



HAL
open science

Wavefield characteristics and spatial incoherency - a comparative study from Argostoli rock- and soil-site dense seismic arrays

Afifa Imtiaz, Vincent Perron, Fabrice Hollender, Pierre-Yves Bard, Cécile Cornou, Angkeara Svay, Nikos Theodoulidis

► To cite this version:

Afifa Imtiaz, Vincent Perron, Fabrice Hollender, Pierre-Yves Bard, Cécile Cornou, et al.. Wavefield characteristics and spatial incoherency - a comparative study from Argostoli rock- and soil-site dense seismic arrays. *Bulletin of the Seismological Society of America*, 2018, 108 (5A), pp.2839-2853. 10.1785/0120180025 . cea-02339740

HAL Id: cea-02339740

<https://cea.hal.science/cea-02339740>

Submitted on 16 Mar 2020

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.

1 **Wavefield characteristics and spatial incoherency – a**
2 **comparative study from Argostoli rock- and soil-site dense**
3 **seismic arrays**

4 by

5 A. Imtiaz^{1,3}, V. Perron^{2,3,6}, F. Hollender², P-Y. Bard³, C. Cornou³, A. Svay⁴, N. Theodoulidis⁵

6
7 Corresponding Author:
8 Afifa Imtiaz
9 Bureau de Recherches Géologiques et Minières (BRGM)
10 Seismic and Volcanic Risks Unit
11 Risks and Prevention Division
12 3, Avenue Claude Guillemin - BP 36009 - 45060 Orléans cedex 2, France
13 Tel : +33(0)2 38 64 30 95
14 E-mail : a.imtiaz@brgm.fr
15 ORCID iD: <https://orcid.org/0000-0001-6687-8690>

16
17 ⁽¹⁾ BRGM, F-45060 Orléans, France
18 ⁽²⁾ CEA, DEN, F-13108 St Paul lez Durance Cedex, France
19 ⁽³⁾ Univ. Grenoble Alpes, Univ. Savoie Mont-Blanc, CNRS, IRD, IFSTTAR, ISTerre,
20 Bâtiment OSUG C, CS 40700, 38058 GRENOBLE Cedex 9, France
21 ⁽⁴⁾ IMSIA, University Paris-Saclay, France
22 ⁽⁵⁾ Institute of Engineering Seismology and Earthquake Engineering (ITSAK), GR-
23 55102 Thessaloniki, Greece
24 ⁽⁶⁾ IRSN, F-92260 Fontenay-aux-Roses, France

25 **ABSTRACT**

26 The current article focuses on the results obtained from the analysis of seismic events recorded
27 by a dense array located on a rock site at Argostoli in the Cephalonia Island of Greece. The
28 objective of the study is threefold: (1) to explore to what extent the non-direct, diffracted
29 surface waves influence the seismic wavefield at a rock site, (2) to investigate the loss of
30 coherency of ground motions, and (3) to compare the results with those from a previously
31 studied similar array located at an adjacent small, shallow sedimentary valley (soft-soil site).
32 Both arrays consist of 21 velocimeters encompassing a central station in four concentric circles
33 with diameters ranging from 10 to 180 m at the soft-soil and 20 to 360 m at the rock site. The
34 seismic datasets under consideration include 40 or more events occurring around the site at
35 epicentral distances up to 200 km having magnitudes ranging from 2 to 5. The seismic
36 wavefields at both sites are analyzed by using the MUSIQUE algorithm: the backazimuth and
37 slowness of the dominant incoming waves are extracted and the Love and Rayleigh waves are
38 identified. Lagged coherency is estimated from the dataset, and the results are averaged for
39 station pairs located at four separation distance intervals, 10–20, 20–30, 30–40 and 80–90 m.
40 The results indicate that coherency at the rock site is, generally, larger than that from the soft-
41 soil, especially at frequencies below 5 Hz. At soil site, lower coherency is observed for pairs
42 along the valley-width direction while no such directional dependence is observed at the rock.
43 The wavefield analysis shows that whilst about 40–60% of the seismogram energy at the soft-
44 soil could be associated to diffracted surface waves (Love and Rayleigh) appearing mainly
45 from the valley-width directions, only about 20% of energy at the rock site could be
46 characterized as diffracted surface waves. Comparison with the widely-quoted parametric
47 models reveals that the observed decay of coherency at both sites has little correlation with
48 those from the models. These significant differences observed between the results of the rock
49 and soil array indicate that the spatial incoherency is largely site dependent and is likely to be
50 closely related with the formation of locally generated wavefield.

51

52 *Keywords: Wavefield characteristics; Spatial Coherency; Dense Array; Rock Site; Soft-soil*
53 *site; Cephalonia.*

54 INTRODUCTION

55 The effects of spatial variability of earthquake ground motion (SVEGM), in other words the
56 amplitude and phase variability of ground motion observed at two adjacent positions, have
57 been of great interest for a long time in the design and analysis of large and extended structures.
58 Though such variation is, generally, attributed to the wave passage delay, spatial incoherence
59 and local site effects, the spatial incoherence caused by the effects of complex wave
60 propagation and scattering has increasingly caught interest of the research community. It is a
61 common practice to characterize the spatial phase variations by coherency functions from the
62 stochastic analysis of dense seismic arrays. These functions represent the degree of correlation
63 between ground motion measurements at two adjacent positions as an exponential decay with
64 increasing frequency and interstation distance.

65

66 A great deal of existing literature into SVEGM has been devoted to the development of
67 coherency models (see Zerva and Zervas 2002 and Liao 2006 for a review). However, the
68 generalizability of these models remains uncertain due to the lack of their association with the
69 physical site parameters. Much of the current literature is dedicated to the analysis of rather
70 large dimension arrays (>100 m), located mostly at soil sites, and relevant for distances longer
71 than the dimensions of most structures. Besides, complex wave propagation and scattering
72 effects at rock-type ground conditions have received scant attention as the seismic wavefield
73 is expected to be dominated by direct body waves at stiffer sites. As a result, there remains a
74 paucity of research on the physical implications of short-distance spatial incoherency from
75 rock-site dense arrays even though spatial incoherency has also been observed at rock sites
76 owing to geological complexities such as weathering and shallow fracturation (Somerville et
77 al. 1991; Steidl et al. 1996). In reality, horizontally extended structures (e.g., bridges) may be
78 supported at different site conditions, soil or rock or a combination, and thus be exposed to

79 different degrees of differential motions. Therefore, extending our attention to rock sites
80 alongside soil sites could be instrumental in understanding the SVEGM as well as evaluating
81 the site specific risk of infrastructures. The present study attempts to contribute to this area of
82 research by presenting the results obtained from the seismic events recorded at two nearby
83 dense arrays, located at a rock and a soft-soil site, at Argostoli in the Cephalonia Island of
84 Greece. It aims at investigating to what extent non-direct, diffracted surface waves contribute
85 to the seismic wavefield at both sites, and at examining the site-dependence of the spatial
86 variation. The part of the research concerning the soft-soil site is drawn from Imtiaz (2015) and
87 Imtiaz et al. (2014, 2015, 2017a, b). Hence, this paper first provides a brief literature review of
88 the rock-site coherency models, presents the dense array deployed at Argostoli rock site,
89 focuses on the results obtained from the corresponding data analysis, and proceeds on their
90 comparison with those from the soil site.

91

92 The two dense arrays consist of 21 broadband velocimeters, placed in four circles, at diameters
93 20 to 360 m (rock) and 10 to 160 m (soft-soil), centered on a reference station. Two sets of
94 earthquakes (40 events for the rock and 46 for the soft-soil array), with local magnitudes M_L
95 2–5, occurring within an epicentral distance (R_{epi}) of 200 km, are utilized to estimate the
96 apparent propagation characteristics of the waves and to identify the energies carried out by
97 Love and Rayleigh waves. The ‘lagged’ coherency is calculated as a measure of random
98 variability of the Fourier phase between any two stations in the array. Median coherency is
99 obtained for four separation distance intervals (10–20, 20–30, 30–40, and 80–90 m), common
100 to both arrays, and compared with some widely-quoted coherency models developed for rock
101 and soil sites.

102 COHERENCY FROM ROCK SITES

103 In current literature, very few coherency models are available for rock sites (e.g. Cranswick
104 1988; Menke et al. 1990; Toksoz et al. 1991; Abrahamson 2007). Schneider et al. (1992)
105 compared coherency estimations from a set of rock- and soil-site arrays with the coherency
106 function developed by Abrahamson et al. (1991) based on a soil site data (LSST, Taiwan array)
107 for separation distance shorter than 100 m. The authors observed that the soil coherency fits
108 well with the LSST model while the rock coherency seems comparable only when the site is
109 not affected by topographic variations. Abrahamson (2007) studied the same set of arrays but
110 removed the ones affected by the topographic variations. The author compared the respective
111 average coherency from the soil, soft-rock and hard-rock sites, and observed a clear site
112 dependence. At interstation distance 15–30 m, the coherency increases with the stiffness of the
113 site. At 50–70 m, the soil and soft-rock estimates become comparable and the hard-rock
114 remains higher. Zerva and Stephenson (2011) investigated seismic records from an array
115 located on a variable site condition (valley and soft-rock) and observed that the complex wave
116 propagation pattern affects not only the valley but also the surrounding soft-rock. The authors
117 then highlighted the significance of irregular subsurface topography and formation of surface
118 waves in the physical understanding of the spatial variation. Konakli et al. (2014) estimated
119 coherency from the USGS Parkfield rock array for the 2004 Parkfield mainshock and compared
120 with some existing models. Their result revealed that the observed variability of coherency
121 decay rate, as a function of frequency and interstation distance, was not reproduced by the
122 semi-analytical model of Luco and Wong (1986). At smaller interstation distances (<100 m),
123 the estimates were lower than those from the empirical model of Ancheta et al. (2011), updated
124 from Abrahamson et al. (1991), and the trend reverses at longer distances (>300 m). The
125 authors concluded, once again, by indicating a rather complex dependence of coherency on
126 effects related to source, propagation, topography and site, and the difficulty associated with

127 the task of explaining them by means of general, simplified processes. In a recent study, Svay
128 et al. (2017) estimated plane-wave coherency from an another subset of data consisting of
129 smaller magnitude (M_L 2.7-3.6), local events (R_{epi} 10-30 km) from the Argostoli rock array and
130 observed that both the Luco and Wong (1986) and Abrahamson (2007) models do not provide
131 a satisfactory match with estimated coherency values. A better fit is obtained when site-specific
132 decay parameters are considered in the functional forms of the parametric models.

