Statistical analysis of time-distance relationship between volcanic eruptions and great earthquakes in Japan

Md. Mahmudul Alam and Masaaki Kimura

Department of Physics and Earth Sciences, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan

Volcanic eruptions preceded by seismicity, and great earthquakes (magnitude, $M \geq 7.6$) within 100 km depth during a period of 123 years from 1880 to 2003 along the Pacific and Philippine Sea Plate in Japan are investigated. Correlation coefficients show that the relationship between earthquakes and eruptions ($r = 0.99$), and time and distance ($r = -0.89$) are highly correlated with 1% level of significance. The time-distance relationship between major eruptions and great earthquakes are shown by the model, $Y = 40.15 - 14.53 \log(X)$, where Y and X indicate time (time interval from the starting time of a major eruption to the occurrence of the earthquake) and distance (distance from the volcano to the epicenter of the shock), respectively. Statistical analysis based on the relation shows that the eruptions occurred earlier prior to the concerned shock if the epicenter of the earthquake is nearer to the respective volcanic activity. This relation is recognized by various statistical testing procedures. Based on this relation, the crustal strain accumulation time in the estimated epicenter should be known. This suggests that eruptions may be a precursor for enough strain accumulation in the epicentral region where the increased accumulated regional strain may squeeze up magmas before breaking the crust. If this is true then the occurrence time of the shock may be predicted by a major eruption.

Key words: Least squares method, correlation, regression, regional strain, eruptions, earthquakes.

1. Introduction

Just as the satellite cloud atlas is important in weather forecasting, the tectonic stress field in the earth's crust is very important in earthquake prediction. Although nobody can directly get such a map by using present techniques, the tectonic stress pattern within a seismic region can be indirectly inferred from different kinds of information. Among them, volcanic eruptions may most directly influence the nature of earthquake generating stresses.

Japan is frequently affected by earthquakes. Nearly one tenth of the earthquakes on the earth occur in or around the Japanese islands. Large interplate earthquakes occur along the plate boundaries off the Pacific coast of the Japan islands. Interplate earthquakes within the continental plate take place in the upper crust beneath the Japanese islands and along the coast in the Sea of Japan. It is a widely accepted idea that most large, shallow earthquakes along island arcs result from active subduction and collisions among four lithospheres plates (Pacific plate, North American plate, Eurasian plate and Philippine Sea plate) in this region (Ishida, 1989; Seno *et al.*, 1993, 1996).

Tectonic strain accumulates in the lithosphere of the preseismic stage and is released by the shocks (Kimura, 1978a). Nakamura (1975) suggested that contractional strain generated by regional crustal stress around a magma reservoir can squeeze up magma within an open conduit, causing a summit eruption on one hand and the formation of dike resulting in flank eruptions through the increase of core pressure on the other hand. If the eruptions are influenced by the regional tectonic stresses causing earthquakes, some spatial and temporal relations between large interplate earthquakes and eruptions can be expected along the island arc systems.

A number of researchers have pointed out possible relationships that existed between eruptive activity and seismic activity since early times, although nobody can be sure about the physical mechanisms connecting the volcanic activities with seismicity. MacGregor (1949), for instance, inferring from statistical studies, suggests that a temporal relation exists between the local seismic activity and volcanic eruptions in the Caribbean volcanic arc. Such local seismic activity is thought to be directly involved in volcanic eruptions. Through his studies around the Japanese and New Hebrides areas, Blot (1956, 1972) showed that the deep seismic activity migrates from a greater depth to a shallower one and finally results in volcanic eruptions (Blot process). On the basis of statistical and worldwide studies, Latter (1971) states that Blot process would probably be a secondary phenomenon and that the relationship would be primarily the correlated sequence of seismic and volcanic events resulting from periods of tectonic instability and perhaps increased tensional conditions which affect very wide areas of the earth's surface for periods of several months to several years at a time. On the other hand, many scientists have pointed out that there exists some physical relation between volcanic activity and tectonic seismicity (Tokarev, 1971; Yokoyama, 1971; Kaminuma, 1973).

The high number of correlations suggests that they are largely governed by common physical process caused by re-
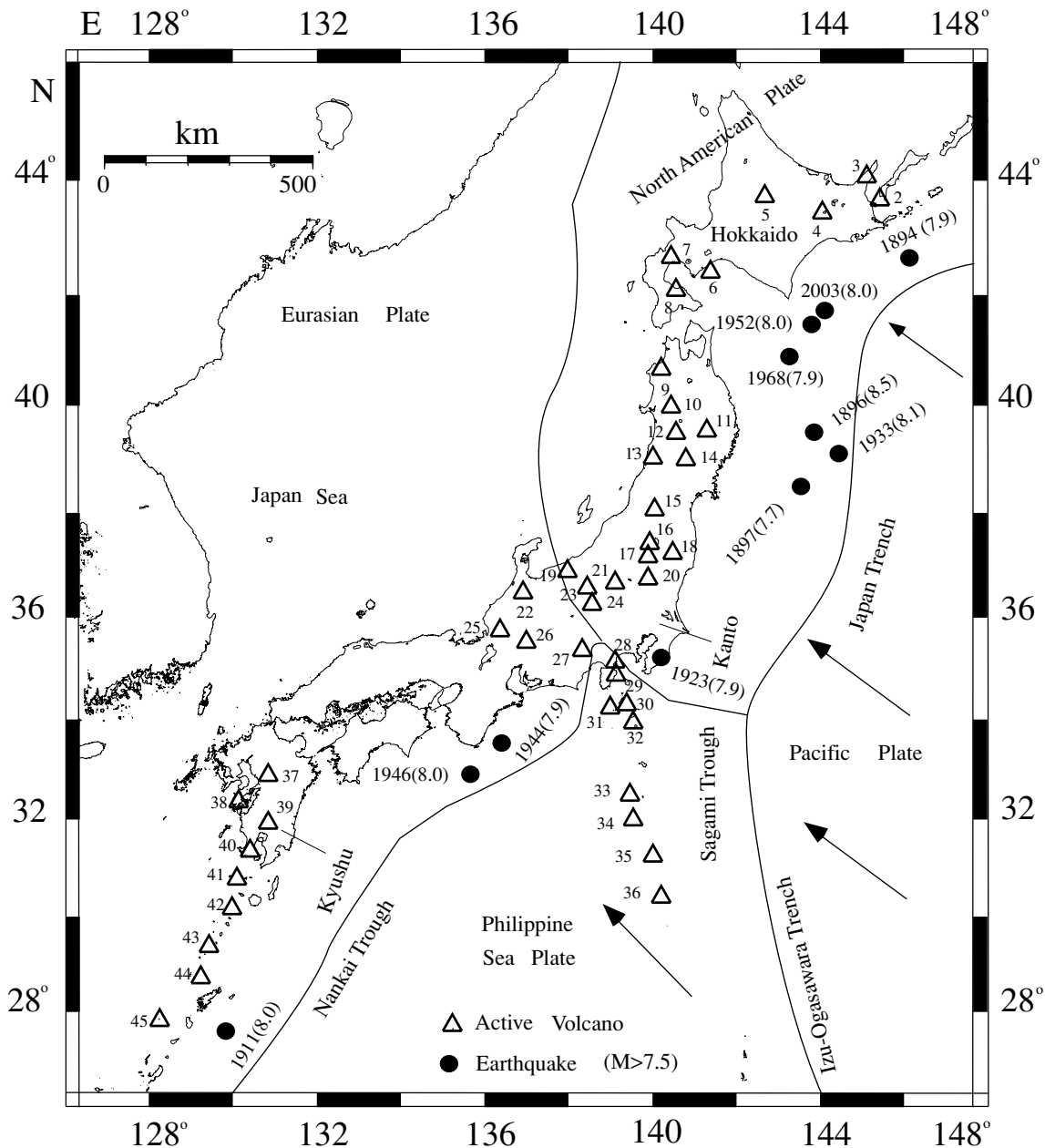
Fig. 1. Distribution of active volcanoes (triangles) and great earthquakes (solid circles) of M ≥ 7.6 since 1880 along the plate boundary of the Pacific and Philippine Sea Plate in Japan. Plate boundaries are taken from the work of Kimura (1994). Arrows indicate directions of plate motion (Wessel and Kroenke, 2000). Name of the volcanoes: (2) Raususan, (3) Shiretoko-Iozan, (4) Meakandake, (5) Tokachidake, (6) Tarumaesan, (7) Usuzan, (8) Hokkaido-Komagatake, (9) Iwakisan, (10) Akita-Yakeyama, (11) Iwatesan, (12) Akita-Komagatake, (13) Tyokaisan, (14) Kurikomayama, (15) Zaozan, (16) Azumayama, (17) Bandaisan, (18) Adatarayama, (19) Niigata-Yakeyama, (20) Nasudake, (21) Nikko-Shiranesan, (22) Yakedake, (23) Kusatsu-Shiranesan, (24) Asamayama, (25) Hakusan, (26) Ontakesan, (27) Fujisan, (28) Hakoneyama, (29) Teishi, (30) Oshima, (31) Kozushima, (32) Miyakejima, (33) Hachijojima, (34) Bayonnaise, (35) Sumisujima, (36) Izu-Torishima, (37) Asosan, (38) Unzendake, (39) Kirishimayama, (40) Sakurajima, (41) Tokara-Io, (42) Kuchinoerabujima, (43) Nakanoshima, (44) Suwanosejima and (45) Okinawa-Torishima.

