



HAL
open science

Humming trains in seismology: an opportune source for probing the shallow crust

Laura Pinzon-Rincon, François Lavoué, Aurélien Mordret, Pierre Boué,
Florent Brenguier, Philippe Dales, Yehuda Ben-Zion, Frank Vernon,
Christopher Bean, Daniel Hollis

► To cite this version:

Laura Pinzon-Rincon, François Lavoué, Aurélien Mordret, Pierre Boué, Florent Brenguier, et al.. Humming trains in seismology: an opportune source for probing the shallow crust. *Seismological Research Letters*, 2021, 92 (2A), pp.623-635. 10.1785/0220200248 . hal-03174914

HAL Id: hal-03174914

<https://hal.univ-grenoble-alpes.fr/hal-03174914>

Submitted on 25 May 2021

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 **Humming trains in seismology: an opportune source for probing**
2 **the shallow crust**

3
4 Laura Pinzon-Rincon^{1,*}, François Lavoué¹, Aurélien Mordret¹, Pierre Boué¹, Florent
5 Brenguier¹, Philippe Dales^{1,6}, Yehuda Ben-Zion², Frank Vernon³, Christopher J. Bean⁴, and
6 Daniel Hollis⁵

7
8 1 - ISTERre, Univ. Grenoble Alpes, CNRS, Grenoble, France

9 *Corresponding author: laura-alejandra.pinzon-rincon@univ-grenoble-alpes.fr Université Grenoble
10 Alpes, ISTERre, CS 40700, 38058 GRENOBLE Cedex 9, France

11 2 - Department of Earth Sciences and Southern California Earthquake Center,
12 University of Southern California, Los Angeles, CA, USA

13 3 - IGPP, University of California, San Diego, La Jolla, CA, USA

14 4 - Geophysics Section, Dublin Institute for Advanced Studies, Dublin, Ireland

15 5 - Sisprobe, Grenoble, France

16 6 - Institute of Mine Seismology, USA

17 =

18 September 18, 2020

19 **Abstract**

20 Seismologists are eagerly seeking new and preferably low-cost ways to map and track changes in the
21 complex structure of the top few kilometers of the crust. By understanding it better they can build on
22 what is known regarding important, practical issues. These include telling us whether imminent
23 earthquakes and volcanic eruptions are generating tell-tale underground signs of hazard, about
24 mitigation of induced seismicity such as from deep injection of waste water, how the Earth and its
25 atmosphere couple, and where accessible natural resources are. Passive seismic imaging usually relies
26 on blind correlations within extended recordings of Earth's ceaseless "hum" or coda of well-mixed,
27 small vibrations. In this paper we are proposing a complementary approach. It is seismic
28 interferometry using opportune sources - specifically ones not stationary in time and moving in a
29 well-understood configuration. Its interpretation relies on accurate understanding how these sources
30 radiate seismic waves, on precise timing, on careful placement of pairs of listening stations, and on
31 seismic phase differentiation (surface and body waves). Massive freight trains were only recently
32 recognized as just such persistent, powerful cultural (human activity-caused) seismic sources. One
33 train passage may generate tremor with an energy output of a magnitude 1 earthquake and be
34 detectable for up to 100 km from the track. We discuss source mechanisms of train tremor and review
35 basic theory on seismic interferometry with opportune sources. Finally, we present case studies of
36 body- and surface-wave retrieval as an aid to mineral exploration in Canada and to monitoring of a
37 Southern California fault zone. We believe noise recovery from this new signal source, together with
38 dense data acquisition technologies such as nodes or Distributed Acoustic Sensing, will deeply
39 transform our ability to monitor activity in the shallow crust at sharpened resolution in time and
40 space.

41 **1 Introduction**

42 Vehicle traffic was long seen mainly as a pervasive source of nuisance noise that degrades seismic
43 records ([Douze and Laster, 1979](#)). But the recent and intriguing discovery of tremor from trains
44 startled seismologists. Studies soon followed on detection and characterization of these signals ([Riahi
45 and Gerstoft, 2015](#); [Li et al., 2018](#); [Green et al., 2017](#); [Fuchs et al., 2018](#); [Inbal et al., 2018](#)) as well as
46 source modeling ([Lavoué et al., 2020](#)). Earlier studies ([Nakata et al. \(2011\)](#); [Quiros et al. \(2016\)](#); [Chang
47 et al. \(2016\)](#)) proposed using traffic noise and seismic interferometry for both body- and surface-wave
48 imaging. These studies were, however, limited to highly local sources of background cultural noise
49 and near-surface applications.

50 In a fortuitous attempt to gather non-volcanic tremors (NVT) along the San Andreas Fault
51 Zone in Southern California, [Inbal et al. \(2018\)](#) discovered extended tremor sequences that shared
52 puzzling similarities with NVTs. But they traced the new discovery to massive freight trains running
53 along the nearby Coachella Valley. They could detect them as much as 100 km from the rails.
54 [Brenguier et al. \(2019\)](#) calculated that a single 1-km-long freight train rolling through a 10-km-long
55 railway section radiates energy equivalent to a magnitude 1 earthquake. By further using seismic
56 interferometry for correlation of this underfoot train noise ([Brenguier et al. \(2019\)](#); [Dales et al. \(2020\)](#))
57 showed it possible to extract useful information on the Earth's crustal structure and temporal changes
58 down several kilometers and that provides a potential alternative to costly monitoring of active
59 sources such as hydraulic thumping or explosives ([Tsuji et al., 2018](#)).