133 **ARGOSTOLI DENSE SEISMIC ARRAYS**

134 Located in the north-westernmost boundary of the Aegean plate, the Ionian Island of
135 Cephalonia is considered to be one of the most active seismo-tectonic regions in Europe. The
136 high seismicity of the area is associated with the dextral Cephalonia Transform Fault (CTF),
137 situated in the northwestern boundary of the island (Figure 1a) and capable of generating
138 earthquake magnitudes up to $M = 7.4$ (Louvari et al. 1999). The rate of seismic slip in the CTF
139 is observed to be varying from 7 to 30 mm/yr (Anzidei et al. 1996; Hollenstein et al. 2006).
140 Historical data show that more than 10 earthquakes of magnitudes between 6.5 and 7.5
141 occurred in the area between 1900 and 1998 causing major destruction. Two major earthquakes
142 (moment magnitude M_w 6.1 and 6.0, and hypocentral depth ~ 10 km) hit the area on January
143 26 and February 3, 2014 (Karakostas et al. 2015; Theodoulidis et al. 2016). These events were
144 associated with the CTF and occurred within 20 km distance of the Argostoli town, damaging
145 a significant number of engineering structures. They were also followed by numerous
146 aftershocks up to M_w 5.5. A post seismic campaign was conducted at Argostoli area (Figure
147 1a) within the framework of the SINAPS@ project, funded by the French Research Agency
148 (ANR). A dense two-dimensional (2D) dense array, referred as “rock array” in this work
149 (Figure 1b,c), intended to study short distance spatial variability of ground motion, was
150 deployed on the rock formation from February 6 to March 10, 2014 (Perron et al. 2018). This

151 network complements the dataset recorded by a geometrically similar, smaller-size array,
152 referred as “soft-soil array” here (Figure 1b,d), on the nearby small, shallow, sedimentary
153 valley of Koutavos-Argostoli. The array was operational from September 2011 to April 2012
154 under the framework of EU-NERA (European Union - Network of European Research
155 Infrastructures for Earthquake Risk Assessment and Mitigation) 2011–2014 project (Imtiaz
156 2015). The rock array was deployed on a Cretaceous massive limestone (karstified and
157 fractured) formation located about 2 km southeast of the soft-soil array. Figure 1b displays the
158 geological formation of the study area after Cushing et al. (2016). Geophysical surveys led to
159 the determination of the V_{S30} (i.e., the harmonic mean of the shear-wave velocities over the
160 30 m of soil) as around 250 m/s for the soft-soil site, and 830 m/s for the rock site.

161

162 The rock array consists of 21 three-component (3C), broadband velocimeters (Güralp
163 CMG6TD sensors with 30 s eigenperiods with integrated digitizers) belonging to different
164 SINAPS@ partners (Perron et al. 2017). The stations are placed on the same geological unit,
165 along four circles of diameters 20, 60, 180 and 360 m around the central station B0R0 (Figure
166 1c). On each circle, the stations branch off from B0R0 in N67°E, N139°E, N218°E, N286°E
167 and N356°E directions. The signals are digitized at 200 samples per second; hence, the Nyquist
168 frequency is 100 Hz. Figure 1d shows the configuration of the soft-soil array. This array also
169 consists of 21 3C, broadband velocimeters (Güralp CMG40T sensors with eigenperiods
170 between 30 and 60s) connected to Nanometrics Taurus digitizers, belonging to the French
171 SISMOBRESIF national pool of portable seismic instruments. The stations are positioned in
172 four concentric circles, with diameters 10, 30, 80 and 160 m, around the central station A00.
173 Five stations, branching off from A00 in five directions, N39°E, N112°E, N183°E, N255°E
174 and N328°E, are placed on each concentric circle. The resolvable frequency range for both the
175 arrays could be considered as 1–20 Hz for slowness over 0.0004 s/m.

176 Both arrays provided a rich database by recording more than thousand earthquakes. The rock
177 array recorded nearly 1850 good quality earthquakes with local magnitude ranging from 1 to
178 5, at epicentral distances up to 300 km, mainly corresponding to the seismic sequence of the
179 aforementioned two consecutive earthquakes occurred in 2014, including some tens of events
180 at greater distances and more varied azimuths (Hollender et al. 2015). A subset of 40 events
181 with very good signal to noise ratio, local magnitude (M_L) 2 to 5, hypocentral depth (H) 4 to
182 40 km, epicentral distance (R_{epi}) 5 to 200 km, recorded by at least 20 stations of the array, is
183 selected for the analysis. The events are chosen such that a homogeneous distribution of
184 epicentral distance, magnitude and azimuthal coverage could be achieved, considering that
185 most of the events occurred along the CTF in the west and some local $M_L > 3.5$ events were
186 saturated. The origin time, location and the magnitude of the events have been taken from the
187 catalog of NOA (National Observatory of Athens) and Karakostas et al. (2015). The location
188 of the selected events are presented in Figure 2a, and their characteristics are provided in Table
189 1. It is to be noted that in this study, the entire seismogram (duration beginning from the P-
190 wave arrival up to the end of coda) is used for the wavefield analysis while the time window
191 representing mainly the most energetic phase carried out by the S-wave is used for coherency
192 estimation. Both durations are also given in Table 1. Figure 2b presents the durations of the
193 selected events for coherency analysis as a function of respective hypocentral distances. A
194 summary of the number of events grouped according to different parameters (distance,
195 magnitude, backazimuth) is given in Table 2. Finally, the results from the analysis of rock array
196 data are compared to those from a set of 46 events with similar characteristics (M_L 2-5 and R_{epi}
197 up to 200 km) recorded by the soft-soil array. The details of this database is provided in Imtiaz
198 (2015) and (Imtiaz et al. 2017b).

199

200 **WAVEFIELD ANALYSIS**

201 The MUSIQUE algorithm (Hobiger et al. 2012, 2016) is used to perform the wavefield
202 characterization of the array data. MUSIQUE is a combination of the "classical" Multiple
203 Signal Characterization (MUSIC) (Schmidt 1986; Goldstein and Archuleta 1987) and the
204 quaternion-MUSIC (Miron et al. 2005, 2006) methods, that allows not only the extraction of
205 backazimuth and apparent phase velocity (or slowness) of the dominant waves crossing the
206 array but also their characterization as Love or Rayleigh waves. Detailed description of the
207 MUSIQUE algorithm for 3C signals is available in Hobiger (2011) and Hobiger et al. (2016).
208 Only a brief summary of the method is presented in the current article.

209

210 The basic principle of the method is to separate the signal and noise subspaces and then to
211 estimate the signal parameters. Let us assume a dataset recorded by an array of N single-
212 component sensors. As a first step, MUSIQUE calculates the cross-correlation matrix (in time
213 domain) or cross spectral matrix (in frequency domain) for each frequency and time windows
214 of interest, and defines the eigenstructure of the covariance matrix. The eigenvectors
215 corresponding to the K strongest eigenvalues define the signal subspace ($K < N$) while the $N-K$
216 weakest eigenvalues define the noise subspace. Then, from the set of array manifold vectors,
217 the vectors that give the minimum projection onto the noise subspace are determined through
218 the search of the maxima of the directional function (the so-called MUSIC spectrum). Once the
219 signal vectors are determined, propagation direction and phase velocity of the dominant signals
220 are computed as the functions of time and frequency. Quaternions, an extension of complex
221 numbers into four dimensions (e.g. Ward 1997), are then used to characterize polarization
222 parameters of an incident wave. The quaternion-MUSIC algorithm merges both the complex-
223 valued data vectors of radial and vertical components into a single data matrix so that the phase
224 information and the sense of rotation of the particle motion remains naturally preserved, hence,

225 allowing the distinction between retrograde and prograde Rayleigh wave motions. Although
226 the classical MUSIC is able to identify multiple sources, the MUSIQUE code is presently
227 limited to a single wave contribution, i.e. to the most dominant source ($K=1$) as it requires
228 projecting the horizontal signals with respect to the identified wave backazimuth in order to
229 estimate the polarization parameters.

230

231 As a first step of the data analysis, the entire length of signal is divided into smaller sub-
232 windows of five periods as a function of frequency. Here 200 log-based steps between 1 to
233 20 Hz frequency are used and the sub-windows are overlapped by 50%. A spectral averaging
234 is done over five frequency samples centered on each frequency step in order to assure a well-
235 conditioned covariance matrix. Then the slowness and backazimuth (θ) of the most dominant
236 incident wave for each sub-window of the signal is determined from the MUSIC spectrum. The
237 radial and transverse components are computed by projecting the east-west (EW) and north-
238 south (NS) components of the signals along the identified backazimuth (θ) and its orthogonal
239 direction. During the post-processing, the results from the sub-windows containing energy less
240 than the median energy of all the analyzed windows of an event are filtered out. Additionally,
241 the sub-windows having signal to noise ratio less than 5 and slowness outside the range of
242 0.0004 to 0.008 s/m (apparent velocity 125 m/s to 2500 m/s) are also filtered out. A detailed
243 description of the post-processing calculations is given in Imtiaz (2015).

244

245 Figure 3 shows EW component of the velocity time series of an $M_L=3.5$ earthquake (ID # 11
246 in Table 1) recorded by all the stations of the rock array. The event occurred at a hypocentral
247 depth (H) of 31 km, an epicentral distance (R_{epi}) 13 km and, and a backazimuth (θ) $N348^\circ E$.
248 Figure 4a illustrates the retrieved backazimuths of the dominant incident waves from all the
249 analyzed sub-windows as a function of time. The colorbar corresponds to the ‘normalized

250 energy' of the respective windows, defined as the energy carried out by the sub-window
251 divided by the squared Fourier amplitude. On this plot, the P and S wave arrivals and their
252 respective durations could be followed through the energy concentration of the results. It can
253 also be observed that the energetic incident waves are not necessarily arriving from the source
254 direction (N348°E, marked by the red horizontal line); they are rather scattered along other
255 directions even in the earlier part of the seismogram (<10 s). Figure 4b presents the 2D
256 histogram of backazimuth distribution as a function of frequency. In order to group the results,
257 72 grid points have been considered between N0°E and N360°E angles (interval = 5°) for the
258 backazimuth axis and 31 log-based grid points between 1 and 20 Hz for the frequency axis.
259 The colorbar indicates the summed 'normalized energy' of the analyzed sub-windows falling
260 into the backazimuth-frequency grids. The histogram demonstrates that most of the energetic
261 waves at frequencies lower than ~6 Hz are appearing from north to north-east (roughly N330°E
262 to N30°E), which corresponds mainly to the event's backazimuth direction. The scattered
263 energetic waves are observed at higher frequencies (>6 Hz).

264 *Decomposition of diffracted surface waves*

265 Next, the results corresponding to the direct arrivals (considered as waves coming from
266 backazimuth $\pm 20^\circ$ direction) are eliminated for each event in order to separate the diffracted
267 wavefield and the estimates from all the individual events are summed. During the wavefield
268 analysis, the dominant incident waves are identified as Love or Rayleigh waves based on the
269 energy content of the analyzed signal window. A sub-window is characterized as Love- or
270 Rayleigh-dominant, respectively, when the estimated transverse energy or the summation of
271 the radial and vertical energy is more than 70% of the total energy of that window. If none of
272 these criteria is fulfilled, no wave is identified for the sub-window under consideration. In order
273 to summarize the observations from all the events and to focus on the surface wave
274 composition, only the results corresponding to the windows 'identified' as Love or Rayleigh

275 are considered for further analysis. Hence, the ‘non-identified’ windows are also subtracted in
276 the process.

277

278 Figure 5a shows the summary of the backazimuth distribution of the entire diffracted wavefield
279 (waves ‘identified’ as Love and Rayleigh + waves remained ‘unidentified’) from all the 40
280 events recorded by rock array. Same as Figure 4b, 72 grid points have been considered between
281 $N0^{\circ}E$ and $N360^{\circ}E$ angles (interval = 5°) for the backazimuth axis and 31 log-based grid points
282 between 1 and 20 Hz for the frequency axis. The colorbar indicates the summed ‘normalized
283 energy’ of all the sub-windows from the dataset falling into the backazimuth-frequency grids.
284 It seems that the diffracted waves are arriving approximately from the north-south direction
285 over the entire frequency range for the rock array. The summary of the diffracted waves
286 identified as Love and Rayleigh is presented in Figure 5b and Figure 5c, respectively. While a
287 weak dominance of Rayleigh waves arriving from the north-south direction is evident at higher
288 frequencies (>5 Hz), an even weaker dominance of Love waves from the north-east is observed
289 at lower frequencies (<5 Hz). On the contrary, the similar results of the wavefield analysis
290 from the soft-soil array (Imtiaz et al. 2014) presented in Figure 5d-e show the predominance of
291 significant diffracted surface waves, propagating along SW-NE direction (Figure 5d), beyond
292 the fundamental frequency (~ 1.5 Hz) of the valley. The lower frequencies (1-3 Hz) seem to be
293 dominated by Love waves (Figure 5e) while the higher frequencies by both Love and Rayleigh
294 waves (Figure 5f).

