



HAL
open science

Effects of Nonlinear Soil Behavior on Kappa (κ): Observations from the KiK-Net Database

Chunyang Ji, Ashly Cabas, Luis Fabian Bonilla, Céline Gelis

► To cite this version:

Chunyang Ji, Ashly Cabas, Luis Fabian Bonilla, Céline Gelis. Effects of Nonlinear Soil Behavior on Kappa (κ): Observations from the KiK-Net Database. Bulletin of the Seismological Society of America, 2021, 111 (4), pp.2138 - 2157. 10.1785/0120200286 . hal-04402168v1

HAL Id: hal-04402168

<https://univ-eiffel.hal.science/hal-04402168v1>

Submitted on 18 Jan 2024 (v1), last revised 30 Jan 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/350640355>

Effects of Nonlinear Soil Behavior on Kappa (κ): Observations from the KiK-Net Database

Article in *Bulletin of the Seismological Society of America* · April 2021

DOI: 10.1785/0120200286

CITATIONS

14

READS

542

4 authors:



Chunyang Ji

North Carolina State University

8 PUBLICATIONS 38 CITATIONS

SEE PROFILE



Ashly Cabas

North Carolina State University

37 PUBLICATIONS 171 CITATIONS

SEE PROFILE



Luis Fabian Bonilla

Université Gustave Eiffel

140 PUBLICATIONS 3,693 CITATIONS

SEE PROFILE



Celine Gelis

Institut de Radioprotection et de Sûreté Nucléaire (IRSN)

89 PUBLICATIONS 1,409 CITATIONS

SEE PROFILE

Abstract

22
23 Soil nonlinear behavior is often triggered in soft sedimentary deposits subjected to strong ground
24 shaking and has led to catastrophic damage to civil infrastructure in many past earthquakes.
25 Nonlinear behavior in soils is associated with large shear strains, increased material damping ratio
26 and reduced stiffness. However, most investigations of the high-frequency spectral decay
27 parameter κ , which captures attenuation, have focused on low-intensity ground motions inducing
28 only small shear strains. Because studies of the applicability of the κ -model when larger
29 deformations are induced are limited, this paper investigates the behavior of κ (both κ_r per record
30 and site-specific κ_0 estimates) beyond the linear-elastic regime. Twenty stations from the KiK-net
31 database, with time-average shear wave velocities in the upper 30 m between 213 and 626 m/s, are
32 used in this study. We find that the classification scheme used to identify ground motions that
33 trigger soil nonlinear behavior biases estimates of κ_0 in the linear and nonlinear regimes. A hybrid
34 method to overcome such bias is proposed considering proxies for in situ deformation (via the
35 shear strain index) and ground shaking intensity (via peak ground acceleration). Our findings show
36 that soil nonlinearity affects κ_r and κ_0 estimates, but this influence is station-dependent. Most κ_0
37 at our sites had a 5-20% increase at the onset of soil nonlinear behavior. Velocity gradients and
38 impedance contrasts influence the degree of soil nonlinearity and its effects on κ_r and κ_0 . Moreover,
39 we observe that other complexities in the wave propagation phenomenon (e.g., scattering and
40 amplifications in the high-frequency range) impose challenges to the application of the κ_0 -model,
41 including the estimation of negative values of κ_r .

42

Introduction

43
44 The anelastic attenuation of seismic waves as they travel through sedimentary deposits is a
45 function of the deformations induced, which in turn depend on the material properties (e.g.,
46 plasticity of the soil) and the intensity of the ground shaking. Material damping ratio, ξ , is
47 commonly used in geotechnical earthquake engineering to quantify viscous and anelastic energy
48 dissipation in soils subjected to dynamic loading. Empirical models of ξ often have a constant
49 minimum value (known as the minimum shear-strain damping, ξ_{\min}) for small shear strains
50 considered in the linear-elastic regime (e.g., Darendeli, 2001). Yet, values of ξ increase as larger
51 shear strains are induced in soil deposits by stronger ground excitations (Idriss et al. 1978, Seed et
52 al. 1986, Darendeli 2001). The characterization of ξ across a wide range of strains is essential to
53 model the effects of local soil conditions on earthquake ground motions.

54
55 The high-frequency spectral decay parameter κ was introduced by Anderson and Hough (1984)
56 based on the Fourier spectrum characteristics of a ground motion's shear-wave window recorded
57 directly in the field, which makes it an observable parameter that quantifies total attenuation (e.g.,
58 energy dissipation caused by scattering and anelasticity). Estimates of κ have proven useful in
59 multiple applications, from stochastic modeling of ground motions (Boore 2003) to the
60 development of host-to-target adjustments of ground motion models (e.g., Campbell 2003, Al Atik
61 et al. 2014). The site-specific, distance-independent component of κ , known as κ_0 (Anderson 1991),
62 is defined as a site parameter that captures the attenuation due to the propagation of seismic waves
63 through near-surface materials. The relationship between κ_0 and ξ_{\min} has been investigated in
64 previous studies (e.g., Cabas et al. 2017; Ktenidou et al. 2015; Xu et al. 2019) for weak motion
65 data, but the quantification of κ and κ_0 beyond the linear-elastic regime remains unsolved.

66

67 Most studies on individual estimates of κ (i.e., the value measured from the observed Fourier
68 Amplitude Spectra (FAS) per record following the Anderson and Hough (1984) approach),
69 hereafter referred to as κ_r , and its site-specific component κ_0 have used ground motion records that
70 do not trigger nonlinear behavior at the sites of interest or that are not considered influenced by
71 the site's nonlinear response (e.g., Ktenidou et al. 2013, Van Houtte et al. 2011, Laurendeau et al.
72 2013, Edwards et al. 2015, Perron et al. 2017, Xu et al. 2019). However, nonlinear soil behavior
73 has often been responsible for increasing the damage potential of strong ground motions in past
74 earthquakes (e.g., Darragh and Shakal, 1991, Trifunac and Todorovska, 1994, Bonilla et al., 2011,
75 Rong et al., 2016). Understanding near-surface attenuation effects in the nonlinear regime is then
76 necessary for a thorough assessment of seismic hazards and risks imposed to civil infrastructure
77 (Anderson and Hough 1984). Hence, this paper investigates the relationships among κ , shear
78 strains and ground motion intensity to understand the behavior of κ at the onset of nonlinear soil
79 behavior.

80

81

Background

82 The first paper that attempted to connect soils' nonlinear response and κ_r was conducted by Yu et
83 al. (1992), where the authors compared two simulated records: one from a linear site response
84 analysis, and the other from a time-domain nonlinear site response analysis. Yu et al. (1992)
85 observed that the value of κ_r estimated with the Anderson and Hough (1984) approach and
86 corresponding to the motion affected by soil nonlinearity was smaller. However, later studies
87 found a positive correlation between κ_r (and κ_0), strain amplitudes and the intensity of ground
88 shaking (e.g., Durward et al 1996; Lacave-Lachet et al 2000; Dimitriu et al. 2001).

89

90 Durward et al. (1996) found that κ_r values were a function of peak ground velocity (PGV, varying
91 from 0.01 to 1 m/s), which was used as a proxy for deformation. Values of κ_r were computed for
92 more than 60 records observed at 23 sites in the Imperial Valley, California based on the
93 acceleration spectrum approach (Anderson and Hough, 1984). Durward et al. (1996) hypothesized
94 that soil nonlinearity had affected κ_r because higher κ_r values correlated well with higher PGVs.
95 Moreover, Lacave-Lachet et al. (2000) analyzed ground motions from the 1995 Kobe earthquake
96 in Japan (i.e., the main shock and aftershocks), and found that κ_r values increased with increasing
97 peak ground acceleration (PGA). Hence, Lacave-Lachet et al. (2000) proposed to use κ_r to detect
98 the onset of soil nonlinearity. Dimitriu et al. (2001) investigated the dependency between site-
99 specific κ_0 and ground shaking intensity. Values of κ_r for 23 ground motions (i.e., 46 horizontal
100 components with values of κ_r reported for each individual component) were computed at two
101 adjacent sites in Lefkas, western Greece, based on the acceleration spectrum method. Dimitriu et
102 al. (2001) provided evidence that κ_0 was a proxy for soil nonlinearity based on the observed
103 dependency between κ_0 and ground shaking amplitudes, which were represented by mean
104 horizontal acceleration in the S-wave window (MGA), PGA, and PGV. Positive correlations were
105 found between the 46 κ_0 values and MGA, PGA, and PGV in log-scale, while a negative
106 correlation was observed between κ_0 and the site dominant-resonance frequency. However, Van
107 Houtte et al. (2014) identified an opposite correlation between estimates of κ_0 (computed as the
108 individual measured κ_r with epicentral distance less than 30 m) and PGA at hard sites (i.e., with
109 V_{s30} varying from 422 to 1073 m/s) using ground motions from the 2010-2011 Canterbury
110 earthquake sequence in New Zealand. The authors suggested further investigations to understand
111 the associated physical mechanism.

112

113 There are still few and contradicting observations of the effects of nonlinearity on κ_r and κ_0
114 estimates (Ktenidou et al. 2015). Previous studies only considered a very limited database of
115 ground motions. This paper takes advantage of the unique Japanese Kiban-Kyoshin network (KiK-
116 net), which is rich in high-quality ground motions, to further investigate the effects of soil
117 nonlinearity on κ_r and κ_0 . More specifically, we explore the dependence of κ_r and κ_0 on ground
118 shaking intensity (i.e., weak, moderate or strong ground motions as parameterized by PGA), and
119 on the level of shear strains induced in near-surface materials at 20 KiK-net stations. First, we
120 present the conceptual basis for the relationship between κ , shear strains and ground motion
121 intensity. Then, we describe our database and methods, starting with the identification of an
122 appropriate classification scheme for linear and nonlinear ground motions. The analysis of the
123 effects of ground shaking amplitudes on κ_r at each study site follows. Lastly, we compare the ratio
124 of nonlinear and linear site-specific κ_0 across all selected stations to assess the variation of near-
125 surface attenuation estimates from the linear-elastic to the nonlinear regime.

126

127 **Conceptual basis for the interpretation of κ beyond the linear-elastic regime**

128 The induced strain level in a given soil layer is a function of the material properties, and the
129 amplitude and frequency content of the incoming wavefield at the site. Stronger ground shaking
130 results in larger-strain responses, which produce an increase in material damping ratio (in
131 combination with a reduction of shear modulus). The short wavelength of high frequency waves
132 allows for multiple cycles of shearing in near-surface sedimentary layers, which makes them more
133 sensitive to the effects of a higher material damping ratio. Thus, we hypothesize that stronger

134 ground shaking inducing larger deformations in sedimentary deposits will affect estimates of the
135 high-frequency spectral decay parameter κ .

136

137 Figure 1 serves as an example to illustrate this hypothesis. The acceleration FAS and empirical
138 transfer function (ETF) corresponding to a pair of ground motions recorded at depth and at the
139 ground surface at FKSH14 are shown. FKSH14 is one of the 20 KiK-net station analyzed in this
140 work (see Database description). One of the ground motion pairs has a low ground shaking
141 intensity (with a surface PGA of less than 0.1 m/s^2 for both horizontal components), while the
142 other one has a higher ground shaking intensity (with a surface PGA of 0.71 m/s^2 and 0.25 m/s^2
143 for the H_1 and H_2 components, respectively). Values of κ_r per record are also estimated for both
144 pairs. To minimize the bias from path effects and isolate local site effects on κ_r values per record,
145 the selected weak and strong ground motions in Figure 1 correspond to events with similar focal
146 depths, azimuths and epicentral distances (i.e., the focal depths, azimuths and epicentral distances
147 are 5 km, 319° , and 15.04 km, respectively for the low-intensity event, and 5 km, 323° , and 15.79
148 km, respectively for the high-intensity event). The moment magnitudes of the events associated
149 with the weak and strong ground motions in Figure 1 are 4.0 and 5.1, respectively. To reduce the
150 variability associated with the calculation of κ_r from the acceleration spectrum method, we use the
151 same frequency window for all ground motions (Edwards et al. 2011).

152

153 Larger values of κ_r at the ground surface (e.g., 0.071 s for the horizontal component H_1) are
154 obtained for the high-intensity ground motion compared to the corresponding κ_r values for the
155 low-intensity ground motion (e.g., 0.057 s for H_1). Likewise, the high-intensity ground motion
156 results in a larger κ_{TF} (Drouet et al. 2010), which is computed on the decaying portion of the

157 empirical transfer function at high frequencies and is equivalent to $\Delta\kappa$ (i.e., $\kappa_{r_sur} - \kappa_{r_bor}$) estimates
158 when using the same frequency band for calculation purposes. The value of κ_{TF} corresponding to
159 the high-intensity and low-intensity motions are 0.026 s and 0.019 s, respectively. Meanwhile, the
160 ETF corresponding to the high-intensity event shows lower amplifications at higher frequencies
161 (e.g., amplification ranges from 1 to 2 between 15 and 30 Hz for H_1 approximately) than its
162 counterpart for the low-intensity ground motion (i.e., amplification ranges from 2 to 3 between 15
163 and 30 Hz for H_1 approximately), which reflects the stronger influence of increased material
164 damping ratio on high frequencies.

165
166 Values of κ_r and $\Delta\kappa$ (or κ_{TF} in Figure 1) are identical for the high- and low-intensity records for
167 the H_2 component. It must be noted that the difference in PGA between the low- and high- intensity
168 surface ground motions is less significant for the H_2 components than that for the H_1 components,
169 while the PGAs at depth are very similar (i.e., the borehole PGAs for the high-intensity motion are
170 0.06 and 0.07 m/s^2 for H_1 and H_2 components, respectively; the borehole PGAs for the low-
171 intensity motion are 0.01 m/s^2 for both H_1 and H_2 components). This may explain why $\Delta\kappa$ values
172 are the same for the low and high intensity motions in the H_2 component direction. Additionally,
173 this observation also hints that the near-surface attenuation and site effects may be affected by the
174 ground motion directionality as shown by Ji et al. (2020). Finally, soil nonlinearity is commonly
175 observed at shallower layers (Régnier et al. 2013), and their effects on κ_r at borehole are smaller
176 than at surface. Thus, the κ_r at borehole should be less affected by soil nonlinearity because it
177 nonlinear soil behavior is less likely to be triggered at depth (e.g., the borehole V_s at FKSH14 is
178 1210 m/s with borehole sensor depth of 147 m). In Figure 1, values of κ_r at borehole are similar
179 for the high- and low-intensity motions. The changes in κ_r shown in Figure 1 exemplify the need

180 to further investigate the influence of the onset of soil nonlinearity on κ_r as well as the potential
181 implications on κ_0 values at a specific site.

182

183

Database description

184 In this study, we use ground motions from the KiK-net database, which provides high quality
185 strong ground motions recorded at more than 600 stations installed uniformly across Japan. Each
186 station possesses a pair of sensors, one at the surface and another one at depth that is typically
187 between 100 to 200 m deep. The sampling frequency of the observed acceleration series is either
188 100 or 200 Hz. The P- and S-wave velocity profiles are measured by downhole PS logging and
189 available at the KiK-net website (see Resources and Data). The earthquake information, including
190 the seismic moment magnitude M_w , focal depth and epicenter location are provided by the
191 broadband seismography network (F-net) catalog. The dataset used in this work was processed by
192 Bahrampouri et al. (2020) with an automated protocol, which corrects the baseline and removes
193 the background noise with a bandpass/high-pass acausal filter. The low-frequency filter corner
194 frequency is determined based on a required value of the (signal+noise)/noise equal to 3,
195 corresponding to a signal-to-noise ratio (SNR) of 2.0 (Boore and Bommer, 2005); this frequency
196 ranges between predetermined minimum and maximum values of 0.05 and 0.5 Hz, respectively.
197 The minimum high-frequency filter corner frequency is determined for a SNR of 1.0 (Douglas and
198 Boore 2011). The minimum bandpass width is 60% of the range defined from zero to the Nyquist
199 frequency. Further detailed descriptions on the ground motion processing are available in
200 Bahrampouri et al. (2020).

201

202 We use surface and borehole records (only horizontal components) in this paper. The criteria
203 applied to select records are as follows: (1) epicentral distance is less than 150 km, (2) the SNR
204 ratio is larger than 3.0 at each frequency from 1.0 to 30 Hz, (3) focal depth is less than 35 km (Ji
205 et al. 2020), and (4) the seismic wave path does not cross the Japanese volcanic belt (Nakano et al.
206 2015). Thus, twenty stations with more than five nonlinear ground motions (the definition of
207 nonlinear ground motions is described next in the Methods section) are used in this work (with 18
208 stations having more than 10 nonlinear records and 8 stations having more than 15 nonlinear
209 records, see Table 1). Table 1 provides local soil conditions and the number of ground motions at
210 each selected site. The locations of selected stations are shown in Figure 2a, while the magnitude
211 and distance distribution of selected ground motions is provided in Figure 2b.