60 This paper reviews basic concepts and examples of the application of seismic interferometry to
61 train tremors. Its special focus is on long-range body-wave retrieval for crustal exploration and
62 monitoring (Fig. [1](#)). The Green's function is the elastic impulse response of the ground between a
63 seismic source and a seismic receiver, i.e., the signature of the ground structure encapsulated by its
64 effects on the velocity and other behavior of a signal as it traveled. Seismic interferometry is often

65 able, by correlating diffuse coda or seismic noise, to retrieve the Green's function between two seismic
66 sensors by turning one sensor into a virtual source. The impact of Green's function retrieval in recent
67 decades is revolutionary (e.g., [Campillo and Paul, 2003](#); [Shapiro et al., 2005](#); [Snieder, 2004](#);
68 [Wapenaar, 2004](#)). It spurred publication of at least 2000 seismology papers in the last 15 years. One
69 payoff from seismic interferometry and Green's function retrieval is improved crustal imaging
70 through correlation of pervasive surface wave noise generated in the oceans in the period range from
71 1 to 20 seconds. Recent studies have also unveiled the possibility of reconstructing body-waves at
72 global (e.g., [Poli et al., 2012](#); [Boué et al., 2013](#)) and local scales ([Draganov et al., 2009](#); [Nakata et al.,](#)
73 [2015](#); [Olivier et al., 2015](#); [Nakata et al., 2016](#)).

74 A perfect application of Green's function retrieval and seismic interferometry requires
75 correlation of either a fully diffused seismic wavefield or noise signals generated from all around the
76 studied region, including at depth ([Wapenaar, 2004](#)). In practice these demands are never met.
77 Seismologists must live with or find work-arounds for partial reconstructions and potentially biased
78 wave travel times ([Snieder et al., 2006](#); [King and Curtis, 2012](#)). Moving trains are welcome, opportune
79 sources of noise on well-mapped railways. It is essential that they be rigorously assessed for seismic
80 interferometry. Train traffic noise cannot be blindly correlated without considering the effects of
81 irregular source distribution on body-wave retrieval.

82 In this paper we first describe typical train noise signals, discuss recent models of mechanisms
83 that create train seismic radiations, and provide a map of the predicted extent of useful train noise in
84 the contiguous US. Second, we propose a methodological framework focusing our approach on the
85 stationary zones ([geographical area where we observe constructive interferences when cross-](#)
86 [correlating signals between two stations, Snieder, 2004](#)) and propose a signal processing strategy for
87 applying seismic interferometry to train noise with focus on long-range body-wave retrieval. We
88 finally review two recent case studies regarding mineral exploration in Canada and tectonic fault

89 monitoring in Southern California.

90

91 **2 The sound of trains in the Earth**

92 As noted, massive freight trains generate seismic waveforms with striking similarity to episodic
93 tectonic tremors. These may be from such events as slow-slip fault motion (Fig. 2 top). As [Inbal et al.](#)
94 [\(2018\)](#) report, the identity of the sources as manmade was not obvious because freight train traffic
95 often lacks cultural diurnal or weekly modulation, and typical train speed (25 m/s or 90km/h) is also
96 in the range of reported tectonic tremor migration velocity at depth. Train hum has however a distinct
97 signature with clear spectral lines above 1 Hz ([Fuchs et al., 2018](#)) illustrated in Fig. 2 for a train signal
98 recorded in Canada about 3 km from the railway (first case study presented below, see Section 5).

99 The engineering community has studied train-induced ground vibrations thoroughly to damp
100 them and mitigate potential hazards. Several source mechanisms are under study (e.g. [Connolly et al.](#)
101 [2015](#)) including quasi-static excitation due to axle loads, and dynamic interactions among trains,
102 tracks, and ground. In a recent study, [Lavoué et al. \(2020\)](#) showed that excitation due to axle loads is
103 the main mechanism producing the spectral characteristics of seismic signals at intermediate to long
104 distances from the railway (from hundreds of meters to tens of kilometers, [Fuchs et al., 2018](#); [Inbal et](#)
105 [al., 2018](#); [Li et al., 2018](#); [Brenquier et al., 2019](#)). One may then model train-generated seismic signals
106 by considering only the vertical forces due to loading applied by axles on the railroad ties (commonly
107 called sleepers) along the railway ([Krylov and Ferguson, 1994](#); [Lavoué et al., 2020](#)).

108 [Lavoué et al. \(2020\)](#) conclude that the spectral lines arise from complex interactions of
109 periodic loads through the regularly spaced wheels on the even more evenly separated sleepers. The
110 frequencies of these spectral lines depend on train geometry (i.e. train car length and wheel spacing
111 within each car), spacing between sleepers, and train velocity. We provide an open-source code to
112 assess the frequency response of a specific train (see link in Data and Resources). With most trains,
113 the dominant spectral lines are expected in the 1 to 20 Hz range, which is ideal both for high-
114 frequency surface wave tomography of the near subsurface and for crustal body-wave imaging and

115 monitoring (wavelengths not too large and scattering not too strong, [Brenquier et al., 2019](#)).

116 Our ability to predict the long-range, body-wave Peak Ground Velocity (PGV) of a moving train
117 tremor – the physical motion in the medium as signals go through it - is crucial to image formations
118 and monitoring any changes with seismic interferometry. [Lavoué et al. \(2020\)](#) propose that train
119 tremor PGV is directly proportional to the wagon weight for a given train length, and is a square-root
120 function of train length for constant wagon weight. Faster trains also generate higher PGVs. Moreover,
121 the ground stiffness beneath railways controls high-frequency content and amplitude of excitation
122 (trains traveling across rock or stiff soil generate higher-frequency and higher- amplitude signals).
123 This ground stiffness parameter may also reflect a coupling between the rail track and the ground.
124 While maximum detection distance may be limited (a few kilometers) in sedimentary basins due to
125 attenuation and weak excitation but, again, it can reach almost 100 km on a hard-rock substratum. In
126 southern California, for instance, [Inbal et al. \(2018\)](#) observed a freight train tremor signal from as far
127 as 90 km from the railway. At 45 km from the railway, they estimated a PGV of about 6×10^{-8} m/s.
128 By applying a simple correction for intrinsic attenuation and geometrical spreading for body (P)
129 waves, we now estimate that the level of PGV for a specific Coachella Valley train would be of the
130 order of 5×10^{-7} m/s at 10 km and 5×10^{-6} m/s at 1 km. These values are quite low. Even high-
131 sensitivity seismometers may record such train signals at long distances only in quiet environments.
132 Nevertheless, [Brenquier et al. \(2019\)](#) demonstrated that Coachella Valley train noise supported
133 seismic body-wave interferometry - with data recorded by an array of geophones (nodes) - as much
134 as 60 km from the railway (see Section [6](#)).