295

296 Figure 6a-d illustrate the arithmetic mean $\pm 1\sigma$ of diffracted and direct, Rayleigh and Love
297 energy as a percentage of the ‘total analyzed energy’ (all direct and diffracted waves) for the
298 rock and soft-soil array. On an average, about 20% of the ‘total analyzed energy’ from the rock
299 array could be characterized as diffracted Love and Rayleigh waves (Figure 6a). Love wave

300 composition seems to be slightly higher than Rayleigh only up to ~2 Hz frequency. In case of
301 the soft-soil array (Figure 6b), the wavefield is found to be composed of a much higher
302 proportion of surface waves, average ranging from 40% to 60%, over the entire analyzed
303 frequency range. A clear frequency dependence of the energy distribution is also observed:
304 lower frequencies (1–2.5 Hz) are dominated by Love waves, while certain narrow bands of
305 higher frequencies are dominated either by Love or Rayleigh waves. Correlatively, only up to
306 5% and 5–10% of the ‘total analyzed energy’ could be characterized as direct surface waves
307 for the rock and soft-soil arrays, respectively (Figure 6c,d). At both sites, more than 80% of the
308 mean surface wave energy (Love + Rayleigh) correspond to the diffracted waves while only
309 around 20% to the direct waves (Imtiaz 2015; Sbaa et al. 2017).

310 **COHERENCY ANALYSIS**

311 *Estimation of lagged coherency*

312 Coherency is a complex-valued function and characterizes the variation in Fourier phases
313 between two ground motions. By definition, lagged coherency is supposed to remove the
314 effects of systematic variation coming from the time delay in wave arrivals, known as the
315 ‘wave-passage effect’. Therefore, the two time histories under consideration are aligned using
316 the time lag that leads to the largest correlation of the motions. In this study, the records from
317 each station in the array are aligned with respect to the central station (B0R0 for the rock and
318 A00 for the soil array). However, no remarkable time lag (generally <2% of the coherency
319 window) is observed for both arrays. The duration of the most energetic, S-wave dominated
320 signal window is selected for coherency analysis by visually inspecting the records of the
321 central station. Details of the time-window selection procedure and the coherency estimation
322 are provided in Imtiaz et al. 2017 (a, b).

323

324 The selected time windows of the signals are cut and tapered by using 5% cosine bell window
 325 at both ends. The cross- and auto-spectra of the motions are calculated and averaged over 11
 326 frequency points by using a Hamming ($M=5$) window. Finally, the lagged coherency between
 327 a pair of stations is estimated as the modulus of the complex coherency, i.e., the ratio between
 328 the smoothed cross spectrum (\bar{S}_{jk}) of the pair of stations (j, k), normalized by the geometric
 329 mean of the smoothed power spectra at the respective stations ($\bar{S}_{jj}, \bar{S}_{kk}$), as function of the
 330 interstation distance (ξ) and frequency (1 to 25 Hz), and given by,

331

$$332 \quad |\bar{\gamma}_{jk}(\xi, f)| = \frac{|\bar{S}_{jk}(f)|}{\sqrt{|\bar{S}_{jj}(f)\bar{S}_{kk}(f)|}} \quad (1)$$

333

334 The values of lagged coherency range from 0 to 1 for fully uncorrelated to fully correlated
 335 motions. When the interstation distance or frequency tends to zero the estimates tend to 1.
 336 Theoretically, at higher frequencies and larger interstation distances coherency tends to zero.
 337 However, in actual calculations the modulus of the complex coherency tends to a constant
 338 value significantly larger than zero due to the level of smoothing applied to the inherent noise
 339 in the data. The median value of this constant is considered as 0.33 when an 11-point Hamming
 340 window is applied for frequency smoothing (Abrahamson 1992). As the coherency estimates
 341 below this value are uninterpretable, it is termed as the ‘resolvability threshold’ in this article.

342

343 The lagged coherency estimated from the EW component of the velocity seismograms (Figure
 344 3) of the event no.11 is shown in Figure 7a-d. The station pairs consist of the central station
 345 and the other stations located along the five array-branch directions (N 356, N 67, N 139, N
 346 213 and N 286) on the concentric circles. For example, the station pairs considered in the Figure
 347 7a are B0R0-B1R1, B0R0-B2R1, B0R0-B3R1, B0R0-B4R1 and B0R0-B5R1 (Figure 1c).
 348 Thus the four interstation or separation distances considered in Figure 7 correspond to the radii

349 of the concentric circles in the array, 10, 30, 90 and 180 m. As expected, the coherency is
350 observed to decay with increasing frequency and separation distance. No clear variation is
351 observed along the different direction of the pairs at a certain separation distance.

352 *Median from the dataset*

353 As the variability of coherency estimates is skewed, it is preferable to perform the statistical
354 analysis on the ATANH (or, \tanh^{-1}) transformation in order to produce approximately normally
355 distributed data (Harichandran 1991; Abrahamson et al. 1991). The ‘median resolvability
356 threshold’ of lagged ATANH coherency is 0.34 when 11-point Hamming window is applied
357 (Abrahamson 1992).

358

359 Coherency analysis from the soft-soil array (Imtiaz et al. 2017b) showed that when the median
360 coherency is derived by averaging all station pairs irrespective of their orientations, the two
361 geometric horizontal components (EW, NS) as well as their rotation along any physical
362 direction result in almost similar estimates. The authors also suggested that the source related
363 parameters had negligible effect on the loss of coherency of the dataset. Therefore, the
364 horizontal components of the ground motion are not rotated and simply the EW and NS
365 components are considered for presenting the results. Lagged coherency of the three
366 components from all the available pairs of stations (max. 210 pairs per event) is estimated for
367 the entire dataset. Table 3 shows the number of ‘ideally available’ (assuming data from all the
368 21 stations in the array are available) pairs of stations per event for the groups of separation
369 distances between 10 and 100 m. Four separation distance ranges (10–20, 20–30, 30–40 and
370 80–90 m) that include maximum number of pairs from both arrays have been chosen to present
371 the results. Median of the coherency estimates from the pairs available at a given distance range
372 for a single event is defined as the ‘individual median’. Then the ‘global median’ of the entire
373 dataset at that given separation distance range is derived by combining all the available pairs

374 from all the events. Table 4 shows the number of total pairs considered from the entire dataset
375 for the estimation of ‘global median’ coherency.

376

377 The average coherency estimates from the rock array in terms of the ‘individual median’ of
378 each event (thin black curves) along with the ‘global median’ (thick black curve) of all events,
379 as a function of frequency, are presented in Figure 8a-d for the EW component and in Figure
380 8e-h for Z (vertical) components. The corresponding ‘global median’ from the soft-soil array
381 is also plotted in thick red curve. The measure of variability or standard deviation (σ) about the
382 respective ‘global median’ curve is also provided by dashed lines. Here the σ is calculated in
383 terms of Median Absolute Deviation (MAD), defined as the median of the absolute deviations
384 (i.e., residuals) about the median of the observations (see Huber 1981; Imtiaz et al. 2017b). A
385 logarithmic scale is used for the frequency axis to highlight the results that are less affected by
386 the resolvability threshold. However, the estimated median curves are observed to be well
387 above the threshold. The coherency seems to be larger on rock than on the soil, at least up to
388 5 Hz for the EW and 10 Hz for the Z components, and at all separation distances.

389

390 A directional variation of coherency estimates, depending on the orientation of the station-
391 pairs, was observed in the results from the soft-soil array (Imtiaz et al. 2017b). Therefore,
392 median coherency estimates from the horizontal components of the arrays are compared, by
393 grouping the pairs lying in five array-branch directions at different separation distances. The
394 results from the EW component are shown in Figure 9a-d for the rock array and in Figure 9e-
395 h for the soft-soil array. In case of the soft-soil array, a consistent directional dependence of
396 coherency is observed at all separation distances, the highest coherency being in the N 328 and
397 the lowest in the N 255 directions corresponding to the valley-parallel and -perpendicular axes,

398 respectively (Imtiaz et al. 2017b). However, no remarkable orientation dependence of
399 coherency is observed for the rock array.

400 **COMPARISON WITH EXISTING PARAMETRIC MODELS**

401 *Selected parametric models*

402 A great variety of functional forms of lagged coherency is available in the literature although
403 very few are based on rock data. Some widely quoted parametric models are selected in this
404 article in order to compare with the median estimates from Argostoli data. Two semi-empirical
405 and two empirical models, applicable for short interstation distances (<100 m), pertinent to soil
406 and rock sites, are chosen and briefly presented below:

407

408 (a) The semi-empirical model of Luco and Wong (1986): It is based on the analysis of shear
409 waves propagating through a random medium and expressed as a double exponential decay
410 function of interstation distance (ξ) in m and angular frequency (ω) in rad/sec,

$$411 \quad |\gamma(\xi, \omega)| = \exp(-\alpha^2 \omega^2 \xi^2) \quad (2)$$

412 The coherency drop parameter is given by $\alpha = v/V_s$, where V_s is the average shear-wave
413 velocity of the medium along the wave travel-path and v is a constant related to the medium
414 properties. Typical values of α are suggested within 2×10^{-4} to 3×10^{-4} sec/m. In this study, rather
415 than estimating α from the dataset, the median value 2.5×10^{-4} sec/m is taken.

416

417 (b) The semi-empirical model of Menke et al. (1990): It is one of the very few rock site
418 coherency models available in the existing literature, expressed as an exponentially decaying
419 function of separation distance (ξ) in m and frequency (f) in Hz,

$$420 \quad |\gamma(\xi, f)| = \exp(-\alpha f \xi) \quad (3)$$

421 where, α value is proposed within the range of 4×10^{-4} to 7×10^{-4} sec/m. Again, the median
 422 value of $\alpha = 5.5 \times 10^{-4}$ sec/m is considered in this study. The functional form of this model is
 423 almost similar to that of Luco and Wong (1986) except that the latter expresses a squared
 424 exponential decay of coherency. The Menke et al. (1990) model is seemingly the square root
 425 of the Luco and Wong. Hence, the Luco and Wong model shows a larger decrease at higher
 426 frequencies and larger separation distance, but the decay starts relatively later as "alpha"
 427 coefficient is smaller.

428

429 (c) The empirical model of Abrahamson (2007): Abrahamson proposed a purely empirical
 430 plane-wave coherency model for rock sites based on 78 earthquakes recorded at the Pinyon
 431 Flat array, which is expressed as,

$$432 \quad |\gamma^{PW}(\xi, f)| = \left[1 + \left(\frac{f \tanh(a_1 \xi)}{f_c(\xi)} \right)^{n_1(\xi)} \right]^{-1/2} \left[1 + \left(\frac{f \tanh(a_1 \xi)}{a_2} \right)^{n_2} \right]^{-1/2} \quad (4)$$

433 The coefficients are given as:

434 $a_1=0.4$, $a_2=40$, $n_1(\xi)=3.8-0.04\ln(\xi+1)+0.0105[\ln(\xi+1)-3.6]^2$, $n_2=16.4$ and $f_c=27.9-4.82$
 435 $\ln(\xi+1)+1.24[\ln(\xi+1)-3.6]^2$ for the horizontal component; and

436 $a_1=0.4$, $a_2=200$, $n_1(\xi)=2.03+0.41\ln(\xi+1)-0.078[\ln(\xi+1)-3.6]^2$, $n_2=10$ and $f_c=29.2-5.20$
 437 $\ln(\xi+1)+1.45[\ln(\xi+1)-3.6]^2$ for the vertical component.

438

439 The plane-wave coherency is calculated by taking the real part of the smoothed cross-spectrum
 440 after aligning the ground motions on the best plane-wave speed. As such, it includes the random
 441 variations in the wave passage effect and should be lower than the lagged coherency, especially
 442 at higher frequencies and longer interstation distances.

