Application of Two-way Fixed Effects Regression on the Error Analysis of Seismic Magnitude Estimation: Hypotheses, Theory and Testing

Guo-quan Zhang[1]

1 Wenxian Emergency Management Bureau, Jiaozuo, Henan, China

https://orcid.org/0000-0002-9468-6132(GZ)

Correspondence to: Guo-quan Zhang, zgq083516@163.com

Key Points:

1. A statistical method is introduced to numerically solve the relative values of source, path and site effects
2. The magnitude uncertainty caused by source, path and site effects is determined and compared quantitatively for Japan’s KiK-net
3. A series of check tests are designed to examine the robustness of the decoupling algorithm

Abstract Seismic magnitude has been an indistinct concept since modern seismology was established. The author discusses it thoroughly , provides an general expression of definition that is only relevant to the source characteristics and explains the magnitude conversion problem physically. Based on the essential understanding of magnitude estimation problems, the author introduces a method called two-way fixed effects regression, which has been applied in seismology since Keiiti Aki studied the properties of coda waves. A series of check tests are carefully designed to examine the hypotheses of this method and then the author applies it to the error analysis of the peak ground acceleration magnitude. The source, path and site errors are estimated by this approach and the robustness of separation is checked. The acquired knowledge about the errors in magnitude estimation may help us to improve the precision of magnitude estimation during earthquake early warning.

Plain Language Summary

The purpose of seismology is to explain all the phenomena associated with earthquakes by studying the laws behind them. To find these laws, the first effort should be to describe the many earthquakes. However, in general, deep insight into these laws is needed to adequately characterize earthquakes. Seismic magnitude is a quantity introduced soon after modern seismology was established to categorize earthquakes but is still a concept that troubles geophysicists, as there are many fundamental seismological problems that have not yet been answered thoroughly. In this paper, the author tries to give a strict definition of magnitude and analyze the errors in it quantitatively from a statistical perspective. Two-way fixed effects regression allows part of the relation to be nonlinear; thus, it is an appropriate tool to explore the complicated laws behind earthquake phenomena. This paper explains the theoretical assumptions and physical meanings of this statistical technique and applies it to separate the fixed effects of different factors that affect the amplitude of seismic record.

1. Introduction

The seismic magnitude is an estimated parameter of the relative size of an earthquake (Bormann et al., 2013). The 'estimated parameter' here suggests that (1) this value contains a considerable error and often not the direct measure of the earthquake size; (2) this value is a parameter estimated from a certain sample (not the population), for instance, the amplitudes recorded by a local, nationwide, or even worldwide seismographic network but anyway from limited points on the ground; and (3) the magnitude of an earthquake may be multivalued, whereas the size should be only, as different observers can estimate the size from different perspectives, such as the duration of the quake (duration magnitude M_d), the maximum amplitude of the body waves (body wave magnitude M_b) or surface waves (surface wave magnitude M_s). However, the word 'size' in this definition is ambiguous and consensus-less. Bormann (2013) tends to regard the radiated energy (energy magnitude M_e) as the size of an earthquake, as this value relates directly to the potential damage of an earthquake. On the other hand, the United States Geology Survey (USGS) practically uses the work done by the stress (moment magnitude M_w) as the preferred published magnitude (USGS, 2023a). Moreover, the difference between these two fundamental physical quantities is a scale factor known as the seismic efficiency.

Although the seismic magnitude is statistical and error-significant, we have little knowledge about the statistical properties of the magnitude or its error. For a theoretical point-shaped source, the amplitude of a certain earthquake at a certain station is relevant to the size of the earthquake, the source spectrum and the dominant frequency of the observed component, the angle between the earth's surface and the fault plane or the focal mechanism, the attenuation on the propagation path, the radiation pattern and the direction of the station relevant to the source, and the site amplification of the station, while the current magnitude formula usually estimates the magnitude from the amplitude after nonexact and empirical corrections of attenuation. In other words, the errors from the prevailing magnitude estimating techniques are typically variable-omitting biases, as they take too few factors into account.

The bias from omitting source-related properties cannot be reduced by averaging the measured values of multiple stations, while it is popular that some official agencies and scholars treat the standard deviation as the uncertainty of the magnitude. On the website of the USGS, one can easily find such instances (USGS, 2023b). In fact, the uncertainty consists of fixed bias (systematic uncertainties) and stochastic error (random uncertainties) (Taylor, 2022), and the estimation of magnitude from multiple measurements is not definitely an unbiased estimation. Furthermore, different magnitudes of the same earthquake may be inconsistent. ('Consistency' here is also a statistic term. This means that the estimation converges to the true value as the quantity of samples increases. Obviously, the true value should be only. Therefore, we generally state that all the consistent estimations are consistent, as they have the same limit.) Seggern (1970) reported that the focal mechanism exerts a different influence on body and surface waves: the variation in the focal mechanism may enhance one while suppressing the other. This will be incomprehensible if both the body wave magnitude and the surface wave magnitude are consistent estimations of the earthquake size.

The bias in magnitude estimation is annoying but nearly inevitable. Physically, this is partly caused by the regular deployment of seismometers, which are all installed near the surface of the earth and violate the principle of random sampling. Mathematically, there are various methods to address fixed bias or endogeneity and obtain a consistent estimation. In this paper, the author will apply two-way fixed effects regression (2-FER) to the magnitude estimation to treat this endogeneity problem.

Numerous seismometers have acquired massive amounts of seismic data on earthquakes. However, during the regression of the magnitude estimating formula, geophysical researchers are attempting to determine a calibration function that is used to correct the attenuation on the path by records over decades and even centuries (Bormann, 2013). As discussed before, it will not help to obtain a fine estimation to simply increase the data volume and wastes the massive acquired information to finally obtain only one or two parameters in the equation. A large amount of data (called panel data in statistics in our problem) provides the possibility for conducting complex statistical methods and motivates us to mine the information inside the raw data and interpret it.

Some researchers have realized that simple linear regression is not a proper tool to analyze the magnitude problem, while they introduce orthogonal regressions to the magnitude regression problem (Bormann, 2013; Das et al., 2018). However, a good orthogonal regression requires knowledge of the error deviation ratio between the magnitude and the earthquake size (Das et al., 2018), but this knowledge is not yet quite clearly derived. The error analysis of magnitude is also important because without this knowledge, it is difficult to distinguish errors and fallacies. For example, the difference in the local magnitudes of earthquakes on the junction between Sichuan and Yunnan (in Changning County in particular, where shale gas is extracted) given by these two provincial networks is sometimes as large as 0.6. Is this difference a fallacy? Or what extent of difference is acceptable? The heads of the Instrument Operation and Maintenance Groups of these two provinces both denied that they had set wrong instrument responses in the software and they did have incentives to hide their faults if they actually existed.

Fixed effects regression (FER) is a popular statistical method (Allison, 2009) in many applications, such as econometrics (Greene, 2011) and biostatistics (Diggle, 2002; Gardiner, 2009), to treat some regression problems of panel data or longitudinal data, as it is simple but effective. Seismic observations naturally generate panel data, as we observe identical earthquakes by multiple stations, which constitute a matrix, a table or a panel. Numerous geophysicists have adopted fixed effects models spontaneously to address seismology problems but may not specifically consult related statistical materials (Phillips & Aki,1970; Joyner & Boore 1981; Brillinger & Preisler, 1984; Andrews, 1986; Iwata & Irikura, 1986; Takemura et al., 1991; Moya et al., 2003;Wang & Shearer, 2019; Torres-Sánchez & Castro, 2023 and so on).

The above papers are focused on the research of coda waves, P-waves or S-waves, which are recorded by a strong motion network or broadband network. Certainly, there is no essential distinction between the strong motion network and the Earthquake Early Warning (EEW) network. Therefore, the relevant method and conclusions are also applicable to EEW. For instance, the major revision of the EEW magnitude estimation formula by JMA (Aketagawa et al., 2010) was heavily based on the conclusions of Joyner & Boore (1981) from strong motion records.

Magnitude estimating techniques in EEW are mainly traditionally semiphysical (amplitude-based [Wu & Zhao, 2006], period-based [Allen & Kanamori, 2003; Atefi et al., 2017] and combined [Wang et al., 2022]) or purely mathematical or statistical (artificially intelligent [AI] [Song et al., 2023; Zhu et al., 2023, Furumura & Oishi, 2023]). The result of AI methods looks much better than that based on traditional seismology and shows great potential in EEW. On the other hand, it is debatable whether one can determine the magnitude before rupture completion (Münchmeyer et al., 2022). People hope that early rupture signals can predict the final earthquake size because of its enormous social value and have obtained results as quickly as AI methods (Colombelli et al., 2014, 2020), but others argue that this optimistic result is an unreal vision due to the selection of a small data set and accidental errors (Trugman et al., 2019; Meier et al., 2021) and oppose rupture determinism in a probabilistic view. If the fallacy of Colombelli (2014, 2020) is caused by the careless treatment of error, then is it possible that the evidence of earthquake evolution similarity is caused by noisy observations and is not as reliable as the noise buries precursors in return? Otherwise, why is the AI's result so astonishing? Did the AI methods cheat? Or are there some empirical laws of the rupture process still unrecognized? What is the upper limit of the performance of the EEW system? What's the theoretical form of the optimal estimation of the earthquake size or the ground motion with limited information? Can we obtain it in a comprehensive way? Or can the traditional methods generate results as fast and as accurately as the current AI at least? Would AI completely replace the weak-looking traditional seismological methods?

The author will not answer the many debatable questions above in this paper. Instead, the author will focus on the property of the magnitude error as a first step toward these answers, try to reduce the error of the traditional magnitude estimating method, and take the peak ground acceleration (PGA) recorded by the KiK-net in Japan to illustrate it.

