

Pilz, M., Cotton, F., Zhu, C. (2022): How much are sites affected by 2-D and 3-D site effects? A study based on single-station earthquake records and implications for ground motion modelling. - Geophysical Journal International, 228, 3, 1992-2004.

<https://doi.org/10.1093/gji/ggab454>

How much are sites affected by 2-D and 3-D site effects? A study based on single-station earthquake records and implications for ground motion modelling

Marco Pilz¹,¹ Fabrice Cotton^{1,2} and Chuanbin Zhu¹

¹Helmholtz Centre Potsdam – German Research Centre for Geosciences GFZ, Telegrafenberg, 14467 Potsdam, Germany. E-mail: pilz@gfz-potsdam.de

²University of Potsdam, Institute for Geosciences, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Accepted 2021 October 29. Received 2021 October 28; in original form 2021 May 22

SUMMARY

1-D site response analysis dominates earthquake engineering practice, while local 2-D/3-D models are often required at sites where the site response is complex. For such sites, the 1-D representation of the soil column can account neither for topographic effects or dipping layers nor for locally generated horizontally propagating surface waves. It then remains a crucial task to identify whether the site response can be modelled sufficiently precisely by 1-D analysis. In this study we develop a method to classify sites according to their 1-D or 2-D/3-D nature. This classification scheme is based on the analysis of surface earthquake recordings and the evaluation of the variability and similarity of the horizontal Fourier spectra. The taxonomy is focused on capturing significant directional dependencies and interevent variabilities indicating a more probable 2-D/3-D structure around the site causing the ground motion to be more variable. While no significant correlation of the 1-D/3-D site index with environmental parameters and site proxies seems to exist, a reduction in the within-site (single-station) variability is found. The reduction is largest (up to 20 per cent) for purely 1-D sites. Although the taxonomy system is developed using surface stations of the KiK-net network in Japan as considerable additional information is available, it can also be applied to any (non-downhole array) site.

Key words: Earthquake ground motions; Earthquake hazards; Site effects; Wave propagation.

1. INTRODUCTION

When a site undergoes seismic shaking, local site effects have the consequences in terms of modifications in amplitude, frequency content and duration of ground-motion. These phenomena can be rather peculiar and they vary from site to site depending on the local geologic and topographic site conditions. The importance of such site response is widely acknowledged. Hence, for a quantitative evaluation of seismic hazard it is important to assess such factors as precisely as possible.

The thicknesses of the layers of the sedimentary cover and their *S*-wave velocities are two of the most commonly used parameters to determine seismic site response. A common approach to this end is to assess the local subsurface structure from available geological data sets, boring explorations, and/or other non-invasive seismic methodologies of underground explorations. In turn, various ground response simulation methods are then used to predict the level of modification of ground motion at a target site. Herein, it is generally assumed that *P* and *S* waves are often fully decoupled, meaning that the vertical component is comprised of *P* waves only while on the

horizontal components only *S* waves exists. The site response is then considered to be dominated by vertically propagating, horizontally polarized *SH* waves through a laterally homogenous medium, allowing to capture some aspects of local effects like the influence of impedance contrasts, 1-D resonance effects, and soil non-linearity (for equivalent linear and non-linear analyses).

Engineering approaches often overlook the effects of lateral variations, meaning that when the hypothesis of a laterally infinite body with horizontal layering vanishes, wave propagation cannot be modelled correctly as 1-D. On the other hand, empirical site response models used in ground-motion analysis only partially account for non-1-D sites since they are based on data sets containing all sorts of sites and wave propagation effects. These lateral heterogeneities, however, can have a significant impact on the level of ground-motion. Multiple empirical studies have illustrated that 1-D site-response approaches, despite their broad usage in practice, often fail to be accurate (e.g. Thompson *et al.* 2012; Kaklamanos *et al.* 2013; Kim & Hashash 2013; Zalachoris & Rathje 2015; Kaklamanos & Bradley 2018; Li *et al.* 2018; Pilz & Cotton 2019; Tao & Rathje 2020; Zhu *et al.* 2020) and more sophisticated approaches

are needed to capture 2-D/3-D site effects (i.e. topographic effects, slopes along with impedance contrasts, the geometry of the soft soil layers, basin-edge induced surface waves, 2-D basin resonance effects). While many different methods with varying degrees of complexity have been used for empirically identifying such complex site effects, these approaches generally rely on the availability of recordings at an adequately nearby reference site and/or detailed velocity models at the site of interest. The application of complex and expensive 2-D/3-D analysis, however, will not be necessary if methods exist to confirm the 1-D nature of a site. Such *a priori* evaluation of the 1-D nature of sites has been recently been identified as an important research need in engineering seismology (Stewart & Afshari 2021).

For engineering applications, for which simplicity is a virtue, our motivation is then to assess the validity of the 1-D assumption based on single-station surface recordings only. In other words, our goal is then to develop an 1-D/3-D index which will be based on surface measurements only. These investigations need high-quality data on site and environmental conditions. KiK-net (Kiban Kyoshin network) strong-motion seismographs in Japan provide an excellent basis also due to the large number of available strong-motion seismographs (Okada *et al.* 2004; Aoi *et al.* 2011; National Research Institute for Earth Science & Disaster Resilience 2019). Hereunto, we first describe a conceptual framework focusing on the azimuthal variability and similarity of the horizontal Fourier spectra, and then we investigate the ability to detect 2-D/3-D effects based on proxy parameters. Results are compared with previous findings on 2-D/3-D effects at Japanese network sites. Finally, we will quantify the impact of 2-D/3-D effects on ground-motion modelling.

2. SINGLE-STATION PARAMETRIZATION OF 2-D/3-D SITE EFFECTS

When looking at single-station recordings in complex environments, it has often been observed that the spectra of the two horizontal components are different in shape and amplitude though with different impacts on these two measures. Uebayashi (2003) and Uebayashi *et al.* (2012) have shown that the variation of the peak amplitude of the single-station horizontal-to-vertical (H/V) is generally more sensitive to any irregularity in the subsurface structure than the actual peak frequency itself. On the contrary, other studies indicate that the peak frequency for complex site models tends to be higher than the fundamental resonance frequency based on a 1-D velocity model while the amplitude is less affected (e.g. Steimen *et al.* 2003; Roten *et al.* 2006; Barnaba *et al.* 2010). Using a modified Vicis-symmetric parameter, Matsushima *et al.* (2017) studied the different shapes of single-station H/V curves from seismic noise at Onahama Port, Japan, attributing their variability to 2-D heterogeneities in the subsurface. Moreover, even when 2-D resonances occur, generally on deep and narrow valleys, the single-station approach can still be sufficient to discern between the 1-D versus 2-D resonance behaviour of a site (Sgattoni & Castellaro 2020).

The visibility of such 2-D/3-D effects on the horizontal component is frequency-dependent: the maximum amplification is generally expected to occur for frequencies at which the incident wavelength is comparable to the length and/or width of the morphological relief (corresponding to the longitudinal or transverse profile in a bi-dimensional resonance case). Since 2-D/3-D effects generally involve different kinds of seismic waves (e.g. diffracted body and

surface waves in areas characterized by strong lateral variations of the thickness of the sedimentary, Guillier *et al.* 2006) and since these effects might show up during different times of the wave window, a simple time-window scheme will hamper a proper identification. Therefore, we will analyse these effects in the spectral domain. However, as there is no optimal single distance metric that can be used for all types of sites for quantifying differences in the horizontal spectral components, complementary procedures are required accounting for differences both in spectral shape and amplitude (an overview on various distance and similarity measures is given, for example, by Cha 2007; Deza & Deza 2009; Crooks 2018).

For measuring the amplitude difference, that is measuring how far apart are the two horizontal components, we rely on the normalized divergence distance d (Sjøvold 1977)

$$d = \frac{1}{N} \sum_{f_1}^{f_2} \frac{[NS(f) - EW(f)]^2}{[NS(f) + EW(f)]^2}. \quad (1)$$

Herein, the squared difference at each frequency value f is used for the calculation of the total distance, that is this definition follows the L_2 norm and is based on the Euclidean distance between the horizontal components. Compared to other Euclidean distance measures, for peaked data (i.e. amplification peaks in the horizontal spectra), d outperforms other groups of measures in terms of its discriminating power (e.g. Huet & Hancock 1996). In eq. (1), f_1 and f_2 define the frequency range over which the comparison is carried out using N spectral values. NS and EW describe the frequency-dependent horizontal Fourier amplitude spectra (FAS). Each term in the summation is squared, so the relationship directly assesses any differences in amplitude. From a theoretical point of view, d takes values between zero and one, being zero if (and only if) the two components are identical.

Complementarily, we select the Jensen Shannon difference J from the Shannon entropy family to compare the shapes of the two spectra,

$$J = \frac{1}{2} \left(\sum_{f_1}^{f_2} \left(NS(f) \ln \left(\frac{2NS(f)}{NS(f) + EW(f)} \right) \right) + \sum_{f_1}^{f_2} \left(EW(f) \ln \left(\frac{2EW(f)}{NS(f) + EW(f)} \right) \right) \right), \quad (2)$$

where J describes the mean relative entropy between the variation of the two horizontal spectra and their respective mean (Lin 1991). Herein, entropy refers to the degree of disorder (or randomness) in a situation. It quantifies how distinguishable the shapes of the two spectra are from each other. It is equal to zero only if the two components are indistinguishable, that is identical.

