



**HAL**  
open science

## Impact of spatial variability of shear wave velocity on the lagged coherency of synthetic surface ground motions

E. El Haber, C. Cornou, D. Jongmans, Fernando Lopez-caballero, D. Youssef Abdelmassih, T. Al-Bittar

### ► To cite this version:

E. El Haber, C. Cornou, D. Jongmans, Fernando Lopez-caballero, D. Youssef Abdelmassih, et al.. Impact of spatial variability of shear wave velocity on the lagged coherency of synthetic surface ground motions. *Soil Dynamics and Earthquake Engineering*, 2021, 145, pp.106689. 10.1016/j.soildyn.2021.106689 . hal-03176493

**HAL Id: hal-03176493**

<https://hal.science/hal-03176493v1>

Submitted on 22 Mar 2023

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Impact of spatial variability of shear wave velocity on the lagged coherency of synthetic**  
2 **surface ground motions** E. EL HABER<sup>1,2,\*</sup>, C. CORNOU<sup>1</sup>, D. JONGMANS<sup>1</sup>, F. LOPEZ-  
3 CABALLERO<sup>3</sup>, D. YOUSSEF ABDELMASSIH<sup>2</sup>, T. AL-BITTAR<sup>2</sup>

4  
5 <sup>1</sup>*Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTERre, 38000 Grenoble, France.*

6 <sup>2</sup>*Scientific Research Center in Engineering, (CSRI), Faculty of Engineering, Geotechnical Engineering,*  
7 *Lebanese University, Lebanon.*

8 <sup>3</sup>*MSSMat laboratory, CNRS UMR 8579, CentraleSupélec Paris-Saclay Univ., 3 Rue Joliot-Curie, 91190 Gif-*  
9 *Sur-Yvette, France.*

10

11 **Abstract**

12 *The spatial incoherence of ground motion during an earthquake can have a significant effect on the dynamic*  
13 *response of engineering structures such as bridges, dams, nuclear power plants and lifeline facilities. The main*  
14 *objective of this paper is to study the effect of anisotropic heterogeneities in a soil layer overlying homogeneous*  
15 *bedrock on the lagged coherency of surface ground motion. A set of numerical experiments is performed based*  
16 *on 2D spatial variability of shear-wave velocities modeled as a homogeneous stationary random field and*  
17 *discretized by the EOLE method (Expansion Optimal Linear Estimation). Seismic ground motions were*  
18 *simulated using FLAC<sup>2D</sup> software in the 1-25 Hz band for a plane wave excitation with SV polarization. The soil*  
19 *is characterized by horizontal and vertical autocorrelation distances ranging between 5 and 20 m and 1 and 2*  
20 *m, respectively, and a coefficient of variation of the shear-wave velocity varying between 5% and 40%. The*  
21 *synthetic seismograms calculated for 9 parameter sets (100 realizations each) clearly show seismic waves*  
22 *scattering and surface waves diffracted locally by the ground heterogeneities, generating large spatial*  
23 *variations in coherence mainly controlled by the coefficient of variation of shear-wave velocity. Consistently*  
24 *with existing models and experimental data, the numerical coherency curves decrease with frequency and*  
25 *receiver distance, however at a rate which is lower than that observed in the experimental data. This difference*  
26 *is probably due to intrinsic attenuation that is not accounted for in the simulations and/or to our 2D simulations*  
27 *that do not reproduce the complete wavefield. The numerical average coherency curves for each parameter set*

---

\* [elias.el-haber@univ-grenoble-alpes.fr](mailto:elias.el-haber@univ-grenoble-alpes.fr); <https://orcid.org/0000-0002-0864-7314>; Tel: +33 6 29 83 77 28

38 *exhibit maxima within narrow frequency bands* caused by the vertically trapped body waves and surface wave  
39 *propagation properties within the average ground model. This interpretation is supported by experimental data*  
30 *recorded in the Koutavos-Argostoli valley (Greece).*

31 *Key words: coherency, spatial variability, seismic response, random field, autocorrelation.*

32

## 33 **1 Introduction**

34 Large earthquakes have often showed spatial variations in damages at the scale from tens to hundreds of meters  
35 (e.g. Northridge in 1994 [1] and Christchurch in 2011 [2]). Such spatial distribution can be related to spatial  
36 variations of earthquake ground motion (SVEGM; [3]). SVEGM may have remarkable impact on extended and  
37 large structures such as bridges and lifelines (see [4] for a summary) and should be considered in the study of  
38 their responses and design -[5, 6, 7, 8 and 9]. With regard to seismic design of bridges, SVEGM has been  
39 recently incorporated in codes, e.g. EUROCODE8 and Caltrans design procedure [10], that highlighted the  
40 importance of assessing it at various spatial scales. SVEGM can be caused by the source mechanism and  
41 radiation (magnitude, slip distribution, directivity effects, ...), path effects (regional heterogeneities, ...) and site  
42 effects including both topographical and lithological effects [among others; 11, 12; 13; 14, 15, 16, 17, 18]. In  
43 addition to these various causes of SVGEM, shallow layers may exhibit spatial heterogeneities resulting from the  
44 natural processes of erosion and sedimentation or from human activities (constructions, mining deposits, fills ...)  
45 [19, 20, 21, 22, 23], which can lead to significant variations in the soil seismic response even at small distances  
46 [24]. To account for these local variations, probabilistic modeling approaches using random fields [25] have  
47 been widely applied in geotechnical studies [among others, 26, 27, 28]. In earthquake engineering, several  
48 studies carried out over the past 20 years have emphasized the effect of the variability of seismic soil properties  
49 on SVEGM mainly considering 1D stochastic seismic wave propagation [among other, 29, 30, 31, 32], few  
50 studies focusing on 2D stochastic wave propagation [among others; 33, 34, 35, 36, 37, 33].

51 While the full SVEGM estimation implies determining the phase and amplitude variability [38], the phase  
52 variability evaluation has deserved more attention from the scientific community. Indeed, many authors [e.g, 39,  
53 40, 41, 4, 42] have used the complex function called coherency to characterize the phase variability of surface  
54 ground motion. Coherency models provide a quantitative measure of the similarity between recordings of the  
55 same earthquake at different locations and frequencies [44] and can be used to incorporate SVEGM in the

56 seismic design of structures [4, 42]. Empirical coherency models have been proposed based on the analysis of  
57 ground motions recorded by dense arrays [among others, 44, 45, 40, 46, 47]. Alternatively, other studies have  
58 focused on analytical calculations and/or ground motion modeling in isotropic heterogeneous media. They have  
59 resulted in semi-empirical coherency models that take into account characteristics of the soil spatial variability in  
60 terms of coefficient of variation of shear modulus and related fluctuation scale [48, 49, 50]. However, modeling  
61 the soil as an isotropic medium can be unrealistic since the vertical variability of soil is generally greater than the  
62 horizontal variability due to natural geological processes that result in a greater distance of horizontal than  
63 vertical fluctuation [51, 52, 53]].

64 The main objective of this paper is to study the effect of the ground spatial variability on the lagged coherency,  
65 by performing 2D numerical modeling using statistical parameters found in the literature. We based our study on  
66 the geometry and parameters determined for the alluvial plain of Nahr Beirut (Lebanon) [54] and considered a  
67 simple 2D model consisting of a soil layer with anisotropic spatial variability overlying a homogeneous elastic  
68 bedrock. This model was already used in the paper of El Haber et al. [37] for studying the effect of  
69 heterogeneous soil properties on surface ground motion spatial variability. In addition to studying the effect of  
70 the spatial variability of soil properties ( $V_s$ ) on the spatial coherency properties (average and standard deviation),  
71 the numerical results also help to better understand some abrupt variation of coherency within narrow frequency  
72 band observed in Koutavos-Argostoli experimental data [55].

## 73 **2 Spatially variable $V_s$ profile generation and wave propagation simulation**

74 In probabilistic approaches, the random field theory is widely used to model the spatial variability of  
75 geotechnical and geophysical soil properties [56]. In this paper, the shear-wave velocity ( $V_s(x,z)$ ) is modeled as  
76 a random field defined by a probability density function (*pdf*) described by a mean ( $\mu_{V_s}$ ), a coefficient of  
77 variation (*COV*) and an autocorrelation function  $\rho(\theta_x, \theta_z)$ ,  $\theta_x$  and  $\theta_z$  being the autocorrelation distances in the  
78 horizontal and vertical directions of the 2D medium, respectively. The autocorrelation distance is the distance  
79 over which the estimated values of  $V_s$  are not correlated any more. *COV* is defined as the ratio of the standard  
80 deviation of  $V_s$  over  $\mu_{V_s}$ .

81 The numerical models and ground motion simulation scheme used in this study are identical to those described  
82 in the paper of El Haber et al. [37], who numerically studied the effect of 2D soil spatial variability on the spatial  
83 variability of earthquake ground motion. The reader may refer to Section 2 of this paper for details on how to

84 introduce ground structure spatial variability into wave propagation simulation. Here, we just recall the main  
85 characteristics of the model and the ground motion modeling methods.

86 The soil properties for the parametric study is shown in Table 1. It consists of a 15.5 m thick soil layer with an  
87 average  $\mu_{V_s}=220$  m/s, overlying a homogeneous bedrock with  $V_s=1000$  m/s.  $V_s(x,z)$  in the soil layer is modeled  
88 as a 2D random field and the statistical parameters are chosen in a range covering the values found in the  
89 literature ( $5\% < COV < 40\%$ ;  $5 < \theta_x < 20$  m;  $1 < \theta_z < 2$  m). The 9 probabilistic models with the different  
90 parameter sets are given in Table 1.

91 To discretize the random field  $V_s$  described by a lognormal probability density function (*pdf*) and determine  $V_s$   
92 at the center of each mesh (i.e. midpoint method) [57], we use the Expansion Optimal Linear Estimation (*EOLE*)  
93 [58] and the ordinary kriging method [59].

94 For each  $V_s$  realization, the seismic ground motions are calculated using the finite difference code (FLAC<sup>2D</sup>)  
95 [60] at 166 receivers located on the surface with a spacing of 1 m by applying a seismic shear stress excitation  
96 (vertical incident plane wave of type *SV*) at the base of the 2D model after initializing the stresses. Only the  
97 horizontal displacements are fixed along the lateral boundaries of the model, while the horizontal and vertical  
98 movements are fixed at its base. Free field boundaries are applied to the side edges of the model. A quiet  
99 boundary (absorbing effect defined in FLAC<sup>2D</sup> [60]), is applied to the model base in order to model a flexible  
100 base that absorbs the energy emitted by the waves reflected on the surface and arriving to the model base. The  
101 source time function is a pseudo-Dirac having a flat Fourier amplitude spectrum equals to 1 between 1 and 25  
102 Hz. No damping is considered in the analysis in order to focus on the effects of elastic properties variability,  
103 only.

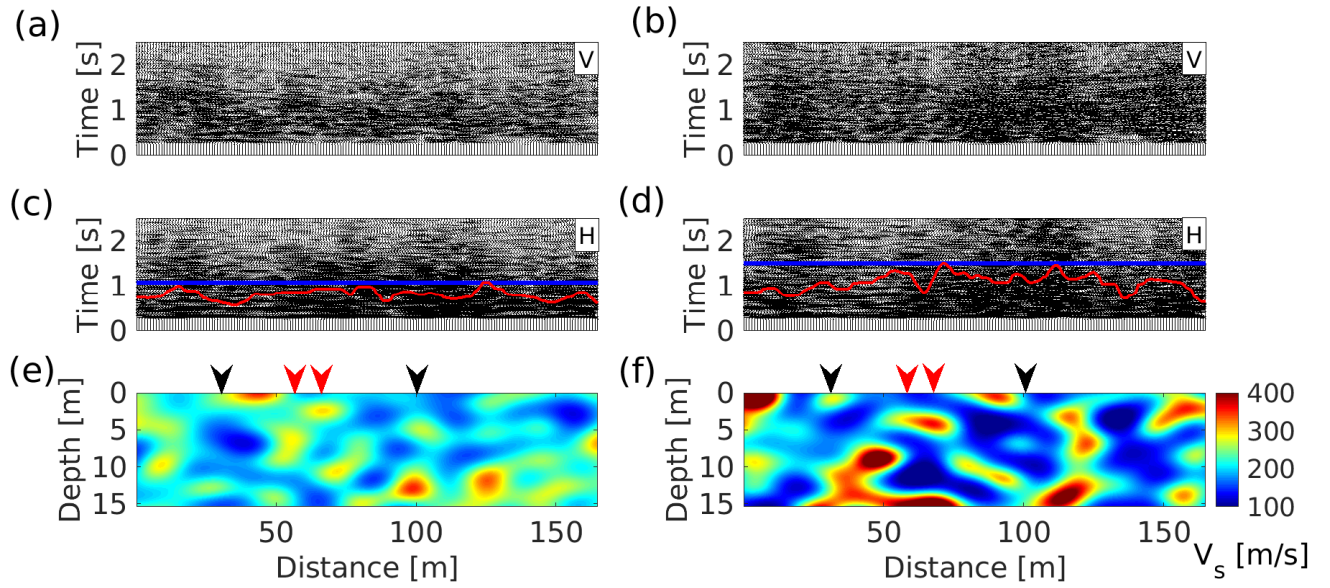
104 In probabilistic modeling approaches, a large number of discretized  $V_s$  realizations must be simulated to ensure  
105 the statistical convergence of the average and standard deviation estimators of any surface ground motion  
106 parameter. This convergence is considered to be reached when the fractions of the relative variation of the  
107 average and standard deviation between realization  $i$  and realization  $i+1$  are smaller than 5% (for more details  
108 refer to El Haber et al. [37]). In this study, for each parameter set, 100  $V_s$  realizations are sufficient to ensure the  
109 statistical convergence of the lagged coherency. The fundamental resonance frequency of the soil layer  
110 characterized by a mean  $V_s=220$  m/s and a thickness of 15.5 m is  $f_{0D} = 3.54$  Hz; this case will be referred to as  
111 the deterministic model in the following.