135 At shorter distances small-amplitude body waves might be barely visible in the raw data,
136 either because surface waves hide them or because they are below the ambient noise level. But it
137 should be possible to extract body waves from the correlations of train tremors by stacking data from
138 several passages (see Section [5](#)). Using train tremors for seismic interferometry thus depends both on

139 detection limits (instrument sensitivity and local noise level) and on reliably recognizable features in
140 train signals.

141 Observation from previous studies ([Inbal et al., 2018](#); [Brenquier et al 2019](#)) persuade us that
142 50 km is a typical maximum distance range for detecting tremors generated from large North
143 American freight trains. That led us to look into the spatial extent of detectable train tremor in the
144 entire contiguous US plus southern Canada (Fig. [3](#)). The map displays the main freight railway routes.
145 The swathes in colors represent high tonnage routes. Their width (100 km) is a rough guide to
146 potential long-range train tremor detection scope. This map does not take into account the reduced
147 detectability of signals in urban areas due to intense local noise and in sedimentary basins with strong
148 attenuation compared to the Southern California-Coachella Valley reference.

149 Noteworthy is that the Coachella Valley, a stretch of Sonoran Desert northwest of the Salton
150 Sea, is a singularly apt place to find practical uses for ground vibrations of massive trains. Its Union
151 Pacific RR tracks are a prime corridor to and from the ports of Los Angeles and Long Beach, the
152 Western Hemisphere's busiest seaport complex. Dozens of trains go through daily. The average length
153 is more than 2.5 km with more than 100 cars, often including multiple engines front and back. Rail
154 enthusiasts visit to make, and often to post on YouTube, mesmerizing videos of the immense steel
155 caravans rumbling by (see Data and Resources).

156 Annual freight tonnage (Fig. [3](#)) is a proxy for the number of trains travelling on rail sections.
157 Assuming an average train length of 2 km and a weight of 15 kilotons (according to statistics derived
158 from the public waybill samples, 2018¹), a tonnage of 100 MT/year corresponds to about 18 trains
159 per day. The number of trains per day will affect ability to stack the reconstructed body waves from
160 seismic interferometry. It also affects the temporal resolution needed in monitoring applications (see
161 Section [6](#)). This map highlights the potential of using trains as sources of opportunity. Potential
162 application may be in Cascade volcanoes, the Southern California's San Andreas Fault system, induced

163 seismicity (e.g. Oklahoma), and resource exploration and monitoring (minerals, water).

164

165 [1https://prod.stb.gov/wp-content/uploads/PublicUseWaybillSample2018.zip](https://prod.stb.gov/wp-content/uploads/PublicUseWaybillSample2018.zip)

166

167 **3 Seismic interferometry with opportune sources**

168 Seismic interferometry is a general term embracing all methodologies aiming to infer seismic

169 responses from the correlation of seismic signals observed at multiple receiver locations (e.g.,

170 [Wapenaar et al., 2010a,b](#)). To turn sensors into virtual sources, this concept has been refined in

171 seismology and seismic exploration, mostly in the last 20 years, based on the pioneering studies of

172 random fields or vertical planar wave autocorrelation ([Aki, 1957](#); [Claerbout, 1968](#)) and the time-

173 reversal principle in acoustics ([Fink, 1997](#)).

174 To retrieve a Green's function using the correlation or an equivalent operator the theory

175 heavily relies on either a stationary phase condition (e.g., [Snieder, 2004](#); [Roux et al., 2005](#)) and/or an

176 equipartition of modes defining a diffuse field (e.g., [Sánchez-Sesma and Campillo, 2006](#)). The

177 stationary phase condition implies that the correlation function's convergence towards the Green

178 function requires the presence of sources (or scatterers) in line with two carefully placed receivers. In

179 a 2D homogeneous medium, these stationary points define a hyperbolic area, outward from the

180 receiver pair, with an aperture that is frequency dependent (the lower the frequency, the broader the

181 calculated source region). Also known as Fresnel zones, these "kernels" are clues to the reliability of

182 the correlation's implied source locations. In 3D and for both surface and body wave retrieval, the

183 requirement of equipartition remains. Full Green function retrieval demands sources evenly

184 distributed, along an arbitrarily shaped surface enclosing the two sensors (e.g., [Wapenaar, 2004](#);

185 [Wapenaar and Fokkema, 2006](#)). However, even with a clearly dominant distribution of sources at the

186 free surface, several studies confirmed the feasibility of retrieving body waves (e.g., [Draganov et al.](#),

187 [2009, 2013](#)), and even explicitly using traffic noise ([Nakata et al., 2011](#)).

188 Each of the possible phases (or wave types) included in the Green's function has its own
189 source sensitivity. The main contributors to a particular phase are sources within its stationary phase
190 area. We can therefore measure a specific phase between two receivers by correlation of a source
191 within its stationary phase zone including the surface. The following case studies investigated P
192 waves from moving trains, and emerging from the interference between a direct P recorded at the
193 first station, and a PP (redirected once by a buried layer or formation edge) recorded by a second
194 station after a rebound below the first one. One can do the same with S waves, Fig. [4b](#).