443

444 (d) The empirical model of Ancheta et al. (2011) : Abrahamson et al. (1991) proposed an
 445 empirical model based on the estimation of lagged coherency from 15 earthquakes recorded at

446 a small array (LSST) located at a soil site in Lotung, Taiwan. This model allows investigations
 447 of coherency for the horizontal components of the ground motion and for station separation
 448 distance smaller than 100 m. Ancheta et al. (2011) proposed a slight modification in the first
 449 term of this model, by comparing an analysis from a set of events recorded by the Borrego
 450 Valley differential array (BVDA) located on an alluvium site, in order to correct the negative
 451 bias at separation distance (" ξ ") smaller than 30 m. The resulting model is described as,

$$452 \quad |\gamma(\xi, f)| = \tanh \left\{ (3.79 - 0.499 \ln(\xi)) \left[\exp\{(-0.115 - 0.00084\xi)f\} + \frac{f^{0.878}}{3} \right] + 0.35 \right\} \quad (5)$$

454 *Observation vs the selected parametric models*

455 The estimated median coherency from the Argostoli rock and soft-soil arrays are compared at
 456 four separation distances with the aforementioned parametric models, and are illustrated in
 457 Figure 10a-d for the horizontal and in Figure 10e-h for the vertical component. Here the
 458 horizontal coherency is the arithmetic mean of the EW and NS component estimates. Instead
 459 of using the ATANH transformation, the results are presented in terms of coherency ($|\gamma|$) in
 460 order to facilitate the interpretation. The models are computed for distances 15, 25, 35 and
 461 85 m. Note that the semi-empirical (Luco and Wong 1986; Menke et al. 1990) and empirical
 462 plane-wave (Abrahamson 2007) coherency models approach zero while the lagged coherency
 463 estimates from the data and the empirical model (Ancheta et al. 2011) tend to the resolvability
 464 threshold at higher frequencies. Thus, they are not comparable at those frequencies.

465
 466 The rock coherency seems to be in fair agreement with the models over all the frequencies at
 467 10–20 m, and up to 3 Hz frequency at other distances. As interstation distance increases, the
 468 models tend to overestimate the actual observations. The coherency values are always
 469 overestimated by the Abrahamson's hard-rock model. This may be due to the fact that the
 470 Argostoli rock array is situated on a lightly softer formation ($V_{s30}=830$ m/s) compared to the

471 data the model is based on ($V_{s30}=1030$ m/s). However, the soil-site model of Ancheta et al.
472 (2011) also overestimates the rock array coherency at all the distances above 20 m.

473

474 One fundamental issue with the semi-empirical models is that they consider a constant decay
475 rate with the wavelength whereas the real data exhibit varying decay tendencies with respect
476 to frequency and interstation distance. Therefore, the double exponential decay of Luco and
477 Wong (1986) model is observed to fit very well with the rock estimates at 80–90 m (up to about
478 6 Hz frequency) but not at shorter interstation distances. At shorter distances, coherency from
479 the Argostoli data are observed to decay much faster compared with the semi-empirical models.
480 This issue is somewhat taken care of by the empirical models but the overall decay rate in the
481 data still remains faster than the models. Coherency estimates from the Argostoli soil array
482 seems to be relatively poorly predicted by all the models. At shorter interstation distances (up
483 to 40 m) and higher frequencies (>5 Hz) the observed rock coherency is, surprisingly, slightly
484 lower than the soil. This could be attributed to the soil heterogeneities associated with the rock
485 site (karst filled with decalcification clay, fractured zone) whose characteristic "wavelength"
486 of variation is rather weak (metric to decametric).

487

488 Similar conclusions could be drawn for rock array data in case of the vertical component
489 (Figure 10e-h). However, the vertical coherency from soil array seems to be decaying at a much
490 faster rate up to about 7–8 Hz and shows a reverse tendency afterwards, especially at shorter
491 interstation distances (10–40 m). This increasing trend is totally absent in the rock data. A
492 possible explanation could be that the vertical coherencies at the soil site might be increased at
493 some higher frequencies because of the presence of converted S-P waves or Rayleigh waves
494 with higher phase velocities.

495 **DISCUSSION AND CONCLUSIONS**

496 The current study performs a twin analysis of wavefield characterization and spatial
497 incoherency of seismic motion recorded at two dense arrays of stations, one located on a rock
498 site, the other on a soft-soil site. As for the rock array, a large number (40) of small to moderate,
499 shallow, local and regional events with relatively even distribution in magnitude, distance and
500 backazimuth around the site, are analyzed. The loss of coherency is estimated for interstation
501 distance ranging from 10 to 360 m. The median of the estimates at different interstation
502 distance ranges (<100 m) have been examined and compared with those estimated from a
503 similar dataset recorded at a nearby soil site as well as with some of the existing parametric
504 models. The key results derived from the first part of the work, that is, the analysis of the
505 wavefield composition, reveal that some wave scattering over the frequency range of 1–20 Hz
506 is arriving from the north-south and north-east directions of the rock site for almost all the
507 events. However, only 20% of the total seismogram energy could be associated to diffracted
508 surface waves (Love and Rayleigh). As for the soft-soil array, about 40–60% of the analyzed
509 seismogram energy could be characterized as diffracted surface waves associated to 2D or 3D
510 geometrical effects, appearing mainly from the valley-width directions, with clear frequency-
511 dependent dominance of Love and Rayleigh waves (Imtiaz et al. 2014). Besides, more than
512 80% of the total surface wave energy is observed to be carried out by the diffracted waves at
513 both sites. Hence, it is evident that the formation of diffracted surface waves is much more
514 significant, especially on the sedimentary valley (soft-soil site) compared to the rock site, and
515 may play an important role in causing the higher loss of coherency and its directional
516 variability.

517

518 In the second part of the work, the estimated median lagged coherency from the rock data
519 exhibits very similar decaying tendencies for all the three components (EW, NS and Z). The

520 coherency is also found to be, generally, higher than that from the nearby soil site. However,
521 at shorter interstation distances and higher frequencies, rock coherency seems to be slightly
522 lower than that of soil. This could likely be attributed to the karstic-type weathering of the
523 Cretaceous limestone of the investigated rock site. This weathering shows strong lateral
524 lithological variation, with alternation at metric to decametric scale, of massive outcropping
525 limestone areas with zones dominated by decalcifying clay content. Further investigation is
526 suggested to better quantify this observation. The directional variability of lagged coherency,
527 depending on the orientation of the station pairs, observed at the soft-soil site appears to be
528 negligible for the rock site. Unlike the soft-soil site, coherency from vertical component seems
529 to be comparable with the horizontal at the rock site. Comparison of the median coherency
530 estimated from the arrays with the empirical models does not show any consistent correlation.
531 Overall, the rock coherency is overestimated by all the parametric models even by the empirical
532 one based on the soil data. Apparently, the decay rate of lagged coherency, as a function of
533 frequency, seems to be not only site dependent but also interstation distance dependent. The
534 fastest decay rate from the Argostoli data (both soft-soil and rock sites) is not well-captured by
535 the empirical models. On the other hand, the variability in the decay rate, especially with
536 respect to the interstation distance, is not addressed by the semi-empirical models. The findings
537 of the present work reaffirm the site dependent nature of spatial variation and emphasize the
538 need for more rigorous efforts on studying different sites in order to better understand the
539 spatial variability of ground motion.

540 **DATA AND RESSOURCES**

541 The data used in the present work are open. The data from the rock and soft-soil sites can be
542 obtained on the data repositories detailed in Perron et al. (2017) and Theodoulidis et al. (2017),
543 respectively.

544 **ACKNOWLEDGEMENTS**

545 This work was carried out under the SINAPS@ project that receives French funding managed
546 by the National Research Agency under the program “Future Investments” (SINAPS@
547 reference: ANR-11-RSNR-0022). SINAPS@ is a “Seism Institute” project
548 (<http://www.institut-seism.fr/en/>). We would like to thank all the contributors of this work,
549 especially for helping on the field during the SINAPS@ post-seismic campaign, Régis
550 Cottureau (ECP), Marc Cushing (IRSN), Alberto Frau (CEA/EMSI), Sébastien Hok (IRSN),
551 Philippe Langlaude (CEREMA), Aurore Laurendeau (CEA/LDG), Armand Mariscal (ISTerre)
552 and Alexandros Savvaidis (EPPO-ITSAK), as well as our local help on the Cephalonia Island,
553 Agis Konidaris (TEI) and Chrisostomos Andreou.

554 **REFERENCES**

- 555 Abrahamson NA (2007) Program on Technology Innovation: Effects of Spatial Incoherence
556 on Seismic Ground Motions. Rpt. No. EPRI 1015110, Palo Alto, CA
- 557 Abrahamson NA (1992) Spatial Variation of Earthquake Ground Motion for Application to
558 Soil-Structure Interaction. Rpt. No. EPRI TR-100463, Tier 1, Palo Alto, CA
- 559 Abrahamson NA, Schneider JF, Stepp JC (1991) Empirical Spatial Coherency Functions for
560 Application to Soil-Structure Interaction Analyses. *Earthq Spectra* 7:1–27. doi:
561 10.1193/1.1585610
- 562 Ancheta TD, Stewart JP, Abrahamson NA (2011) Effects of Surface Geology on Seismic
563 Motion. In: 4th IASPEI / IAEE International Symposium, University of California Santa
564 Barbara. California,
- 565 Anzidei M, Baldi P, Casula G, et al (1996) Repeated GPS surveys across the Ionian Sea:
566 evidence of crustal deformations. *Geophys J Int* 127:257–267. doi:
567 <https://doi.org/10.1111/j.1365-246X.1996.tb04718.x>
- 568 Cranswick E (1988) The information content of high-frequency seismograms and the near-
569 surface geologic structure of “hard rock” recording sites. *Pure Appl Geophys PAGEOPH*
570 128:333–363. doi: 10.1007/BF01772604
- 571 Cushing EM, Hollender F, Guyonnet-Benaize C, et al (2016) Close to the lair of Odysseus
572 Cyclops : the SINAPS@ postseismic campaign and accelerometric network installation
573 on Kefalonia island – Site effect characterization experiment. In: McCalpin J, Gruetzner C
574 (eds) *Proceedings of the 7th INQUA on Paleoseismology, Active Tectonics and*
575 *Archeoseismology*, Crestone, Colorado, USA, 30 May - 3 June.
- 576 Goldstein P, Archuleta RJ (1987) Array analysis of seismic signals. *Geophys Res Lett* 14:13–
577 16.
- 578 Harichandran RS (1991) Estimating the spatial variation of earthquake ground motion from

579 dense array recordings. *Struct Saf* 10:219–233.

580 Hobiger M (2011) Polarization of surface waves : characterization, inversion and application
581 to seismic hazard assessment. Université de Grenoble

582 Hobiger M, Cornou C, Bard P-Y, et al (2016) Analysis of seismic waves crossing the Santa
583 Clara Valley using the three-component MUSIQUE array algorithm. *Geophys J Int*
584 207:439–456. doi: 10.1093/gji/ggw289

585 Hobiger M, Le Bihan N, Cornou C, Bard P-Y (2012) Multicomponent Signal Processing for
586 Rayleigh Wave Ellipticity Estimation Application to Seismic Hazard Assessment. *IEEE*
587 *Signal Process Mag* 29:29–39–29–39.