The individual Fourier amplitude spectra of each component are calculated by filtering the full-time series with a second-order Butterworth filter at cutoff frequencies of 0.1 and 30 Hz (corresponding to the sensor's high-cut frequency) before tapering it with a 5 per cent-cosine function and smoothing the resulting Fourier spectra with a Konno & Ohmachi (1998) window. The smoothing parameter b is set to 10 for ensuring the smoothing of numerical instabilities while preserving the major features of the spectra. In eqs (1) and (2), the frequency parameters f_1 and f_2 are set to 0.3 and 25 Hz and the individual frequency values are equally spaced in logarithmic scales.

For each event and each station, we quantify d and J by rotating the two horizontal axes by steps of 5° for finding the orientation that strongly relates to directional dependence, that is the direction for which the largest difference between the two horizontal components is observed. For both indices, the corresponding maximum value is chosen. As an example, Fig. 1 shows the azimuthal variation of the

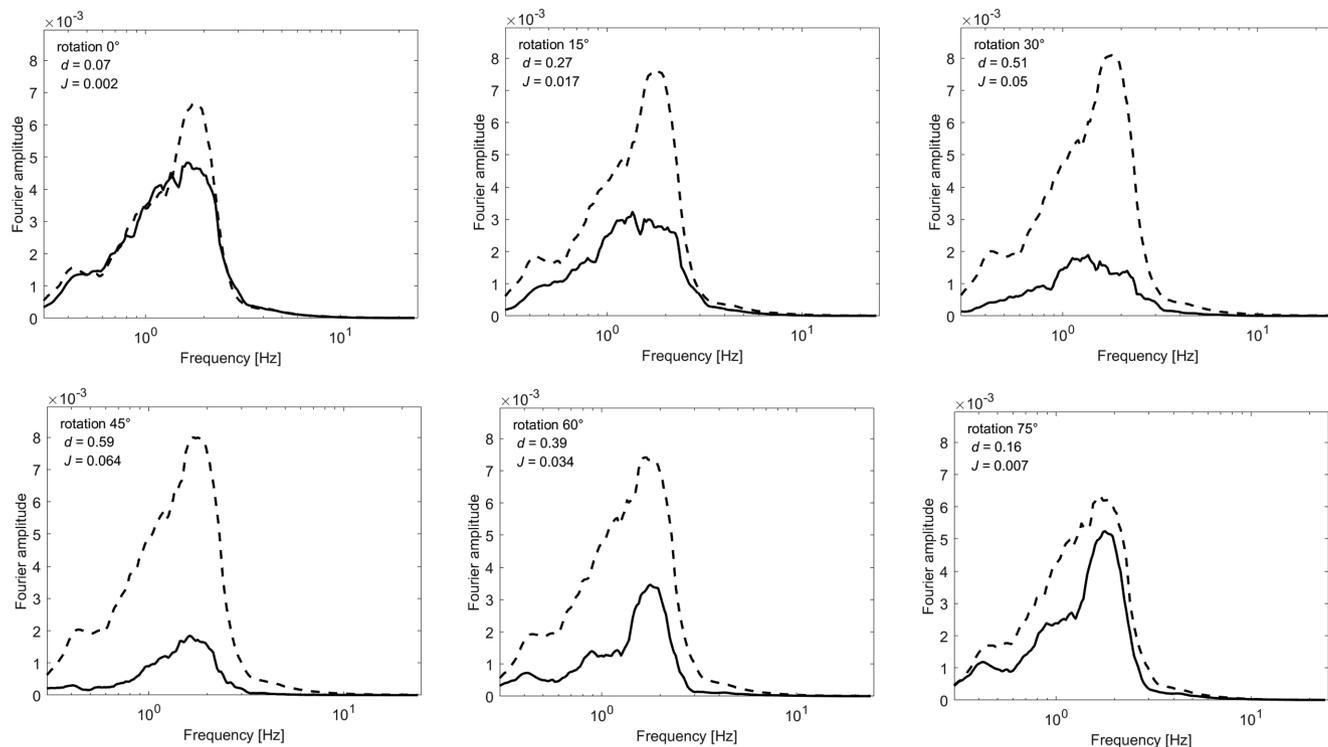


Figure 1. Azimuthal variation of the two horizontal spectra (*NS* component: bold line, *EW* component: dashed line), d and J for the 23 March 2001 event at KiK-net site HDKH06.

two horizontal spectra and the respective variation of d and J for a single event.

For each site we further investigate the interevent variability of the H/V spectral ratios. A similar measure has already been proposed by Thompson *et al.* (2012) using surface-to-borehole spectral ratios although it has recently been shown that surface-to-borehole spectral ratios have a limited usable frequency range with respect to H/V spectral ratios (Zhu *et al.* 2021a). If a large H/V interevent variability is observed at a single site, such differences could be caused by event-specific site effects linked to source-site configurations. This includes (1) the variability in source azimuth, incidence angle and wave field on the input motion in which, for example, shallow events might generate local surface waves which can contribute to azimuthal-dependent amplification in certain frequency bands (Pilz *et al.* 2018), (2) the lateral variability of the site (i.e. surface and subsurface topography and (3) temporally variable nature of soil properties due to soil non-linearity (Régnier *et al.* 2018) and/or climatic factors (Alshembari *et al.* 2020). With various degrees all these effects will negatively influence the ability for 1-D ground-motion modelling at this site. For quantifying the variability of the H/V spectral ratios at a single site, in a first step, the H/V mean and its natural log standard deviation s are calculated from all recordings for 128 discrete frequency values in a frequency range between 0.3 and 25 Hz. The corresponding interevent variability, denoted as σ_{HV} , is then taken as the median across the 128 frequency values in the studied frequency range, that is $\sigma_{\text{HV}} = \tilde{s}_f$.

3. DATA SET

In this study, earthquake recordings from the KiK-net strong-motion network are analysed. The KiK-net stations are each equipped with a pair of surface and downhole sensors while in the present analysis only surface data will be used. The reason is that for KiK-net data

a large number of previous studies are available which will be utilized for validation in a later stage. The ground-motions used in this work are a subset of the motions compiled and processed by Bahrapouri *et al.* (2019) collecting the accelerometric data recorded between 1996 and 2017 at all KiK-net recording sites observed for earthquakes with magnitudes greater than 3.0. For magnitudes larger than 3.5, the values are taken from the F-net catalogue calculated by the Japan Meteorological Agency (JMA). Dawood *et al.* (2016) concluded that the hypocentral location and magnitude from the F-net catalogue can be considered more reliable than information provided by KiK-net.

In addition, we choose only earthquakes with at least three usable recordings for which, in a first step, the peak ground acceleration (PGA) has to be less than 0.1 g for keeping the analysis in the linear range. In the end, the number of usable recordings corresponds to 103 258 motions from 5662 events. The locations of the events and a corresponding magnitude versus hypocentral distance distribution are shown in Fig. 2.

4. OBSERVED SITE PARAMETERS

Fig. 3 illustrates d , J , both of them along with their azimuthal variation, and σ_{HV} exemplarily for two KiK-net sites.

FKSH11 can serve as a typical example of a 1-D site. Besides a narrower confidence interval, smaller and azimuthally less variable values are obtained for both similarity measures although there is a limited number of events for which larger variations in shape are found. Such individual variations might be caused by event-specific phenomena (e.g. transient signals). For HDKH06, located in the coastal plains of southern Hokkaido, both parameters are directionally dependent for individual events, resulting further in a large standard deviation over the entire frequency range, that is a

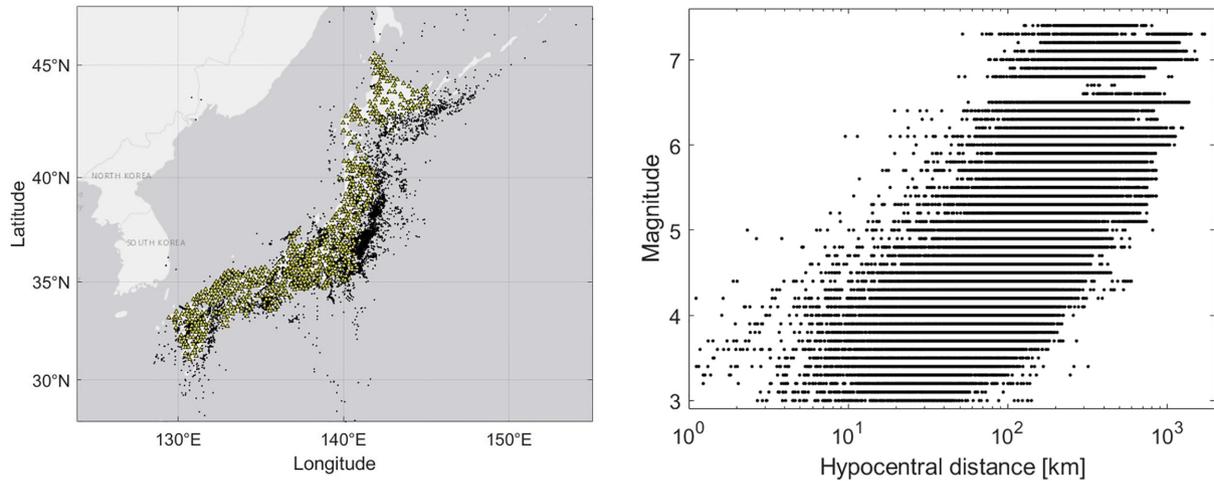


Figure 2. Left-hand panel: spatial distribution of earthquakes (black dots) and KiK-net stations (yellow triangles) used in this study. Right-hand panel: magnitude-distance distribution of the selected earthquake recordings.

large value of σ_{HV} , which, in summary, is pointing at some 2-D/3-D effects at this site.