112 Table 1 : Properties of the two layers (soil and bedrock) for the probabilistic modeling. The bedrock properties  
 113 are fixed, as well as the density  $\rho$  and the P-wave velocity  $Vp$  in the soil layer.  $Vs$  in the soil is considered as a  
 114 random field and characterized by four statistical parameters ( $\mu_{Vs}$ ,  $COV$ ,  $\theta_x$  and  $\theta_z$ ). Nine parameter sets were  
 115 defined, combining variations in  $COV$ ,  $\theta_x$  and  $\theta_z$  (see text for details) (Modified from El Haber et al. [37]).

<i>Layers</i>	<i>Properties</i>						<i>Parameter set #</i>
	$\rho$ (Kg/m3)	$Vp$ (m/s)	$\mu_{Vs}$ (m/s)	$COV$ (%)	$\theta_x$ (m)	$\theta_z$ (m)	
<i>Soil</i>	1600	1500	220	5	5	2	1
					10	2	2
					5	2	3
				20	10	1	4
					2	2	5
					20	2	6
				40	5	2	7
					10	1	8
					2	2	9
<i>Bedrock</i>	2500	3000	1000	---	---	---	---

116

117



118

119 Figure 1: Seismic responses for the two  $V_s$  realizations of parameter sets #5 (left column) and # 9 (right  
 120 column). a) and b) vertical velocities ( $V$ ) computed for the  $V_s$  random field realization shown in e) and f),  
 121 respectively. c) and d) horizontal velocities ( $H$ ) ; e) and f)  $V_s$  realizations for the parameter sets #5 ( $COV=20\%$   
 122  $\theta_x = 10\text{m}$  and  $\theta_z = 2\text{m}$ ) and #9 ( $COV = 40\%$   $\theta_x = 10\text{m}$  and  $\theta_z = 2\text{m}$ ), respectively (see Table 1). The red curves  
 123 correspond to the time ( $t_{75\%}$ ) when 75 % of the total energy (eq. 2) is reached for each seismogram, while the  
 124 blue lines show the maximum  $t_{75\%}$  value calculated for all seismograms of the given realization. The black and  
 125 red arrows show the receivers located at 30 m and 100 m, and at 55 m and 65 m, respectively.

126 The synthetic ground motions calculated for one  $V_s$  realization for the two parameter sets #5 ( $COV = 20\%$   $\theta_x$   
 127  $= 10\text{m}$  and  $\theta_z = 2\text{m}$ ) and #9 ( $COV = 40\%$   $\theta_x = 10\text{m}$  and  $\theta_z = 2\text{m}$ ) (Table 1) are shown in Figure 1. For each  
 128 parameter set, the  $V_s$  distribution in the sediment layer and the corresponding horizontal and vertical  
 129 seismograms computed on the surface are presented. The two  $V_s$  distribution shown in Figure 1(e) and (f)  
 130 highlight the effect of the  $COV$  on the generated  $V_s$  profiles, with larger velocity contrasts for parameter set #9  
 131 ( $COV=40\%$ ) than for parameter set #5 ( $COV=20\%$ ). The seismograms (Figure 1a to d) are highly affected by the  
 132 spatial variation of  $V_s$ , showing larger amplitude and longer duration at receivers located above low velocity  
 133 zones, especially for  $COV = 40\%$ .

134 **3 Lagged coherency estimation**

135 **3.1 Definition**

136 The coherency is a normalized complex function given by the ratio between the smoothed-cross spectral density  
137 ( $S_{jk}$ ) and the smoothed-power spectral density ( $S_{jj}$  and  $S_{kk}$ ) of two signals recorded at location  $j$  and  $k$  [44]:

$$\bar{\gamma}_{jk}(d_{jk}, f) = \frac{\bar{S}_{jk}(f)}{\sqrt{\bar{S}_{jj}(f) \cdot \bar{S}_{kk}(f)}} \quad (1)$$

138

139 where  $d_{jk}$  is the distance between location  $j$  and  $k$ , and  $f$  is the frequency.

140 From the complex coherency function presented in Eq. 1, several real coherency functions can be used: lagged  
141 coherency, unlagged coherency, and plane-wave coherency [61]. The lagged coherency ( $\gamma_{jk}^L$ ) is the amplitude of  
142 the complex coherency function ( $\bar{\gamma}_{jk}$ ) (Eq. 1). At low frequency and small inter-receiver distance, ground  
143 motions are close and the corresponding lagged coherency tends to 1. In contrast, for large distance and high  
144 frequency,  $\gamma_{jk}^L$  decreases down to 0.

145 After pre-aligning the signals in time by using the time lag leading to the maximum cross-correlation, the lagged  
146 coherency is the most commonly used indicator to measure the seismic motion similarity [44, 45, 46, 47].  
147 Indeed, its use is assumed to eliminate the deterministic variation in surface ground motion caused by the wave  
148 propagation, keeping only the stochastic variation [61] We use the lagged coherency in this paper to evaluate the  
149 impact of the spatial variability of soil properties on the phase variability of synthetic ground motions.

150 **3.2 Coherency function smoothing**

151 In the lagged coherency estimation, a frequency smoothing process of the seismic time histories spectra is  
152 mandatory [44]. Indeed, when no smoothing window is used, the phase difference terms disappear from the  
153 covariance spectra and the lagged coherency estimate is unity for any frequency and any receiver distance [62].  
154 This process also controls the statistical properties (i.e variance and bias) and the resolution of the coherency  
155 estimates. The increase of the smoothing bandwidth leads to a decrease in coherency estimation uncertainty  
156 while decreasing however the resolution. Thus, the smoothing type and parameter have to be well chosen to find



157 a balance between uncertainty and resolution of the estimated coherency. According to Harichandran [41], the  
 158 most common smoothing type windows (e.g. hamming, cosine, triangular, ...) yield similar results as long as the  
 159 equivalent spectral bandwidths are the same. Abrahamson et al. [61] proposes an 11-point Hamming window  
 160 ( $2M + 1$ ; where  $M$  is the Hamming parameter) for the spectral smoothing of the time windows containing less  
 161 than 2000 samples when the coherency estimates are to be used in structural analysis with damping coefficient of  
 162 5%. In this study, the cross- and auto-spectral densities in equation Eq. 1 are smoothed using an 11-point ( $M =$   
 163 5) Hamming window, corresponding to a spectral bandwidth of 1.95 Hz. The synthetic ground motions  
 164 calculated at the surface are resampled to a rate of 50 Hz with 256 samples.

### 165 **3.3 Time window selection**

166 The shear (S-) wave part of seismograms is commonly used to estimate the coherency function as it is  
 167 considered as the the most damaging seismic phase for civil engineering structures [43, 45, 63]. However, the  
 168 identification of S-waves in seismograms is sometimes difficult because they are often accompanied by other  
 169 waves due the scattering and diffraction in the heterogeneous medium [63]. Although the selection of the S-wave  
 170 window can be made visually, Abrahamson [43] proposed an automatic process based on the duration of the  
 171 normalized Arias Intensity  $AI$  (Eq. 2):

$$AI_{i,n}(\tau) = \frac{\int_0^\tau v_{i,n}^2 dt}{\int_0^{end} v_{i,n}^2 dt} \quad (2)$$

172

173 where  $v_{i,n}$  corresponds to the horizontal velocity component calculated at the receiver  $i$  of the realization  $n$  and  $\tau$   
 174 indicates time.

175 The criterion used by Abrahamson [43] is that the S-wave time window corresponds to  $AI$  between 0.1 and 0.75,  
 176 the lower limit being chosen to remove the P-wave firstly arriving at the surface. As we generate a plane wave of  
 177 SV type, the S-wave time window in this study is defined for  $AI$  between 0 and 0.75 (Eq. 2).

178 As shown in Figure 1c and d, the horizontal velocities are highly affected by the shallow  $V_s$  structure, leading to  
 179 significant variations in duration and amplitude. For the horizontal component, the time corresponding to  $AI$   
 180 equal to 0.75 ( $t_{75\%}$ ) are shown in red for the two  $V_s$  realizations in Figure 1c and d. The rise in  $COV$  (40%)  
 181 clearly increases  $t_{75\%}$  (Figure 1d), particularly for receivers located above superficial low velocity zones. In

182 literature, the lagged coherency estimation is based on the selection of the same time window for all the  
 183 recordings of the same event [43, 64]. Accordingly, the time window used to estimate the lagged coherency  
 184 between seismograms of the same realization was chosen as the maximum  $t_{75\%}$  (blue line) of this realization  
 185 (Figure 1c and 1d). Even though  $t_{75\%}$  is calculated using seismograms on the horizontal component, the same  
 186 time window is used for the lagged coherency estimation of both the horizontal and vertical component.

### 187 **3.4 Average and standard deviation**

188 For statistical analysis on the coherency estimates, normally distributed data are preferable [41]. Hence, the  
 189  $atanh$  of lagged coherency is used to produce an approximately normal distribution with a bias that can be  
 190 estimated and removed [65].

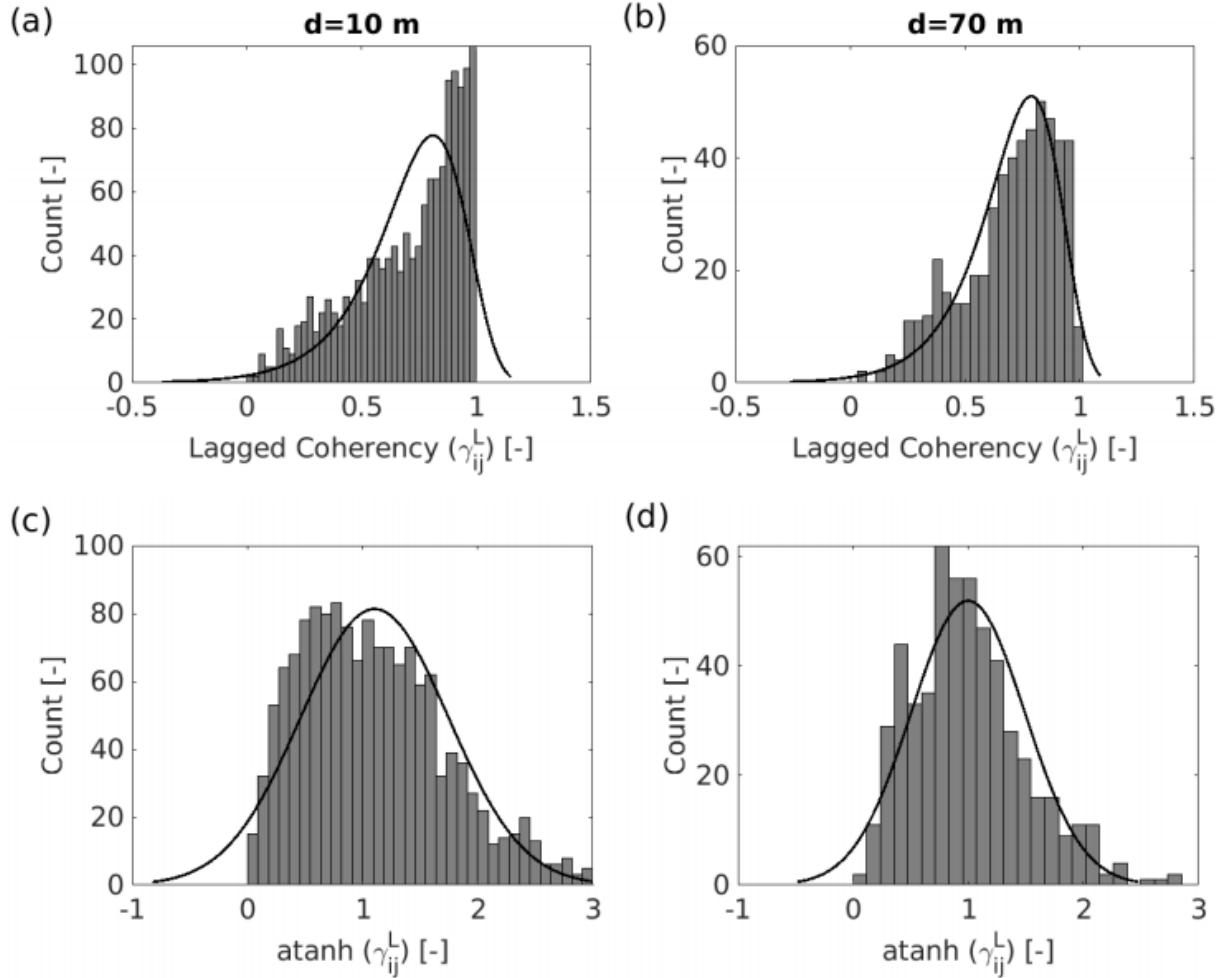
191 Following Harichandran [41] and Abrahamson et al. [61], the statistical analyses of the lagged coherency are  
 192 performed on  $atanh(\gamma)$  instead of  $\gamma$  and the average and standard deviation of the lagged coherency are  
 193 calculated as follow:

$$\gamma = \gamma(d, f)|_{mean} = \tanh\left(E(Atanh(\gamma_{jk}^L))\right) \quad (3)$$

$$\sigma(\gamma) = \gamma(d, f)|_{std} = \tanh\left(\sqrt{S(Atanh(\gamma_{jk}^L))}\right) \quad (4)$$

194 where  $E$  is the mathematical expectation,  $S$  is the variance defined by  $S(X) = E(X^2 - E(X)^2)$  for a variable  $X$   
 195 and  $std$  denotes the standard deviation.