195 Useful interference occurs if the seismic sources (trains) satisfy the stationary phase criterion:

$$196 \quad \Delta t = t_{pp} - t_p \leq t_{green} \pm \frac{T}{4},$$

197 where t_{pp} is the arrival time of the PP wave at the second receiver; t_p is the arrival time of the P
198 wave at the first station t_{green} is the arrival time of the P wave between the two receivers and T is the
199 dominant period. Note that using somehow controlled sources to retrieve body-wave response
200 through interferometry is similar to daylight imaging developed by ([Schuster et al., 2004](#)) or to the
201 virtual source approach discussed by ([Bakulin and Calvert, 2006](#)) for borehole imaging. For train
202 signals, we need to characterize the source and of course take into account that the sources are
203 moving.

204 One reason train signals are practical for interferometric studies is that we can easily detect,
205 or learn in advance, that a train is coming. If a railway is sufficiently close to a targeted area, a single
206 train's motion could illuminate many azimuths and potentially different depths. Figure [4a](#) shows an
207 example of geometry in Marathon (Ontario, Canada). There a railway essentially surrounds a
208 temporary array put in to assess an ore deposit (detailed in the following section). By selecting station
209 pairs aligned with train locations (illustrated for two positions by red and blue stars), one can
210 potentially illuminate the ore body from a broad azimuth range. Figure [4b](#) to 4d are schematics of

211 several P-wave interference scenarios, each with a pair of stations. They offer a perfect ballistic
212 interference between a diving P and PP wave (Figure [4b](#)) leading to a directly measurable diving P
213 wave between the two receivers; and a classical scenario of a scattered wavefield from which we
214 expect some random source energy to transit between the two receivers (Figure [4c](#)); See also a more
215 problematic interference between two diving wave, or a head wave recorded at the two stations
216 (Figure [4d](#)). Instances of this last scenario are sometime regarded as spurious correlations or virtual
217 refractions ([Dong et al., 2006](#); [Snieder et al., 2006](#); [Mikesell et al., 2009](#)). Although not included within
218 the impulse response between the two stations, this last correlation feature might be useful for
219 imaging if it is well distinguished from expected diving waves ([Dong et al., 2006](#)).

220 We decided to try to illuminate specific ray paths by using a data processing workflow,
221 starting with the selection of short time windows including specific train passages. This presumably
222 could be extended to any kind of seismic tremors and should help extract body waves between well-
223 selected pairs of stations useful for imaging and monitoring studies.

224

225 **4 Strategy for data processing**

226 Standard noise-correlation workflow typically removes strong transient events such as earthquakes
227 and then correlates the entire remaining time series recorded at different sensors ([Bensen et al.,](#)
228 [2007](#)). With opportune sources including train traffic we propose a novel workflow. It includes source
229 characterization with signal and station pair selections as alternatives to blind correlation. We thus
230 aim to improve the signal-to-noise ratio (SNR) of the reconstructed correlation functions and the
231 temporal resolution of monitoring studies. This approach is illustrated in sections [5](#) and [6](#) for imaging
232 and monitoring applications, respectively. Figure [5](#) summarizes the five main stages of our data
233 processing in comparison to the classical method of continuous blind data correlation.

234 The workflow's steps:

235 - Identify opportune source signatures in the continuous data and, if possible, locate these
236 sources perhaps by distance but at least in azimuth. As shown in section 2, the modeling of opportune
237 sources helps reveal the temporal and spectral content of the generated wavefield. Standard (short-
238 time average window / long-time average window) and more advanced techniques (e.g., [Meng et al.](#)
239 2019; [Kong et al., 2019](#)) detect these transient events. Array processing techniques (e.g., Cheng et al
240 2020) can be used to locate their sources.

241 - Station pair selection: With source location estimates in mind we can narrow down the
242 options for station pairs. For a given signal time window we use only station pairs for which the train
243 source is in a stationary phase zone. During a train passage, the energy carried by its seismic signal
244 reaches an array of sensors from a range of directions. Figure 4a illustrates two train positions at
245 different times (red and blue stars) and the associated selected stations for pair-wise correlations
246 (red and blue dots).

247 - Compute cross-correlations after proper time windowing and station pairs selection.

248 - Stack (by events, by azimuth): To improve SNR, we stack the cross-correlations over
249 different events. Cultural sources such as train traffic have the advantage of reliability and frequent
250 repetition.

251 - Measurement and analysis: Depending on the type of studies, various approaches such as
252 travel time measurements can enhance imaging and monitoring applications.

253
254 **5 Body- and surface-wave retrieval from correlations of train tremors applied to mineral**
255 **exploration**

256 We investigated a region near Marathon, Ontario, Canada (see Fig. 6b) where potential targets
257 include a high concentration of platinum group metals and minor Cu in a gabbro intrusion.

258 Reconstruction of high-frequency body-waves from train noise correlations was of significant

259 interest. A reason is such signals' sharp sensitivity to seismic velocity contrasts at depth, offering a
260 clear path to imaging geological boundaries. In the fall of 2018 we 1020 seismic stations in a backbone
261 array and a dense station line (see Fig. [6b](#)). We recorded 30 straight days of seismic signals.

262 [Dales et al. \(2020\)](#) showed that the main generators of high-frequency seismic noise in
263 Marathon are freight trains to the southwest. They reinforced earlier evidence that by selectively
264 using train traffic noise, one retrieves body waves better than does correlating a more extended or full
265 noise record. [Dales et al. \(2020\)](#) stacked correlations over 1 month, selecting all periods during which
266 the ambient noise came from the direction aligned with a dense W-E line of sensors. Their study is
267 illustrative but the results did not allow them to perform 3D imaging. We moved a step further by
268 separating and binning noise azimuths for virtual source retrieval in different directions. Following
269 the workflow proposed in section [4](#), we detected train passages, inferred the positions and azimuths
270 of the trains relative to the array, carefully selected station pairs and time windows for correlations,
271 and finally stacked by train passage and azimuth.