212

213

Method

214 Identification of the onset of nonlinearity

215 Identifying ground motions that trigger soil nonlinear behavior is key to evaluate κ estimates
216 beyond the linear-elastic regime empirically. The shear-strain index ($I_\gamma = PGV/V_s$), which is a
217 proxy for the in-situ deformation, and PGA, which describes the peak amplitude of the ground
218 motion, are commonly used to differentiate linear from nonlinear ground motions (e.g., Xu et al.
219 2019, Cabas et al. 2017, Wang et al. 2019). Moreover, the correlation between PGA and I_γ has
220 been shown to be an effective proxy to capture in-situ stress-strain relationships. This correlation
221 has been characterized via the classic hyperbolic model, which fits empirical observations
222 (Chandra et al. 2014, 2015, Guéguen et al. 2018). However, there is lack of consensus regarding
223 the sufficiency and efficiency of existing proxies associated with the onset of nonlinear behavior.
224 For example, Xu et al. (2019) assumed that records with I_γ less than 0.01% are linear ground

225 motions, while Cabas et al. (2017) adopted 0.1%. On the other hand, Ktenidou et al. (2013) chose
226 a PGA of 0.1 m/s² as the threshold for linear ground motions. Régnier et al. (2013) conducted a
227 statistical analysis on the KiK-net dataset to understand nonlinear site response at their stations.
228 They characterized linear soil behavior as that associated with motions with a PGA at depth less
229 than 0.1 m/s².

230

231 We develop appropriate criteria to identify nonlinear ground motions based on PGA and I_γ . In this
232 paper, the shear-strain index at the surface ($I_{\gamma,0}$) is defined as follows:

$$233 \quad I_{\gamma,0} = PGV_{rotD50} / V_{s,0} \quad (1)$$

234 where $V_{s,0}$ is the shear-wave velocity at the ground surface, and PGV_{rotD50} is the median PGV for
235 all rotated surface ground motions following the approach of Boore (2010). The use of PGV_{rotD50}
236 rather than the PGV from recorded ground motion horizontal components can minimize
237 directionality effects. The use of $I_{\gamma,0}$ (Equation 1) as an indicator of the onset of soil nonlinearity
238 has some limitations though. First, the selection of single V_s and PGV values to capture the depth-
239 dependent deformation in the profile may underestimate or overestimate the level of nonlinearity
240 experienced across the whole column. Thus, $I_{\gamma,0}$, as defined in Equation 1, serves simply as a proxy
241 for a representative shear strain to take place in the sedimentary deposit of interest. Notably, there
242 is no consensus regarding the most appropriate choice of V_s for I_γ estimates. For example, V_{s30} is
243 a commonly used site proxy and hence often selected to compute I_γ (e.g., Kim et al. 2016, Guéguen
244 et al. 2018). The equivalent V_s measured between two successive sensors with seismic
245 interferometry by deconvolution has also been adopted for I_γ estimates (e.g., Chandra et al. 2015,
246 2016, Wang et al. 2019). Second, values of $I_{\gamma,0}$ are not directly comparable to shear strains
247 measured in the laboratory, not only because I_γ is a proxy and not a measured value, but also

248 because the dissipation of seismic energy as captured in the laboratory may not fully represent the
249 attenuation mechanisms taking place in the field (Cabas et al. 2017).

250

251 By applying the classic hyperbolic model to describe the correlation between PGA_{rotD50} and $I_{\gamma,0}$,
252 we find that there is no unique threshold to identify nonlinear ground motions across all study sites.
253 Figure 3 provides examples of theoretical hyperbolic fitting curves at four stations (i.e., IBRH16,
254 IBRH17, IBRH20 and IWTH21) with varying V_{s30} values (from 244 to 626 m/s) to demonstrate
255 the limitations associated with using a single parameter to identify nonlinear ground motions at
256 multiple sites. Scatter points represent PGA_{rotD50} and $I_{\gamma,0}$ pairs from recorded ground motions at
257 the sites of interest, while the lines correspond to the fitting curves from the hyperbolic model. It
258 is observed that the same deformation at various sites would be triggered by different levels of
259 ground shaking (e.g., $I_{\gamma,0}$ of 0.05% will be caused by a PGA_{rotD50} around 1 m/s² at a NEHRP D site,
260 such as IBRH20 with V_{s30} of 244 m/s, and by a PGA_{rotD50} of 2 m/s² at a NEHRP C site, such as
261 IBRH16 with V_{s30} of 626 m/s). On the other hand, if nonlinearity is assumed to be triggered when
262 the PGA_{rotD50} is larger than a predetermined threshold, different levels of $I_{\gamma,0}$ will be associated
263 with the onset of soil nonlinearity. Hence, in this work, we propose a hybrid method (further
264 described in the next section) based on both, the intensity of the excitation and in-situ deformation
265 to classify ground motions.

266

267 **Linear, transitional, and nonlinear datasets**

268 Surface and borehole ground motions are considered separately herein. Régnier et al. (2013)
269 considered that records with borehole PGA less than 0.1 m/s² cannot trigger nonlinear site response
270 at the ground surface. They also observed that soil nonlinearity is mainly triggered at superficial

271 layers. In this study, there are only 3% records with borehole PGA larger than 0.1 m/s^2 . Therefore
272 the borehole records are assumed to remain in the linear-elastic regime (i.e., they do not trigger
273 nonlinear behavior at depth). Surface records are separated into three sub-datasets, namely linear,
274 transitional (i.e., soil's behavior is between the linear-elastic and nonlinear regimes), and nonlinear
275 ground motions. First, we define a threshold based on $I_{\gamma,0}$ to differentiate linear from transitional
276 records, which is hereafter referred to as $I_{\gamma,0,l}$. Likewise, a transitional threshold, $I_{\gamma,0,t}$, is defined to
277 separate transitional and nonlinear ground motions. The linear $I_{\gamma,0,l}$ threshold is defined as the onset
278 of soil nonlinearity by visual inspections of the corresponding $\text{PGA}_{\text{rotD50}}$ versus $I_{\gamma,0}$ curve, and
279 corresponds to the point where $\text{PGA}_{\text{rotD50}}$ values begin to increase at a higher rate with increasing
280 $I_{\gamma,0}$. The transitional $I_{\gamma,0,t}$ threshold captures when the soil nonlinearity becomes more apparent,
281 which corresponds to the point where the second change in slope of the $\text{PGA}_{\text{rotD50}}$ versus $I_{\gamma,0}$ curve
282 takes place. Figure 4 provides an example of the selection of $I_{\gamma,0,t}$ and $I_{\gamma,0,l}$ at station MYGH10. The
283 threshold separating the linear and transitional ground motions is $I_{\gamma,0,l} = 0.001\%$, while the
284 threshold separating transitional and nonlinear ground motions is $I_{\gamma,0,t} = 0.007\%$.

285
286 A maximum $\text{PGA}_{\text{rotD50}}$ of 0.25 m/s^2 , which is the value adopted by Régnier et al. (2016) to define
287 low-amplitude motions, is chosen as an additional constraint to avoid linear ground motions being
288 erroneously included into the nonlinear dataset. Thus, linear, transitional, and nonlinear datasets
289 are defined as follows:

- 290 ○ Linear ground motions: records with $I_{\gamma,0}$ less than the $I_{\gamma,0,l}$ threshold.
- 291 ○ Nonlinear ground motions: records with (a) $I_{\gamma,0}$ larger than the $I_{\gamma,0,t}$ threshold and (b)
292 $\text{PGA}_{\text{rotD50}}$ larger than 0.25 m/s^2 .
- 293 ○ Transitional ground motions: records not classified as either linear or nonlinear.

294

295 The validity of the proposed linear, transitional, and nonlinear datasets is tested by examining the
296 behavior of the shear modulus, G against $I_{\gamma,0}$ at the study sites. The reduction of G for empirical
297 ground motions is estimated as follows (after Guéguen et al., 2019):

$$298 \quad \frac{G}{G_{\max}} = \frac{PGA_{rotD50}}{PGV_{rotD50}/V_{s,0}} \left/ \left(\frac{PGA_{rotD50}}{PGV_{rotD50}/V_{s,0}} \right)_{\max} \right. \quad (2)$$

299 The term $\left(\frac{PGA_{rotD50}}{PGV_{rotD50}/V_{s,0}} \right)_{\max}$ is computed from the corresponding average ratio of records with

300 $I_{\gamma,0}$ less than 0.001%, which is the predetermined threshold of $I_{\gamma,0,1}$ for the linear-elastic deformation
301 limit in this work. Figure 5 shows the G/G_{\max} versus $I_{\gamma,0}$ curves at all study sites. Even though
302 Figure 5 does not directly correspond to laboratory-based curves, it serves as a first order
303 verification of the distinct behavior corresponding to the linear, transitional, and nonlinear ground
304 motions identified with the categorization scheme proposed herein. One challenge when
305 interpreting these data is the characterization of the soil volume being sampled when using I_{γ} ,
306 which is related to the frequency band that PGV is acting on. Identified linear ground motions
307 mainly have G/G_{\max} values around 1 (G/G_{\max} values higher than 1 result from using mean G values
308 as a proxy for G_{\max}), while the ratios corresponding to the nonlinear dataset are generally less than
309 1 due to the onset of soil nonlinearity. Notably, values of G/G_{\max} associated with the transitional
310 dataset are between the linear and nonlinear datasets. It is not clear whether the site response
311 associated with records identified as transitional could be equivalent to a linear-elastic or a
312 nonlinear response, because associated G/G_{\max} values vary within a single station and across
313 stations. Hence, the characterization as transitional is deemed appropriate herein.

314

315 **κ_{r_AS} estimates**

316 We use the acceleration spectrum approach (Anderson and Hough, 1984) to estimate κ_{r_AS} . To
317 minimize the variability introduced by the selection of the S-wave window, the entire time series
318 is used. Because compatibility with engineering analysis such as geotechnical site response
319 analysis and ground motion models is also desired, and such applications use complete time series,
320 calculating κ_{r_AS} using the entire time series is convenient. Moreover, the differences of κ_{r_AS}
321 values measured from S-wave window and the entire time series are not significant in most cases
322 (Ji et al. 2020). In this study, a subset of randomly selected ground motions is used to further assess
323 potential discrepancies between using the S-wave window relative to the entire time series, and no
324 significant differences are observed. Hence, the entire time histories are used.

325

326 The variability in estimates of κ_{r_AS} is a function of the selection of the frequency band (Edwards
327 et al. 2015, Perron et al. 2017) among other factors (Ji et al. 2020, Ktenidou et al. 2013). Moreover,
328 soil nonlinearity affects low and high frequencies differently. The onset of nonlinear soil behavior
329 can influence high frequencies first (Bonilla et al. 2011, Bonilla and Ben-Zion, 2020) because
330 larger shear strains are induced in softer, thinner layers located at shallower depths (i.e., in a profile
331 with increasing stiffness with depth). Hence, we compute κ_{r_AS} based on a pre-determined fixed
332 frequency window ($[f_1, f_2]$). The pre-determined f_1 corresponds to the maximum value between 1.4
333 f_c (where f_c is the earthquake source corner frequency of each record) and $1.4 f_0$ (where f_0 is the
334 site's predominant frequency). Indeed, if f_1 is lower than f_0 , the value of κ_{r_AS} will be biased by the
335 site amplification in the high-frequency range (Parolai and Bindi 2004). On the other hand, the f_c
336 requirement is added to reduce the effects of the earthquake source. The value of f_2 is set to be 25

337 Hz due to consideration of KiK-net instrument's response (Aoi et al., 2004, Fujiwara et al. 2004,
338 Oth et al. 2011, Laurendeau et al. 2013). These limits ensure a broad frequency bandwidth for κ
339 calculations of at least 10 Hz per record, which reduces potential bias from local amplification
340 effects (Parolai and Bindi, 2004; Ktenidou et al. 2016). The arithmetic average of the resulting
341 κ_{r_AS} estimated from two orthogonal horizontal components is used because it is not affected by
342 the record azimuth and its implementation can reduce the variability in κ_{r_AS} caused by ground
343 motion directionality (Ji et al. 2020). However, a prescribed, fixed-frequency band does not
344 guarantee the most appropriate linear regression for the high-frequency spectral decay in all cases.
345 Thus, we further investigate the performance of the fixed-frequency window and the effects of the
346 frequency band selection for weak and strong ground motion records by comparing individual
347 κ_{r_AS} values estimated from a pre-determined fixed frequency window with their counterparts,
348 κ_{r_auto} , resulting from an automated algorithm which does account for the most appropriate linear
349 regression.

350

351 The automated procedure used in this paper follows a similar protocol as those presented in
352 Sonnemann, and Halldorsson (2017) and Pilz et al. (2019), which focus on finding an appropriate
353 frequency band ($[f_l, f_2]$) to describe the linear decay in the high frequency range over a relatively
354 broad frequency window. As part of the automated protocol, the minimum f_l is selected as the
355 maximum value between $1.4f_0$, and $1.4f_c$. To ensure a minimum frequency bandwidth of 10 Hz,
356 the maximum value of f_l is 15 Hz and the minimum f_2 corresponds to $(f_l + 10)$ Hz. With 0.5 Hz
357 increments in f_l and f_2 , f_l is varied from the maximum value between $1.4f_0$ and $1.4f_c$ to 15 Hz, while
358 f_2 changes from $(f_l + 10)$ Hz to 25 Hz. Going through all the possible combinations of f_l and f_2 , the
359 frequency range with the minimum root mean square error over the frequency bandwidth is set as

360 the optimal frequency band. The errors are computed with the following equation after Pilz et al.
361 (2019):

$$362 \quad P = \frac{RMS}{\sqrt{\Delta f}} \quad (3)$$

363 Where Δf is the frequency bandwidth, and RMS is the root mean square error between the fitting
364 line and smoothed FAS. The FAS is smoothed with the Konno-Ohmachi filter with a coefficient
365 of 40 (Konno and Ohmachi, 1998). It should be noted that this automated procedure returns $\kappa_{r,auto}$
366 values associated with an appropriate regression for a broad frequency band. However, the changes
367 of the FAS shape in high frequencies caused by site effects or soil nonlinearity (e.g., bumps or
368 multiple linear trends) may not be properly accounted for by the automated algorithm.

369
370 Figure 6 compares κ_{r_AS} and $\kappa_{r, auto}$ for all selected ground motions at FKSH14, where overall
371 similar κ_{r_AS} and $\kappa_{r, auto}$ estimates are observed and discrepancies are more significant at the ground
372 surface than at depth. The remaining stations also show an acceptable agreement between the two
373 methods at the surface and at depth. However, there are few records that show significant
374 differences between κ_{r_AS} and $\kappa_{r, auto}$ (e.g., one surface record has κ_{r_AS} of about 0.025 s, while $\kappa_{r, auto}$
375 $\kappa_{r, auto}$ is almost 0.05 s). This reflects the variability of κ_r as a function of the frequency band
376 selected. For example, there are 13% of records at AICH17 showing clear variations with the
377 frequency band selection based on the coefficient of variation (COV) obtained for all frequency
378 windows evaluated by the automatic procedure (i.e., COV larger than 0.15). Differences between
379 κ_{r_AS} and $\kappa_{r, auto}$ are mainly caused by some empirical FAS shapes, for example, when multiple
380 linear decaying trends are present in the high-frequency range. The latter cannot be captured by
381 the single linear decay assumption within the κ -model. More research is needed to study how more

382 complex wave propagation effects in the high frequency range can be captured. Testing the
383 appropriateness of the κ -model as introduced by Anderson and Hough (1984) for these cases is
384 outside the scope of this study, but it constitutes an area of relevant future research.