2. Data

The author selects 228 earthquakes with magnitudes larger than 4 (Figure 1a) recorded by the KiK-net in Japan (National Research Institute for Earth Science and Disaster Resilience, 2019). The KiK-net is composed of 699 stations (Figure 1b) with paired seismometers (one in the borehole and another on the ground). Trugman (2019) selected the population of observed earthquakes to avoid the suspicion of selecting a particular data set or to ensure the objectivity of his conclusions. However, to cover the entire of Japan as evenly as possible and avoid too large weights of small earthquakes, the author selects 228 earthquakes manually preferably in the area where earthquakes occur infrequently and selects fewer than 6 earthquakes in every 0.1 magnitude unit (m.u.). Unilaterally, the author thinks this subjective selection will not harm the objectivity of this research. The program of this research is open (please refer to the Data Availability Statement), and it is easy to check the result on the entire data set if one has such a data set in hand. Moreover, the source parameters and peak ground acceleration (PGA) in the following study are from the files that are directly downloaded from the official K- & KiK-net website. The author adopts them without roundly checking and believes the JMA completely.

In addition, 6 channels at each station consist of KiK-net. Channels 1-3 constitute the seismometer in the borehole, while Channels 4-6 constitute the seismometer on the ground. Channels 1 and 4 record the NS component, Channels 2 and 5 record the EW component, and Channels 3 and 6 record the UD component.

Why most researchers prefer to estimate the magnitude by peak ground displacement (PGD) (for example, Wu & Zhao, 2006) but the author of this paper chooses the PGA? PGD actually estimates the magnitude better than peak ground velocity (PGV), and PGV does it better than PGA (Zhang et al., 2023). We know the PGD is the integration of PGV and PGV the integration of PGA, so will it be better if we continue to integrate PGD? What will happen if we integrate unlimited times? These questions relate to a question we raised in the Introduction, namely, what is the optimal estimation of the earthquake size. The author believes that the answer lies in another domain, for example, the frequency domain maybe.

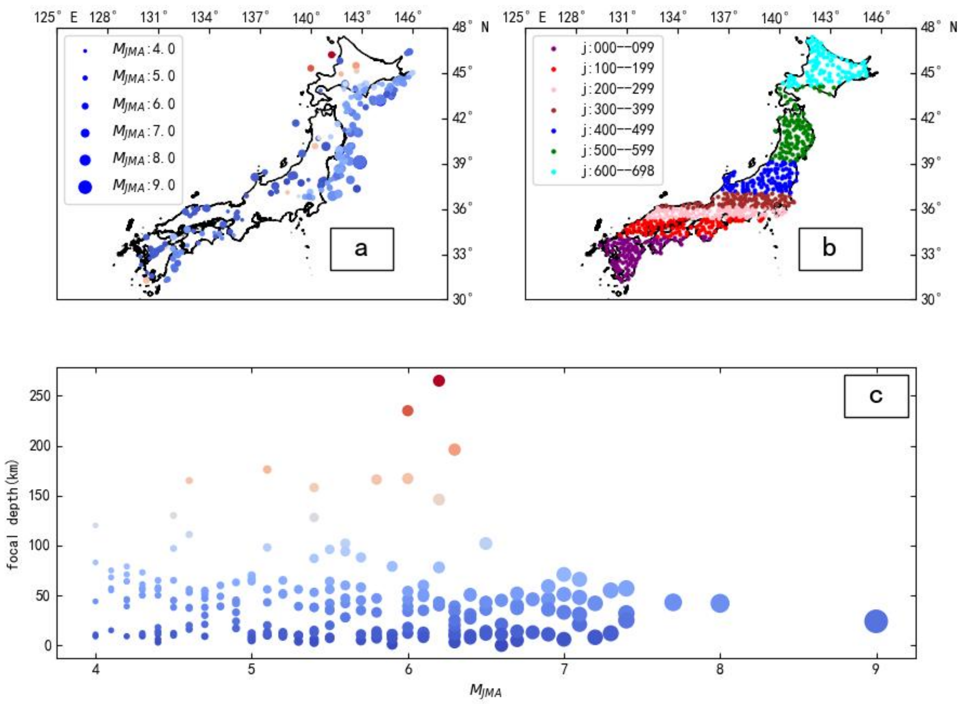

Figure 1. (a) Spatial distribution of the 228 selected earthquakes. The warmer color indicates a deeper focal depth (see subplot [c]). (b) Spatial distribution of the 699 stations in KiK-net. They are numbered from south to north. (c) The depths of the 228 earthquakes determined by JMA. The warmer color indicates a deeper focal depth.

We know that an integration operator is transformed to a division operator by Fourier transform, which is a characteristic that is often used to solve differential equations, including the wave equation:

$$g(t) = \int h(t)\,dt \Leftrightarrow G(f) = \frac{H(f)}{f}$$

where G is the Fourier transform of g and H is the Fourier transform of h. Therefore, if h is monofrequent, integrations of h are meaningless and abundant as we multiply a constant before it each time. However, the real seismic signal is multifrequent, and its integration will suppress the high-frequency component due to a larger divisor. Unlimited integrations of the seismic signal will make it converge to the amplitude at 0 frequency (M_w in fact). Therefore, the author of this paper thinks integration is a fancy operation in magnitude estimation and makes it more complicated. If it is truly necessary, then filtering can replace it completely. The integration estimates a more accurate magnitude because of the narrower signal bandwidth or lower frequency component. It is not clear which of these two factors is dominant. Here, we just shallowly analyzed this issue and will avoid touching the essence. Below, we concentrate on the purpose of this paper.

3. Mathematical Modeling

We assume that the amplitude of earthquake i at station j in the frequency domain is:

$$A_{ij}(f) = S_i(f)P_{ij}(f)L_j(f), \text{ for each } (i,j) \text{ in } I \times J \quad (1)$$

where $A_{ij}(f)$ can be the observed maximum amplitude of a certain phase, f is the frequency, $S_i(f)$ is the source effect of earthquake i $(i = 1, 2, 3...I)$, $L_j(f)$ is the site effect at station j $(j = 1, 2, 3...J)$, $P_{ij}(f)$ is the path effect determined by both earthquake i and station j, and set $I \times J$ is the population of integer pairs (i,j) (Figure 2a).

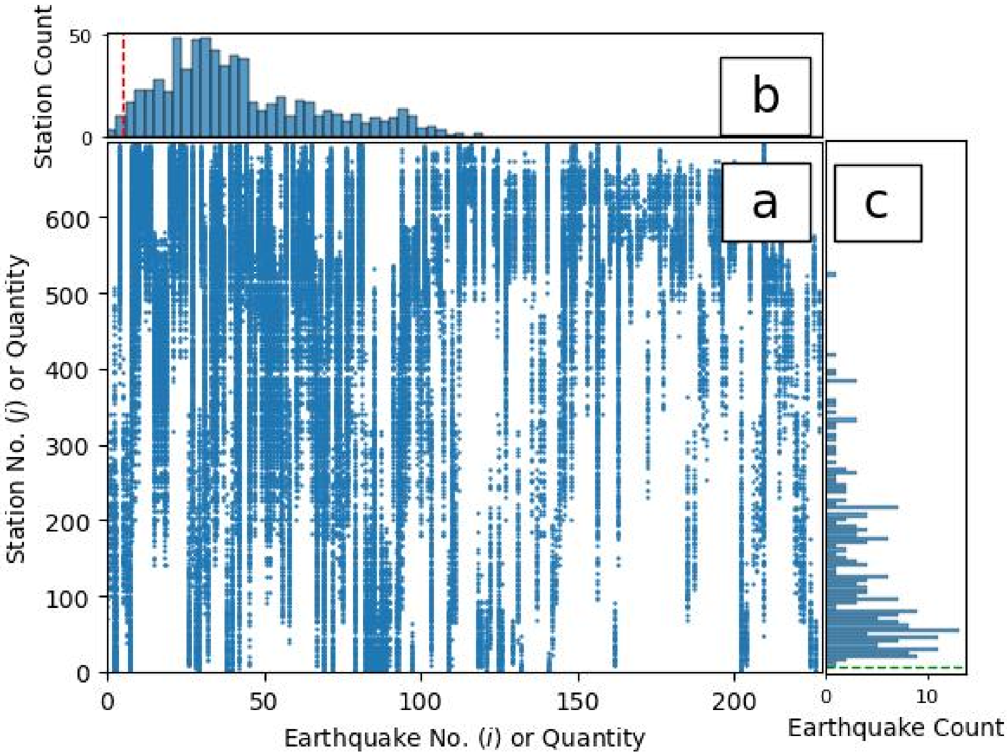

Figure 2. The joint distribution and marginal distribution of the 228 earthquakes and the 699 stations in Figure 1. (a) The scatter plot of integer pair (i, j) in which i is the earthquake number and j is the station number. (b) The distribution describing how many earthquakes are recorded by the stations. The x axis is the earthquake quantity, the y axis is the station count corresponding to it, and point (x, y) means there are y stations that record exactly x earthquakes. There are 13 stations that record fewer than 5 earthquakes (stations numbered 4, 94, 100, 105, 131, 165, 210, 273, 287, 321, 376, 430 and 515, namely, the bars on the right of the red dotted line). (c) The distribution describing how many stations that one earthquake triggered. The y axis is the station quantity, the x axis is the earthquake count corresponding to it, and point (x, y) means there are x earthquakes that are recorded by exactly y stations. All earthquakes are recorded by more than 10 stations (it is blank below the green dotted line).

In this paper, the author will try not to talk about frequency. We assume that the frequency is a constant during the following derivation.

Taking the logarithm of both sides of Equation (1) and ignoring the frequency, we obtain:

$$\lg A_{ij} = s_i + p_{ij} + l_j, \text{ for each } (i, j) \text{ in } I \times J \quad (2)$$

where $s_i = \lg S_i$, $r_i = \lg R_i$ and $p_{ij} = \lg P_{ij}$.