272 A more detailed workflow follows:

273 - We first generated a catalog of train passages with the covariance matrix method proposed
274 by [Seydoux et al. \(2016\)](#). This method uses the spatial coherence of the signals to detect emergent
275 signals in the noise. We applied the procedure to the entire data set day by day and detected 207 train
276 passages in 30 days. We retained for study approx. 180 events after skipping overlapping trains.
277 Beamforming concluded that the array receives energy from each train for approx. 80 minutes.

278 - Second, we extracted train signals from the rest of the recording, and selected station pairs in
279 line with train positions. To determine position, we did beamforming within 1-minute-long windows
280 using data filtered between 8 and 16 Hz (Fig. [6-d](#) and e, the right side shows 6 beamforming panels for
281 6 different events at two different times). Each panel corresponds to a one-minute beamforming time
282 window and one single train passage. We saw that with properly selected time windows for each

283 event, we got a tight and usable ranges of azimuth. We assumed that the main source of energy was
284 the train and noted the maximum beam power. We back-projected this signal onto the railway to
285 locate each train minute by minute. Figure [6-b](#) (red and blue cross) shows the train position from the
286 first beamforming panel (i.e., one single train). We then selected station pairs that are in line with the
287 train position for each minute, always taking the station closest to the railway as a virtual source (red
288 and blue arrows in Fig. [6b](#)). We applied an azimuthal filter of ± 5 degrees for each station pair with
289 respect to the train position.

290 - Third, we cross-correlated the selected station pairs minute by minute without overlap and
291 for each event (i.e., train passage). Filters excluded signals outside 15 to 40Hz to avoid surface waves.
292 We stacked cross-correlations according to their inter-station distances and collected them in
293 distance-binned correlation gathers for the selected station pairs (second step). In contrast, Figure [6a](#)
294 shows the stack of one-minute cross-correlation for a quiet period (i.e., non-train passage),
295 highlighting the absence of coherent wave propagation in this rather high-frequency window.

296 - In the fourth and last step we stacked events sharing the same train azimuth. We stacked
297 these correlation gathers into a reference azimuthal gather. We converged on a stable reference stack
298 from 6 train passages. We showed that by applying the workflow explained in section 4 we only
299 needed one minute of data and stacking over the 6 events to retrieve body-waves. Figures [6-d](#) and e,
300 left side, show the stacked section over 6 train passages with one-minute data segments.

301 We retrieved two dominant arrivals with an apparent velocity of 3.8 km/s and 7 km/s. There
302 are uncertainties, but we suggest that the first arrival is a P-wave, and the second one is probably a
303 mix of S- and surface waves. One can use both P- and S-waves plus high-frequency surface waves
304 jointly for imaging the subsurface. We need further analysis to assess the types of body-waves (direct,
305 refracted) and how one can use velocity variations in the azimuth's function for 3D imaging.

306

307 **6 Retrieving long-range body waves from train-tremor correlations to monitor the San Jacinto**
308 **Fault Zone**

309 Following studies by [Nakata et al. \(2015\)](#) and [Nakata et al. \(2016\)](#) of high-frequency body- wave
310 retrieval using dense seismic receiver arrays [Takano et al. \(2020\)](#), [Brenguier et al. \(2020\)](#) and [Zhou](#)
311 [and Paulssen \(2020\)](#) explored ways to monitor temporal changes of ballistic wave velocities. In this
312 section, we employ opportune seismic sources to passively monitor temporal changes and revisit the
313 experiment of [Brenguier et al. \(2019\)](#). Here, the goal was to use ballistic P-waves, reconstructed from
314 ambient vibrations between two dense arrays, to monitor subtle velocity changes at depth within the
315 San Jacinto Fault Zone (SJFZ, parallel to the San Andreas and part of the same fault complex).

316 [Brenguier et al. \(2019\)](#) showed that standard ambient noise correlation processing can retrieve
317 high-frequency direct P-waves that traveled between two arrays, one at Piñon Flat Observatory (PFO)
318 and the other on the Cahuilla Reservation (CIR, Fig. [7](#)). The main sources of these P-waves were
319 Coachella Valley freight trains traversing the Coachella Valley about 30 km to the East-North-East of
320 PFO. [Brenguier et al. \(2019\)](#) used full records of ambient noise to obtain stable direct P-wave
321 seismograms. We showed that, by carefully selecting time-windows where most of the energy is
322 generated by trains, the quality and spatiotemporal stability of the reconstructed P-waves rose. As
323 described in Figure [5](#), the standard three-step noise correlation computation workflow was replaced
324 by a four-step procedure to correlate only the main source of opportune energy i.e., here, trains. Our
325 workflow:

326 - First, we built our train catalog for the period of interest (July 22 to August 11 of 2018) using
327 three broadband stations (MGE, IDO, and THM of the CI network, Fig. [7b](#)) near the railway in the
328 Coachella Valley.) After band-pass filtering the continuous data between 0.75-5 Hz we slant-stacked
329 the envelopes of the continuous seismograms with apparent velocities of plus or minus ~95 km/h
330 (dashed blue and orange lines in Fig. [7c](#)). This procedure detected trains passing through the

331 stationary phase zone (Fig. [7b](#)) both North to South and South to North.

332 - With the catalog in hand, we sorted broad time-windows, those with and without train
333 tremors (large green and red shaded rectangle in Fig. [7c](#) respectively).

334 - In the third step, to analyze the dense nodal array data, we cross-correlated the selected
335 time-windows, (green rectangles in Fig [7c](#)) and [8a\)i](#)), filtered between 3 and 10 Hz and using non-
336 overlapping data segments of 30 min.