588 Hollender F, Perron V, Imtiaz A, et al (2015) Close to the lair of Odysseus Cyclops : the
589 SINAPS@ postseismic campaign and accelerometric network installation on Kefalonia
590 Island. In: *Proceedings of the 9th Colloque National AFPS 2015, Marne-La-Vallée,*
591 *France, 30 November - 2 December.*

592 Hollenstein C, Geiger A, Kahle H-G, Veis G (2006) CGPS timeseries and trajectories of crustal
593 motion along the West Hellenic Arc. *Geophys J Int* 164:182–191. doi:
594 <https://doi.org/10.1111/j.1365-246X.2005.02804.x>

595 Huber P (1981) *Robust statistics.* New York: John Wiley

596 Imtiaz A (2015) Seismic wave field, spatial variability and coherency of ground motion over
597 short distances : near source and alluvial valley effects. PhD Thesis. Université Grenoble
598 Alpes, English. <NNT : 2015GREAU002>. <tel-01148138>. [https://tel.archives-](https://tel.archives-ouvertes.fr/tel-01148138)
599 [ouvertes.fr/tel-01148138](https://tel.archives-ouvertes.fr/tel-01148138)

600 Imtiaz A, Cornou C, Bard P-Y (2017a) Sensitivity of Ground Motion Coherency to the Choice
601 of Time Windows from the Argostoli Dense Seismic Array. *Bull Earthq Eng* (under Rev
602 0:1–30.

603 Imtiaz A, Cornou C, Bard P-Y, Hobiger M (2014) Diffracted Wave Field And Coherency

604 Analysis: An Example From Dense Array Network In Argostoli Basin, Cephalonia,
605 Greece. In: Proceedings of the Second European Conference on Earthquake Engineering
606 and Seismology (2ECEES), Istanbul, Turkey, 24-29 August. p 12, Paper # 2351

607 Imtiaz A, Cornou C, Bard P-Y, Zerva A (2017b) Effects of site geometry on short-distance
608 spatial coherency in Argostoli, Greece. *Bull Earthq Eng* 1–35. doi:
609 <https://doi.org/10.1007/s10518-017-0270-z>

610 Imtiaz A, Cornou C, Bard P-Y, Zerva A (2015) Spatial coherence of seismic ground motion
611 and geometric structure of the sub-surface: an example in Argostoli, Greece. In:
612 Proceedings of the 9th Colloque National AFPS 2015, Marne-La-Vallée, France, 30
613 November - 2 December.

614 Karakostas V, Papadimitriou E, Mesimeri M, et al (2015) The 2014 Kefalonia Doublet (M6.1
615 and M6.0), Central Ionian Islands, Greece: seismotectonic implications along the
616 Kefalonia transform fault zone. *Acta Geophys* 63:1–16. doi: 0.2478/s11600-014-0227-4

617 Konakli K, Der Kiureghian A, Dreger D (2014) Coherency analysis of accelerograms recorded
618 by the UPSAR array during the 2004 Parkfield earthquake. *Earthq Eng Struct Dyn*
619 43:641–659. doi: 10.1002/eqe.2362

620 Louvari E, Kiratzi A, Papazachos BC (1999) The Cephalonia Transform fault and its extension
621 to western Lefkada Island (Greece). *Tectonophysics* 308:223–236–223–236.

622 Luco J, Wong H (1986) Response of a rigid foundation to a spatially random ground motion.
623 *Earthq Eng Struct Dyn* 14:891–908.

624 Menke W, Lerner-Lam AL, Dubendorff B, Pacheco J (1990) Polarization and coherence of 5
625 to 30 Hz seismic wave fields at a hard-rock site and their relevance to velocity
626 heterogeneities in the crust. *Bull Seism Soc Am* 80:430–449.

627 Miron S, Le Bihan N, Mars J (2005) High resolution vector-sensor array processing using
628 quaternions.

629 Miron S, Le Bihan N, Mars J (2006) Quaternion-MUSIC for vector-sensor array processing.
630 IEEE Trans Signal Process 54:1218–1229.

631 Perron V, Hollender F, Mariscal A, et al (2018) Accelerometer, velocimeter dense-array, and
632 rotation sensor datasets from the Sinaps@ post-seismic survey (Cephalonia 2014-2015
633 aftershock sequence). Seism Res Lett (in press):0–0.

634 Sbaa S, Hollender F, Perron V, et al (2017) Analysis of rotation sensor data from the Sinaps@
635 Kefalonia (Greece) post-seismic experiment — Link to surface geology and wave-field
636 characteristics. Earth, Planets Sp. doi: <https://doi.org/10.1186/s40623-017-0711-6>

637 Schmidt RO (1986) Multiple emitter location and signal parameter estimation. IEEE Trans
638 Antennas Propag 34:276–280–276–280.

639 Schneider JF, Stepp JC, Abrahamson NA (1992) The spatial variation of earthquake ground
640 motion and effects of local site conditions. Proc theTenth World Conf Earthq Eng Madrid,
641 Spain 2:967–972.

642 Steidl JH, Tumarkin AG, Archuleta RJ (1996) What is a reference site? Bull Seismol Soc Am
643 86:1733–1748.

644 Svay A, Perron V, Imtiaz A, et al (2017) Spatial coherency analysis of seismic ground motions
645 from a rock site dense array implemented during the Kefalonia 2014 aftershock sequence.
646 Earthq Eng Struct Dyn. doi: 10.1002/eqe.2881

647 Theodoulidis N, Cultrera G, Cornou C, et al (2017) Basin effects on ground motion: the case
648 of a high-resolution experiment in Cephalonia (Greece). Bull Earthq Eng. doi:
649 <https://doi.org/10.1007/s10518-017-0225-4>

650 Theodoulidis N, Karakostas C, Lekidis V, et al (2016) The Cephalonia, Greece, January 26
651 (M6.1) and February 3, 2014 (M6.0) earthquakes: near-fault ground motion and effects
652 on soil and structures. Bull Earthq Eng 14:1–18. doi: 10.1007/s10518-015-9807-1

653 Toksoz MN, Dainty AM, Charrette III EE (1991) Spatial variation of ground motion due to

- 654 lateral heterogeneity. *Struct Saf* 10:53–77.
- 655 Zerva A, Stephenson WR (2011) Stochastic Characteristics of Seismic Excitations at a Non-
- 656 Uniform (Rock and Soil) Site. *Soil Dyn Earthq Eng* 31:1261–1284.

657 LIST OF FIGURE CAPTIONS

658 Figure 1 : Location of the study area and the arrays. (a) The Argostoli study area in the
659 Cephalonia Island. (b) Two arrays on the geological map of the area (Cushing et al. 2016).
660 Layout of the (c) rock and (d) soft-soil array.

661 Figure 2 : Dataset recorded by the rock-site array. (a) Location of the selected events with
662 respect to the site. The white circles show the other events recorded by the array. (b) Duration
663 of the selected time windows for coherency analysis.

664 Figure 3 : EW component of the velocity time series from an earthquake (ID #11 in Table 1:
665 $M_L = 3.5$, $Baz = N 348$, $R_{epi} = 13$ km, $H = 31$ km) recorded by the stations in the array. The time
666 window between the red lines are considered for the coherency analysis of the event.

667 Figure 4 : Backazimuth distribution of the dominant incoming waves from the event no. 11
668 (Table 1) as a function of (a) time and (b) frequency. Red line marks the event-backazimuth.

669 Figure 5 : Cumulative histogram of the backazimuth, extracted from the wavefield analysis of
670 the selected dataset, considering the entire diffracted wavefield, and diffracted waves identified
671 as Love and Rayleigh waves, respectively, for the (a)-(c) rock and (d)-(f) soft-soil array.

672 Figure 6 : Repartition of energy (mean $\pm 1\sigma$) carried out by identified Rayleigh (R) and Love
673 (L) waves, from the wavefield analysis of the selected dataset, considering the (a)-(b) diffracted
674 and (c)-(d) direct waves, for the rock and soft-soil array. Here, the waves coming from the
675 source-backazimuth $\pm 20^\circ$ direction are represented as the direct waves and the rest are
676 considered as the diffracted waves.

677 Figure 7 : Lagged coherency estimated from the EW component of velocity time series (Figure
678 3) of the event no. 11 recorded by the rock array. Here, coherency is presented for the pairs
679 between the central station (B0R0 in Figure 1c) and the other stations lying along five array-
680 branch directions at four separation distances (a) 10 m, (b) 30 m, (c) 90 m and (d) 180 m.

681 Figure 8 : Median ATANH coherency estimated from the (a)-(d) EW and (e)-(h) Z component
682 of the velocity time series at four separation distances. The corresponding lagged coherency
683 values are marked at the right-side axis. Thin grey curves represent the ‘individual median’ of
684 the events recorded by the rock array and thick black curves represent their ‘global median’.
685 Red thick curves show the ‘global median’ estimated from the dataset recorded by the soft-soil
686 array. The dashed curves show the respective $\pm 1\sigma$ bound of the ‘global median’ ATANH
687 coherency for the two arrays. Grey horizontal lines mark the coherency resolvability threshold.

688 Figure 9 : EW component of the ‘global median’ ATANH coherency, considering only the
689 station pairs lying along the five array-branch directions at four separation distances, from the
690 datasets recorded by the (a)-(d) rock and (e)-(h) soft-soil arrays.

691 Figure 10 : Comparison of the ‘global median’ coherency estimated from Argostoli rock and
692 soft-soil arrays with the existing semi-empirical (‘LW 1986’ for Luco and Wong 1986; ‘M
693 1990’ for Menke et al. 1990) and empirical (‘A 2007’ for Abrahamson 2007; ‘A 2011’ for
694 Abrahamson et al. 2011) coherency models for the (a)-(d) horizontal and (e)-(h) vertical
695 component of the ground motion at four separation distances.