gional stresses in terms of plate tectonics (Nakamura, 1971, 1975; Kanamori, 1972). A probable relationship between volcanic activity and large earthquakes was first shown in the Kanto area when the Oshima volcano was in active stage and a few years later the two earthquakes occurred concurrently in that region (Kimura, 1976). Based on this information we build up a model to verify the time-distance relation between volcanic activities and large earthquakes exists. Further, the magmatic and seismic activity of central Japan influences us to predict time-distance relationships between volcanic eruptions and great earthquakes in Japan.

2. Data

The area selected in this paper is based on available data. Scientific studies of volcanoes have been going on since the nineteenth century in Japan (Suwa, 1970). Data based on eruptions that have been occurring in Japan from 1880 till today are used in this study. The data on eruptions are taken from the secondary sources: (i) *the Catalogue of Active Volcano of the World Including Solfatara Fields*, IAVCEI (International Association of Volcanology and Chemistry of the Earth's Interior); (ii) *the Bulletin of Volcanic Eruptions*, Volcanological Society of Japan, (IAVCEI); (iii) *the List of the*

Table 1. Name (serial number according to Fig. 1), location and eruption year of volcanoes and earthquakes during 1880 to 2003 with their characteristics and damages due to eruptions are shown. 0; related eruptions, x; non-related eruptions, and ∗; same eruptions used for previous earthquakes. Symbols are as usual meaning.

Volcano	Location	Eruption year and type of eruption	Earthquake year, place, and location	Distance	Year	Remark
5.Tokachidake	43°25′N, 142° 41′E	1889 (o ↑)		290	5	0
3.Shiretoko-Iozan	44°08′N, 145°10′E	1889 (∞)		190	5	0
6.Tarumaesan	42°41′N, 141°23′E	1894 (o ↑)		380	1	0
		1894.3.22(7.9), Nemuro, 42°30′N, 146°00′E				
20.Nasudake	37°07′N, 139°58′E	1881 (o ↑ ⊠)		440	15	x
17.Bandaisan	37°36′N, 140°05′E	1888 (⌐o↑⊠ †)		400	8	x
5.Tokachidake	43°25′N, 142°41′E	1889 (o ↑)		450	7	∗
10.Akita-Yakeyama	39°58′N, 140°46′E	1890 unknown		280	6	0
12.Akita-Koma	39°45′N, 140°48′E	1890-91 (o ↑)		280	6	0
6.Tarumaesan	42°41′N, 141°23′E	1894 (o ↑)		420	2	∗
16.Azumayama	37°44′N, 140°15′E	1894(93-96) (o ↑ †)		380	2	0
15.Zaozan	38°08′N, 140°18′E	1894 (o↑)		360	2	0
		1896.6.15(8.5), Sanriku, 39°30′N, 144°00′E				
20.Nasudake	37°07′N, 139°58′E	1881 (o ↑ ⊠)		320	16	∗
17.Bandaisan	37°36′N, 140°05′E	1888 (o↑⌐⊠ †)		290	9	∗
12.Akita-Koma	39°45′N, 140°48′E	1890-91 (o ↑)		290	7	∗
16.Azumayama	37°44′N, 140°15′E	1894 (93-96) (↑ †)		280	4	∗
15.Zaozan	38°08′N, 140°18′E	1894 (o↑)		260	3	∗
10.Akita-Yakeyama	39°58′N, 140°46′E	1890 unknown		290	7	∗
23.Kusatsu-Shirane	36°38′N, 138°32′E	1897 (o↑)		460	1	0
		1897.8.5(7.7), Sendai, 38°12′N, 142°06′E				
45.Okinawwa-Tori	27°51′N, 128°15′E	1903 (o ↑)		170	8	0
42.Kuchinoerabu	30°26′N, 130°13′E	1906 unknown		270	5	0
		1911.6.15(8.0), Amami, 28°00′N, 130°00′E				
30.Oshima	34°43′N, 139°24′E	1912 (o ↑ ⇒)		80	11	0
24.Asamayama	36°24′N, 138°32′E	1913 (↑ †)		130	10	0
22.Yakedake	36°13′N, 137°35′E	1915 (↑⌐⊠)		170	8	0
35.Sumisujima	31°27′N, 14 0°02′E	1916 (∧)		440	7	x
15.Zaozan	38°08′N, 140°18′E	1918(18-23) unknown		330	5	0
		1923.9.1(7.9), Kanto, 35°06′N, 139°30′E				
15.Zaozan	38°08′N, 140°18′E	1918(18-23) unknown		380	15	∗

Worlds Active Volcanoes, Special Issue of Bulletin of Volcanic Eruptions; (iv) Tokyo Astronomical Observatory (*Science Almanac*) (v) Meteorological Office, and (vi) some published papers on volcanic eruptive activities. Epicenter of the great earthquakes within 100 km depth and magnitude of M ≥ 7.6 are taken from both JMA (Japan Meteorological Agency) and Chronological Scientific. Events are ignored if information does not coincide with both JMA and Chrono-

Table 1. (continued).

8.Hokkaido-Koma	42°04′N, 140°41′E	1919 (o↑)		450	14	x
11.Iwatesan	39°51′N, 141°00′E	1919 (⇒)		310	14	x
8.Hokkaido-Koma	42°04′N, 140°41′E	1922 (o↑)		450	11	x
8.Hokkaido-Koma	42°04′N, 140°41′E	1929 (o↑ ⊠ †)		450	4	x
5.Tokachidake	43°25′N, 142°41′E	1928 (o↑)		490	5	x
12.Akita-Koma	39°45′N, 140°48′E	1932 (∞↑)		320	1	0
1933.3.3(8.1), Sanriku, 39°12′N, 144°30′E						
24.Asamayama	36°24′N, 138°32′E	1935-6 (↑ violent)		340	9	x
23.Kusatsu-Shirane	36°38′N, 138°32′E	1937-39 (o↑)		360	7	x
24.Asamayama	36°24′N, 138°32′E	1938 (o→)		340	6	x
32.Miyake	34°05′N, 139°32′E	1940 (∞ ↑ ⊠ † ⇒)		270	4	0
24.Asamayama	36°24′N, 138°32′E	1941(37-47) (↑ †)		340	3	0
23.Kusatsu-Shirane	36°38′N, 138°32′E	1942 unknown		360	2	0
1944.12.7(7.9), Tonankai, 33°48′N, 136°36′E						
23.Kusatsu-Shirane	36°38′N, 138°32′E	1932 (→o↑ †)		490	14	*
24.Asamayama	36°24′N, 138°32′E	1935-36 (↑ violent)		460	11	*
23.Kusatsu-Shirane	36°38′N, 138°32′E	1937-39 (o↑)		490	9	*
24.Asamayama	36°24′N, 138°32′E	1941(37-47) (↑ †)		460	5	*
24.Asamayama	36°24′N, 138°32′E	1938 (o→)		460	8	*
23.Kusatsu-Shirane	36°38′N, 138°32′E	1942 unknown		490	4	*
34.Bayonnaise	31°55′N, 139°55′E	1946 (∧)		420	1	0
40.Sakurajima	31°35′N, 130°40′E	1946 (↑ ⇒)		490	1	0
1946.12.21(8.0), Nankai, 33°00′N, 135°36′E						
12.Akita-Koma	39°45′N, 140°48′E	1932 (∞↑)		360	20	*
11.Iwatesan	39°51′N, 141°00′E	1934 (↑)		340	18	x
8.Hokkaido-Koma	42°04′N, 140°41′E	1935-42 (o↑)		280	17	x
3.Shiretoko-Iojima	44°08′N, 145°10′E	1936 (→∞)		270	16	x
6.Tarumaesan	43°25′N, 142°41′E	1944 unknown		240	8	0
14.Kurikomayama	38°57′N, 140°47′E	1950 o↑		420	2	0
7.Usuzan	42°32′N, 140°51′E	1944-45 (∞ ↑)		280	7	0
10.Akita-Yakeyama	39°58′N, 140°46′E	1949 (o↑)		350	3	0
1952.3.4(8.0), Tokachi, 41°48′N, 144°06′E						
4.Meakandake	43°23′N, 144°01′E	1962 (o)		300	6	0
10.Akita-Yakeyama	39°58′N, 140°46′E	1957 (o→→)		250	11	x
6.Tarumaesan	42°41′N, 141°23′E	1958 unknown		290	10	x
5.Tokachidake	43°25′N, 142°41′E	1962 (o↑ †)		310	6	0
1968.5.16(7.9), Tokachi, 40°42′N, 143°36′E						
5.Tokachidake	43°25′N, 142°41′E	1985 (o↑)		220	18	x
4.Meakandake	43°23′N, 144°01′E	1988 (o↑)		180	15	x

Table 1. (continued).