Once the individual parameters d , J and σ_{HV} are available for all KiK-net sites, they are standardized separately. For every parameter, we subtract the respective parameter mean and divide the results by the standard deviation to shift each distribution to have a mean of zero and a standard deviation of one. The three standardized parameters are summed before rescaling the corresponding sum Σ to $[0, 1]$, that is

$$I = \frac{\Sigma}{\max(\Sigma) - \min(\Sigma)}. \quad (3)$$

The previous steps do not change the distribution but they will only affect the range of the data. The final index I assumes a value in the range 0 to 1 with 0 indicating a fully 1-D behaviour while an index of 1 represents the certainty of significant horizontal variability at this site.

For the analysed KiK-net sites, Fig. 4 plots the distribution of the three site parameters and the final value of I . Details on the site-specific parametrization for all sites are provided in Table S1.

d and σ approximately follow a normal distribution while for J and correspondingly for I the distributions are slightly asymmetric with a tendency towards smaller values. While there is only a limited number of sites with extreme values for I (54 sites with $I \leq 0.2$, 16 sites with $I \geq 0.8$), the vast majority of sites (almost 90 per cent) cannot be characterized as purely 1-D or purely 2-D/3-D. However, for grouping sites as 1-D or 2-D/3-D sites, we refrain from using strict threshold values since this would mean that such classification would be depending on our choices. On the contrary, the use of a probability index offers the possibility to set user- and application-specific confidence levels.

5. EFFECTIVENESS OF THE PROPOSED 1-D/3-D TAXONOMY

Many recent studies have already evaluated the accuracy of 1-D site response modelling on the basis of comprehensive data sets available for the KiK-net strong-motion network. Pilz & Cotton (2019) systematically assessed 2-D/3-D site effects relying on the closeness of the empirical amplification function to the theoretical

transfer function for vertically propagating SH waves. In the frequency range between the site's fundamental frequency and 25 Hz, Spearman's correlation coefficient s was calculated between the two functions. Sites with coefficients close to 1, indicating a good match between the two functions, were considered as 1-D sites while a low value of s indicates a site with prominent 2-D/3-D effects. A threshold of $s = 0.6$ was set to distinguish 1-D sites from sites with 2-D/3-D effects. Fig. 5 depicts I from eq. (3) against the previously proposed Spearman correlation coefficient s for the 354 KiK-net sites. For 354 of 689 KiK-net sites a sufficient match between the measured and the calculated fundamental resonance frequency has been observed and only these sites were used for further analysis by Pilz & Cotton (2019).

As can be seen in Fig. 5, sites with small values of I tend to have larger values of s (i.e. both values indicating a 1-D behaviour) and vice versa. The Pearson correlation coefficient r , measuring the strength of the relationship between both parameters I and s ($-1 \leq r \leq 1$), takes a value of -0.67 . This value is statistically significant with a level of two sigma when testing the null hypothesis that there is not a significant linear correlation between I and s , that is $r = 0$. Both I and s aim at identifying 1-D and 2-D/3-D sites but both variables rely on different data sets (While I is based on surface recordings only, s is calculated from a spectral inversion with respect to reference rock conditions and requires KiK-net velocity profiles). Therefore, Fig. 5 might indicate that surface recordings can be sufficient for identifying 2-D/3-D effects.

Tao & Rathje (2020) recently presented a taxonomy to assess the suitability of 1-D analysis considering the justification of peaks in the H/V spectral ratio and the presence of true and pseudo resonances at a site. For minimizing any parametric uncertainty in their modelling approach, Tao & Rathje (2020) re-inverted the available velocity profiles at some KiK-net sites. For 28 studied KiK-net sites, 21 are classified by Tao & Rathje (2020) to be modelled well by 1-D analysis. Based on this study, the corresponding list of 1-D sites takes a mean of $I = 0.40 \pm 0.16$ and 18 of these 21 sites are characterized by I of less than 0.5. Also for the seven non-1-D sites of Tao & Rathje (2020), a consistent result is found: I takes much higher values with a mean of $I = 0.52 \pm 0.08$.

Thompson *et al.* (2012) have presented another thorough approach for developing a taxonomy to identify sites that can be

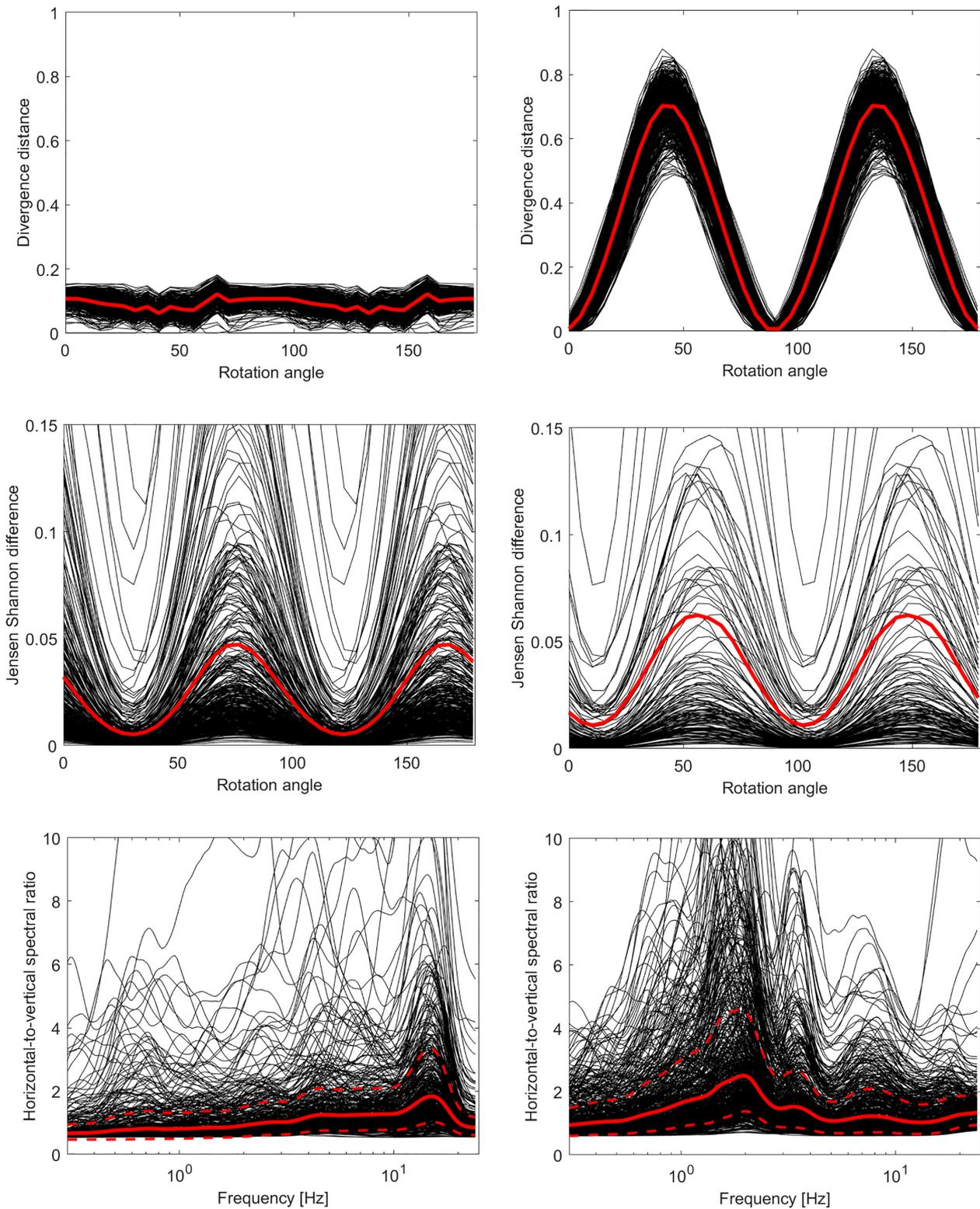


Figure 3. Divergence distance (top row), Jensen–Shannon difference (middle row) and earthquake H/V spectral ratios (bottom row) at KiK-net sites FKSH11 (left-hand panel, 286 records) and HDKH06 (right-hand panel, 176 records). Top and middle row presents d and J against rotation angle of the two horizontal axes. Black lines indicate individual events while the red lines show the mean. For the H/V spectral ratios the dashed line represents one standard deviation.