The real site amplification is also related to the direction of incidence of the waves (Papageorgiou, 1991), so it relates both to i and j. Denote the mean of site amplification or site effect as l_j, and we assume l_j is a fixed value or does not change much with i:

$$l_j = \overline{l_{ij}^{i}} = l_{ij} - \varepsilon_{ij}^{l} \quad (3)$$

We denote:

$$p_{ij} = g_{ij} + h_{ij} + \varepsilon_{ij}^{p} + \varepsilon_{ij}^{l} \quad (4)$$

where g_{ij} is the geometrical spreading, h_{ij} is the anelastic attenuation and ε_{ij}^{p} is an error term caused by the omission of the radiation pattern, as this factor is determined both by i and j. Certainly, the 2-D radiation pattern error is related to the orientation of the path vector, so it is comprehensible to classify this omitted variable bias as part of the path effect.

The geometrical spreading g_{ij} is supposed to be dominated by the hypocentral distance r_{ij} if the source is simply point-shaped:

$$g_{ij} = g(r_{ij}) + \varepsilon_{ij}^{g} \quad (5)$$

where ε_{ij}^{g} is the uncertainty caused by the irregular shape of the earth.

Similarly, the anelastic attenuation h_{ij} is also supposed to be dominated by the hypocentral distance r_{ij} if the source is simply point-shaped:

$$h_{ij} = h(r_{ij}) + \varepsilon_{ij}^{h} \tag{6}$$

where ε_{ij}^{h} is the uncertainty caused by the heterogeneity and anisotropy of the earth.

We denote the attenuation function as:

$$p(r_{ij}) = g(r_{ij}) + h(r_{ij}) \tag{7}$$

and define the magnitude corresponding to the selected phase as:

$$M_i = s_i - p(0) \tag{8}$$

We rigorousize the seismic magnitude definition in consideration of (a) Richter (1935) used the amplitude recorded by the stations at an epicentral distance of 100 kilometers, and this value is influenced not only by the earthquake size but also by the local attenuation within 100 km, so we use s_i, which is the logarithmic amplitude at the source, to characterize the earthquake size instead. (b) The limit of the amplitude at the source is infinite, so we subtract it by another infinitely large quantity $p(0)$ to obtain a finite quantity.

As previously mentioned, magnitude is an estimation of the earthquake size S_i:

$$M_i = d(S_i) \tag{9}$$

where $d(*)$ is a function to be determined. Certainly, S_i can be regarded as another kind of magnitude scale, and we usually presume that the relation between different magnitude scales is linear (Das, 2018):

$$d(S_i) = \alpha S_i + \beta + \varepsilon_i \tag{10}$$

where α and β are constants and ε_i is the uncertainty caused by the source spectrum and focal mechanism for a point-shaped source, as discussed in the Introduction. However, if one of the two is saturated, the linearity may be broken. Furthermore, please note that if the fault exhibits a certain scale, the rupture process is more appropriate to describe the earthquake than the source spectrum theory, so Equation (10) should be reconsidered.

As discussed in the Introduction, there is currently no strict definition of earthquake size. In this paper, regarding the convenience of our study, we adopt the magnitude published by JMA as the earthquake size:

$$S_i = M_{\mathrm{JMA}\_i} \tag{11}$$

Equation (2) is clearly undetermined because there are more variables to be determined than observed variables. The key is that the amount of p_{ij} is as large as $\lg A_{ij}$. However, we know that not all the solutions of Equation (2) are physically reasonable. At least, $g(r_{ij})$ and $h(r_{ij})$ should look well regulated. Castro adds a smooth condition on p_{ij} to solve Equation (2) (for the latest example, Torres-Sánchez & Castro, 2023). However, because p_{ij} contains error and is not smooth at all, Castro plays a good trick in that he selects a portion of p_{ij} and requires the distance between two closest p_{ij} larger than 10 km to ensure that p_{ij} is almost smooth. As Castro represents the attenuation or calibration function by a sequence rather than a function that contains only one or two parameters, this method is named the nonparametric method. The disadvantage of the nonparametric method is that the acquired calibration function is applicable only to the distance within which stations exist and cannot be extrapolated, and as the limitation of distance between p_{ij}, it seems difficult to reduce random error by increasing the data volume, at least it has not fully utilized the data. On the other hand, the parametric method is much more popular and presupposes the functional forms of the calibration function from physical models or intuitions. By replacing p_{ij} with only one or two unknown parameters, Equation (2) will be much easier to solve. However, this method is subjective and not strict most of the time, which makes the extrapolation poor. In this paper, we adopt a parametric method and will show that our residual sequence is stationary to the distance, and then the fitting function is supposed to be extrapolated well.

Equation (2) is also undetermined because its coefficient matrix is rank deficient. This feature is well specified in most FER and dummy variable regression materials. We just briefly explain it: if s_i, l_j, p_{ij} is a set of solutions of Equation (2), then $s_i - \gamma, l_j + \gamma - \delta, p_{ij} + \delta$ also satisfies Equation (2). Certainly, $h(0)$ should be 0, and it seems that $p_{ij|r_{ij}=0} - g(0)$ should be small, but there is still one degree of freedom in it. That is, the solution space of Equation (2) is at least one-dimensional rather than a point. As only the relative values of s_i and l_j can be determined by Equation (2), if one would like to obtain the absolute value, one should achieve it by other means, for example, part of feathers of the source spectrum (Moya et al., 2003; Wang & Shearer, 2019) or the site amplification of a certain station (Andrews, 1986) is known. As the frequency is not discussed in this paper and the magnitude is a relative value, we arbitrarily set $\gamma = \overline{l_j}$, as in Torres-Sánchez & Castro (2023).

We denote the path error as:

$$\varepsilon_{ij} = \varepsilon_{ij}^{p} + \varepsilon_{ij}^{g} + \varepsilon_{ij}^{h} + \varepsilon_{ij}^{r} \tag{12}$$

Assume:

$$g(r_{ij}) =- a \lg r_{ij} \tag{13}$$

and:

$$h(r_{ij}) =- b r_{ij} \tag{14}$$

Rewrite Equation (2):

$$\lg A_{ij} = s_i + l_j - a \lg r_{ij} - b r_{ij} + \varepsilon_{ij} \tag{15}$$

Expand each term in Equation (15):

$$\left[\begin{array}{cc|cccc|cccc}
-\lg r_{11} & -r_{11} & 1 & & & & 1 & & & \\
-\lg r_{12} & -r_{12} & 1 & & & & & 1 & & \\
\vdots & \vdots & \vdots & & & & & & \ddots & \\
-\lg r_{1J} & -r_{1J} & 1 & & & & & & & 1 \\
\hline
-\lg r_{21} & -r_{21} & & 1 & & & 1 & & & \\
-\lg r_{22} & -r_{22} & & 1 & & & & 1 & & \\
\vdots & \vdots & & \vdots & & & & & \ddots & \\
-\lg r_{2J} & -r_{2J} & & 1 & & & & & & 1 \\
\hline
\vdots & \vdots & & & \ddots & & \vdots & \vdots & \vdots & \vdots \\
\hline
-\lg r_{I1} & -r_{I1} & & & & 1 & 1 & & & \\
-\lg r_{I2} & -r_{I2} & & & & 1 & & 1 & & \\
\vdots & \vdots & & & & \vdots & & & \ddots & \\
-\lg r_{IJ} & -r_{IJ} & & & & 1 & & & & 1
\end{array}\right]
\left[\begin{array}{c} a \\ b \\ \hline s_1 \\ s_2 \\ \vdots \\ s_I \\ \hline r_1 \\ r_2 \\ \vdots \\ r_J \end{array}\right]
=
\left[\begin{array}{c} \lg A_{11} \\ \lg A_{12} \\ \vdots \\ \lg A_{1J} \\ \hline \lg A_{21} \\ \lg A_{22} \\ \vdots \\ \lg A_{2J} \\ \hline \vdots \\ \hline \lg A_{I1} \\ \lg A_{I2} \\ \vdots \\ \lg A_{IJ} \end{array}\right] \tag{16}$$

$\lg(r_{ij})$ and r_{ij} are certainly dependent, although not linearly dependent. Regardless, this will still lead to a matter called colinearity, as we have explained previously (Zhang, et al., 2023). The author will elaborate on this matter later in the Test of Path Effect. a、b and s_i、r_j are heterogeneous, and it looks strange to put them together. Therefore, a weighted regression is more reasonable, but for the sake of simplicity, we still choose an ordinary regression below. We can solve the underdetermined Equation (16) by determining the generalized inverse of its coefficient matrix or singular value decomposition (Parolai et al., 2004). The coefficient matrix is a sparse matrix and can be solved quickly with an ordinary computer. Particularly, to achieve a stable solution, the author solves it by ridge regression, i.e., Tikhonov regularization with a ridge parameter of $\lambda = 0.0001$ in this paper.

Below, we denote the normalized coefficient as the original symbol with an apostrophe such as a', b', β', ε_i', ε_j', ε_{ij}' and so on. They are equal to the original variables divided by α except $\alpha' = 1/\alpha$.

4. Check Tests

When we model the magnitude regression problem, we simplify it by many ideal theoretical hypotheses that are almost impossible to satisfy strictly in practice. It is essential to design tests to check the influence of these dissatisfactions and to what extent these hypotheses are approximately satisfied. The basic idea is very simple: no observations, no testing. As we observe identical physical quantities much more times than necessary, we can design check tests and infer the influence of the unmeasured variables from these tests. In addition, without analyzing the properties of the coefficient matrix and simply from artificial experiments, we can infer the properties of the above model.

4.1. Test of Source Effect

In the Introduction, we have pointed out that the traditional earthquake magnitude estimation method takes too few variables into account. However, its results have been used for a very long time, and they seem reliable to some extent. How large is the difference between them? Is improvement necessary? In this part, we try to quantitatively analyze this issue. As there is no fixed effect in the traditional method, we denote it as the 0-FER method. The method that considers only the source effect and ignores the station effect (1-FER) is detailed in our former paper (Zhang et al., 2023). Their statistical hypotheses are detailed in Appendix A.

