Bulletin of Earthquake Engineering (2023) 21:5375–5401
https://doi.org/10.1007/s10518-023-01749-0

ORIGINAL ARTICLE

New ground-motion prediction equations for significant duration of subduction intraslab and interface earthquakes in Japan

Yinan Zhao[1,2] · Maosheng Gong[1,2] · Jingyang Tan[1,2] · Zhanxuan Zuo[1,2] · Jia Jia[1,2]

Received: 22 August 2022 / Accepted: 24 July 2023 / Published online: 1 August 2023

Abstract
To date, multiple ground-motion prediction equations (GMPEs) for the significant duration (D_S) of shallow crustal earthquakes have been proposed, but there are few GMPEs for subduction intraslab and interface earthquakes in Japan, and corresponding spatial correlation models have not been published. To address this issue, we first select ground-motion records with moment magnitude $4 \leq M_w \leq 9$, rupture distance $R_{rup} \leq 300$ km, and peak ground acceleration ≥ 10 gal based on the K-NET and KiK-net databases. Then, for intraslab and interface earthquakes in the subduction zone, based on previous works, the traditional source duration term is simplified, a depth term is added, and new GMPEs for D_S of earthquakes in the subduction zone in Japan are developed. The rationality and reliability of the prediction model proposed in this study are verified by residual analysis and comparison with the previous models. Finally, according to the intra-event residuals, a spatial correlation model of D_S is established by using a semivariogram and exponential model. The results show that D_S of subduction interface earthquakes is larger than that of subduction intraslab earthquakes on the whole, but the spatial correlation coefficient of subduction interface earthquakes decreases more slowly with increasing separation distance than that of subduction intraslab earthquakes, and the significant duration of the spatial correlation is related to site effects and path effects. The results provide a reference for the development of GMPEs based on nonergodic assumptions, regional seismic hazard analysis and loss assessment.

Keywords Prediction equations · Significant duration · Subduction intraslab · Subduction interface · Spatial correlation

✉ Maosheng Gong
gmshiem@163.com

[1] Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, No. 29, Xuefu Road, Harbin 150080, Heilongjiang, China

[2] Key Laboratory of Earthquake Disaster Mitigation, Ministry of Emergency Management, No. 29, Xuefu Road, Harbin 150080, Heilongjiang, China

1 Introduction

Amplitude, frequency content and duration jointly describe the basic characteristics of ground motion, but research on ground-motion duration started relatively late compared to that on the amplitude and frequency content. In most engineering seismic analyses, the influence of the amplitude and frequency content is considered more than the duration because when the engineering structure is in the elastic response stage, the duration has little effect on its seismic response. However, when the engineering structure is in the inelastic response stage, even if the maximum deformation of the structure does not exceed the limit deformation, the structure may collapse due to the loss of energy storage capacity reaching a certain limit. The effect of duration on the seismic response of the structure cannot be ignored. Therefore, to fully describe the effect of ground motion on structural seismic response, the duration should be considered.

The influence of the duration effect on structural response is still controversial (Du and Wang 2017); most studies have reported that the duration has a negligible effect on the peak structural responses (e.g., Rahnama and Manuel 1996; Cosenza et al. 2004; Iervolino et al. 2006), but it is generally accepted that there is a positive correlation between the duration and cumulative damage measures (e.g., Meskouris 1983; Uang and Bertero 1990; Reinoso et al. 2000; Bommer et al. 2004), as described by Hancock and Bommer (2006). In recent years, with the development of performance-based seismic design theory, researchers have paid increasing attention to the influence of ground-motion duration on the seismic capacity of structures. Chandramohan et al. (2016) examined the influence of duration on the collapse capacities of a five-story steel moment frame and a reinforced concrete bridge pier; the results showed that the collapse capacity of the steel moment frame is 29% lower, and for bridge pier, it is 17% lower, when using long duration records. Bravo-Haro and Elghazouli (2018) examined the influence of the ground-motion duration on the seismic response of steel moment frames with due consideration for cyclic degradation effects; the results indicated collapse capacity reductions up to 40% due to the duration influence. Bravo-Haro et al. (2020) presented a detailed investigation into the seismic response of nondeteriorating and deteriorating single degree-of-freedom systems controlled by P-Δ effects, with due account for the duration influence; they found that the duration effect is significant, increasing with an increasing structural period and decreasing with an increasing P-Δ effect. Liapopoulou et al. (2020) explored the effects of the ground-motion duration and acceleration pulse on the collapse capacity of ductile single-degree-of-freedom systems and showed that the collapse ability of flexible bilinear systems can be reduced by 60% due to the duration effect at low level P-Δ levels. Du et al. (2020) quantitatively investigated the duration effect on structural collapse assessment by using hazard-consistent ground-motion suites; the results showed that a longer duration record is more likely to reduce the structural collapse capacity. These prior researches indicate that the influence of the ground-motion duration on the collapse capacities of structures cannot be ignored. Therefore, it is recommended to consider this duration in addition to its intensity and frequency content in structural design and seismic risk assessment.

There are more than 30 definitions of ground-motion duration, and the significant duration (D_S) is the most commonly used in the field of earthquake engineering and the one most studied by researchers (Chandramohan et al. 2016; Du and Wang 2017). Ground-motion prediction equations (GMPEs) are one of the critical components in probabilistic seismic hazard assessment (PSHA), and the GMPEs of the duration are mainly developed according to the attenuation relationship (Akkar et al. 2018; Ameri et al. 2017; Ebrahimian

et al. 2019; Zolfaghari and Darzi 2019). The GMPE of the duration was first proposed by Esteva and Rosenblueth (1964). In the following 20 years, some scholars also proposed their own GMPEs. The forms of these GMPEs are simple, and the influencing factors considered are relatively few (Housner 1975; Trifunac and Brady 1975; Dobry et al. 1978; McGurie and Barnhard 1979; Kamiyama 1984). Abrahamson and Silva (1996) first expressed the source duration (D_{source}) as a function of the moment magnitude (M_{w}) and stress drop ($\Delta\sigma$) based on the Brune source model, and the influence of rupture distance (R_{rup}) and site type (S) were added to their GMPEs of D_S (AS96). Hernandez and Cotton (2000) developed a GMPE for $D_{S5\text{-}95}$ based on ground-motion databases in California and Italy, but the coupling effect of magnitude and path terms was not included. Reinoso and Ordaz (2001) proposed a GMPE for $D_{S2.5\text{-}97.5}$ using ground-motion data of normal faults in the subduction zone of Mexico and combined the duration with random vibration theory to predict the response spectrum. On the basis of the work of Abrahamson and Silva, Kempton and Stewart (2006) established a GMPE (KS06) for $D_{S5\text{-}75}$ and $D_{S5\text{-}95}$ using the ground-motion data of global shallow crustal earthquakes near active plate margins. Different from the previous GMPEs, KS06 takes the time-averaged shear-wave velocity to 30 m (V_{S30}) as the site effect parameter and considers the influence of the depth to shear wave velocity of 1500 m/s for the site ($Z_{1.5}$) on D_S. Bommer et al. (2009) used the NGA-West1 database to put forward GMPEs for D_B (bracketed duration), D_U and D_S (BSA09). In addition to M_{w}, R_{rup} and V_{S30}, the influence of the fault type F_{rv} was added to the GMPEs of D_B and D_U, and the influence of the depth to the top of rupture (Z_{tor}) was incorporated into the GMPE. BSA09 considered the coupling of magnitude and path, but the basin depth was not included. Yaghmaei-Sabegh et al. (2014) proposed a GMPE of D_S for Iran. Due to the lack of fault type, stress drop, depth of rupture surface and other data in Iran, their GMPE includes only the magnitude, hypocentral distance and site type. Lee and Green (2014) proposed a GMPE of D_S (LG14) in stable continental regions and compared and analyzed its difference with active shallow crustal region motions. Afshari and Stewart (2016) used the NGA-West2 database to develop GMPEs (AS16) for $D_{S20\text{-}80}$, $D_{S5\text{-}75}$, and $D_{S5\text{-}95}$, and AS16 was found to be an improvement over KS06. Du and Wang (2017) also developed a new GMPE (DW17) based on the NGA-West2 database, which improved the model prediction for small- and medium-magnitude earthquakes and far-field earthquakes. Anbazhagan et al. (2017) used intraslab regions from Canada, Australia, the Indian Peninsula and the central and southern United States to present GMPEs for D_B and D_S, and Xu and Wen (2018) established a GMPE of D_S for mainland China based on strong earthquake data in China. López-Castañeda and Reinoso (2021) and Jaimes and García-Soto (2021) successively proposed GMPEs (JG21) for D_S of ground motion for interface and intraslab earthquakes in the Mexican subduction zone. The results showed that D_S of subduction interface and intraslab earthquakes is significantly different from those in other parts of Mexico and that these two types of earthquakes need to be considered separately when establishing prediction models. Bahrampouri et al. (2021) proposed GPMEs of D_S (BRG21) for intraslab, interface, and shallow crustal earthquakes based on the KiK-net database, which compensated for the lack of research on the GMPEs of duration in subduction zones, but the effect of depth was not incorporated in these equations. In addition, the magnitude term in the exponential form in the equation may not easily accommodate regression analysis under different data sets, and the residuals have prominent trends when $M_{\text{w}} < 5$ and $R_{\text{rup}} < 100$ km.

As the GMPEs of ground-motion parameters have been successively proposed, the spatial correlation of ground-motion parameters has gradually attracted attention. In terms of seismic hazard analysis, the traditional single-site analysis method cannot evaluate the probability of simultaneous earthquakes at multiple sites, and the spatial correlation model

of ground motions is beneficial to develop the seismic hazard analysis from a single site to multiple sites (Albarello and Peruzza 2017). Therefore, scholars have begun to analyze the spatial correlation characteristics of ground-motion parameters through the residuals of GMPEs. Boore et al. (2003) analyzed the spatial variability of the PGA using the main shock records of the 1994 Northridge earthquake in the United States. The results showed that the spatial variability in PGA decreases with increasing distance between the two sites. Goda and Hong (2008) studied the spatial correlation of the PGA, PGV and pseudo-spectral acceleration p-$S_a(T=0.3$ s, 1.0 s, 3.0 s) using the records of Southern California in the United States and Chi-Chi in Taiwan, China. The results showed that the spatial correlation decreases with increasing separation distance. The spatial correlation of ground motion is still one of the key research topics of scholars, and an increasing number of spatial correlation models have been proposed (Wang and Takada 2005; Hong et al. 2009; Jayaram and Baker 2009; Goda and Atkinson 2010; Du and Wang 2013; Bradley 2014; Foulser-Piggott and Goda 2015), but the spatial correlation curves of individual earthquake events are more variable. However, to date, research on the spatial correlation of ground-motion duration has been very rare, especially for the Japanese subduction zone.

This study analyzes the effect of M_w, R_{rup}, V_{S30}, δZ_1 (the difference between the depth to shear wave velocity of 1000 m/s for the site and a median depth conditional on V_{S30}) and other factors on D_S (the geometric mean of significant durations in two horizontal directions) by developing new GMPEs of Japanese subduction intraslab and subduction interface earthquakes based on K-NET and KiK-net databases. The magnitude term is simplified, the magnitude is calibrated by a bilinear function (Xing and Zhao 2021), and the depth (Z_{tor}) term is added in the new GMPEs. Through residual analysis and comparison with existing prediction models, the rationality and reliability of the prediction model proposed in this study are proven. Finally, according to the intra-event residuals of the new GMPEs, the spatial correlation of D_S is analyzed, the spatial correlation coefficient of D_S is calculated by using semivariograms, and the corresponding continuous spatial correlation function is obtained by fitting an exponential model. This work provides a reference for the development of next-generation GMPEs based on nonergodic assumptions, the spatially distributed lifeline engineering, and regional seismic hazard and loss assessment.