modelled well by 1-D analysis. The taxonomy is based on the correlation coefficient between the theoretical and the empirical transfer function, considering further the variability of the latter. Based on their study, 16 of 100 sites have been classified as LG (Low

interevent variability and Good match between the empirical and theoretical transfer function), that is sites ideal for validation of 1-D constitutive models. Also in our study, most of these sites show a very low value of I while only for a few sites (IBRH10, IBRH17,

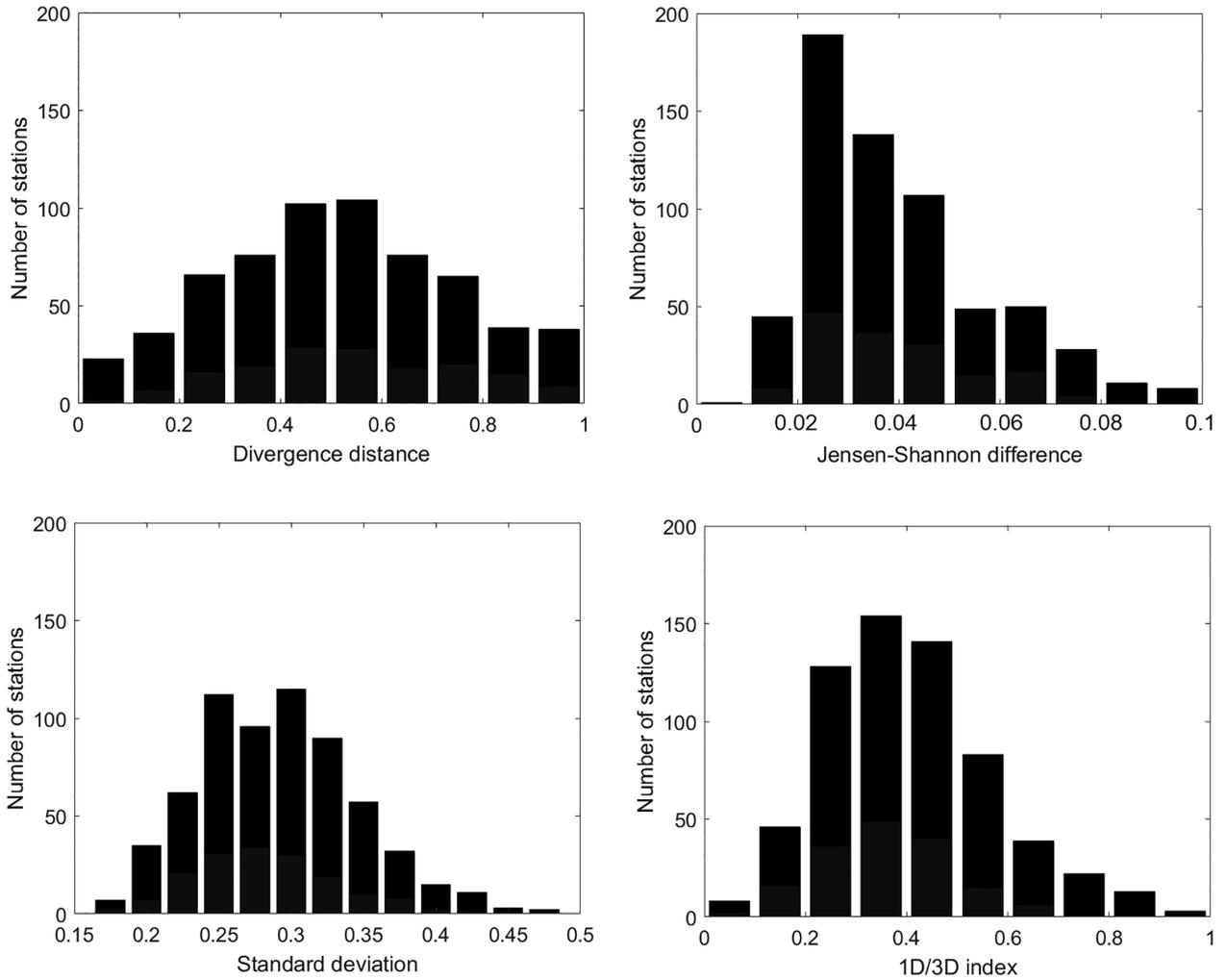


Figure 4. Quantitative distribution of the number of stations for d (top left-hand panel), J (top right-hand panel), σ (bottom left-hand panel) and I (bottom right-hand panel).

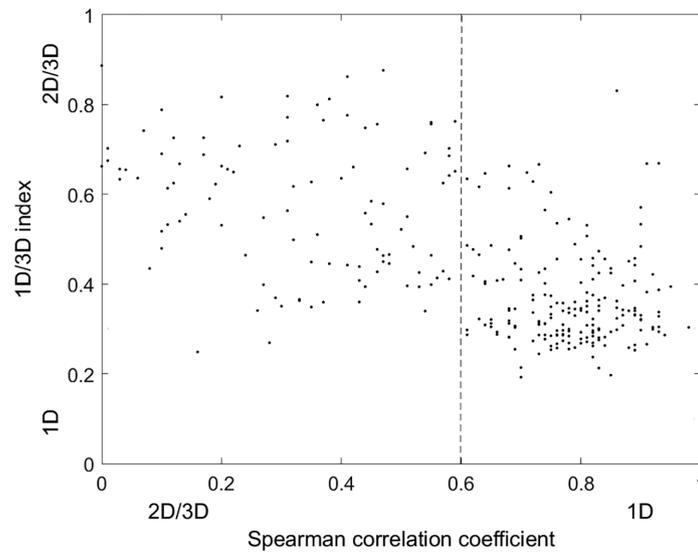


Figure 5. I against Spearman correlation coefficient for 354 KiK-net sites of Pilz & Cotton (2019). The dashed line indicates the threshold used by Pilz & Cotton (2019) for distinguishing 1-D and 2-D/3-D sites.

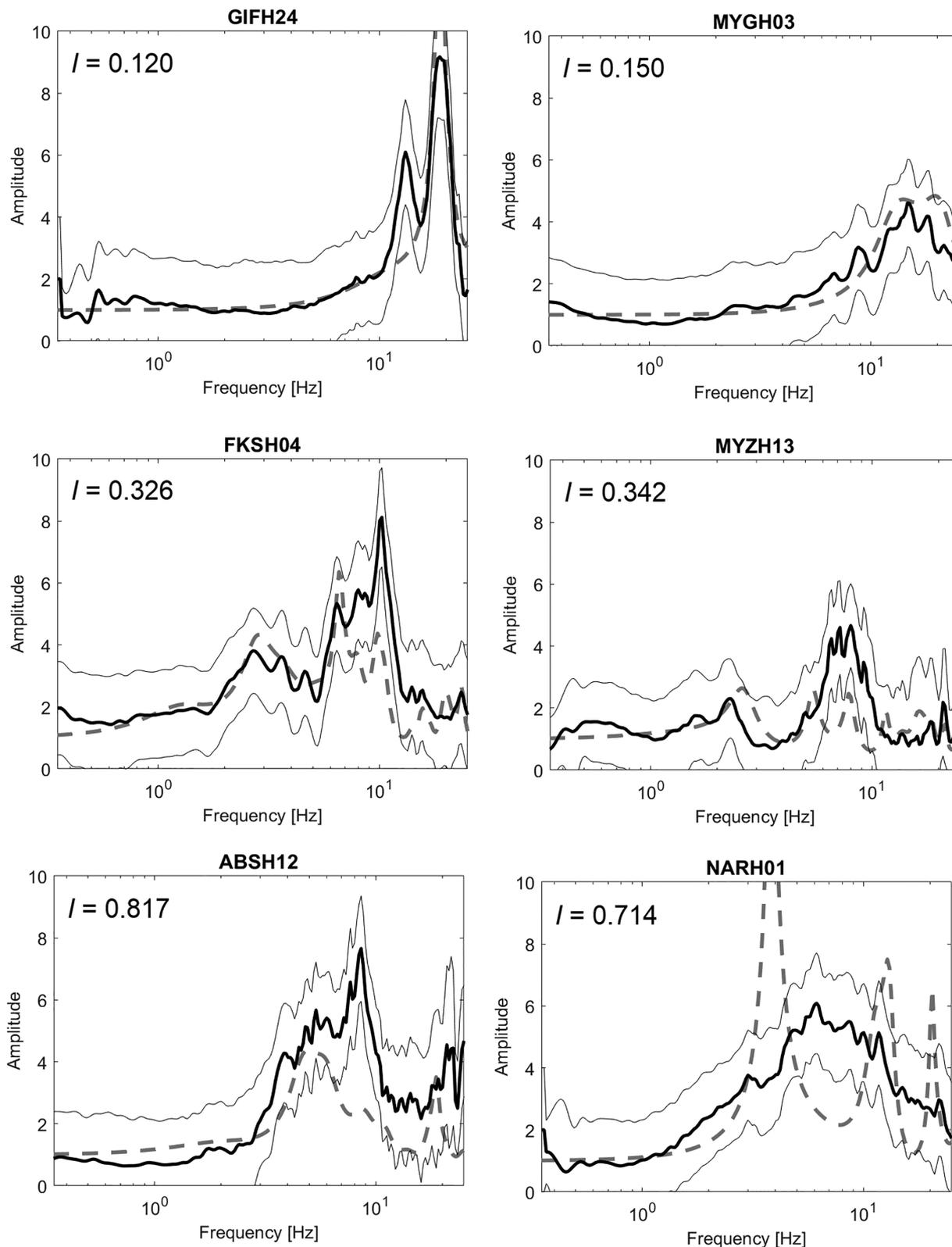


Figure 6. Theoretical 1-D SH transfer function (grey dotted lines) and surface-to-borehole spectral ratio plus/minus one standard deviation corrected for the downgoing wave effect (black lines). *I* represents the 1-D/3-D index. GIFH24 and MYGH03 (top line) are representative of 1-D sites. FKSH04 and MYZH13 (middle line) represent sites in the transition range between 1-D and 2-D/3-D sites while ABSH12 and NARH01 are characteristic for 2-D/3-D sites (bottom line).

