GFZ
Helmholtz-Zentrum
POTSDAM

HELMHOLTZ-ZENTRUM POTSDAM
DEUTSCHES
GEOFORSCHUNGSZENTRUM

Zhu, C., Pilz, M., Cotton, F. (2020): Evaluation of a novel application of earthquake HVSR in site-specific amplification estimation. - Soil Dynamics and Earthquake Engineering, 139, 106301.

https://doi.org/10.1016/j.soildyn.2020.106301

Institional Repository GFZpublic: https://gfzpublic.gfz-potsdam.de/

Evaluation of a novel application of earthquake HVSR in site-specific amplification estimation

Chuanbin Zhu[a,*], Marco Pilz[b], Fabrice Cotton[a,b]

[a] GFZ German Research Centre for Geosciences, Telegrafenberg, 14473, Potsdam, Germany
[b] University of Potsdam, Institute for Earth Sciences, Karl-Liebknecht-Str. 24-25, 14476, Potsdam, Germany

Article Info

Keywords:
Site amplification
HVSR
Ground response analysis
KiK-net
Earthquake

Abstract

Ground response analyses (GRA) model the vertical propagations of SH waves through flat-layered media (1DSH) and are widely carried out to evaluate local site effects in practice. Horizontal-to-vertical spectral ratio (HVSR) technique is a cost-effective approach to extract certain site-specific information, e.g., site fundamental frequency (f_0), but HVSR values cannot be directly used to approximate the levels of S-wave amplifications. Motivated by the work of Kawase et al. (2019), we propose a procedure to correct earthquake HVSR amplitudes for direct amplification estimations. The empirical correction compensates HVSR by generic vertical amplification spectra categorized by the vertical fundamental frequency (f_{0v}) via k-means clustering. In this investigation, we evaluate the effectiveness of the corrected HVSR in approximating observed linear amplifications in comparison with 1DSH modellings. We select a total of 90 KiK-net (Kiban Kyoshin network) surface-downhole sites which are found to have no velocity contrasts below their boreholes and thus of which surface-to-borehole spectral ratios (SBSRs) can be taken as their empirical transfer functions (ETFs). 1DSH-based theoretical transfer functions (TTFs) are computed in the linear domain considering uncertainties in V_S profiles through randomizations. Five goodness-of-fit metrics are adopted to gauge the closeness between observed (ETF) and predicted (i.e., TTF and corrected HVSR) amplifications in both amplitude and spectral shape over frequencies from f0 to 25 Hz. We find that the empirical correction to HVSR is highly effective and achieves a "good match" in both spectral shape and amplitude at the majority of the 90 KiK-net sites, as opposed to less than one-third for the 1DSH modelling. In addition, the empirical correction does not require a velocity model, which GRAs require, and thus has great potentials in seismic hazard assessments.

1. Introduction

Ground response analyses (GRA, see Acronyms) are widely carried out in current routine engineering practice to evaluate the effects of local geological conditions on earthquake-induced ground motions. GRA models vertically incident SH waves propagating through a laterally homogenous medium assuming either linear, equivalent linear or fully nonlinear stress-strain behavior (also referred to as "1DSH" modelling hereafter). GRA can capture some aspects of site effects, e.g., the influence of impedance contrast, 1D resonance, and nonlinearity (for equivalent and nonlinear analyses).

Thompson et al. [1] investigated the effectiveness of GRA by comparing the theoretical transfer functions (TTFs) based on linear 1DSH modellings with empirical transfer functions (ETFs) estimated from surface-to-borehole (S/B) spectral ratios (SBSR) at 100 KiK-net (Kiban Kyoshin network) sites [2,3]. A "good" match (Pearson's sample correlation coefficient $r > 0.6$) was only found at 18% of the sites. Several other such investigations were also carried out by, for instance, Refs. [4–8]. They reported different success rates of 1DSH modelling in duplicating observed site responses, which will be revisited in Discussion, but there is no doubt that GRA fails many sites due to a multitude of uncertainties involved in the process. To acquire accurate site responses by numerical simulations, firstly, one needs a ground model that correctly depicts the material properties (shear-wave velocity, mass density, damping, modulus, etc.) and their spatial distributions in a three-dimensional space (geometry). One also needs a set of laws to govern how various materials in the physical model behave under cyclic loadings as in reality (constitutive model). Besides, modelers also need to estimate the wavefield impinging at the model boundaries. All these factors contribute to the deviation of prediction from observation.

However, after analyzing the weak-motion blind predictions at the Turkey Flat Test Area, California, Cramer [9], attributed the mismatch primarily to the inaccuracy in the geotechnical models. Real et al. [10] also concluded that geotechnical models may be more important than the method used to calculate site responses. For the strong-motion blind predictions at Turkey Flat, Kwok et al. [11]

* Corresponding author.
E-mail address: chuanbin.zhu@gfz-potsdam.de (C. Zhu).
https://doi.org/10.1016/j.soildyn.2020.106301

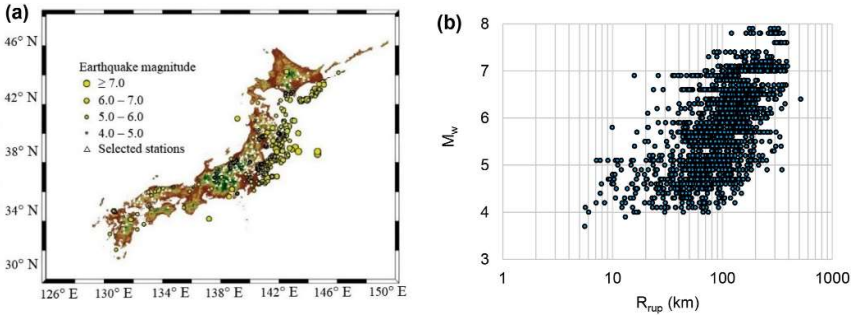

Figure 1. *(a) Spatial distribution of earthquakes and 207 KiK-net stations used in this study, and (b) M_w–R_{rup} distribution of the 1840 selected earthquake recordings.*

reported under-prediction bias at high frequencies and blamed the 1D soil model. Using KiK-net surface-downhole recordings, Kaklamanos and Bradley [4] found that material and geometrical properties of the 1D ground models contributed more to the discrepancy between simulated and observed amplifications. In addition, GRA has an inherent limitation arising from its lateral homogenous assumption and thus entirely overlooks complex site effects, e.g., laterally propagating surface waves, focusing and multi-dimensional resonance (e.g., Zhu et al. [12,13]).

Nakamura [14] proposed an alternative site-specific approach to quantify site effects, i.e., the horizontal-to-vertical (H/V) spectral ratio (HVSR) technique. HVSR does not need ground models and can be applied to either microtremor or earthquake recordings (e.g., Lermo and Chavez-Garcia [15]). Hence, the HVSR method is rather cost-effective and has been the focus of numerous investigations as reviewed by Bard [16]. There is a consensus that HVSR resembles the empirical transfer function (ETF) in shape (alignment of peaks and troughs) and can thus be utilized to reliably reveal the (horizontal) fundamental frequency (f_0) of a site. However, HVSR amplitudes tend to underestimate the actual levels of amplification, which hinders its broader application in site-effects evaluation.

Kawase et al. [17] proposed a double correction procedure to HVSR of microtremor recordings for a direct estimation of S-wave amplification. The first correction involves deriving earthquake HVSRs from microtremor HVSRs based on predefined generic spectral ratios between them. In a similar manner, the second step gives amplification estimates from earthquake HVSRs corrected by their binned average spectral ratios. Kawase et al. [17] have shown good matches between estimated and observed amplifications at 100 K-NET and KiK-net sites. This is quite promising for the direct application of HVSR in site-specific seismic hazard assessments. However, what is missing is whether the HVSR-based empirical approach is advantageous over the currently used site-specific approach, i.e., 1DSH modelling in characterizing site effects. Therefore, we devote this article to addressing this issue. It is achieved by comparing the goodness-of-fit (GoF) of TTFs with that of corrected HVSRs to observed amplifications at 90 KiK-net sites.

In this paper, firstly, we introduce our database which consists of 1840 earthquake recordings at 207 KiK-net vertical downhole sites. Then we describe a procedure to exclude the sites which are identified as having large velocity contrast below the downhole station, and thus SBSRs of the remaining sites (total number: 90) can be taken as their ETFs. Next, we compute the TTFs of the 90 sites, considering uncertainties in soil properties by randomizing about each base-case V_S profile through Monte Carlo simulations. Meanwhile, we obtain site amplifications by correcting HVSR after a description of the empirical correction procedure developed in this study. This is followed by a systematical comparison of the closeness of corrected HVSRs and TTFs to ETFs over frequencies from f_0 to 25 Hz. The effectiveness of the 1DSH modelling and the HVSR-base empirical correction are discussed.

2. Data

KiK-net is a strong-motion recording network, which consists of approximately 700 surface-downhole recording pairs of seismographs throughout Japan (Ⓔ, see Data and Resources). In a preceding investigation [18] on linear site effects at KiK-net stations, we selected earthquake ground-motions from a KiK-net database processed by Dawood et al. [19]. We considered earthquakes with a moment magnitude (M_w) in the range between 3.5 and 8.0. Then we only kept recordings (a) with a rupture distance (R_{rup}) up to 400 km, (b) having a high-pass corner frequency $f_c \leq 0.12$ Hz, (c) with a signal-to-noise ratio of at least 3.0 in the frequency range from $2f_c$ to 30 Hz, and (d) without being significantly affected by soil nonlinearity using the threshold proposed by Fujimoto and Midorikawa [20]. Meanwhile, we applied the following two criteria to screen seismic stations: (a) having at least three such records, and (b) with the downhole sensor installed in a layer at least 100 m deep from the ground surface and having a shear-wave velocity larger than 800 m/s. The second criterion was to mitigate the potential implications of downgoing waves on borehole recordings. In the end, we had 1840 seismograms from 207 KiK-net stations

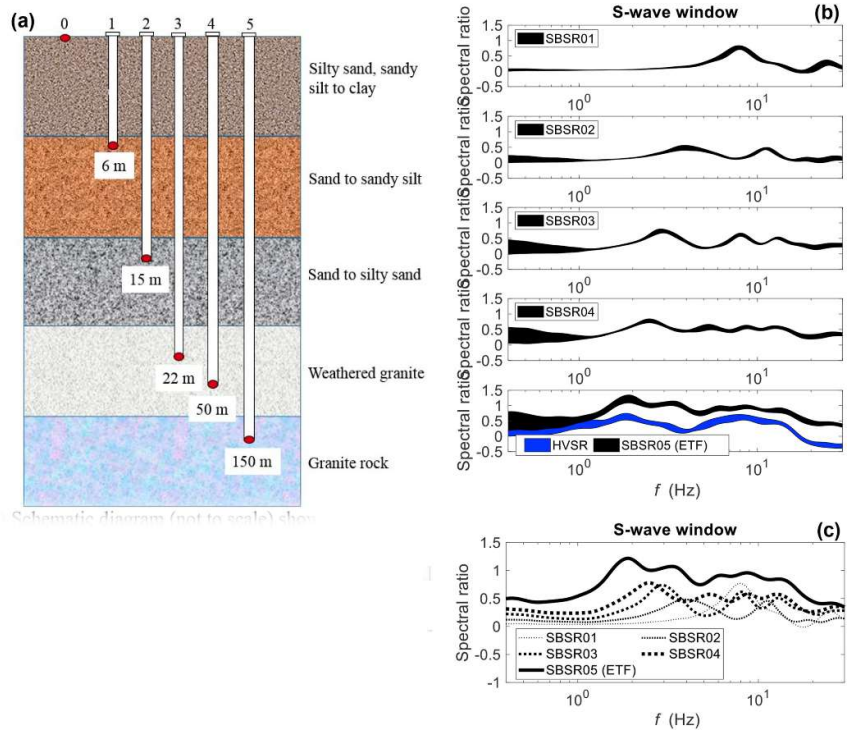

Figure 2. *(a) Schematic diagram (not to scale) showing the depths of accelerometers at the Garner Valley Downhole Array (GVDA); (b) SBSRs and surface HVSR, SBSRs are computed for the surface station with reference to different downhole stations (1–5), and the widths of the ratios' curves represent ± one standard deviation (1std); and (c) mean SBSRs. Spectral ratios are in log10 scales.*

in the database [18].