Figure 3 a-c give the results of the a posteriori estimate of magnitudes in the three models, namely, the normalized source term $(s_i - \beta)/\alpha$. As the distribution of the source error is nonnormal and even nonsymmetric (Figure 3 e-f), we adopt median regression in the 1-FER and 2-FER models instead of LSM regression. The LSM regression will produce the minimum deviation $D\varepsilon_i$, but the author is afraid that the result is biased in such cases. As more effects are taken into account in the models, the regression seems more reliable with less $D\varepsilon_i$ and a good-looking distribution of ε_i. In other words, the discrete points are more concentrated. However, the shapes of discrete points are similar in Figure 3 a-c. This result indicates that the energy of the site effect will leak to the source effect if not corrected; however, this leak is stochastic but not *i*-uncorrelated, which means that the dotted line will be deflected with an unknown but possibly small angle (i.e., the varying slope α). In addition, this means that the additional conditions $\sum_j l_j = 0$ and $\sum_i s_i = 0$ in condition (AC3) are not strictly met, so $D\varepsilon_i$ differs slightly. Please note that the custom deviation is generated by the difference between the measured data and its mean, and it assumes that the true value is unmeasurable, while the root mean square error (RMSE) is generated by the difference between the measured data and the true value (namely, M_{JMA}). The relation between deviation $D\varepsilon$ and RMSE $\sqrt{E\varepsilon^2}$ is:

$$E\varepsilon^2 = D\varepsilon + (E\varepsilon)^2$$

The above equation shows that the error consists of two parts: a random error $D\varepsilon$ and a fixed bias $E\varepsilon$.

The linearity of discrete points in Figure 3 is obvious. We have found that in Yunnan, the P-wave EEW magnitude is not saturated even at a surface wave magnitude of 6.4 (Zhang et al., 2023). The author has been expecting to observe a plateau level and preparing to adopt a nonlinear regression and extremely worried how to estimate the earthquake size in Yunnan as he supposed the P-wave EEW magnitude would saturate at approximately 5 m.u. If so, there will be no P-wave maximum amplitude difference between an earthquake of 6 m.u. and 7 m.u. However, now, it seems an unnecessary worry. In fact, the results of JMA also violate the saturation theory (Aketagawa et al., 2010; Kiyomoto et al., 2010). Melgar (2015) noted that the PGD from GPS will not be saturated. The further discussion of magnitude saturation is beyond the purpose of this paper. We may discuss this issue in detail at a future date.

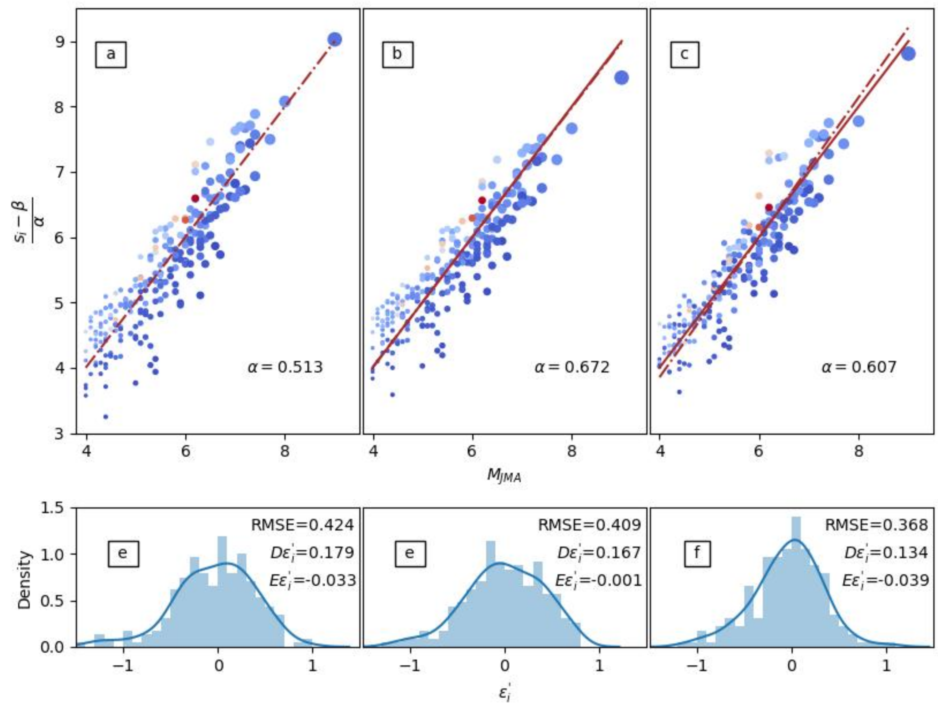

Figure 3. (a-c) The regression of the source effect of the 0-FER, 1-FER and 2-FER models. The dotted lines are LSM regression models, while the solid lines are median regression models. (e-f) The source term error of the 0-FER, 1-FER and 2-FER models, i.e., the distance between the discrete points and the dotted line in the corresponding subplot above. Except for special instructions, all the results below are from Channel 3, i.e., the UD direction of the seismometers in the borehole.

If the source error, site error (we regard the whole site term as error here as generally it is not corrected in conventional magnitude estimating) and the path error are randomly combined, in statistics, we have the conservation of energy:

$$D(\varepsilon_i + r_j + \varepsilon_{ij}) = D\varepsilon_i + Dr_j + D\varepsilon_{ij}$$

In our 2-FER model, the above equation is approximately held (Figure 4e):

$$0.340 \approx 1.03 * 0.340 = 0.349 = 0.134 + 0.097 + 0.118$$

The coefficient 1.03 indicates that the endogeneity has not been eliminated clean; otherwise, it will be 1. Energy leakage still exists. The author thinks it is caused primarily by the nonrandom pairing of i and j (Figure 1). The total error of magnitude is not normal (black line in Figure 5), and we have pointed out that it is not quite proper to represent the error by only its standard error as if it was a normal distribution (Zhang et al., 2023). By the 2-FER technique, the author thinks that the three error components are successfully separated, and the path error looks normal, as expected (Figure 4d). Therefore, he estimates that the three errors contribute comparably to the uncertainty of magnitude, and it is necessary to handle any of them cautiously. It is quite imprecise to evaluate the magnitude uncertainty of an earthquake by its deviation, but RMSE will be a much better evaluation. Although it is difficult to obtain the RMSE just at the moment the earthquake occurs, their confounding will lead to an overoptimistic estimation of error.

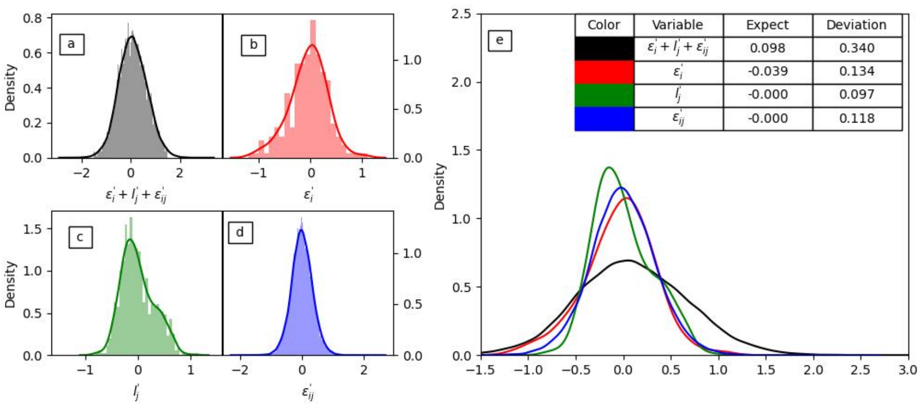

Figure 4. (a-d) The distributions of the 4 variables listed in the table of subplot (e) and their parameters. (e) Comparison and parameters of the 4 distributions.

The source and path errors cannot be corrected when an earthquake just occurs. If one applies the above three models to the magnitude calculation, only the site effect and path attenuation can be corrected. We denote the left error as ε. Please note that the relation between the parameters of the 2-FER model in Figure 4 and of Figure 5 is:

$$D\varepsilon = D(\varepsilon_i + \varepsilon_{ij}) \approx D\varepsilon_i + D\varepsilon_{ij}$$

The magnitude error energy (deviation) drops considerably after the source effect is introduced. Surprisingly, the LSM (0-FER) model generates a result with a larger error. This is typically due to the regression's objective, and the author has explained it in Appendix A. Another explanation is that the raw data are endogenous and the basic assumption of LSM is not satisfied. For a long time, we used LSM to address magnitude problems, pinning our hope on increasing the data volume and long-term observations, but without seriously checking whether the hypotheses of the model are met on earth.

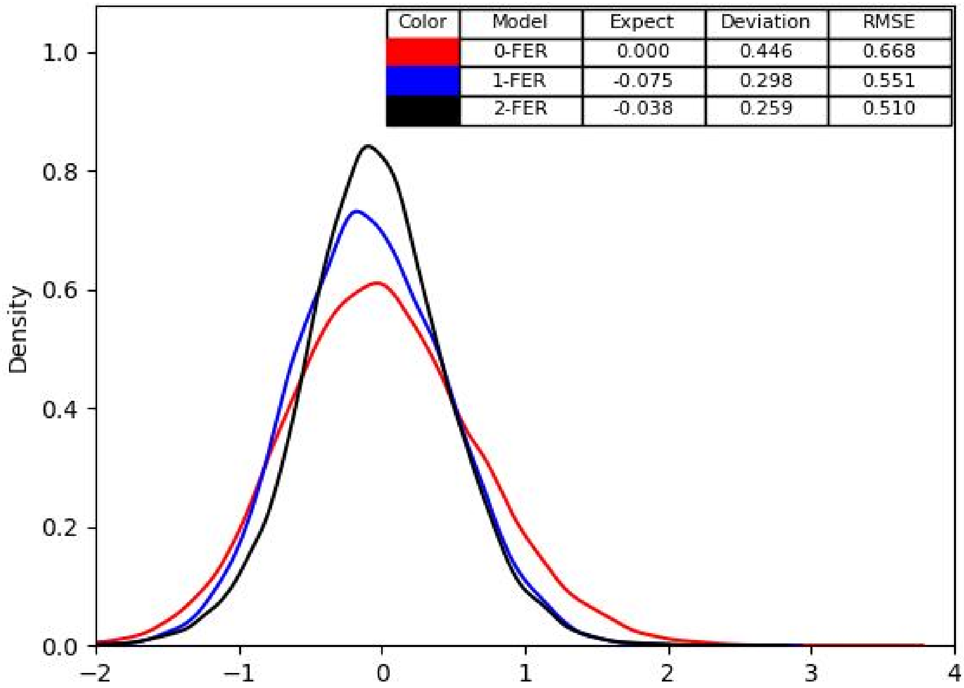

Figure 5. The distributions and single-station magnitude errors and their parameters of the three models.