337 - We next stacked the cross-correlations, according to their inter-station separations, into
338 distance-binned correlation gathers. We highlight only correlation gathers for the causal part of the
339 correlations (from PFO to CIR), for a time-window centered at the P-wave arrival time. These 30 min
340 correlation gathers were further pruned based on three quality criteria seen in their vespagram,
341 indicators of the different waves' apparent velocities observed in the correlation gathers. ([Davies et](#)
342 [al., 1971](#)). Figure [8a\)ii](#) shows the vespagrams associated with the correlation gathers in the upper
343 panels (Fig. [8a\)i](#)). The three quality criteria were: 1) SNR1, the ratio between the maximum
344 vespagram amplitude in the [0.13-0.2] s/km slowness (inverse of velocity [5-7.5] km/s velocity)
345 window (dashed black rectangle in the leftmost vespagram panel, Fig. [8a\)ii](#) and the root-mean-
346 squared (RMS) amplitude of the rest of the vespagram. 2) SNR2, the ratio between the maximum
347 vespagram amplitude in the [0.13-0.2] s/km slowness \times [4.5-6] s travel-time window (solid black
348 rectangle in the leftmost vespagram panel, Fig. [8a\)ii](#) and the RMS in the rest of the [0.13-0.2] s/km
349 slowness window. 3) MaxAmp, the peak vespagram amplitude in the [0.13-0.2] s/km slowness \times [4.5-
350 6] s travel-time window. SNR1 is used to reject gathers exhibiting phases with apparent velocities
351 different from the expected apparent velocity of a direct P-wave. SNR2 is used to reject gathers
352 exhibiting energetic spurious phases with arrival times that are either too early or too late, even
353 though their apparent velocity is correct. We used the MaxAmp criteria to reject gathers for which the
354 expected P- wave phase is not energetic enough or is too large for a train signal, indicating the

355 detection of an earthquake located in the Fresnel zone (Fig. [8a\)i](#)). For this specific situation, we set
356 thresholds to be sure the conditions $SNR1 \geq 2.5$, $SNR2 \geq 1.5$, and $0.15 \leq MaxAmp \leq 4.0$ were met for a
357 correlation gather to be selected (green boxes in Fig. [8a\)iii](#)). The actual values for SNR1, SNR2 and
358 MaxAmp are shown below each vespagram in Figure [8a](#).

359 - In the last step, we stacked the selected correlation gathers into daily gathers and a reference
360 gather including every selected gather for the whole period of interest (Fig. [8a\)vi](#)). Ultimately, we
361 used less than 20% of the full dataset for the monitoring measurements (Fig. [8b](#)).

362 To quantify the improvement of the signals using the opportune sources approach, we
363 measured the ratio of SNRs between a reference gather computed with all the data (Fig. [8c\)i](#), similar
364 to [Brenguier et al. \(2019\)](#) and the reference gather from selected train windows shown in Fig. [8a\)iv](#).
365 We performed this operation for each waveform in the gathers.

366 The results (Fig. [8c\)ii](#)) show that the opportune source concentration improves the SNR of the
367 P-wave signal by an average of more than 25%. This has important implications for monitoring. As
368 [Silver et al. \(2007\) showed](#), the SNR is the main factor controlling the precision of a time delay
369 measurement between two similar waveforms; such precision scales linearly with the SNR. Therefore,
370 carefully selecting train signals before correlation allows us to improve the precision of the
371 monitoring measurements. The 30 min long segments of continuous data used here to discretize the
372 study period could be decreased and adapted even more closely to the train signals, which in turn
373 should allow even larger SNR improvements. This final process is still ahead of us but the
374 methodology proved effective.

375 -The final step of the workflow was to measure seismic velocities. Different approaches were
376 available. We chose to measure relative time-shift between the seismograms resulting from the slant-
377 stack at 6 km/s of the daily gathers and the reference gather (black and red traces in Fig. [8b](#),
378 respectively). We measured the instantaneous time-delay $\delta t(t)$ between the traces in the 3-10 Hz

379 frequency band using the cross-wavelet transform algorithm of [Mao et al. \(2020\)](#). Although a time-
380 delay was determined for each sample of the waveform, we only show δt values for the direct P-wave.
381 Here, the time-shifts we found are shorter than 0.1% of the propagation time, corresponding to time-
382 shift shorter than 5 ms between the daily and reference seismograms. These time-shifts can be
383 translated into relative velocity changes with the relation $\delta v/v = -\delta t/t$, using the absolute travel-time
384 t of the slant-stacked P-wave. We obtained velocity changes on the order of $\pm 0.1\%$. The meaning of
385 these values is difficult to know because it will take a lot more work to understand the exact
386 sensitivity of the reconstructed P-wave and the different trade-offs among the source and structure
387 sensitivities (see Fig. 4). We plan to estimate 3D spatial sensitivity kernels for these retrieved travel
388 time perturbations and correct for shallow, environmental velocity changes. Thus, we shall see
389 whether we will soon be observing and locating any places where changes in seismic velocity at a few
390 kilometers depth occur within the San Jacinto Fault Zone.

391

392 **7 Discussion and conclusions**

393 We see great opportunity for exploiting any available massive freight train noise recovery to improve
394 crustal imaging and monitoring dramatically. We describe applications to North America but our
395 conclusions have global ramifications, especially in such countries and regions as China, Europe,
396 Japan, and India. All have large freight railway systems, often with high speed passenger lines too. The
397 latter are lighter than freight trains and generate less energetic tremors to be sure, but applications
398 might be found in near-surface environmental or engineering studies.

399 For all its potential, to put heavy freight train noise to work for seismic imaging and
400 monitoring reasons is of course limited to regions near railways. It also requires trains traveling at
401 rather high speed. But generally, this paper presents a workflow for using other and more local
402 sources of cultural noise, including car and truck traffic, wind farms, and natural sources such as surf

403 break or tectonic, volcanic tremor, as opportune sources of useful seismic noise.