2 Strong ground-motion databases

The strong ground-motion data in this study come from the K-NET and KiK-net databases developed by the National Research Institute for Earth Science and Disaster Prevention (NIED) of Japan. The ground-motion records of nearly the past 25 years from January 1997 to July 2021 are included and selected according to the following standards: (a) select earthquake events with $4 \le M_w \le 9$ and $R_{rup} \le 300$ km; (b) identify PGA ≥ 10 gal (both horizontal components); (c) discard records with abnormal waveforms, including clearly aberrant waveforms, incomplete waveforms (mainly manifested as missing P waves), and multiple earthquake events included in one waveform; (d) discard records with incomplete information, including records with missing site information and focal mechanism solutions; and (e) the number of records contained in a single earthquake event is not less than 10.

Since the ground-motion records obtained from K-NET and KiK-net are unprocessed and there are different degrees of baseline wander and noise pollution, the records must be baseline corrected and filtered before used. The baseline correction method proposed

by Ohsaki (2008) is used. Before filtering, P-wave arrival picking is performed on each record to extract the pre-event noise, the Fourier amplitude spectrum of the ground-motion signal and the pre-event noise signal is calculated, and the frequency range $[f_{c1}, f_{c2}]$ of the signal-to-noise ratio (SNR) ≥ 3 is determined (Bozorgnia 2020). This study stipulates that the high-pass cutoff frequency f_{c1} should be between 0.05 and 0.5 Hz and that the low-pass cutoff frequency f_{c2} should not be lower than 25 Hz (Bahrampouri et al. 2020). For the two horizontal components of one record, f_{c1} takes the maximum high-pass cutoff frequency of the two components, and f_{c2} takes the minimum low-pass cutoff frequency of the two components. After determining the filtering frequency range, zero-padding is performed on each record, and a 4th-order Butterworth acausal filter is used for filtering (Boore 2005; Boore and Akkar 2003).

The latitude and longitude coordinates, depth, and M_w of the sources in this study are from the Japan Meteorological Agency (JMA) Unified Hypocenter Catalog and F-net Mechanism Research. The source information for the events in 2005 and later are taken from the JMA Unified Hypocenter Catalog with K precision, and other events are given in F-net mechanism research. The latitude and longitude coordinates of the stations are derived from the site list given by K-NET and KiK-net, and we use the method proposed by Boore et al. (2011) to calculate V_{S30}. The fault models provided by the United States Geological Survey (USGS), K-NET, SRCMOD database, and published papers (see electronic supplement) are used in this study, allowing R_{rup} and Z_{tor} to be calculated. For earthquake events without a fault model, an empirical formula is used to estimate R_{rup} and Z_{tor} (Scherbaum et al. 2004; Mai et al. 2005; Kaklamanos et al. 2011). The global subduction-zone geometry model Slab 2.0 provided by Hayes et al. (2018) is used to consider the location of the hypocenter relative to the depth of the surface of the subducting plate, and the earthquake events are classified according to the earthquake classification method of the Japanese subduction zone proposed by Zhao et al. (2015). It should be pointed out that the classification of some earthquake events is based on the research results of NGA-Subduction Project (Stewart 2020).

According to the abovementioned ground-motion records selection and data processing, 68,256 records are finally selected, including 27,086 subduction intraslab records and 10,209 subduction interface records. The distribution of M_w versus R_{rup} is shown in Fig. 1, and the distribution of the number of stations versus V_{S30} is shown in Fig. 2. An interesting phenomenon in Fig. 2b is that two large-magnitude earthquakes are limited to data with R_{rup} less than 250 km, shorter than smaller-magnitude earthquakes. There are two reasons to explain this phenomenon: on the one hand, the distribution of stations is not always uniform, which leads to a lack of records over some distance ranges; on the other hand, the records with large

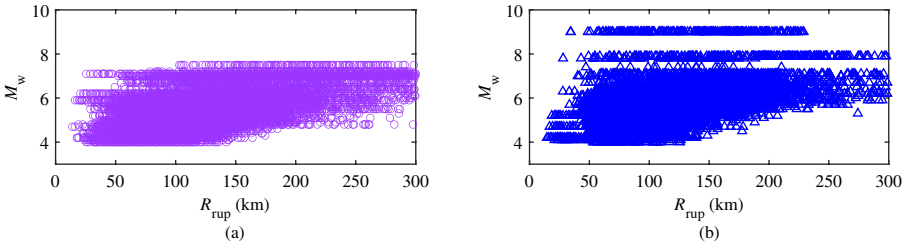

Fig. 1 Distribution of M_w versus R_{rup}: **a** subduction intraslab earthquakes and **b** subduction interface earthquakes

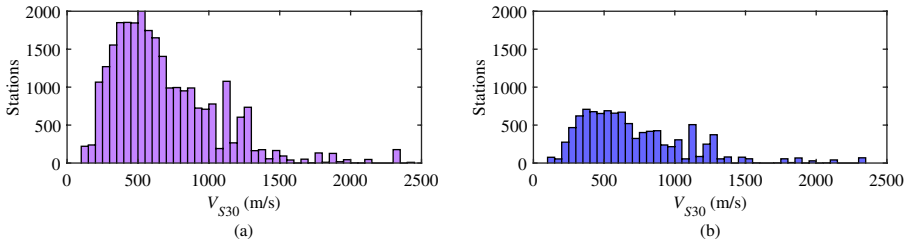

Fig. 2 Histogram of the number of stations versus V_{S30}: **a** subduction intraslab earthquakes and **b** subduction interface earthquakes

distance have a relatively low SNR and poor data quality. To ensure the high quality of the database in this study, these low-quality data have been removed.

Two components (E-W and N-S) of the accelerograms in the database and their Husid plots are shown in Fig. 3. Even though long pre-event and post-event noise (approximately 20–30 s) signals appear in the ground-motion acceleration records, the slope of the Husid value is small and has almost no effect on D_S.

3 Prediction models

The ground-motion duration D is usually written as the sum of the magnitude term F_M, the path term F_{path} and the site term F_{VS30} (Kempton and Stewart 2006), and the expression is as follows:

$$\ln D = F_M + F_{\text{path}} + F_{VS30} \tag{1}$$

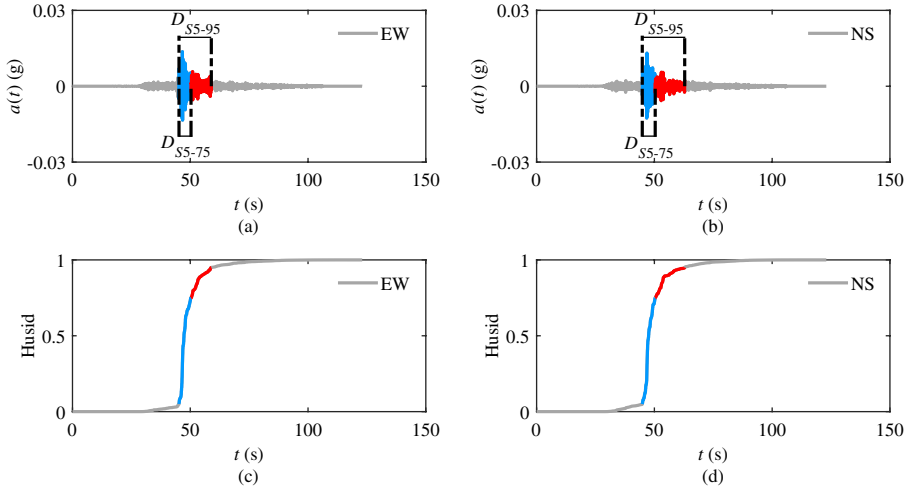

Fig. 3 Accelerograms (TCG0142107270519) in the database and their Husid plots: **a**, **b** accelerograms and **c**, **d** Husid plots

With the continuous improvement in GMPEs, an increasing number of influencing terms have been considered, including basin and focal depth. The influence of the basin effect on the ground motion is very significant. This is because when the seismic wave enters the basin from the soil, the difference in the impedance of the medium amplifies the ground motion, and at the same time, the body waves reaching the critical incident angle are completely reflected in the basin. Surface waves propagating inside the basin are generated. After complex superposition and interference of seismic waves, the ground motion is amplified, and the duration is also significantly increased. Depth is also an indispensable factor in the relationship with ground-motion attenuation. Zhao et al. (2016a, b, c) noted that the effect of depth on ground motion in the subduction zone is strong. There are many different types of earthquakes that can be generated over a wide range of depths, so changes in depth can have a significant effect on duration. Based on the K-NET and KiK-net databases, the influence of the magnitude, path, V_{S30}, basin and depth to the top of rupture on D_S of the ground motion are deeply explored, and a new GMPE of D_S is proposed in this study:

$$\ln D_S = F_{\text{source}} + F_{\text{path}} + F_{\text{site}} + \delta B + \delta W = F_M + F_{\text{path}} + F_{VS30} + F_{\text{basin}} + F_{\text{depth}} + \delta B + \delta W \tag{2}$$

where, the source term F_{source} includes the magnitude term F_M and the depth to the top of the rupture term F_{depth}, the site term F_{site} includes the V_{S30} term F_{VS30} and the basin term F_{basin}, δB is the inter-event residual, and δW is the intra-event residual. The coefficients in the GMPE are regressed using random effects regression analysis (Abrahamson and Youngs 1992), and the *t*-distribution statistics of each coefficient and the AIC value of the GMPE are calculated (Akaike 1974). When a new explanatory variable is added to the GMPE, only when $|t_{\alpha/2}| \geq 1.96$ (significance level $\alpha = 0.05$) and the AIC value decreases upon adding the new explanatory variable can the new explanatory variable be significant, and the goodness of fit of the GMPE is higher; otherwise, the new explanatory variable needs to be removed.

3.1 Path term

By considering the coupling relationship between magnitude and distance, the path term cannot be obtained directly from the distribution of D_S and R_{rup}. Therefore, the distribution of D_S on R_{rup} is grouped according to different M_{w} ranges, and the path term under different M_{w} ranges is given. Figure 4 shows the R_{rup} distribution of $D_{S5\text{-}95}$ under each M_{w} range. The R_{rup} distribution of $D_{S5\text{-}75}$ is similar to that of $D_{S5\text{-}95}$ and is given in the electronic supplement due to space limitations (the results for $D_{S5\text{-}75}$ in later sections are also included in the electronic supplement). By fitting the data with a piecewise linear function, it can be found that when $M_{\text{w}} < M_4^*$, there is a significant difference in the slope of the piecewise linear function on both side of the reference distance R^*, and the path term is as follows:

$$F_{\text{path}} = \begin{cases} C_1 \ln R_{\text{rup}}, & R_{\text{rup}} \leq R^* \\ C_1 \left[\ln R^* + C_2 \left(\ln R_{\text{rup}} - \ln R^*\right)\right], & R_{\text{rup}} > R^* \end{cases} \tag{3}$$

$$C_2 = \begin{cases} 0, & M_{\text{w}} < M_3^* \\ \frac{M_{\text{w}} - M_3^*}{M_4^* - M_3^*}, & M_3^* \leq M_{\text{w}} < M_4^* \\ 1, & M_{\text{w}} \geq M_4^* \end{cases} \tag{4}$$

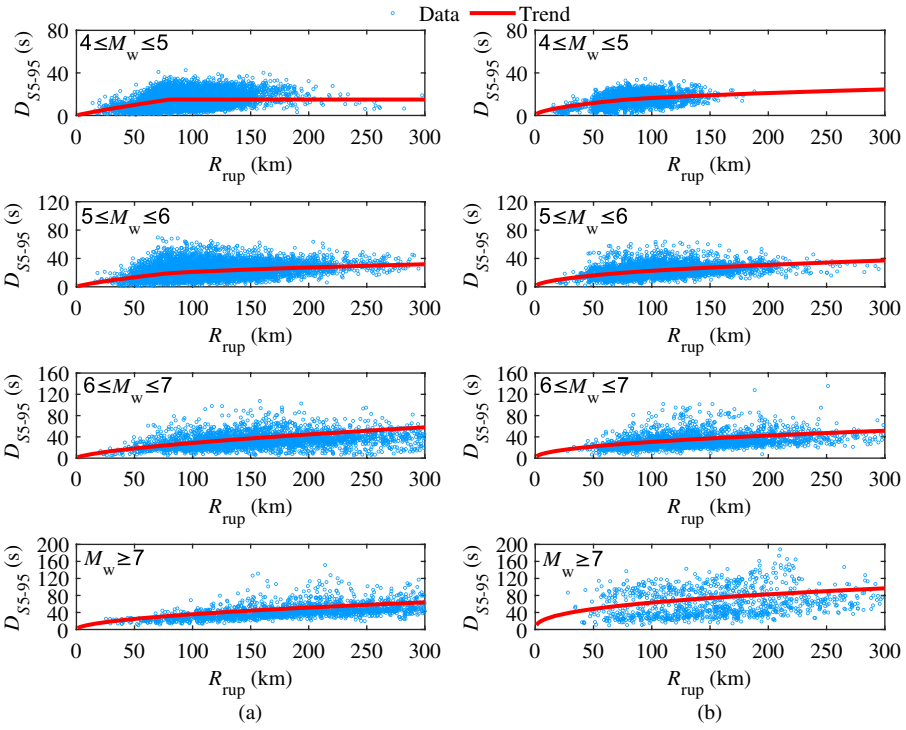

Fig. 4 Distribution of $D_{S5\text{-}95}$ versus R_{rup} under different M_w: **a** subduction intraslab earthquakes; **b** subduction interface earthquakes

where C_1 and C_2 are model parameters, M_3^* and M_4^* are reference magnitudes, and C_2 is determined by the relationship between M_w, M_3^* and M_4^*.