385

386 Negative κ_{r_AS} values are excluded from this work. Overall, there are 14 of 20 stations with less
387 than 20% negative κ_{r_AS} values (i.e., arithmetic means of two horizontal components), and 1 of 20
388 stations with more than 50% negative κ_{r_AS} estimates. To further understand the observed negative
389 κ_{r_AS} values, the corresponding FAS are reinspected. These FAS generally show bumps or multiple
390 linear decays in high frequencies. The pre-determined frequency band used in this study is then
391 not able to capture the linear decay appropriately and leads to negative κ_{r_AS} values. Moreover,
392 amplifications at higher frequencies may also affect estimates of κ_{r_AS} and lead to negative values.
393 The original κ_r model (i.e., the linear decay of FAS in log-linear scale for high frequency ranges
394 per record proposed by Anderson and Hough, 1984), requires the site response at the site of interest
395 to be almost flat in the high frequency range. Complex in-situ site conditions that lead to high
396 frequency amplification (e.g., heterogeneities in the near-surface, and/or shallow impedance
397 contrasts) can challenge this simple linear decay model. Our observations suggest the need to
398 further explore the limitations and simplified assumptions of the Anderson and Hough (1984) κ_r
399 model for it to be applicable or extended to more complex environments and stronger ground
400 shaking. Accounting for the discrepancies between actual field conditions and assumptions
401 suggested by the Anderson and Hough (1984) κ model could reduce the large variability observed
402 in κ_{r_AS} estimates. Further research should focus on evaluating the limitations of the Anderson and
403 Hough (1984) model and its potential modification to capture more complex wave propagation
404 patterns in heterogeneous media, especially near the surface. In this work, such an investigation is

405 not included, but future efforts of the authors envision the assessment of negative κ_{r_AS} estimates
406 observed in this work as a first step to improve the existing κ model.

407

408 **κ_0 -model**

409 κ_{r_AS} is generally modeled with contributions from a site-specific component (κ_0), a path
410 component (κ_R), and a source component (κ_s). The source component, κ_s , is often assumed to be
411 negligible and its contribution is reduced by using a dataset with sufficient records (Van Houtte et
412 al. 2011, Ktenidou et al. 2014). A linear distance-dependency model is commonly applied to
413 capture the path component κ_R , which represents source-to-site effects or regional attenuation
414 (Hough et al. 1988, Anderson, 1991, Ktenidou et al. 2013, Boore and Campbell, 2017). Thus, the
415 most commonly accepted model is described below:

$$416 \quad \kappa_{r_AS} = \kappa_0 + \kappa_R \times R_e \quad (4)$$

417 where κ_0 and κ_{r_AS} are in units of time (s), κ_R is in units of second per kilometer (s/km) and R_e
418 refers to epicentral distance in km. This model is valid when a unique source-to-site path is
419 assumed for each record along with a homogeneous and frequency-independent seismic quality
420 factor Q (Knopoff 1964). In this paper, the assumption of a unique source-to-site path is supported
421 by using ground motions with R_e less than 150 km (e.g., Palmer and Atkinson 2020, Cabas et al.
422 2017, Ktenidou et al. 2013). In addition, ground motions whose travel path crosses Japan's
423 volcanic belt are not included in our database to minimize the likelihood of seismic waves
424 propagating through regions with varying Q values (Pei et al. 2009, Nakano et al. 2015).

425

426 The model described by Equation (4) is straightforward to apply when only surface linear ground
427 motion datasets are used. However, the incorporation of nonlinear and borehole ground motions

428 adds complexity to the estimation of regional attenuation as captured by κ_R . In this paper, we
 429 assume that soil nonlinearity is triggered near the surface, which is consistent with previous studies
 430 showing that nonlinear behavior occurs mostly in the superficial soil layers (i.e., Régnier et al.,
 431 2013; Bonilla et al., 2019; Qin et al., 2020). Thus, nonlinear soil behavior is treated as a site
 432 contribution rather than a path contribution, and we assume the regional attenuation to be identical
 433 for linear and nonlinear ground motions recorded at the ground surface and at depth.

434
 435 Analogous to the formulation suggested by Douglas et al. (2010) for soil and rock sites, we propose
 436 a model based on Equation (4), which includes linear and nonlinear surface and borehole records:

$$437 \quad \kappa_{r_AS} = N_1 \kappa_{0_depth} + N_2 \kappa_{0_lin_sur} + N_3 \kappa_{0_nl_sur} + N_4 \kappa_{0_tran_sur} + \kappa_R \times R_e \quad (5)$$

438 where κ_{0_depth} is the site-specific κ_0 at depth (i.e., depth of borehole sensor), and $\kappa_{0_lin_sur}$, $\kappa_{0_tran_sur}$
 439 and $\kappa_{0_nl_sur}$ are the site-specific linear, transitional and nonlinear κ_0 at the surface. The coefficients
 440 N_1 , N_2 , N_3 , and N_4 are defined as follows:

$$441 \quad N_1 = \begin{cases} 1 & \text{for dataset at depth} \\ 0 & \text{otherwise} \end{cases}$$

$$442 \quad N_2 = \begin{cases} 1 & \text{for linear dataset at surface} \\ 0 & \text{otherwise} \end{cases}$$

$$443 \quad N_3 = \begin{cases} 1 & \text{for nonlinear dataset at surface} \\ 0 & \text{otherwise} \end{cases}$$

$$444 \quad N_4 = \begin{cases} 1 & \text{for transitional dataset at surface} \\ 0 & \text{otherwise} \end{cases}$$

445 The parameter N_4 in Equation (5) only takes a value of 1 when ground motions classified as
 446 transitional are considered independently and not as part of the linear or nonlinear datasets. That
 447 means that when transitional ground motions are excluded from analysis (i.e., AP1, see Table 2)
 448 or included into either the linear (i.e., AP2, see Table 2) or nonlinear (i.e., AP3, see Table 2)
 449 datasets, the coefficient N_4 will be equal to zero.

450

451

Results and Discussion

452 Effects of soil nonlinearity on empirical κ_{r_AS}

453 First, we study the influence of soil nonlinearity on κ_{r_AS} estimates per record at each site. It should

454 be noted that soil nonlinearity is commonly triggered at shallower soil layers (Régner et al. 2013),

455 so we only focus on surface records in this section. Figure 7 depicts calculated κ_{r_AS} values at the

456 surface against the corresponding $I_{\gamma,0}$ values at FKSH14 ($V_{s30} = 237$ m/s) and MYGH10 ($V_{s30} =$

457 348 m/s). As described in Equation (4), κ_{r_AS} is affected by both local site conditions and path

458 effects in the context of a linear-elastic deformation analysis. Hence, the colorbar in Figure 7

459 represents varying epicentral distances, and the sizes of markers represent the corresponding

460 PGA_{rotD50} . An overall increasing trend of κ_{r_AS} with increasing intensity of ground shaking (either

461 evidenced by increased PGA or $I_{\gamma,0}$ values) is observed at FKSH14 for events that share similar

462 epicentral distances. A slightly decreasing trend is observed for short-distance records with R_e less

463 than about 50 km and high PGA_{rotD50} . However, κ_{r_AS} values corresponding to those shorter

464 distance and higher PGA_{rotD50} events (i.e., the largest circles in Figure 7a) are larger than their

465 counterparts for low-intensity motions (i.e., the smallest circles in Figure 7a) regardless of the R_e .

466 To the best knowledge of the authors, this observation has not been reported before and could be

467 associated with the depth of influence of κ_r . Variations in κ due to strong nonlinear effects may be

468 a function of a more significant contribution of the site to the overall attenuation, which may not

469 be necessarily the case for smaller amplitude events. There might be several mechanisms of

470 attenuation combined, and their contributions as captured by κ need to be further investigated.

471 Similar trends are observed at other seven sites with V_{s30} less than 400 m/s, which include AICH17,

472 CHBH13, FKSH11, IBRH20, IWTH26, MYGH07, and TCGH16, and at KMMH12 with V_{s30}
473 greater than 400 m/s.

474

475 The increasing κ_{r_AS} trend with increasing PGA_{rotD50} and $I_{\gamma,0}$ is not as significant at MYGH10
476 (Figure 7b), which has relatively stiffer site conditions than FKSH14. Either no correlation or a
477 slightly decreasing trend is found at other stiff sites with V_{s30} greater than 400 m/s (i.e., FKSH21,
478 NIGH12, NGNH29, NIGH07, KMMH01, and IBRH16), and at four softer sites with V_{s30} between
479 300 and 400 m/s (i.e., IWTH21, FKSH18, FKSH19, and IBRH17). We note that the number of
480 available nonlinear records for the R_e ranges at the sites where the decreasing trend is observed is
481 rather limited. Additional nonlinear records at those sites are necessary (i.e., stronger intensity
482 ground motions) to further evaluate the contributing factors to a potential decreasing trend in κ_{r_AS}
483 values. However, in general, we observe that positive correlations between κ_{r_AS} and the intensity
484 of ground shaking are more significant at softer sites (e.g., TCGH16 with V_{s30} of 213 m/s) than at
485 stiffer sites (e.g., KMMH12 with V_{s30} of 408 m/s). These data support that the onset of soil
486 nonlinearity can affect κ_{r_AS} estimates, but such influence is station-dependent. The level of soil
487 nonlinearity can be unique at each site (for a similar intensity of ground shaking) because of the
488 characteristics of shallow geologic structures (e.g., differences in velocity gradients and seismic
489 impedance contrasts) and the location of low shear-wave velocity layers. Thus, subsurface
490 conditions can play a key role on the effects of nonlinearity on κ_{r_AS} . We observe the same patterns
491 shown in Figure 7 when using our results from the automated procedure to compute κ_{r_auto} .

492

493 **Effects of soil nonlinearity on the empirical κ_0 -model**

494 Linear, transitional, and nonlinear ground motion datasets are used in this section to evaluate the
495 κ_0 -model beyond the linear-elastic regime. We explore four approaches (i.e., AP1 to AP4) to
496 incorporate records within the transitional dataset into the κ_0 -model presented in Equation (5).
497 Table 2 summarizes how the identified linear, transitional, and nonlinear datasets are used to
498 estimate $\kappa_{0_lin_sur}$ and $\kappa_{0_nl_sur}$.

499
500 Figures 8 and 9 present the resulting κ_0 models from each approach at FKSH14 and MYGH10,
501 respectively. Figure 8 shows that κ_{r_AS} and κ_0 values corresponding to the nonlinear ground
502 motions (regardless of the selected approach to construct the nonlinear dataset) are larger than their
503 linear counterparts at FKSH14. However, results at the stiffer station presented in Figure 9 show
504 little disagreement between κ_{r_AS} and κ_0 values corresponding to the linear and nonlinear motions
505 (regardless of the approach to construct each dataset).

506
507 Variations in $\kappa_{0_lin_sur}$ estimates are observed as a function of the approach considered to construct
508 the linear datasets (see specific values in Table 3 for FKSH14 and MYGH10, the results for other
509 selected stations are available in the electronic supplement). Similarly, variations in $\kappa_{0_nl_sur}$ values
510 are also found across the different approaches to define the nonlinear datasets. At FKSH14 (Figure
511 8), $\kappa_{0_lin_sur}$ estimates are more variable as a function of the dataset definitions (with a maximum
512 difference of 15.83% across approaches AP1 to AP4), compared to $\kappa_{0_nl_sur}$ values (with a
513 maximum difference of 5.10%). In addition, Figure 8 (d) shows that data points corresponding to
514 the transitional dataset are more compatible with their counterparts within the nonlinear dataset,
515 which may indicate that at FKSH14, the level of nonlinearity induced by the transitional dataset is

516 closer to that induced by the ground motions in the nonlinear dataset. Other study sites such as
517 AICH17 ($V_{s30} = 314$ m/s) and IWTH21 ($V_{s30} = 521$ m/s) also show that $\kappa_{0_lin_sur}$ estimates are more
518 sensitive to dataset selections. In contrast, variations of $\kappa_{0_lin_sur}$ and $\kappa_{0_nl_sur}$ across datasets at
519 MYGH10 (Figure 9) are small, with maximum differences of only 1.56% and 2.53%, respectively.
520 Large differences in $\kappa_{0_nl_sur}$ estimates across datasets are observed at eight sites, but the limited
521 number of nonlinear records at some of those sites may be the main contributing factor (e.g., there
522 are only six nonlinear records at IWTH21, which results in a maximum difference of 47.05% for
523 $\kappa_{0_lin_sur}$ and 11.17% for $\kappa_{0_nl_sur}$). Adding transitional records to either the linear or the nonlinear
524 dataset at such sites can bias the regression model. In general, differences in the κ_0 -model as a
525 function of the selected dataset are observed in 45% of our study sites (with differences in $\kappa_{0_lin_sur}$
526 or $\kappa_{0_nl_sur}$ values greater than 10%). This is a relevant observation because it demonstrates the
527 importance of selecting appropriate ground motions even for typical κ_0 estimations (i.e., in the
528 linear-elastic regime) at a given site.

529
530 Our findings suggest that the development of a κ_0 -model beyond the linear-elastic regime requires
531 an evaluation of the definition of what constitutes linear and nonlinear ground motion datasets.
532 The identification of transitional ground motion datasets in this study allows us to assess which
533 records provide estimates of κ_{r_AS} that are closer to either the linear or the nonlinear behavior at
534 different sites. Differences in behavior triggered by the records within the transitional database
535 may be caused by unique local site conditions (i.e., the level of soil nonlinearity developed at each
536 site) or by limitations of the simplified definition used herein to classify transitional records (i.e.,
537 as a function of PGA and I_T). Identifying appropriate linear and nonlinear datasets for κ_{r_AS}
538 estimations requires further research to provide consistent models of near-surface attenuation that

539 can more effectively be implemented from small to large shear strains. However, the site-specific
540 response at a site of interest may impose challenges in determining appropriate dataset
541 classifications based on a simple, generalized criterion.

542

543 Figure 10 provides ratios of $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ at the 20 study sites against the corresponding time-
544 average V_s value in the top 5 m (V_{s5}), 10 m (V_{s10}), and 30 m (V_{s30}). The ratios are computed based
545 on the AP3 and AP4 approaches to construct linear and nonlinear datasets (Table 2). Larger ratios
546 are observed at softer sites regardless of the dataset chosen (i.e., AP3 and AP4) for the κ_0 -model.
547 Differences between κ_0 values in the linear and nonlinear regimes seem to be reconciled at sites
548 with higher V_{s5} (> 300 m/s), V_{s10} (> 300 m/s) and V_{s30} (> 400 m/s) values, where the ratios fluctuate
549 more closely around unity particularly when using AP3. The trend of increasing ratios of
550 $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ with softer site conditions is better captured by V_{s5} and V_{s10} than by V_{s30} , because
551 soil nonlinearity is more likely triggered at shallower and softer layers (Régner et al. 2013). Hence,
552 lower V_s layers may dominate soil nonlinearity effects on κ_0 . Thus, site proxies that can
553 characterize such near-surface layers may be more informative when evaluating nonlinear soil
554 effects on κ_0 .

555

556 When grouping transitional and nonlinear ground motions (i.e., AP3), most stations result in ratios
557 of $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ larger than one, which can be interpreted as the signature of soil nonlinearity
558 on the near-surface attenuation estimates (i.e., near-surface attenuation increases with increasing
559 deformations as soil nonlinearity is triggered). These findings are consistent with the behavior of
560 material damping ratio observed in dynamic laboratory testing of soils (i.e., increased damping
561 ratio with increasing shear strain; Darendeli 2001; Menq 2003; Ishibashi and Zhang 1993). When

562 treating linear, transitional, and nonlinear datasets independently (i.e., AP4), there are 12 sites with
563 ratios larger than one. The instances where ratios of $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ are lower than one may result
564 from the limited nonlinear records available at those sites coupled with the uncertainties associated
565 with $\kappa_{0_lin_sur}$ (e.g., Ji et al., 2020).

566

567 Overall, the variations observed in the $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ ratio support our hypothesis that soil
568 nonlinearity plays a role on the estimates of near-surface attenuation from recorded ground
569 motions. This effect is station-dependent, and further research is needed to identify the most
570 appropriate parameter or vector of parameters capable of capturing the influence of nonlinear soil
571 behavior on near-surface attenuation. Moreover, the relatively weaker correlation between V_{s30}
572 and the $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ ratio evidences the challenges in connecting site conditions and soil
573 nonlinearity via a single site parameter. Multiple parameters that can describe attenuation and
574 impedance effects from the shallow and deep geologic structures should be investigated. The
575 $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ ratio corresponding to IWTH21 ($V_{s30} = 521$ m/s) is not shown in Figure 10 because
576 it is very large (i.e., approximately 1.8). This observation may result from uncertainties associated
577 with κ_{r_AS} values propagating to estimates of κ_0 when the fixed frequency band approach is applied
578 for all records without consideration of the optimal linear decay trend. In fact, the corresponding
579 $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ ratio when implementing the automated procedure is approximately 0.90 for this
580 station. Finally, Figure 10 shows less scatter in $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ ratios when using datasets defined
581 by AP3. In addition, using AP3 results in ratios either larger than one or approaching one for most
582 stations (i.e., only FKSH19 and KMMH01 results in a ratio lower than unity), which is consistent
583 with our conceptual basis for increased attenuation with the onset of nonlinear soil behavior.