5.Tokachidake	43°25′N, 142°41′E	1988-89 (o↟ ↑ →)		220	15	x
4.Meakandake	43°23′N, 144°01′E	1996, 98 unknown		180	7	0
8.Hokkaido-Koma	42°04′N, 140°41′E	1998 (o→)		280	5	0
7.Usuzan	42°32′N, 140°51′E	2000-01 (↟ ↑ →)		280	3	0
10. Akita-Yakeyama	39°58′N, 140°46′E	1997 (↟ ↑ →)		340	6	x
2003.9.26(8.0), Tokachi, 41°47′N, 144°05′E						

logical Scientific (2002).

The volcanoes have their own specific stages, characterized generally by active and dormant stages. The degree of eruptive activities are graded by Kimura (1978b) and major eruptions are designated to represent eruptions in this paper. Major eruption suggests the maximum eruptive activity in the active stage (P_2), and very much corresponds to great eruptions (Kimura, 1976; 1988). The duration of the active stage is approximately 20–30 years usually identified in Japan (Kimura, 2003).

Determination of the magnitudes of eruptions is based on the category of IAVCEI classification, depending on the volume (V) of the erupted materials as: (a) $V < 1 \times 10^4$ m^3—small eruption, (b) $V = 1 \times 10^4$–10^7 m^3—medium eruption, and (c) $V > 1 \times 10^7$ m^3—great eruption. Active volcanoes, great earthquakes and plate boundaries are shown in Fig. 1, and magnitudes, durations and lengths of all events from epicenters to volcanoes are shown in Table 1.

3. Methods
3.1 Selection criterion

(i) Activity of the Oshima volcano is directly related to the occurrence of the great Kanto earthquake in 1923 and Boso-oki earthquake in 1953 along the Sagami Trough (Fig. 2(a)). The altitude of the Aburatsubo area in southern Kanto was dropped before the earthquakes in 1923 and 1953. Since the land mass dropped, the floor of the summit crater of the Oshima volcano rose as much as 400 m, and the volcano erupted. The earthquake occurred almost concurrently. The crater's floor fell when the land of southern Kanto area rose again after the earthquake. It is suggested that the increased compressional crustal stress along the trough squeezes up magma beneath the Mihara-yama, and consequently, the large earthquakes occur to release compressional strain along the trough (Fig. 2(b)). This activity is regarded as a contributing factor to both a major eruptions and earthquakes (Kimura, 1976).

T_1 and T_2 represent time intervals between major eruptions of the Oshima volcano in 1912 and the earthquake of 1923, and the eruptions (Oshima) of 1950 to the earthquake of 1953 as shown in Fig. 2(b). We noticed that the variation of time intervals T_1 (11 years) is larger than T_2 (3 years). But D_1 (distance between the Oshima volcano and the earthquake of 1923; 80 km) is smaller than D_2 (distance between the Oshima volcano to the earthquake of 1953; 230km: Fig. 2(a)) which shows that time interval decreases as distance increases (Fig. 2). The relationship between time and distance is almost linear. A log-linear relationship between time and distance is also found for multiple eruptions to single earthquake in central Japan (Kimura, 2003). This occurs because variations of time intervals strongly suggest that the crustal strain migrates from the area where crustal rupture may appear in future; this was pointed out using all of the related eruptions and large earthquakes in central Japan (Kimura, 2003).

Activity of volcanic eruptions, epicenters, magnitudes and distribution of seismic intensity of great earthquakes were examined to justify the time-distance relationships observed by Kimura (1994) from 1880 to 1993. We used this to provide a testing ground for different kinds of statistical modeling and analytical procedures to understand more precisely the time-distance relationships between volcanic eruptions and great earthquakes in Japan. As a result, eruptions occurred within 20 year and maximum 500 km from the epicenter of great earthquakes along the Pacific and Philippine Sea Plate in the Japanese territory (128° to 147°E and 28° to 46°N) are chosen (Table 1). Eruptions and earthquakes in the Kanto region are the basis for selection of a model which shows a good correlation between time and log distance. Goodness of fit test is applied which recognize the statistical significance of the model. Treating this model as a standard and consequently extended to all events one after another, to other parts of Japan.

(ii) The oldest eruption should be the nearest of the respective epicenter (Fig. 3). If there are several eruptions of a volcano within 20 years before the occurrence of a great earthquake we test the oldest eruption first, if this eruption time fits insignificantly our model, we then look for the next eruption time and consequently search for all eruptions within this range.

If several volcanoes erupted in the same year, the distance between each volcano and the shock should follow the equation: $B = A + 0.5A$, where A is the distance between the nearest volcano and the shock. Events are omitted if distance from volcano to epicenter exceed B.

Usually, events are chosen according to their age of eruptions following a sequence from the oldest to the youngest (Figs. 3–4) and also the distance. For example, according to criterion (i), five eruptions (Tokachi, 1988; Meakan, 1996; Akita-Yakeyama, 1997; Hokkaido-Koma, 1998 and Usu, 2000) were found before the occurrence of the 2003 earthquake. The shortest distance from the epicenter of the 2003 earthquake to the Meakan volcano (1996) was 180 km. Since the eruption of the Tokachi volcano (1988) is older
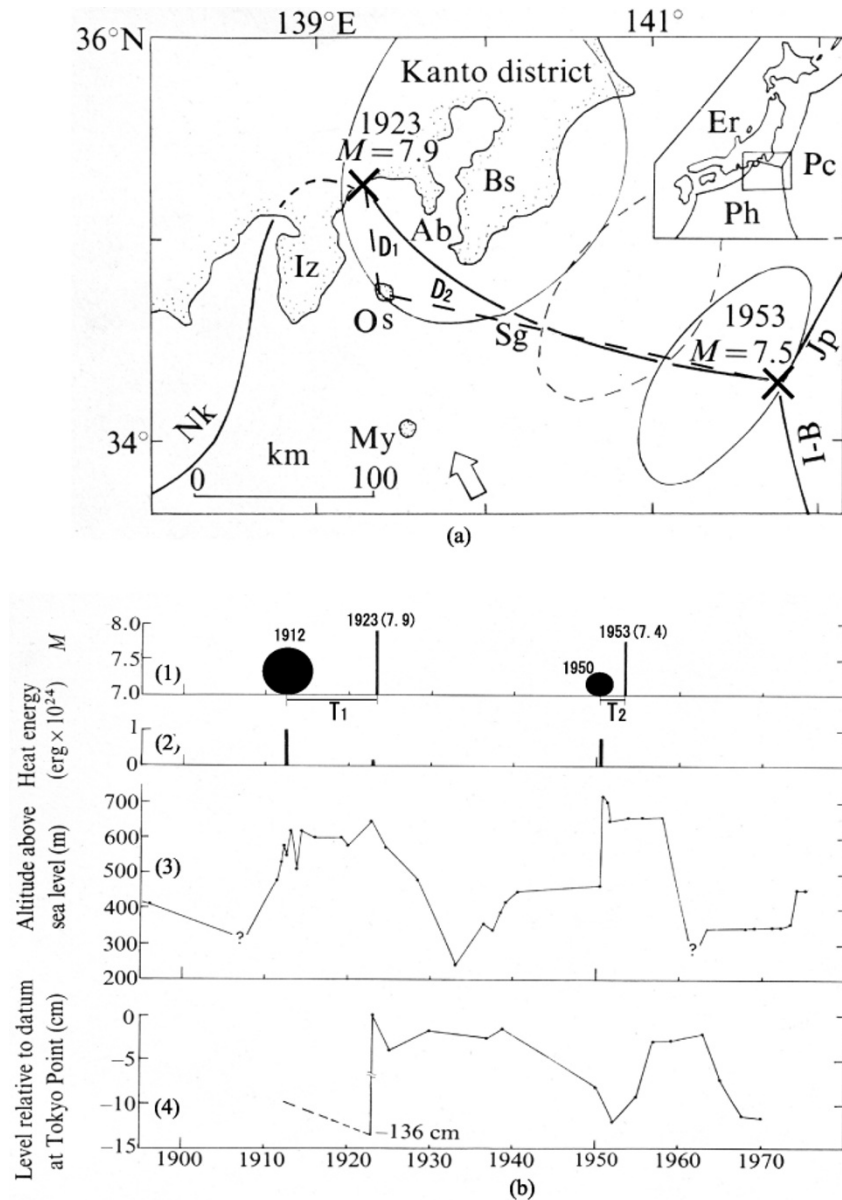
Fig. 2. (a) Large earthquakes (X = epicenter; solid and broken out lines show aftershock areas and seismicity gap, respectively) that have occurred along
the Sagami Trough since 1880. Direction of slip of the northern Philippine Sea Plate (arrow), Sagami Trough (Sg), Nankai Trough (Nk), Japan Trance
(Jp), Izu-Bonin Trence (I-B), Oshima (Os), Miyakejima (My), Aburatsubo (Ab), Boso Peninsula (Bs), Izu Pensinula (Iz), Philippine Sea Plate (Ph),
Pacific Plate (Pc) and Eurasian Plate (Er) are shown. D_1 and D_2 are the distance between the Oshima volcano and the earthquakes of 1923, and 1953.
(b) Relationship between large eruptions (dots) and change in level of floor of summit crater of Mihara-yama (line), and large earthquakes (vertical
bars) that have occurred along the Sagami Trough. (1) Record of large earthquakes and eruptions along Sagami Trough; (2) Heat energy discharged
by large eruptions of Mihara-yama; (3) Change in levels of floor of the summit crater of Mihara-yama (after Kimura, 1976) and (4) Change in level at
Aburatsubo (Solid line measured with precise leveling relative to datum at Tokyo Point, and the broken line was deduced from tidal data). T_1 and T_2 are
the time intervals between the Oshima volcano and occurrence of the shocks of 1923 and 1953, respectively.