The red lines in Fig. 4 are the regression lines after determining the source and path terms and do not represent the final results. The same is true for the red solid line in Figs. 5, 6, 8 and 10. The specific form of the GMPEs is subject to the final result obtained by regression analysis.

3.2 Magnitude term

Abrahamson and Silva (1996), Kempton and Stewart (2006), and Afshari and Stewart (2016) used the Brune source model to express the source duration (i.e., magnitude term) as a function of the stress drop $\Delta\sigma$ as follows:

$$F_M = \frac{\left(\Delta\sigma/M_0\right)^{-\frac{1}{3}}}{4.9\cdot 10^6\beta} = \frac{1}{4.9\cdot 10^6\beta}\left\{\frac{\exp\left[b_1+b_2\left(M_w-M^*\right)\right]}{10^{1.5M_w+16.05}}\right\}^{-\frac{1}{3}} \tag{5}$$

where M_0 is the seismic moment, β is the shear wave velocity at the source (usually taken as 3.2 km/s), b_1 and b_2 are regression coefficients, and M^* is the reference magnitude. Although there is theoretical support for expressing the source effect of duration in this form, the stress drop is difficult to obtain accurately in practice. Afshari and Stewart

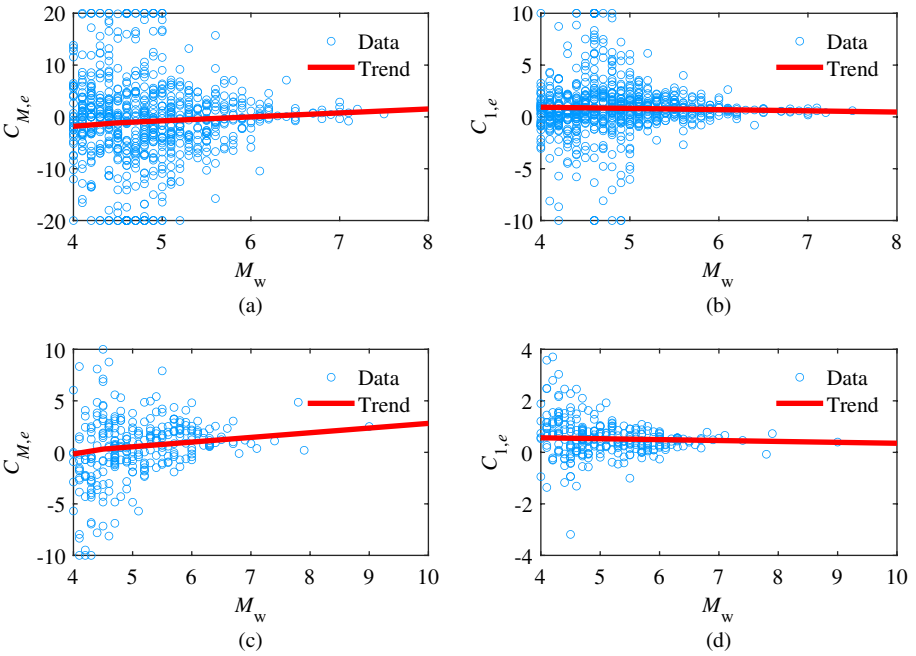

Fig. 5 Distribution of $C_{M,e}$ and $C_{1,e}$ versus M_w: **a**, **b** subduction intraslab earthquakes; **c**, **d** subduction interface earthquakes

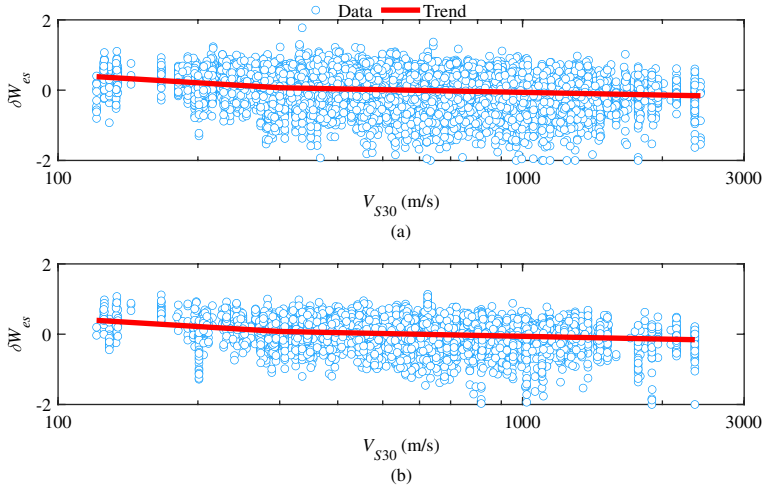

Fig. 6 Distribution of δW_{es} versus V_{S30}: **a** subduction intraslab earthquakes; **b** subduction interface earthquakes

(2016) reported that the duration and distance of each earthquake can be linearly fitted to obtain a linear function relationship, and then the duration when the distance is 0 is taken as the source duration. However, we have concluded that the source duration and stress drop obtained in this way are inaccurate, as after calculating, it is found that the

source duration and stress drop of some earthquake events are negative (Afshari and Stewart pointed this out in their paper as well). The reason is that the relationship between duration and distance is not strictly linear within the whole range of R_{rup} but approximate linear or piecewise approximate linear; thus, using a linear function to fit the source duration leads to large errors. Moreover, the magnitude term given by Eq. (5) is relatively complex. Although Bahrampouri et al. (2021) adopted a variant of Eq. (5), the magnitude term is still in exponential form, which may not converge in regression analysis, and Bommer et al. (2009) concluded that the magnitude term should be equivalent to a simpler form to facilitate regression. Therefore, the present study adopts the variant of the magnitude term given by Bommer et al. (2009) and uses the method of Xing and Zhao (2021) to calibrate the magnitude. First, we assume that the GMPE is of the following form:

$$\ln D_{S,es} = C_{M,e} + F_{\mathrm{path}} \tag{6}$$

where the subscripts e and s represent the number of earthquake event and the station, respectively. The site and depth are not considered for now, and the magnitude term is assumed to be a constant C_M; that is, the magnitude term of the eth earthquake event is $C_{M,e}$, and then the magnitude term is calibrated. In addition, to study whether there is a coupling phenomenon between the magnitude and the distance, the path term coefficient $C_{1,e}$ of the eth earthquake event will also be calibrated. Figure 5 shows the distribution of $C_{M,e}$ and $C_{1,e}$ versus M_{w} for $D_{S5\text{-}95}$, $C_{M,e}$ is piecewise linear versus M_{w} and $C_{1,e}$ is also related to M_{w}. $C_{M,e}$ shows an upward trend with increasing M_{w}, while $C_{1,e}$ shows a decreasing trend. Therefore, the coefficients of the magnitude and path terms can be written as:

$$F_M = \begin{cases} a_1 + a_2 M_{\mathrm{w}}, & 4 \le M_{\mathrm{w}} < M_1^* \\ a_1 + a_2 M_1^* + a_3\left(M_{\mathrm{w}} - M_1^*\right), & M_{\mathrm{w}} \ge M_1^* \end{cases} \tag{7}$$

$$C_1 = \begin{cases} a_4 + a_5 M_{\mathrm{w}}, & M_{\mathrm{w}} \le M_2^* \\ a_4 + a_5 M_2^* + a_6\left(M_{\mathrm{w}} - M_2^*\right), & M_{\mathrm{w}} > M_2^* \end{cases} \tag{8}$$

where C_1 is a model parameter in Eq. (3), a_1 to a_6 are regression coefficients, and $M_1{}^*$ and $M_2{}^*$ are reference magnitudes obtained from Fig. 5a, c and b, d by piecewise linear fitting, respectively.

3.3 V_{S30} term

The early GMPEs expressed the site effect as the site type, but in recent years, they usually incorporate V_{S30} to represent the site effect. V_{S30} is a smooth and continuous parameter that can provide a smooth transition from a rock site to a soft soil site when describing site effects (Kamai et al. 2016). Kempton and Stewart (2006), Afshari and Stewart (2016), and Bahrampouri et al. (2021) chose V_{S30} as the site effect parameter, so the present paper also uses V_{S30} to describe the impact of the site effect (i.e., V_{S30} term) on D_S; the basin term is presented in the next section.

Figure 6 shows the distribution of V_{S30} versus intra-event residual δW_{es} of various types of earthquakes after determining the magnitude and path terms. With increasing V_{S30}, δW_{es} gradually decreases, and the slope of the trend line changes when V_{S30} is greater than $V_{S30}{}^*$, where $V_{S30}{}^*$ is a reference value calculated from Fig. 6 by piecewise linear fitting. Therefore, the expression of the V_{S30} term is as follows:

$$F_{VS30} = \begin{cases} a_7 \ln V_{S30}, & V_{S30} \le V_{S30}^* \\ a_7 \ln V_{S30}^* + a_8 \left(\ln V_{S30} - \ln V_{S30}^* \right), & V_{S30} > V_{S30}^* \end{cases} \quad (9)$$

Figure 7 shows the scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by V_{S30}. As V_{S30} decreases, the slope of scatter points becomes larger, and the peak value of the probability distribution histogram gradually moves to longer durations, which indicates that the duration has a negative correlation with V_{S30}.

3.4 Basin term

It is difficult to quantify the basin effect due to the complexity. At present, only the basin depth has been considered in GMPEs: Z_n (the depth to shear wave velocity of n km/s for the site, e.g., Z_1, $Z_{1.5}$ and $Z_{2.5}$) and δZ_n (the difference between the depth to shear wave velocity of n km/s for the site and a median depth conditional on V_{S30}, e.g., δZ_1 $\delta Z_{1.5}$, and $\delta Z_{2.5}$). Z_n can be obtained from the J-SHIS website. Note that J-SHIS directly provides Z_1 and $Z_{1.5}$, and $Z_{2.5}$ is calculated by modified Akima cubic Hermite interpolation.