696

697 **AUTHORS' AFFILIATIONS, ADDRESSES**

698 1. Afifa Imtiaz

699 BRGM, F-45060 Orléans, France

700 a.imtiaz@brgm.fr

701

702 2. V. Perron

703 CEA, DEN, F-13108 St Paul lez Durance Cedex, France

704 Univ. Grenoble Alpes, Univ. Savoie Mont-Blanc, CNRS, IRD, IFSTTAR, ISTerre, Bâtiment

705 OSUG C, CS 40700, 38058 Grenoble Cedex 9, France

706 IRSN, F-92260 Fontenay-aux-Roses, France

707 vincent.perron.mail@gmail.com

708

709 3. F. Hollender

710 CEA, DEN, F-13108 St Paul lez Durance Cedex, France

711 fabrice.hollender@cea.fr

712

713 4. P-Y. Bard

714 Univ. Grenoble Alpes, Univ. Savoie Mont-Blanc, CNRS, IRD, IFSTTAR, ISTerre, Bâtiment

715 OSUG C, CS 40700, 38058 Grenoble Cedex 9, France

716 pierre-yves.bard@univ-grenoble-alpes.fr

717

718 4. C. Cornou

719 Univ. Grenoble Alpes, Univ. Savoie Mont-Blanc, CNRS, IRD, IFSTTAR, ISTerre, Bâtiment

720 OSUG C, CS 40700, 38058 Grenoble Cedex 9, France

721 cecile.cornou@univ-grenoble-alpes.fr

722

723 6. A. Svay

724 IMSIA, University Paris-Saclay, France

725 angkeara.svay@polytechnique.edu

726

727 7. N. Theodoulidis

728 Institute of Engineering Seismology and Earthquake Engineering (ITSAK), GR-55102

729 Thessaloniki, Greece

730 ntheo@itsak.gr

731 TABLES

732 Table 1: Characteristics¹ of the selected events from the rock array (Figure 1 c) data

ID	YY-MM-DD_hhmmss	PGV (m/s)	M _L	H (km)	R _{epi} (km)	Baz (N°)	Ts (s)	T (s)	Missing station(s)
1	2014-02-06_180549	1.918	3.9	8	16.3	299.9	2.21	26	-
2	2014-02-07_012153	1.134	4.2	20	104.4	80.5	6.62	50	-
3	2014-02-08_031503	2.295	3.5	18	16.4	267.7	1.79	21	-
4	2014-02-08_213038	0.038	2.4	16	36.9	268.2	2.22	42	-
5	2014-02-10_100754	0.463	2.6	17	19.0	93.9	1.9	21	-
6	2014-02-10_102933	0.043	2.5	10	48.8	240.0	1.54	31	B2R3
7	2014-02-12_221352	0.161	3.1	13	30.8	349.0	2.39	31	-
8	2014-02-13_064249	0.744	3.2	7	10.6	264.6	1.91	31	-
9	2014-02-13_171411	0.125	2.7	15	31.0	244.0	1.77	31	-
10	2014-02-14_003600	0.017	2.8	10	128.9	81.2	2.82	41	-
11	2014-02-14_012313	0.490	3.5	13	30.7	348.9	2.83	31	B2R3
12	2014-02-15_011709	0.091	2.6	3	69.0	3.5	2.87	20	B2R3
13	2014-02-16_174243	0.034	2.0	18	15.8	110.9	1.99	17	B2R3
14	2014-02-17_050452	0.104	2.4	20	40.1	230.4	1.72	32	B2R3
15	2014-02-18_045311	0.067	2.6	15	32.5	342.6	2.09	31	-
16	2014-02-18_232338	0.212	2.7	16	36.4	240.7	2.79	26	-
17	2014-02-19_130252	0.049	2.6	8	39.3	8.8	2.79	31	-
18	2014-02-20_141936	0.157	3.0	21	9.7	7.7	2.53	20	B2R4
19	2014-02-20_171321	0.020	2.4	28	169.2	134.8	5	30	-
20	2014-02-23_024435	0.017	2.1	4	64.8	6.1	4.36	31	B2R2
21	2014-02-25_045354	0.097	2.7	13	90.4	145.4	3.2	31	-
22	2014-02-25_174743	0.028	1.9	21	11.8	106.9	2.25	21	-
23	2014-02-25_175828	0.017	2.5	18	74.1	101.1	5.87	51	-
24	2014-02-25_202722	0.026	2.9	17	62.8	124.4	4.48	51	-
25	2014-02-27_142314	0.077	2.6	17	31.1	251.1	1.96	31	-
26	2014-02-27_170311	0.026	2.7	10	53.3	359.9	3.36	31	-
27	2014-02-28_221354	0.020	3.4	21	175.0	87.6	14.5	120	-
28	2014-03-01_065830	0.027	1.7	24	105.1	92.7	3.6	29	-
29	2014-03-01_132540	0.025	2.0	6	35.7	5.4	2.4	31	-
30	2014-03-02_080146	0.020	2.2	15	33.7	126.5	2.99	42	-
31	2014-03-02_194201	0.233	3.3	38	33.7	139.5	2.5	41	-
32	2014-03-03_224948	0.080	2.8	12	30.1	348.2	1.65	31	-
33	2014-03-04_171008	0.059	2.8	8	53.3	1.8	2.85	47	-
34	2014-03-05_131346	3.187	3.4	20	21.0	252.2	3.13	21	-
35	2014-03-05_155453	0.043	2.7	11	43.3	358.7	2.59	41	-
36	2014-03-06_144115	0.015	2.0	18	34.1	136.0	2.48	30	-
37	2014-03-07_174356	0.027	3.9	16	177.2	145.0	15	120	-
38	2014-03-09_035823	0.038	2.2	23	39.0	239.1	3.2	32	-
39	2014-03-09_115801	0.032	2.5	9	171.8	79.8	5.1	38	B1R4
40	2014-03-10_055054	0.033	2.0	7	27.1	10.9	2.12	21	B1R4

¹ ID is the earthquake index, YY-MM-DD_hhmmss is the date and time of the earthquake occurrence, PGV is the Peak Ground Velocity, M_L is the local magnitude, H is the hypocentral depth, R_{epi} is the epicentral distance, Baz is the earthquake backazimuth, Ts is the duration of the time window for coherency estimation and T is the signal duration for wavefield analysis.

733 Table 2: The distribution of backazimuth, epicentral distance and magnitude of the selected
 734 dataset from the rock array

Backazimuth quadrant	Nb. of events	Epicentral Distance (km)	Nb. of events	Magnitude (M_L)	Nb. of events
NE	11	0-30	9	~2-3	29
NW	12	30-60	19	3-4	10
SE	10	60-100	5	4-5	1
SW	7	100-200	7		
Total	40		40		40

735

736 Table 3: Number of station pairs ideally available, when no station is missing, per event at
 737 different separation distances (up to 100 m) for the rock and soft-soil arrays

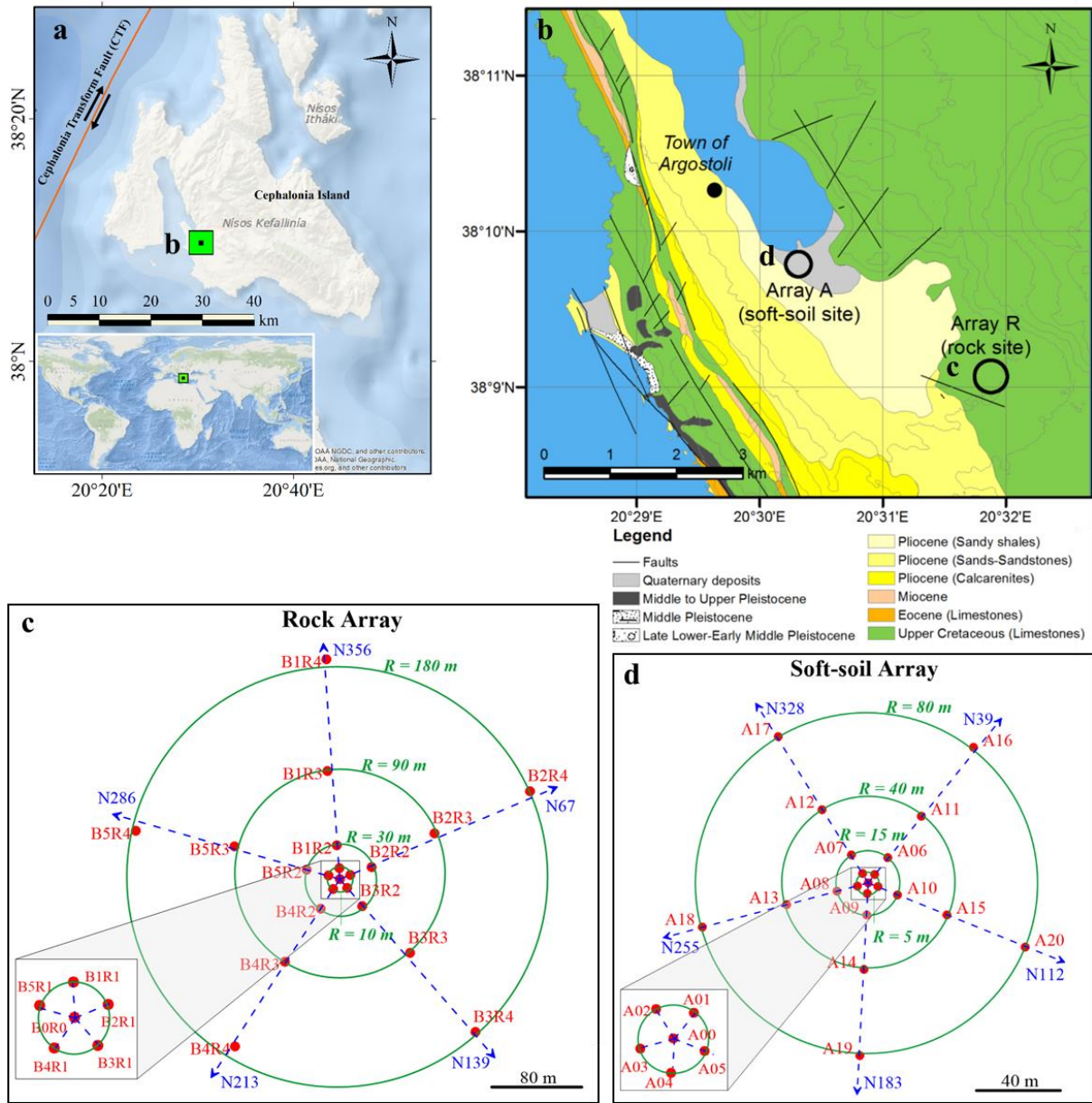
Separation Distance (m)	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Rock Array	19	15	16	0	7	3	2	24	15
Soft-soil Array	33	10	31	19	10	5	43	12	15

738

739 Table 4: Total number of actually available station pairs (considered for the median estimation)
 740 at the selected four separation distances for the rock and soft-soil arrays.

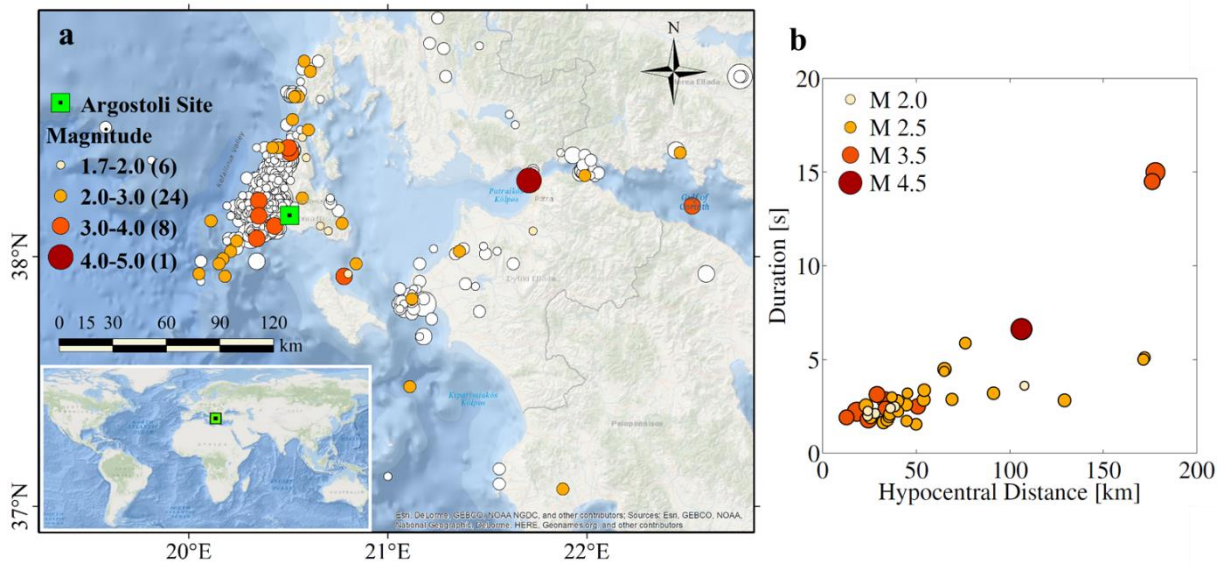
Separation Distance (m)	10-20	20-30	30-40	80-90
Rock Array (total pairs from 40 events)	759	597	636	928
Soft-soil Array (total pairs from 46 events)	1574	414	1241	430

741



743
 744 Figure 1 : Location of the study area and the arrays. (a) The Argostoli study area in the
 745 Cephalonia Island. (b) Two arrays on the geological map of the area (Cushing et al. 2016).
 746 Layout of the (c) rock and (d) soft-soil array.