584 Therefore, we adopt the AP3 approach (which includes transitional records into nonlinear dataset)
585 to evaluate predictions of near-surface attenuation in the next section of this paper.

586

587 **Effects of soil nonlinearity on predicted near-surface attenuation**

588 Site-specific $\kappa_{0_lin_sur}$ or $\kappa_{0_nl_sur}$ values from Equation (5) allow for the comparison of empirical
589 estimates of near-surface attenuation, but these two parameters represent the average attenuation
590 of all records in linear and nonlinear regimes. Thus, in this section, we introduce the predicted
591 near-surface attenuation at zero-distance (κ_{0_pred}), which is expected to capture the attenuation
592 contributed by the superficial soil layers per event by removing the path contributions from κ_{r_sur} .
593 κ_{0_pred} is modeled as:

$$594 \quad \kappa_{0_pred} = \kappa_{r_AS_sur} - R_e \bullet \kappa_R \quad (6)$$

595 where $\kappa_{r_AS_sur}$ refers to the individual κ_{r_AS} value for a surface ground motion, and the path-
596 component, κ_R , corresponds to the values derived with Equation (5) at each site of interest. We
597 assume that by removing the effect of the path-component κ_R from $\kappa_{r_AS_sur}$ values per record, the
598 remaining κ_{0_pred} becomes an approximation to the attenuation contributed by the shallower
599 sedimentary deposits per event. Thus, we can explore how the near-surface attenuation changes
600 with the various input ground motion amplitudes at the site of interest.

601

602 Figure 11 provides comparisons between κ_{0_pred} , ground shaking intensity, and deformation as
603 captured by PGA_{rotD50} and $I_{7,0}$ at FKSH14 and MYGH10. Both colors and sizes of markers
604 represent the PGA_{rotD50} values per record. The red dashed-lines result from a local regression model
605 characterizing the κ_{0_pred} versus $I_{7,0}$ data. Triangles and circles represent the linear and nonlinear

606 ground motions (identified with AP3), respectively. Values of κ_{0_pred} first increase and then
607 decrease with increasing PGA_{rotD50} and $I_{\gamma,0}$ at FKSH14. This behavior is also observed at other 7
608 sites (i.e., AICH17, CHBH13, FKSH11, IWTH26, KMMH12, MYGH07, TCGH16). Even with
609 the decreasing trend for large $I_{\gamma,0}$, κ_{0_pred} is still generally higher than its counterpart in the linear
610 regime (i.e., the means of linear and nonlinear κ_{0_pred} are 0.050 sec and 0.0605 sec, respectively).
611 Overall, κ_{0_pred} values at FKSH14 corresponding to larger deformations and higher PGA_{rotD50} are
612 larger than those corresponding to weaker ground motions. In contrast, only a weak correlation to
613 the intensity of ground shaking and deformation in situ is observed at MYGH10. These results are
614 consistent with our estimations of κ_r shown in Figure 7. Soil nonlinear behavior can influence
615 near-surface attenuation as captured by κ_r and κ_0 , and local site conditions may play a key role in
616 this process. The remaining approaches explored in this study (i.e., AP1, AP2, and AP4) provide
617 similar results as those shown in Figure 11.

618
619 Figure 12 compares the probability distribution of κ_{0_pred} values from the linear and nonlinear
620 datasets (AP3 case) at FKSH14 and MYGH10. The resulting κ_{0_pred} values are fitted with a
621 Gaussian distribution and the corresponding probability density functions (PDFs) are represented
622 by red lines. A shift to the right (i.e., toward larger κ_{0_pred} values) of the theoretical PDF is observed
623 at FKSH14 as ground motions from the linear and nonlinear datasets are considered. The mean
624 κ_{0_pred} estimates change from 0.05 s for the linear dataset to 0.0605 s for the nonlinear dataset at
625 FKSH14 (i.e., a difference of 21%). In contrast, the variation of mean κ_{0_pred} between linear and
626 nonlinear datasets at MYGH10 is 4%. Most of our study sites have either a significant increase in
627 their mean κ_{0_pred} when using the nonlinear dataset (i.e., an increase of more than 20%) or only a
628 slight increase. There are only 4 stations that show a decrease in their mean κ_{0_pred} values with

629 respect to the linear dataset when using the nonlinear one (i.e., FKSH19, KMMH01, NIGH07, and
630 NIGH12). Statistical hypothesis tests (i.e., t-test) are conducted to analyze whether there is a
631 statistically significant difference between the means of linear and nonlinear κ_{0_pred} distributions.
632 Considering a critical value of 5%, p-values at each station are shown in Table 4. Nine out of 20
633 stations display statistically significant differences between their mean κ_{0_pred} corresponding to the
634 linear and nonlinear datasets. Table 4, Figures 11 and 12 show that soil nonlinear behavior can
635 affect κ_{0_pred} at the sites selected in this study, although this influence is station-dependent. At the
636 stations that display apparent effects of nonlinearity on κ , an increasing trend in predicted near-
637 surface attenuation with increasing ground shaking intensity and/or increasing deformation is
638 observed.

639 **Conclusions**

640 In this work, we investigated the influence of soil nonlinear behavior on κ_{r_AS} values per record
641 and site-specific κ_0 estimates at 20 stations selected from the KiK-net database. To avoid potential
642 bias on our results due to the calculation process, we also examined the effects of the frequency
643 band selection on κ_{r_AS} estimates, and the differences between using the S-wave window or the
644 entire time series FAS. We compared results from a predetermined fixed-frequency window
645 approach with an automated procedure that considers multiple frequency windows. The latter is
646 capable of finding the optimal frequency band per record for all records at each site. The selection
647 of a common, fixed and broader frequency band for κ_{r_AS} estimations reduced the scatter and bias
648 in the data, while providing reasonable estimations of κ_{r_AS} . On the other hand, values of κ_{r_AS}
649 computed from the S-wave window FAS were reasonably similar to their counterparts based on
650 the entire time series FAS. Hence, the analyses presented in this paper were conducted with κ_r

651 values estimated by the fixed-frequency band approach and the FAS corresponding to the entire
652 time series.

653

654 A consistent identification of ground motions that trigger nonlinear behavior in sedimentary
655 deposits is also necessary to quantify near-surface attenuation beyond the linear-elastic regime.

656 Based on the examination of an in-situ stress-strain proxy, namely the correlation between

657 PGA_{rotD50} and $I_{\gamma,0}$, we found that the variation of shear strains with ground shaking intensity at the

658 onset of nonlinear soil behavior is site-specific. A unique threshold for a single parameter, whether

659 it is PGA_{rotD50} or $I_{\gamma,0}$, was not able to capture the onset of soil nonlinearity at our study sites in a

660 consistent manner across all sites. Hence, we proposed a hybrid method to classify linear and

661 nonlinear ground motions considering both, PGA_{rotD50} and $I_{\gamma,0}$, which resulted in linear, transitional,

662 and nonlinear datasets at each site.

663

664 Increasing κ_{r_AS} values with increasing PGA_{rotD50} or $I_{\gamma,0}$ for ground motions with similar epicentral

665 distances were observed at about half of our study sites. This trend was more consistently observed

666 at softer sites. Additionally, we found that κ_0 -models could be biased by the definition of linear

667 and nonlinear ground motion datasets. Hence, we studied the effects of ground motion

668 categorization and proposed a hybrid classification scheme for linear and nonlinear records. We

669 defined transitional ground motions as those associated with soil behavior between the linear-

670 elastic and nonlinear regimes. Even though more research is necessary to define robust

671 classification schemes for linear and nonlinear ground motions, we observed that including the

672 transitional motions into the nonlinear dataset reduced the variability associated with κ estimations

673 at our study sites.

674

675 Our results also revealed differences between $\kappa_{0_lin_sur}$ (i.e., κ_0 corresponding to the linear-elastic
676 regime) and $\kappa_{0_nl_sur}$ (i.e., κ_0 for the nonlinear regime) at most sites when implementing the κ_0 -
677 model using ground motions classified by AP3 (which includes transitional records into the
678 nonlinear dataset). Such differences were more prevalent among softer sites. Site parameters such
679 as V_{s5} , V_{s10} , and V_{s30} were used in this study to investigate the influence of soil conditions on the
680 effects of nonlinearity on κ_0 . Considering that high frequencies have short wavelengths, and that
681 nonlinear soil behavior is triggered in low velocity layers more often located at a shallow depth,
682 site proxies such as V_{s5} and V_{s10} may be more informative than V_{s30} when assessing effects of
683 nonlinearity on κ . For instance, large V_{s30} values do not imply that all near-surface layers have a
684 large V_s . The ratio of $\kappa_{0_nl_sur}$ and $\kappa_{0_lin_sur}$ decreases and approaches one for increasing V_{s5} , V_{s10} ,
685 and V_{s30} , when using the AP3 method to define nonlinear ground motion datasets.

686

687 The hypothesis posed and tested in this paper focused on the effects of ground shaking intensity
688 on induced shear strains in sedimentary deposits and associated consequences on the attenuation
689 experienced by seismic waves (particularly in the high frequency range). In general, we find that
690 soil nonlinear behavior can affect estimates of κ_{r_AS} and κ_0 , but our results show that this influence
691 is station-dependent. This is reasonable because the wave propagation of short wavelength waves
692 is highly affected by heterogeneities in the soil or rock, local geologic structures, and topography.
693 Moreover, the level of soil nonlinearity can be distinct at a given site (even when site
694 parameterizations such as V_{s30} are similar and the considered intensity of ground shaking is also
695 similar) because of the complexities of the in situ subsurface conditions (e.g., differences in
696 velocity gradients and seismic impedance contrasts). We note that 2D/3D site effects may affect

697 ground motions recorded at six of our stations. The influence of soil nonlinearity on κ values
698 computed at these stations (i.e., AICH17, CHBH13, FKSH21, IBRH20, KMMH01, KMMH12;
699 based on the classification of Thompson et al. 2012 and Pilz and Cotton 2019) may be masked by
700 the combined effects of wave scattering and topographic effects. Further research is necessary to
701 evaluate the contributions of the aforementioned mechanisms on κ estimates at stations subjected
702 to 2D/3D site effects. Likewise, future work should focus on collecting and analyzing additional
703 strong ground motion data to identify local site conditions more conducive to generate significant
704 changes in near-surface attenuation as captured by κ_0 when nonlinear soil behavior is triggered.

705
706 Complexities in the wave propagation phenomenon driven by scattering effects and amplification
707 in the high-frequency range can result in negative estimates of κ_r . In this study, we obtained
708 negative κ_{r_AS} estimates when multiple linear decaying trends, bumps, and high frequency
709 amplifications affected the corresponding FAS spectral shape. The identification of multiple linear
710 decays in the high-frequency range supports previous work on the bias in κ_{r_AS} associated with the
711 selection of the frequency band. The bumps and amplifications in the high frequencies present in
712 the FAS of some of the ground motions in our database hint that the site response may not be
713 approximately flat within the frequency range of interest for κ_{r_AS} calculation. Considering that a
714 flat site response is one of the assumptions of the Anderson and Hough (1984) κ -model, further
715 research is needed to overcome this limitation at sites where this is not the case. This work not
716 only provides evidence of the need to understand and quantify κ in both, the linear and nonlinear
717 regimes, but it also presents the limitations of the current κ -model when it comes to characterizing
718 attenuation when conditions deviate from the original assumptions embedded in the Anderson and
719 Hough (1984) model.

720

Data and Resources

721 Accelerograms and geotechnical data are downloaded from the KiK-net network at
722 <http://www.kyoshin.bosai.go.jp> (last accessed May 2020). The earthquake information is available
723 from F-net network at <http://www.fnet.bosai.go.jp/top.php> (last accessed May 2020). The
724 supplemental material to this article includes two tables and three figures. The tables provide the
725 κ_0 -model results when using the datasets defined by AP1, AP2, AP3, and AP4, and the κ_{0_pred}
726 results estimated with datasets defined by AP3. The three sets of figures presented depict the results
727 corresponding to our 20 stations as follows:

- 728 • Surface PGA_{rotD50} against $I_{\gamma,0}$ (i.e., results analogous to those presented in Figure 4 for
729 MYGH10).
- 730 • Surface κ_{r_AS} estimates and their corresponding PGA_{rotD50} , $I_{\gamma,0}$ and R_c values for selected
731 ground motions at each study site (i.e., results analogous to those presented in Figure 7 for
732 FKSH14 and MYGH10).
- 733 • Estimated surface κ_{0_pred} and their corresponding ground shaking intensity and in situ
734 deformation characterized by PGA_{rotD50} against $I_{\gamma,0}$, respectively (i.e., results analogous to
735 those presented in Figure 11 for FKSH14 and MYGH10).

736

737

Acknowledgements

738 The authors would like to appreciate that the National Research Institute for Earth and Disaster
739 Prevention (NIED) provides the ground motion and site information. The authors acknowledge Dr.
740 Adrian Rodriguez-Marek and Mahdi Bahrampouri for sharing their processed dataset. We also are
741 thankful to Associate Editor Stefano Parolai and two anonymous reviewers for the useful
742 comments. This work was sponsored by the U.S. Geological Survey under Grant Number

743 G19AP00058 and the NCSU Internationalization Seed Grant. The views and conclusions
744 contained in this document are those of the authors and should not be interpreted as representing
745 the opinions or policies of the U.S. Geological Survey.

746

747

Reference

748 Anderson, J. G., & Hough, S. E. (1984). A model for the shape of the Fourier amplitude spectrum
749 of acceleration at high frequencies. *Bull Seismol Soc Am.* **74**, no. 5, 1969-1993.

750 Anderson, J.G. (1991). A preliminary descriptive model for the distance dependence of the spectral
751 decay parameter in southern California. *Bull Seismol Soc Am.* **81**, no. 6, 2186-2193.

752 Al Atik, L., Kottke, A., Abrahamson, N. and Hollenback, J., (2014). Kappa (κ) scaling of ground-
753 motion prediction equations using an inverse random vibration theory approach. Bulletin
754 of the Seismological Society of America, 104, no. 1, 336-346.

755 Aoi, S., T. Kunugi, and H. Fujiwara (2004). Strong-motion seismograph network operated by
756 NIED: K-NET and KiK-net, J. Japan Assoc. Earthq.Eng. 4, no. 3, 65–74.

757 Bahrapouri, M., Rodriguez-Marek, A., Shahi, S., & Dawood, H. (2020). An update database for
758 ground motion parameters for KiK-net records. *Earthquake Spectra*.

759 Bonilla, L.F., Tsuda, K., Pulido, N., Régnier, J. and Laurendeau, A., 2011. Nonlinear site response
760 evidence of K-NET and KiK-net records from the 2011 off the Pacific coast of Tohoku
761 Earthquake. *Earth, planets and space*, 63(7), p.50.

762 Bonilla, L.F., Guéguen, P. and Ben-Zion, Y., 2019. Monitoring Coseismic Temporal Changes of
763 Shallow Material during Strong Ground Motion with Interferometry and
764 Autocorrelation Monitoring Coseismic Temporal Changes of Shallow Material during

765 Strong Ground Motion. *Bulletin of the Seismological Society of America*, 109(1), pp.187-
766 198.

767 Bonilla, L.F. and Ben-Zion, Y., 2020. Detailed space-time variations of the seismic response of
768 the shallow crust to small earthquakes from analysis of dense array data. *Geophysical*
769 *Journal International*.

770 Boore, D.M., 2003. Simulation of ground motion using the stochastic method. *Pure and applied*
771 *geophysics*, 160(3-4), pp.635-676.

772 Boore, D.M. and Bommer, J.J., 2005. Processing of Strong-Motion Accelerograms: Needs,
773 Options and Consequences. *Soil Dynamics and Earthquake Engineering*, 25(2), 93-115.

774 Boore, D. M. (2010). Orientation-independent, nongeometric-mean measures of seismic intensity
775 from two horizontal components of motion. *Bull Seismol Soc Am.* 100, no. 4, 1830-1835.

776 Boore, D.M. and Campbell, K.W., 2017. Adjusting central and eastern North America ground-
777 motion intensity measures between sites with different reference-rock site conditions.
778 *Bulletin of the Seismological Society of America*, 107(1), pp.132-148.

779 Campbell, K.W., 2003. Prediction of strong ground motion using the hybrid empirical method and
780 its use in the development of ground-motion (attenuation) relations in eastern North
781 America. *Bulletin of the Seismological Society of America*, 93(3), pp.1012-1033.