In this study, it is found that δZ_1 and Z_1 are selected as the basin depth indicators to have the best fitting effect (the AIC value of the GMPE is small). Considering that both Afshari and Stewart (2016) and Bahrampouri et al. (2021) used δZ_1 as the basin depth index, we also choose δZ_1 to describe the basin effect. Figure 8 shows the distribution of the intra-event residual δW_{es} versus δZ_1 after considering the magnitude, path and V_{S30} terms. When δZ_1 is less than $\delta Z_1{}^*$, δW_{es} increases with increasing δZ_1, and greater than $\delta Z_1{}^*$, δW_{es} is almost constant. Therefore, the basin effect term is written as follows:

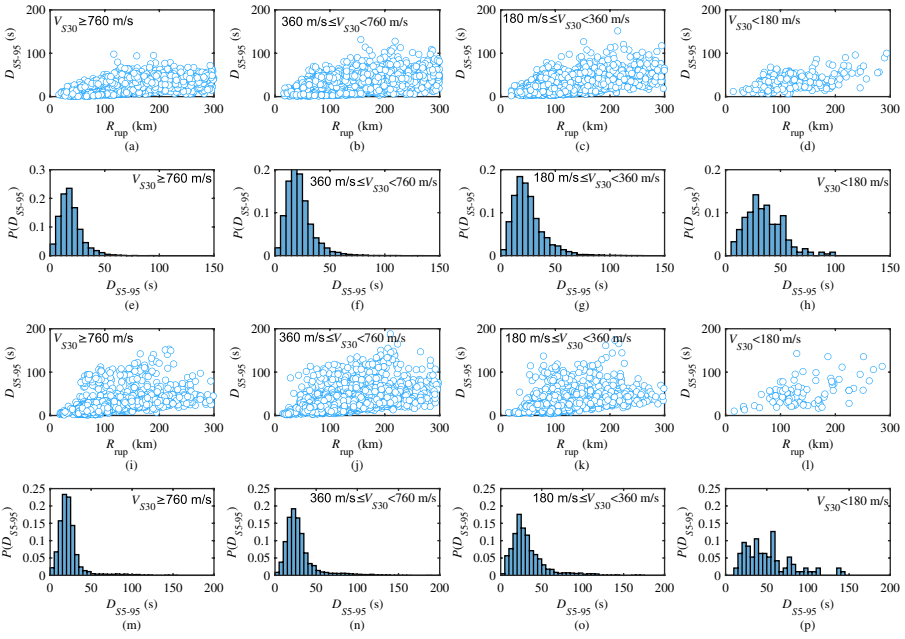

Fig. 7 Scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by V_{S30}: **a–h** subduction intraslab earthquakes and **i–p** subduction interface earthquakes

$$F_{\text{basin}} = a_9 \min\left(\delta Z_1, \delta Z_1^*\right) \tag{10}$$

where $\delta Z_1{}^*$ is a reference value calculated from Fig. 8 by piecewise linear fitting.

Figure 9 shows a scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by δZ_1. As δZ_1 decreases, the slope of scatter points becomes smaller, and the peak value of the probability distribution histogram gradually moves to shorter durations, which indicates that the duration has a positive correlation with δZ_1.

3.5 Depth term

The influence of depth on ground-motion D_S is not reflected in most studies, but Z_{tor} is used in BSA09 and DW17, so we also consider the depth term in the GMPE. Figure 10 shows the distribution of the inter-event residual δB_e versus Z_{tor} after considering the magnitude, path, V_{S30} and basin terms. It can be found that δB_e basically shows a decreasing trend with increasing Z_{tor}. For subduction interface earthquakes, δB_e does not change with increasing Z_{tor} within the range of 50 km. When Z_{tor} is greater than 50 km, δB_e shows a decreasing trend. The expression of the depth term is as follows:

$$F_{\text{depth}} = a_{10} \max\left(Z_{\text{tor}}, Z_{\text{tor}}^*\right) \tag{11}$$

here, Z_{tor}^* equals 50 km for interface earthquakes and 0 for intraslab earthquakes.

Figure 11 shows a scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by Z_{tor}. As Z_{tor} decreases, the slope of scatter points becomes larger, and the peak value of the probability distribution histogram gradually moves to longer durations, which indicates that the duration has a negative correlation with Z_{tor}.

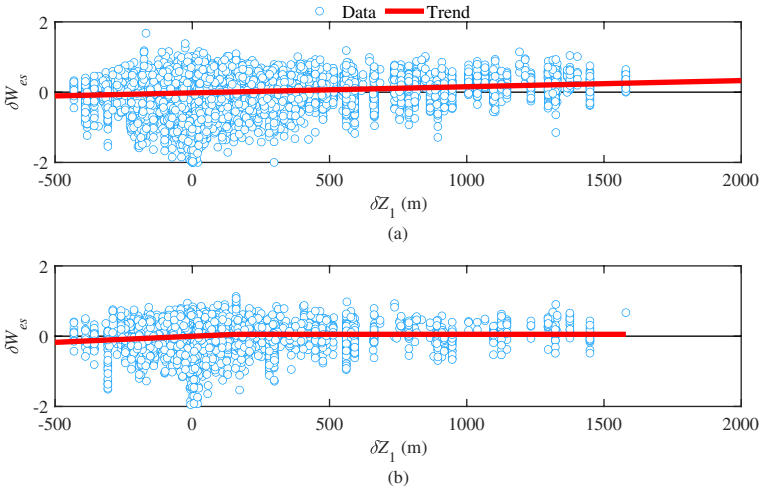

Fig. 8 Distribution of δW_{es} versus δZ_1: **a** subduction intraslab earthquakes; **b** subduction interface earthquakes

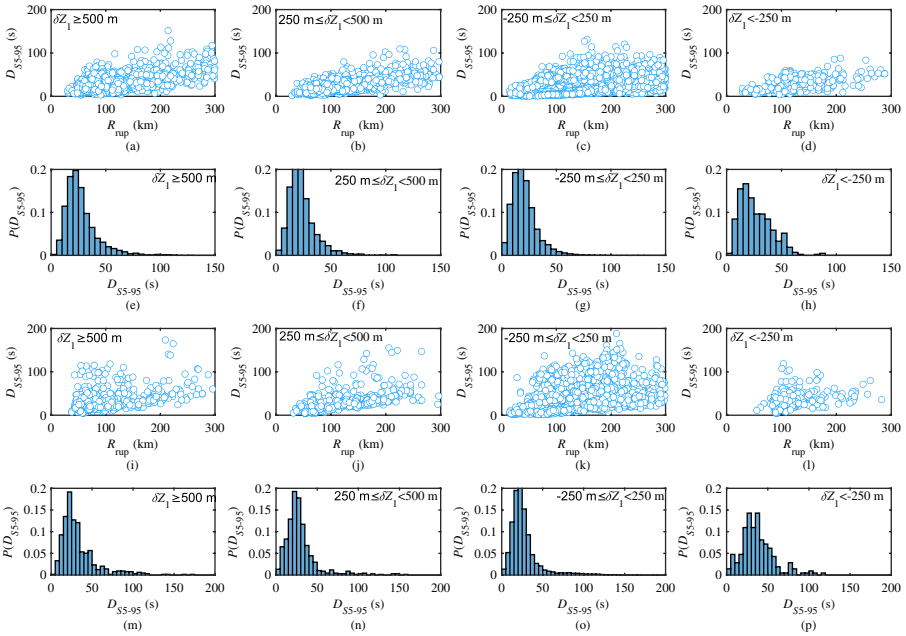

Fig. 9 Scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by δZ_1: **a**–**h** subduction intraslab earthquakes and **i**–**p** subduction interface earthquakes

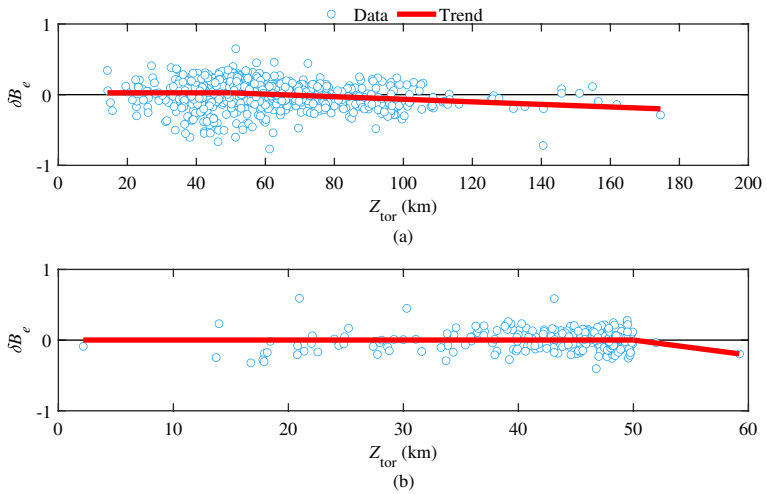

Fig. 10 Distribution of δB_e versus Z_{tor}: **a** subduction intraslab earthquakes; **b** subduction interface earthquakes

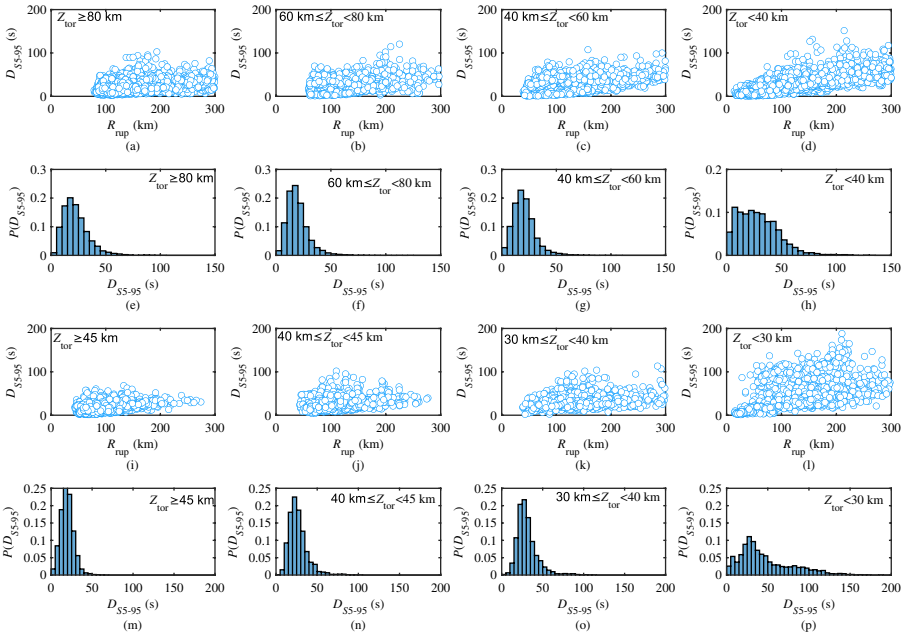

Fig. 11 Scatter plot of $D_{S5\text{-}95}$ versus R_{rup} and probability distribution of $D_{S5\text{-}95}$, which are grouped by Z_{tor}: **a**–**h** subduction intraslab earthquakes and **i**–**p** subduction interface earthquakes

3.6 Model performance

The intra-event residual δW_{es} is further decomposed into the site-to-site residual $\delta S2S_s$ and the intra-event single-site residual δWS_{es} (Baltay et al. 2017; Chen and Faccioli 2013). Thus, Eq. (2) is rewritten as:

$$\ln D_{S,es} = \ln \hat{D}_{S,es} + \delta B_e + \delta W_{es} = \ln \hat{D}_{S,es} + \delta B_e + \delta S2S_s + \delta WS_{es} \tag{12}$$

where $D_{S,es}$ is the observation value of the duration, $\hat{D}_{S,es}$ represents the predicted value calculated by the GMPE, the standard deviation of δW_{es} is σ, the standard deviation of δB_e is τ, and the corresponding standard deviations of $\delta S2S_s$ and δWS_{es} are ϕ_{s2s} and ϕ_{ss}, respectively. The random effect model is used to regress the GMPE proposed in this study. The regression results of each coefficient, standard deviation and *t*-distribution statistics are shown in Tables 1 and 2. Table 3 shows the AIC values of the GMPEs considering different explanatory variables. From the *t*-distribution statistics of explanatory variables given in Tables 1 and 2, the values are all greater than 1.96, and the AIC values gradually decrease as the number of explanatory variables increases, which proves that each explanatory variable in the GMPE is significant and that the fitting effect of the prediction equation gradually tends to be optimal. In the GMPE, $M_1{}^*$, $M_2{}^*$, $M_3{}^*$, $M_4{}^*$, R^*, $V_{S30}{}^*$, $\delta Z_1{}^*$ and $Z_{\text{tor}}{}^*$ are obtained by observation and trial and error, so these parameters have no standard deviation.