196 At a frequency of 3.54 Hz (the fundamental frequency of the deterministic model) and for the parameter set #9  
 197 ( $COV = 40\%$   $\theta_x = 10m$  and  $\theta_z = 2m$ ; Table 1) that exhibits the largest  $V_s$  range, Figure 2a and b show the  
 198 distribution of the lagged coherency  $\gamma_{jk}^L$  from 100  $V_s$  realizations and two inter-receiver distance,  $d = 10$  m and  $d$   
 199  $= 70$  m, respectively. It is clear that  $\gamma_{jk}^L$  is not normally distributed regardless the inter-receiver distance, while  
 200 the distribution of  $atanh(\gamma_{jk}^L)$  shown in Figure 2c and d can be fitted by a normal distribution.



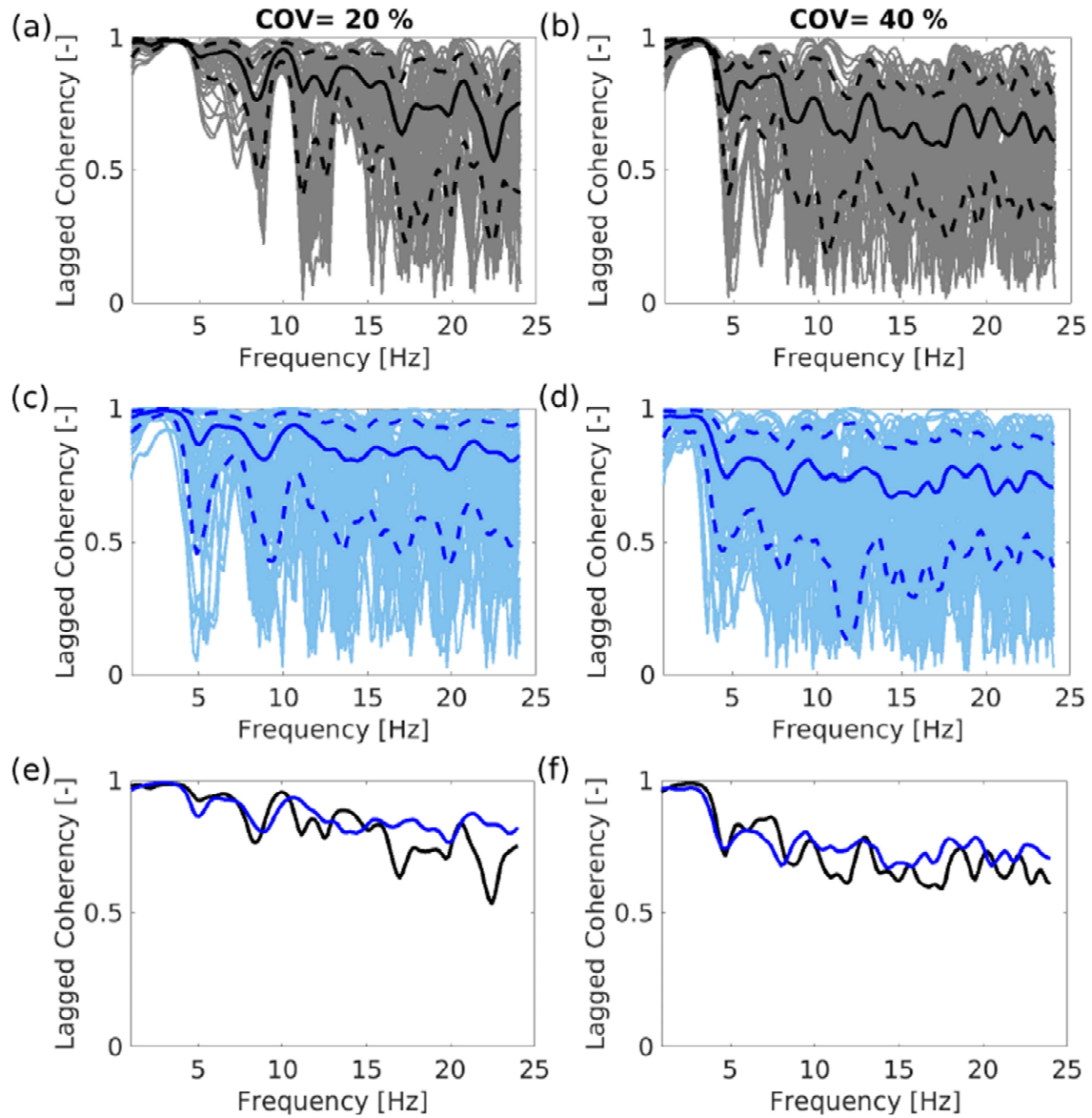
201

202 *Figure 2: Statistical distribution of the lagged coherencies: a) and b) Histograms of the*  
 203 *lagged coherencies ( $\gamma_{jk}^L$ ) calculated between receivers separated respectively by  $d = 10$  m*  
 204 *(left column) and  $d = 70$  m (right column) for all realizations of the parameter set #9 (COV*  
 205 *=40%  $\theta_x = 10$ m and  $\theta_z = 2$ m) and for a frequency of 3.54 Hz. c) and d) Histograms of  $\text{atanh}$*   
 206 *transformation of  $\gamma$  in a) and b) respectively. a) and b) are fitted by an Extreme Value*  
 207 *distribution and c) and d) are fitted by a Normal distribution (continuous lines).*

### 208 3.5 Stationarity of lagged coherency

209 An important aspect in the use of random field modeling approach is to ensure the stationarity of the output  
 210 average estimates. We thus compare the lagged coherency variability obtained when using 100 stochastic Vs  
 211 profiles for one receiver pair and when using all receivers pair for a given realization. Figure 3a and b show the  
 212 coherencies calculated for all pairs of receivers separated by a distance  $d = 70$  m for the two realizations

213 presented in Figure 1e and f, respectively, while Figure 3c and d shows the coherencies calculated between  
214 receivers located at 31 m and 101 m (black receivers in Figure 1e and f) for the 100 realizations of the two  
215 parameter sets #5 and #9 (Table 1), respectively. The bold solid and dashed curves represent the average  
216 coherency +/- one standard deviation using Eq. 2 and Eq. 3. The four figures show that the coherency curves  
217 vary with the location of the receivers on the surface and the  $V_s$  realization. However, the comparison of the  
218 average coherency in Figure 3e and f indicates almost similar coherency values. This comparison clearly outlines  
219 the lagged coherency stationarity by using the *EOLE* random field modeling approach and indicates that the  
220 estimated average coherency curves are robust. In the following sections, the average and standard deviation of  
221 the coherency will be computed for all  $V_s$  realizations, using all receivers except those located within 30 m of the  
222 model boundaries.



223

224 *Figure 3: Lagged coherency curves (horizontal velocity) calculated for parameter sets #5 (left*  
 225 *column) and #9 (right column), (see Table 1). a) and b) curves for all the receiver pairs*  
 226 *separated by a distance  $d = 70$  m (in grey) for the two  $V_s$  realizations shown in Figure 1e*  
 227 *and Figure 1f, respectively. c) and d) curves computed between the receivers at 30 and 100m*  
 228 *(in light blue) for the 100  $V_s$  realizations. The bold solid and dashed curves represent the*  
 229 *average and the average  $\pm$  one standard deviation, respectively. e) and f) superposition of*  
 230 *the average curves shown in a) and c), and in b) and d), respectively.*

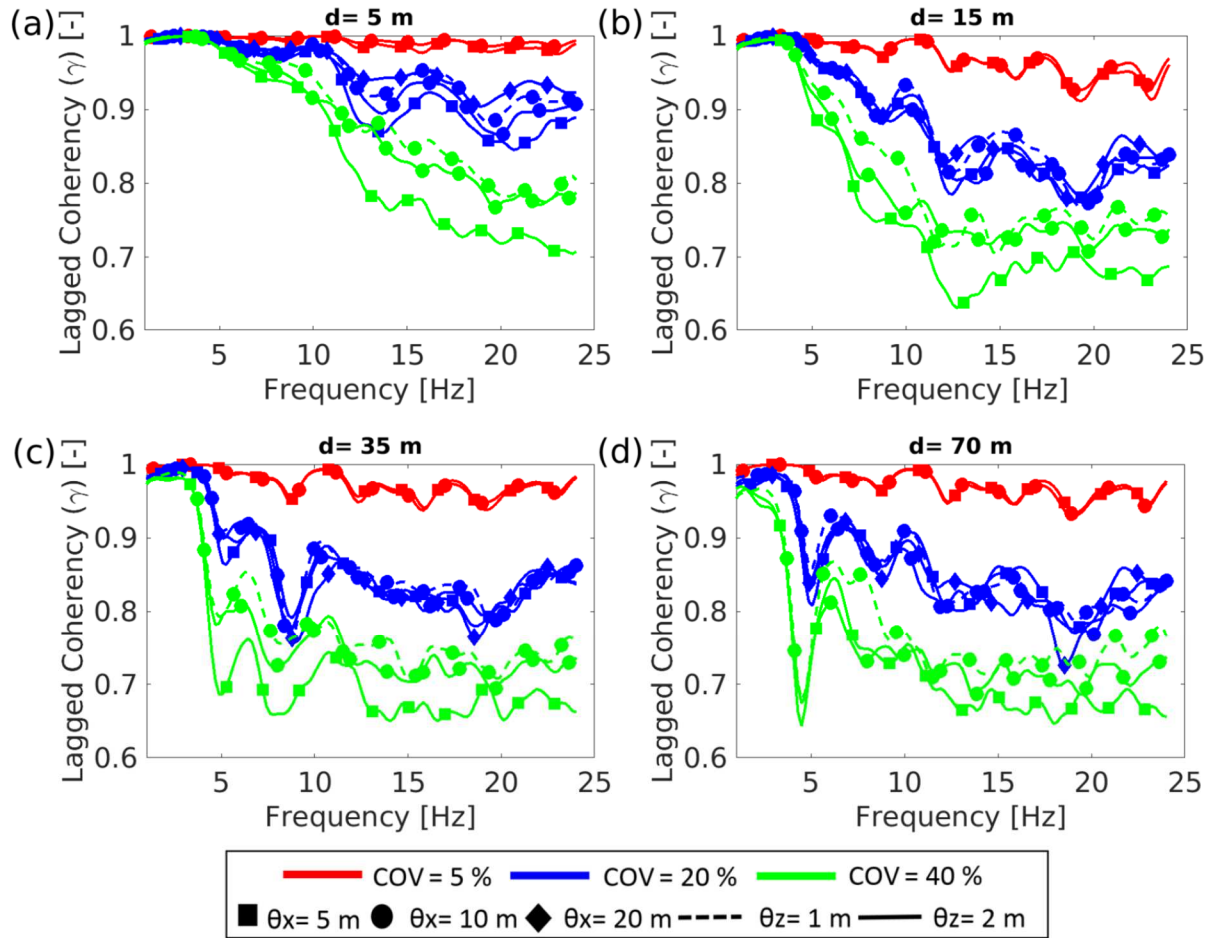
231

232 **4 Effect of the ground heterogeneities on the average lagged coherency**

233 Our study is focused on the lagged coherency computed for the horizontal component of the ground motion,  
234 which is the most studied in the literature (Imtiaz et al.,2018a) because of its impact on the behavior of structures  
235 and the importance given to it in seismic design codes. Results along the vertical component will be briefly  
236 commented and given in appendix A2.

237 The average lagged coherency ( $\gamma$ ) of the horizontal component is calculated for all the parameter sets given in  
238 Table 1 and shown in Figure 4 for 4 receivers distances (from  $d = 5$  to 70 m). The main trend is that  $\gamma$  decreases  
239 with the frequency and  $d$ , with a much more pronounced decrease effect when  $COV$  increases. The effect of the  
240 horizontal and vertical autocorrelation distances ( $\theta_x$  and  $\theta_z$ ) is weak and can only be observed for large  $COV$   
241 (40%) (Figure 4) and for  $d=5$  m and  $COV = 20\%$  (Figure 4a), with the trend that  $\gamma$  decreases with the decrease in  
242  $\theta_x$ . All average coherency curves computed for large  $d$  values (Figure 4c and d) exhibit two minima at 4.8 Hz  
243 and 8.5 Hz, respectively, which are more pronounced for large  $COV$  values.