In the present research, we utilize the same dataset as Zhu et al. [18]. The spatial distributions of the selected earthquakes and KiK-net stations are shown in Fig. 1a. The M_w–R_{rup} distribution of the selected records is presented in Fig. 1b. We refer interested readers to Zhu et al. [18] for more details on site data, e.g., sediment thickness, average shear-wave velocity and f_0. Complete waveforms (P, S and S-coda waves) of each record are utilized and are tapered with a Tukey window at both ends of the truncated time series (5%). Fourier amplitude spectra (FAS) are computed for all three components of each recording and then are smoothed using the Konno-Ohmachi function [21] with a smoothing
coefficient $b = 20$. The two horizontal components (NS and EW) of each record are combined using their geometrical mean prior to the calculation of spectral ratios (e.g., SBSR or HVSR).

3. Methods

3.1. Empirical transfer function (ETF)

ETF is estimated from SBSR, which is the Fourier spectrum of the ground motion recorded on the ground surface divided by that at the borehole. For a recording k ($k = 1,2,\dots,N$, where N is the total number of recordings at a given station, and $N \geq 3$ in this study), S/B spectral ratios in horizontal and vertical directions are derived according to Eqs. (1) and (2), respectively:

$$SBSR_k(f) = \frac{H_k(f)}{H_{b,k}(f)} \tag{1}$$

$$SBSR_{v,k}(f) = \frac{V_k(f)}{V_{b,k}(f)} \tag{2}$$

Similarly, H/V spectral ratios at the ground surface and borehole are calculated by Eqs. (3) and (4), respectively:

$$HVSR_{v,k}(f) = \frac{H_k(f)}{V_{b,k}(f)} \tag{3}$$

$$SBSR_{b,k}(f) = \frac{H_{b,k}(f)}{V_{b,k}(f)} \tag{4}$$

where $H_k(f)$ and $H_{b,k}(f)$ are the geometrical means of the smoothed Fourier spectra of the two horizontal components recorded on the ground surface and at the borehole, respectively, during event k. Correspondingly, $V_k(f)$ and $V_{b,k}(f)$ are those for the vertical direction. Subscript v denotes "vertical" and is used to differentiate from its horizontal (default) counterpart; likewise, subscript b means "borehole". We assume the logarithmic value of spectral ratio at a certain frequency is a variable following a Gaussian distribution, thus its statistical quantities, mean and standard deviation, can be obtained using Eqs. (5) and (6), respectively:

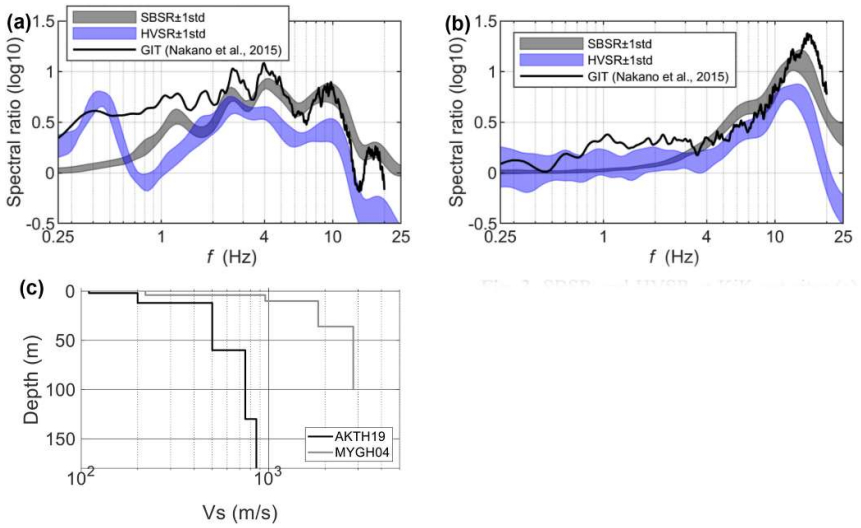

Figure 3. *SBSR and HVSR at KiK-net sites (a) AKTH19 (with velocity contrast below borehole) and (b) MYGH04 (without velocity contrast below borehole), as well as their (c) V_S profiles. Red curves are amplification estimates derived by Nakano et al. [23] through the generalized inversion technique (GIT). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)*

$$\log_{10} SR(f) = \frac{\sum_{k=1}^{N} \log_{10} SR_k(f)}{N} \quad (5)$$

$$std_{SR}(f) = \sqrt{\frac{\sum_{k=1}^{N} (\log_{10} SR_k(f) - \log_{10} SR(f))^2}{N-1}} \quad (6)$$

where SR represents the average of generic Fourier spectral ratios, including SBSR, $SBRS_v$, $HVSR$, and $HVSR_b$.

It is expected that deep velocity structures affect the amplification spectra in the low-frequency range. However, using a numerical approach, Yamanaka et al. [22] showed that deep velocity configurations could influence the amplifications at high frequencies as well. When the underlying layer having higher V_S is removed, the levels of amplification at high frequencies are lower than those of the complete model. To verify this, we capitalize on the Garner Valley Downhole Array (GVDA) in Southern California. GVDA consists of both surface (station code: 0) and downhole (station code: 1–5) accelerometers at depths of 6, 15, 22, 50, and 150 m, respectively (Fig. 2a). Parameters of selected earthquakes are listed in Table S1 (Ⓔ, see Data and Resources). We compute the SBSRs for the surface station (0) with reference to different downhole stations (1–5). The deepest downhole station (5) is embedded in a competent granite bedrock with $V_S = 1632$ m/s, and thus its corresponding SBSR05 is taken as the ETF of the site. Average SBSRs (log10) and their variations are shown in Fig. 2b and are compared in Fig. 2c. Fig. 2c clearly shows that SBSR01~04 significantly underestimate SBSR05 (ETF), suggesting that neglecting deep velocity variations causes a considerable underestimation of site responses at both low and, more interestingly, high frequencies although downward waves may also be somewhat responsible. Our empirical evidence is consistent with the numerical results by Yamanaka et al. [22]. Therefore, only when there is no abrupt impedance contrast (IC) below the downhole sensor at a site, can SBSR be used as its ETF.

Even though we only include KiK-net stations with boreholes deeper than 100 m and reaching or penetrating the layer with $V_S = 800$ m/s, we cannot rule out the possibility of having other deeper velocity changes that can influence site responses in the investigated frequency range 0.25–25 Hz. Hence, we further screen these 207 sites by examining the potential existence of major ICs below downhole sensors. The presence of deeper ICs at a site is gauged based on the difference between SBSR and surface HVSR at frequencies below f_0 determined from SBSR. HVSR is used since it carries the information on all ICs beneath a surface station. As exemplified in $SBSR_{b,k}$, there is an apparent peak on HVSR below 1.0 Hz at site AKTH19 whereas such a peak is completely absent on its SBSR. This suggests the presence of a deeper IC, and thus the V_S logging data (Fig. 3c) is just a partial representation of the complete model at AKTH19. On the contrary, site MYGH04 (Fig. 3b) is deemed as having no further impedance discontinuity, and its V_S profile (Fig. 3c) between surface and borehole stations can be considered as a complete representation.

Based on the above analysis, we define a site as having a major IC below its borehole if there exists a frequency (below f_0) at which HVSR amplitude exceeds SBSR value by 20% or more. The threshold value (20%) is defined subjectively, and future study is recommended for a more stringent setting. This screening is contingent on the useable frequency range, i.e., 0.25–25 Hz in this study, thus we could not identify large ICs which could cause HVSR peaks at $f < 0.25$ Hz, even though such ICs still affect amplifications at $f > 0.25$ Hz. However, by this criterion, we can partition the 207 KiK-net sites into two groups: sites

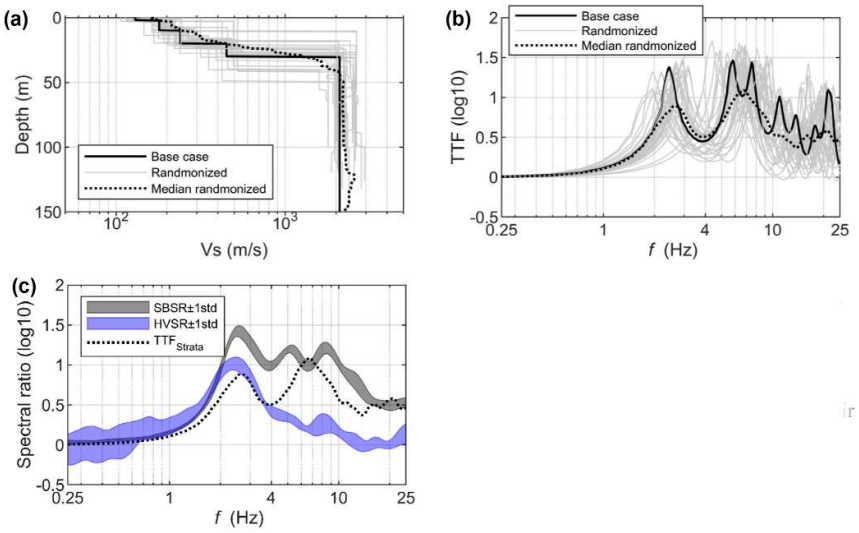

Figure 4. *Site effects at KiK-net site IBRH11. (a) Base-case profile as well as 30 randomized profiles and their median; (b) TTF_{Strata} of the base-case profile as well as TTF_{Strata}s of the 30 random profiles and their median; and (c) SBSR (ETF), HVSR and median TTF_{Strata}.*

with (total number: 117 or 57%) and without (total number: 90 or 43%) major ICs below boreholes. No significant trends are observed with regards to V_{S30} and the presence of deeper ICs. In the following analyses, we only utilize the 90 sites without major deeper ICs. At these sites, the strata between the surface and downhole sensors can be considered as a complete representation of the site, and thus its corresponding SBSR can be taken as its ETF. VS-depth plots of the 90 sites are provided as Electronic Supplements to this article (Ⓔ File Spectral Ratios, see Data and Resources).

Nakano et al. [23] separated site effects from the Fourier spectra of ground motions observed at KiK-net stations through the generalized inversion technique (GIT, Andrews [24]). At sites AKTH19 and MYGH04, their amplification spectra are compared with our SBSR results in Fig. 3a and b, respectively. It should be noted that the GIT-based amplifications are referenced to the seismological outcrop rock (V_S = 3.45 km/s). At AKTH19 (Fig. 3a), SBSR significantly underestimates GIT estimates at low frequencies, which is expected given that we have identified the presence of a deeper IC at this site. However, at MYGH04 (Fig. 3b), SBSR and GIT results match quite well in both spectral shape and amplitude even though SBSR is referenced to borehole bedrock (V_S, b = 2.83 km/s). A good match can also be seen at many other sites at which we conduct such a comparison (not presented here for brevity), which justifies our use of SBSR as ETF. However, the good match at relatively high frequencies might be a result of the cancellation of two competing effects. SBSR, on one hand, tends to overestimate amplifications due to destructive interference between up- and down-going waves at borehole; on the other hand, SBSR may underestimate GIT results due to its smaller reference V_S (>0.8 km/s versus 3.45 km/s).