Figures 12 and 13 show the distribution of residual of each explanatory variable in the GMPEs for $D_{S5\text{-}95}$. The scatters of residuals are relatively evenly distributed on both sides of the zero line, and the trend line basically coincides with the zero line. The distribution

Table 1 Regression results of the GMPE for D_S of subduction intraslab earthquakes

Coefficient	$D_{S5\text{-}75}$	$\lvert t \rvert$	$D_{S5\text{-}95}$	$\lvert t \rvert$
a_1	-11.9265 ± 0.2970	40.1612	-5.1012 ± 0.2312	22.0647
a_2	2.0165 ± 0.0505	39.9206	1.4659 ± 0.0333	43.9996
a_3	1.3734 ± 0.0552	24.8837	0.9928 ± 0.0363	27.3397
a_4	2.4221 ± 0.0712	34.0146	1.7249 ± 0.0478	36.1230
a_5	-0.2500 ± 0.0118	21.1746	-0.1721 ± 0.0078	22.0589
a_6	–	–	–	–
a_7	-0.1453 ± 0.0071	20.4869	-0.4843 ± 0.0221	21.9342
a_8	-0.5517 ± 0.0495	11.1358	-0.1264 ± 0.0048	26.5824
a_9	0.0008 ± 0.00003	25.0966	0.0003 ± 0.00001	37.6294
a_{10}	-0.0032 ± 0.0002	20.9632	-0.0022 ± 0.0001	19.4385
M_1^*	4.5	–	4.5	–
M_2^*	–	–	–	–
M_3^*	4.5	–	4.5	–
M_4^*	6	–	6.5	–
R^*	80	–	80	–
V_{S30}^*	1300	–	300	–
δZ_1^*	160	–	Inf	–
Z_{tor}^*	0	–	50	–
σ_{T}	0.6455	–	0.4571	–
T	0.2304	–	0.1919	–
ϕ_{ss}	0.4713	–	0.3055	–
ϕ_{s2s}	0.3761	–	0.2807	–

of residuals in some explanatory variables is biased, which may be due to two reasons: (1) the data corresponding to the residual migration part are relatively sparse, which cannot provide sufficient data support for the regression analysis, and (2) the lack of accurate site information and fault models makes the data have deviations compared with the actual situation. However, in general, the deviation of the residuals is not large and in most cases is unbiased. Figure 14 shows the D_S prediction values of D_S with different magnitudes and different site conditions, the D_S increases with increasing M_{w} and R_{rup}, and decreases with increasing V_{S30}. Due to the influence of δZ_1 and Z_{tor} is not as significant as other parameters on D_S, thus, this study does not give the prediction curves of D_S under different δZ_1 and Z_{tor}. However, from the regression results, D_S increases with increasing δZ_1, and decreases with increasing Z_{tor}. The above rules are consistent with the existing research results. Figure 15 shows the variation in D_S of various types of earthquakes with the R_{rup} under different magnitudes. The prediction values of D_S of subduction intraslab and interface earthquakes are significantly different, the prediction values of subduction interface earthquakes are generally longer than those of intraslab earthquakes. Figure 14a, c and Fig. 15 show that at $M_{\text{w}}5$, there is a prominent abrupt change in the slope of D_S of subduction intraslab earthquakes, and at $M_{\text{w}}7$, the slope of D_S of subduction intraslab earthquakes is slightly larger than that of interface earthquakes. Therefore, it can be concluded that the differences in earthquake types should be considered when discussing issues related to the duration, such as seismic risk assessment and structural response.

Table 2 Regression results of the GMPE for D_S of subduction interface earthquakes

Coefficient	$D_{S5\text{-}75}$	$\mid t \mid$	$D_{S5\text{-}95}$	$\mid t \mid$
a_1	-4.7865 ± 0.3335	14.3541	0.8164 ± 0.4015	2.0334
a_2	0.9455 ± 0.0540	17.5248	0.9526 ± 0.0415	22.9323
a_3	0.8244 ± 0.0610	13.5154	0.4076 ± 0.0362	11.2697
a_4	1.2397 ± 0.1077	11.5066	0.6095 ± 0.0451	13.5180
a_5	-0.0693 ± 0.0214	3.2380	-0.0285 ± 0.0076	3.7576
a_6	-0.1220 ± 0.0124	9.8145	–	–
a_7	-0.1527 ± 0.0116	13.1780	-0.5460 ± 0.0352	15.5193
a_8	-0.7252 ± 0.0658	11.0254	-0.1782 ± 0.0070	25.3721
a_9	0.0006 ± 0.00006	9.3028	0.0006 ± 0.00003	17.2369
a_{10}	-0.0198 ± 0.0015	13.1091	-0.0247 ± 0.0056	4.3729
M_1^*	5.5	–	4.5	–
M_2^*	4.5	–	–	–
M_3^*	0	–	1.5	–
M_4^*	6	–	6	–
R^*	80	–	80	–
$V_{S30}{}^*$	1300	–	300	–
$\delta Z_1{}^*$	90	–	160	–
$Z_{tor}{}^*$	40	–	50	–
σ_T	0.5512	–	0.4018	–
τ	0.2006	–	0.1515	–
ϕ_{ss}	0.4070	–	0.2588	–
ϕ_{s2s}	0.3129	–	0.2673	–

Table 3 AIC values of the GMPE considering different explanatory variables

Explanatory variables	Subduction intraslab		Subduction interface	
	$D_{S5\text{-}75}$	$D_{S5\text{-}95}$	$D_{S5\text{-}75}$	$D_{S5\text{-}95}$
M_w, R_{rup}	48,993.54	27,566.59	16,679.94	8284.98
M_w, R_{rup}, V_{S30}	48,966.99	27,530.57	16,661.41	8280.66
M_w, R_{rup}, V_{S30}, δZ_1	48,905.95	27,447.53	16,643.66	8274.06
M_w, R_{rup}, V_{S30}, δZ_1, Z_{tor}	48,861.85	27,415.92	16,624.67	8274.04

3.7 Comparison to prior models

Figure 16 shows the comparison between the D_S prediction model in this paper and the results of previous work. The prediction results of the subduction intraslab and interface earthquake models in this paper are basically between those of the BRG21 model and the JG21 model. However, the predicted values of the models proposed in this study are smaller than those of the BRG21 model. The reasons may be as follows: (1) in this study, we select records with PGA $\geq$ 10 gal, but BRG21 does not limit the PGA. (2) R_{rup} of the data set in this paper is within 300 km, and the R_{rup} of the subduction intraslab and interface earthquakes of BRG21 reaches 500 km. Since the discreteness of D_S increases with

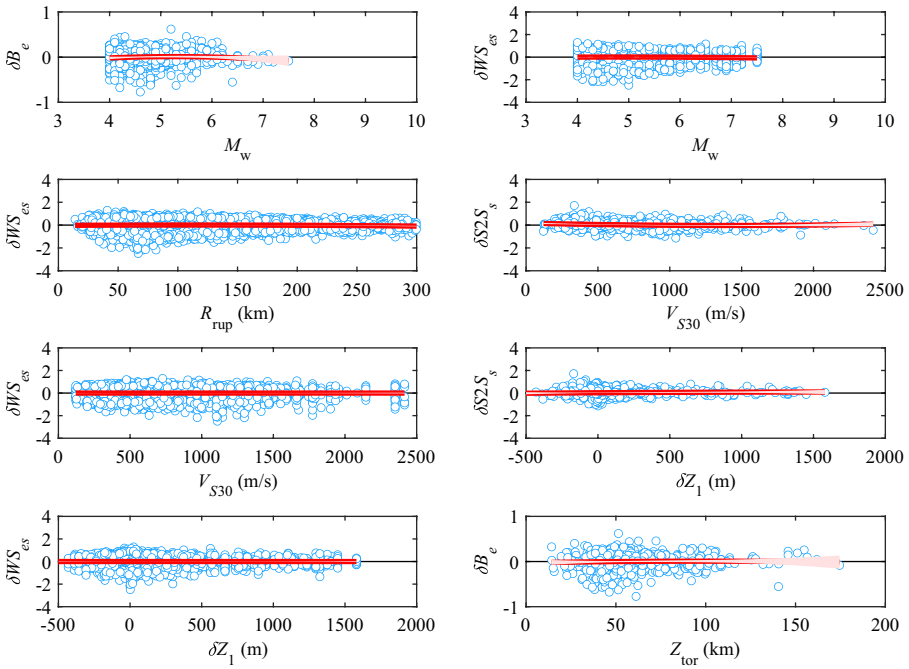

Fig. 12 Distribution of residuals of the GMPE for $D_{S5\text{-}95}$ of subduction intraslab earthquakes versus explanatory variables (the blue dots are residual values, the red solid line is the trend line, and the red area is the 95% confidence interval)

increasing distance, there is great uncertainty in a prediction of D_S at large distance, and when the distance exceeds 300 km, the PGA attenuates to a very low level, which has little impact on the structure. Data beyond 300 km are unnecessary for the follow-up work of this paper. (3) The selection of records in this paper is relatively strict. BRG21 requires SNR ≥ 2, while this paper requires SNR ≥ 3 to ensure higher data quality. BRG21 does not have a clear limit on the number of records for a single earthquake event, but to ensure the statistical significance of the data, it is required that the records of a single earthquake event should not be less than 10 in this study. These constraints cause some data to be ignored, so that the number of records used in this paper is less than that of BRG21, but the database in this study has higher quality than that of BRG21. In addition, the way the focal mechanism solution is obtained can also affect the dip angles of the fault. The criteria for judging fault types, earthquake types, and the selection of the geometry model of the subduction zone are different, resulting in differences in the final earthquake classification. In this study, the geometry model of the subduction zone is selected as the latest Slab 2.0, and the criteria for judging fault types and earthquake types are also the most commonly used and convincing.

Although there are differences between BRG21 and the models in this study, at present, there are very few GMPEs for D_S of the ground motion in the subduction zone of Japan. The results and conclusions of this paper need to be tested after the research results in this area are continuously enriched and improved. However, we simplifie the traditional magnitude term and directly uses M_w to recalibrate it, which reduces the error caused by calculating the source

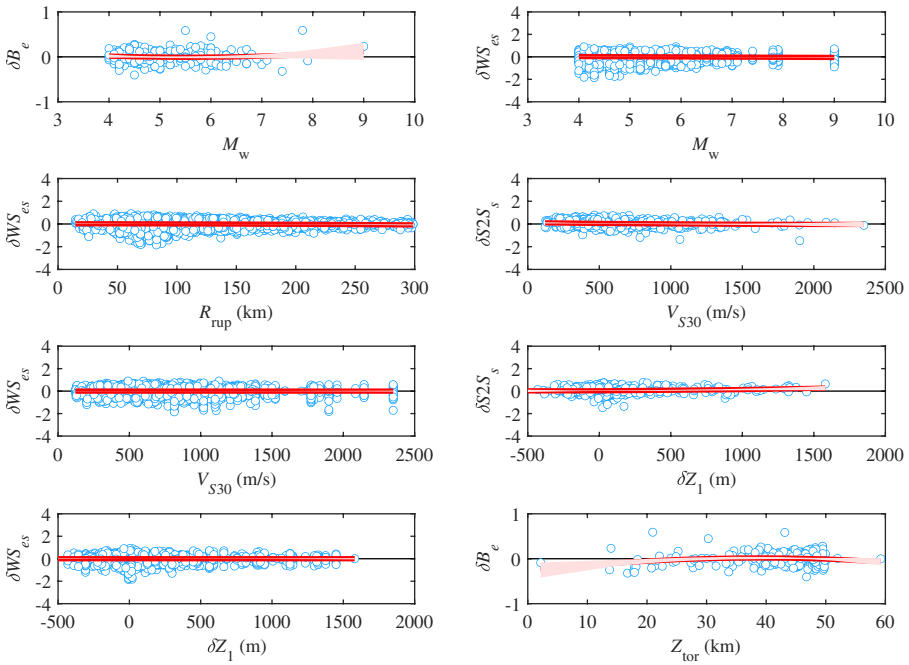

Fig. 13 Distribution of residuals of the GMPE for $D_{S5\text{-}95}$ of subduction interface earthquakes versus explanatory variables (the blue dots are residual values, the red solid line is the trend line, and the red area is the 95% confidence interval)

duration from the stress drop. We also consider a focal depth term in the GMPEs, which helps improve the prediction accuracy of the model. In addition, the authors also investigated the influence of fault type, volcanic prearc and postarc distances and volcanic area distance (Zhao et al. 2016a, b, c) on D_S, but the results show that the above factors have no significant influence on D_S of this database. Therefore, the terms corresponding to the above factors are not added to the GMPE. With the continuous expansion of the database, the GMPE in this study is planned to be updated, and the authors plan to improve the database and the GMPEs in future work.