404 Although promising, this work poses important, practical challenges that the field must
405 confront. Most important is to improve understanding of the retrieved body and surface waves'
406 spatial sensitivity to crustal structures when combining seismic interferometry with opportune
407 sources. In contrast to actively-controlled and placed sources, measurements of travel times or
408 temporal travel time perturbations using more irregular sources can improve sensitivity not only to
409 the structure between the receivers but contrarily can also blur the overall picture due to interference
410 in areas between the noise source and the receivers. This latter downside may induce misleading
411 interpretations of velocity or velocity change measurements.

412
413 A drawback of examples in this paper is that they used so many sensors (hundreds). Train
414 vibrations cost seismologists nothing but recording them is not yet easy. One solution to overcome
415 these limitations is to find a way to use permanent, single seismic stations instead of costly temporary
416 arrays. One potential initial approach is to deploy dense but temporary seismic arrays around
417 permanent seismic stations. This may help to identify useful phases emanating from noise
418 correlations of opportune sources. A hope is that we learn enough to extract the needed information
419 on a long-term basis with permanent stations alone.

420 One additional major advance would be to couple Distributed Acoustic Sensing data ([Zhan,](#)
421 [2020](#)) to seismic interferometry with opportune sources, as described by [Dou et al. \(2017\)](#) for car
422 traffic and near-surface applications. This indicates potential for reconstructing widespread virtual
423 sources along fiber-optics from correlations of both short- and long-range opportune sources. Success
424 will open the path to many applications including water resource management in the near-surface
425 and earthquake studies at greater depth.

426

427 **8 Data and Resources**

428 The Marathon dataset will be released in June 2021. It will either be hosted online or freely sent on
429 external hard disks upon request via the website for passive seismic techniques for environmentally
430 friendly and cost-efficient mineral exploration (PACIFIC) (<https://www.pacific-h2020.eu>). The San
431 Jacinto array data are available on request to Florent Brenguier. The broadband seismic data used in
432 this study originate from the Southern California Earthquake Center, Caltech. Dataset.
433 doi:10.7909/C3WD3xH1.

434 Open-source codes reproducing [Lavoué et al. \(2020\)](#)'s results are available at [https://gricad-](https://gricad-gitlab.univ-grenoble-alpes.fr/pacific/publications/2020_Lavoue-et-al_SRL_supplemental-material)
435 [gitlab.univ-grenoble-alpes.fr/pacific/publications/2020_Lavoue-et-al_SRL_supplemental-material](https://gricad-gitlab.univ-grenoble-alpes.fr/pacific/publications/2020_Lavoue-et-al_SRL_supplemental-material).

436 Maps are made with Natural Earth. Free vector and raster map data @ naturalearthdata.com.

437 Figure 3 is based on a map published by the US Department of Transportation
438 (<https://railroads.dot.gov/sites/fra.dot.gov/files/inline-images/0209.png>), built from the
439 (confidential) waybill samples 2010 established by the US Surface Transportation Board, which we
440 could unfortunately not access directly.

441 Coachella Valley train video can be found at
442 https://www.youtube.com/watch?v=pE0LYuf7_F8

443

444 **9 Acknowledgments**

445 We acknowledge Nick Arndt, Meysam Rezaeifar, Olivier Coutant, John McBride, Christopher Johnson,
446 Pieter Share and Florian Fuchs, for useful discussions and for supporting data acquisition. This
447 work has received funding from the European Union's Horizon 2020 research and innovation
448 program under grant agreement No 776622 (PACIFIC), from the European Research Council (ERC)
449 under grant agreement No 817803 (FaultScan) and from the Department of Energy (Award DE-
450 SC0016520). We are grateful to the reviewers for their useful comments.

451 This work was performed in part at the University of California Natural Reserve System Boyd
452 Deep Canyon Desert Research Center Reserve DOI: (DOI: 10.21973/N3V66D). We gratefully
453 acknowledge the Cahuilla Band of Mission Indians Reservation for graciously allowing us to deploy
454 instruments on tribal land.

455 **References**

- 456 Aki, K., 1957: Space and time spectra of stationary stochastic waves, with special reference to
457 microtremors. *Bulletin of the Earthquake Research Institute, University of Tokyo*, 35 (3),
458 415–456.
- 459 Bakulin, A., and R. Calvert, 2006: The virtual source method: Theory and case study. *Geophysics*,
460 71 (4), SI139–SI150.
- 461 Bensen, G., M. Ritzwoller, M. Barmin, A. L. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang,
462 2007: Processing seismic ambient noise data to obtain reliable broad-band surface wave
463 dispersion measurements. *Geophysical Journal International*, 169 (3), 1239–1260.
- 464 Boué, P., P. Poli, M. Campillo, H. Pedersen, X. Briand, and P. Roux, 2013: Teleseismic correlations
465 of ambient seismic noise for deep global imaging of the Earth. *Geophysical Journal*
466 *International*, 194 (2), 844–848, doi:10.1093/gji/ggt160.
- 467 Brenguier, F., and Coauthors, 2019: Train traffic as a powerful noise source for monitoring active
468 faults with seismic interferometry. *Geophysical Research Letters*, 46 (16), 9529–9536.
- 469 Brenguier, F., and Coauthors, 2020: Noise-based ballistic wave passive seismic monitoring. part 1:
470 body waves. *Geophysical Journal International*, 221 (1), 683–691.
- 471 Campillo, M., and A. Paul, 2003: Long-range correlations in the diffuse seismic coda. *Science*, 299
472 (5606), 547–549, doi:10.1126/science.1078551.
- 473 Chang, J. P., S. A. L. de Ridder, and B. L. Biondi, 2016: High-frequency Rayleigh-wave tomography
474 using traffic noise from Long Beach, California. *Geophysics*, 81 (2), B43– B53,
475 doi:10.1190/geo2015-0415.1.
- 476 Cheng, Y., Ben-Zion, Y., Brenguier, F., Johnson, C. W., Li, Z., Share, P. E., ... & Vernon, F. (2020). An
477 Automated Method for Developing a Catalog of Small Earthquakes Using Data of a Dense Seismic
478 Array and Nearby Stations. *Seismological Society of America*, 91(5), 2862-2871.