3.2. Theoretical transfer function (TTF)

TTF is obtained based on linear 1DSH assumptions. Since ETF is referenced to borehole conditions, TTF is thus computed according to Eq. (7) as the ratio of the amplitude of harmonic motion on the ground surface ($A_0 + B_0$) to that of the total wavefields within the bedrock at the borehole depth z ($A_z + B_z$):

$$TTF\ (f) = \frac{A_0 + B_0}{A_z + B_z} \tag{7}$$

where A_0 and B_0 represent the amplitudes of incident waves and free-surface reflection, respectively, $A_0 = B_0$; and A_z and B_z denote the amplitudes of upward and downward waves at the depth z.

Given that empirical spectral ratios are averaged over at least three recordings (Eq. (5)), to enable an impartial comparison, we devote efforts to accounting for the variation in TTF associated with uncertainties in V_S profiles. Thus, we compute TTF in the present research using Strata [25], denoted by TTF_{Strata} hereafter. V_S variability is commonly separated into aleatory variability and epistemic uncertainty. Aleatory variability arises from the spatially variable nature of soil properties, and it is often accounted for by using alternative base-cases, namely upper-and lower-range values corresponding to the 10th and 90th percentiles of a normal distribution, in addition to the mean base-case.

In comparison, epistemic uncertainty stems from our inability to perfectly measure these soil properties and is typically considered through correlated random perturbations to base-case values (e.g., Teague et al. [26]. However, some studies (e.g., Teague et al. [26]) demonstrated that upper-and lower-range profiles did not match the empirically-derived site signature, e.g., resonant frequency and gave unrealistic site responses. Thus, in this

Table 1
Toro [27] site variation model parameters.

Model	Parameter	Value
Velocity	Standard deviation σ_{ln}	0.15
	Correl. coeff. at the surface (ρ_0)	0.99
	Correl. coeff. at 200 m (ρ_{200})	1
	Change in correl. with depth (Δ)	8
	Depth intercept (d_0)	0
	Exponent (b)	0.16
Layer thickness	Coefficient (a)	1.98
	Initial (b)	10.86
	Exponent (c)	-0.89
Depth to bedrock	Standard deviation σ_{ln}	0.15

investigation, we do not use alternative base-cases but only consider the epistemic uncertainty in soil properties by randomizing about mean base-case ground models through Monte Carlo simulations.

At a given site, Monte Carlo simulations estimate its median site response and variation by generating a large sample of site profiles with soil properties selected from a defined statistical distribution and then computing their site responses. The Toro [27] V_S randomization model is widely employed to describe the statistical distribution and correlation of soil properties (V_S, the thickness of each layer, and depth to bedrock). Since downhole array sites, no exception to KiK-net sites, are typically characterized by only one single measure with no estimates of uncertainty, we thus adopt default/recommended statistical properties in Strata, as listed in Table 1.

At each site, we consider a total number of 30 realizations that are needed to achieve a stable median value [28]. Since there is no density information available in KiK-net database, an empirical equation is utilized to obtain mass density (ρ_s, unit: kg/m^3) from V_s (unit: m/s): $\rho_s = 0.914V_s^{0.122}$ which was proposed by Wang et al. [29] for K-NET profiles. Shear-wave effective quality factor (Q) is determined from V_s (unit: m/s) using the Campbell model ([30], Eq. (8)). Eq. (8) was initially proposed for the Central and Eastern United States but has been applied in many other regions, including Japan (e.g., Cabas et al. [31]). Q is converted to small-strain soil damping (D_{min}) through Eq. (9).

$$Q = 7.17 + 0.0276V_S \tag{8}$$

$$D_{min} = \frac{1}{2Q_S} \tag{9}$$

Fig. 4a illustrates the 30 randomly generated V_s profiles developed for a KiK-net site IBRH11. The median of the 30 random profiles roughly coincides with the base-case. Fig. 4b displays the TTF_{Strata}s of the 30 random soil models and their median, as well as the TTF_{Strata} of the base-case profile. Comparing the randomized median with the base-case site responses, randomization smooths the amplification curve and remarkably lowers individual peaks, which is also observed by Rathje et al. [32]; among others. The reduction in median amplifications is related to the way in which variabilities of soil properties are modeled [28]. In Fig. 4c, TTF_{Strata} is compared with empirical spectral ratios SBSR and HVSR. Both 1DSH modelling and HVSR roughly give correct $f_0 = 2.6$ Hz, which corresponds to the large velocity jump at the depth of 30 m.

3.3. Goodness-of-fit metrics

To quantify the resemblance in shape (alignment of peaks and troughs) between observed and predicted amplifications in various frequency ranges, we use Pearson's sample correlation coefficient r, Spearman's rank correlation coefficient ρ, and Kendall's rank correlation coefficient τ (Table 2). Pearson's r measures the strength of linear correlation between two variables, but it requires two variables to be approximately normally distributed and to have no extreme outliers. We adopt Pearson's r to enable a direct comparison with other studies, e.g., Thompson et al. [1]. Besides, two nonparametric measures of rank correlation (whether linear or not), Spearman's ρ and Kendall's τ, are also employed since they do not require data normality and are insensitive to outliers. The Spearman correlation between two variables is equal to the Pearson correlation between the rank values of those two variables. All three coefficients range from -1 to 1. A value of 1 represents a total positive correlation, 0 is no correlation, and -1 denotes a total negative correlation.

In addition to the spectral shape, we also investigate the GoF in spectral amplitudes between observed and predicted amplifications at different frequencies. We adopt the Index of Agreement d [33] and Mean Absolute Error (MAE, Table 2) to gauge the closeness in amplitudes in relative and absolute terms, respectively. Values of d vary between 0 and 1. An agreement value of $d = 1$ indicates a perfect match, and 0 indicates no agreement at all. MAE, which quantifies the absolute difference between two variables, is recommended by Legates and McCabe [34] to be used along with a relative measure, i.e., d in this case.

4. Empirical correction to earthquake HVSR

For a pair of surface and downhole recording stations, we express the HVSR on the ground surface as:

$$\begin{aligned} HVSR(f) &= \frac{H(f)}{H_b(f)} \cdot \frac{H_b(f)}{V_b(f)} \cdot \frac{V_{b(f)}}{V(f)} \\ &= \frac{HVSR_b(f)}{SBSR_v(f)} \cdot SBSR(f) \end{aligned} \tag{10}$$

For the surface HVSR to represent the horizontal amplifications, Nakamura [14] made two major assumptions. The first assumption is that, in a firm substratum, wave propagation is even in all directions, namely, $HVSR_b$

Table 2
Goodness-of-fit (GoF) metrics used in this investigation.

Goodness-of-fit metric	Expression	Range	Measure	Interpretation
Pearson's r	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{(y_i-\bar{y})^2}}$	[-1, 1]	Linear relationship	Measure the closeness in shape (alignment of peaks and troughs)
Spearman's ρ	$\frac{cov(rg_x, rg_y)}{\sigma_{rg_x}\sigma_{rg_y}}$	[-1, 1]	Ordinal relationship	
Kendall's τ	$\frac{2[\sum_{i<j} sgn(y_i-y_j)]}{n(n-1)}$	[-1, 1]	Ordinal relationship	
Index of Agreement d	$1-\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(\lvert y_i-\bar{x}\rvert+\lvert x_i-\bar{x}\rvert)^2}$	[0, 1]	Degree of difference (relative)	Measure the difference in amplitude
Mean Absolute Error MAE	$\frac{\sum_{i=1}^{n}\lvert y_i-x_i\rvert}{n}$	-		Degree of difference (absolute)

Note: n, sample size ($n = 90$ in this research); x and y, vectors of observations and predictions, respectively; x_i and y_i, ith ($i = 1,2,\dots,n$) element of x and y, respectively; $\bar{x}$ and $\bar{y}$, Means of x and y, respectively; rg_x and rg_y, rank vectors of x and y, respectively; σ_{rg_x} and σ_{rg_y}, standard deviations of rg_x and rg_y, respectively; cov, covariance operation; sgn, the sign function.

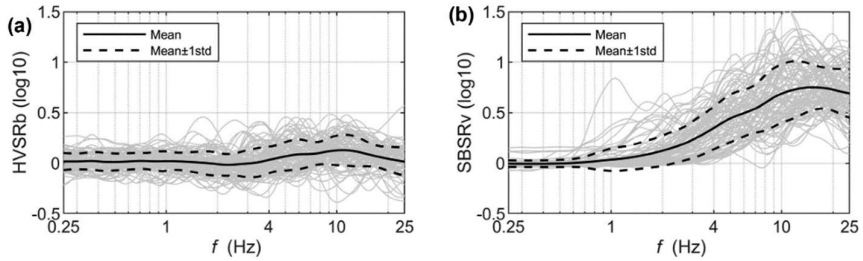

Figure 5. *Spectral ratios (a) $HVSR_b$s and (b) $SBSR_v$s at the 90 KiK-net sites. In each plot, each thin solid line represents the average spectral ratio curve at a site (Eq. (5)).*

is equal to unity (within a factor of 2, the usual uncertainty when using spectral ratios, e.g., King and Tucker [35]):

$$H_b(f) = V_b(f) \text{ or } HVSR_b = 1.0 \tag{11}$$

This assumption was verified by Nakamura [14] experimentally using microtremor measurements at downhole stations. For earthquake recordings, many studies (e.g., Ref. [36,37], confirmed that $HVSR_b$ could be approximated by a constant close to unity. In Fig. 5a, we examine the fluctuations of $HVSR_b$ with frequency at the 90 selected stations. It shows that $HVSR_b$ in the aggregate can be represented by a near-unity constant in the investigated frequency range, and the variation with frequency is insignificant. Hence, this assumption ($HVSR_b = 1.0$) can be considered to be "true" for our dataset. Additionally, Fig. 5a also substantiates the effectiveness of our "no IC below borehole" screening.

The second assumption is that vertical components of seismic waves are not influenced by local site effects, namely

$$V_b(f) = V(f) \text{ or } SBSR_v = 1.0 \tag{12}$$

To verify this assumption, we plot the $SBSR_v$s at the 90 KiK-net sites in Fig. 5b. Rather pronounced amplifications in the vertical direction can be observed, especially at relatively high frequencies. This clearly demonstrates that the second assumption ($SBSR_v = 1.0$) does not hold at most sites.