Note that although R_{rup} of the database used in this study reaches 300 km, R_{rup} of most earthquake events (especially for $M_w < 5.5$) is much less than 300 km and basically reaches only 200 km; the data beyond 200 km are relatively scarce. Moreover, with increasing distance, the discreteness of duration also gradually increases, especially when $M_w \geq 6$ and $R_{rup} \geq 200$ km, and the distribution range of duration can reach tens of seconds. Therefore, we suggest that the distance range of the GMPEs proposed in this study be within 200 km. The distance range of the database should reach 350 km or even 400 km to ensure higher accuracy in predicting D_S of R_{rup} close to 300 km.

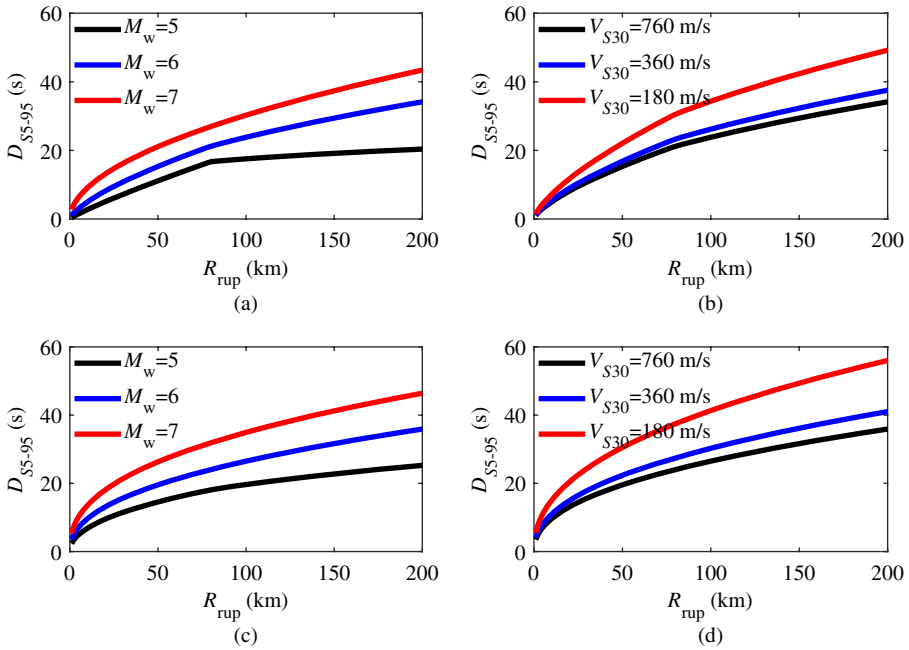

Fig. 14 Predicted values of $D_{S5\text{-}95}$ ($\delta Z_1 = 100$ m, $Z_{tor} = 25$ km): **a** subduction intraslab, $V_{S30} = 760$ m/s; **b** subduction intraslab, $M_w = 6$; **c** subduction interface, $V_{S30} = 760$ m/s; **d** subduction interface, $M_w = 6$

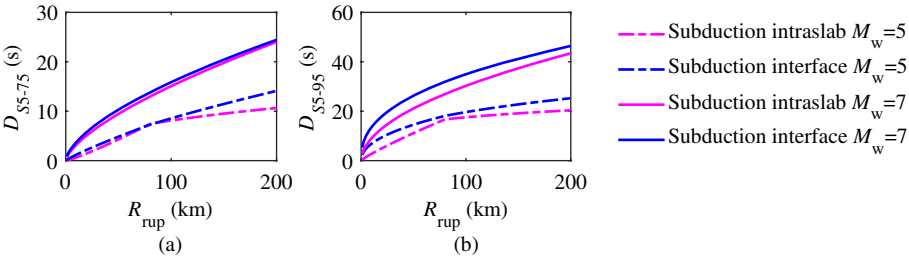

Fig. 15 Predicted values of D_S under different magnitudes: **a** $D_{S5\text{-}75}$; **b** $D_{S5\text{-}95}$ ($V_{S30} = 760$ m/s, $\delta Z_1 = 100$ m, $Z_{tor} = 25$ km)

4 Spatial correlation of intra-event residuals

4.1 Semivariogram and spatial correlation coefficient

As discussed, the GMPE is a ground-motion probability distribution model established based on parameters such as magnitude, distance, and site. The inter-event residual in the GMPE describe the average deviation away from the median GMPE for any one event, while the intra-event residual is the remaining variation of a single record about the event-specific average GMPE. For the same event, closer sites show stronger

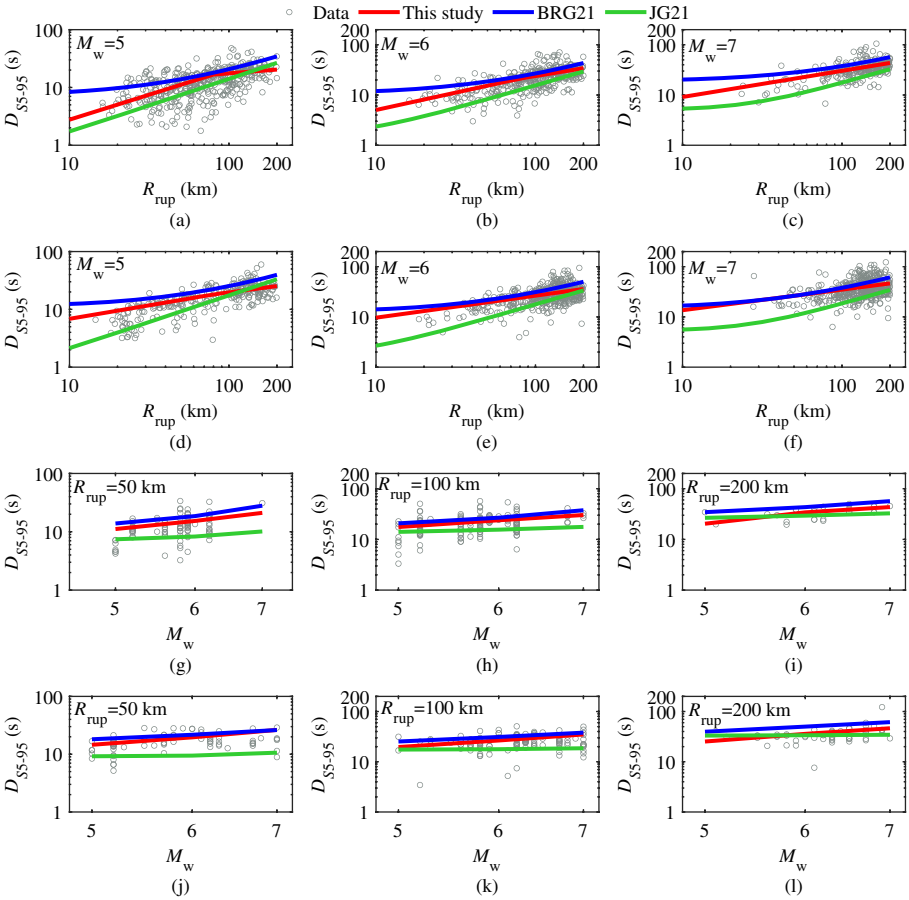

Fig. 16 Comparison of predictions of different models for significant duration with results from this study: **a**, **b**, **c**, **g**, **h**, **i** subduction intraslab earthquakes and **d**, **e**, **f**, **j**, **k**, **l** subduction interface earthquakes (V_{S30} = 760 m/s, δZ_1 = 100 m, Z_{tor} = 25 km)

correlation of the ground-motion parameters, constituting spatial correlation of ground motion (Chen et al. 2020; Goda and Hong 2008; Jayaram and Baker 2009).

Evaluating the correlation of intra-event residuals is usually divided into the following steps: (a) calculating the intra-event residuals of the prediction model; (b) fitting the semivariogram of the residual in the event; and (c) calculating the spatial correlation coefficient of the residual in the event. Fitting the semivariogram (step (b)) is the most critical step. A semivariogram is one of the indicators describing spatial correlation in geostatistics. If the semivariogram depends only on the separation distance h rather than the actual position, the semivariogram is considered to be second-order stationary, and the expression is as follows:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left(\varepsilon_{ij} - \varepsilon_{ik}\right)^2 \tag{13}$$

where ε_{ij} and ε_{ik} represent the intra-event residuals of the jth and kth sites of the ith earthquake event, respectively, $i=1, 2, \ldots, N(h)$, and $N(h)$ is the number of stations $[a_i, b_i]$ satisfying $h-\Delta h/2 \le |a_i-b_i| \le h+\Delta h/2$. To ensure the reliability of the statistics, $\Delta h=2$ km. The above formula is the method of moments proposed by Matheron (1962). Cressie (1985) proposed a more robust estimator, which is relatively insensitive to discrete values:

$$\hat{\gamma}(h) = \frac{1}{2}\left\{\frac{\left[\frac{1}{|N(h)|}\sum_{i=1}^{N(h)}|\varepsilon_{ij}-\varepsilon_{ik}|^{0.5}\right]^4}{0.457+\frac{0.494}{|N(h)|}}\right\} \tag{14}$$

Therefore, Eq. (14) is selected as the estimator of the semivariogram in this paper. $N(h)$ should be as large as possible. Esposito and Iervolino (2012) and Du and Wang (2013) suggested that $N(h)$ should be at least 30, while Wagener et al. (2016) suggested that $N(h)$ should be at least 100 to ensure more reliable and representative estimation results. The parameter should not be too small, at least not less than 30. However, when is $N(h)$ too large, it is difficult to accurately evaluate the spatial correlation coefficients because there are too few data records that meet the requirements. Based on repeated reductions in $N(h)$ and corresponding evaluation, we suggest that $N(h)=50$ is more appropriate.