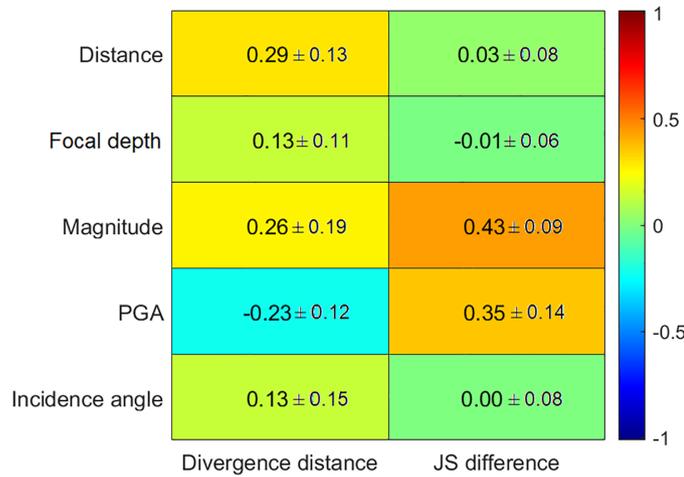


Figure 7. Pearson correlation coefficient r of d and J with hypocentral distance, focal depth, moment magnitude, site-specific PGA and incidence angle averaged over all analysed stations.

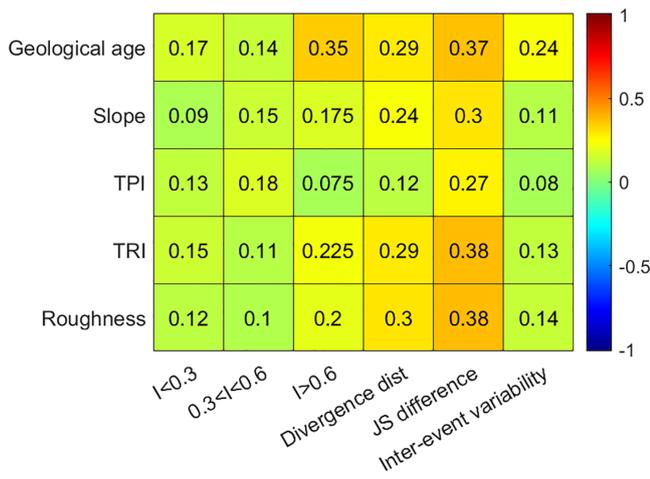


Figure 8. Pearson correlation coefficient r between various ranges of I and individual site parameters d , J and σ with site condition proxies from Zhu *et al.* (2021b). TPI represents the topographic position index and TRI represents the terrain ruggedness index.

IWTH08, KRSH07, KRSH10, NMRH04 and TKCH08) significant differences occur, meaning that for these sites I takes values larger or much larger than 0.5. Here the difference is mainly caused by a significantly higher interevent variability for the corresponding sites in this study as we assign rather large values for σ_{HV} for the corresponding sites with a mean σ_{HV} of 0.29 ± 0.07 (for the distribution of σ_{HV} see Fig. 4, for individual values of σ_{HV} see Table S1). A visual inspection, however, indicates the validity of a large σ_{HV} for the modelled sites.

A large number (53 of 100 studied sites) have been characterized as LP sites (Low interevent variability and Poor level of match between the empirical and theoretical transfer function) by Thompson *et al.* (2012). For these sites, however, I mostly takes values of less than 0.5, that is tending towards 1-D, with a mean of 0.42 ± 0.15 . As noted by Tao & Rathje (2019), correlation coefficients are particularly sensitive to the modelled level of damping. In turn, this might have caused a site not being considered as 1-D due to a poor fit between the empirical and theoretical transfer function. This means that in this situation the underlying evaluation of the goodness-of-fit

might not only be affected by the modelling uncertainty (e.g. 2-D/3-D effects, complex incident wave fields) but also by the parametric uncertainty of the parameters entering in the ground-motion model.

By comparing the theoretical transfer function based on linear 1-D SH modelling with the site-specific empirical transfer function estimated from surface-to-downhole spectral ratios, Fig. 6 provides a comparison of both functions and I for representative KiK-net sites. A comparison for all sites is provided as an electronic supplement to this article. For each site, surface-to-downhole spectral ratios are calculated by dividing the horizontal Fourier spectrum of ground-motion at the surface by that at the downhole. Theoretical transfer functions are computed in the linear domain as the ratio of the amplitude of harmonic motion on the ground surface to that of the total wavefield at the depth of the downhole sensor.

GIFH24 and MYGH03 are characteristic of 1-D sites for which a good match in shape and amplitude is found. For sites in the transition range between the 1-D and the 2-D/3-D regime (with intermediate values for I) generally show less similarity between the two functions. For most sites with large values for I , a significant mismatch between the theoretical and empirical transfer functions can be found. Such mismatch is generally seen at sites having (1) a considerable level of amplification for frequencies higher than the site's fundamental frequency (ABSH12 in Fig. 6) since 2-D/3-D effects are most significant in this frequency band or (2) a significant discrepancy between the measured and the calculated fundamental resonance frequency (NARH01).

While Fig. 6 might indicate that I can directly be derived by comparing the empirical and theoretical transfer functions, there is no direct link between the two functions and I . Zhu *et al.* (2020) recently studied the variation of the theoretical transfer function associated with uncertainties in velocity profiles concluding that 1-D SH-based ground response analysis exhibits poor performance in properly predicting site amplifications at frequencies larger than ~ 3 Hz achieving a sufficient agreement only at less than one-third of the examined sites. Moreover, at most sites we also cannot rule out the possibility of having deeper velocity contrasts below the downhole sensor which might alter the site's fundamental frequency but also influence the empirical site responses in the entire studied frequency range.

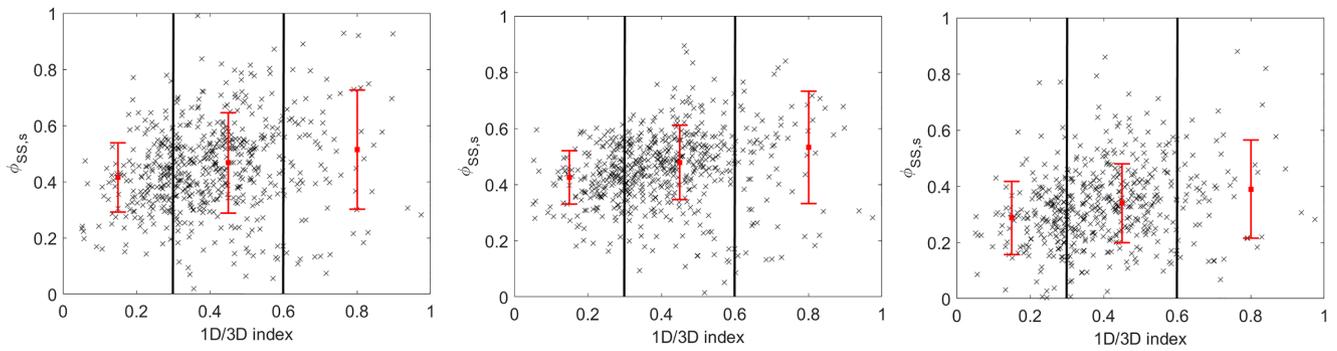


Figure 9. Dependency of the site-specific event-and-site corrected variability ($\phi_{SS,s}$) on I for periods of $T = 0.02$ s (left-hand panel), 0.2 s (middle panel) and 2 s (right-hand panel). The crosses indicate $\phi_{SS,s}$ estimated at each station. The black lines indicate ranges of I with the range-specific mean and standard deviation given in red.