Therefore, the first assumption (Eq. (11)) is valid whereas the second one (Eq. (12)) is rejected, then according to Eq. (10), prominent vertical amplifications ($SBSR_v$) are the main reason for the underestimation of HVSR to SBSR at relatively high frequencies. Vertical amplifications are associated with the propagations of P-waves regardless of the chosen window (P-or S-wave window), and these P-waves primarily come from the conversions of non-vertically travelling SV waves at major layer interfaces (e.g., Parolai et al. [38]). SV-to-P conversions divert wave energies to the vertical direction and cause the underestimation of HVSR to the horizontal amplifications. To utilize HVSR for a direct estimation of horizontal amplifications, we need to compensate HVSR by this underestimation controlled by $SBSR_v$. Thus, substituting Eq. (11) into Eq. (10), we can obtain horizontal amplifications via the following formula:

$$SBSR(f) = HVSR(f) \cdot SBSR_v(f) \tag{13}$$

The above equation suggests that the horizontal amplification function at a given site can be obtained through correcting its surface HVSR by its $SBSR_v$. However, Eq. (13) has a very limited engineering implication since it contains the site-specific $SBSR_v$, which is often unattainable in practice. Kawase et al. [17] proposed to use generic $SBSR_v$ as a substitute for the site-specific one in Eq. (13). Following Kawase et al. [17]'s approach, we utilize a

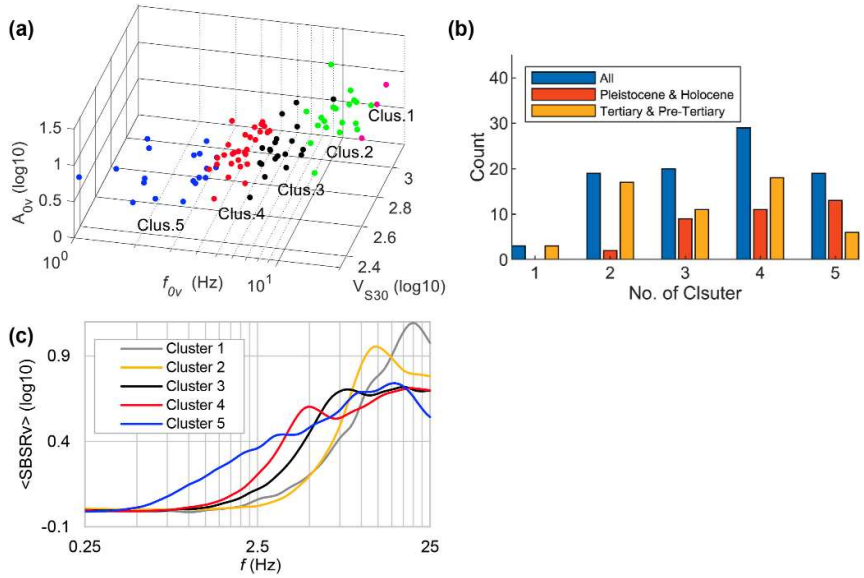

Figure 6. *k-means clustering of the 90 KiK-net sites. Distributions of these clustered sites in (a) f_{0v}_V_{S30}_A_0 space, and (b) each geological unit according to the classification by Wakamatsu et al. [39]; and (c) average $SBSR_v$ for each cluster, i.e., $\langle SBSR_v \rangle$.*

binned average vertical amplification spectrum $\langle SBSR_v \rangle$ to correct HVSR for horizontal amplification estimates, referred to as pseudo surface-to-borehole spectral ratio (pSBSR):

$$pSBSR(f) = HVSR(f) \cdot < SBSR_v(f) > \qquad (14)$$

where $< >$ represents the averaging operation. To construct generic vertical correction spectra $\langle SBSR_v \rangle$ for our study, we first analyze each individual $SBSR_v$ curve. We find that, as shown in Fig. 6a (Ⓔ, see Data and Resources), A_{0v} scales strongly with f_{0v} (primary) and V_{S30} (secondary), analogous to the horizontal amplification [18]. A_{0v} denotes the vertical amplification at the vertical fundamental frequency f_{0v}, and V_{S30} is the average shear-wave velocity in the top 30 m.

Next, based on f_{0v}, V_{S30} and A_{0v}, we partition the 90 sites into k number of mutually exclusive clusters by an unsupervised technique, k-means clustering [40,41]. A value of $k = 5$ is found to be the optimal number of clusters. Each site characterized by V_{S30}, f_{0v}, and A_{0v} (Fig. 6a) is assigned to a cluster by minimizing the distance from the data point to the mean location of its assigned cluster. The resultant clusters 1–5 contain 3, 29, 19, 20 and 19 sites, respectively and can be distinguished based solely on f_{0v}. Ranges of f_{0v} (in Hz) for clusters 1–5 are [17.0, 19.8], [10.8, 15.0], [6.2, 9.4], [3.7, 5.6], and [1.1, 3.5], respectively. As shown in Fig. 6a, from Clusters 1 to 5, f_{0v}, V_{S30} and A_{0v} follow a decreasing trend.

Fig. 6b depicts the distribution of these clustered sites in each geological unit (rock age) which is collected from the Japan Engineering Geomorphologic Classification Map [39]. Considering the sample size, herein we adopt a binary geological classification: (i) Pleistocene, Holocene and Quaternary (volcanic), and (ii) Tertiary and Pre-Tertiary. Clusters 1–4 are dominated by sites from old geologies which tend to have relatively shallow and stiff sites (with high f_{0v}). In contrast, the majority of sites in Cluster 5 are from relatively young geologies which tend to host soft and deep sites with low f_{0v}. After the clustering, we then derive the average $SBSR_v$ for sites within each cluster, i.e., $\langle SBSR_v \rangle$. As shown in Fig. 6c, $\langle SBSR_v \rangle$ curves are distinctly different. As the cluster number increases (from 1 to 5), the first peak and its corresponding frequency decrease. $\langle SBSR_v \rangle$ for each cluster can be found in Table S2 (Ⓔ, see Data and Resources).

Our correction procedure is, in principle, motivated by the one proposed by Kawase et al. [17]; but there are several differences. For instance, we assume that $HVSR_b = 1.0$ (Eq. (11)) is reasonable whereas Kawase et al. [17] considered the fluctuation of $HVSR_b$ with frequency. We make this assumption since it enables us to utilize $SBSR_v$ in the correction (Eq. (13)), rather than $V(f)/H_b(f)$ as adopted by Kawase et al. [17]. We prefer $SBSR_v$ because it has a clear physical meaning, namely the vertical amplifications, which facilitates the subsequent clustering. Another striking difference is that our amplifications are referenced to borehole bedrock whereas that of Kawase et al. [17] is conditioned on seismological outcrop rock. The above two prime differences deter a direct comparison of the correction spectra from the two investigations. In Fig. 7, we present a few examples to show the efficacy of the empirical correction (Eq. (14)) in comparison with other techniques (Fig. 8).

Fig. 7a and b depict HVSR, TTF_{Strata}, pSBSR and $HVSR*SBSR_v$ (Eq. (13)) in comparison with SBSR (ETF) at KiK-net sites IWTH27 and IBRH13 from Clusters 2 and 4, respectively. To enable a better view, we only display HVSR, TTF_{Strata}, pSBSR and SBSR in Fig. 7c–f. Comparing with HVSR, pSBSR has an obviously improved match with

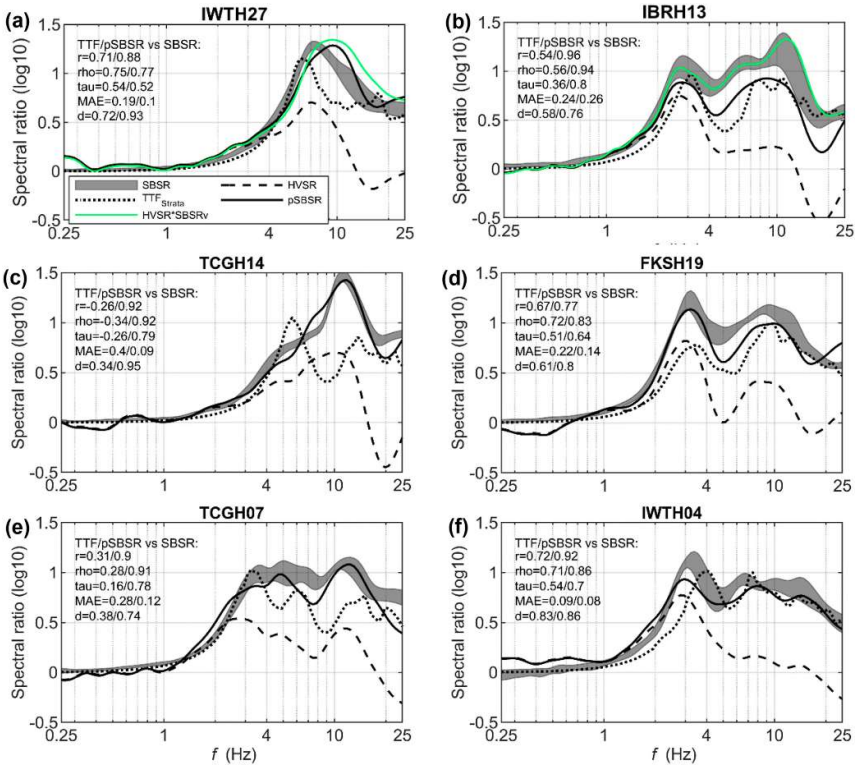

Figure 7. *HVSR, TTF_{Strata} and pSBSR (Eq. (14)) and HVSR*$SBSR_v$ (Eq. (13)) in comparison with SBSR (ETF) at sites (a) IWTH27 from Cluster 2, and (b) IBRH13 from Cluster 3. To have a better view, we simplify the plots for sites (c) TCGH14 (Cluster 1), (d) FKSH19 (Cluster 4), (e) TCGH07 (Cluster 4), and (f) IWTH04 (Cluster 3). HVSR is the horizontal-to-vertical spectral ratio; SBSR and $SBSR_v$ are the surface-to-borehole spectral ratios in the horizontal and vertical directions, respectively; pSBSR denotes pseudo SBSR, i.e., corrected HVSR; TTF and ETF represent theoretical and empirical transfer function, respectively. The widths of the ratios' curves represent ± one standard deviation.*

SBSR (ETF), and the improvement is especially notable at relatively high frequencies. Fig. 7 exemplifies the efficacy of the novel application of HVSR after correction. Such plots for all the 90 sites are provided as an Electronic Supplements to this article (Ⓔ File Spectral Ratios, see Data and Resources).

5. Comparison of site-specific amplification estimates

In this section, we systematically compare the performance of pSBSR with TTF_{Strata} in approximating ETF at the 90 KiK-net sites, as illustrated in Fig. 8. Though HVSR is not directly used for amplification estimation in practice, it is included in the comparison to assess to what extent the corrected HVSR is advantageous over HVSR. However, the approximation by Eq. (13) is not considered here since it needs a site-specific $SBSR_v$ which is often unavailable and necessitates the usage of its clustered average $\langle SBSR_v \rangle$ in Eq. (14). Hence, we only evaluate the closeness of HVSR, pSBSR, and TTF_{Strata} to SBSR (Fig. 8). GoF statistics (Table 2) are computed in eight frequency bands which correspond to 0.25–1.0, 1.0–2.0, 2.0–4.0, 4.0–8.0, 8.0–12.0, 12.0–16.0, 16.0–20.0, 20.0–25.0 Hz, respectively. To compare the effectiveness of any two techniques, we plot their difference (Δ) in GoF at the midpoint of each frequency band in Fig. 9. Since r, ρ and τ all measure the similarity in spectral shape, for brevity, we only present the results for r. For the same reason, we only display results of MAE for amplitude.

Fig. 9a and b compare HVSR and corrected HVSR (or pSBSR) in spectral shape and amplitude, respectively. Fig. 9a shows that the median value of Δr is nearly zero regardless of the variation in frequency. Because the shape of HVSR is known to match that of SBSR well, $\Delta r = \sim 0$ suggests that the corrected HVSR inherits the good match. This is because the clustering and averaging operation (Fig. 6) result in a relatively smooth correction which retains site-specific features carried by HVSR.