The continuous semivariogram can be fitted from the estimated value of Eq. (14). There are many commonly used semivariogram fitting models based on 2nd-order stationarity and isotropy. In this study, the exponential model is selected because the exponential model is considered to be the model with the best fitting effect (Jayaram and Baker 2009). The exponential model selected is represented by the following formula:

$$\gamma(h) = a\left[1-\exp\left(-\frac{3h}{b}\right)\right] \tag{15}$$

where a and b are the sill and the range of the semivariogram model, respectively. The sill a is equal to the variance in the intra-event residual ε, and the range b is the separation distance corresponding to the semivariogram equal to 0.95 times the sill. The intra-event residual is normalized; that is, the intra-event residual ε is divided by its standard deviation φ. Then, Eq. (15) can be simplified as:

$$\gamma(h) = 1-\exp\left(-\frac{3h}{b}\right) \tag{16}$$

The choice of the standard deviation φ is extremely important here, and Esposito and Iervolino (2012) use the standard deviation of the intra-event residuals regressed from the GMPEs to be normalized as φ. However, since the standard deviation of intra-event residual of each earthquake event is different, using a fixed standard deviation leads to a biased intra-event correlation. Therefore, each earthquake event should use its own standard deviation (Foulser-Piggott and Goda 2015; Schiappapietra and Douglas 2020). After normalization, the spatial correlation coefficient can be calculated:

$$\rho(h) = 1-\gamma(h) = \exp\left(-\frac{3h}{b}\right) \tag{17}$$

Table 4 Fitting values of ranges for semivariogram exponential models

Earthquake type	Subduction intraslab				Subduction interface			
D_S	$D_{S5\text{-}75}$		$D_{S5\text{-}95}$		$D_{S5\text{-}75}$		$D_{S5\text{-}95}$	
Residual	δWS	δW	δWS	δW	δWS	δW	δWS	δW
Range b	64	78	71	80	101	101	90	96

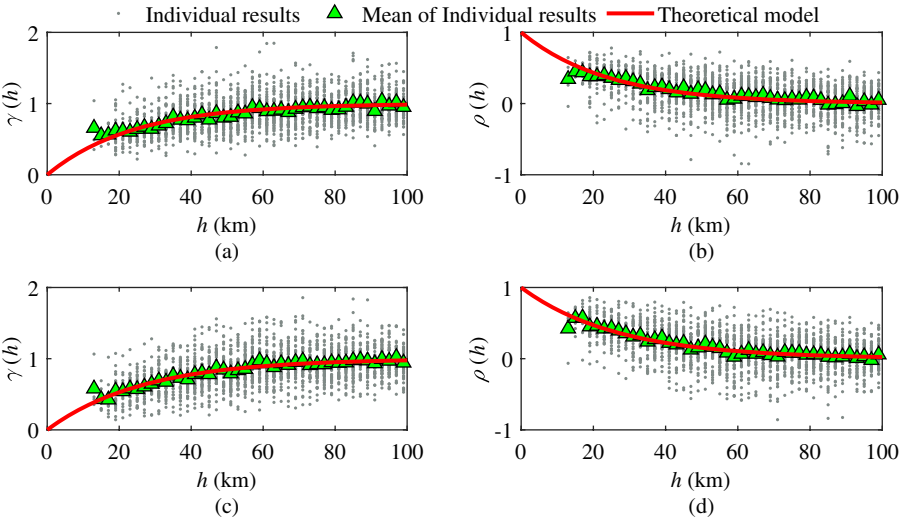

Fig. 17 Fitting results of the semivariogram and spatial correlation coefficient for $D_{S5\text{-}95}$ of subduction intraslab earthquakes: **a** fitting results of the semivariogram (based on δWS); **b** spatial correlations (based on δWS); **c** fitting results of the semivariogram (based on δW); **d** spatial correlations (based on δW)

Since the intra-event residual δW can be decomposed into the site-to-site residual $\delta S2S$ and the intra-event single-site residual δWS according to Eq. (12), the semivariogram can be used not only to measure the spatial correlation of intra-event residual but also to analyze the spatial correlation of intra-event single-site residual. Therefore, this study also gives the spatial correlation analysis results based on the intra-event residual δW and the intra-event single-site residual δWS.

4.2 Results of spatial correlation models

The semivariogram values at each separation distance of various types of earthquakes are fitted, referring to Foulser-Piggott and Goda (2015), and the fitting value of range b is shown in Table 4. Figures 17 and 18 show the fitting results of the semivariogram and spatial correlation coefficient for $D_{S5\text{-}95}$. The spatial correlation coefficient of D_S decreases rapidly with increasing separation distance; when the separation distance reaches 100 km, the spatial correlation coefficient is reduced to approximately 0. In addition, the spatial correlation curve of D_S of a single earthquake event is basically the same as that of the combined earthquake event, but the spatial correlation curve of a single event has a large variability, especially when the separation distance is large. Both the spatial correlation coefficient based on δWS and that based on δW show the characteristics of decay with

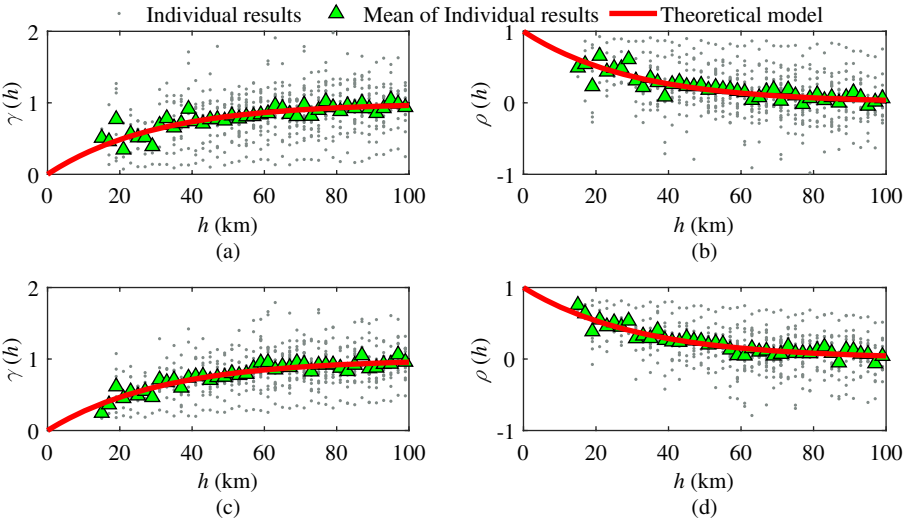

Fig. 18 Fitting results of the semivariogram and spatial correlation coefficient for $D_{S5\text{-}95}$ of subduction interface earthquakes: **a** fitting results of the semivariogram (based on δWS); **b** spatial correlations (based on δWS); **c** fitting results of the semivariogram (based on δW); **d** spatial correlations (based on δW)

increasing separation distance, indicating that the spatial correlation is related not only to the site effect but also to factors unrelated to site conditions such as path effects. The above conclusions are consistent with the results of the spatial correlation analysis of the Arias intensity by Foulser-Piggott and Goda (2015). Figure 19 shows the comparison of spatial

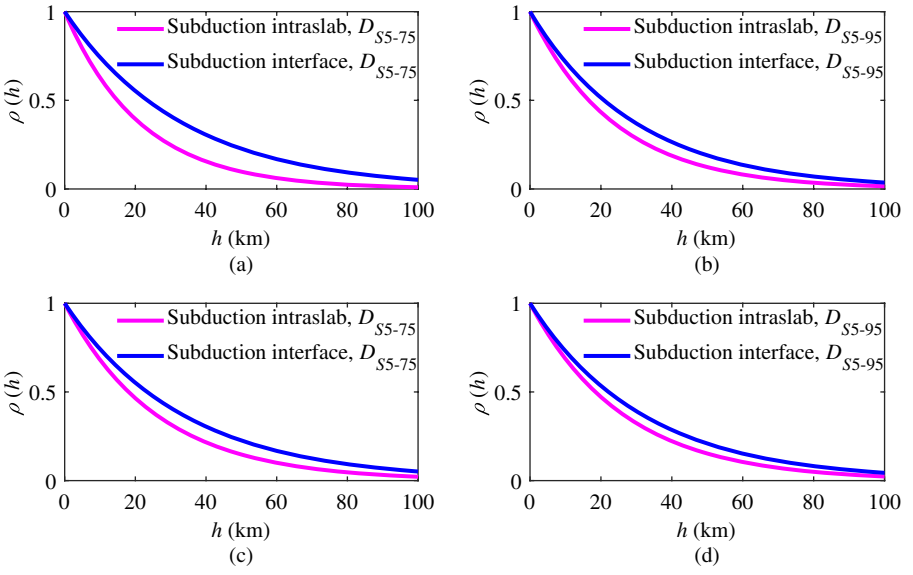

Fig. 19 Comparison of D_S spatial correlation curves: **a**, **b** Spatial correlations based on δWS and **c**, **d** spatial correlations based on δW

correlation curves for D_S of different types of earthquakes. With increasing separation distance, the attenuation rate of the spatial correlation coefficient of subduction interface earthquakes grows slower than that of subduction intraslab earthquakes.

5 Conclusions

In this study, for subduction intraslab and interface earthquakes in Japan, we first consider the effects of M_w, R_{rup}, V_{S30}, δZ_1, and Z_{tor} on D_S. Then, GMPEs for $D_{S5\text{-}75}$ and $D_{S5\text{-}95}$ are developed, and the rationality and reliability of these GMPEs are proven by residual analysis and comparison with previous models. Finally, based on the intra-event residuals, the spatial correlation of D_S is briefly analyzed. The results show that (1) D_S increases with increasing M_w, R_{rup} and δZ_1 and decreases with increasing V_{S30} and Z_{tor}, and the path term is dependent on M_w. (2) D_S of subduction interface earthquakes is greater than that of subduction intraslab earthquakes on the whole; and under large magnitudes, the slope of D_S of subduction intraslab earthquakes is slightly greater than that of subduction interface earthquakes; the differences of the earthquake types should be taken into account when discussing issues related to the duration of ground motions such as seismic risk assessment and structural response. (3) The spatial correlation of D_S is related not only to site effects but also to factors unrelated to site conditions, such as path effects, and the spatial correlation coefficient decreases with increasing separation distance. When the separation distance reaches 100 km, the spatial correlation coefficient decreases to approximately 0, but the spatial correlation coefficient attenuation rate of the subduction interface earthquake is slower than that of the intraslab subduction earthquake. (4) The variation trend of the spatial correlation curve of a single earthquake event is basically consistent with the overall average spatial correlation curve of earthquake events, but the spatial correlation curve of a single earthquake event has great variability, especially for large separation distances.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10518-023-01749-0.

Acknowledgements The authors greatly appreciate support by the National Natural Science Foundation of China (Grant No. U2139209, Grant No. 52178514), the Scientific Research Fund of Institute of Engineering Mechanics, China Earthquake Administration (Grant No. 2019A01, Grant No. 2021EEEVL0301) and the Heilongjiang Touyan Innovation Team Program. The authors acknowledge the Kyoshin Network (K-NET) and Kiban Kyoshin Network (KiK-net) for providing the strong ground-motion data and site list, the Japan Meteorological Agency (JMA) for the information regarding the earthquake events, the Global Centroid-Moment-Tensor (CMT) and Full Range Seismograph Network (F-net) of Japan for the focal mechanism solutions, the United States Geological Survey (USGS), K-NET, KiK-net and SRCMOD for the stochastic finite fault models, and the Japan Seismic Hazard Information Station (J-SHIS) for the subsurface structure models.

Funding This work is supported by the National Natural Science Foundation of China (Grant No. U2139209, Grant No. 52178514), the Scientific Research Fund of Institute of Engineering Mechanics, China Earthquake Administration (Grant No. 2019A01, Grant No. 2021EEEVL0301) and the Heilongjiang Touyan Innovation Team Program.