244



245

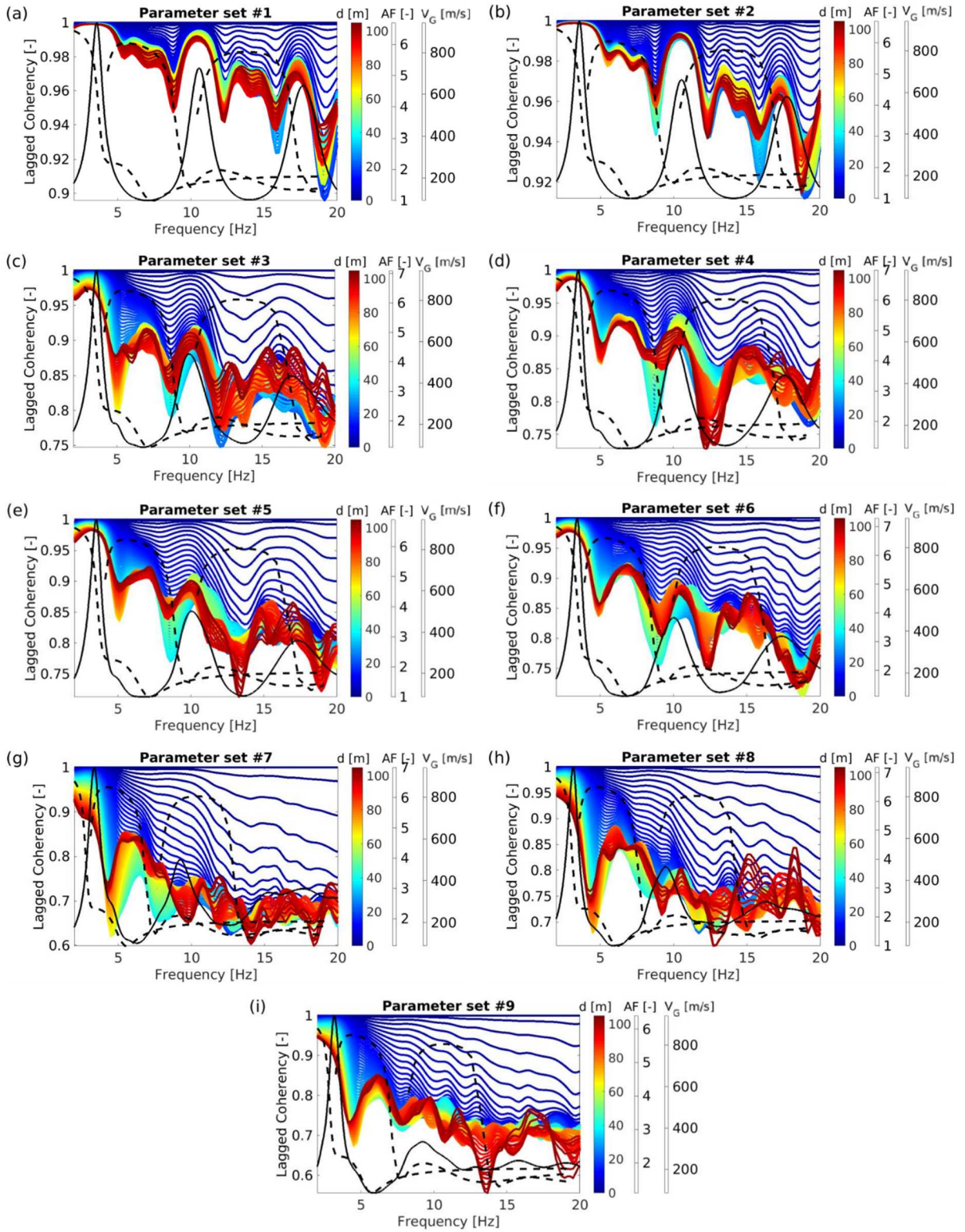
246 Figure 4: Average lagged coherency ( $\gamma$ ) calculated using the horizontal velocity, for different values of  $COV$ ,  $\theta_x$   
 247 and  $\theta_z$  for the 100 realizations computed by using the *atanh* transformation (see text for details) for four  
 248 receivers' distances: a)  $d = 5$  m, b)  $d = 15$  m, c)  $d = 35$  m and d)  $d = 70$  m.

249 Figure 5 presents the average coherency curves calculated for all inter-receiver distance and for each of the 9  
 250 parameter sets (Table 1). Minima and maxima of the average coherency within the same narrow frequency bands  
 251 are systematically observed whatever the parameter set considered, particularly for large distance ( $d > 25$  m).  
 252 This consistency suggests that curve extrema (minima and maxima) are controlled by the average properties of  
 253 the soil layer. To understand these results, an equivalent 1D model per parameter set is defined to reproduce the  
 254 measured fundamental resonance frequency picked on the average amplification function  $AF(f)$  from the 100  
 255 realizations. The calculation process is described in Appendix 1. From these equivalent models, the group  
 256 velocity dispersion curves ( $V_G$ ) of the Rayleigh waves are computed using the method proposed by Dziewonski  
 257 et al. [66] and improved by Levshin et al. [67]. The average amplification function  $AF(f)$  and the group velocity  
 258 dispersion curves of the first higher modes are plotted in Figure 5a to 5j for the 9 parameter sets. The computed

259 amplification curves exhibit several peaks (around 3.5, 10.5 and 17.5 Hz, i.e. at the resonance frequencies of the  
260 soil column) with a decrease in the peak amplitude with frequency and *COV*. As there is no attenuation in the  
261 numerical ground motion simulation, this attenuation effect results from the scattering generated by the  
262 heterogeneities. The larger the *COV* is the higher the apparent attenuation. Comparing the average coherency and  
263 amplification curves, we can see that the amplification peaks approximately coincide with maxima of the  
264 coherency curve. At resonance frequencies indeed, the vertical reverberation of S-waves are in-phase, leading to  
265 high  $\gamma$  values. Between the resonance frequencies,  $\gamma$  tends to decrease due to wave scattering, particularly for  
266 large *COV* values (Figure 5h to 5j). However, there is a systematic coherency increase (slight to strong,  
267 depending on *COV*) between the fundamental resonance frequency and first harmonic. This increase coincides  
268 with the Airy phase frequency of the fundamental mode of Rayleigh wave group velocity (minimum in the  
269 dispersion curve in Figure 5a to i), which corresponds to a very energetic in-phase motion [68] and lead to high  $\gamma$   
270 values. In our case, the Airy phase frequency of the first Rayleigh wave mode is in the same range as the second  
271 resonance frequency (10.5 Hz).

272 A similar study is conducted for the vertical component of the ground motion. The average lagged coherency  
273 curves are given in Appendix 2 for all the parameter sets and receivers distances ranging from  $d = 5$  to 70 m. The  
274 main results outline that coherency curves exhibit sinusoidal shapes for small *COV* (5%), however rapidly  
275 damped with frequency for large *COV* (20% and 40%). The autocorrelation distances have small influence on  
276 the coherency curves. The first maximum of the coherency curves again coincides with the fundamental  
277 frequency of the deterministic model ( $f_0 = 3.54$  Hz). As the input motion is a pure SV wave, this result also  
278 outlines that scattered S-waves by soil heterogeneities distribute energy on the vertical component.





279

280 *Figure 5: Average lagged coherency ( $\gamma$ ) for different receivers' distances in color curves, the*  
 281 *average amplification function (AF) in solid black and the Rayleigh group velocity curves*  
 282 *( $V_G$ ) for the three first modes in dashed black calculated for the nine parameter sets in table*

283 1. The coherency scale is different for each parameter set. The vertical axis of  $AF$  and  $V_G$  are  
284 displayed to the right of the distance ( $d$ ) colorbar.

285

## 286 5 Comparison with experimental data and existing models

### 287 5.1 Comparison between numerical and experimental data

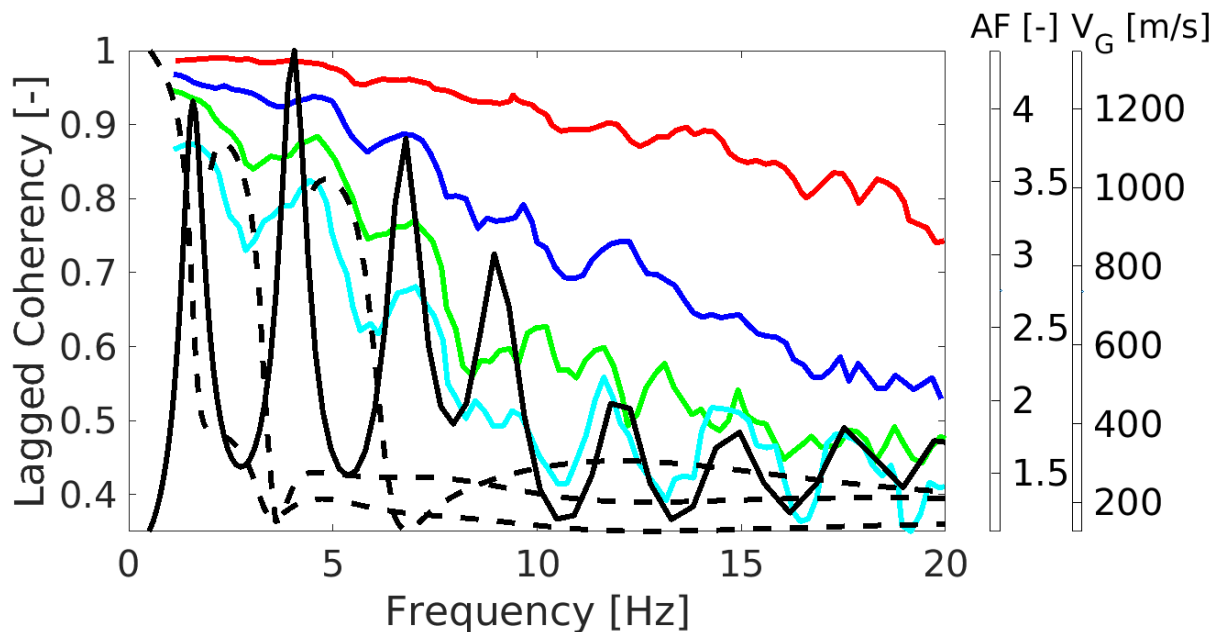
288 A very dense seismological array was deployed on the site of Koutavos-Argostoli (Greece), which is a small  
289 valley characterized by a soil thickness of maximum 80 m and a resonance frequency of about 1.5 Hz [69, 70].  
290 Below the array, shear-wave velocity in the sediment gradually increases from 150 m/s to 400 m/s and the  
291 seismic bedrock is at 60 m depth [63]. Using a set of 46 earthquakes, with magnitude ranging from 2 to 5 and  
292 epicentral distances up to 200 km, Imtiaz et al. [55] calculated average lagged coherencies for distances between  
293 receivers ranging from 10 to 90 m. The coherency curves for the horizontal component (Figure 6) exhibit shapes  
294 with a strong decrease with frequency and receivers' distances, leading to coherency values lower than 0.5.,  
295 suggesting that the soil layer is heterogeneous. However, the intrinsic attenuation, which is not considered in the  
296 numerical simulations, and our 2D simulations that do not account for the complete wavefield, could also play a  
297 role on the coherency curve shape at high frequency and explain the low coherency values obtained in the  
298 experimental data.

299 The other point to emphasize is the progressive apparition of minima and maxima at specific frequencies with  
300 the receiver distance increase. Between 2 and 8 Hz, the experimental lagged coherency curves (Figure 6) exhibit  
301 three maxima at about 2, 4.8 and 6.5 Hz, for the larger receivers' distances. Other extrema are also observed at  
302 higher frequencies, especially for receiver distances of 30-40 m and 80-90 m. In the same figure, the theoretical  
303 amplification function and the dispersion curves of the three first modes (group velocity, Rayleigh waves)  
304 computed from shear-wave velocity profiles [63, 71] are plotted. In this case, the Airy phase frequencies of the  
305 first two Rayleigh wave modes are close to the first resonance harmonic, while the third Rayleigh wave mode  
306 frequency is in the same range as the second harmonic. Similarly to the numerical results, the maxima of the  
307 lagged coherency in the low frequency range (here 0-8 Hz) approximately coincide with the first three peaks of  
308 amplification and the Airy phase of the Rayleigh waves. These theoretical and experimental results obtained for  
309 a sediment layer over bedrock suggest that the maxima within narrow frequency range of coherency in the low

310 frequency range (up to two to three times the fundamental resonance frequency of the site) are mainly influenced  
 311 by the average 1D soil properties (mean  $V_s$  and thickness) that control the Airy phase and resonance occurrence  
 312 frequencies. At these frequencies, the S and Rayleigh waves are propagating in phase and carry most of the  
 313 seismic energy, explaining the maxima of the lagged coherency. At higher frequencies, seismic waves are highly  
 314 diffracted leading to low average coherency values.