Fig. 9b depicts ΔMAE in various frequency ranges. For frequencies 0.25–2.0 Hz, the median ΔMAE is zero with a very small percentile value. This is because sites with $f_{0v} > 2.0$ Hz constitute the majority (83 out of 90 sites) of our dataset, and thus $\langle SBSR_v \rangle$ is insignificant at frequencies lower than 2.0 Hz (Fig. 6c), and subsequently, there is a negligible correction. For frequencies higher than 2.0

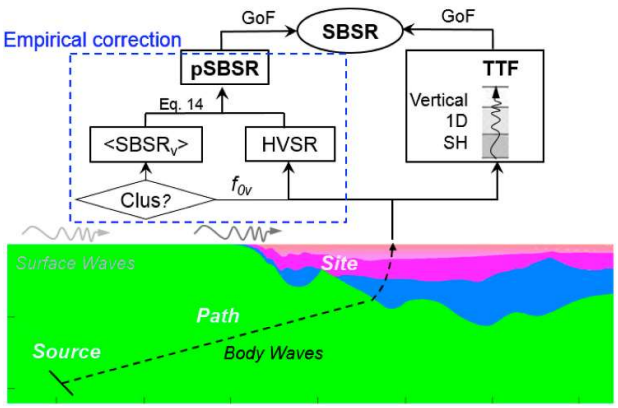

Figure 8. *Illustration of techniques used in site effects quantification.*

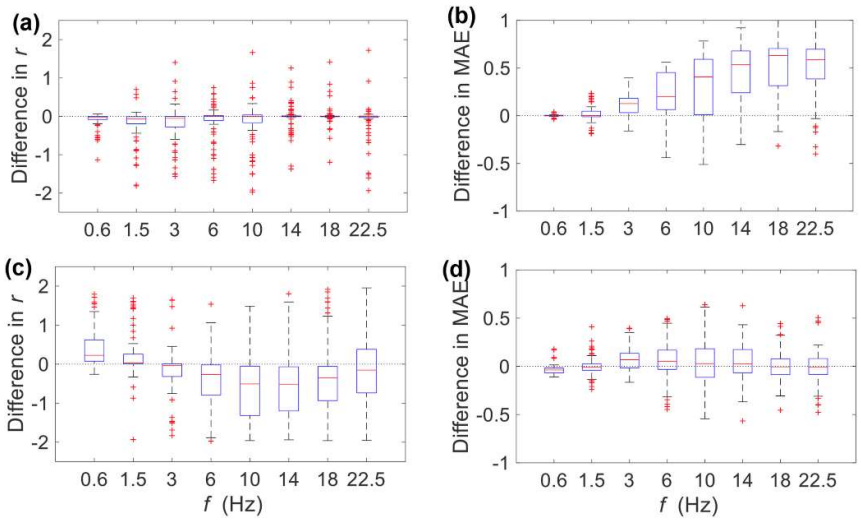

Figure 9. *Comparison of closeness to SBSR (ETF) between (a, b) HVSR and pSBSR, and (c, d) TTF_{Strata} and pSBSR using Pearson's r (the first column) and Mean Absolute Error MAE (the second column) to examine closeness in spectral shape and amplitude, respectively. Correlation coefficients are calculated in eight frequency bands: 0.25–1.0, 1.0–2.0, 2.0–4.0, 4.0–8.0, 8.0–12.0, 12.0–16.0, 16.0–20.0 and 20.0–25.0 Hz. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.*

Hz, ΔMAE is larger than zero and increases significantly with frequency (Fig. 9b). This implies that, compared to ETF, corrected HVSR has a smaller margin of error than HVSR, evidencing the effectiveness of the correction procedure in improving the match with ETF in amplitude, especially at high frequencies (>~10.0 Hz).

The corrected HVSR has shown an apparent advantage over HVSR. It is also intriguing to know whether the corrected HVSR performs better than TTF_{Strata} based on 1DSH modelling which is widely conducted in seismic hazard analyses. We thus compare TTF_{Strata} and corrected HVSR in Fig. 9c and d. For frequencies lower than f_0, both TTF_{Strata} and ETF follow a monotonic change in amplitude with frequency, resulting in an artificially good match with r values close to 1.0 at most sites [1]. Thus,
we will not compare TTF_{Strata} and pSBSR in the low-frequency band 0.25–2.0 Hz. However, for frequencies higher than 2.0 Hz, Δr and ΔMAE consistently suggest that corrected HVSR has an obviously better match with ETF in both spectral shape (Fig. 9c) and amplitude (Fig. 9d) than TTF_{Strata}. This demonstrates that the HVSR-based empirical approach is also advantageous over the site-specific GRA.

6. Discussion

6.1. Effectiveness of 1DSH modeling

Fig. 9 only shows the difference in GoF quantities in different frequency ranges, henceforth we investigate the actual GoF values computed over the whole frequency range from f_0 to 25 Hz. GoF results of the 90 sites are provided in Table S3 (Ⓔ, see Data and Resources). Fig. 10a–c depicts the values of Pearson's r between TTF_{Strata} and ETF against site characterization proxies, including topographic slope, V_{S30}, and f_0, respectively. Topographic slopes at KiK-net sites are derived from a 1-arc (approximately 30 m) resolution digital elevation model. It is worth noting that r gauges the similarity in spectral shape, and $r > 0.6$ is often regarded as the threshold for a "good match" (e.g., Thompson et al. [1]). According to this criterion, we find a "good match" ($r > 0.6$) at 27% (24 out of 90) of our selected sites.

Comparing with similar studies on KiK-net sites, the pioneering work by Thompson et al. [1] reported a "good match" (r > 0.6) at 18 out of 100 sites regardless of interevent variability. Later Kaklamanos and Bradley [4], found the same percentage (18% or 21 out of 114 sites) of "good match" ($r > 0.6$) sites. Pilz and Cotton [6] defined a "good match" using Spearman's $\rho > 0.6$ and considered the reliability of V_S profiles. However, if we count the total number of examined sites, the percentage of "good match" ($\rho > 0.6$) is 28% (196 out of the 689). In addition to the Japanese sites, Afshari and Stewart [7] investigated 21 vertical arrays in California and identified 24% of sites having a "good match" ($r > 0.6$).

Nevertheless, a significantly high success rate was reported by Laurendeau et al. [5] who found 108 out of 152 KiK-net sites (or 71%) having a "good" match ($r > 0.6$).

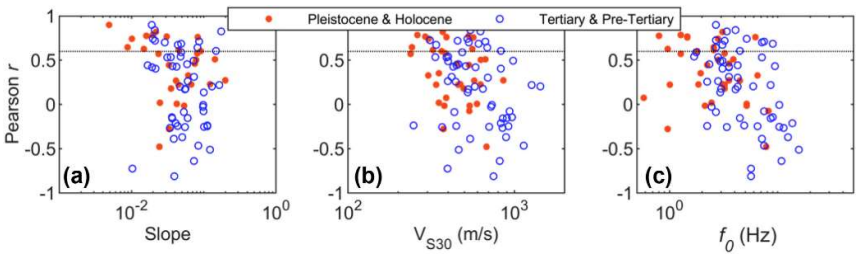

Figure 10. *Pearson's r between TTF_{Strata} and ETF in the frequency range from f_0 to 25 Hz at the 90 selected KiK-net sites.*

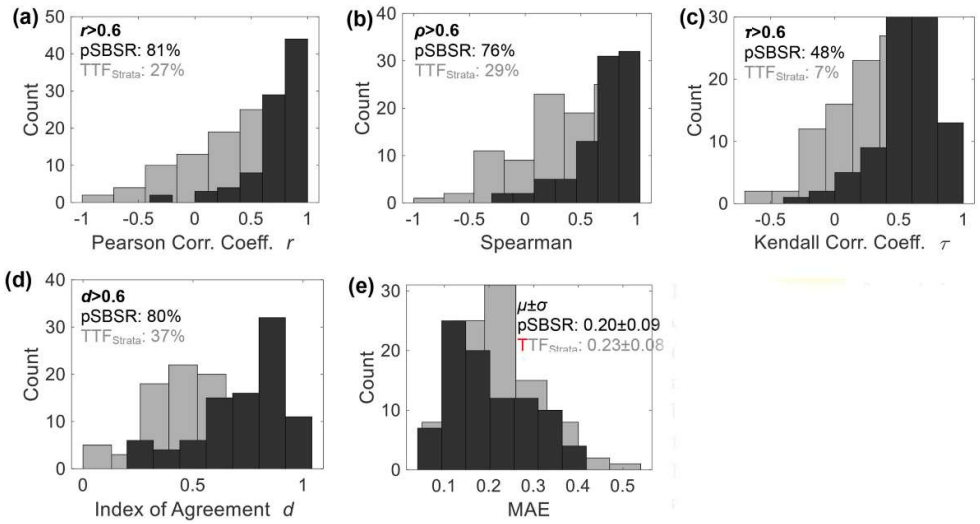

Figure 11. *Histograms of correlation coefficients (a) Pearson's r, (b) Spearman's ρ, (c) Kendall's τ, (d) Index of Agreement d, and (e) Mean Absolute Error MAE (log10) between pSBSR and ETF (black) in the frequency range from f0 to 25 Hz for the 90 KiK-net sites. Histograms for TTF_{Strata} (grey) are superimposed.*

This success rate is much higher than that in this study and others (e.g., Pilz and Cotton [6]). It may be because Laurendeau et al. [5] computed r over a frequency range from $f_0/2$ whereas others used f_0 as the lower-bound. Computing r over a lower frequency range gives an artificially better match due to the monotonic change of both SBSR and TTF for $f < f_0$. Another reason may be that Laurendeau et al. [5] smoothed their TTFs, which would also improve the fit.

However, the success rates in the present study and those of [1,4,6,7] are consistent and are below one third. This indicates a limited efficacy of 1DSH modellings in duplicating observed amplifications. The causes of its overall poor performance are multifaceted. As shown in Fig. 10a–c, Pearson's r exhibits a decreasing trend with slope, V_{S30}, and f_0, respectively, implying poorer fits at steeper and stiffer sites which tend to be from relatively old geological units. Considering that most KiK-net sites are on thin sediments or weathered rocks at the margin of basins, the violation of 1D assumption and complex geological conditions might be responsible for the low success rate. Although randomizing V_S profiles accounts for the spatial variations in soil properties to a certain degree, we did not apply any site-specific constraints on the randomness. Excluding randomized profiles that are inconsistent with measured site signatures, e.g., f_0, is one way to achieve informed randomization and can potentially improve the match (e.g., Teague et al. [26]).

In addition to multi-dimensional effects, in a preceding study [42], we found around 30% KiK-net sites at which its theoretical f_0 was out of 0.5–2.0 times of its empirical one, suggesting a flawed 1D geotechnical model, e. g., imperfect representation of vertical layering and faulty estimates of material properties. Many previous studies (e.g., Refs. [4, 9–11]), blamed the defects in ground models for the biased prediction. Therefore, besides physical limitations of 1D modelling, imperfect KiK-net velocity profiles might be another main culprit curtailing the efficacy of 1DSH modelling.