The plot of errors versus distance (Figure 6) explains why LSM (0-FER) cannot generate a result with the minimum error from another perspective. The distance between the discrete points and the attenuation curve is the error terms in the brackets of Models (AM0), (AM1) and (AM2), i.e., the hypothetical path error. In the 0-FER model (Figure 6a) and 1-FER model (Figure 6b), obviously, the attenuation curve does not pass through the center of the discrete points. This means that the path effect is not corrected appropriately, and the close station will underestimate the magnitude, while the far station will overestimate it. To achieve a steady result, we will be forced to calculate the station at various distances to average the errors. The distortion is due to the leakage of energy of source and site terms and makes the total error in Figure 5 enlarged. The 2-FER model overcomes this shortcoming best, and as the curve passes through the center of the points (Figure 6c), we can deduce that this form of attenuation function is suitable for solving our physical problem.

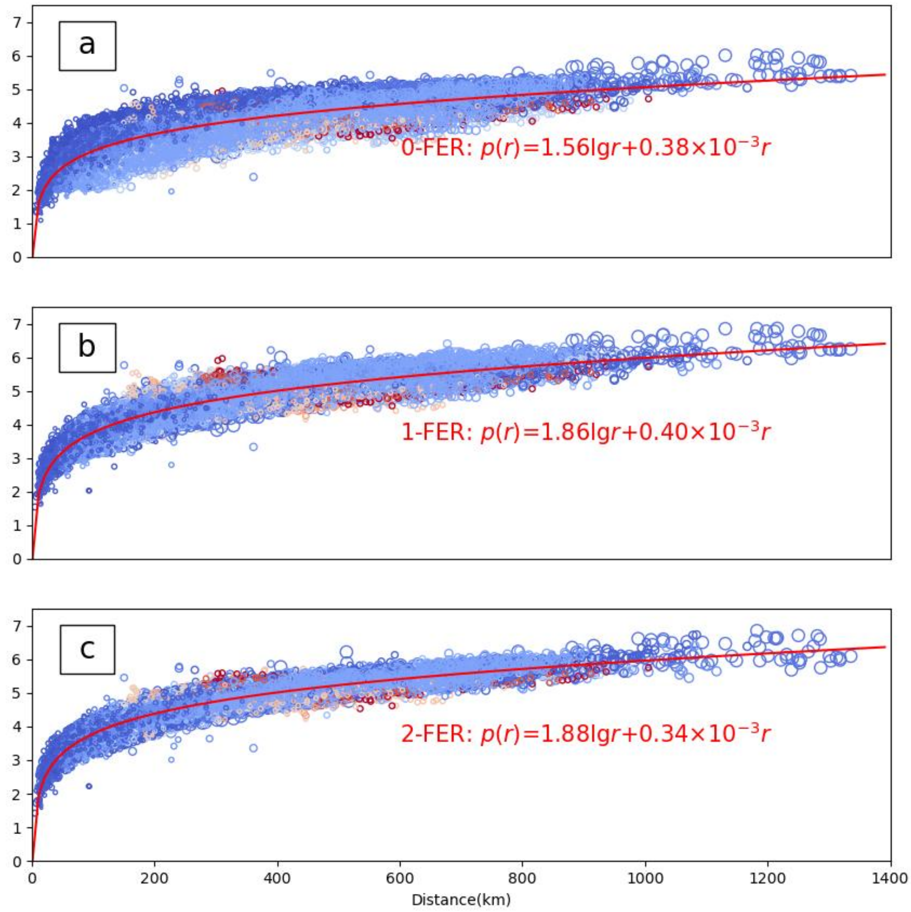

Figure 6. The regression of attenuation functions (solid red curve) of the three models. The colors of the points are the same as those in Figure 1c, indicating the depth of the source.

Regarding Figure 6 as evidence, the author asserts that with suitable modeling, the effects of source, path and site can be reliably separated. As the distance increases, the variance becomes homogeneous, and the stochastic process becomes stationary. After attenuation is compensated, it is not necessary to reject far stations for their precision.

4.2. Test of Site Effect

We know that the model with more parameters tends to produce more accurate results naturally mathematically, and a good result does not definitely indicate that a model is more physically reasonable. We call this phenomenon overfitting. In this part, we focus on the stability of l_j or the change in l_j over time and evaluate the overfitting by the model's generalization ability.

One of our important hypotheses is that l_j is steady and does not vary with i too much. We check this hypothesis by removing the data after 2015. Then, 13759 of 29290 records are removed. As mentioned in Appendix A, more station effects will be unreliable as the data lessen, and the ridge regression of unreliable stations will generate values close to 0. After these stations are removed, 661 stations remain. The difference between the site effects in the two sets is approximately 2*0.003/(0.037+0.039)=7.9% (Figure 7b). There is a strong linear correlation between them (Figure 7a), which indicates that the result before 2015 is applicable up to 2023. Thus, we prove that the stability assumption holds approximately.

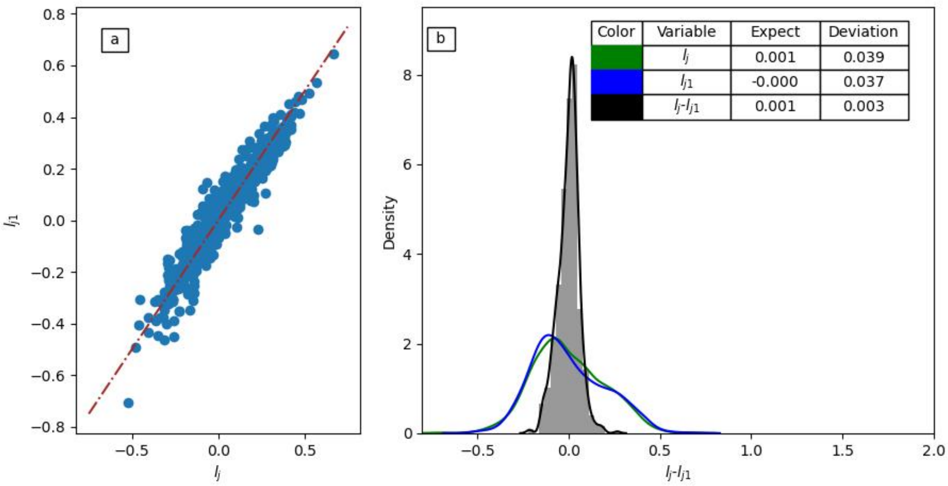

Figure 7. (a) The relation between the subset site effect l_{j1} and complete set site effect l_j. (b) The difference between l_j and l_{j1}.

4.3. Test of Path Effect

As previously mentioned, we use only two parameters to characterize the attenuation. Conversely, the nonparameter method introduces a large number of parameters. The real attenuation may be piecewise (Motaghi & Ghods, 2012), and it may not be enough to describe the attenuation by two parameters. This will result in another problem - underfitting. In addition, the effects of source, path and site are coupled, which means inappropriate assumptions on one of them will influence the others, and results in that the attenuation function cannot approximate the path effect well.

The attenuation of three components of one seismometer should be the same if the effect separating makes sense. We have checked this for the seismometers on the ground, and the coefficients in $p(r_{ij})$ change slightly. An accident occurred for the seismometers in the borehole: the coefficients changed slightly significantly (Figure 8 a-c). However, the relative trends of the curves are comparable (Figure 8d). It reminds us that only the relative value of three effects can be determined as mentioned in Mathematical Modeling. The individual parameter in $p(r_{ij})$ does not make sense physically, and it is only an approximation for the real attenuation. One should not regard the anelastic attenuation coefficient in this paper as the real property of the Earth's crust. Due to the collinearity of $g(r_{ij})$ and $h(r_{ij})$, we cannot separate them numerically. In the three channels, the regressed attenuation curves pass through the center of the points. The author has tried to fix the coefficient before $\lg r_{ij}$ at one, and then the attenuation curve does not coincide with the curve of piecewise medians, i.e., it does not pass through the center. Therefore, not all assumptions of the form of the attenuation function can ensure that the regressed curve performs well with distance r_{ij}. Judged by our naked eyes, it appears that the underfitting in Figure 8 is slight and acceptable. Thus, subjectively, the author believes that two parameters are sufficient to characterize the attenuation. The parameter method in this paper does not work worse than the nonparameter method.