315 These results contradict the few previous analyses suggesting that the minimum in the coherency functions  
 316 coincide with the fundamental frequency of the site [72, 73]. Based on this interpretation and on the Luco and  
 317 Wong [48] model, Zerva and Harada [74] proposed a new coherency model depending on the average and  
 318 standard deviation of the estimated fundamental frequency of the medium and exhibiting a coherency drop at the  
 319 fundamental frequency. Further studies involving detailed experimental data from other sites will need to be  
 320 carried out to understand these differences in interpretation.

321 Imtiaz et al. [63, figure 9 in their paper] also provided experimental average coherency curves for the vertical  
 322 component, which show a first maximum at the fundamental frequency and a significant downward trend at  
 323 higher frequencies. This curve shape is consistent with the numerical results for high  $COV$  values ( $\approx 40\%$ ;  
 324 appendix 2), again suggesting that the soil cover at Koutavos-Argostoli is spatially highly heterogeneous.



325  
 326 *Figure 6: Experimental average lagged coherency ( $\gamma$ ) curves in color, the average*  
 327 *amplification function (AF) curve in solid black and the Rayleigh group velocity curves ( $V_G$ )*

328 *for the three first modes in dashed black calculated in the valley of Koutavos-Argostoli*  
 329 *(Imtiaz et al., 2017). Coherency curves are shown for four distances between receivers: 10-20*  
 330 *m (red curve), 20-30 m (blue curve), 30-40 m (green curve) and 80-90 m (cyan curve). The*  
 331 *vertical axis of AF and  $V_G$  are displayed to the right of the distance ( $d$ ) colorbar.*

332

## 333 5.2 Comparison with existing models

334 The computed coherency curves (horizontal motion) are compared to four parametric coherency models used in  
 335 the literature for engineering application: the theoretical coherency model of Uscinski [75], the semi-empirical  
 336 model of Luco and Wong [48], the semi-empirical model of Sato et al. [49] and the empirical model of  
 337 Abrahamson et al. [60]. These coherency models are usually used

338 The coherency model of Uscinski [75] is derived from the analysis of shear waves propagating a distance  $H$   
 339 through a random medium. The lagged coherency for a pair of ground motions as a function of inter-receiver  
 340 distance ( $d$ ) and angular frequency ( $\omega$ ) is given by:

$$\gamma(d, \omega) = \exp\left(-\omega^2 \theta H \frac{COV^2}{V_s^2} \left(1 - \exp\left(\frac{d^2}{\theta^2}\right)\right)\right) \quad (6)$$

341

342 where  $\theta$  is the correlation length of random inhomogeneities along the path,  $COV$  is a measure of the relative  
 343 variation of elastic properties,  $\omega$  is the angular frequency,  $V_s$  is an estimate of the elastic wave velocity and  $d$  is  
 344 the receiver distance. In this study,  $H$  is the depth of the heterogeneous soil layer.

345 By simplifying this coherency model, Luco and Wong [48] proposed to express the coherency for receiver  
 346 distance ( $d$ ) and an angular frequency ( $\omega$ ) as follows:

$$\gamma(d, \omega) = \exp\left(-\omega^2 \left(\frac{COV}{V_s} \sqrt{\frac{H}{\theta}}\right)^2 d^2\right) \quad (7)$$

347

348 From theoretical considerations on seismic wave propagation in random heterogeneous media [49] gave the  
 349 functional form of the coherency model as:

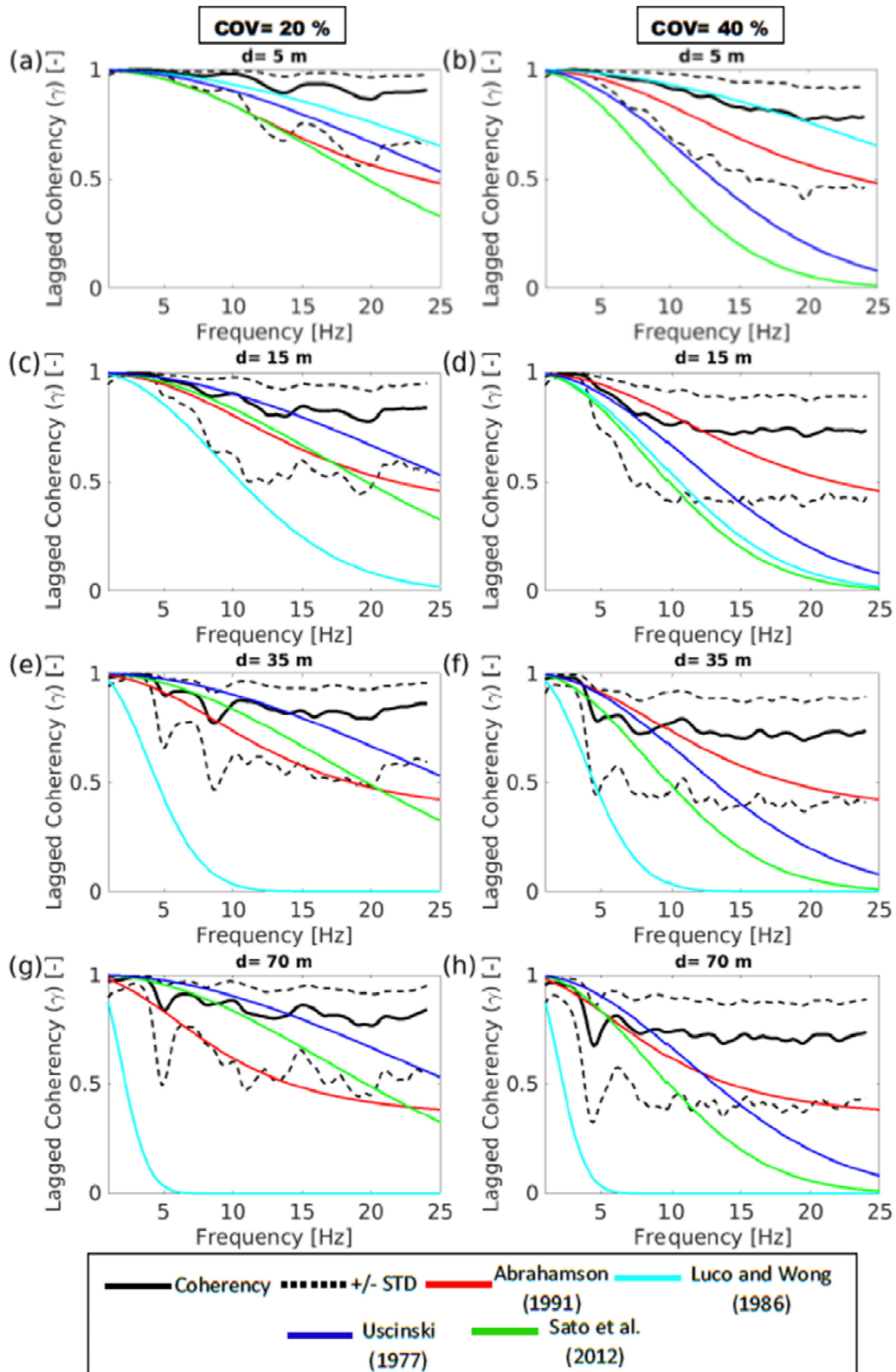
$$\gamma(d, \omega) = \exp\left(-\sqrt{\pi}\omega^2\theta H \frac{COV^2}{V_s^2}\left(1 - \exp\left(\frac{d^2}{\theta^2}\right)\right)\right) \quad (8)$$

350

351 Finally, the majority of the empirical coherency models in the literature are elaborated for receiver distance  
 352 larger than 100 m. However, Abrahamson et al. [61] coherency model was developed from the horizontal  
 353 seismograms recorded by a small array (LSST) in Taiwan and allows the estimation of the lagged coherency for  
 354 receiver distance smaller than 100 m. The lagged coherency is given by:

$$\gamma(d, f) = \tanh\left((2.54 - 0.012d)\left(\exp((-0.115 - 0.00084d)f) + \frac{f^{-0.878}}{3}\right) + 0.35\right) \quad (9)$$

355 where  $f$  is the frequency.



356

357 *Figure 7: The average lagged coherency +/- one standard deviation (in black) for parameter*  
 358 *set #5 (COV =20%  $\theta_x =10$ m and  $\theta_x =2$ m) (left column) and parameter set #9 (COV =40%  $\theta_x$*   
 359 *=10m and  $\theta_x =2$ m) (right column) in Table 1 for 4 receiver distances: (a) and (b) for  $d =5$  m,*

360 (c) and (d) for  $d = 15$  m, (e) and (f) for  $d = 35$  m and (g) and (h) for  $d = 70$  m compared with  
361 the coherency models of Abrahamson et al. [61] in red, Luco and Wong [48] in cyan,  
362 Uscinski [75] in blue and Sato et al. [49] in green.

363 Figure 7 compares the average lagged coherency ( $\gamma$ ) computed for the two parameter sets #5 and #9 having the  
364 same autocorrelation distances ( $\theta_x=10$ m and  $\theta_z=2$ m) and different *COV* (20% and 40%, respectively), with the  
365 aforementioned four coherency models for the four receivers' distances:  $d = 5$  m,  $d = 15$  m,  $d = 35$  m and  $d = 70$   
366 m. Although the coherency model of Luco and Wong [48] correctly predicts the coherency for small  $d$  and large  
367 *COV* (Figure 7b), it largely underestimates the coherency for large  $d$  (Figure 7g and h). Furthermore, even if the  
368 model of Sato et al. [49] provides lower coherencies than the one of Uscinski [75], these 2 coherency models are  
369 very similar to each other. They are not appropriate for very small distances ( $d = 5$  m) (see Figure 7a and b), but  
370 they better estimate the coherency for the parameter set having a *COV* of 20% (left column) than the one with a  
371 *COV* of 40% (right column) especially for low frequencies. Finally, the coherency model of Abrahamson et al.  
372 [61] gives relatively a good estimate of the coherency at low frequencies (below 7 Hz), but underestimate it at  
373 high frequencies. It seemingly fits better the parameter set #9 (*COV* =40%) than parameter set #5 (*COV* =20%),  
374 being within the computed average coherency +/- one standard deviation range. However, none of the four  
375 coherency models predict the coherency extrema in narrow frequency ranges observed for the synthetic  
376 simulations and at the Koutavos-Argostoli array. It outlines the need of further development of coherency  
377 models able to reproduce both the general decay of coherency with frequency of distance and extrema in narrow  
378 frequency bands.

379

## 380 **6 Effects of the ground heterogeneities on the standard deviation of the coherency**

381 As outlined in previous sections, decrease of  $\gamma$  with frequency and inter-receiver distance is clearly controlled by  
382 *COV*, while influence of the vertical and horizontal autocorrelation distances are not significant except for large  
383 *COV* and small inter-receiver distance (Figure 4). Interestingly, the influence of *COV* on  $\gamma$  is not observed on the  
384 standard deviation ( $\sigma(\gamma)$ , Eq. 4) whatever the parameter sets. Figure 8a to d presents  $\sigma(\gamma)$  for four inter-receiver  
385 distances ranging from 5 to 70 m : for a given *COV*,  $\sigma(\gamma)$  curves are rather similar for the different receiver  
386 distances, being only slightly influenced by the auto-correlation distances  $\theta_x$  and  $\theta_z$ . At low frequency (around 5  
387 Hz),  $\sigma(\gamma)$  exhibit similar values for all *COV* and autocorrelation distances, while  $\sigma(\gamma)$  tends to decrease with *COV*

388 at high frequencies. This observation is surprising in view of Figure 3, which shows that individual coherency  
389 curves are more variable for large *COV* values. The standard deviation of  $\sigma(\gamma)$  should intuitively increase with  
390 *COV*, which is not the case because of the *atanh* transformation of lagged coherency (Eq. 4)

391 In order to better account for larger variability of individual lagged coherency curves for larger *COV* in standard  
392 deviation of average lagged coherency, we propose to use another distribution function to fit the lagged  
393 coherency curves. Given that the lagged coherencies exhibit a maximum equal to unity for all the receiver-pairs  
394 and for different frequencies (Figure 3), the Extreme Value (*EV*) distribution (Type I) function is a good  
395 candidate for random variables having a minimum/maximum extreme [76]. The *EV* function is defined by:

396

$$f(x) = \frac{1}{\sigma'} e^{-e^{-\frac{x-\mu'}{\sigma'}} - \frac{x-\mu'}{\sigma'}} \quad (5)$$