6.2. Effectiveness of HVSR correction

Given the major problems with 1DSH modelling, it entails a search for an alternative site-specific approach to quantify site response in some applications. HVSR after the correction has shown an encouraging advantage over TTF_{Strata} (Fig. 9). In this section, however, we discuss its effectiveness in absolute terms. Fig. 11 displays the histograms of various GoF measures for corrected HVSR, i.e., pSBSR. The results for TTF_{Strata} are superimposed. All closeness quantities are calculated over the frequencies from f_0 to 25 Hz (Ⓔ Table S3). For all relative GoF metrics, i.e., r, ρ, τ, and d, we set the threshold value of a good match at 0.6 as adopted by Thompson et al. [1].

Fig. 11a–d illustrates that pSBSR achieves a good match at 81%, 76%, 48% and 80% of the 90 KiK-net sites using r, ρ, τ, and d, respectively. This demonstrates that the corrected HVSR can reproduce the observed amplification function in either shape or amplitude at most of our investigated stations. Comparing with TTF_{Strata}, the empirical correction realizes a surge in success rate by 40–50% depending on the GoF metric in use. Fig. 11a devidence the remarkable effectiveness of the empirical correction, as well as its considerable advantage over 1DSH modelling.

As aforementioned, we define a good fit using a single GoF indicator. However, as pointed out by Tao and Rathje [8]; this could lead to misjudgments because one GoF metric alone cannot well evaluate the level of the match in both amplitude and spectral shape. Considering that a good match in both needs to be achieved in practice, thus

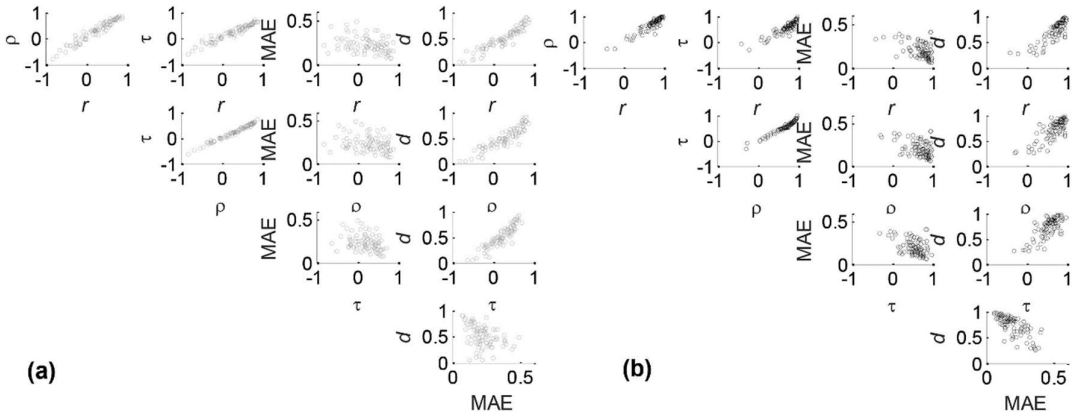

Figure 12. *Scatterplots of various GoF metrics (Pearson's r, Spearman's ρ, Kendall's τ, Index of Agreement d, and Mean Absolute Error MAE) for (a) TTF_{Strata} vs. ETF; and (b) pSBSR vs. ETF.*

we use dual GoF parameters, one for amplitude and the other for spectral shape, to define a good match. Correlations between GoF indicators are examined in Fig. 12. r is strongly correlated with ρ and τ, and thus r is selected as the GoF measure for spectral shape. For amplitude, both d and MAE are adopted. In addition, the threshold value is also varying. Hence, a good match is defined using two GoF pairs, (r, d) and (r, MAE) with varying thresholds.

Table 3 lists the percentages of good match sites. Albeit different definitions, the success rate of TTF_{Strata} in matching ETF is consistently below one third (out of 90 sites) whereas it is no less than 50% for pSBSR. For the empirical correction method, both the assumption $HVSR_b$ = 1.0 and the use of generic correction spectra are responsible for the remaining unsuccessful sites. Nevertheless, Table 3 consolidates that the HVSR-based empirical correction approach has a considerable advantage over 1DSH modelling.

Fig. 11e depicts the histograms of MAE (in log10 scales). The mean values of MAE for TTF_{Strata} is 0.23 which is larger than that for pSBSR, 0.20. To realize a direct comparison, we present the residuals of TTF_{Strata}, as well as pSBSR about ETF against continuous frequencies in Fig. 13a and b, respectively. Frequencies are normalized to f_0. Fig. 13c presents the average and standard deviation of residuals in the frequency domain without normalization. It can be seen from Fig. 13a that there exists an obvious negative bias at $f/f_0 = 1.0$, indicating that TTF_{Strata} tends to overestimate the fundamental-mode site response. A similar trend has been reported in many previous studies (e.g., Kaklamanos and Bradley [4]). This underestimation reflects the inability of TTF_{Strata} to properly model the downgoing wave effects which are less significant due to multidimensional scattering (e.g., Ref. [1,8]), than what 1DSH modelling predicts.

However, at relatively high frequencies (Fig. 13a and c), TTF_{Strata} on average substantially underestimates the observed amplification, for which Kaklamanos and Bradley [4] blamed the simple and coarse V_S profiles used in numerical modellings. Very recently, Kaklamanos et al. [43] demonstrated that applying depth-dependent V_S gradients within layers could effectively mitigate the underpredictions at sites with coarse profiles. Meanwhile, the high-frequency underestimations might also be related to damping values adopted in simulations. However, Cabas et al. [31], and Xu et al. [44] revealed that Q-based damping estimates using Eqs. (8) and (9) tended to underestimate the kappa (κ_0)-consistent damping ratios based on field observations at KiK-net sties. According to their findings, the Q-based damping should be increased to be compatible with field estimates of attenuation, which would predictably lead to larger high-frequency underestimations. Thus, more works are needed to better constrain site-specific attenuation parameters.

For pSBSR, Fig. 13b and c illustrate a negative bias at high frequencies, implying that pSBSR tends to overestimate observed amplifications. This overestimation is mainly due to the assumption, $HVSR_b$ = 1.0 (Eq. (11)) used in the correction. $HVSR_b$ is larger than 1.0 at some frequencies (Fig. 5a), thus the correction (Eq. (14)) overcompensates HVSR, inducing an overestimation of pSBSR to ETF (Fig. 13c). However, this bias can be readily eliminated by considering the fluctuation of $HVSR_b$ with frequency, rather than a constant, in the correction procedure. Even so, pSBSR has a considerably less amount of bias and lower level of uncertainty than TTF_{Strata}.

Table 3
Success rates of TTFStrata and pSBSR in reproducing SBSR under different definitions of "good match".

Estimation	$r > 0.60$	$r > 0.60$ $d > 0.60$	$r > 0.65$ $d > 0.65$	$r > 0.60$ MAE < 0.25	$r > 0.65$ MAE < 0.20
TTF_{Strata}	27%	27%	18%	22%	14%
pSBSR	81%	76%	68%	62%	50%

6.3. Regional dependency and soil nonlinearity

Fig. 14a and b presents three example applications of the empirical correction at sites that are not included in the dataset based on which the average correction spectra are developed. Their V_S structures are shown in Fig. 14c.

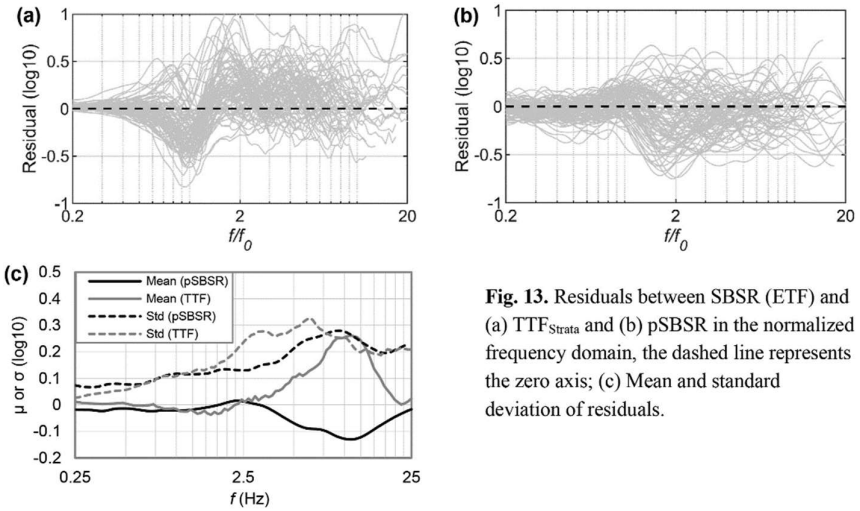

Fig. 13. Residuals between SBSR (ETF) and (a) TTF_{Strata} and (b) pSBSR in the normalized frequency domain, the dashed line represents the zero axis; (c) Mean and standard deviation of residuals.

Figure 14. *Residuals between SBSR (ETF) and (a) TTF_{Strata} and (b) pSBSR in the normalized frequency domain, the dashed line represents the zero axis; (c) Mean and standard deviation of residuals.*

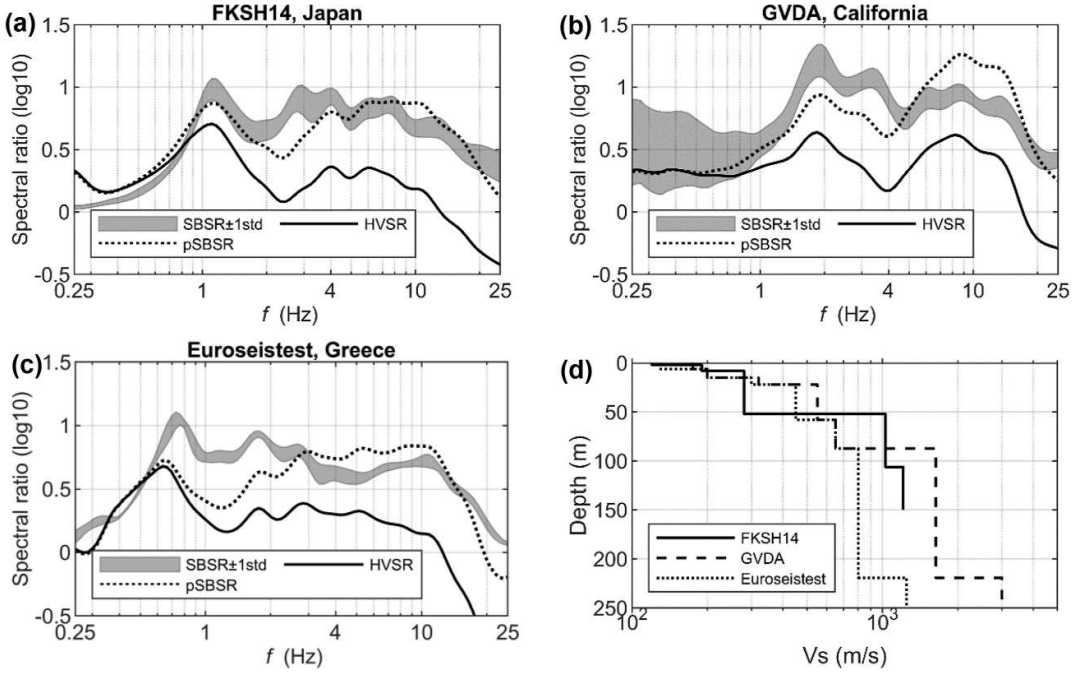

Figure 13. *Application of the empirical correction at (a) KiK-net site FKSH14 in Japan, (b) GVDA in California, and (c) Euroseistest in Greece. V_S profiles are displayed in (c) with a zoom-in view on the shallow parts. The three sites are not among the selected 90 sites.*