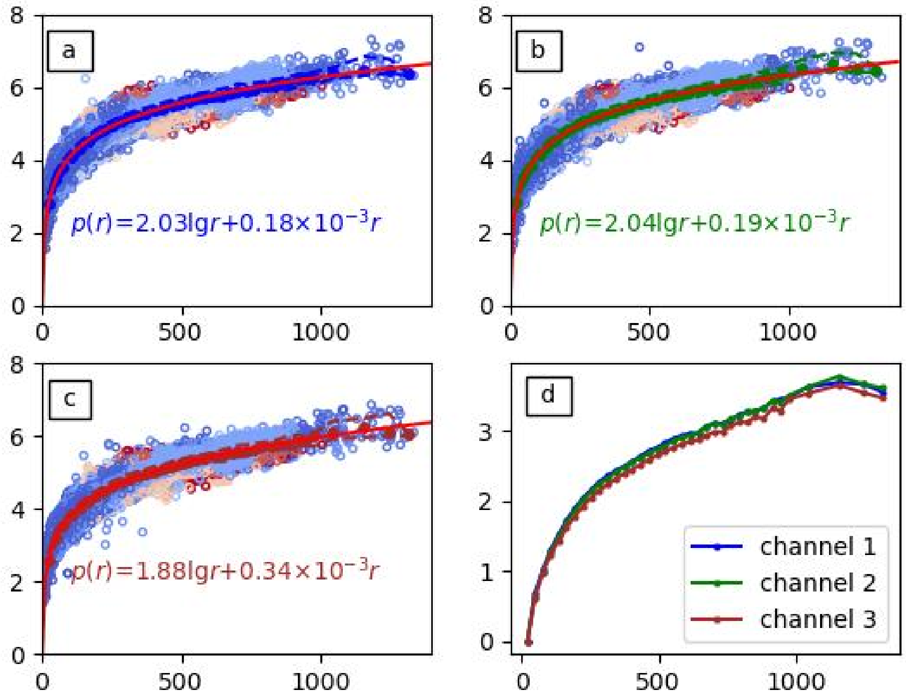

Figure 8. (a-c) The regression of the attenuation of Channels 1, 2 and 3. The red line represents the regressed attenuation, the line with circle markers represents the piecewise median regression, and the dotted lines represent the upper and lower quartiles. (d) The comparison of three piecewise median regressions that have been moved down so start at 0.

5. Extension and Discussion

(1) In this paper, we adopt the 2-FER model to obtain an estimation of the earthquake size, i.e., the magnitude of JMA. The author used the complete set and a subset to show the stability of the site effect, because the author found the complementary set is of low quality by plotting a figure like Figure 2. One should ensure that an earthquake contains enough stations and that a station is covered enough times before adopting the 2-FER model so that condition (AC3) is satisfied. Otherwise, one should choose a less precise model such as 1-FER or 0-FER, especially when the data volume is finite.

(2) The author has avoided talking all the issues about frequency in this paper, but it does not mean it is not important. Instead, perhaps it is closer to the essence of the

magnitude estimation issue as we analyzed in Data. Initially, the author supposed that only the site effects of the two seismometers are different, and those of the source and path are identical. However, the separated path effects are different (Figure 9a). The author hypothesized that this should be due to the different dominant frequency components for seismometers in the borehole and on the ground. If so, the difference in the attenuation should mainly be due to the anelastic attenuation and should be linear. However, it seems to be logarithmic (Figure 9b). We leave this issue to the future.

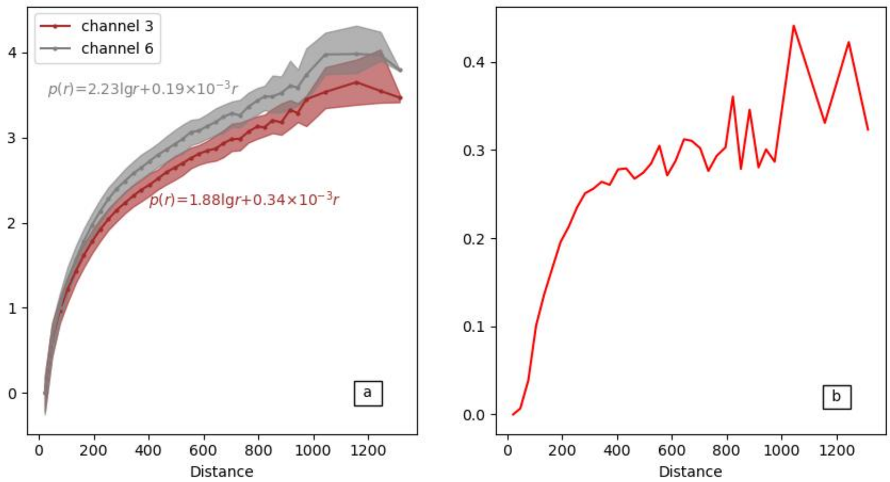

Figure 9. (a) The piecewise attenuation function of Channels 3 and 6. The shaded area is between the upper and lower quartiles, i.e., 50% of the points fall in this area. (b) The difference between the attenuation of channel 6 and channel 3.

(3) We know that the site effect is caused by coherence emphasis and is strongly related to sediment thickness and wave frequency. In this paper, the site effect seems not very prominent, but that is possibly due to the instrument characteristics. When estimating the magnitude, we prefer a frequency at which the site effect is weak; however, when estimating the ground motion, we are more concerned about the frequency at which the site effect is strong. On the other hand, for large earthquakes, it is not enough to simplify the source to a point. Böse et al. (2012, 2018) applied an image identification method to estimate the fault length in real time. The author

especially stresses which equations are based on the point source assumption in Mathematical Modeling. It is everyone's aim who specializes in EEW to predict ground motion in real time. The author personally hopes these careful assumptions will be conducive to his future research. JMA has applied the site effect correction of PGD to P-wave attenuation and found virtually no use (Kiyomoto et al., 2010). This is expected as the dominant frequency is different. However, it is essential to design tests to study whether the site effects of P-waves and S-waves at the same frequency are identical.

(4) We have mentioned the collinearity of $g(r_{ij})$ and $h(r_{ij})$ in our former works (Zhang et al., 2023) and wishfully thought the coefficient in $g(r_{ij})$ should be 1 from a point source model in the homogeneous medium of a semi-infinite half-space. The author apologizes that this is a mistake, and it is proven to be wrong by the mid-field data of Japan's KiK-net. Please note that the coefficients do not make sense in physics in his papers actually; they are actually a mathematical approximation.

(5) The RMSE of the multistation-average magnitude of the PGD magnitude formula by JMA is 0.309 after correction of the site effect and 0.319 before (Kiyomoto et al., 2010). The corresponding error to this paper is approximately 0.368. The author is not sure if JMA calculates the magnitude after simulation, but according to the published literature, its initial equation is from the manual earthquake catalog (Aketagawa, 2010). Regardless, the author infers that the estimation precision will be higher in the frequency domain, both because the original physical quantities are frequency related and α of long-period waves is larger. This paper focuses on the basic properties of errors in magnitude estimation, and it is still far from the optimal estimation of magnitude. Different magnitudes are measured with different components in the waves; thus, the magnitude estimation technique actually estimates one component according to the measurement of another component. In addition to using a specific frequency in the waveform to estimate the magnitude, it seems better to use the whole

observed bandwidth to predict an unobserved component. This idea is called spectral magnitude (Duda & Yanovskaya, 1993), and it may be useful to reduce the source error ε_i. Before looking for the optimal map from the source spectral to magnitude (namely, the field of functional analysis), the author thinks the work of this paper is essential and fundamental.

(6) One may notice that in Figure 3c, the colors of points above the line are mostly light cool or warm, whereas those below are heavily cool. However, in Figure 6c, the warm points are on both sides of the line. Therefore, it looks the magnitude relates to depth. Gutenberg and Richter (1956) thought that attenuation also relates to depth. As the source and path effects are coupled, we are not sure that the depth dependency is numerically caused by attenuation or source. In addition, the data set we used is too small to determine this dependency (Figure 10). JMA has used a linear term to correct the effect of focal depth during EEW (Aketagawa et al., 2010; Kiyomoto et al., 2010).

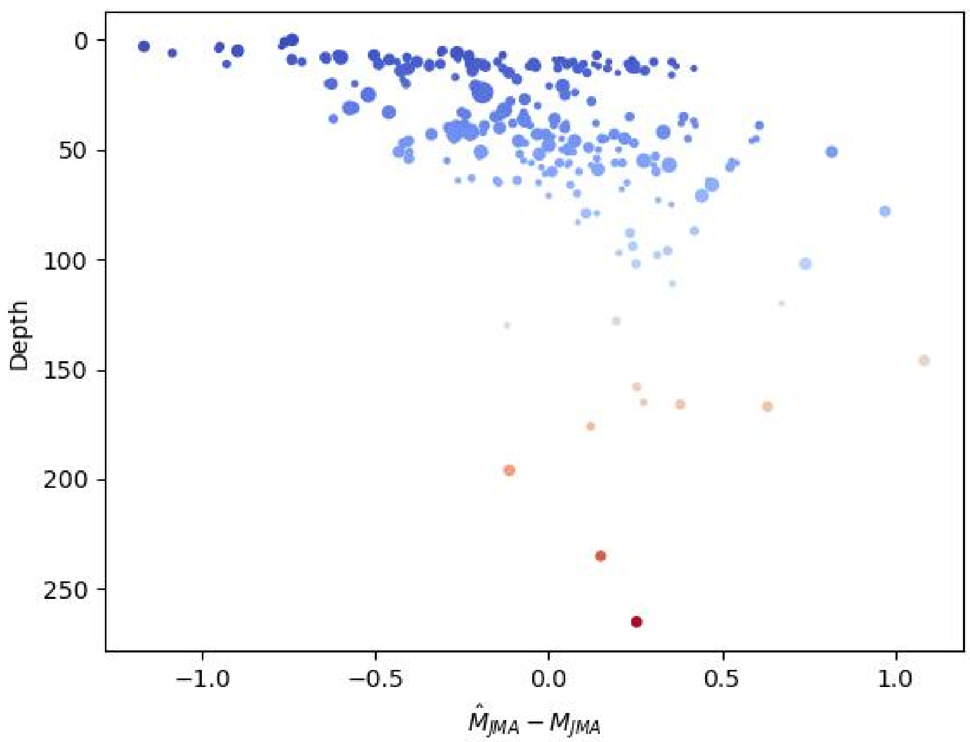

Figure 10. Single-earthquake magnitude estimation error versus focal depth. The size of the circle is directly proportional to the magnitude, and a warmer color indicates a deeper depth.