782 Cabas, A., Rodriguez-Marek, A., & Bonilla, L. F. (2017). Estimation of Site-Specific Kappa (κ_0)-
783 Consistent Damping Values at KiK-Net Sites to Assess the Discrepancy between
784 Laboratory-Based Damping Models and Observed Attenuation (of Seismic Waves) in the
785 Field. *Bull Seismol Soc Am.* **107**, no. 5, 2258-2271.

786 Chandra, J., Guéguen, P. and Bonilla, L.F., 2014. Application of PGV/VS proxy to assess
787 nonlinear soil response—from dynamic centrifuge testing to Japanese K-NET and KiK-net
788 data. In Second European Conference on Earthquake Engineering and Seismology,
789 Istanbul Aug (pp. 25-29).

790 Chandra, J., Guéguen, P., Steidl, J.H. and Bonilla, L.F., 2015. In situ assessment of the G– γ curve
791 for characterizing the nonlinear response of soil: Application to the Garner Valley
792 downhole array and the wildlife liquefaction array. *Bulletin of the Seismological Society*
793 *of America*, 105(2A), pp.993-1010.

794 Chandra, J., Gueguen, P. and Bonilla, L.F., 2016. PGA-PGV/Vs considered as a stress–strain
795 proxy for predicting nonlinear soil response. *Soil dynamics and earthquake engineering*, 85,
796 pp.146-160.

797 Darragh, R.B. and Shakal, A.F., 1991. The site response of two rock and soil station pairs to strong
798 and weak ground motion. *Bulletin of the Seismological Society of America*, 81(5),
799 pp.1885-1899.

800 Darendeli M. B. (2001). Development of a new family of normalized modulus reduction and
801 material damping curves, Ph.D. Thesis, University of Texas at Austin, Austin, Texas.

802 Dimitriu, P., Theodulidis, N., Hatzidimitriou, P., & Anastasiadis, A. (2001). Sediment non-
803 linearity and attenuation of seismic waves: a study of accelerograms from Lefkas, western
804 Greece. *Soil Dynam Earthquake Eng.* **21**, no.1, 63-73.

805 Durward, J.A., Boore, D.M., Joyner, W.B., Durward, J.A., Boore, D.M. and Joyner, W.B., 1996.
806 The amplitude dependence of high-frequency spectral decay: constraint on soil

807 nonlinearity. In Proc, of the International Workshop on Site Response Subjected to Strong
808 Earthquake Motions (pp. 82-103).

809 Drouet, S., F. Cotton, and P. Gueguen (2010). VS30, κ , regional attenuation and Mw from
810 accelerograms: Application to magnitude 3–5 French earthquakes, *Geophys. J. Int.* 182,
811 880–898.

812 Douglas, J., Gehl, P., Bonilla, L.F. and Gélis, C., (2010). A κ model for mainland France. *Pure*
813 *Appl Geophys.* 167, no.11, 1303-1315.

814 Douglas, J. and Boore, D.M., 2011. High-frequency filtering of strong-motion records. *Bulletin of*
815 *Earthquake Engineering* 9(2), 395-409.

816 Edwards, B., Ktenidou, O.J., Cotton, F., Abrahamson, N., Van Houtte, C. and Fäh, D., (2015).
817 Epistemic uncertainty and limitations of the $\kappa 0$ model for near-surface attenuation at hard
818 rock sites. *Geophys J Int.* **202**, no. 3, 1627-1645.

819 Fujiwara, H., S. Aoi, T. Kunugi, and S. Adachi (2004). Strong-motion observation networks of
820 NIED: K-NET and KiK-net, National Research Institute for Earth Science and Disaster
821 Prevention. Garcia, D., Wald, D. J., & Hearne, M. G. (2012). A global earthquake
822 discrimination scheme to optimize ground-motion prediction equation selection. *Bull*
823 *Seismol Soc Am.* 102, no.1, 185-203.

824 Guéguen, P., Bonilla, L.F. and Douglas, J., 2019. Comparison of Soil Nonlinearity (In Situ Stress–
825 Strain Relation and G/Gmax Reduction) Observed in Strong-Motion Databases and
826 Modeled in Ground-Motion Prediction Equations Comparison of Soil Nonlinearity
827 Observed in Strong-Motion Databases and Modeled in GMPEs. *Bulletin of the*
828 *Seismological Society of America*, 109(1), pp.178-186.

829 Hough, S.E., Anderson, J.G., Brune, J., Vernon III, F., Berger, J., Fletcher, J., Haar, L., Hanks, L.
830 and Baker, L., 1988. Attenuation near Anza, California. *Bulletin of the Seismological*
831 *Society of America*, 78(2), pp.672-691.

832 Idriss, I.M., Dobry, R.U. and Sing, R.D., 1978. Nonlinear behavior of soft clays during cyclic
833 loading. *Journal of geotechnical and geoenvironmental engineering*, 104.

834 Ishibashi, I. and Zhang, X., 1993. Unified dynamic shear moduli and damping ratios of sand and
835 clay. *Soils and foundations*, 33(1), pp.182-191.

836 Ji, C., Cabas, A., Cotton, F., Pilz, M. and Bindi, D., 2020. Within-Station Variability in Kappa:
837 Evidence of Directionality Effects. *Bull Seismol Soc Am.*, 110(3), pp.1247-1259.

838 Knopoff, L. (1964). *Q*, *Rev. Geophys.* 2, no. 4, 625–660.

839 Ktenidou, O. J., Gélis, C., & Bonilla, L. F. (2013). A study on the variability of kappa (κ) in a
840 borehole: Implications of the computation process. *Bull Seismol Soc Am.* **103**, no.2A, 1048-
841 1068.

842 Ktenidou, O. J., Cotton, F., Abrahamson, N. A., & Anderson, J. G. (2014). Taxonomy of κ : A
843 review of definitions and estimation approaches targeted to applications. *Seismol Res Lett.*
844 **85**, no.1, 135-146.

845 Ktenidou, O.J., Abrahamson, N.A., Drouet, S. and Cotton, F., (2015). Understanding the physics
846 of kappa (κ): Insights from a downhole array. *Geophys J Int.* 203, no. 1, 678-691.

847 Ktenidou, O.J., Abrahamson, N., Darragh, R. and Silva, W., 2016. A methodology for the
848 estimation of kappa (κ) from large datasets, example application to rock sites in the NGA-
849 East database, and implications on design motions. PEER Report 2016, 1.

850 Konno, K. and Ohmachi, T., 1998. Ground-motion characteristics estimated from spectral ratio
851 between horizontal and vertical components of microtremor. *Bulletin of the Seismological*
852 *Society of America*, 88(1), pp.228-241.

853 Laurendeau, A., Cotton, F., Ktenidou, O. J., Bonilla, L. F., & Hollender, F. (2013). Rock and stiff-
854 soil site amplification: Dependency on VS 30 and kappa (κ 0). *Bull Seismol Soc Am.* 103,
855 no. 6, 3131-3148.

856 Lacave-Lachet, C., Bard, P.Y., Gariel, J.C. and Irikura, K., 2000. Straightforward methods to
857 detect non-linear response of the soil. Application to the recordings of the Kobe earthquake
858 (Japan, 1995). *Journal of seismology*, 4(2), pp.161-173.

859 Menq, F.Y., (2003). Dynamic properties of sandy and gravelly soils , Ph.D. Thesis, University of
860 Texas at Austin, Austin, Texas.

861 Nakano, K., Matsushima, S., & Kawase, H. (2015). Statistical properties of strong ground motions
862 from the generalized spectral inversion of data observed by K-NET, KiK-net, and the JMA
863 Shindokei network in Japan. *Bull Seismol Soc Am.* **105**, no. 5, 2662-2680.

864 Oth, A., Parolai, S. and Bindi, D., (2011). Spectral analysis of K-NET and KiK-net data in Japan,
865 Part I: Database compilation and peculiarities. *Bull Seismol Soc Am.* 101, no.2, 652-666.

866 Parolai, S. and Bindi, D., (2004). Influence of soil-layer properties on k evaluation. *Bulletin of the*
867 *Seismological Society of America*, 94(1), pp.349-356.

868 Pilz, M. and Cotton, F., 2019. Does the One-Dimensional Assumption Hold for Site Response
869 Analysis? A Study of Seismic Site Responses and Implication for Ground Motion
870 Assessment Using KiK-Net Strong-Motion Data. *Earthquake Spectra*, 35(2), pp.883-905.

871 Pilz, M., Cotton, F., Zaccarelli, R. and Bindi, D., (2019). Capturing Regional Variations of Hard-
872 Rock Attenuation in Europe. *Bull Seismol Soc Am*.

873 Perron, V., Hollender, F., Bard, P. Y., Gélis, C., Guyonnet-Benaize, C., Hernandez, B., & Ktenidou,
874 O. J. (2017). Robustness of kappa (κ) measurement in low-to-moderate seismicity areas:
875 Insight from a site-specific study in Provence, France. *Bull Seismol Soc Am*. 107, no.5,
876 2272-2292.

877 Pei, S., Cui, Z., Sun, Y., Toksöz, M.N., Rowe, C.A., Gao, X., Zhao, J., Liu, H., He, J. and Morgan,
878 F.D., 2009. Structure of the upper crust in Japan from S-wave attenuation tomography.
879 *Bulletin of the Seismological Society of America*, 99(1), pp.428-434.

880 Qin, L., Ben-Zion, Y., Bonilla, L.F. and Steidl, J.H., 2020. Imaging and Monitoring Temporal
881 Changes of Shallow Seismic Velocities at the Garner Valley Near Anza, California,
882 Following the M7. 2 2010 El Mayor-Cucapah Earthquake. *Journal of Geophysical*
883 *Research: Solid Earth*, 125(1), p.e2019JB018070.

884 Régnier, J., Cadet, H., Bonilla, L.F., Bertrand, E. and Semblat, J.F., 2013. Assessing nonlinear
885 behavior of soils in seismic site response: Statistical analysis on KiK-net strong-motion
886 data. *Bulletin of the Seismological Society of America*, 103(3), pp.1750-1770.

887 Régnier, J., Cadet, H. and Bard, P.Y., 2016. Empirical Quantification of the Impact of Nonlinear
888 Soil Behavior on Site Response. *Bulletin of the Seismological Society of America*, 106(4),
889 Behavior on Site Response. *Bulletin of the Seismological Society of America*, 106(4),
890 pp.1710-1719.

891 Rong, M., Wang, Z., Woolery, E.W., Lyu, Y., Li, X. and Li, S., 2016. Nonlinear site response
892 from the strong ground-motion recordings in western China. *Soil Dynamics and*
893 *Earthquake Engineering*, 82, pp.99-110.

894 Seed, H.B., Wong, R.T., Idriss, I.M. and Tokimatsu, K., 1986. Moduli and damping factors for
895 dynamic analyses of cohesionless soils. *Journal of geotechnical engineering*, 112(11),
896 pp.1016-1032.

897 Sonnemann, T. and Halldorsson, B., 2017, June. Towards an Automated Kappa Measurement
898 Procedure. In *International Conference on Earthquake Engineering and Structural*
899 *Dynamics* (pp. 39-52). Springer, Cham.

900 Thompson, E.M., Baise, L.G., Tanaka, Y. and Kayen, R.E., 2012. A taxonomy of site response
901 complexity. *Soil Dynamics and Earthquake Engineering*, 41, pp.32-43.

902 Trifunac, M.D. and Todorovska, M.I., 1996. Nonlinear soil response—1994 Northridge,
903 California, earthquake. *Journal of geotechnical engineering*, 122(9), pp.725-735.

904 Van Houtte, C., Drouet, S., & Cotton, F. (2011). Analysis of the origins of κ (kappa) to compute
905 hard rock to rock adjustment factors for GMPEs. *Bull Seismol Soc Am.* **101**, no. 6, 2926-
906 2941.

907 Van Houtte, C., Ktenidou, O.J., Larkin, T. and Holden, C., (2014). Hard-site κ_0 (kappa)
908 calculations for Christchurch, New Zealand, and comparison with local ground-motion
909 prediction models. *Bull Seismol Soc Am.* **104**, no. 4, 1899-1913.

910 Xu, B., Rathje, E.M., Hashash, Y., Stewart, J., Campbell, K. and Silva, W.J., 2020. κ_0 for soil
911 sites: Observations from Kik-net sites and their use in constraining small-strain damping
912 profiles for site response analysis. *Earthquake Spectra*, 36(1), pp.111-137.

913 Yu, G., Anderson, J.G. and Siddharthan, R.A.J., 1993. On the characteristics of nonlinear soil
914 response. Bulletin of the Seismological Society of America, 83(1), pp.218-244.

915 Zalachoris, G. and Rathje, E.M., 2015. Evaluation of one-dimensional site response techniques
916 using borehole arrays. Journal of Geotechnical and Geoenvironmental Engineering,
917 141(12), p.04015053.

918 Chunyang Ji
919 Department of Civil, Construction, and Environmental Engineering,
920 3101 Fitts-Woolard Hall, NC State University
921 Raleigh, NC 27695, the United States
922 Email: cji3@ncsu.edu
923
924 Ashly Cabas
925 Department of Civil, Construction, and Environmental Engineering,
926 3175 Fitts-Woolard Hall, NC State University
927 Raleigh, NC 27695, the United States
928 Email: amcabasm@ncsu.edu
929
930 Luis Fabian Bonilla
931 GERS-SRO, Univ. Gustave Eiffel,
932 F-77447 Marne-la-Vallée, France
933 Email: luis-fabian.bonilla-hidalgo@univ-eiffel.fr
934
935 Céline Gelis
936 IRSN
937 Paris Fontenay-aux-Roses, France
938 Email: celine.gelis@irsn.fr

939
940

Table 1. Local soil conditions, number of ground motions per dataset, predetermined fixed-frequency bandwidth and thresholds for shear strain index, I_γ at all study sites.

Station	V_{s30} ¹ (m/s)	V_{s0} ² (m/s)	$V_{s,depth}$ ³ (m/s)	Hole Depth (m)	Number of linear records	Number of transitional records	Number of nonlinear records	1 st resonant frequency (Hz)	Predominant frequency (Hz)	f_1 ⁴ (Hz)	f_2 ⁵ (Hz)	$I_{\gamma,0,l}$ ⁶ (%)	$I_{\gamma,0,t}$ ⁷ (%)
AICH17	314	150	2200	101	23	27	12	4.07	4.07	12.65	25.00	0.001	0.003
CHBH13	235	220	2920	1300	139	49	11	1.78	1.78	8.95	25.00	0.001	0.003
FKSH11	240	110	700	115	148	140	26	1.51	9.98	13.97	25.00	0.001	0.003
FKSH14	237	120	1210	147	114	221	28	1.12	4.15	10.05	25.00	0.001	0.007
FKSH18	307	140	2250	100	158	103	16	2.59	5.69	8.95	25.00	0.001	0.003
FKSH19	338	170	3060	100	185	95	21	3.27	3.27	10.05	25.00	0.001	0.003
FKSH21	365	200	1600	200	60	17	8	3.90	3.90	12.65	25.00	0.001	0.003
IBRH16	626	140	2050	300	137	81	15	1.71	7.08	10.05	25.00	0.001	0.003
IBRH17	301	90	2300	510	117	177	18	0.93	9.30	13.01	25.00	0.001	0.007
IBRH20	244	180	1200	923	133	86	11	0.27	0.27	8.95	25.00	0.001	0.007
IWTH21	521	150	2460	100	39	24	6	5.27	5.27	7.38	25.00	0.001	0.003
IWTH26	371	130	680	108	79	32	11	2.12	10.17	14.24	25.00	0.001	0.003
KMMH01	575	150	1900	100	94	24	15	4.03	9.98	13.97	25.00	0.001	0.003
KMMH12	410	210	1000	123	134	34	11	3.27	8.17	11.44	25.00	0.001	0.003
MYGH07	366	130	740	142	59	39	11	0.93	8.61	12.06	25.00	0.001	0.003
MYGH10	348	110	770	205	229	132	16	0.95	10.66	14.93	25.00	0.001	0.007
NGNH29	465	150	1040	110	81	38	16	1.95	6.93	10.05	25.00	0.001	0.003
NIGH07	528	200	1600	106	29	10	11	4.12	7.08	10.05	25.00	0.001	0.003
NIGH12	553	240	780	110	29	9	11	2.00	5.00	12.65	25.00	0.001	0.003
TCGH16	213	80	680	112	112	334	35	1.27	4.81	11.27	25.00	0.001	0.007

941 ¹ V_{s30} : time averaged shear-wave velocity in the top 30 m of the soil profile

942 ² V_{s0} : shear-wave at the ground surface

943 ³ $V_{s,depth}$: shear-wave velocity at the depth of the borehole sensor

- 944 4f_l : the lower frequency limit to estimate individual κ_r
- 945 5f_2 : the upper frequency limit to estimate individual κ_r
- 946 ${}^6I_{\gamma,0,l}$: the shear-strain index threshold to separate linear and transitional datasets
- 947 ${}^7I_{\gamma,0,t}$: the shear-strain index threshold to separate transitional and nonlinear datasets

948

949 **Table 2.** Ground motion datasets constructed via alternative approaches (AP1 to AP4) explored
 950 in this study to implement the κ_0 -model.