than the Meakan but the distance (220 km) is greater than
the Meakan which fails to fulfill criterion (ii) (Fig. 3(h)).
The same thing also happened for the Akita-Yakeyama vol-
cano (Table 1, Figs. 3–4). Further test for residual analysis is
needed for justification.

(iii) If several earthquakes occur within a very short time
interval (within one or two year) and within a very short dis-
tance, each eruption is thought to be related with only one
earthquake. Both earthquakes in 1896 and 1897 are related
to the eruptions of Akita-Yakeyama (1890), Akita-Koma
(1890), Azuma (1894), and Zao (1894). We have taken all
these eruptions for the earthquake of 1896 (Figs. 3(b) and 4).

(iv) Compare to the model of the Kanto region, events are
selected according to residual analysis and outlier test (stan-
dard residuals should belong within the limit of 3.0). Erup-
tions of Tokachidake (1988) and Akita-Yakeyama (1997)
events for the earthquakes of 2003 (Figs. 3(h) and 4) ex-
ceed the error limit and so we do not use these events for
our model.

Time relation between eruptions and earthquakes is also
a factor for selection criterion (Fig. 4). We can verify our
selected events according to Fig. 4. Related and non-related
events (eruptions occurred within 20 years and within 500
km from the epicenter of great earthquakes) are shown by

Table 2. Major eruptions and their related great earthquakes (M≥7.6) in Japan. Distance (km); Distance from a volcano to the epicenter of the respective earthquake, Time (year); Time interval between major eruption and great earthquake, and Magnitude; Magnitude of the earthquake.

No.	Volcano	Eruption year	Earthquake year	Magnitude	Distance (km)	Time interval
5	Tokachidake	1889	Nemuro	7.9	290	5
3	Shiretoke-Iozan	1889	1894		190	5
6	Tarumaesan	1893			380	1
12	Akita-Komagatake	1890	Sanriku	8.5	280	6
10	Akita-Yakeyama	1890	1896		280	6
16	Azumayama	1894			380	2
15	Zaozan	1894			360	2
23	Kusatsu-Shirane	1897	Sendai 1897	7.7	460	1
45	Okinawa-Torishima	1903	Amami	8.0	170	8
42	Kuchinoerabujima	1906	1911		270	5
30	Oshima	1912	Kanto	7.9	80	11
24	Asamayama	1913	1923		130	10
22	Yakedake	1915			170	8
15	Zaozan	1918			330	5
12	Akita-Komagatake	1932	Sanriku 1933	8.1	320	1
32	Miyake	1940	Tonankai	7.9	270	4
23	Kusatsu-Shirane	1942	1944		360	2
24	Asamayama	1941			340	3
34	Bayonnaise	1946	Nankai	8.0	420	1
40	Sakurajima	1946	1946		490	1
6	Tarumaesan	1944	Tokachi	8.0	240	8
7	Usuzan	1945	1952		280	7
10	Akita-Yakeyama	1949			350	3
14	Kurikomayama	1950			420	2
5	Tokachidake	1962	Tokachi	7.9	310	6
4	Meakandake	1962	1968		300	6
4	Meakandake	1996	Tokachi	8.0	180	7
8	Hokkaido-Koma	1998	2003		280	5
7	Usuzan	2000			280	3

open circle and open triangle, respectively. No events (open triangles) are found between circle and bold bar (earthquake year) which primarily agree to our unbiased selection procedure. Figure 4 shows a reverse correlation between time and distance, that is, the time interval decreases as distance increases among the selected events. We also took into consideration plate motion direction and seismic blocking when selection was done. Considering all the factors described above, selected events are shown in Table 2.

3.2 Regressions and model building

Regression analysis is a statistical technique for investigating and modeling the relationship between variables. Applications of regression are numerous and occur in almost every field. In fact, regression analysis may be the most widely used statistical technique. We have a number of sample observations on time and distance between eruptions and great earthquakes in Japan. Plotted observations are shown in Fig. 5 which is called a scatter diagram. This display clearly suggests a relationship between time and distance; in fact, the impression is that the data points generally, but not exactly, fall along a line. If we let Y represent time and X_1 represent log distance, then the equation relating to these two variables is,

$$Y = \beta_0 + \beta_1 X_1 \qquad (1)$$

where β_0 is the intercept and β_1 is the slope. Now the data points do not fall exactly on a line, so Equation (1) should be modified to account this. Let the difference between the observed value of Y and the line ($Y = \beta_0 + \beta_1 X_1$) be an error ε. ε is a statistical error that is, it is a random variable that accounts for the failure of the model to fit the data exactly. Thus, a more plausible model for the time data is,