6. Conclusion

In this paper, the author introduces a statistical technique named 2-FER. It is difficult and not enough to analyze the mathematical property of this technique in theory. For example, its coefficient matrix is not full rank; therefore, its condition number cannot tell us if the relative value of the solution is robust or not. The author designs a series of check tests to carefully examine hypotheses of this mathematical approach and conclude that the relative value of the source, path and site effects is reliable. Thus, 2-FER is a proper tool to analyze the errors in magnitude estimation. After checking the deviation additivity and stationarity of the error process with distance r_{ij}, the author asserts that the endogeneity of the original issue is finely addressed, the energy leaks slightly and the different errors in the estimated magnitude are decoupled correctly. In addition, the author compares them quantitatively and explains in detail which physical factors impact each of the three errors.

The focal mechanism affects both the source and path effects. The dip affects the source effect, and the trend affects the path effect. The author has noted that the path error constituents are complicated by Equation (12). The author is not sure if it is mainly from the focal mechanism or radiation pattern and if its accuracy is sufficient to determine the radiation pattern in real time.

The simple point source assumption is not a suitable assumption to address near-field ground motion prediction, which is an important issue in EEW. The station density of the EEW Network is sufficient to determine a complex source in real time (Böse et al., 2012, 2018). A large earthquake usually consists of a few main ruptures and thus can be regarded as a few moderate earthquakes. If the waves from different moderate sources can be separated, then we can replace a large earthquake with a few points. Then, the analysis of large earthquakes during EEW will be largely simplified and closer to the physical matter. We have proven that 2-FER is a powerful tool to handle point sources. However, it is nearly impossible to separate waves from different orientations in only one channel, but it may be possible for three-component seismic data (Lei, 2005). AI methods may also help in this area.

The author makes an effort to analyze the magnitude error as he hopes to determine the rupture as quickly as possible from minor differences during EEW in the future. The regression of Equation (9) is not necessarily linear. Thus, it does not require us to measure the maximum amplitude, and we may research the time evolution of PGD more carefully with 2-FER. As this evolution is nonlinear and error-sensitive, mainstream geophysicists disagree with rupture determinism in a probabilistic view (Trugman et al., 2019; Meier et al., 2021). Thus, methods with a higher precision are quite crucial no matter whether to support it or against it in a physical way.

Although the author is introducing 2-FER for EEW purposes, the analysis method in this paper is not limited. To decouple the errors successfully, the key is a proper choice of attenuation function when applying it to surface waves or P-waves in other regions. The error process stationarity can help us to determine if the attenuation is appropriate. However, the author thinks that current tests of the application of 2-FER in seismology are not sufficient. Before being adopted, careful and further inspection is still essential.

Appendix A: Preconditions and Objectives of Different Regressions

The statistical model of 0-FER is:

$$\mathrm{lg}A_{ij} = \alpha S_i - a\mathrm{lg}r_{ij} - br_{ij} + \beta + (\varepsilon_{ij} + l_j + \varepsilon_i) \tag{AM0}$$

The objective of Regressing Model (AM0) is to find the minimum of:

$$\min_{a,b} \sum_{i,j} (\mathrm{lg}A_{ij} - \alpha S_i + a\mathrm{lg}r_{ij} + br_{ij} - \beta)^2 \tag{AO0}$$

Additionally, to achieve reliable $a\mathrm{lg}r_{ij} + br_{ij}$ by the least minimum square method (LSM), we should constrain:

$$\sum_{i,j}(\varepsilon_{ij}+l_j+s_i)\to 0 \tag{AC0}$$

The statistical model of 1-FER is:

$$\lg A_{ij}=\sum_{k=1}^{N}s_i\delta_{ik}-a\lg r_{ij}-br_{ij}+(\varepsilon_{ij}+l_j) \tag{AM1}$$

The objective of Regressing Model (AM1) is to find the minimum of:

$$\min_{\mathbf{s},a,b}\sum_{i,j}(\lg A_{ij}-\sum_{k=1}^{N}s_i\delta_{ik}+a\lg r_{ij}+br_{ij})^2 \tag{AO1}$$

where $\mathrm{s}=(s_1,s_2...s_N)$ and δ_{*k} is the Kronecker function, which is also referred to as a dummy variable in statistics.

Additionally, to achieve reliable s_i by LSM, we should constrain:

$$\sum_{j}(\varepsilon_{ij}+l_j)\to 0,\text{for each } i \tag{AC1}$$

The statistical model of 2-FER is:

$$\lg A_{ij}=\sum_{k=1}^{N}s_i\delta_{ik}+\sum_{k=1}^{M}l_j\delta_{jk}-a\lg r_{ij}-br_{ij}+(\varepsilon_{ij}) \tag{AM2}$$

The objective of Regressing Model (AM2) is to find the minimum of:

$$\min_{\mathbf{s},\mathbf{l},a,b}\sum_{i,j}(\lg A_{ij}-\sum_{k=1}^{N}s_i\delta_{ik}-\sum_{k=1}^{M}l_j\delta_{jk}+a\lg r_{ij}+br_{ij})^2 \tag{AO2}$$

where $\mathbf{l}=(l_1,l_2...l_M)$.

Additionally, to achieve reliable s_i and r_j by LSM, we should constrain:

$$\begin{cases}\sum_{i}\varepsilon_{ij}\to 0,\text{for each } j\\ \sum_{j}\varepsilon_{ij}\to 0,\text{for each } i\end{cases} \tag{AC2}$$

The objective of ridge regression is to find the minimum of:

$$\min_{\mathbf{s},\mathbf{l},a,b}\sum_{i,j}(\lg A_{ij}-\sum_{k=1}^{N}s_i\delta_{ik}-\sum_{k=1}^{M}l_j\delta_{jk}+a\lg r_{ij}+br_{ij})^2+\lambda(\mathbf{s}^2+\mathbf{l}^2+a^2+b^2)$$

Therefore, the ridge regression will find an estimator that not only produces a result with less error but also smaller coefficients. In this paper, the author has not rejected the data that may harm condition (AC2) (namely, has not removed stations covered less than 5 times, as mentioned in Figure 2c). The author chooses ridge regression to make the result look not that strange, as there will be some outliers for unreliable r_j. This operation covers up the bad results but please note this is not fabrication or falsification and the author just does not want the few abnormally large values mask the true trends in the following figure.

The logical relationships between the three preconditions are as follows:

$$\left.\begin{array}{r}\left.\begin{array}{r}AC2\\ \sum_j l_j=0\end{array}\right\}\Rightarrow AC1\\ \sum_i s_i=0\end{array}\right\}\Rightarrow AC0 \tag{AC3}$$

Please note that as we concern only the relative value in this paper, $l_j = 0$ here means l_j is a constant that is uncorrelated to j, as is s_i.

In the 0-FER model, if we denote

$$\alpha\widehat{S_{ij}} = \lg A_{ij} + a\lg r_{ij} + br_{ij} - \beta$$

where $\widehat{S_{ij}}$ is the estimated size of earthquake i at station j. Then, the objective function can be rewritten as:

$$\min_{a,b}\sum_{i,j}\alpha(\widehat{S_{ij}} - S_i)^2$$

This objective prefers a small α. If one wants an optimal magnitude estimation, i.e., the desired objective is

$$\min_{a',b'}\sum_{i,j}(\widehat{S'_{ij}} - S_i)^2 \tag{AC0'}$$

the Model (AM0) should be rewritten as

$$S_i = \alpha'\lg A_{ij} + a'\lg r_{ij} + b'r_{ij} - \beta' - (\varepsilon'_{ij} + l'_j + \varepsilon'_i) \tag{AM0'}$$

The Model (AM0') minimizes the error of the single-station magnitude instead of the amplitude, while Model (AM1) minimizes the error of the single-earthquake or multistation-average magnitude. The hyperplane defined by (AM0') performs even slightly better than (AM2), although the author did not plot it in Figure 5. The objective (AC0') requires (AM0') to be the best estimation of single-station magnitude mathematically, but the author does not think it is physical, as it overestimates the magnitude of far stations even more seriously than (AM0). On the other hand, the model (AM0') performs slightly worse than (AM0) in the view of single-eathquake magnitude error in the data set of this paper, but the coefficients differ greatly, which means that the ordinary linear regressions are not robust.

JMA explained that it is their preference to fix the coefficient before $\lg A_{ij}$ at one (Aketagawa et al., 2010). Wu and Zhao (2006) also adopt the model form of (AM0). However, this form makes the error term appear small as it multiplies the real error term $\varepsilon_i^{'}$ by $\alpha < 1$, especially for a short time window such as 3 seconds and then the α is quite small (Wu & Zhao, 2006). JMA found that the regression model calculated by stations in 200 km has a better RMSE in 500 km than the model in 500 km. This result is comprehensible, as the α in the model of 200 km is larger. JMA thought the model with the larger error is better, as it is from a larger data volume (Aketagawa et al., 2010).

To remain consistent with the former works, the author also fixes the coefficient before $\lg A_{ij}$ at one. However, please note that the author does not think it is a definitely good tradition and it brings an overoptimistic looking result. To reduce its influence, the author did not regress on Equation (10) directly in the 1-FER and 2-FER models. Instead, in this paper, the author regressed on the following model:

$$S_i = \alpha' M_i + \beta' + \varepsilon_i^{'}$$

Acknowledgments

The regressions in this paper were completed with statsmodels and sklearn, the seismic data were preprocessed with obspy, and the drawings in this paper were plotted with matplotlib and seasborn. The author would like to thank the developers of these open-source third-party packages of Python.

Data Availability Statement

The intermediate data and programs used for processing and plotting are available in Mendeley Data, V1 (https://data.mendeley.com/datasets/y4sjgg7vxs/1,doi: 10.17632/y4sjgg7vxs.1). The original waveform data can be downloaded from the official site of K-& KiK-net (https://www.kyoshin.bosai.go.jp).

