{\text{max}} - \log\Delta\sigma_2$	-26.08	0.90	0.41, 0.69, 0.17
	$\log L_s - D_{\text{ave}} - \log\Delta\sigma_2$	-22.15	0.90	0.66, -0.72, -0.18
	$\log L_{\text{sub}} - D_{\text{max}} - \log\Delta\sigma_2$	18.17	0.82	0.67, -0.36, -0.15
	$\log L_{\text{sub}} - D_{\text{ave}} - \log\Delta\sigma_2$	13.56	0.80	0.66, -0.44, -0.11
	$\log L_s - D_{\text{max}} - \log\Delta\sigma_1$	22.73	0.83	0.64, 0.05, 0.41
	$\log L_s - D_{\text{ave}} - \log\Delta\sigma_1$	5.68	0.84	0.66, -0.26, 0.22
	$\log L_{\text{sub}} - D_{\text{max}} - \log\Delta\sigma_1$	19.99	0.82	0.67, 0.18, 0.48
	$\log L_{\text{sub}} - D_{\text{ave}} - \log\Delta\sigma_1$	13.59	0.80	0.66, -0.08, 0.29

$$\text{Mw} = 1.62\log L_s + 0.76\Delta\sigma_2 + 3.07 \quad (3.6)$$

However, the dynamic stress drop $\Delta\sigma_2$ is a value obtained by the spectrum of the seismic wave record observed after the occurrence of an earthquake. Therefore, it is an inappropriate parameter for the estimation of future earthquakes. Our alternative approach is to find a proxy parameter of dynamic stress drop $\Delta\sigma_2$. Figures 3.3 and 3.4 show the relationship between slip rate and $\Delta\sigma_2$ and between recurrence interval and $\Delta\sigma_2$, respectively. Figure 3.3 shows that the dynamic stress drop $\Delta\sigma_2$ is large when the average slip rate S is small and the dynamic stress drop $\Delta\sigma_2$ becomes small when the average slip rate S is large. One problem arising when slip rate is used for prediction of earthquake magnitude is that the average slip rate S has a small value of 0.01–1 mm/year in the active fault catalog, and most of them that are calculated from Eq. (3.5) need age and displacement data derived from tectonic landforms in field surveys. On the other hand, the average recurrence interval R could be derived in both Eq. (3.4) and trench excavations in field surveys. Figure 3.4 shows a relationship between recurrence interval and dynamic stress drop $\Delta\sigma_2$ in which longer recurrence intervals correlated with larger stress drops.

We next conducted regression analyses with the average recurrence interval R determined from average displacement D_{ave} and average slip rate in Table 3.1. Moment magnitude M_w was set as the response variable, and surface rupture length L_s and average recurrence interval R were set as the explanatory variables. The

Fig. 3.3 Relationship between average slip rate and the dynamic stress drop $\Delta\sigma_2$ in Table 3.1

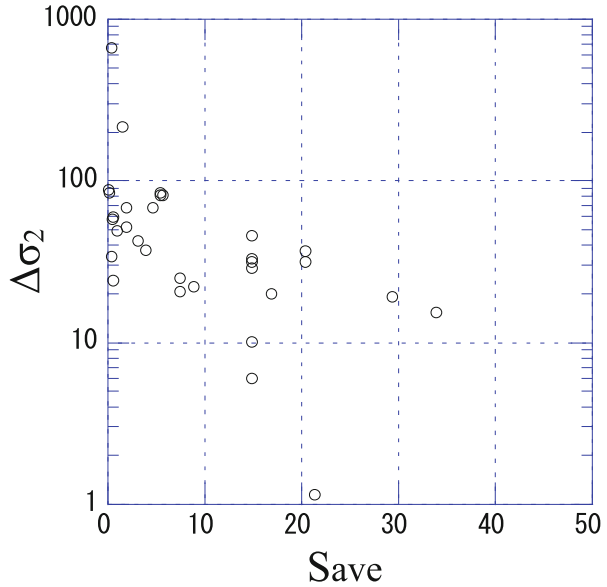
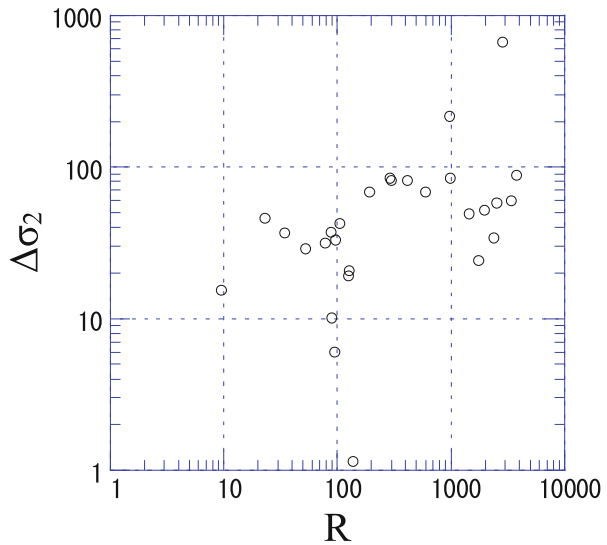