$$(i) \text{ Model-1} : Y = \beta_0 + \beta_1 X_1 + \varepsilon. \qquad (2)$$

The variable X_1 (log distance) is called predictor or regressor variable and Y (time) as the response variable. If more than one regressor variable is involved in the model, the model is called multiple linear regression models. The Model-1 can be represented as a multiple regression model for adding one more regressor variable X_2 (Magnitude) that is available in
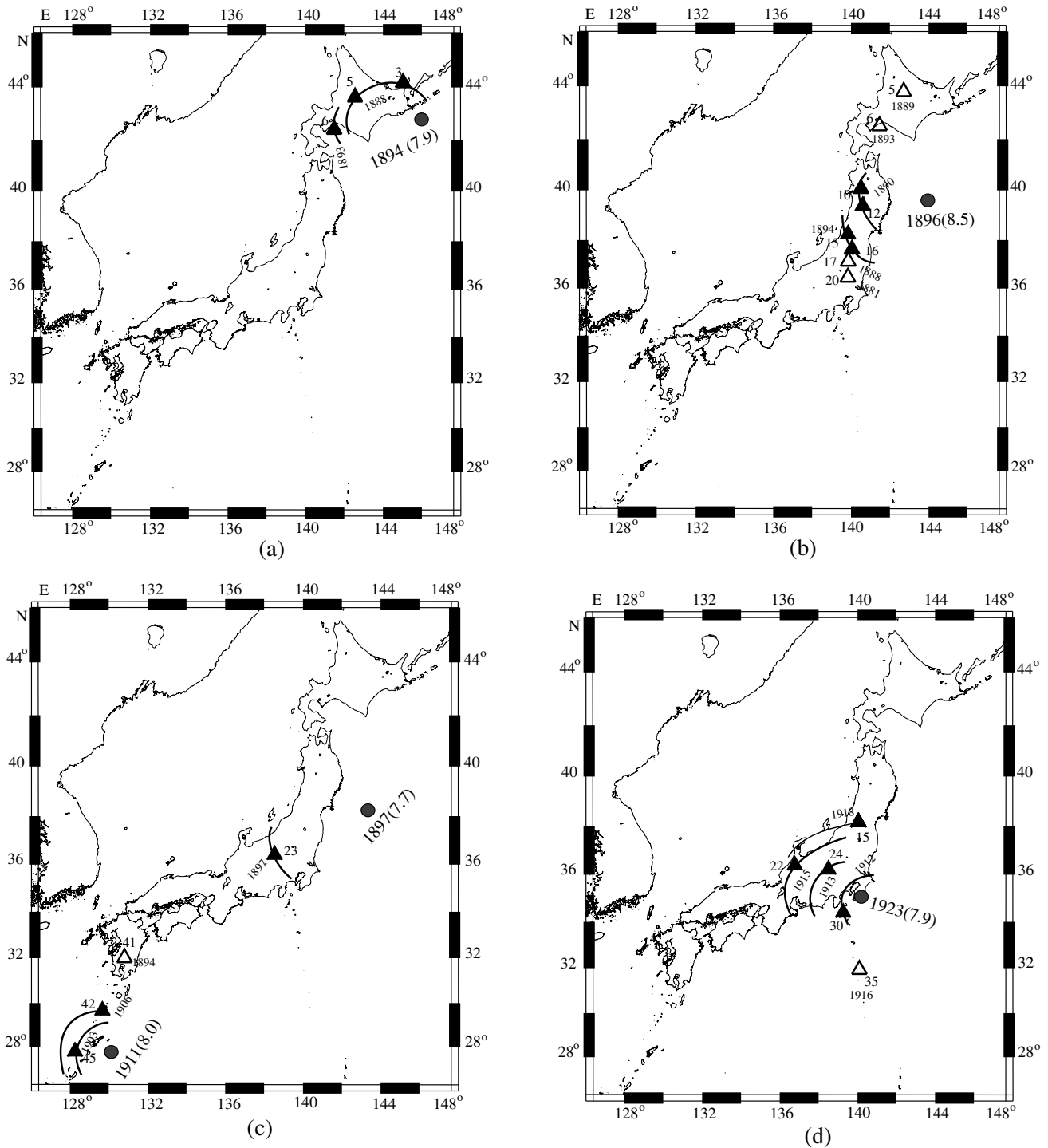
Fig. 3. Distance relations among eruptions and great earthquakes. Volcano/volcanoes with their age of eruptive activities within 20 years and 500 km from the epicenter of respective great earthquake during the period of 1880 to 2003 are shown. Solid triangles are thought to be related eruptions of respective shocks. Open triangles are the events that fulfill partly but not all of our criterions. Numerals aside the volcanoes are the eruption year and volcanoes number according to Fig. 1. Strain migration directions from the epicentral region are shown by contour lines. (a) Eruptive activities and seismicity during 1880–1894, (b) Eruptive activities and seismicity during 1895–1896, (c) Eruptive activities and seismicity during 1897–1911, (d) Eruptive activities and seismicity during 1912–1932, (e) Eruptive activities and seismicity during 1933–1945, (f) Eruptive activities and seismicity during 1946–1967, (g) Eruptive activities and seismicity during 1968–2002, and (h) Eruptive activities and seismicity during 2003.

our collection,

(ii) Model-2 : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon.$ (3)

If we have k regressor variables, the multiple regression models can be written as,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon. \quad (4)$$

The errors are uncorrelated normally and independently distributed (NID) with mean zero and unknown variance (σ^2) i.e., $\varepsilon \sim \mathrm{NID}(0, \sigma^2)$.

3.2.1 Estimation of parameters by using least squares method Suppose that $n > k$ observations are available. The parameters $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ are unknown and must be
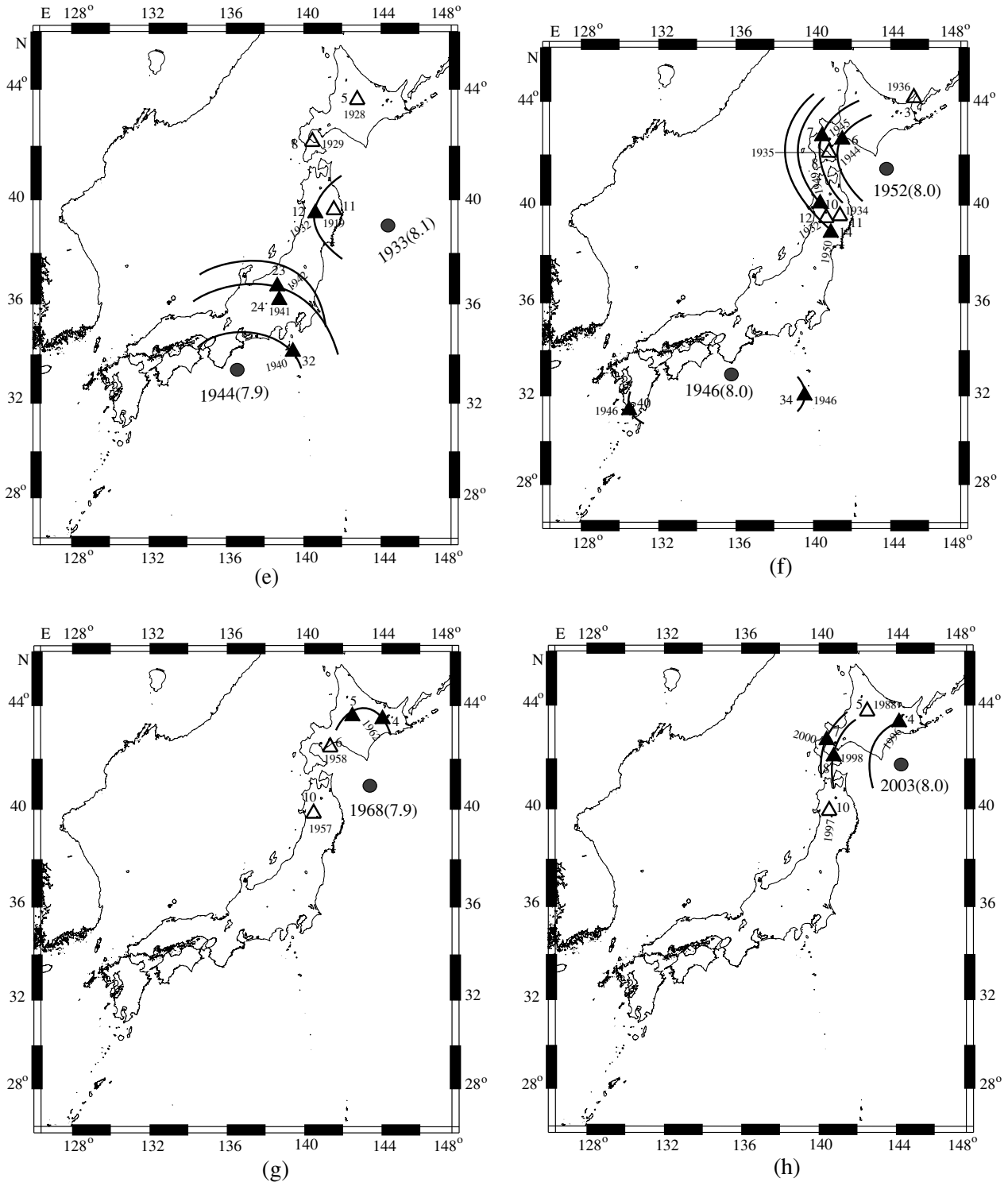
Fig. 3. (continued).

estimated using sample data. For models in which some transformation of any function is linear in the parameters, least squares estimation can be used to estimate the parameters of the model. That is, we will estimate the parameters in Eq. (4) so that the sum of the squares of the differences between the observations (Y_i, $i = 1, 2, \ldots, n$) and the log linear model is a minimum. If the errors are normally and independently distributed with mean zero and constant variance (σ^2), the unknown parameters of the above models can be estimated using least squares method as: $\hat{\beta} = (X'X)^{-1}X'Y$, where $\hat{\beta}$ is a vector of order $(k + 1) \times 1$, X is a matrix of order $n \times (k+1)$ and Y is a vector of order $n \times 1$. Here, $k+1$ and n are the number of parameters and number of sample observations, respectively (Montgomery *et al.*, 2001).