747

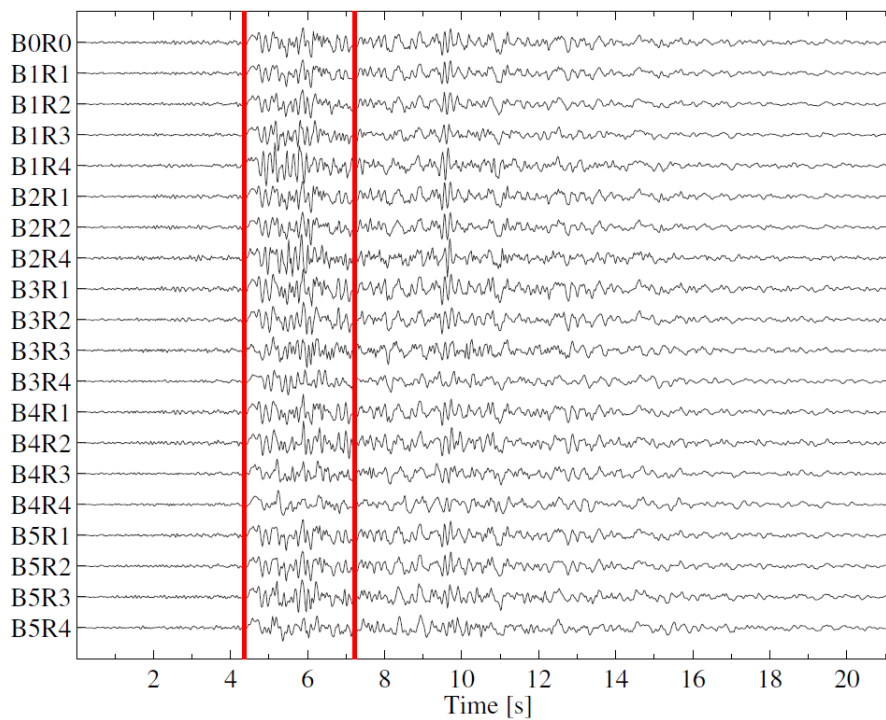


748

749 Figure 2 : Dataset recorded by the rock-site array. (a) Location of the selected events with

750 respect to the site. The white circles show the other events recorded by the array. (b) Duration

751 of the selected time windows for coherency analysis.

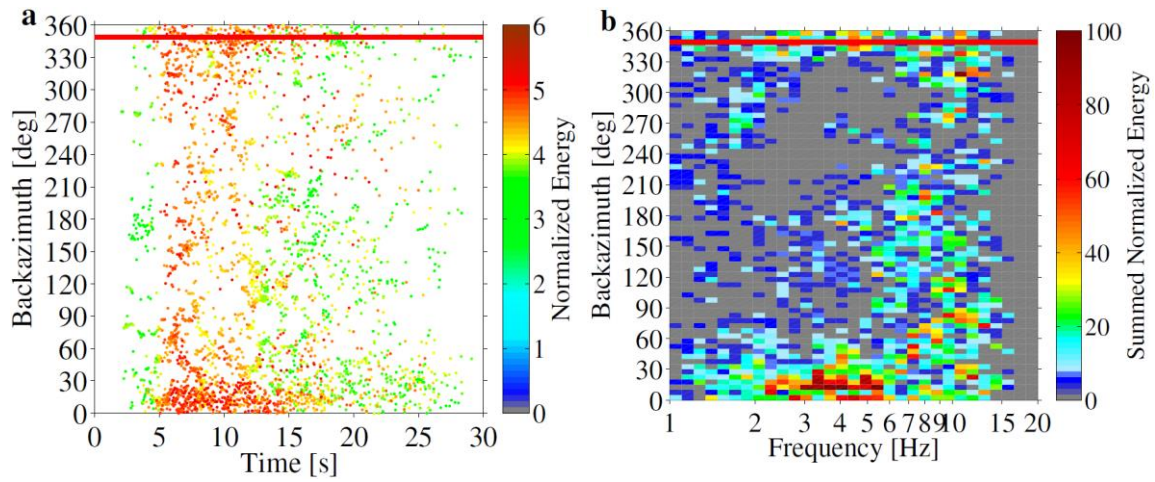


752

753 Figure 3 : EW component of the velocity time series from an earthquake (ID #11 in Table 1:

754 $M_L = 3.5$, $Baz = N 348$, $R_{epi} = 13$ km, $H = 31$ km) recorded by the stations in the array. The time

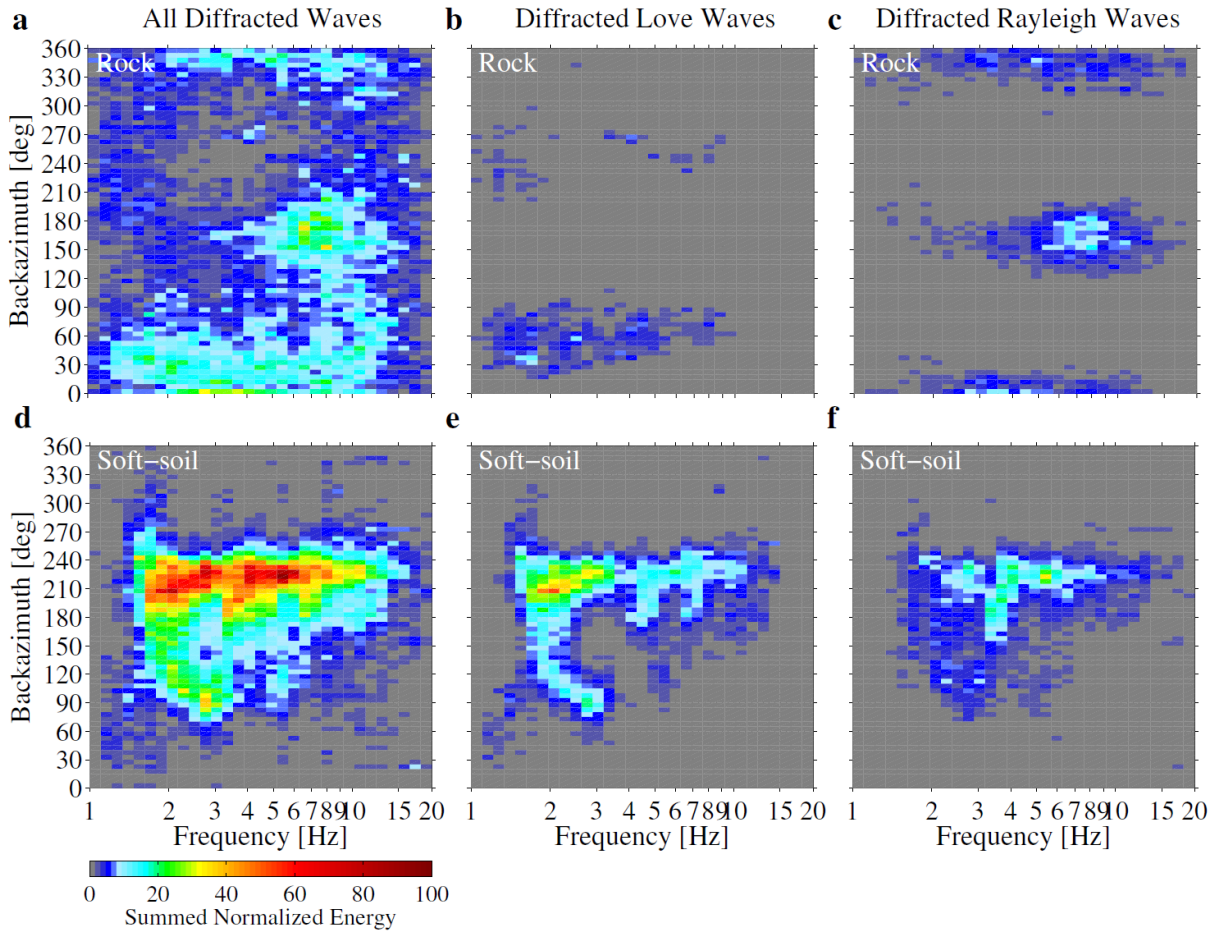
755 window between the red lines are considered for the coherency analysis of the event.



756

757 Figure 4 : Backazimuth distribution of the dominant incoming waves from the event no. 11

758 (Table 1) as a function of (a) time and (b) frequency. Red line marks the event-backazimuth.



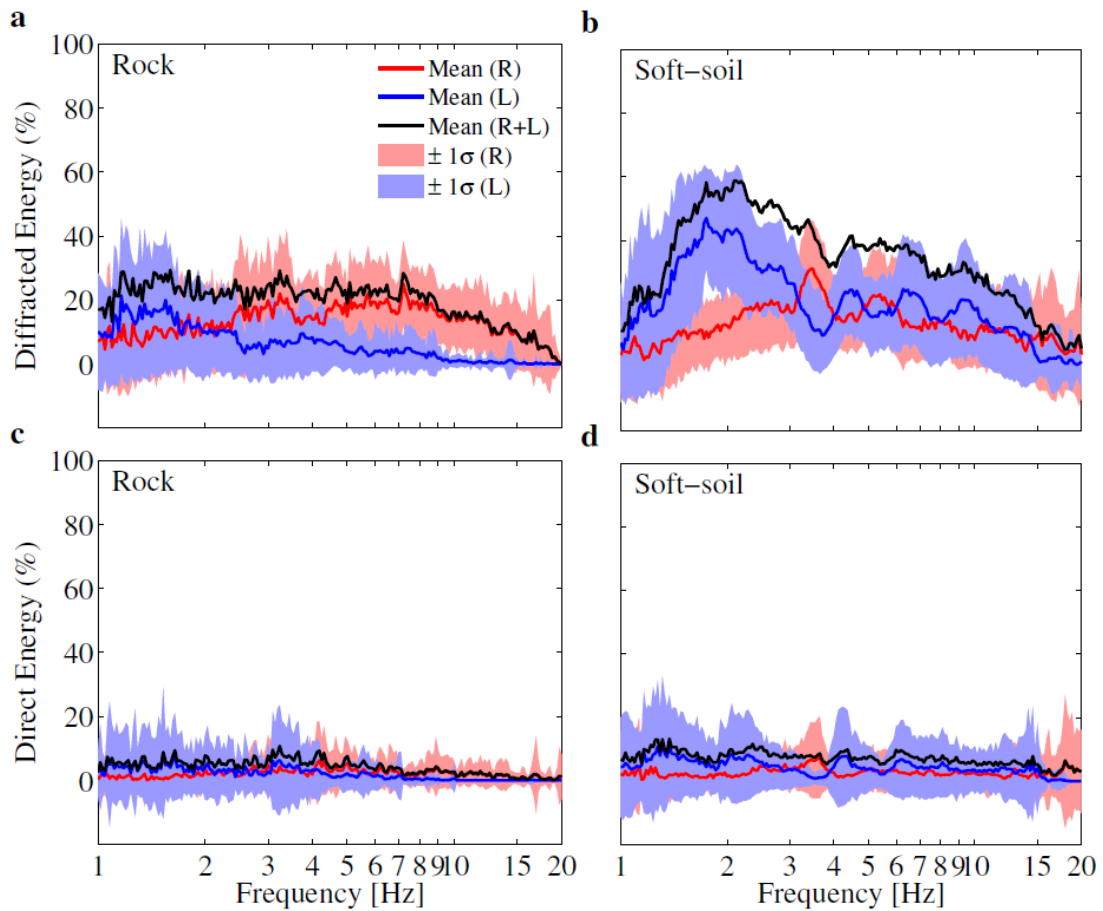
759

760

761 Figure 5 : Cumulative histogram of the backazimuth, extracted from the wavefield analysis of

762 the selected dataset, considering the entire diffracted wavefield, and diffracted waves identified

763 as Love and Rayleigh waves, respectively, for the (a)-(c) rock and (d)-(f) soft-soil array.