Approach	κ_0 _lin_sur	κ_0 _nl_sur	κ_0 _tran_sur	κ_0 _depth	κ_0 -model
AP1	Linear dataset	Nonlinear dataset	--	Borehole dataset	Equation (5)
AP2	Linear and transitional datasets	Nonlinear dataset	--	Borehole dataset	Equation (5)
AP3	Linear dataset	Nonlinear and transitional datasets	--	Borehole dataset	Equation (5)
AP4	Linear dataset	Nonlinear dataset	Transitional datasets	Borehole dataset	Equation (5)

951

952 **Table 3.** Site-specific κ_0 values obtained from different dataset definitions at stations FKSH14
 953 and MYGH10.

	Approach	κ_0 _lin_sur (s)	κ_0 _nl_sur (s)	κ_0 _tran_sur (s)
FKSH14	AP1	0.0488	0.0633	--
	AP2	0.0565	0.0638	--
	AP3	0.0500	0.0607	--
	AP4	0.0499	0.0638	0.0602
	Maximum difference	15.83%	5.10%	
MYGH10	AP1	0.0567	0.0563	--
	AP2	0.0568	0.0560	--
	AP3	0.0559	0.0572	--
	AP4	0.0560	0.0558	0.0575
	Maximum difference	1.56%	2.53%	

954

955

956

957

958

959 **Table 4.** T-test results for κ_{0_pred} values estimated with linear and nonlinear datasets defined by
 960 AP3 (which includes records categorized as transitional into the nonlinear dataset).

Station	V_{s30} (m/s)	p-value	Statistically different
		Linear κ_{0_pred} vs. nonlinear κ_{0_pred} *	
AICH17	314	10.37%	--
CHBH13	235	<0.01%	Yes
FKSH11	240	1.07%	Yes
FKSH14	237	<0.01%	Yes
FKSH18	307	2.36%	Yes
FKSH19	338	57.66%	--
FKSH21	365	47.12%	--
IBRH16	626	13.80%	--
IBRH17	301	55.92%	--
IBRH20	244	5.97%	--
IWTH21	521	9.23%	--
IWTH26	371	0.62%	Yes
KMMH01	575	0.32%	Yes
KMMH12	410	3.38%	Yes
MYGH07	366	0.04%	Yes
MYGH10	348	7.45%	--
NGNH29	465	72.74%	--
NIGH07	528	88.97%	--
NIGH12	553	66.61%	--
TCGH16	213	<0.01%	Yes

961
 962 *The transitional dataset is included into the nonlinear dataset (i.e., datasets following the criteria of AP3).

963 **List of Figure Captions:**

964 **Figure 1.** Comparisons between a weak and a strong ground motion recorded at FKSH14 ($V_{s30} =$
965 237 m/s). The M_w and R_e are 4 and 15 km for the low-intensity ground motion, while 5.1 and 15
966 km for the high-intensity ground motion. The frequency window ([10.05, 30] Hz) applied in this
967 plot is picked manually. The left and right columns correspond to analyses conducted on the
968 horizontal components H_1 and H_2 , respectively. The color version of this figure is available only
969 in the electronic version of this article.

970 **Figure 2.** (a) Locations of selected Japanese recording stations in this study, and (b) magnitude
971 and distance distribution of selected ground motions.

972 **Figure 3.** Hyperbolic models fitted to observed PGA_{rotD50} and $I_{\gamma,0}$ data at four study sites. The V_{s30}
973 for IBRH16, IBRH17, IBRH20, and IWTH21 are 626, 301, 244, and 521 m/s, respectively.

974 **Figure 4.** Surface PGA_{rotD50} against $I_{\gamma,0}$ at MYGH10 ($V_{s30} = 348$ m/s). The red dot-dashed lines
975 present the linear ($I_{\gamma,0,l}$) and transitional ($I_{\gamma,0,t}$) thresholds for $I_{\gamma,0}$. The color version of this figure is
976 available only in the electronic version of this article.

977 **Figure 5.** G/G_{max} versus $I_{\gamma,0}$ at study sites. The G_{max} is computed from the average values of

978 $\left(\frac{PGA_{rotD50}}{PGV_{rotD50}/V_{s,0}} \right)$ for records with $I_{\gamma,0}$ less than 0.001%. The colors represent PGA_{rotD50} values. The

979 color version of this figure is only available in the electronic version of this article. The color

980 version of this figure is available only in the electronic version of this article.

981 **Figure 6.** Comparisons of individual κ_r estimates from our automated algorithm, $\kappa_{r,auto}$, and the
982 fixed-frequency band method, κ_{r_AS} at FKSH14 ($V_{s30} = 237$ m/s) for surface (left) and borehole
983 (right) records. The color version of this figure is available only in the electronic version of this
984 article.

985 **Figure 7.** Surface κ_{r_AS} estimates and their corresponding PGA_{rotD50} , $I_{\gamma,0}$ and R_e values for selected
986 ground motions recorded at (a) FKSH14 ($V_{s30} = 237$ m/s) and (b) MYGH10 ($V_{s30} = 348$ m/s).
987 Different colors represent varying epicentral distances per record, and the size of markers indicate
988 the corresponding PGA_{rotD50} . The color version of this figure is only available in the electronic
989 version of this article. The color version of this figure is available only in the electronic version of
990 this article.

991 **Figure 8.** κ_0 -model at FKSH14 ($V_{s30} = 237$ m/s) from datasets defined by (a) AP1, which only
992 considers the linear and nonlinear datasets, (b) AP2, where transitional records are included as part
993 of the linear dataset, (c) AP3, where transitional records are included as part of the nonlinear
994 dataset, and (d) AP4, where the linear, transitional, and nonlinear datasets are considered
995 separately. The color version of this figure is available only in the electronic version of this article.

996 **Figure 9.** κ_0 -model at MYGH10 ($V_{s30} = 348$ m/s) with datasets defined by (a) AP1, which only
997 considers the linear and nonlinear datasets, (b) AP2, where transitional records are included as part
998 of the linear dataset, (c) AP3, where transitional records are included as part of the nonlinear
999 dataset, and (d) AP4, where the linear, transitional, and nonlinear datasets are considered
1000 separately. The color version of this figure is only available in the electronic version of this article.
1001 The color version of this figure is available only in the electronic version of this article.

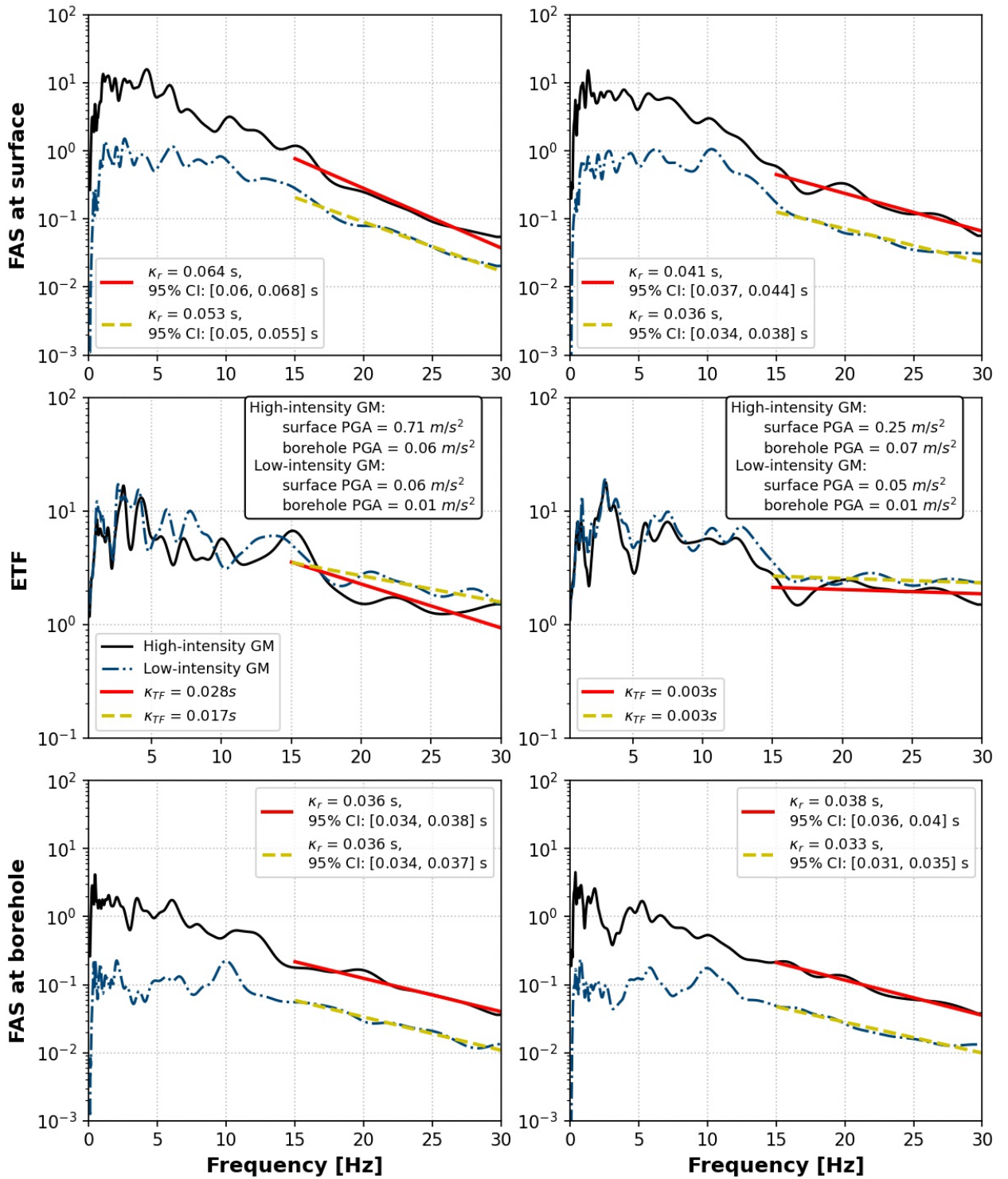
1002 **Figure 10.** Ratio of $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ at study sites estimated using the dataset definitions based on
1003 AP3 (left panel) and AP4 (right panel) against to V_{s5} , V_{s10} , and V_{s30} . The color version of this
1004 figure is available only in the electronic version of this article.

1005 **Figure 11.** Estimated surface κ_{0_pred} and their corresponding ground shaking intensity and in situ
1006 deformation characterized by PGA_{rotD50} and $I_{\gamma,0}$, respectively at (a) FKSH14 ($V_{s30} = 237$ m/s) and
1007 (b) MYGH10 ($V_{s30} = 348$ m/s). Both, the color and the size of markers represent varying PGA_{rotD50}

1008 values. Triangles and circles represent the linear and nonlinear datasets defined by AP3. The red
1009 dashed lines depict the local regression model based on the κ_{0_pred} and $I_{\gamma,0}$ data. The color version
1010 of this figure is only available in the electronic version of this article. The color version of this
1011 figure is available only in the electronic version of this article.

1012 **Figure 12.** Observed distribution of κ_{0_pred} at (a) FKSH14 ($V_{s30} = 237$ m/s) and (b) MYGH10 (V_{s30}
1013 = 348 m/s) for the linear and nonlinear datasets. The red lines depict the theoretical probability
1014 density function (PDF) fitted with a Gaussian distribution. The color version of this figure is
1015 available only in the electronic version of this article.

1016



1018

1019

1020

Figure 1. Comparisons between a weak and a strong ground motion recorded at FKSH14 ($V_{s30} = 237$ m/s). The M_w and R_e are 4 and 15 km for the low-intensity ground motion,

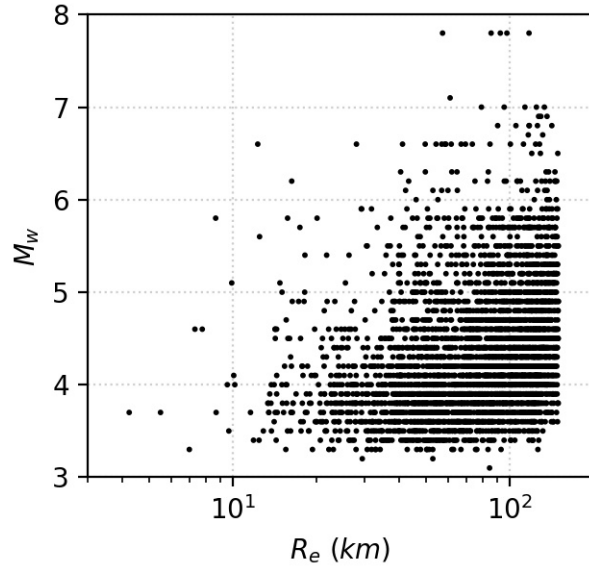
1021 while 5.1 and 15 km for the high-intensity ground motion. The frequency window ([10.05,
1022 30] Hz) applied in this plot is picked manually. The left and right columns correspond to
1023 analyses conducted on the horizontal components H_1 and H_2 , respectively. The color
1024 version of this figure is available only in the electronic version of this article.

1025

1026

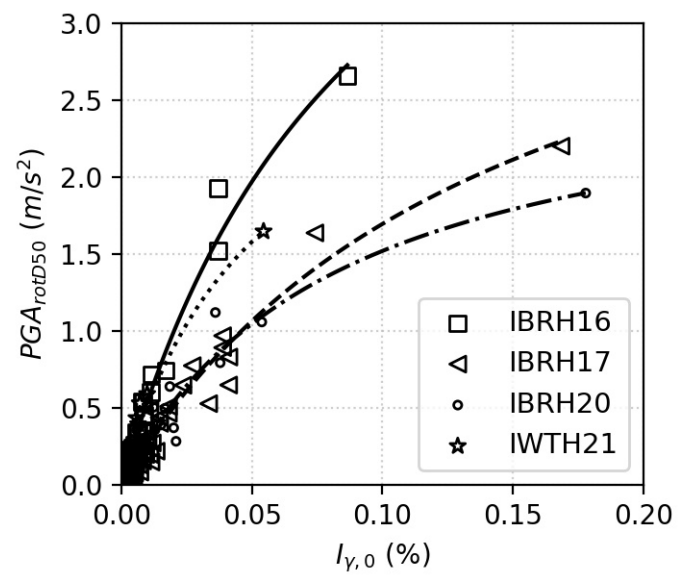


1027



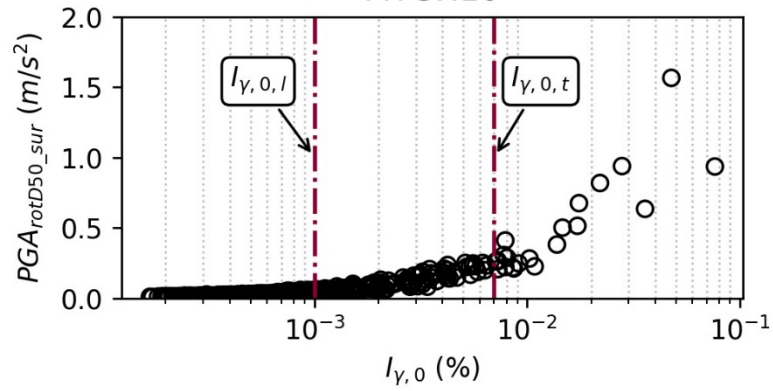
1028
 1029
 1030
 1031
 1032

Figure 2. (a) Locations of selected Japanese recording stations in this study, and (b) magnitude and distance distribution of selected ground motions.



1033
 1034
 1035
 1036

Figure 3. Hyperbolic models fitted to observed PGA_{rotD50} and $I_{y,0}$ data at four study sites. The V_{s30} for IBRH16, IBRH17, IBRH20, and IWTH21 are 626, 301, 244, and 521 m/s, respectively.