3.2.2 Test for significance of regression The test for significance of regression is a test to determine if there is a linear (log linear) relationship between the response Y and any of the regressor variables $X_1, X_2, \ldots, X_k$. This
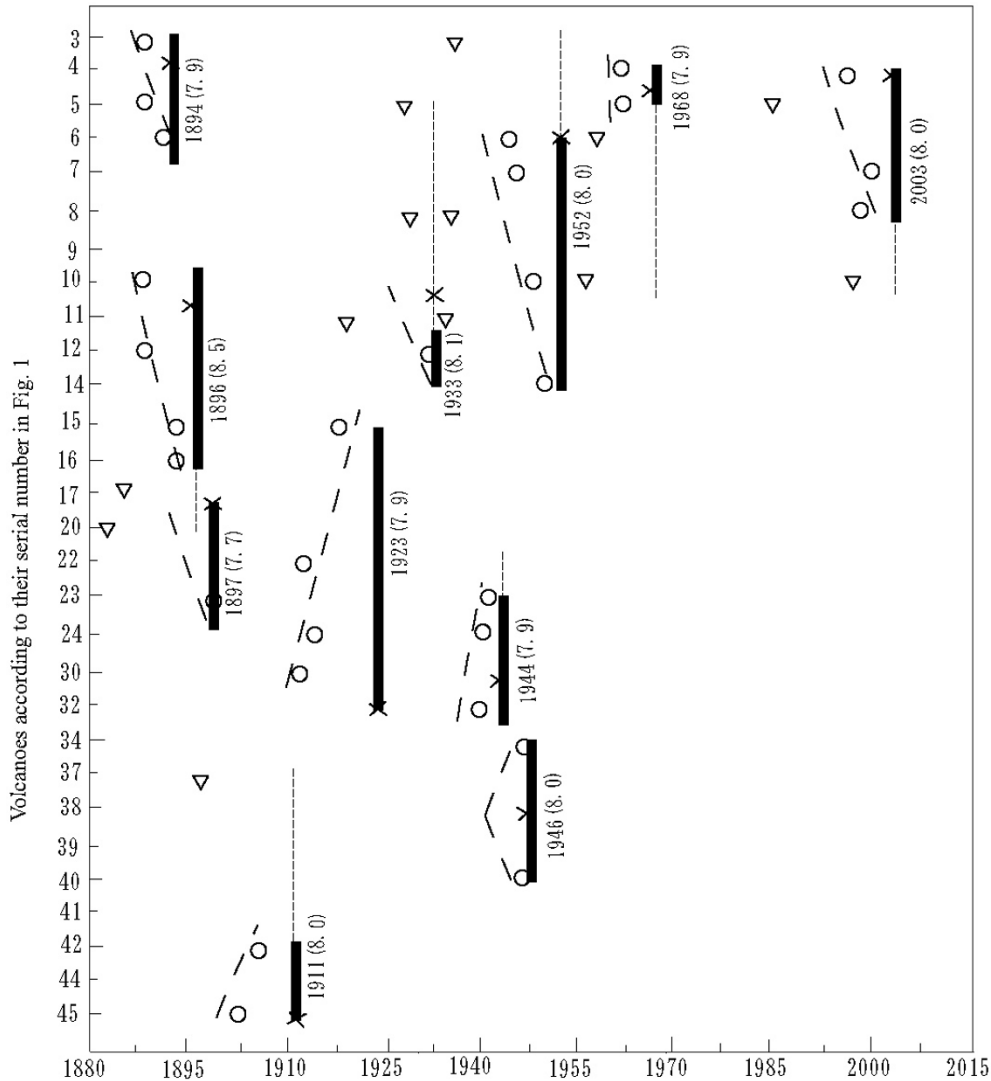
Fig. 4. Time relation between eruptions and great earthquakes. Earthquakes (bold bars), related eruptions (open circle), and non-related eruptions (open triangle) according to Table 1 are shown. Bold and broken lines are the distance (maximum 500 km) between eruptions and earthquakes. Thick broken lines around the open circles suggest the reverse correlation between time and distance from the approximate epicenter (X).

procedure is often thought of as an overall test of model adequacy. The appropriate hypotheses are,

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \qquad (5)$$
$$H_1 : \beta_i \neq 0 \text{ for at least one } i.$$

Rejection of this null hypothesis implies that at least one of the regressors ($X_1, X_2, \ldots, X_k$) contributes significantly to the model. If the null hypothesis is true, then SSR (sum of squares regression)/σ^2 follows a χ_k^2 distribution, which has the same number of degrees of freedom as number of regressor variables in the model. Also SS_{Res} (sum of squares regression)/$\sigma^2 \sim \chi_{n-k-1}^2$, and that SS_R and SS_{Res} are independent. By the definition of F statistic,

$$F_0 = \frac{SS_R/k}{SS_{Res}/(n-k-1)} = \frac{MS_R}{MS_{Res}} \sim F_{k,n-k-1} \qquad (6)$$

where n and k are the total number of observations and parameters respectively, MS_{Res} and MS_R represent the mean sum of squares residual and mean sum of squares regression,

respectively. A test of size α for the null hypothesis is given by rejecting H_0 if $F_0 > F_{\alpha;k,n-k-1}$, where $F_{\alpha;k,n-k-1}$ denotes the $\alpha\%$ point of the F-distribution with k and $n - k - 1$ degrees of freedom (Hocking, 2003).

3.2.3 Test on individual regression coefficients Once we have determined that at least one of the regressors is important, a logical question becomes which one(s). Adding a variable to a regression model always causes the sum of squares to decrease. We must decide whether the increase in the regression sum of squares is sufficient to warrant using the additional regressior in the model. The addition of a regressor also increases the variance of the fitted value ($Y = $ time), so we must be careful to include only regressors that are of real value explaining the response. The hypothesis for testing the significance of any individual regression coefficient, such as β_j are

$$H_0 : \beta_j = 0 \qquad (7)$$
$$H_1 : \beta_j \neq 0 \qquad j = 1, 2, \ldots, k.$$

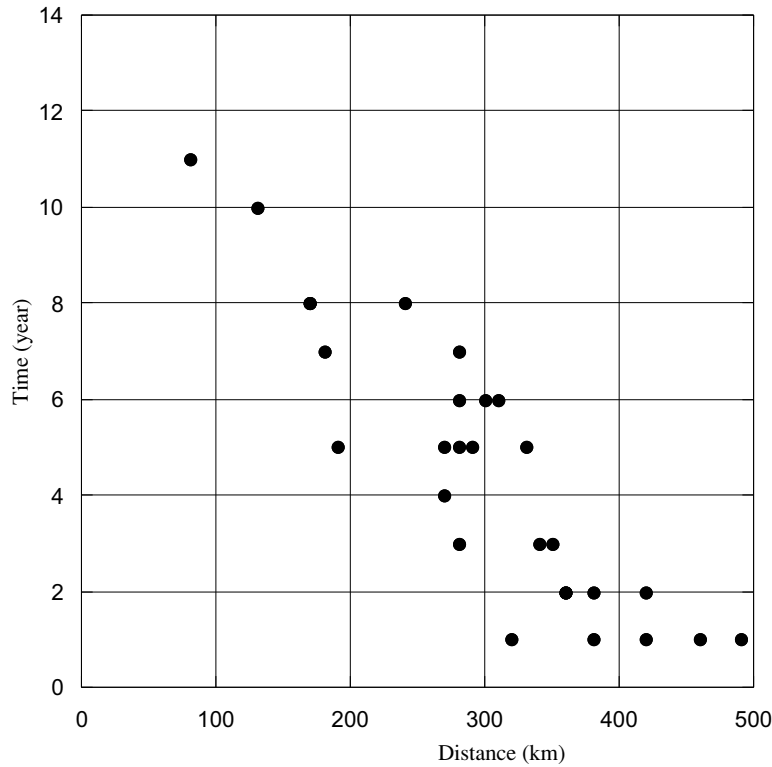If $H_0 : \beta_j = 0$ is not rejected, then this indicates that

Fig. 5. Scatter diagram showing the distribution of events. The distribution pattern is approximately log-linear (linear), that is, time and distance may have a log-linear (or linear) relationship.

the regressor X_j can be deleted from the model. The test statistic for this hypothesis is $t_0 = \{\hat{\beta}_j - E(\hat{\beta}_j)\}/se(\hat{\beta}_j) \sim t$-distribution with degrees of freedom $(n - k - 1)$. The null hypothesis $H_0 : \beta_j = 0$ is rejected if $|t_0| > t_{\frac{\alpha}{2},n-k-1}$. Here, $t_{\frac{\alpha}{2},n-k-1}$ denotes the value of the t distribution such that $\Pr(t > t_{\frac{\alpha}{2},n-k-1}) = \frac{\alpha}{2}$. k, n, se and α are the number of parameters, number of sample observations, standard error and level of significance, respectively. This is really a partial test because the regression coefficient $\hat{\beta}_j$ depends on all of the other regressor variables X_i $(i \neq j)$ that are in the model. Thus, this is a test of the contribution of X_j given the other regressors in the model (Montgomery *et al.*, 2001).