764

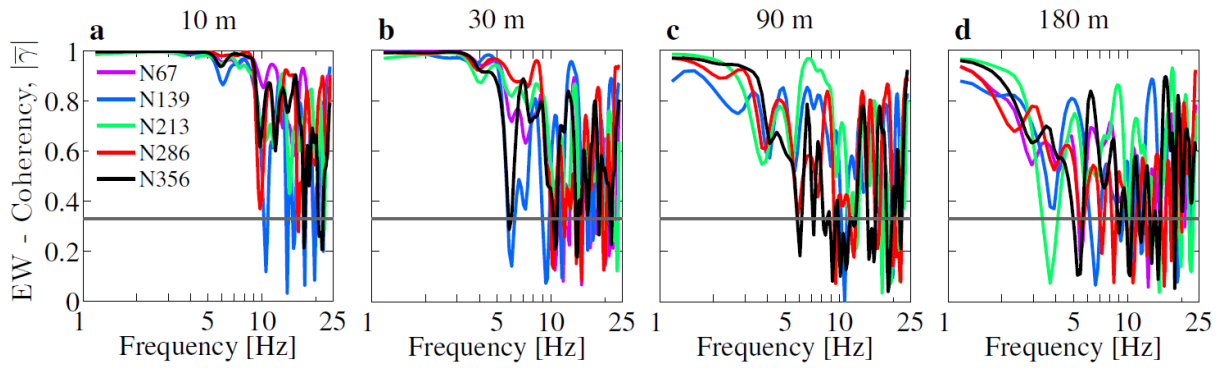
765 Figure 6 : Repartition of energy (mean $\pm 1\sigma$) carried out by identified Rayleigh (R) and Love
 766 (L) waves, from the wavefield analysis of the selected dataset, considering the (a)-(b) diffracted
 767 and (c)-(d) direct waves, for the rock and soft-soil array. Here, the waves coming from the
 768 source-backazimuth $\pm 20^\circ$ direction are represented as the direct waves and the rest are
 769 considered as the diffracted waves.

770

771

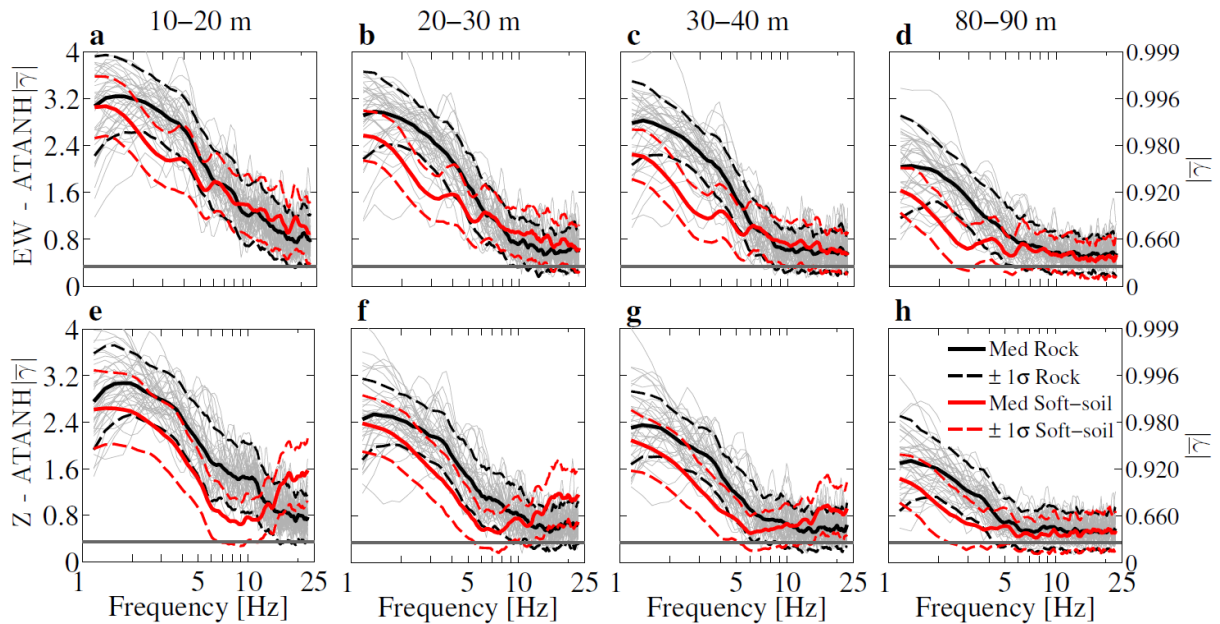
772

773



774

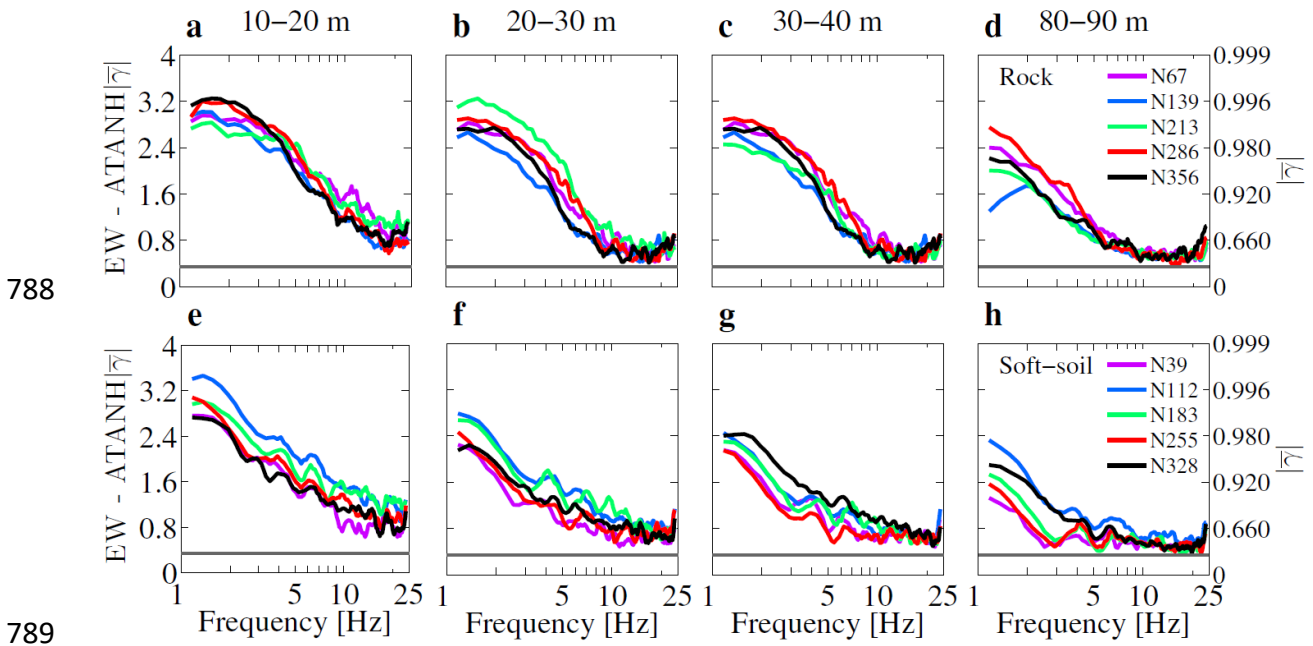
775 Figure 7 : Lagged coherency estimated from the EW component of velocity time series (Figure
 776 3) of the event no. 11 recorded by the rock array. Here, coherency is presented for the pairs
 777 between the central station (BOR0 in Figure 1c) and the other stations lying along five array-
 778 branch directions at four separation distances (a) 10 m, (b) 30 m, (c) 90 m and (d) 180 m.



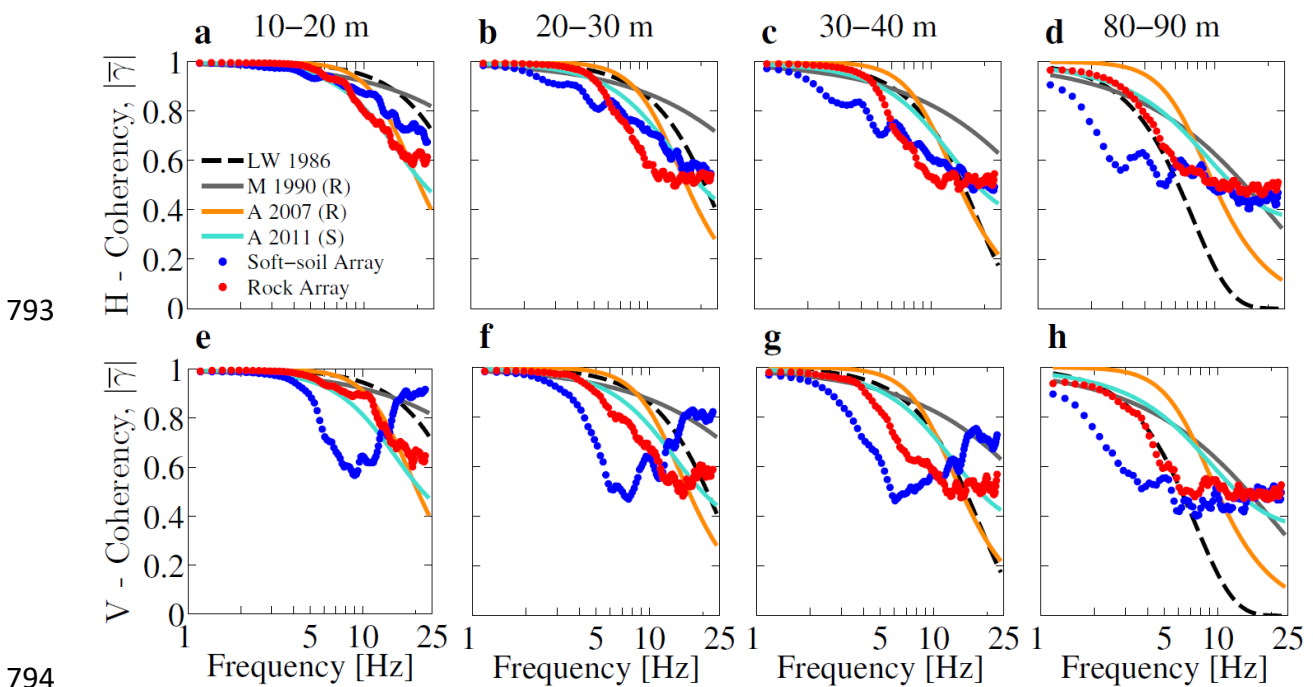
779

780

781 Figure 8 : Median ATANH coherency estimated from the (a)-(d) EW and (e)-(h) Z component
 782 of the velocity time series at four separation distances. The corresponding lagged coherency
 783 values are marked at the right-side axis. Thin grey curves represent the ‘individual median’ of
 784 the events recorded by the rock array and thick black curves represent their ‘global median’.
 785 Red thick curves show the ‘global median’ estimated from the dataset recorded by the soft-soil
 786 array. The dashed curves show the respective $\pm 1\sigma$ bound of the ‘global median’ ATANH
 787 coherency for the two arrays. Grey horizontal lines mark the coherency resolvability threshold.



790 Figure 9 : EW component of the ‘global median’ ATANH coherency, considering only the
 791 station pairs lying along the five array-branch directions at four separation distances, from the
 792 datasets recorded by the (a)-(d) rock and (e)-(h) soft-soil arrays.



795 Figure 10 : Comparison of the ‘global median’ coherency estimated from Argostoli rock and
 796 soft-soil arrays with the existing semi-empirical (‘LW 1986’ for Luco and Wong 1986; ‘M
 797 1990’ for Menke et al. 1990) and empirical (‘A 2007’ for Abrahamson 2007; ‘A 2011’ for
 798 Abrahamson et al. 2011) coherency models for the (a)-(d) horizontal and (e)-(h) vertical
 799 component of the ground motion at four separation distances.