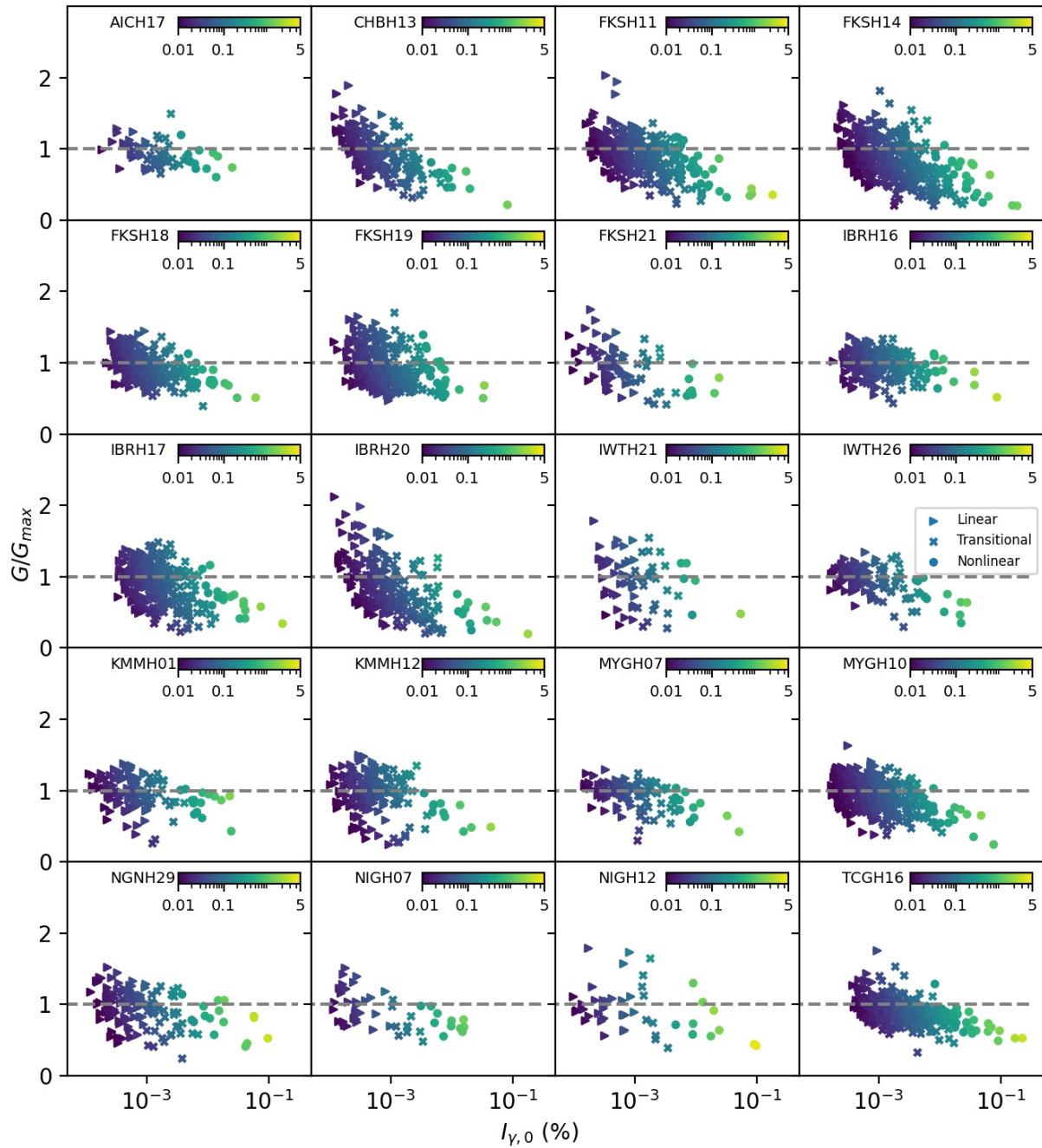


1037

1038 **Figure 4.** Surface PGA_{rotD50_sur} against $I_{\gamma,0}$ at MYGH10 ($V_{s30} = 348$ m/s). The red dot-dashed lines
 1039 present the linear ($I_{\gamma,0,l}$) and transitional ($I_{\gamma,0,t}$) thresholds for $I_{\gamma,0}$. The color version of this figure

1040 is available only in the electronic version of this article.

1041



1042

1043

Figure 5. G/G_{\max} versus $I_{\gamma,0}$ at study sites. The G_{\max} is computed from the average values of

1044

$$\left(\frac{PGA_{rotD50}}{PGV_{rotD50}/V_{s,0}} \right)$$

for records with $I_{\gamma,0}$ less than 0.001%. The colors represent PGA_{rotD50} values (in

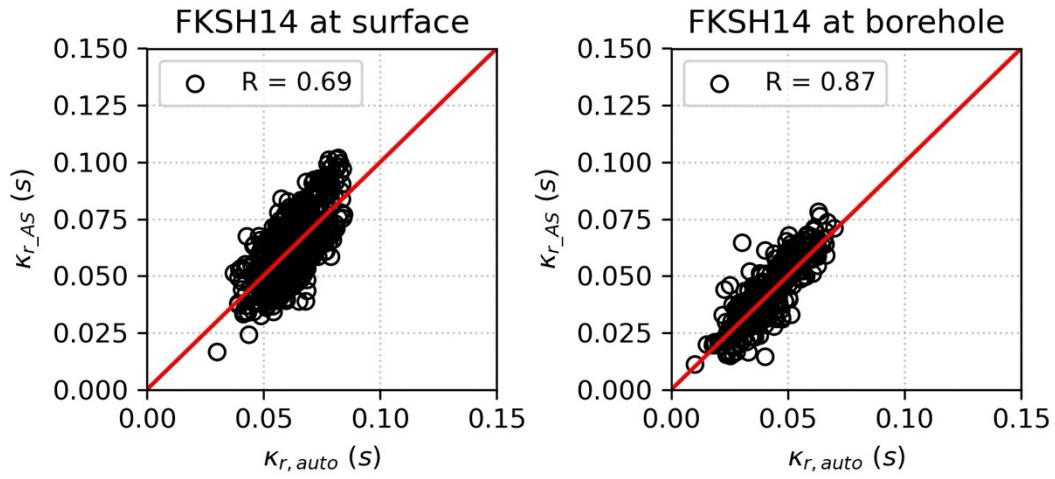
1045

units of m/s^2). The color version of this figure is only available in the electronic version of this

1046

article. The color version of this figure is available only in the electronic version of this article.

1047



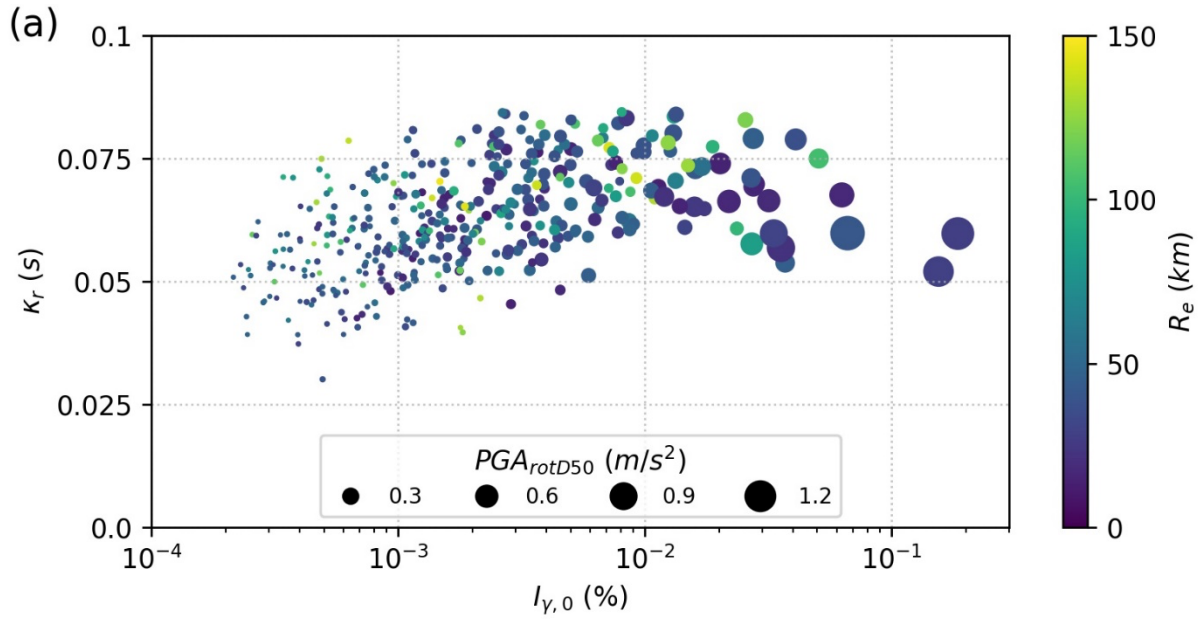
1048

1049 **Figure 6.** Comparisons of individual κ_r estimates from our automated algorithm, κ_{r_auto} , and the
1050 fixed frequency band method, κ_{r_AS} at FKSH14 ($V_{s30} = 237$ m/s) for surface (left) and borehole
1051 (right) records. The color version of this figure is available only in the electronic version of this

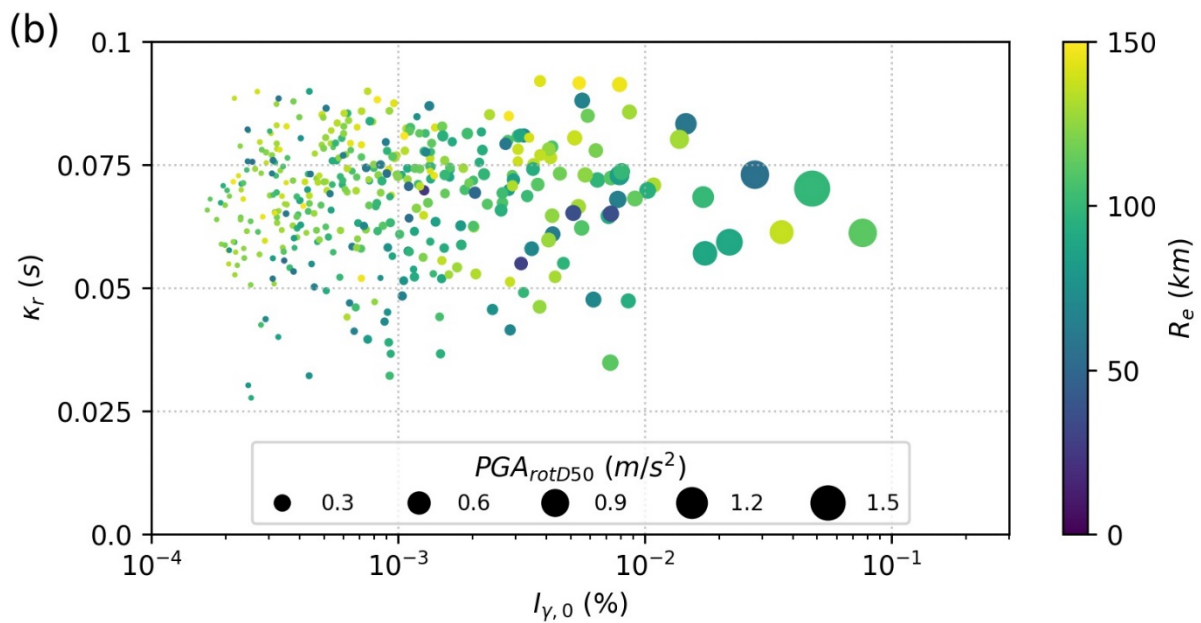
1052

article.

1053



1054



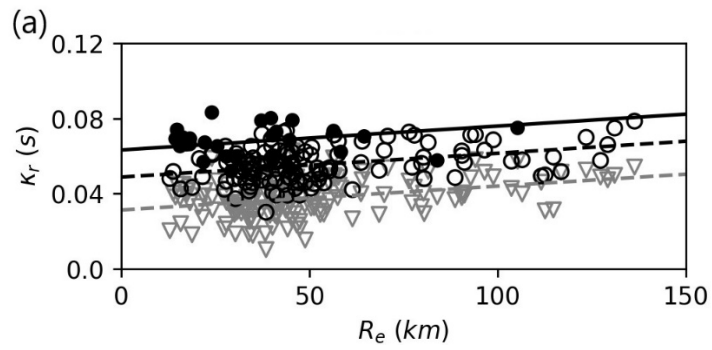
1055

1056 **Figure 7.** Surface κ_{r_AS} estimates and their corresponding PGA_{rotD50} , $I_{\gamma,0}$ and R_e values for
 1057 selected ground motions recorded at (a) FKSH14 ($V_{s30} = 237$ m/s) and (b) MYGH10 ($V_{s30} = 348$
 1058 m/s). Different colors represent varying epicentral distances per record, and the size of markers
 1059 indicate the corresponding PGA_{rotD50} . The color version of this figure is only available in the
 1060 electronic version of this article. The color version of this figure is available only in the
 1061 electronic version of this article.

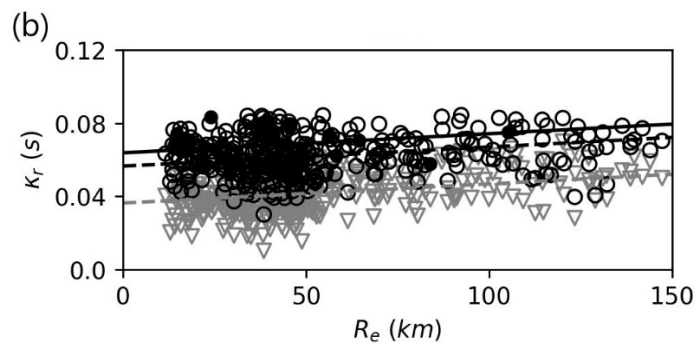
1062

1063

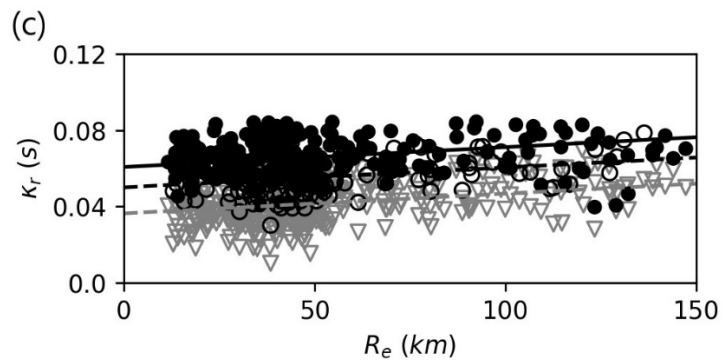
1064



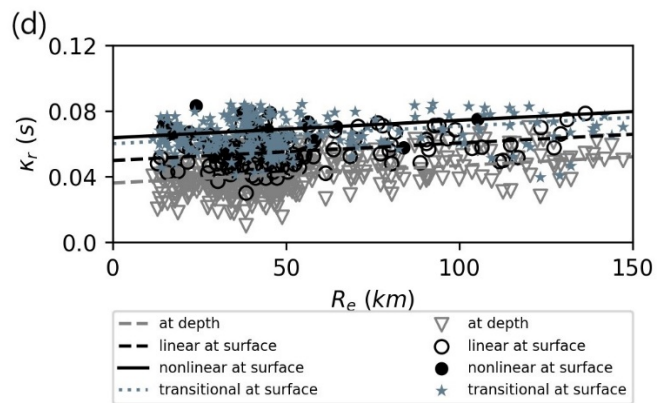
1065



1066

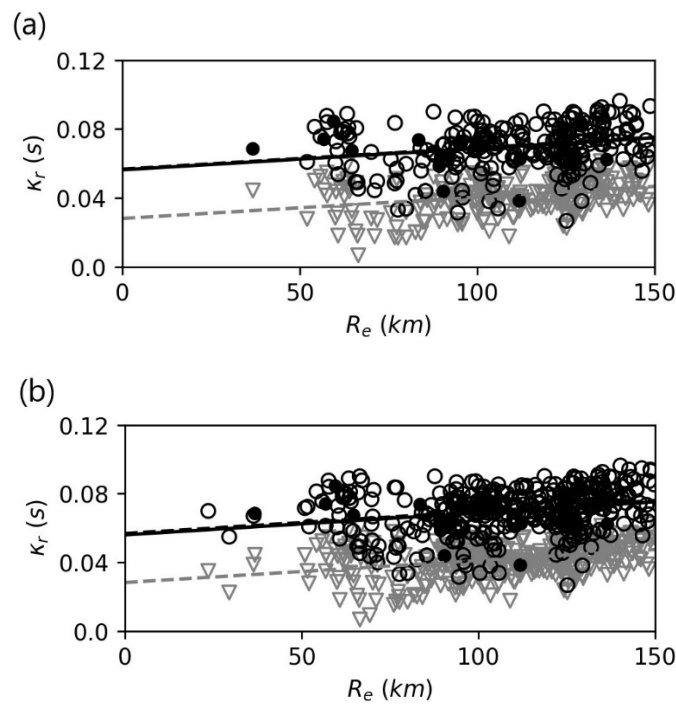


1067



1068

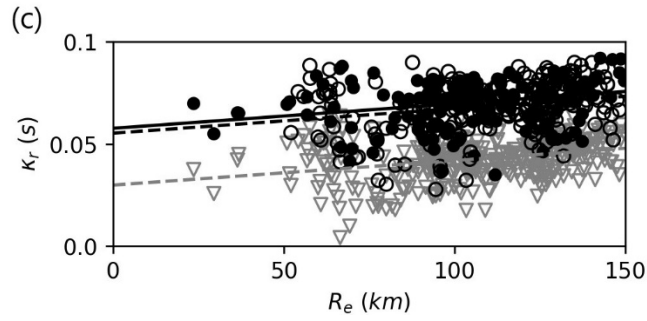
1069 **Figure 8.** κ_0 -model at FKSH14 ($V_{s30} = 237$ m/s) from datasets defined by (a) AP1, which only
1070 considers the linear and nonlinear datasets, (b) AP2, where transitional records are included as
1071 part of the linear dataset, (c) AP3, where transitional records are included as part of the nonlinear
1072 dataset, and (d) AP4, where the linear, transitional, and nonlinear datasets are considered
1073 separately. The color version of this figure is only available in the electronic version of this
1074 article.