397

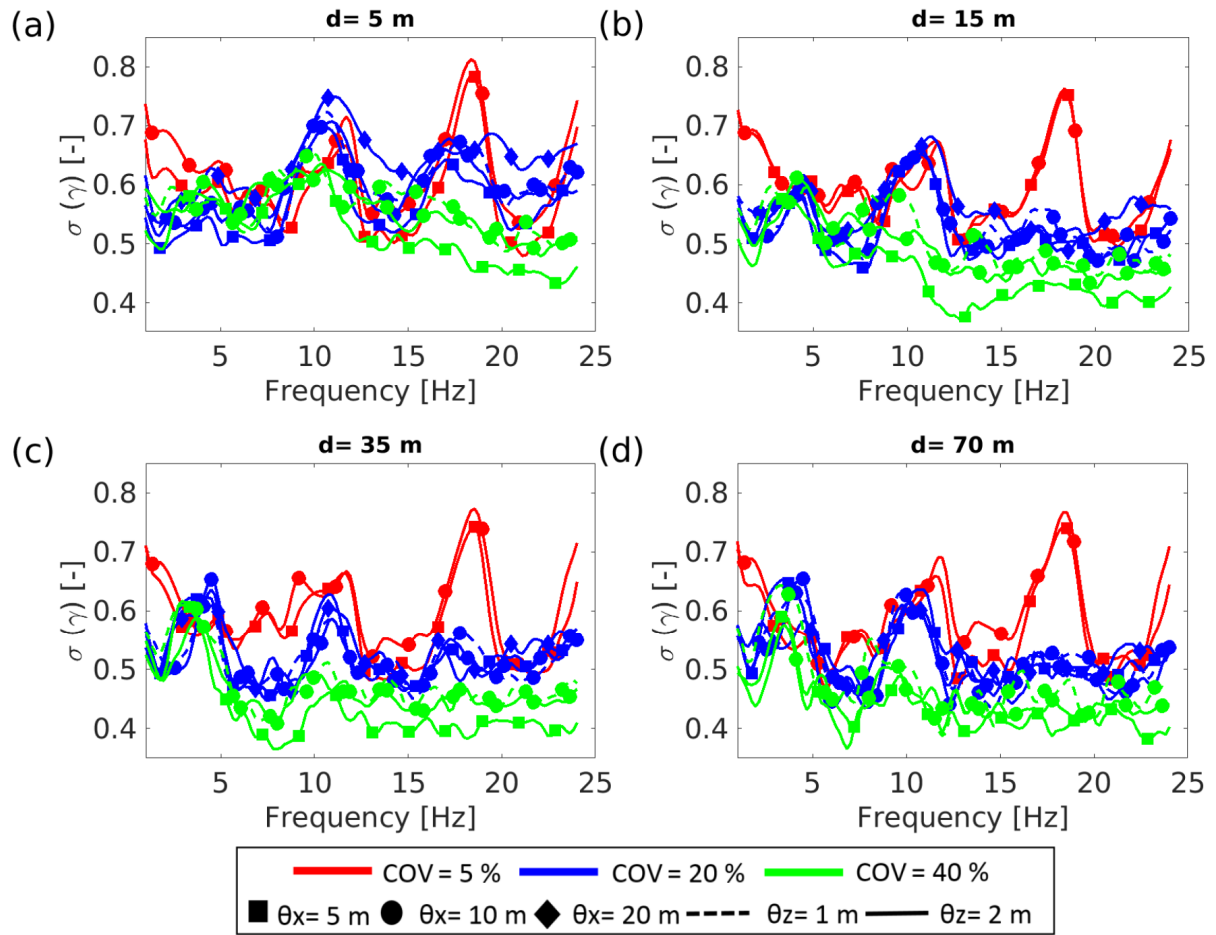
398 where  $\mu'$  and  $\sigma'$  are the two parameters that can be associated to  $\mu$  and  $\sigma$  of the normal distribution function,  
399 respectively.

400 Figure 2(a) and (b) illustrates the fit of lagged coherency with the *EV* distribution

401

402





403

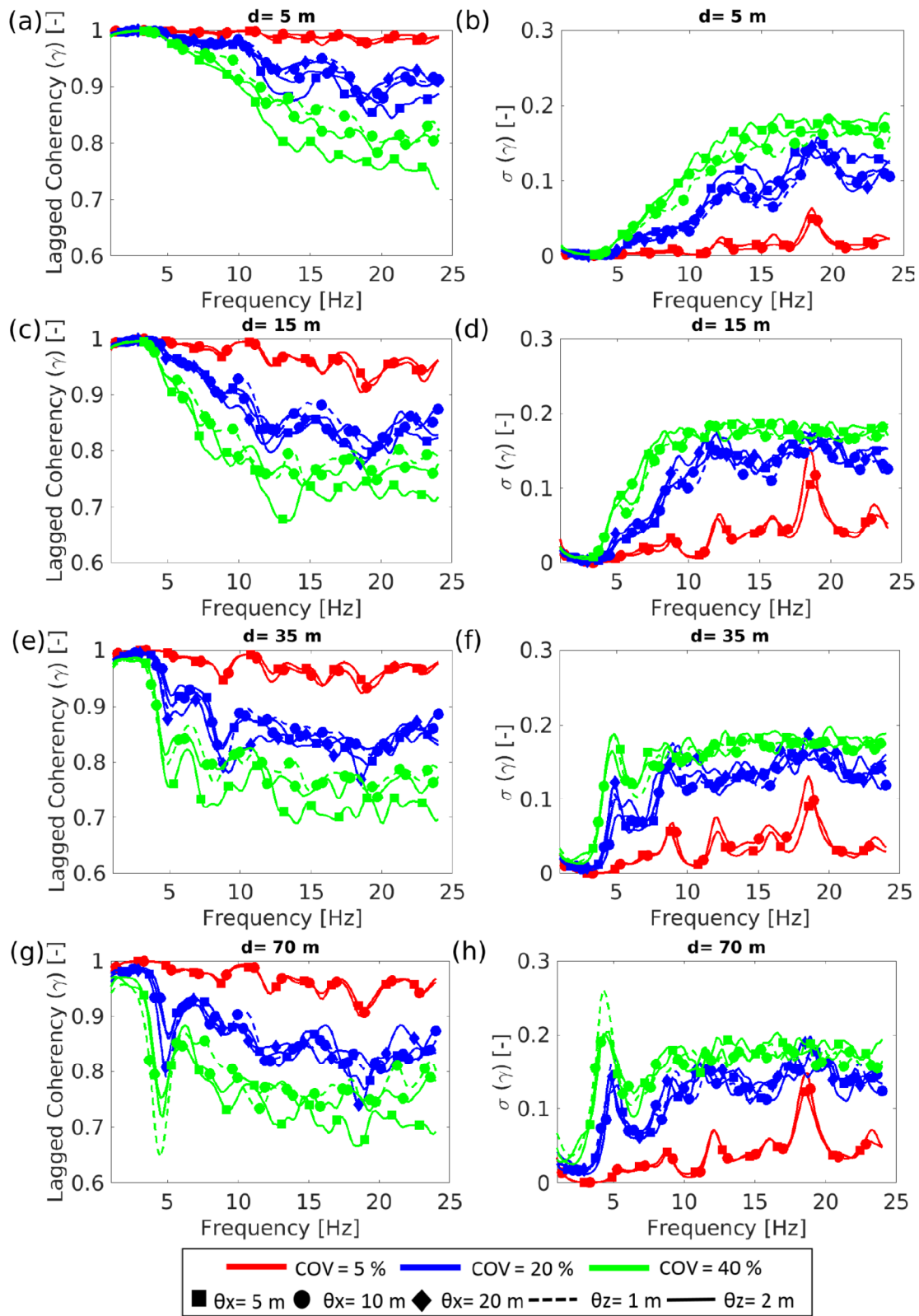
404 *Figure 8: The standard deviation of the lagged coherency (horizontal component) calculated*  
 405 *using the atanh transformation for different values of COV,  $\theta_x$  and  $\theta_z$ , and for four receivers'*  
 406 *distances: a)  $d = 5$  m, b)  $d = 15$  m, c)  $d = 35$  m and d)  $d = 70$  m.*

407 Figure 9 shows the average and the standard deviation of the lagged coherency calculated by using the *EV*  
 408 distribution function for 4 receiver distances ranging from 5 to 70 m. The average coherency curves (left  
 409 column) are similar to those shown in Figure 4 obtained using the *atanh* transformation with slightly lower  
 410 values and with minima at the same frequencies.

411 Standard deviations calculated using the *EV* distribution (Figure 9, right column) clearly outlines the increase of  
 412  $\sigma(\gamma)$  with *COV* for all frequencies and receiver distances and, for large inter-receiver distances, at the frequencies  
 413 corresponding to the coherency minima.

414 These results outline the interest of using the *EV* distribution to picture the effects of ground heterogeneities both  
415 on the average lagged coherency and the related standard deviation. It also opens perspective to infer *COV* at a  
416 site based on the joint properties of the average and variability of the lagged coherency.

417



418

419 *Figure 9: The average ( $\gamma$ ) (left column) and the standard deviation ( $\sigma(\gamma)$ ) (right column) of*  
 420 *the lagged coherence calculated for the horizontal component and using the Extreme Value*  
 421 *distribution. Computations are made for different values of COV,  $\theta_x$  and  $\theta_z$  for receiver*

422 *distances of  $d = 5$  m (a) and (b),  $d = 15$  m (c) and (d),  $d = 35$  m (e) and (f) and  $d = 70$  m (g)*  
423 *and h).*

## 424 7 Conclusions

425 A numerical probabilistic study for a simple velocity structure (an elastic heterogeneous sedimentary layer over a  
426 homogeneous half-space) is performed in order to investigate the effect of the  $V_s$  spatial variability on the lagged  
427 coherency of ground motion. Nine parameter sets were built by varying the following three  $V_s$  statistical  
428 parameters: the coefficient of variation ( $COV$ ) and the horizontal and the vertical autocorrelation distances ( $\theta_x$   
429 and  $\theta_z$ ). The analysis of the simulated synthetic seismogram at surface shows that the average lagged coherency  
430 values (horizontal component) exhibit the same shape for all parameter sets with a decrease as frequency or -  
431 receiver distance increase. The main parameter controlling the decrease of the average lagged coherency is the  
432 coefficient of variation ( $COV$ ), autocorrelation distances ( $\theta_x$  and  $\theta_z$ ) having only negligible effects except for  
433 small receivers' distances in the high frequency range. For engineering application, geotechnical and geophysical  
434 tests (e.g.:  $V_p$  and  $V_s$  tomography tests) should be carried out to measure the most important statistical  
435 parameters of the soil (the mean and the coefficient of variation) in order to reduce the uncertainty on the  
436 prediction of the surface ground motion variability.

437 Synthetic coherency curves exhibit extrema (maxima or minima) within narrow frequency bands in-between the  
438 resonance frequency of the site up to the second to third overtones, which are not predicted by coherency  
439 models. Such extrema (maxima) are controlled by the average properties of the site, namely resonance  
440 frequencies and Rayleigh wave Airy phase frequencies. Indeed, at those frequencies, seismic motions at surface  
441 are in-phase and therefore lead to large coherencies. This interpretation is very consistent with the experimental  
442 coherencies observed in the valley of Koutavos-Argostoli [55]. Finally, our interpretation is that coherency  
443 decreases with frequency or distance as result of seismic wave scattering at ground heterogeneities, modulated at  
444 specific frequencies by high coherency values due to the body and surface wave propagation within the average  
445 ground model.

446 Comparison between synthetic average coherency curves (horizontal component) and existing coherency models  
447 used for engineering applications outlines that existing models underestimate the coherency for all parameter  
448 sets. However, this probabilistic numerical study is limited to an elastic soil behavior and intrinsic attenuation  
449 could explain at least partly the observed discrepancy. Anyhow, none of the tested existing models are enable to

450 capture the variation of coherencies within narrow frequency bands. Thus, new coherency models shall be  
451 developed in the future to better reproduce these sudden variations that may have significant importance for  
452 soil-structure interaction applications.

453 Moreover, we showed that the *atanh* transformation of individual coherencies recommended by Abrahamson et  
454 al. [61] to compute the standard deviation is not relevant if one wants to relate statistical parameters of the  
455 underground velocity structure and lagged coherency in terms of average and standard deviation. Instead, we  
456 propose a new measure of average and standard deviation based on the Extreme Value (*EV*) distribution (Type  
457 I). We show that this distribution provides average coherency close to the one inferred from *atanh*  
458 transformation and standard deviation clearly sensitive to the *COV*.

459 Finally, the simple case of a sedimentary layer over a bedrock does not represent the real site conditions, even  
460 though it provided interesting findings regarding the relationships between underground heterogeneities and the  
461 ground motion variability observed on surface. To confirm our results, further numerical simulations, including  
462 3D ground models, should be carried out for a set of typical ground structure with various *V<sub>s</sub>* profiles, various  
463 seismic impedances and sediment thickness, considering also attenuation and non-linear soil behavior. Dedicated  
464 measurements at well-known instrumented sites will also help assessing the relationships between ground spatial  
465 heterogeneities and observed lagged coherencies in terms of average value and standard deviation.