3.2.4 Test for correlation coefficient The sample correlation coefficient is a measure of the linear association between Y (Time) and X (Distance). The estimator of the population correlation ρ is the sample correlation coefficient $r = \frac{S_{XY}}{\sqrt{(S_{XX}SS_T)}}$, where S_{XX}, SS_T and S_{XY} is the sum of squares X, sum of squares total and sum of the products of X and Y, respectively. It is often useful to test the hypothesis that the correlation coefficient equals zero, that is,

$$H_0 : \rho = 0 \qquad (8)$$
$$H_1 : \rho \neq 0.$$

The appropriate test statistic for this hypothesis is, $t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ which follows the t distribution with degrees of freedom $n - 2$ if $H_0 : \rho = 0$ is true. Therefore, we would reject the null hypothesis if $|t_0| > t_{\frac{\alpha}{2},n-2}$, where n and α are the number of sample observations and level of significance, respectively.

3.2.5 Confidence interval estimation The width of the confidence intervals of the parameters is a measure of the overall quality of the regression line. If the errors are normally and independently distributed, then the sampling distribution with some transformation follows t distribution with $n - 2$ degrees of freedom. Therefore, a $100(1 - \alpha)$ percent confidence interval of the parameters are given by,

$$\Pr\{\hat{\beta} - t_{\frac{\alpha}{2},n-2}se(\hat{\beta}) \leq \beta \leq \hat{\beta} + t_{\frac{\alpha}{2},n-2}se\hat{\beta}\} = 100(1 - \alpha).$$
$$(9)$$

These confidence intervals have the usual frequency interpretation. That is, if we were to take repeated samples of the same size at the same X levels and construct, for example, 95% of those intervals will contain the true value of the parameters. The limits of the parameters depend on the value of α (level of significance).

4. Results

In this paper, a simple log linear regression model is fitted to investigate the time-distance relationships between volcanic eruptions and great earthquakes in Japan. Model-1 includes variables time and log distance, and Model-2 includes variables time, log distance and magnitude.

Coefficient of determination (R^2) is approximately 0.8808 that is 88.08% of the variability in time (Y) is accounted for by the regression Model-1 and 88.15% by the Model-2 (Table 3). Addition of one more variable (Magnitude; Model-2) in the model results (Table 3) a slightly higher value of R^2 ($R^2 = 0.8815$) than Model-1 ($R^2 = 0.8808$). In general, R^2 always increases when a regressor is added to the model, regardless of the value of the contribution of that variable. Therefore, it is difficult to judge whether an increase in R^2

Fitted model: Y = 40.151 - 14.534 Log(X)
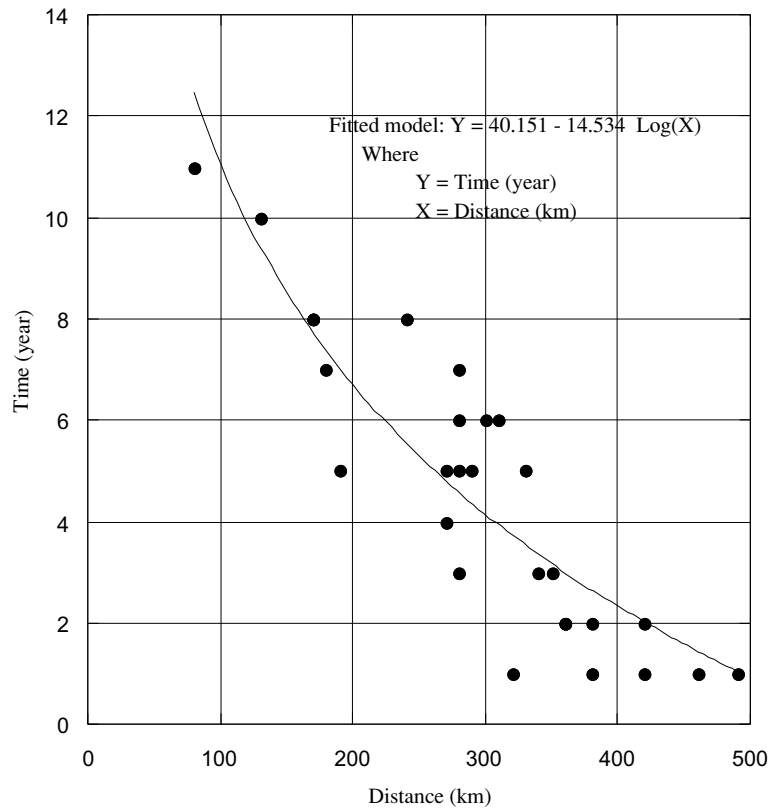Where
Y = Time (year)
X = Distance (km)

Fig. 6.   Time-distance relationship between eruptions and great earthquakes in Japan. Negative association between time and distance is shown by the curve. That is, time interval decreases as distance increases.

Table 3.   Model summary. R^2 shows the percentage of variation that can be explained by the models.

Model	Coefficient of determination, R^2
Model - 1	0.8808
Model - 2	0.8815

is really telling us anything important. Additional variable may have significant influence if the addition of the variable reduces the residual mean square. In our analysis, residual mean square is smaller in Model-1 (1.87) than Model-2 (1.93). Therefore, additional variable (Magnitude) may not have significance influence on time-distance relationship. Further test is required for justification (Tables 4–5).

Table 4 represents the analysis of variance test significance of regression. Failing to reject the null hypothesis (Eq. (5)) implies that there is no significant relationship between time and log distance. On the other hand, rejection of null hypothesis implies that log distance is of value in explaining the variability in time. However, rejection of null hypothesis could mean either that the log linear model is adequate or that even though there is a log linear effect of distance, better result could be obtained with the addition of more regressor variables or of higher order polynomial terms in log distance (X). Our calculated results in Table 4 ($P = 1 \times 10^{-6}$) suggest that the overall regression is highly significant indicating time may have a significant relationship with log distance. Addition of a variable (Magnitude; Model-2) shows that the P-value of the overall regression test (Model-2, Table 4) is also very small ($P = 1 \times 10^{-6}$) indicating time may have re-

lation to log distance and/or magnitude. However, this does not necessarily imply that the relationship found is an appropriate one for predicting time as a function of log distance and magnitude. Further tests of model adequacy are required (Table 5).

We performed tests on individual regression coefficients (Model-2, Table 5) to determine whether a significant influence of magnitude on time-distance relationship exists. Since absolute value of the test statistic t_0 ($t_0 = 0.33$) in Table 5 is less than the true value of t ($t_{0.025,27} = 2.05$; this value is found in any t-distribution table), we may not reject $H_0 : \beta_2 = 0$ and conclude that the magnitude, or X_2, contributes insignificantly to the model given that log distance or X_1, is in the model. This implies magnitude may have a very little influence on the time-distance relationship although statistically insignificant ($P = 0.74$ which is greater than 0.05). Linear regression between time and magnitude, and log distance and magnitude are also tested but found no significant association between them. So we may delete magnitude from the model and the proposed model for the time-distance relationship should be: $Y = 40.15 - 14.53 \log(X)$, where Y and X represent time and distance, respectively (Montgomery et al., 2001).

Table 4.   Analysis of variance test.   Table showing the mean squares of regression and residual along with F statistic for testing H_0: $\beta_1 = \beta_2 = \ldots = \beta_k = 0$. P-values are used for hypothesis testing.

Source	Sum of Squares	Degrees of Freedom	Mean Square	Test Statistic (F_0)	P
Model - 1					
Regression	174.39	1	174.39	93.36	1×10^{-6}
Residual	50.43	27	1.87		
Total	224.83	28			
Model - 2					
Regression	174.61	2	87.30	45.20	1×10^{-6}
Residual	50.22	26	1.93		
Total	224.83	28			

Table 5.   Table shows the standard errors of the estimates and intercepts along with the t statistic for testing H_0: $\beta_0 = 0$, H_0: $\beta_1 = 0$ and H_0: $\beta_2 = 0$. The P-values for the test for significance of individual regression coefficients (distance and/or magnitude) are reported. 95% confidence intervals for the parameters are also shown. LB and UB are the lower and upper limits of the parameters, respectively.