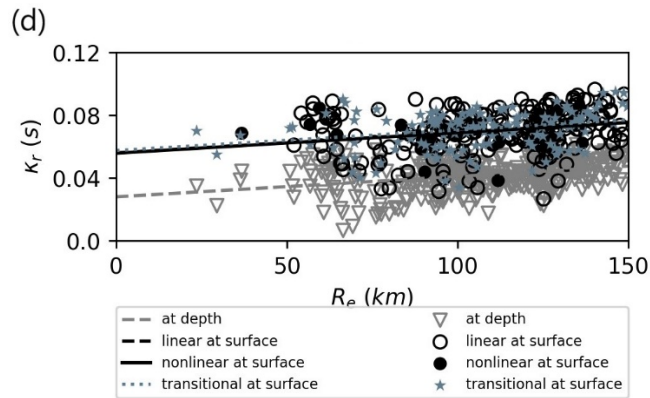
1077

1078

1079

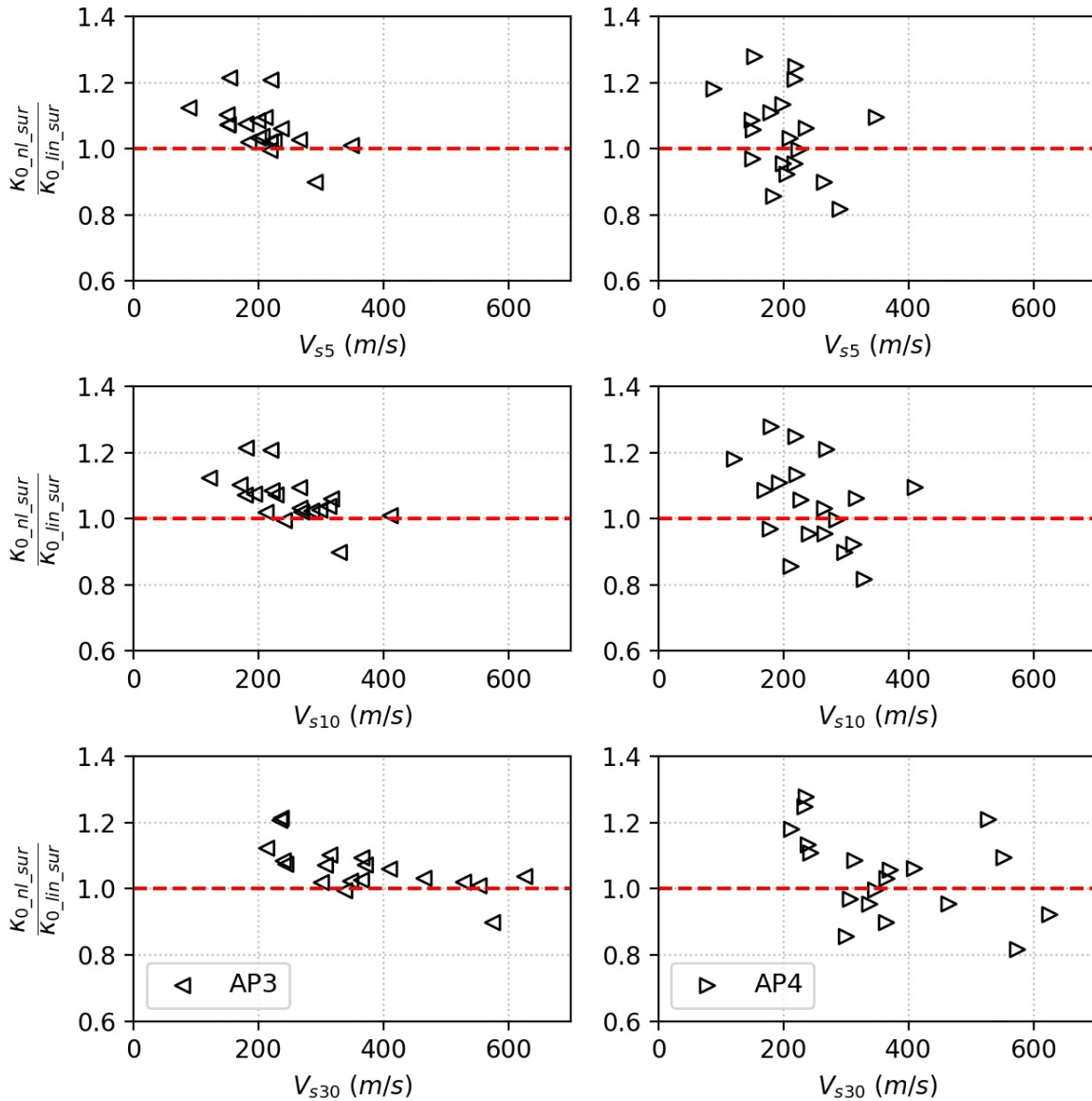


1080



1081 **Figure 9.** κ_0 -model at MYGH10 ($V_{s30} = 348$ m/s) with datasets defined by: (a) AP1, which only
1082 considers the linear and nonlinear datasets, (b) AP2, where transitional records are included as
1083 part of the linear dataset, (c) AP3, where transitional records are included as part of the nonlinear
1084 dataset, and (d) AP4, where the linear, transitional, and nonlinear datasets are considered
1085 separately. The color version of this figure is available only in the electronic version of this
1086 article.

1087



1088

1089 **Figure 10** Ratio of $\kappa_{0_nl_sur}/\kappa_{0_lin_sur}$ at study sites estimated using the dataset definitions based on

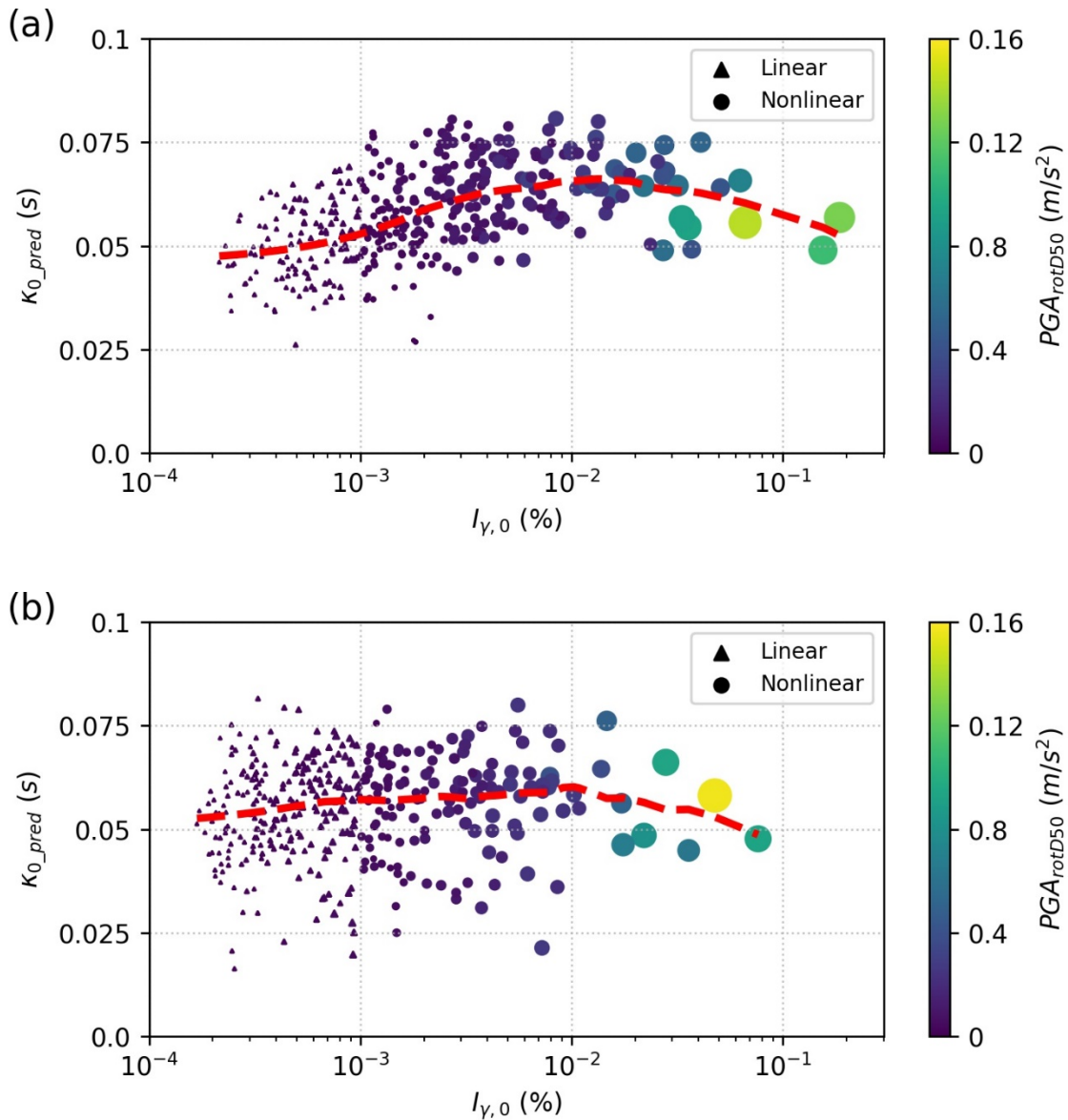
1090 AP3 (left panel) and AP4 (right panel) against to V_{s5} , V_{s10} , and V_{s30} . The color version of this

1091 figure is available only in the electronic version of this article.

1092

1093

1094



1095

1096

1097 **Figure 11.** Estimated surface κ_{0_pred} and their corresponding ground shaking intensity and in situ
 1098 deformation characterized by PGA_{rotD50} and $I_{\gamma,0}$, respectively at (a) FKSH14 ($V_{s30} = 237$ m/s) and

1099 (b) MYGH10 ($V_{s30} = 348$ m/s). Both, the color and the size of markers represent varying

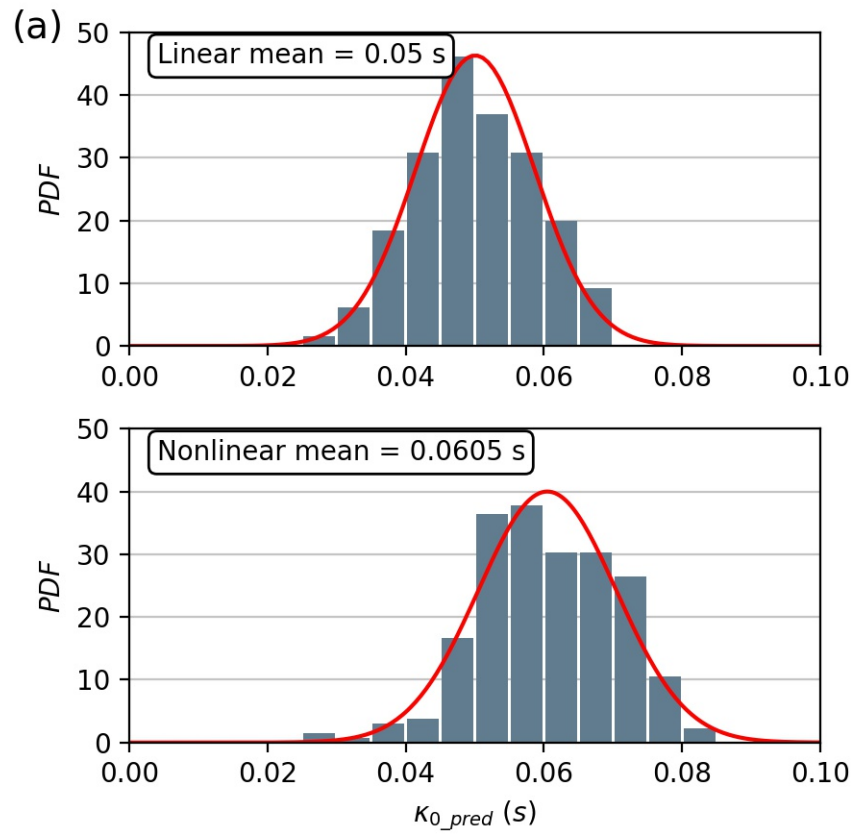
1100 PGA_{rotD50} values. Triangles and circles represent the linear and nonlinear datasets defined by

1101 AP3. The red dashed lines depict the local regression model based on the κ_{0_pred} and $I_{\gamma,0}$ data. The

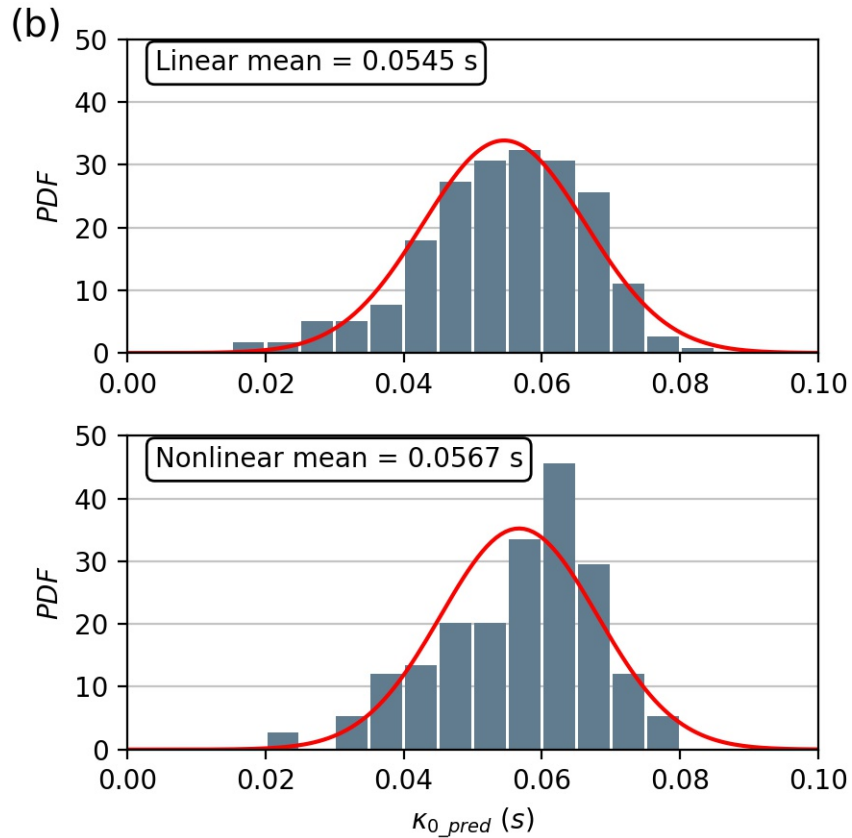
1102 color version of this figure is only available in the electronic version of this article. The color

1103 version of this figure is available only in the electronic version of this article.

1104



1105



1106

1107

1108 **Figure 12.** Observed distribution of κ_{0_pred} at (a) FKSH14 ($V_{s30} = 237$ m/s) and (b) MYGH10

1109 ($V_{s30} = 348$ m/s) for the linear and nonlinear datasets. The red lines depict the theoretical

1110 probability density function (PDF) fitted with a Gaussian distribution. The color version of this

1111 figure is available only in the electronic version of this article.