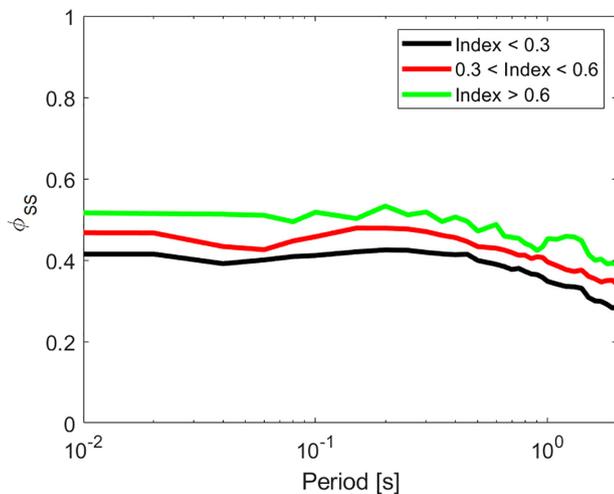


Figure 10. Remaining standard deviation after removing the site-specific site response prediction for different ranges of I .

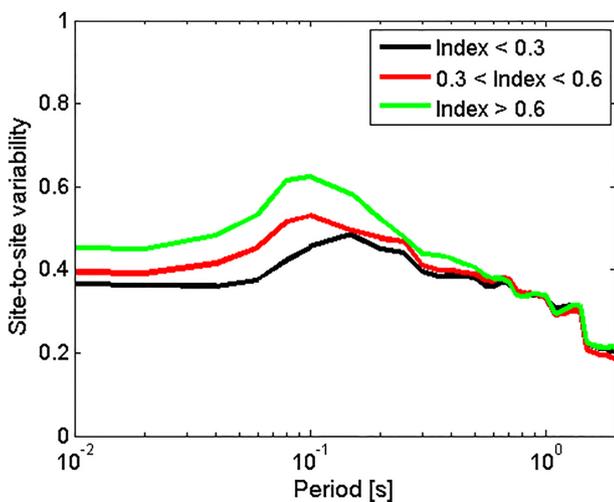


Figure 11. Site-to-site variability ϕ_{S2S} for different ranges of I .

6. CORRELATION OF THE 1-D/3-D INDEX WITH EVENT AND SITE PROXIES

From the variability of the individual parameters, as seen in Fig. 3, the question arises of how repeatable the apparent 2-D/3-D effects are with respect to different seismic sources and different path effects. In principle, several factors like source mechanism, geometrical spreading and attenuation contribute to the observed spectra at a station. From a theoretical point of view, we would expect a range-dependent pattern in the waveform complexity caused by varying focal depth and radiation pattern as the travel path increases from the range of simple direct crustal phases (within a few tens of kilometres) to reverberate multipathed P phases (beyond 150 km). For particular and well-studied sites, Riepl *et al.* (1998), Raptakis *et al.* (2000) and Maufroy *et al.* (2016) have already indicated a high sensitivity of ground-motion to focal depth and distance of local events, especially a high sensitivity of the occurrence of 2-D/3-D site effects to backazimuth. The path duration (i.e. the lengthening of the ground-motion duration with distance), on the other hand, will contribute to the dependency of the frequency content of ground-motion decay but has been found to remain a second-order effect (Dujardin *et al.* 2016). This implies if different events have a different frequency content, their relative I might be different. If this were the case, we would expect the site-specific parameters to be correlated with event-specific ones.

Fig. 7 plots the Pearson correlation coefficient r of d and J with hypocentral distance, focal depth, moment magnitude taken from the NIED catalogue, PGA and incidence angle. Again the values are statistically significant at a confidence level of 2σ .

For assessing the incidence angle, in a first step we use the S -wave velocity model of Nakajima *et al.* (2001) for determining the ray parameter p between each earthquake and the downhole sensor at each station. Because of the large depth of all studied events with respect to the depth of the downhole sensors, the variations in downhole sensor depths will not have a significant influence on p . Given the simplified velocity models available at each KiK-net site, the angle of incidence θ at the surface sensor for the wave propagating from the downhole station is then calculated by

$$\cos \theta = \sqrt{1 - v_s^2 p^2} \quad (4)$$

in which v_s is the average S -wave velocity between the downhole

and the surface sensor taken from the logging data at the KiK-net sites. Although the velocity model of Nakajima *et al.* (2001) is specific for northeastern Japan, any regional velocity perturbation will have negligible influence on θ .

As shown in the correlation matrix, we observe almost no correlation between various parameters. Only between magnitude and PGA on the one hand and J on the other hand a weak correlation is found. One reason might be that the velocity pulse period is a function of the event's moment magnitude, meaning that large magnitude events go along with large pulse periods (e.g. Bray & Rodriguez-Marek 2004). In this period range, directivity effects, that is focusing of energy along the fault in the direction of rupture, have frequently been observed, in turn causing some variation in the two horizontal directions. Similar considerations also apply for PGA. While in the high-frequency range, there is only a weak dependence on the radiation pattern, this latter effect becomes more significant for intermediate and low frequencies (e.g. Vidale 1989) and might be mapped as a difference in the horizontal spectra here. We emphasize, however, that the correlation is rather weak.

Only for a very limited number of stations, a strong correlation ($r > 0.7$) between either d with magnitude (NGNH27, SZOH35), J with magnitude (OSKH04) or J with PGA (NGNH35) is observed but none of these few sites is characterized by a large value of I . Since we are averaging over many events in a wide range of directions and distances (254 median events per site), it is therefore expected that the sometimes observed dependence of the 3-D site response on the event parameters is limited.

Since the presented approach is fully empirical, one cannot specify the primary cause of the site effects under consideration. It might anyhow be advisable to identify site condition proxies and their derived topographic and/or environmental attributes for assessing if a site is more susceptible for 2-D/3-D effects. To this regard, Zhu *et al.* (2021b) have recently presented an open-source and systematic site database for 1742 K-NET and KiK-net sites in Japan. The data base contains various topographic and geological parameters and proxies inferred from regional models or maps which have been derived for all network sites in a systematic way.

We apply the following site conditions proxies: (1) geological age [i.e. Holocene, Pleistocene, Quaternary (volcanic rocks), Tertiary and pre-Tertiary], (2) slope of topography from a 15'' digital terrain model, (3) topographic position index (TRI) representing the difference between the elevation of the central pixel and the mean of the surrounding pixels, (4) terrain ruggedness index (TRI) quantifying the mean of the difference in elevation between the central pixel and each of the surrounding pixels and (5) roughness which is defined as the largest difference in elevation between the central pixel the surrounding ones. Fig. 8 plots the Pearson correlation matrix between site condition proxies and site parameters.

Among the parameters, no significant linear correlations between any environmental and site parameters are found as r is always smaller than 0.4. This indicates that the occurrence of the mapped effects is not limited to specific environmental conditions although there is, as expected, some weak linear correlation for the occurrence of 2-D/3-D effects at steeper slopes in rougher terrain. Since the results represent an average over different environmental conditions, it might be unlikely to infer the site behaviour from the environmental parameters alone.

Zhu *et al.* (2020) have already assessed the effectiveness of 1-D modelling for 90 KiK-net sites having no significant velocity contrasts below their downhole sensors, meaning that the surface-to-downhole spectral ratios can be taken as their empirical transfer functions. They assessed if a linear correlation between the latter and

1-D theoretical transfer functions exists with respect to various site proxies. Complementarily to Fig. 8, their results indicate a limited efficacy of 1-D modelling in reproducing observed amplifications at steeper and stiffer sites. Since most KiK-net sites are located on thin sediments or weathered rocks at the margin of basins, Zhu *et al.* (2020) conclude that complex geological site conditions in line with the violation of the 1-D assumption might be responsible for the low success rate.

7. VARIABILITY OF GROUND-MOTION

Due to the increasing amount of strong ground-motion data, empirical site amplification functions can be obtained from ground-motion prediction equation (GMPE) residuals for use in site-specific Probabilistic Seismic Hazard Assessment (PSHA). It is most often considered—at least implicitly—that the site response can be decoupled from source and path effects. The record-specific total residuals Δ can be decomposed into between-event (δB_e) and within-event (δW_{es}) components and the latter can be decomposed in the single-site approach (Atkinson 2006; Rodriguez-Marek *et al.* 2013) into between-site residuals ($\delta S2S_s$) and event-and-site corrected residuals (δWS_{es}),

$$\Delta = \delta B_e + \delta S2S_s + \delta WS_{es}. \quad (5)$$

Herein, the site terms $\delta S2S_s$ represent the systematic average deviation of the observed amplification at a specific site from the median amplification predicted by the model based on a simple site classification. The total variability σ of the data set with respect to the GMPE is given by

$$\sigma = \sqrt{\tau^2 + \phi_{S2S}^2 + \phi_{SS}^2}. \quad (6)$$

The between-event standard deviation is represented by τ . The standard deviation of the between-site residuals is the site-to-site variability denoted by ϕ_{S2S} . ϕ_{SS} is the standard deviation of the event-and-site corrected residual (δWS_{es}) at a given station. Conceptually, ϕ_{SS} is the record-to-record aleatory variability associated with propagation and other sources of variability across all source—site paths in the data set, that is it includes all components of event-to-event and path-to-path variability for a given site. ϕ_{SS} is identical to the single-station standard deviation of Rodriguez-Marek *et al.* (2011).