The first one is a soft site from KiK-net, FKSH14 with $V_{S30} = 237$ m/s and $f_{0v} = 3.5$ Hz (Cluster 5). We correct its HVSR using the $\langle SBSR_v \rangle$ for Cluster 5 (Fig. 6c) and then compare the corrected HVSR (pSBSR) with the observed amplification SBSR (ETF) in Fig. 14a. GoF statistics are $r = 0.69$, $d = 0.81$, and MAE $= 0.12$, indicating a good match in both shape and amplitude using dual indicators (Table 3). The second application site is GVDA in Southern California with $f_{0v} = 3.1$ Hz (Cluster 5). V_{S3} is estimated to be 316 m/s from the profile proposed by Bonilla et al. [36]. We compare its pSBSR and SBSR (ETF) in Fig. 14b. GoF results are $r = 0.52$, $d = 0.71$ and MAE $= 0.22$, which does not qualify as a good match (Table 3). The third one is the Euroseistest site which is in the tectonically active Mygdonian basin, Greece (Table S1). This site has a $V_{S30} = 224$ m/s [45] and $f_{0v} = 1.3$ Hz (Cluster 5). GoF results for pSBSR and SBSR (Fig. 14c) are $r = 0.75$, $d = 0.83$ and MAE $= 0.17$, indicating a good match. MATLAB code and other relevant files for this site are provided as Supplemental Materials (Ⓔ File Euroseistest, see Data and Resources). Based on these above applications, we stress that the correction spectra in this research are developed for Japan, and we caution about their direct application to other regions. Region-specific $\langle SBSR_v \rangle$ might need to be

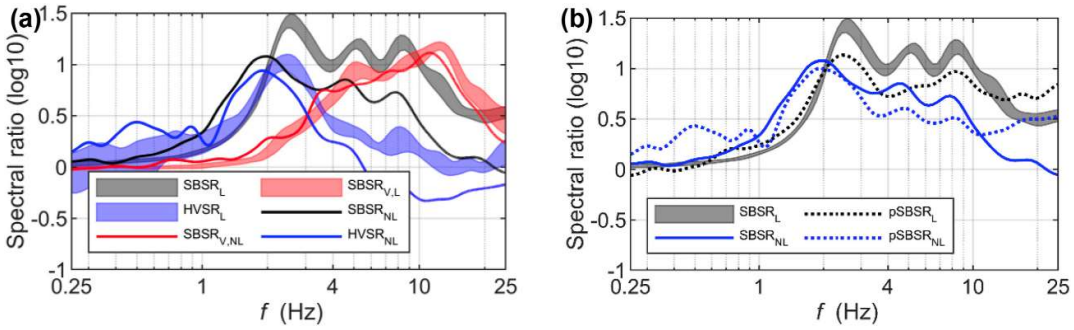

Figure 15. *Application of the empirical correction in estimating nonlinear site response at KiK-net site IBRH11. (a) Empirical spectral ratios using weak and strong ground-motions that correspond to linear (L) and nonlinear (NL) site response; and (b) Empirical corrections to weak- and strong-motion HVSRs by the correction spectrum for Cluster 4 (Fig. 6c). The strong-motion record is from the mainshock of the 2011 Tōhoku earthquake ($M_w = 9.0$). The widths of the ratios' curves represent ± one standard deviation.*

constructed.

Results in sections above are based on relatively weak ground-motion records that were not significantly affected by soil nonlinearity and thus might not applicable to strong motions under which nonlinearity of shallow sediments modifies site responses. The proposed empirical correction approach (Eq. (14)) uses a set of predefined vertical correction spectra ⟨$SBSR_v$⟩ (Fig. 6c) and a site- and event-specific HVSR. The former can be affected by soil nonlinearity but to a much lesser extent than its horizontal counterpart (e.g., Tsai and Liu [46]). In contrast, the latter, i.e., HVSR, is sensitive to ground-motion intensities and thus is utilized in some studies (e.g., Wen et al. [47]), to identify nonlinear site effects. Therefore, Eq. (14) might be also viable under strong ground-motions in certain circumstances even though ⟨$SBSR_v$⟩ is derived from relatively weak motions.

To test the above hypothesis, we investigate the site response at one of our selected KiK-net site IBRH11 (V_{S30} = 242 m/s, Fig. 4a) during the mainshock of the 2011 M_w = 9.0 Tōhoku earthquake, Japan. Corresponding spectral ratios are plotted in Fig. 15a in which average spectral ratios derived previously from weak motions (Fig. 4c) are also superimposed. Comparing the spectral ratios under weak and strong ground shakings, it clearly indicates both SBSR and HVSR are affected by soil nonlinearity during the mainshock whereas $SBSR_v$ only varies to a much lesser degree. Next, we apply the empirical correction approach to strong-motion HVSR. The corrected HVSR ($pSBSR_{NL}$) by the correction spectrum for Cluster 4 (Fig. 6c) is compared with the observed amplification ($SBSR_{NL}$) during the mainshock in Fig. 15b. The corrected weak-motion HVSR is also displayed in Fig. 15b.

Comparing with $pSBSR_L$, $pSBSR_{NL}$ exhibits typical nonlinear features, namely a reduction in peak values and a shift in peak frequencies. In addition, the GoF statistics ($pSBSR_{NL}$ vs. $SBSR_{NL}$) for strong motion case are $r = 0.68$, $\rho = 0.76$, $\tau = 0.57$, $d = 0.72$ and MAE = 0.21, which quantify as a good match ($r > 0.6$ and $d > 0.6$) and are only slightly worse than these ($pSBSR_L$ vs. $SBSR_L$) under weak motions: $r = 0.73$, $\rho = 0.77$, $\tau = 0.60$, $d = 0.76$ and MAE = 0.20. Thus, Fig. 15 suggests that the empirical correction approach has implicitly accounted for soil nonlinearity to some extent due to its use of site- and event-specific HVSR (Eq. (14)). However, the fit between $pSBSR_{NL}$ and $SBSR_{NL}$ (Fig. 15b) is less well at high frequencies (> ~10 Hz). In this frequency range, HVSR is overcompensated by weak-motion ⟨$SBSR_v$⟩ whereas site-specific $SBSR_v$ (Fig. 15a) undergoes a certain degree of nonlinearity during the mainshock. This application at IBRH11 corroborates the validity of the proposed approach under strong motions given that the vertical site response is, more or less, in the linear domain. Nonetheless, such applications at more sites are needed in future studies prior to drawing statistically meaningful conclusions on its reliability under strong motions.

In this study, we take SBSR as observed amplifications, and thus the corrected HVSR or pSBSR is the amplification referenced to site-specific borehole rock conditions. However, the example (station MYGH04) shown in Fig. 3b suggests that SBSR is close to the GIT-based amplifications [23] referenced to outcropping conditions (V_S = 3.45 km/s). This implies that the effects of downgoing waves and the impacts of bedrock property variability may be insignificant at our selected sites given the high V_S and large depth of borehole bedrock. In addition, the existence of prominent vertical amplifications is the underlying reason for the underestimation of HVSR to the actual horizontal amplifications. However, the vertical amplification (e.g., Ref. [46,48], has received far less attention than its horizontal counterpart. Thus, studies are needed to gain a better understanding of the amplification of vertical ground-motions, e.g., the optimal site proxies for its characterization and regional variation.

7. Conclusions

It is known that horizontal-to-vertical spectral ratio (HVSR) shares a similar shape with the horizontal S-wave amplification spectrum but underestimates its amplitudes at relatively high frequencies due to the strong presence of amplifications in the vertical direction. In this

research, we selected earthquake recordings at 90 KiK-net surface-downhole pairs at which no large velocity variation was identified below the downhole sensor and thus of which surface-to-borehole spectral ratio (SBSR) could be taken as its empirical transfer function (ETF). Motivated by the approach originally proposed by Kawase et al. [17]; we developed generic vertical amplification spectra, which were categorized primarily by vertical site frequency (f_{0v}) via k-means clustering. The generic spectra were then utilized to correct HVSR amplitudes for a direct horizontal amplification estimation. The empirical correction preserves the shape of HVSR but compensates HVSR amplitudes by generic vertical amplification spectra.

The main purpose of this study is to compare the performance of this novel application of earthquake HVSR with that of 1DSH modellings in reproducing observed linear amplifications (ETF) at the 90 KiK-net sites. The 1DSH-based theoretical transfer function (TTF) at each site was the median of 30 profiles generated through randomization about the base-case model. Five goodness-of-fit (GoF) statistics were adopted to gauge the closeness of various amplification predictions (i.e., TTF and corrected HVSR) to observation (ETF) in a broad frequency range from horizontal site frequency (f_0) to 25 Hz. These GoF quantities include four relative metrics (Pearson's r, Spearman's ρ, Kendall's τ, and Index of Agreement d) and one absolute measure (Mean Absolute Error, MAE). To assess the closeness in both spectral shape and amplitude, we utilized dual GoF measures to define a "good match".

Key conclusions of this research are:

(1) The 1DSH-based GRA exhibits poor performance in predicting amplifications, achieving a "good match" at less than one-third of the 90 examined sites.
(2) The novel HVSR-based empirical correction has a considerably higher success rate than GRA, realizing a "good match" at the majority of the 90 sites.

The relatively low success rate of 1D ground response analyses may be primarily attributed to violations of 1DSH assumptions and inaccurate KiK-net PS logging data. However, this study shows quite promising results for utilizing the empirically corrected HVSR to approximate site-specific horizontal amplifications. The novel application of HVSR has a considerably better performance than the widely adopted 1DSH modelling in assessing site effects. In addition, the empirical correction to HVSR does not need velocity and damping models which are required by GRA, thus it is rather cost-effective. The findings of this study are based on Japanese ground-motion recordings that are not significantly affected by soil nonlinearity, and thus the HVSR-based empirical correction should be extensively tested in other regions and under strong motions in future investigations. Propagation of uncertainty also needs to be addressed. This study serves as the first step in this direction.

Data and resources

Earthquake recordings and velocity profiles at KiK-net stations were downloaded from the http://www.kyoshin.bosai.go.jp (last accessed on June 05, 2018). GVDA and Euroseistest recordings were downloaded from http://nees.ucsb.edu/facilities/GVDA (last accessed on April 01, 2020). Supplemental content includes an interactive 3D plot (Fig. 6a), three tables (Tables S1–S3) and two zip files (Spectral Ratios and Euroseistest). Table S1 tabulates the parameters of selected earthquakes at the GVDA and Euroseistest. Table S2 contains the correction spectra for the five clusters. Table S3 lists the values of all GoF statistics computed over the frequency range from f_0 to 25 Hz at the 90 selected sites. The zip file "Spectral Ratios" includes plots showing different amplification estimates (HVSR, TTFStrata, pSBSR, and SBSR) and GoF statistics, as well as plots of V_S-depth structures published by NIED at the 90 selected sites. The other zip file Euroseistest contains a MATLAB code, as well as its input and output files for this site.

Credit author statement

Chuanbin Zhu: Conceptualization, Methodology, Investigation, Writing - Original Draft, Revision.

Marco Pilz: Data Curation, Writing - Review & Editing, Supervision, Revision.

Fabrice Cotton: Writing - Review & Editing, Project administration, Funding acquisition, Revision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Acknowledgments

We thank the National Research Institute for Earth Science and Disaster Prevention (NIED), Japan, and the Garner Valley Downhole Array (GVDA) for making the data available. This work is supported by the Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA) project funded by the EU Horizon 2020 Programme under Grant Agreement Number 730900. We also want to thank Prof. Hiroshi Kawase and two anonymous reviewers who provided very constructive feedback on this article, as well as Dr. Zhandong Su for his assistance with figures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.soildyn.2020.106301.