Data availability The databases generated and/or analyzed during the current study are available from the author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abrahamson NA, Silva WJ (1996) Empirical ground motion models. Report to Brookhaven National Laboratory

Abrahamson NA, Youngs RR (1992) A stable algorithm for regression analysis using the random effects model. Bull Seismol Soc Am 82(1):505–510

Afshari K, Stewart JP (2016) Physically parameterized prediction equations for significant duration in active crustal regions. Earthq Spectra 32(4):2057–2081

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Akkar S, Kale Ö, Yakut A, Çeken U (2018) Ground-motion characterization for the probabilistic seismic hazard assessment in Turkey. Bull Earthq Eng 16(8):3439–3463

Albarello D, Peruzza L (2017) Accounting for spatial correlation in the empirical scoring of probabilistic seismic hazard estimates. Bull Earthq Eng 15(6):2571–2585

Ameri G, Drouet S, Traversa P, Bindi D, Cotton F (2017) Toward an empirical ground motion prediction equation for France: accounting for regional differences in the source stress parameter. Bull Earthq Eng 15(11):4681–4717

Anbazhagan P, Sheikh MN, Bajaj K, Mariya Dayana PJ, Madhura H, Reddy GR (2017) Empirical models for the prediction of ground motion duration for intraplate earthquakes. J Seismol 21:1001–1021

Bahrampouri M, Rodriguez-Marek A, Shahi S, Dawood H (2020) An updated database for ground motion parameters for KiK-net records. Earthq Spectra 37(4):505–522

Bahrampouri M, Rodriguez-Marek A, Green RA (2021) Ground motion prediction equations for significant duration using the KiK-net database. Earthq Spectra 37(2):903–920

Baltay AS, Hanks TC, Abrahamson NA (2017) Uncertainty, variability, and earthquake physics in ground-motion prediction equations. Bull Seismol Soc Am 107(4):1754–1772

Bommer JJ, Magenes G, Hancock J, Penazzo P (2004) The influence of strong-motion duration on the seismic response of masonry structures. Bull Earthq Eng 2(1):1–26

Bommer JJ, Stafford PJ, Alarcón JA (2009) Empirical equations for the prediction of the significant, bracketed, and uniform duration of earthquake ground motion. Bull Seismol Soc Am 99(6):3217–3233

Boore DM (2005) On pads and filters: processing strong-motion data. Bull Seismol Soc Am 95(2):745–750

Boore DM, Akkar S (2003) Effect of causal and acausal filters on elastic and inelastic response spectra. Earthq Eng Struct Dyn 32(11):1729–1748

Boore DM, Gibbs JF, Joyner WB, Tinsley JC, Ponti DJ (2003) Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California. Bull Seismol Soc Am 93(6):2737–2751

Boore DM, Thompson EM, Cadet H (2011) Regional correlations of V_{S30} and velocities averaged over depths less than and greater than 30 meters. Bull Seismol Soc Am 101(6):3046–3059

Bozorgnia Y (2020) Data resources for NGA-subduction project. University of California, Berkeley

Bradley BA (2014) Site-specific and spatially-distributed ground-motion intensity estimation in the 2010–2011 Canterbury earthquakes. Soil Dyn Earthq Eng 61–62:83–91

Bravo-Haro MA, Elghazouli AY (2018) Influence of earthquake duration on the response of steel moment frames. Soil Dyn Earthq Eng 115:634–651

Bravo-Haro MA, Liapopoulou M, Elghazouli AY (2020) Seismic collapse capacity assessment of SDOF systems incorporating duration and instability effects. Bull Earthq Eng 18:3025–3056

Chandramohan R, Baker JW, Deierlein GG (2016) Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. Earthq Spectra 32(2):927–950

Chen L, Faccioli E (2013) Single-station standard deviation analysis of 2010–2012 strong-motion data from the Canterbury region, New Zealand. Bull Earthq Eng 11(5):1617–1632

Chen K, Yu YX, Gao MT, Kang CC (2020) Study on spatial correlation of ground-motion: a case study of Napa earthquake. Seismol Geol 42(5):1218–1228 (in Chinese)

Cosenza E, Iervolino I, Manfredi G (2004) On ground motion duration and engineering demand parameters. In: International workshop on performance-based seismic design, concepts and implementation, Bled, Slovenia

Cressie N (1985) Fitting variogram models by weighted least squares. Math Geol 17(5):563–586

Dobry R, Idriss IM, Ng E (1978) Duration characteristics of horizontal components of strong-motion earthquake records. Bull Seismol Soc Am 68(5):1487–1520

Du W, Wang G (2013) Intra-event spatial correlations for cumulative absolute velocity, Arias intensity, and spectral accelerations based on regional site conditions. Bull Seismol Soc Am 103(2A):1117–1129

Du W, Wang G (2017) Prediction equations for ground-motion significant durations using the NGA-West2 database. Bull Seismol Soc Am 107(1):319–333

Du W, Yu X, Ning C (2020) Influence of earthquake duration on structural collapse assessment using hazard-consistent ground motions for shallow crustal earthquakes. Bull Earthq Eng 18(7):3005–3023
Ebrahimian H, Jalayer F, Forte G, Convertito V, Licata V, d'Onofrio A, Santo A, Silvestri F, Manfredi G (2019) Site-specific probabilistic seismic hazard analysis for the western area of Naples, Italy. Bull Earthq Eng 17(9):4743–4796
Esposito S, Iervolino I (2012) Spatial correlation of spectral acceleration in European data. Bull Seismol Soc Am 102(6):2781–2788
Esteva L, Rosenblueth E (1964) Espectros de temblores a distancias moderadas y grandes. Bull Mex Soc Seismic Eng 2:1–18
Foulser-Piggott R, Goda K (2015) Ground-motion prediction models for Arias intensity and cumulative absolute velocity for Japanese earthquakes considering single-station sigma and within-event spatial correlation. Bull Seismol Soc Am 105(4):1903–1918
Goda K, Atkinson GM (2010) Intraevent spatial correlation of ground-motion parameters using SK-net data. Bull Seismol Soc Am 100(6):3055–3067
Goda K, Hong HP (2008) Spatial correlation of peak ground motions and response spectra. Bull Seismol Soc Am 98(1):354–365
Hancock J, Bommer JJ (2006) A state-of-knowledge review of the influence of strong-motion duration on structural damage. Earthq Spectra 22(3):827–845
Hayes GP, Moore GL, Portner DE, Hearne M, Flamme H, Furtney M, Smoczyk GM (2018) Slab2, a comprehensive subduction zone geometry model. Science 362(4610):58–61
Hernandez B, Cotton F (2000) Empirical determination of the ground shaking duration due to an earthquake using strong motion accelerograms for engineering applications. In: Proceedings: 12th world conference of earthquake engineering
Hong HP, Zhang Y, Goda K (2009) Effect of spatial correlation on estimated ground-motion prediction equations. Bull Seismol Soc Am 99(2A):928–934
Housner GW (1975) Measures of severity of earthquake ground shaking. US National Conference on Earthquake Engineering
Iervolino I, Manfredi G, Cosenza E (2006) Ground motion duration effects on nonlinear seismic response. Earthq Eng Struct Dyn 35(1):21–38
Jaimes MA, García-Soto AD (2021) Ground-motion duration prediction model from recorded Mexican interplate and intermediate-depth intraslab earthquakes. Bull Seismol Soc Am 111(1):258–273
Jayaram N, Baker JW (2009) Correlation model for spatially distributed ground-motion intensities. Earthq Eng Struct Dyn 38(15):1687–1708
Kaklamanos J, Baise LG, Boore DM (2011) Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. Earthq Spectra 27(3):1219–1235
Kamai A, Abrahamson NA, Silva WJ (2016) V_{S30} in the NGA GMPEs-regional differences and suggested practice. Earthq Spectra 32(4):2083–2108
Kamiyama M (1984) Effect of subsoil conditions and other factors on the duration of earthquake ground shaking. In: Proceedings: 8th world conference on earthquake engineering, Sanfrancisco, pp 793–800
Kempton JJ, Stewart PJ (2006) Prediction equations for significant duration of earthquake ground motions consideration site and near- source effects. Earthq Spectra 22(4):958–1013
Lee J, Green RA (2014) An empirical significant duration relationship for stable continental regions. Bull Earthq Eng 12(1):217–235
Liapopoulou M, Bravo-Haro MA, Elghazouli AY (2020) The role of ground motion duration and pulse effects in the collapse of ductile systems. Earthq Eng Struct Dyn 49(11):1051–1071
López-Castañeda AS, Reinoso E (2021) Strong-motion duration predictive models from subduction interface earthquakes recorded in the hill zone of the Valley of Mexico. Soil Dyn Earthq Eng 144:106676
Mai PM, Spudich P, Boatwright J (2005) Hypocenter locations in finite-source rupture models. Bull Seismol Soc Am 95(3):965–980
Matheron G (1962) Traité de géostatistique appliquée. Editions Techniques
McGuire RK, Barnhard TP (1979) The usefulness of ground motion duration in prediction of severity shaking. In: Proceedings: 2nd national conference on earthquake engineering, Stanford, California, pp 713–722
Meskouris K (1983) Influence of earthquake strong motion duration on nonlinear structural response. In: Transactions of the 7th international conference on structural mechanics in reactor technology, pp 267–275
Ohsaki Y (2008) Introduction to spectral analysis of ground motion 2nd edn
Rahnama M, Manuel L (1996) The effect of strong motion duration on seismic demands. In: 11th World conference on earthquake engineering, Paper 924

Reinoso E, Ordaz M (2001) Duration of strong ground motion during Mexican earthquakes in terms of magnitude, distance to the rupture area and dominant site period. Earthq Eng Struct Dyn 30(5):653–673
Reinoso E, Ordaz M, Guerrero R (2000) Influence of strong ground-motion duration in seismic design of structures. In: 12th World conference on earthquake engineering, Auckland, Paper 1151
Scherbaum F, Schmedes J, Cotton F (2004) On the conversion of source-to-site distance measures for extended earthquake source models. Bull Seismol Soc Am 94(3):1053–1069
Schiappapietra E, Douglas J (2020) Modelling the spatial correlation of earthquake ground motion: Insights from the literature, data from the 2016–2017 Central Italy earthquake sequence and ground-motion simulations. Earth-Sci Rev 203:103139
Stewart JP (2020) Data resources for NGA-Subduction project. University of California, Berkeley
Trifunac MD, Brady AG (1975) A study on the duration of strong earthquake ground motion. Bull Seismol Soc Am 65(3):581–626
Uang CM, Bertero VV (1990) Evaluation of seismic energy in structures. Earthq Eng Struct Dyn 19:77–90
Wagener T, Goda K, Erdik M, Daniell J, Wenzel F (2016) A spatial correlation model of peak ground acceleration and response spectra based on data of the Istanbul earthquake rapid response and early warning system. Soil Dyn Earthq Eng 85:166–178
Wang M, Takada T (2005) Macrospatial correlation model of seismic ground motions. Earthq Spectra 21(4):1137–1156
Xing H, Zhao JX (2021) A ground-motion prediction equation for the western and the southwestern parts of China based on local strong-motion records and an overseas reference model for the vertical component. Bull Seismol Soc Am 111(6):3314–3331
Xu PB, Wen RZ (2018) The prediction equations for the significant duration of strong motion in Chinese mainland. Acta Seismol Sin 40(6):809–819 (**in Chinese**)
Yaghmaei-Sabegh S, Shoghian Z, Sheikh MN (2014) A new model for the prediction of earthquake ground-motion duration in Iran. Nat Hazards 70(1):69–92
Zhao JX, Zhou SL, Gao PJ, Long T, Zhang YB, Thio HK, Lu M, Rhoades DA (2015) An earthquake classification scheme adapted for Japan determined by the goodness of fit for ground-motion prediction equations. Bull Seismol Soc Am 105(5):2750–2763
Zhao JX, Jiang F, Shi P, Xing H, Huang HF, Hou RB, Zhang YB, Yu PC, Lan XW, Rhoades DA, Somerville PG, Irikura K, Fukushima Y (2016a) Ground-motion prediction equations for subduction slab earthquakes in Japan using site class and simple geometric attenuation functions. Bull Seismol Soc Am 106(4):1535–1551
Zhao JX, Liang X, Jiang F, Xing H, Zhu M, Hou RB, Zhang YB, Lan XW, Rhoades DA, Irikura K, Fukushima Y, Somerville PG (2016b) Ground-motion prediction equations for subduction interface earthquakes in Japan using site class and simple geometric attenuation functions. Bull Seismol Soc Am 106(4):1518–1534
Zhao JX, Zhou SL, Zhou J, Zhao C, Zhang H, Zhang YB, Gao PJ, Lan XW, Rhoades D, Fukushima Y, Somerville PG, Irikura K (2016c) Ground-motion prediction equations for shallow crustal and upper-mantle earthquakes in Japan using site class and simple geometric attenuation functions. Bull Seismol Soc Am 106(4):1552–1569
Zolfaghari MR, Darzi A (2019) Ground-motion models for predicting vertical components of PGA, PGV and 5%-damped spectral acceleration (0.01–10 s) in Iran. Bull Earthq Eng 17(7):3615–3635