Using a mixed-effects regression, Kotha *et al.* (2018) recently derived site-response adjustment factors by analysing the spectral properties of $\delta S2S_s$ with respect to real reference site conditions for KiK-net data. The underlying ground-motion model does not include any site response component in its fixed effects, unlike the common practice of including a function of travel-time based average S -wave velocity in the uppermost 30 m (v_{S30}), meaning that the $\delta S2S_s$ will absorb all average site response. A large number of smaller events ($M_w > 3.5$) and long-distance recordings ($0 \text{ km} < R_{JB} < 600 \text{ km}$) have been included to ensure that the site terms are constrained well and capture the full linear site response while not being biased by non-linear response on soft soils (Kotha *et al.* 2018).

We first focus on the dependency of the standard deviation of δWS_{es} for each station, the site-specific single-station standard deviation ($\phi_{SS,s}$), on I (Fig. 9). While a slightly decreasing trend of $\phi_{SS,s}$ with increasing period is found, there is a tendency of increasing $\phi_{SS,s}$ with increasing I for all considered periods. This means that for sites that tend to be more affected by 2-D/3-D effects, the

site-specific single-station standard deviation is larger with respect to 1-D sites.

The lower within-site variability of ground-motions for 1-D sites is confirmed by Fig. 10 presenting the remaining standard deviation after removing the site-specific site response prediction.

As expected from the dependency of $\phi_{SS,s}$ on I shown in Fig. 9, the single-station variability ϕ_{SS} is lower for 1-D sites: ϕ_{SS} takes values of smaller than or equal 0.43 (in logarithmic scale) for all periods. These values are lower than the corresponding estimates retrieved from literature for similar environments (e.g. Lin *et al.* 2011; Rodriguez-Marek *et al.* 2011; Baltay *et al.* 2017; Subhadra & Mai 2021 and others). For 2-D/3-D sites, on the other hand, ϕ_{SS} remains larger. This trend is almost independent on the period though it is slightly larger at intermediate periods. Both Figs 8 and 9 then indicate that within-station ground-motions may be more variable at 2-D/3-D sites compared to 1-D sites. The physical factors explaining this larger within-station variability (e.g. influence of the incidence angle) remain to be explored.

The variability of the δS_{2S} among the stations, that is the site-to-site variability ϕ_{S2S} , for different ranges of I is shown in Fig. 11.

The site-to-site variability is found to exhibit a notable dependence on I with perceptible frequency-dependent smaller values for ϕ_{S2S} at 1-D sites. The influence of 2-D/3-D effects is observed to become smaller and finally negligible for large periods. For shorter periods ($T < 0.7$ s), a reduced variability of ϕ_{S2S} of more than 10 per cent is found. The largest differences of up to 20 per cent for ϕ_{S2S} between 1-D and 2-D/3-D sites, however, can be observed at intermediate periods. This might be due to the fact that 2-D/3-D effects are most significant between the site's fundamental and the first higher 1-D resonance frequency (e.g. broadening of H/V resonance peaks due to dipping layers, basin-edge induced surface waves generally affect periods slightly shorter than the site's fundamental period). A generic high site-to-site variability is found at a response period around $T = 0.1$ s but this effect can mainly be attributed to the transformation from Fourier spectra (frequency domain) to response spectra via convolution with a single-degree-of-freedom oscillator transfer function (discussed in Stafford *et al.* 2017 and Kotha *et al.* 2017).

8. CONCLUSIONS

In this study, we have proposed a probabilistic indexing scheme to identify sites which are more likely to be affected by 2-D/3-D site effects indicating that more complex ground-motion modelling is required. The scheme is based on single-station earthquake recordings at surface stations only. It is benefitting from the fact that no additional site-specific information nor recordings at nearby reference sites are required. Although we have developed the method solely using strong-motion data from Japan, the approach is not regional-specific and can be applied to other regions with strong-motion networks and recordings.

Three metrics, that is the divergence distance, the Jensen-Shannon difference and the interevent variability are used to account for azimuthal variations in ground-motion spectral shape and amplitude and event-specific site effects linked to particular source-site configurations. The occurrence of 2-D/3-D effects is found to be only partially event-dependent but not identifiable through site condition proxies. Only a limited number of sites do actually qualify as purely 1-D or 2-D/3-D sites, while the majority of sites are influenced by 2-D/3-D effects with varying degrees. As the use of strict quantitative measures may cause an underestimation of the number

of sites modelled well by 1-D analysis, the site-specific probability index can allow to apply user-specific thresholds for grouping sites as 1-D or 2-D/3-D.

Our results indicate that both the within and between-site variabilities are larger for 2-D/3-D sites. In this case, the site response is influenced more significantly by the 2-D/3-D environment, such as bedrock slope, topographical irregularities and geometry of the soft soil layers. For the between-site variability, these effects seem to diminish for longer periods, in particular above the site's fundamental period. The within-site (single-station) variability is lower by up to 20 per cent for 1-D sites at almost all periods. The considerably lower values of ϕ_{SS} for sites which can be accurately be described by 1-D wave propagation indicate that the ground response at these sites is less variable from a physical point of view and can therefore be described more appropriately with standard 1-D parametrizations. The 1-D/3-D index can then facilitate our capacity to evaluate, *a priori*, if classical 1-D site response analysis for a given site may provide an accurate and sufficiently well-constrained response estimate.

ACKNOWLEDGEMENTS

We warmly thank the National Research Institute for Earth Science and Disaster Resilience (NIED), Japan, for making available the high-quality KiK-net data. Ben Edwards, Shinichi Matsushima and the editor Eiichi Fukuyama are also thanked for their constructive comments which helped to improve the manuscript.

DATA AVAILABILITY STATEMENT

The raw earthquake ground-motion records for all stations were obtained from the KiK-net database at <http://www.kyoshin.bo.sai.go.jp/> (last accessed April 2021). Site parameters for KiK-net sites were taken from the database of Zhu *et al.* (2021b) at <https://dataservices.gfz-potsdam.de/panmetaworks/showshort.php?id=0354b209-32f1-11eb-9603-497c92695674> (last accessed May 2021).

REFERENCES

- Alshembari, R., Parolai, S., Boxberger, T., Sandron, D., Pilz, M. & Sy-lacheva, N., 2020. Seasonality in site response: an example from two historical earthquakes in Kazakhstan, *Seismol. Res. Lett.*, **91**(1), 415–426.
- Aoi, S., Kunugi, T., Nakamura, H. & Fujiwara, H., 2011. Deployment of new strong motion seismographs of K-NET and KiK-net, in *Earthquake Data in Engineering Seismology*, pp. 167–186, eds Akkar, S., Gülkan, P. & van Eck, T., Springer.
- Atkinson, G. M., 2006. Single-station sigma, *Bull. seism. Soc. Am.*, **96**(2), 446–455.
- Bahrampouri, M., Rodriguez-Marek, A., Shahi, S. & Dawood, H., 2019. An updated database for ground motion parameters for KiK-net records, *Earthq. Spectra*, **37**, 505–522.
- Baltay, A. S., Hanks, T. C. & Abrahamson, N. A., 2017. Uncertainty, variability, and earthquake physics in ground-motion prediction equations, *Bull. seism. Soc. Am.*, **107**(4), 1754–1772.
- Barnaba, C., Marelllo, L., Vuan, A., Palmieri, F., Romanelli, M., Priolo, E. & Braitenberg, C., 2010. The buried shape of an alpine valley from gravity surveys, seismic and ambient noise analysis, *Geophys. J. Int.*, **180**(2), 715–733.
- Bray, J. D. & Rodriguez-Marek, A., 2004. Characterization of forward-directivity ground motions in the near-fault region, *Soil Dyn. Earthq. Eng.*, **24**(11), 815–828.
- Cha, S.H., 2007. Comprehensive survey on distance/similarity measures between probability density functions, *City*, **1**(2), 1.