References

Allen, R. M., & Kanamori, H. (2003). The potential for earthquake early warning in southern California. Science, 300(5620), 786-789. https://doi.org/10.1126/science.1080912

Allison, P. D. (2009). Fixed Effects Regression Models. America: SAGE Publications.

Atefi, S., Heidari, R., Mirzaei, N., & Siahkoohi, H. R. (2017). Rapid Estimation of Earthquake Magnitude by a New Wavelet-Based Proxy. Seismological Research Letters, 88(6), 1527-1533. https://doi.org/10.1785/0220170146

Aketagawa Tamotsu, et al (2010). Improvement of P-wave magnitude estimation for earthquake early warnings from JMA. Quarterly Journal of Seismology 73. (In Japanese)

Andrews, D. J. (1986). Objective determination of source parameters and similarity of earthquakes of different size. Earthquake source mechanics, 37, 259-267.

Böse M, Smith D E, Felizardo C, Meier M-A, Heaton T H, Clinton J F (2018), FinDer v.2: Improved real-time ground-motion predictions for M2–M9 with seismic finite-source characterization, Geophysical Journal International, Volume 212, Issue 1, January 2018, Pages 725–742, https://doi.org/10.1093/gji/ggx430

Böse M, Thomas H. Heaton, Egill Hauksson (2012), Real-time Finite Fault Rupture Detector (FinDer) for large earthquakes, Geophysical Journal International, Volume 191, Issue 2, November 2012, Pages 803–812, https://doi.org/10.1111/j.1365-246X.2012.05657.x

Brillinger, D. R., & Preisler, H. K. (1984). An exploratory analysis of the Joyner-Boore attenuation data. Bulletin of the Seismological Society of America, 74(4), 1441-1450.

Bormann, P., Wendt, S., DiGiacomo, D. (2013): Seismic Sources and Source Parameters. - In: Bormann, P. (Ed.), New Manual of Seismological Observatory Practice 2 (NMSOP2), Potsdam : Deutsches GeoForschungs Zentrum GFZ, 1-259. https://doi.org/10.2312/GFZ.NMSOP-2_ch3

Colombelli, Simona; Festa, Gaetano; Zollo, Aldo (2020). Early rupture signals predict the final earthquake size. Geophysical Journal International, (), ggaa343–. doi:10.1093/gji/ggaa343

Colombelli, S., Zollo, A., Festa, G., & Picozzi, M. (2014). Evidence for a difference in rupture initiation between small and large earthquakes. Nature Communications, 5, 3958. https://doi.org/10.1038/ncomms4958

Das, Ranjit; Wason, H. R.; Gonzalez, Gabriel; Sharma, M. L.; Choudhury, Deepankar; Lindholm, Conrad; Roy, Narayan; Salazar, Pablo (2018). Earthquake Magnitude Conversion Problem. Bulletin of the Seismological Society of America, 108(4), 1995–2007. doi:10.1785/0120170157

Diggle, Peter J.; Heagerty, Patrick; Liang, Kung-Yee; Zeger, Scott L. (2002). Analysis of Longitudinal Data (2nd ed.). Oxford University Press. pp. 169–171. ISBN 0-19-852484-6.

Duda, S. J., and Yanovskaya, T. B. (1993). Spectral amplitude-distance curves for P-waves: effects of velocity and Q-distribution. Tectonophysics, 217, 255-265.

Furumura, T., & Oishi, Y. (2023). An early forecast of long-period ground motions of large earthquakes based on deep learning. Geophysical Research Letters, 50, e2022GL101774. https://doi.org/10.1029/2022GL

Gardiner, Joseph C.; Luo, Zhehui; Roman, Lee Anne (2009). "Fixed effects, random effects and GEE: What are the differences?". Statistics in Medicine. 28 (2): 221–239. doi:10.1002/sim.3478. PMID 19012297. S2CID 16277040.

Greene, W.H., 2011. Econometric Analysis, 7th ed., Prentice Hall

Gutenberg B, Richter CF (1956) Magnitude and energy of earthquakes. Ann Geofis 9:1–15

Iwata, Tomotaka; Irikura, Kojiro (1986). Separation of Source, Propagation and Site Effects from Observed S-Waves. Zisin (Journal of the Seismological Society of Japan. 2nd ser.), 39(4), 579–593. doi:10.4294/zisin1948.39.4_579 (in Japanese)

Lei Jun (2005). A method for non-orthogonal seismic polarization-vector separation. , 162(3), 965–974. doi:10.1111/j.1365-246x.2005.02709.x

Joyner, W. B. and D. M. Boore (1981)，Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, Bull. Seism. Soc. Am. 71, 2011-2.

Kiyomoto Masashi, et al (2010). Investigation of Technical Issues for Earthquake Early Warning. Quarterly Journal of Seismology 73. (In Japanese)

Meier, M. A., Ampuero, J. P., Cochran, E., & Page, M. (2021). Apparent earthquake rupture predictability. Geophysical Journal International, 225(1), 657-663.

Melgar, D., B. W. Crowell, J. Geng, R. M. Allen, Y. Bock, S. Riquelme, E. M. Hill, M. Protti, and A. Ganas (2015), Earthquake magnitude calculation without saturation from the scaling of peak ground displacement, Geophys. Res. Lett., 42, 5197–5205, doi:10.1002/2015GL064278.

Moya, A. (2000). Inversion of Source Parameters and Site Effects from Strong Ground Motion Records using Genetic Algorithms. Bulletin of the Seismological Society of America, 90(4), 977–992. doi:10.1785/0119990007

Motaghi, K., & Ghods, A. (2012). Attenuation of Ground-Motion Spectral Amplitudes and Its Variations across the Central Alborz Mountains. Bulletin of the Seismological Society of America, 102(4), 1417–1428.

Münchmeyer, J., Leser, U., & Tilmann, F. (2022). A probabilistic view on rupture predictability: All earthquakes evolve similarly. Geophysical Research Letters, 49, e2022GL098344. https://doi.org/10.1029/2022GL098344

National Research Institute for Earth Science and Disaster Resilience (2019), NIED K-NET, KiK-net, doi:10.17598/NIED.0004

Nuttli, O. W.(1973), Seismic wave attenuation and magnitude relations for eastern North America, J. Geophys. Res., 78, 876-885,.

Papageorgiou, A. S., & Kim, J. (1991). Study of the propagation and amplification of seismic waves in Caracas Valley with reference to the 29 July 1967 earthquake: SH waves. Bulletin of the Seismological Society of America, 81(6), 2214-2233.

Phillips, W. Scott, Keiiti Aki; Site amplification of coda waves from local earthquakes in central California. Bulletin of the Seismological Society of America 1986;; 76 (3): 627–648. doi: https://doi.org/10.1785/BSSA0760030627

Parolai, S. (2004). Comparison of Different Site Response Estimation Techniques Using Aftershocks of the 1999 Izmit Earthquake. Bulletin of the Seismological Society of America, 94(3), 1096–1108. doi:10.1785/0120030086

Richter, C. F. (1935). An instrumental earthquake magnitude scale, Bull. Seism. Soc. Am. 25, 1-32.

Seggern , David von (1970); The effects of radiation patterns on magnitude estimates. Bulletin of the Seismological Society of America; 60 (2): 503–516. doi: https://doi.org/10.1785/BSSA0600020503

Song, Jindong; Zhu, Jingbao; Li, Shanyou. MEANet: Magnitude Estimation Via Physics-based Features Time Series, an Attention Mechanism, and Neural Networks. Geophysics. 2023. 88(1): V33-V43.

Takemura, Masayuki; Kato, Kenichi; Ikeura, Tomonori; Shima, Etsuzo (1991). Site Amplification of S-Waves from Strong Motion Records in Special Relation to Surface Geology. Journal of Physics of the Earth, 39(3),
537–552. doi:10.4294/jpe1952.39.537

Taylor, John R. (2022), An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements (Third Edition), USA: University Science Books

Torres-Sánchez, E. M., & Castro, R. R. (2023). P-and S-wave attenuation in the northern region of the Gulf of California, Mexico. Journal of Seismology, 1-17.

Trugman, D. T., Page, M. T., Minson, S. E., & Cochran, E. S. (2019). Peak ground displacement saturates exactly when expected: Implications for earthquake early warning. Journal of Geophysical Research: Solid Earth, 124, 4642–4653. https://doi.org/10.1029/2018JB017093

United States Geology Survey (2023a): Magnitude Types, https://www.usgs.gov/programs/earthquake-hazards/magnitude-types

United States Geology Survey (2023b): Origin of M 4.9 - 115 km SE of Honiara, Solomon Islands, https://earthquake.usgs.gov/earthquakes/eventpage/us7000kwk7/origin/detail

Wang, W., & Shearer, P. M. (2019).An improved method to determine coda-Q, earthquake magnitude, and site amplification: Theory and application to southern California. Journal of Geophysical Research: Solid Earth, 124, 578–598. https://doi.org/10.1029/2018JB015961

Wang, Y., Li, X., Li, L., Wang, Z., & Lan, J. (2022). New magnitude proxy for earthquake early warning based on initial time series and frequency. Seismological Research Letters, 93(1), 216-225.

Wu, Y.M. & Zhao, L., 2006. Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning, Geophys. Res.Lett., 33, L16312, doi:10.1029/2006GL026871

Zhang, G., D. Li, Y. Gao, and J. Yang (2022). A Magnitude Estimation Approach and Application to the Yunnan Earthquake Early Warning Network, Seismol. Res. Lett. 94, 234–242, doi: 10.1785/0220220093.

Zhang, G. (2023). Programs of Two-way Fixed Effects Regression on the Error Analysis of PGA Magnitude (Version 1) [Program]. Mendeley Data. https://doi.org/10.17632/y4sjgg7vxs.1

Zhu, Jingbao; Li, Shanyou; Song Jindong. Hybrid Deep-Learning Network for Rapid On-Site Peak Ground Velocity Prediction. 2022. IEEE Transactions on Geoscience and Remote Sensing. vol. 60, pp. 1-12, 2022, Art no. 5925712