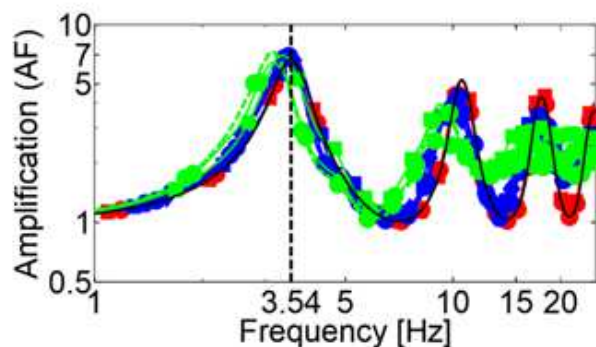
466

## 467 **8 Appendices**

### 468 **Appendix 1: 1D equivalent models**

469 We define 1D equivalent deterministic models for each parameter set, based on the computation of the  
470 amplification function  $AF(f)$ . For each parameter set (Table 1),  $AF(f)$  is the average over the 100 realizations of  
471 the ratio between the Fourier amplitude spectra of the signals recorded at the central receiver and the outcropping  
472 bedrock after smoothing the Fourier amplitude spectra by using Konno and Ohmachi algorithm with  $b=50$  [77].  
473 Figure A1 exhibits the amplification function  $AF(f)$  for the 9 parameter sets [37]. The fundamental resonance  
474 frequencies are close to the deterministic frequency, except for the higher *COV* (40%), for which a slightly lower  
475 frequency is obtained. The equivalent 1D deterministic model for the sediment layer over bedrock is then  
476 defined from the fundamental frequency  $f_0$ , determined from the first peak of the average amplification curve and

477 the average shear velocity ( $\mu_{Vs}^{1D}$ ) computed as the logarithmic average of all  $V_s$  values generated for the 100  
 478 realizations. The equivalent soil thickness  $H$  is then calculated using the relation  $H = \mu_{Vs}^{1D}/4f_0$ . The values of  $f_0$ ,  
 479  $\mu_{Vs}^{1D}$  and  $H$  are given in Table A1 for all parameter sets.



480  
 481 *Figure A1: Average outcrop amplification function at the central receiver (X=83 m) for*  
 482 *different values of COV,  $\theta_x$  and  $\theta_z$  computed using the 100 realizations. The black curve and*  
 483 *black dashed line refer to the amplification and the fundamental frequency of the*  
 484 *deterministic model, respectively [37].*

485  
 486 *Table A1: Properties of the 1D models equivalent to the 2D probabilistic models.*

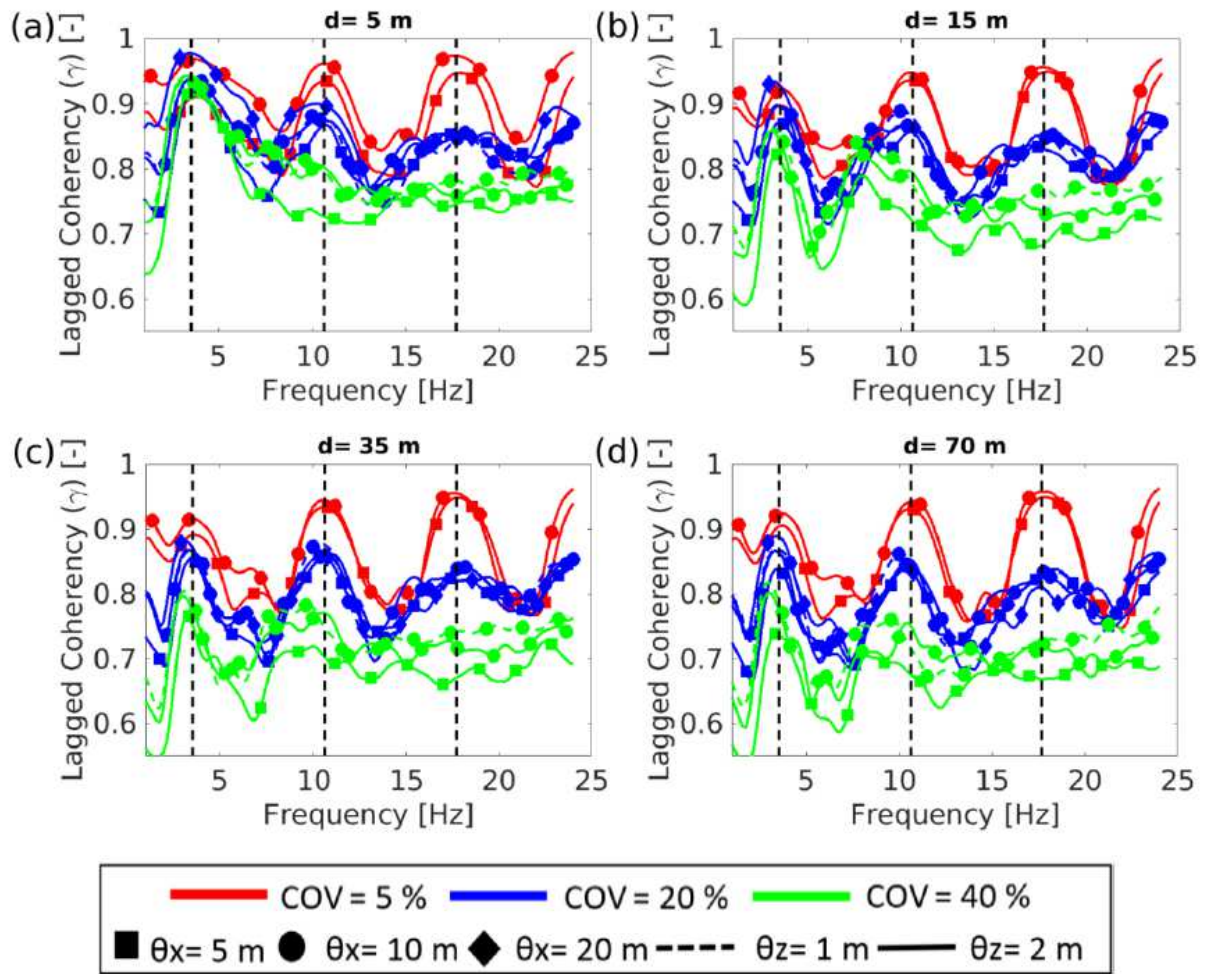
Parameter set #	Average shear-wave velocity $\mu_{Vs}^{1D}$ [m/s]	Fundamental frequency $f_0$ [Hz]	Equivalent soil thickness H $H$ [m]
1	220	3.56	15.45
2	220	3.54	15.54
3	218	3.50	15.57
4	218	3.51	15.53
5	217	3.49	15.54
6	217	3.36	16.15
7	216	2.69	20.07
8	216	2.99	18.06
9	214	2.77	19.31

487  
 488

489 **Appendix 2: Vertical lagged coherency**

490 The lagged coherency ( $\gamma$ ) of the vertical component is calculated for all the parameter sets given in Table 1 and  
491 is shown in Figure 4 for 4 inter-receiver distances (from  $d = 5$  to 70 m). For small COV (5%) the coherency  
492 curves have the shape of a sinusoidal function, while for large COV (20% and 40%) the sinusoidal shape is  
493 attenuated with frequency. For all curves, the average coherency value decreases with the increase in COV. In  
494 contrast, the other parameters (autocorrelation distances) only exhibits slight influence on the coherency curves.  
495 The first maximum of the coherency curves is almost equal to the fundamental frequency of the deterministic  
496 model ( $f_0 = 3.54 \text{ Hz}$ ) and the other maxima are at overtones ( $(n+3)f_0$ , for  $n=0, 1, 2, \dots$ ). As the seismic  
497 excitation is a pure SV wave at the base of the model, no vertical motion is expected at the surface. We interpret  
498 that S waves are scattered by the soil heterogeneities, distributing energy on the vertical component. The energy  
499 in the studied time window is then dominated by the vertical reverberation of the in-phase S-wave, leading to a  
500 maximum coherency at the resonance frequencies. Such sinusoidal shape is also observed on the lagged  
501 coherency calculated on the vertical surface velocity of the Koutavos-Argostoli recordings (Imtiaz et al., 2018a).  
502 However, for this case study, the first and second peaks are observed at 2 and 10 Hz, respectively, at almost  $f_0$   
503 and  $5f_0$  of the site. No peak is observed at  $3f_0$  (the resonant frequency of the third mode) and the observed  
504 oscillation periodicity ( $\sim 8 \text{ Hz}$ ) does not correspond to the site fundamental frequency. Further research is needed  
505 to understand the origin of this oscillation and its relationship to the site characteristics.

506



507

508 *Figure A2: Average lagged coherency ( $\gamma$ ) calculated using the vertical component, for*  
 509 *different values of COV,  $\theta_x$  and  $\theta_z$  for the 100 realizations computed by using the atanh*  
 510 *transformation (see text for details) for four receivers distances: a)  $d = 5$  m, b)  $d = 15$  m, c)  $d$*   
 511  *$= 35$  m and d)  $d = 70$  m. The resonant frequencies of the 3 fundamental modes of the*  
 512 *deterministic model are indicated by black dashed vertical lines.*

513

## 514 9 Acknowledgments

515 This work was supported by Institut de Recherche pour le Développement (IRD), especially the ARTS doctoral  
 516 thesis program, the IRD Young Research Unit JEA1 SAMMOVA and GDRI O-LIFE.



517 **10 References**

- 518 [1] Hall, J. F., Holmes, W. T., & Somers, P. (1994). Northridge earthquake, January 17, 1994. Preliminary  
519 reconnaissance report.
- 520 [2] Chouw, N., & Hao, H. (2012). Pounding damage to buildings and bridges in the 22 February 2011  
521 Christchurch earthquake. *International Journal of Protective Structures*, 3(2), 123-139.
- 522 [3] Shabestari, K. T., & Yamazaki, F. (2003). Near - fault spatial variation in strong ground motion due to  
523 rupture directivity and hanging wall effects from the Chi - Chi, Taiwan earthquake. *Earthquake  
524 engineering & structural dynamics*, 32(14), 2197-2219.
- 525 [4] Zerva, A., & Zervas, V. (2002). Spatial variation of seismic ground motions: an overview. *Applied  
526 Mechanics Reviews*, 55(3), 271-297.
- 527 [5] Der Kiureghian, A., & Neuenhofer, A. (1992). Response spectrum method for multi-support seismic  
528 excitations. *Earthquake Engineering & Structural Dynamics*, 21(8), 713-740.
- 529 [6] Zerva, A. (1992). Seismic loads predicted by spatial variability models. *Structural Safety*, 11(3-4), 227-  
530 243.
- 531 [7] Berrah, M. K., & Kausel, E. (1993). A modal combination rule for spatially varying seismic motions.  
532 *Earthquake Engineering & Structural Dynamics*, 22(9), 791-800.
- 533 [8] Der Kiureghian, A., Keshishian, P., & Hakopian, A. (1997). Multiple support response spectrum  
534 analysis of bridges including the site-response effect and the MSRS code (No. UCB/EERC-97/02).
- 535 [9] Saxena, V., Deodatis, G., & Shinozuka, M. (2000). Effect of spatial variation of earthquake ground  
536 motion on the nonlinear dynamic response of highway bridges. In *Proc of 12th World Conf on  
537 Earthquake Engineering*.
- 538 [10] Liao, S. (2006). Physical characterization of seismic ground motion spatial variation and conditional  
539 simulation for performance-based design.
- 540 [11] Hao, H., Oliveira, C. S., & Penzien, J. (1989). Multiple-station ground motion processing and  
541 simulation based on SMART-1 array data. *Nuclear Engineering and Design*, 111(3), 293-310.
- 542 [12] Somerville, P. G., McLaren, J. P., Sen, M. K., & Helmberger, D. V. (1991). The influence of site  
543 conditions on the spatial incoherence of ground motions. *Structural Safety*, 10(1-3), 1-13.
- 544 [13] Abrahamson, N. A. (1993, October). Spatial variation of multiple support inputs. In *Proc of 1st US  
545 Seminar on Seismic Evaluation and Retrofit of Steel Bridges*.