Predictor	Coefficients	Standard Error	Test Statistic (t_0)	P	95% Confidence Interval	
					LB	UB
Model – 1						
Constant	$\hat{\beta}_0 = 40.15$	3.67	10.89	1×10^{-7}	32.59	47.71
Log distance	$\hat{\beta}_1 = -14.53$	1.50	-9.66	1×10^{-7}	-17.62	-11.45
Model – 2						
Constant	$\hat{\beta}_0 = 36.93$	10.41	3.55	0.001	15.54	58.32
Log distance	$\hat{\beta}_1 = -14.61$	1.55	-9.45	1×10^{-7}	-17.78	-11.43
Magnitude	$\hat{\beta}_2 = 0.42$	1.28	0.33	0.74	-2.20	3.05

Table 6.   Correlation coefficients among distance, earthquake, eruption, magnitude, and time. $\ast\ast$ indicates the significant correlation between the variables at 1% level (according to test statistic t, Eq. (8)).

	Distance	Earthquake	Eruption	Time	Magnitude
Distance	1.000	-0.04	0.04	-0.89**	0.09
Earthquake		1.00	0.996**	0.06	-0.30
Eruption			1.00	-0.03	-0.29
Time				1.00	-0.10
Magnitude					1.00

The negative value of the coefficient β_1 indicates that the eruption occurs prior to the concerned shock if the epicenter of the earthquake is nearer to the respective volcano/volcanoes. Table 5 shows the 95% confidence intervals for the parameters of the models. That is, if we choose repeated samples of the same size then 95% of those intervals will contain the true value of the parameters.

The correlation coefficient between earthquakes and eruptions is 0.99 which indicates that the earthquakes and eruptions are highly correlated with 1% level of significance. Again the correlation coefficient between time and distance is about −0.89 which shows a highly significant negative correlation. But the correlation coefficients between time

and magnitude, and distance and magnitude are −0.10 and 0.09, respectively, are statistically insignificant which mean magnitude may not have significant association with time and/or distance. Other pairs do not show any significant correlation (Table 6). The Variance Inflation Factors ($VIF = 1.00$) suggest that the data used in the analysis is free from multicolinearity. Outliers and the effects of autocorrelations on the data set were also tested.

5.   Discussion and Conclusions

There have been some studies, however, in which triggering of earthquakes by volcanic eruptions (Kimura, 1994) was proposed. We examine here the records of eruptions

and earthquakes to see if there are indeed significant earthquakes following eruptions to show the time-distance relationship between them. Earthquakes usually accompany volcanic eruptions; they may occur before or after the initial eruptions as magma movement stresses the surrounding crust (Linde and Sacks, 1998). Our concern was only to study those earthquakes occurred after the eruptions.

We considered the area of Japan to justify whether the time-distance relationship exists between volcanic activity and great earthquakes. We performed three kinds of models belonging to simple, linear, log linear and logistic regressions. Logistic regressions and a simple log linear regressions provided similar fits. Results of the analysis for log linear regression models are shown in Tables 3–6 and Fig. 6. Statistical analyses suggest that time and log distance is significantly related but magnitude may not have significant relation with the time-distance relation. Correlation coefficients between magnitude and other focal parameters also suggest that magnitude may not have significance association with them. This implies that the earthquake preparation process is larger scale and not a function of the size of the impending event. It also implies that mode of rupture is not controlled by stress and crustal condition which suggest that the size of the earthquakes may not play a significant role for model building. Further theoretical and experimental work including numerical simulation is needed to conform the tectonic cause for happening of both volcanic eruptions and great earthquakes.

All the results also suggest that volcanic activities and occurrence of great earthquakes are closely related to the change of stress activity. Volcanic activity increases prior to the occurrence of large earthquakes due to accumulation of regional strain around the seismic and volcanic region. As a result, volcanoes in critical state erupt but other volcanoes, not in critical state may not erupt at the same time. The nearest volcano of the epicentral area may be affected earlier by the migration of accumulated strain and erupted first. Consequently, other volcanoes are affected and erupted by the same way according to distance. In general, when the volcanoes are erupted, enough strain accumulated in the eventual epicentral region following earthquakes.

In conclusion, it is suggested that the criterion for occurrence of a great earthquake based on timing of a volcanic eruptions has a time-distance relationship. That is, volcanoes nearer to the eventual epicentral region erupt earlier than the others. This result strongly suggests that time-distance relations may help to predict an earthquake before it strikes if the epicentral location can be identified in advance and if the activity of the volcanoes is well monitored.

References

Blot, C., Relations entres seismes profonds et les eruption volcaniques au Japan, *Bull. Volcanol.*, **28**, 25–63, 1956 (in French).

Blot, C., Volcanism et seismes du manteau superieur dam l'Archipel des Nouvelles-Hebrides, *Bull. Volcanol.*, **36**, 446–561, 1972 (in French with English abstract).

Chronological Scientific, *Rika Nenpyo*, edited by National Astronomical Observations, Maruzen Co., Ltd, 2002.

Hocking, R. R., *Methods and Applications of Linear Models*, 2nd ed., 741 pp., John Wiley and Sons, Inc., NJ, 2003.

Ishida, M., Seismicity and tectonics of Japan, in *Earthquakes and Seismotectonics of the Japanese Islands*, edited by T. Hagiwara, pp. 57–85, Kajima Institute Publishing, Tokyo, 1989.

Kaminuma, K., Correlation between volcanic and seismic activities in Kyushu, Japan, Publ. for the 50th Anniversary of the Great Kanto Earthquakes, 1923, pp. 185–197, Earthquake Res. Inst., Tokyo Univ., 1973 (in Japanese with English abstract).

Kanamori, H., Relation between tectonic stress, great earthquake and earthquake swarms, *Tectonophysics*, **14**, 1–12, 1972.

Kimura, M., Major magmatic activity as a key to predicting large earthquakes along Sagami Trough, Japan, *Nature*, **260**, 131–133, 1976.

Kimura, M., *Relation between Eruptions and Large Earthquakes*, pp. 187, Univ. Tokyo Press, Tokyo, 1978a (in Japanese).

Kimura, M., Significant eruptive activities related to large interplate earthquakes in the north western Pacific margin, *J. Phys. Earth*, **26**, S557–S570, 1978b.

Kimura, M., Volcanic eruptions and their tectonic control in Izu-Oshima and Miyakejima volcanoes, Japan, *The Volcanological Society of Japan*, **33** (Special issue-2), S319–S329, 1988.

Kimura, M., Relationship between volcanic eruption and large earthquakes in the vicinity of Japan, *Annuals, Disas. Prev. Inst. Kyoto Univ.*, **37**, B-1, 293–317, 1994 (in Japanese with English abstract).

Kimura, M., *Earth Science for Eruptive Activity and Seismicity*, 293 pp., Ronso Press, Tokyo, 2003 (in Japanese).

Latter, J. H., The interdependence of seismic and volcanic phenomena in the Caribbean volcanic arc, *Bull. Volcanol.*, **35**, 127–142, 1971.

Linde, A. T. and S. Sacks, Triggering of volcanic eruptions, *Nature*, **395**, 888–890, 1998.

MacGregor, A. G., Prediction in relation to seismo-volcanic phenomena in the Caribbean volcanic arc, *Bull. Volcanol.*, **8**, 69–85, 1949.

Montgomery, D. C., E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 3rd ed., 641 pp., John Wiley and Sons, Inc., NY, 2001.

Nakamura, K., Volcano as a possible indicator of crustal strain, *Bull. Volcanol. Soc. Jpn.*, **16**, 63–71, 1971 (in Japanese with English abstract).

Nakamura, K., Volcano structure and possible mechanical correlation between volcanic eruption and earthquakes, *Bull. Volcanol. Soc. Jpn.*, **20**, 229–240, 1975(in Japanese with English abstract).

Seno, T., S. Stein, and A. Gripp, A model for motion of the Philippine Sea plate consistent with NUVEL-1 and geological data, *J. Geophys. Res.*, **98**, 17941–17948, 1993.

Seno, T., T. Sakurai, and S. A. Stein, Can the Okhotsk plate be discriminated from the North American Plate?, *J. Geophys. Res.*, **101**, 11305–11315, 1996.

Suwa, A., *Eruptions in Japanese Islands*, 221 pp., Kodan-sha, Tokyo, 1970 (in Japanese).

Tokarev, P. I., On the focal layer, seismicity and volcanicity of the Kurile-Kamchatka zone, *Bull. Volcanol.*, **35**, 230–242, 1971.

Wessel, P. and L. W. Kroenke, Ontong Java Plateau and Late Neogene changes in Pacific plate motion, *J. Geophys. Res.*, **105**, 28255–28277, 2000.

Yokoyama, I., Volcanic eruptions triggered by tectonic earthquakes, *Bull. Geophys. Inst., Hokkaido Univ.*, **25**, 129–139, 1971 (in Japanese with English abstract).

M. M. Alam (e-mail: ahnaftahmid@yahoo.com) and M. Kimura