Fig. 3.4 Relationship between recurrence interval and the dynamic stress drop $\Delta\sigma_2$ in Table 3.1

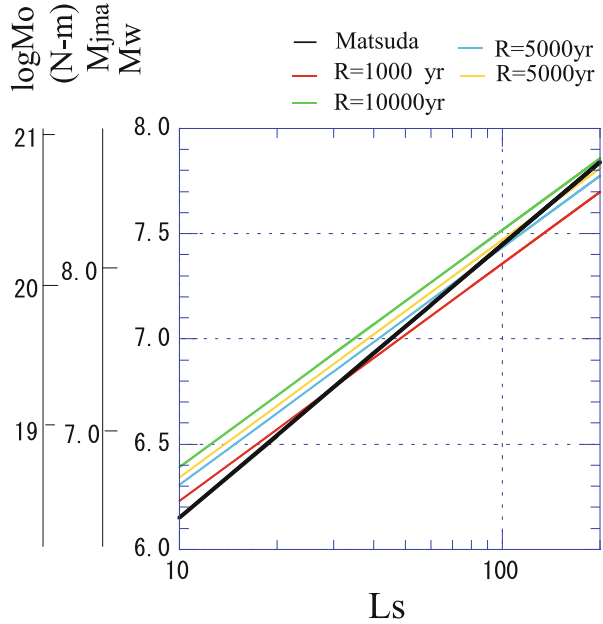


following regression equation was obtained based on the average and standard deviation for each regression coefficient:

$$M_w = 1.13\log L_s + 0.16\Delta\sigma_2 + 4.62 \tag{3.7}$$

This Eq. (3.7) is then compared to Matsuda’s Equation (Eq. 3.1) in Fig. 3.5. Equation (3.7) includes the average recurrence interval R in its explanatory

Fig. 3.5 Comparison of the multiple regression equation in this study with different recurrence intervals and Matsuda's Equation



variables, and thus 1000, 3000, 5000, and 10,000 years are given as typical examples of the average recurrence intervals in Fig. 3.5. For surface rupture length L_s of 20 km, we find that the magnitude obtained with Eq. (3.7) is larger than Matsuda's Equation, for all recurrence interval estimates. In contrast, the magnitude is smaller than that of Matsuda's Equation when L_s is 80 km for average recurrence intervals of less than 3000 years and larger than Matsuda's Equation for average recurrence intervals more than 3000 years. Furthermore, the magnitude is smaller than Matsuda's Equation for an average recurrence interval of 1000 years or less. A difference of 0.2 in magnitude corresponds to a difference of 1000 and 10,000 years in recurrence interval. Therefore, earthquakes of different magnitude in faults of similar lengths can be explained by different recurrence intervals.

3.5 Summary and Conclusion

We have developed new regressions for estimating earthquake magnitude from the fault parameters such as stress drop and recurrence interval. The resulting equation was obtained. The equation shows that for a fault possessing an average recurrence interval of 1000 years or less and a length of 30 km or more, the magnitude estimated from our equation is less than that produced by Matsuda's Equation. Conversely, the magnitude is larger than that of Matsuda for a fault length of 80 km and an average recurrence interval of 3000 years or more. A difference of 0.2 in

magnitude between average recurrence intervals of 1000 years and 10,000 years was also shown.

Acknowledgment The authors thank Ms. Miki Tachibana for data collection. This work is supported by Japan Nuclear Energy Safety Organization (2008–2011) and partly Grant-in-Aid for Scientific Research (18540423).

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19(6):716–723
2. Matsuda T (1975) Magnitude and recurrence interval of earthquakes from a fault (in Japanese). *J Seismol Soc Japan* 28:269–283
3. The Headquarters for Earthquake Research Promotion (2009) A recipe proposed for estimating strong ground motions from specific earthquakes (in Japanese). http://www.jishin.go.jp/main/chousa/09_yosokuchizu/g_furoku3.pdf
4. Irikura K, Miyake H (2001) Prediction of strong ground motion for scenario earthquakes (in Japanese). *J Geogr* 110:849–875
5. Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture area, and surface displacement. *Bull Seismol Soc Am* 84:974–1002
6. Anderson JG et al (1996) Earthquake size as a function of fault slip rate. *Bull Seismol Soc Am* 86:683–690
7. Mohammadioun B, Serva L (2001) Stress drop, slip type, earthquake magnitude, and seismic hazard. *Bull Seismol Soc Am* 91:694–707
8. Stirling MW et al (2002) Comparison of earthquake scaling relations derived from data of the instrumental and pre-instrumental era. *Bull Seismol Soc Am* 92:812–830
9. Cotton et al (2013) What is sigma of the stress drop? *Seismol Res Lett* 84:42–48