- 546 [14]Moczo, P., & Bard, P. Y. (1993). Wave diffraction, amplification and differential motion near strong  
547 lateral discontinuities. *Bulletin of the Seismological Society of America*, 83(1), 85-106.
- 548 [15]Chávez-García, F. J., & Bard, P. Y. (1994). Site effects in Mexico City eight years after the September  
549 1985 Michoacan earthquakes. *Soil Dynamics and Earthquake Engineering*, 13(4), 229-247.
- 550 [16]Field, E. H. (1996). Spectral amplification in a sediment-filled valley exhibiting clear basin-edge-  
551 induced waves. *Bulletin of the Seismological Society of America*, 86(4), 991-1005.
- 552 [17]Graves, R. W., Pitarka, A., & Somerville, P. G. (1998). Ground-motion amplification in the Santa  
553 Monica area: Effects of shallow basin-edge structure. *Bulletin of the Seismological Society of America*,  
554 88(5), 1224-1242.
- 555 [18]Harichandran, R. S. (1999). Spatial variation of earthquake ground motion, what is it, how do we model  
556 it, and what are its engineering implications. Dept. of Civil and Environmental Engineering, Michigan  
557 State Univ., East Lansing, Mich.
- 558 [19]Jenny, H. (1941). *Factors of soil formation: A system of quantitative pedology*, 281 pp.
- 559 [20]Burrough, P. A. (1993). Soil variability revisited. *Soils Fert*, 56(5), 529-562.
- 560 [21]Lacasse, S., & Nadim, F. (1997). Uncertainties in characterising soil properties. *Publikasjon-Norges*  
561 *Geotekniske Institutt*, 201, 49-75.
- 562 [22]Einsele, G. (2000). *Sedimentary basins: evolution, facies, and sediment budget*. Springer.
- 563 [23]Uzielli, M., Nadim, F., Lacasse, S., & Kaynia, A. M. (2008). A conceptual framework for quantitative  
564 estimation of physical vulnerability to landslides. *Engineering Geology*, 102(3-4), 251-256.
- 565 [24]Pagliaroli, A., Lanzo, G., Tommasi, P., & Di Fiore, V. (2014b). Dynamic characterization of soils and  
566 soft rocks of the Central Archeological Area of Rome. *Bulletin of earthquake engineering* , 12 (3),  
567 1365-1381.
- 568 [25]Popescu, R. (1995). *Stochastic variability of soil properties: data analysis, digital simulation, effects on*  
569 *system behavior*. Princeton University.
- 570 [26]Lopez-Caballero, Fernando, & Modaressi-Farahmand-Razavi, A. (2010). Assessment of variability and  
571 uncertainties effects on the seismic response of a liquefiable soil profile. *Soil Dynamics and Earthquake*  
572 *Engineering* , 30 (7), 600-613.
- 573 [27]Youssef Abdel Massih, D. S., Soubra, A. H., & Mao, N. (2010). Reliability-based analysis of strip  
574 footings subjected to an inclined or an eccentric loading. In *GeoFlorida 2010: Advances in Analysis,*  
575 *Modeling & Design* (pp. 2133-2142).

- 576 [28] Al - Bittar, T., & Soubra, A. H. (2014). Efficient sparse polynomial chaos expansion methodology for  
577 the probabilistic analysis of computationally - expensive deterministic models. *International Journal for*  
578 *Numerical and Analytical Methods in Geomechanics*, 38(12), 1211-1230.
- 579 [29] Wang, S., & Hao, H. (2002). Effects of random variations of soil properties on site amplification of  
580 seismic ground motions. *Soil Dynamics and Earthquake Engineering*, 22(7), 551-564.
- 581 [30] Sadouki, A., Harichane, Z., & Chehat, A. (2012). Response of a randomly inhomogeneous layered  
582 media to harmonic excitations. *Soil dynamics and earthquake engineering*, 36, 84-95.
- 583 [31] Rodriguez - Marek, A., Rathje, E. M., Bommer, J. J., Scherbaum, F., & Stafford, P. J. (2014).  
584 Application of single-station sigma and site-response characterization in a probabilistic seismic-hazard  
585 analysis for a new nuclear site. *Bulletin of the Seismological Society of America*, 104(4), 1601-1619.
- 586 [32] Berkane, H. D., Harichane, Z., Çelebi, E., & Elachachi, S. M. (2019). Site dependent and spatially  
587 varying response spectra. *Earthquake Engineering and Engineering Vibration*, 18(3), 497-509.
- 588 [33] Assimaki, D., Pecker, A., Popescu, R., & Prevost, J. (2003). Effects of spatial variability of soil  
589 properties on surface ground motion. *Journal of earthquake engineering*, 7(spec01), 1-44.
- 590 [34] Thompson, E., Baise, L., Kayen, R., & Guzina, B. (2009). Impediments to Predicting Site Response:  
591 Seismic Property Estimation and Modeling Simplifications. *Bulletin of the Seismological Society of*  
592 *America, BSSA* , 99, 2927-2949.
- 593 [35] Pagliaroli, A., Moscatelli, M., Raspa, G., & Naso, G. (2014a). Seismic microzonation of the central  
594 archaeological area of Rome: results and uncertainties. *Bull. Earthq. Eng.* , 12, 1405–1428.
- 595 [36] Stripajova, S., Moczo, P., Kristek, J., Bard, P. Y., Hollender, F., & Sicilia, D. (2018). Extensive  
596 numerical study on identification of key structural parameters responsible for site effects. In  
597 Conference: 16th European Conference on Earthquake Engineering (16ECEE). At: Thessaloniki,  
598 Greece.
- 599 [37] El Haber, E., Cornou, C., Jongmans, D., Abdelmassih, D. Y., Lopez-Caballero, F., & Al-Bittar, T.  
600 (2019). Influence of 2D heterogeneous elastic soil properties on surface ground motion spatial  
601 variability. *Soil Dynamics and Earthquake Engineering*, 123, 75-90.
- 602 [38] Abrahamson, N. A. (1990). Uncertainty in numerical strong motion predictions. In *Proc. Fourth US*  
603 *Nat. Conf. Earthq. Eng. (Vol. 1, pp. 407-416).*

- 604 [39]Matsushima, Y. (1977). Stochastic response of structure due to spatially variant earthquake excitations.  
605 In Proceedings of Sixth World Conference on Earthquake Engineering (Vol. 2, pp. 1077-1082).
- 606 [40]Abrahamson, N., Schneider, J., & Stepp, J., 1990. Spatial variation of strong ground motion for use in  
607 soil structure interaction analyses, in Proceedings of the Fourth U.S. National Conference on  
608 Earthquake Engineering, Palm Springs, CA.
- 609 [41]Harichandran, R. S. (1991). Estimating the spatial variation of earthquake ground motion from dense  
610 array recordings. *Structural Safety*, 10(1-3), 219-233.
- 611 [42]Zerva, A. (2009). Spatial variation of seismic ground motion. Modeling and engineering applications.
- 612 [43]Abrahamson, N. A. (2007). Program on technology innovation: effects of spatial incoherence on  
613 seismic ground motions. EPRI, Palo Alto, CA, 1015110.
- 614 [44]Abrahamson, N. (1992). Spatial variation of earthquake ground motion for application to soil-structure  
615 interaction. Final report (No. EPRI-TR--100463-Tier1). Electric Power Research Inst., Palo Alto, CA  
616 (United States); Bechtel Civil, Inc., San Francisco, CA (United States); Woodward-Clyde Consultants,  
617 Pasadena, CA (United States).
- 618 [45]Harichandran, R. S., & Vanmarcke, E. H. (1986). Stochastic variation of earthquake ground motion in  
619 space and time. *Journal of Engineering Mechanics*, 112(2), 154-174.
- 620 [46]Ancheta, T., Stewart, J., & Abrahamson, N., 2011. Engineering characterization of earthquake ground  
621 motion coherency and amplitude variability, in Proc. 4th IASPEI / IAEE Int. Sym. on Effects of Surface  
622 Geology on Seismic Motion, pp. 23–26.
- 623 [47]Abrahamson, N. A. (2006). Program on technology innovation: spatial coherency models for soil-  
624 structure interaction. EPRI, Palo Alto, CA and the US Department of Energy: Report, 1012969.
- 625 [48]Luco, J. E., & Wong, H. L. (1986). Response of a rigid foundation to a spatially random ground motion.  
626 *Earthquake Engineering & Structural Dynamics*, 14(6), 891-908.
- 627 [49]Sato, H., Fehler, M. C., & Maeda, T. (2012). Seismic wave propagation and scattering in the  
628 heterogeneous earth (Vol. 496). Berlin: Springer.
- 629 [50]Svay, A. (2017). Modélisation de la variabilité spatiale du champ sismique pour les études d'interaction  
630 sol-structure. Phd thesis, Centrale Supélec, Paris.
- 631 [51]DeGroot, D. J. (1996). Analyzing spatial variability of in situ soil properties. In *Uncertainty in the  
632 geologic environment: From theory to practice* (pp. 210-238). ASCE.

- 633 [52]Phoon, K. K., & Kulhawy, F. H. (1999). Characterisation of geotechnical uncertainty. Canadian  
634 Geotechnical Journal, 36(4).
- 635 [53]Liu, C. N., & Chen, C. H. (2010). Spatial correlation structures of CPT data in a liquefaction site.  
636 Engineering Geology, 111(1-4), 43-50.
- 637 [54]Salloum, N. (2015). Evaluation de la variabilité spatiale des paramètres géotechniques du sol à partir de  
638 mesures géophysiques: application à la plaine alluviale de Nahr-Beyrouth (Liban). Phd thesis,  
639 Universite Joseph Fourier, Grenoble.
- 640 [55]Imtiaz, A., Perron, V., Sway, A., Hollender, F., Bard, P. Y., & Theodoulidis, N. (2017, January).  
641 Wavefield characteristics and coherency of seismic ground motion from a rocksite at Argostoli Greece.  
642 In Sixteenth World Conference on Earthquake Engineering (No. 1743).
- 643 [56]Vanmarcke, E. (2010). Random fields: analysis and synthesis. World Scientific.
- 644 [57]Der Kiureghian, A., & Ke, J. B. (1987). The stochastic finite element method in structural reliability. In  
645 Stochastic structural mechanics (pp. 84-109). Springer, Berlin, Heidelberg.
- 646 [58]Li, C.-C., & Der Kiureghian, A. (1993). Optimal discretization of random fields. *Journal of*  
647 *Engineering Mechanics*, 119 (6), 1136-1154.
- 648 [59]Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and analysis of computer  
649 experiments. *Statistical science*, 409-423.
- 650 [60]ITASCA. (2011). FLAC – Fast Lagrangian Analysis of Continua (Vol. Version 7.0). Minneapolis,  
651 USA: Itasca Consulting Group.
- 652 [61]Abrahamson, N. A., Schneider, J. F., & Stepp, J. C. (1991). Empirical spatial coherency functions for  
653 application to soil-structure interaction analyses. *Earthquake spectra*, 7(1), 1-27.
- 654 [62]Jenkins, G. M., & Watts, D. G. (1968). Spectral analysis.
- 655 [63]Imtiaz, A., Perron, V., Hollender, F., Bard, P. Y., Cornou, C., Sway, A., & Theodoulidis, N. (2018a).  
656 Wavefield Characteristics and Spatial Incoherency: A Comparative Study from Argostoli Rock - and  
657 Soil - Site Dense Seismic Arrays. *Bulletin of the Seismological Society of America*.
- 658 [64]Imtiaz, A., Cornou, C., & Bard, P. Y. (2018b). Sensitivity of ground motion coherency to the choice of  
659 time windows from a dense seismic array in Argostoli, Greece. *Bulletin of Earthquake Engineering*,  
660 16(9), 3605-3625.

- 661 [65] Enochson, L. D., & Goodman, N. R. (1965). Gaussian approximations to the distribution of sample  
662 coherence (No. MAC-403-07). MEASUREMENT ANALYSIS CORP LOS ANGELES CA.
- 663 [66] Dziewonski, A., Bloch, S., & Landisman, M. (1969). A technique for the analysis of transient seismic  
664 signals. *Bulletin of the seismological Society of America*, 59(1), 427-444.
- 665 [67] Levshin, A., Ratnikova, L., & Berger, J. O. N. (1992). Peculiarities of surface-wave propagation across  
666 central Eurasia. *Bulletin of the Seismological Society of America*, 82(6), 2464-2493.
- 667 [68] Sheriff, R. E., & Geldart, L. P. (1995). *Exploration seismology*. Cambridge university press.
- 668 [69] Theodoulidis, N., Cultrera, G., Cornou, C., Bard, P. Y., Boxberger, T., DiGiulio, G., ... & Argostoli  
669 NERA Team. (2018a). Basin effects on ground motion: the case of a high-resolution experiment in  
670 Cephalonia (Greece). *Bulletin of Earthquake Engineering*, 16(2), 529-560.
- 671 [70] Theodoulidis, N., Hollender, F., Mariscal, A., Moiriat, D., Bard, P. Y., Konidaris, A., ... & Roumelioti,  
672 Z. (2018b). The ARGONET (Greece) Seismic Observatory: An Accelerometric Vertical Array and Its  
673 Data. *Seismological Research Letters*.
- 674 [71] Imtiaz, A. (2015). *Seismic wave field, spatial variability and coherency of ground motion over short  
675 distances: near source and alluvial valley effects* (Doctoral dissertation, Université Grenoble Alpes).
- 676 [72] Kanasewich, E. R. (1981). *Time sequence analysis in geophysics*. University of Alberta.
- 677 [73] Cranswick, E. (1988). The information content of high-frequency seismograms and the near-surface  
678 geologic structure of "hard rock" recording sites. In *Scattering and Attenuations of Seismic Waves, Part  
679 I* (pp. 333-363). Birkhäuser, Basel.
- 680 [74] Zerva, A., & Harada, T. (1997). Effect of surface layer stochasticity on seismic ground motion  
681 coherence and strain estimates. *Soil Dynamics and Earthquake Engineering*, 16(7-8), 445-457.
- 682 [75] Uscinski, B. J. (1977). *The elements of wave propagation in random media*. McGraw-Hill Companies.
- 683 [76] Gumbel, E. J. (1958). Statistical-theory of extreme values. *Bull Internat Stat Inst*, 36(3), 12-14.
- 684 [77] Konno, K., & Ohmachi, T. (1998). Ground-motion characteristics estimated from spectral ratio between  
685 horizontal and vertical components of microtremor. *Bulletin of the Seismological Society of America*,  
686 88(1), 228-241.

687