Acronyms

GRA	ground response analysis based on 1DSH assumptions
HVSR	horizontal-to-vertical spectral ratio at the ground surface
HVSRb	horizontal-to-vertical spectral ratio at a borehole
SBSR	surface-to-borehole spectral ratio in the horizontal direction
$SBSR_v$	surface-to-borehole spectral ratio in the vertical

	direction
$\langle SBSR_v \rangle$	average surface-to-borehole spectral ratio in the vertical direction
pSBSR	pseudo surface-to-borehole spectral ratio in the horizontal direction
GIT	general inversion technique
TTF	theoretical transfer function
ETF	empirical transfer functions
FAS	Fourier amplitude spectrum
IC	impedance contrast
GoF	goodness-of-fit
MAE	mean absolute error
GVDA	Garner Valley downhole array

References

[1] Thompson EM, Baise LG, Tanaka Y, Kayen RE. A taxonomy of site response complexity. Soil Dyn Earthq Eng 2012; 41: 32–43.

[2] Okada Y, Kasahara K, Hori S, Obara K, Sekiguchi S, Fujiwara H, Yamamoto A. Recent progress of seismic observations networks in Japan – Hi-net, F-net, K-net and KiK-net. Earth Planets Space 2004; 56: xv–xxviii.

[3] Aoi S, Kunugi T, Nakamura H, Fujiwara H. Deployment of new strong motion seismographs of K-NET and KiK-net. In: Akkar S, et al., editors. Earthquake data in engineering seismology, geotechnical, geological, and earthquake engineering, vol. 14; 2011. https://doi.org/10.1007/978-94-007-0152-6_12.

[4] Kaklamanos J, Bradley BA. Challenges in predicting seismic site response with 1D analyses: conclusions from 114 KiK-net vertical seismometer arrays. Bull Seismol Soc Am 2018; 108: 2816–38.

[5] Laurendeau A, Bard PY, Hollender F, Perron V, Foundotos L, Ktenidou OJ, Hernandez B. Derivation of consistent hard rock (1000 < V_S < 3000 m/s) GMPEs from surface and down-hole recordings: analysis of KiK-net data. Bull Earthq Eng 2018; 16: 2253–84.

[6] Pilz M, Cotton F. Does the 1D 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 2019; 2: 883–905.

[7] Afshari K, Stewart JP. Insights from California vertical arrays on the effectiveness of ground response analysis with alternative damping models. Bull Seismol Soc Am 2019; 109: 1250–64.

[8] Tao Y, Rathje E. Taxonomy for evaluating the site-specific applicability of one-dimensional ground response analysis. Soil Dyn Earthq Eng 2020; 128: 105865.

[9] Cramer CH. Weak-motion observations and modeling for the Turkey flat, U.S. site-effects test area near Parkfield, California. Bull Seismol Soc Am 1995; 85: 440–51.

[10] Real CR, Shakal AF, Tucker BE. Turkey Flat, U.S.A. site effects test area: anatomy pf a blind ground-motion prediction test. In: Third international symposium on the effects of surface geology on seismic motion, Grenoble, France, 30 August-1 September 2006 paper number: KN 3; 2006.

[11] Kwok AOL, Stewart JP, Hashash YMA. Nonlinear ground-response analysis of Turkey Flat shallow stiff-soil site to strong ground motion. Bull Seismol Soc Am 2008; 98: 331–43.

[12] Zhu C, Thambiratnam D, Gallage C. Inherent characteristics of 2D alluvial formations subjected to in-plane motion. J Earthq Eng 2017. https://doi.org/ 10.1080/13632469.2017.1387199.

[13] Zhu C, Chávez-García FJ, Thambiratnam D, Gallage C. Quantifying the edge-induced seismic aggravation in shallow basins relative to the 1D SH modelling. Soil Dyn Earthq Eng 2018; 115: 402–12.

[14] Nakamura Y. A method for dynamic characteristics estimation of subsurface using 549 microtremor on ground surface. Q Rep Railway Tech Res Inst 1989; 30: 25–33.

[15] Lermo J, Chavez-Garcia FJ. Site effects evaluation using spectral ratios with only one station. Bull Seismol Soc Am 1993; 83: 1574–94.

[16] Bard PY. Microtremor measurements: A tool for site effect estimation? In: The effects of surface geology on seismic motion. Rotterdam: Balkema; 1999.

[17] Kawase H, Nagashima F, Nakano K, Mori Y. Direct evaluation of S-wave amplification factors from microtremor H/V ratios: Double empirical corrections to "Nakamura" method. Soil Dyn Earthq Eng 2019; 126: 105067.

[18] Zhu C, Pilz M, Cotton F. Which is a better proxy, site period or depth to bedrock, in modelling linear site response in addition to the average shear-wave velocity? Bull Earthq Eng 2019. https://doi.org/ 10.1007/s10518-019-00738-6.

[19] Dawood HM, Rodriguez-Marek A, Bayless J, Goulet C, Thompson E. A flatfile for the KiK-net database processed using an automated protocol. Earthq Spectra 2016; 32: 1281–302.

[20] Fujimoto K, Midorikawa S. Relationship between average shear-wave velocity and site amplification inferred from strong motion records at nearby station pairs. J Jap Assoc Earthq Eng 2006; 6: 11–22. in Japanese with English abstract.

[21] Konno K, Ohmachi T. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bull Seismol Soc Am 1998; 88: 228–41.

[22] Yamanaka H, Chimoto K, Tsuno S, Dhakal YP, Amrouche M, Yamada N, Fukumoto S, Eto K. Estimation of S-wave velocity profiles and site amplification around the K-NET Tsukidate station, Miyagi Prefecture, with Reference to Large PGA During the 2011 off Pacific Coast of Tohoku Earthquake, Japan. J Disaster Res 2012; 7: 682–92.

[23] Nakano K, Matsushima S, Kawase H. Statistical Properties of Strong Ground Motions from the Generalized Spectral Inversion of Data Observed by K-NET, KiK-net, and the JMA Shindokei Network in Japan. Bull Seismol Soc Am 2015; 105: 2662–80.

[24] Andrews DJ. Objective determination of source parameters and similarity of earthquakes of different size. In: Das S, Boatwright J, Scholz CH, editors. Earthquake source mechanics. Washington, D.C: American Geophysical Union; 1986.

[25] Kottke AR, Wang X, Rathje EM. Technical manual for strata, geotechnical engineering center. Austin: Department of Civil, Architectural, and Environmental Engineering, University of Texas; 2013.

[26] Teague DP, Coxb BR, Rathje EM. Measured vs. predicted site response at the Garner Valley Downhole Array considering shear wave velocity uncertainty from borehole and surface wave methods. Soil Dyn Earthq Eng 2018; 113: 339–55.

[27] Toro G. Probabilistic models of the site velocity profiles for generic and site-specific ground-motion amplification studies. Upton, N.Y.: Brookhaven National Laboratory; 1995. p. 147. Technical Report No. 779574.

[28] Li B, Pandy MD, Dai K, Lu Y, Zhou Y, Shi Y, Ding Z. Effects of soil parameter variabilities on the estimation of ground-motion amplification factors. Earthq Spectra 2019; 35: 907–28.

[29] Wang SY, Shi Y, Jiang WP, Yao EL, Miao Y. Estimating site fundamental period from shear-wave velocity profile. Bull Seismol Soc Am 2018; 108: 3431–45.

[30] Campbell KW. Estimates of shear-wave q and k0 for unconsolidated and semi-consolidated sediments in Eastern North America. Bull Seism Soc Am 2009; 99: 2365–92.

[31] Cabas A, Rodriguez-Marek A, Bonilla LF. Estimation of site-specific kappa (κ_0)- consistent damping values at KiK-net sites to assess the discrepancy between laboratory-based damping models and observed attenuation (of seismic waves) in the field. Bull Seismol Soc Am 2017; 107: 2258–71.

[32] Rathje EM, Kottke AR, Trent WL. Influence of Input Motion and Site Property Variabilities on Seismic Site Response Analysis. J Geotech Geoenviron 2010; 136: 607–19.

[33] Willmott CJ. On the validation of models. Phys Geogr 1981; 2: 184–94.

[34] Legates DR, McCabe GJ. Evacuating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour Res 1999; 35: 233–41.

[35] King JL, Tucker BE. Observed variations of earthquake motion over a sediment filled valley. Bull Seism Soc Am 1984; 74: 137–52.

[36] Bonilla LF, Steidl JH, Gariel JC, Archuleta RJ. Borehole Response

Studies at the Garner Valley Downhole Array, Southern California. Bull Seismol Soc Am 2002; 92: 3165–79.
[37] Rong M, Fu LY, Wang Z, Li X, Carpenter NS, Woolery EW, Lyu Y. On the amplitude discrepancy of HVSR and site amplification from strong-motion observations. Bull Seismol Soc Am 2017; 107: 2873–84.
[38] Parolai S, Richwalski SM. The importance of converted waves in comparing H/V and RSM site response estimates. Bull Seism Soc Am 2004; 94: 304–13.
[39] Wakamatsu K, Kubo S, Matsuoka M, Hasegawa K, Sugiura M. Japan engineering geomorphologic classification Map with CD-ROM database. Tokyo: University of Tokyo Press; 2005.
[40] Kotha SR, Cotton F, Bindi D. A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions. Soil Dyn Earthq Eng 2018; 110: 318–29.
[41] Ji K, Wen R, Ren Y, Dhakal Y. Nonlinear seismic site response classification using K-means clustering algorithm: Case study of the September 6, 2018 Mw6.6 Hokkaido Iburi-Tobu earthquake, Japan. Soil Dyn Earthq Eng 2019; 128: 105907.
[42] Zhu C, Cotton F, Pilz M. Testing the depths to 1.0 and 2.5 km/s velocity isosurfaces in a velocity model for Japan and implications for ground motion modelling? Bull Seism Soc Am 2019. https://doi.org/10.1785/0120190016.
[43] Kaklamanos J, Bradley BA, Moolacattu AN, Picard BM. Physical hypotheses for adjusting coarse profiles and improving 1D site-response estimation assessed at 10 KiK-net sites. Bull Seismol Soc Am 2020. https://doi.org/10.1785/0120190263.
[44] Xu B, Rathje EM, Hashash Y, Stewart J, Campbell K, Silva WJ. k0 for soil sites: Observations from KiK-net sites and their use in constraining small-strain damping profiles for site response analysis. Earthq Spectra 2020; 36: 111–37.
[45] Raptakis D, Chavez-Garcia FJ, Makra K, Pitilakis K. Site effects at Euroseistest - I. Determination of the valley structure and confrontation of observations with 1D analysis. Soil Dyn Earthq Eng 2000; 19: 1–22.
[46] Tsai C-C, Liu H-W. Site response analysis of vertical ground motion in consideration of soil nonlinearity. Soil Dyn Earthq Eng 2017; 102: 124–36.
[47] Wen KL, Chang TM, Lin CM, Chiang HJ. Identification of nonlinear site response using the H/V spectral ratio method. Terr Atmos Ocean Sci 2006; 17: 533–46.
[48] Yang J, Sato T. Interpretation of Seismic Vertical Amplification Observed at an Array Site. Bull Seism Soc Am 2000; 90: 275–85.