- Crooks, G., 2018. On measures of entropy and information, Technical Note, Available at: threeplusone.com/on-information.pdf, last accessed 12 February 2021.
- Dawood, H.M., Rodriguez-Marek, A., Bayless, J., Goulet, C. & Thompson, E., 2016. A flatfile for the KiK-net database processed using an automated protocol, *Earthquake Spectra*, **32**(2), 1281–1302.
- Deza, M. & Deza, E., 2009. *Encyclopedia of Distances*, Springer, pp. 297–310.
- Dujardin, A., Courboux, F., Causse, M. & Traversa, P., 2016. Influence of source, path, and site effects on the magnitude dependence of ground-motion decay with distance, *Seismol. Res. Lett.*, **87**(1), 138–148.
- Guillier, B., Cornou, C., Kristek, J., Moczo, P., Bonnefoy-Claudet, S., Bard, P. Y. & Fäh, D., 2006. Simulation of seismic ambient vibrations: does the H/V provide quantitative information in 2D-3D structures, in *Third International Symposium on the Effects of Surface Geology on Seismic Motion*, Grenoble, France, Vol. 30.
- Huet, B. & Hancock, E. R., 1996. A statistical approach to hierarchical shape indexing, *IEE Colloquium on Intelligent Image Databases*, **7**, 1–5.
- Kaklamanos, J. & Bradley, B. A., 2018. Challenges in predicting seismic site response with 1D analyses: conclusions from 114 KiK-net vertical seismometer arrays, *Bull. seism. Soc. Am.*, **108**(5A), 2816–2838.
- Kaklamanos, J., Bradley, B. A., Thompson, E. M. & Baise, L. G., 2013. Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data, *Bull. seism. Soc. Am.*, **103**(3), 1733–1749.
- Kim, B. & Hashash, Y. M., 2013. Site response analysis using downhole array recordings during the March 2011 Tohoku-Oki earthquake and the effect of long-duration ground motions, *Earthq. Spectra*, **29**(S1), S37–S54.
- Konno, K. & Ohmachi, T., 1998. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. seism. Soc. Am.*, **88**(1), 228–241.
- Kotha, S. R., Bindi, D. & Cotton, F., 2017. Site-corrected magnitude- and region-dependent correlations of horizontal peak spectral amplitudes, *Earthq. Spectra*, **33**(4), 1415–1432.
- Kotha, S. R., Cotton, F. & Bindi, D., 2018. A new approach to site classification: mixed-effects ground motion prediction equation with spectral clustering of site amplification functions, *Soil Dyn. Earthq. Eng.*, **110**, 318–329.
- Li, G., Motamed, R. & Dickenson, S., 2018. Evaluation of one-dimensional multi-directional site response analyses using geotechnical downhole array data in California and Japan, *Earthq. Spectra*, **34**(1), 349–376.
- Lin, J., 1991. Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Theory*, **37**(1), 145–151.
- Lin, P. S., Chiou, B., Abrahamson, N., Walling, M., Lee, C. T. & Cheng, C. T., 2011. Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models, *Bull. seism. Soc. Am.*, **101**(5), 2281–2295.
- Matsushima, S., Kosaka, H. & Kawase, H., 2017. Directionally dependent horizontal-to-vertical spectral ratios of microtremors at Onahama, Fukushima, Japan, *Earth, Planets Space*, **69**(1), 96.
- Maufroy, E. et al., 2016. 3D numerical simulation and ground motion prediction: verification, validation and beyond—Lessons from the E2VP project, *Soil Dyn. Earthq. Eng.*, **91**, 53–71.
- Nakajima, J., Matsuzawa, T., Hasegawa, A. & Zhao, D., 2001. Three-dimensional structure of Vp, Vs, and Vp/Vs beneath northeastern Japan: implications for arc magmatism and fluids, *J. geophys. Res.*, **106**(B10), 21 843–21 857.
- National Research Institute for Earth Science and Disaster Resilience, 2019. NIED K-NET, KiK-net, National Research Institute for Earth Science and Disaster Resilience, <https://doi:10.17598/NIED.0004>, last accessed July 2021.
- Okada, Y., Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H. & Yamamoto, A., 2004. Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-net, *Earth, Planets Space*, **56**(8), xv–xxviii.
- Pilz, M. & Cotton, F., 2019. Does the one-dimensional assumption hold for site response analysis? A study of seismic site responses and implication for ground motion assessment using KiK-net strong-motion data, *Earthq. Spectra*, **35**(2), 883–905.
- Pilz, M., Parolai, S., Petrovic, B., Silacheva, N., Abakanov, T., Orunbaev, S. & Moldobekov, B., 2018. Basin-edge generated Rayleigh waves in the Almaty basin and corresponding consequences for ground motion amplification, *Geophys. J. Int.*, **213**(1), 301–316.
- Raptakis, D., Chávez-García, F. J., Makra, K. & Pitilakis, K., 2000. Site effects at Euroseistest—I. Determination of the valley structure and confrontation of observations with 1D analysis, *Soil Dyn. Earthq. Eng.*, **19**(1), 1–22.
- Régnier, J., Bonilla, L. F., Bard, P. Y., Bertrand, E., Hollender, F., Kawase, H. & Verrucci, L., 2018. PRENOLIN: international benchmark on 1D nonlinear site-response analysis—validation phase exercise, *Bull. seism. Soc. Am.*, **108**(2), 876–900.
- Riepl, J., Bard, P. Y., Hatzfeld, D., Papaioannou, C. & Nechtschein, S., 1998. Detailed evaluation of site-response estimation methods across and along the sedimentary valley of Volvi (EURO-SEISTEST), *Bull. seism. Soc. Am.*, **88**(2), 488–502.
- Rodriguez-Marek, A., Cotton, F., Abrahamson, N. A., Akkar, S., Al Atik, L., Edwards, B. & Dawood, H. M., 2013. A model for single-station standard deviation using data from various tectonic regions, *Bull. seism. Soc. Am.*, **103**(6), 3149–3163.
- Rodriguez-Marek, A., Montalva, G. A., Cotton, F. & Bonilla, F., 2011. Analysis of single-station standard deviation using the KiK-net data, *Bull. seism. Soc. Am.*, **101**(3), 1242–1258.
- Roten, D., Fäh, D., Cornou, C. & Giardini, D., 2006. Two-dimensional resonances in Alpine valleys identified from ambient vibration wavefields, *Geophys. J. Int.*, **165**(3), 889–905.
- Sgatonni, G. & Castellaro, S., 2020. Detecting 1D and 2D ground resonances with a single-station approach, *Geophys. J. Int.*, **223**, 471–487.
- Sjøvold, T., 1977. Non-metrical divergence between skeletal populations: the theoretical foundation and biological importance of CAB Smith's mean measure of divergence, *PhD thesis*, University of Stockholm, Sweden.
- Stafford, P. J., Rodriguez-Marek, A., Edwards, B., Kruijver, P. P. & Bommer, J. J., 2017. Scenario dependence of linear site-effect factors for short-period response spectral ordinates, *Bull. seism. Soc. Am.*, **107**(6), 2859–2872.
- Steimen, S., Fäh, D., Kind, F., Schmid, C. & Giardini, D., 2003. Identifying 2D resonance in microtremor wave fields, *Bull. seism. Soc. Am.*, **93**(2), 583–599.
- Stewart, J. P. & Afshari, K., 2021. Epistemic uncertainty in site response as derived from one-dimensional ground response analyses, *J. Geotech. Geoenviron. Eng.*, **147**(1),.
- Subhadra, N. & Mai, P. M., 2021. Estimation of ground-motion variability in Japan, *Bull. seism. Soc. Am.*, submitted.
- Tao, Y. & Rathje, E., 2019. Insights into modeling small-strain site response derived from downhole array data, *J. Geotech. Geoenviron. Eng.*, **145**(7),.
- Tao, Y. & Rathje, E., 2020. Taxonomy for evaluating the site-specific applicability of one-dimensional ground response analysis, *Soil Dyn. Earthq. Eng.*, **128**, 105865.
- Thompson, E. M., Baise, L. G., Tanaka, Y. & Kayen, R. E., 2012. A taxonomy of site response complexity, *Soil Dyn. Earthq. Eng.*, **41**, 32–43.
- Uebayashi, H., 2003. Extrapolation of irregular subsurface structures using the horizontal-to-vertical spectral ratio of long-period microtremors, *Bull. seism. Soc. Am.*, **93**(2), 570–582.
- Uebayashi, H., Kawabe, H. & Kamae, K., 2012. Reproduction of microseism H/V spectral features using a three-dimensional complex topographical model of the sediment-bedrock interface in the Osaka sedimentary basin, *Geophys. J. Int.*, **189**(2), 1060–1074.
- Vidale, J. E., 1989. Influence of focal mechanism on peak accelerations of strong motions of the Whittier Narrows, California, earthquake and an aftershock, *J. geophys. Res.*, **94**(B7), 9607–9613.
- Zalachoris, G. & Rathje, E. M., 2015. Evaluation of one-dimensional site response techniques using borehole arrays, *J. Geotech. Geoenviron. Eng.*, **141**(12),.

- Zhu, C., Cotton, F., Kwak, D. Y., Ji, K., Kawase, H. & Pilz, M., 2021a. Within-site variability in earthquake site response, *Geophys. J. Int.*, submitted.
- Zhu, C., Pilz, M. & Cotton, F., 2020. Evaluation of a novel application of earthquake HVSR in site-specific amplification estimation, *Soil Dyn. Earthq. Eng.*, **139**.
- Zhu, C., Weatherill, G., Cotton, F., Pilz, M., Kwak, D. Y. & Kawase, H., 2021b. An open-source site database of strong-motion stations in Japan: K-NET and KiK-net, *Earthq. Spectra*, **37**, 2126–2149.

SUPPORTING INFORMATION

Supplementary data are available at [GJI](#) online.

Table S1. KiK-net stations with 1-D/3-D site index, divergence

distance d , Jensen–Shannon difference J and interevent standard deviation σ .

Figure S1. Theoretical 1-D SH transfer function for all KiK-net sites (grey dotted lines). The black lines represent the average measured surface-to-borehole spectral ratios plus/minus one standard deviation corrected for the downgoing wave effect. For each site, I represents the 1-D/3-D index.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.