479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502

Claerbout, J. F., 1968: Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, 33 (2), 264–269, doi:10.1190/1.1439927.

Connolly, D., G. Kouroussis, O. Laghrouche, C. Ho, and M. Forde, 2015: Benchmarking railway vibrations — track, vehicle, ground and building effects. *Construction and Building Materials*, 92, 64–81, doi:10.1016/j.conbuildmat.2014.07.042.

Dales, P., and Coauthors, 2020: Virtual Sources of Body Waves from Noise Correlations in a Mineral Exploration Context. *Seismological Research Letters*, 91 (4), 2278–2286, doi:10.1785/0220200023, URL <https://doi.org/10.1785/0220200023>.

Davies, D., E. Kelly, and J. Filson, 1971: Vespa process for analysis of seismic signals. *Nature Physical Science*, 232 (27), 8–13.

Dong, S., J. Sheng, and G. T. Schuster, 2006: Theory and practice of refraction interferometry. *SEG Technical Program Expanded Abstracts 2006*, Society of Exploration Geophysicists, 3021–3025.

Dou, S., and Coauthors, 2017: Distributed acoustic sensing for seismic monitoring of the near surface: A traffic-noise interferometry case study. *Scientific reports*, 7 (1), 1–12.

Douze, E., and S. Laster, 1979: Seismic array noise studies at roosevelt hot springs, Utah geothermal area. *Geophysics*, 44 (9), 1570–1583.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2009: Reflection images from ambient seismic noise. *Geophysics*, 74 (5), A63–A67.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2013: Seismic exploration-scale velocities and structure from ambient seismic noise (> 1 Hz). *Journal of Geophysical Research: Solid Earth*, 118 (8), 4345–4360.

Fink, M., 1997: Time reversed acoustics. *Physics today*, 50 (3), 34–40.

503 Fuchs, F., G. Bokelmann, and the AlpArray Working Group, 2018: Equidistant Spectral Lines in
504 Train Vibrations. *Seismological Research Letters*, 89 (1), 56–66,
505 doi:10.1785/0220170092, URL <https://doi.org/10.1785/0220170092>.

506 Green, D. N., I. D. Bastow, B. Dashwood, and S. E. Nippres, 2017: Characterizing broadband
507 seismic noise in central london. *Seismological Research Letters*, 88 (1), 113–124.

508 Inbal, A., T. Cristea-Platon, J.-P. Ampuero, G. Hillers, D. Agnew, and S. E. Hough, 2018 Sources of
509 Long-Range Anthropogenic Noise in Southern California and Implications for Tectonic
510 Tremor Detection. *Bulletin of the Seismological Society of America*, 108 (6), 3511–3527,
511 doi:10.1785/0120180130, URL <https://doi.org/10.1785/0120180130>.

512 King, S., and A. Curtis, 2012: Suppressing nonphysical reflections in Green’s function estimates
513 using source-receiver interferometry. *Geophysics*, 77 (1), Q15–Q25,
514 doi:10.1190/geo2011-0300.1.

515 Kong, Q., D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade, and P. Gerstoft, 2019: Machine learning
516 in seismology: Turning data into insights. *Seismological Research Letters*, 90 (1), 3–14.

517 Krylov, V., and C. Ferguson, 1994: Calculation of low-frequency ground vibrations from railway
518 trains. *Applied Acoustics*, 42 (3), 199–213, doi:10.1016/0003-682X(94)90109-0.

519 Lavoué, F., O. Coutant, P. Boué, L. Pinzon-Rincon, F. Brenguier, P. Dales, M. Rezaeifar, and C. J. Bean,
520 2020: Understanding seismic waves generated by train traffic via modelling: implications
521 for seismic imaging and monitoring. submitted to *Seismological Research Letters*.

522 Li, C., Z. Li, Z. Peng, C. Zhang, N. Nakata, and T. Sickbert, 2018: Long-Period Long-Duration Events
523 Detected by the IRIS Community Wavefield Demonstration Experiment in Oklahoma:
524 Tremor or Train Signals? *Seismological Research Letters*, 89 (5), 1652–1659,
525 doi:10.1785/0220180081.

526 Mao, S., A. Mordret, M. Campillo, H. Fang, and R. D. van der Hilst, 2020: On the measurement of

527 seismic traveltimes changes in the time–frequency domain with wavelet cross-spectrum
528 analysis. *Geophysical Journal International*, 221 (1), 550–568.

529 Meng, H., Y. Ben-Zion, and C. W. Johnson, 2019: Detection of random noise and anatomy of
530 continuous seismic waveforms in dense array data near Anza California. *Geophysical
531 Journal International*, 219 (3), 1463–1473, doi:10.1093/gji/ggz349.

532 Mikesell, D., K. van Wijk, A. Calvert, and M. Haney, 2009: The virtual refraction: Useful energy in
533 seismic interferometry. *Geophysics*, 74 (3), A13–A17, doi:10.1190/1.3095659.

534 Nakata, N., P. Boué, F. Brenguier, P. Roux, V. Ferrazzini, and M. Campillo, 2016: Body and surface
535 wave reconstruction from seismic noise correlations between arrays at piton de la
536 fournaise volcano. *Geophysical Research Letters*, 43 (3), 1047–1054.

537 Nakata, N., J. P. Chang, J. F. Lawrence, and P. Boué, 2015: Body wave extraction and tomography
538 at long beach, california, with ambient-noise interferometry. *Journal of Geophysical
539 Research: Solid Earth*, 120 (2), 1159–1173, doi:10.1002/2015JB011870.

540 Nakata, N., R. Snieder, T. Tsuji, K. Larner, and T. Matsuoka, 2011: Shear wave imaging from traffic
541 noise using seismic interferometry by cross-coherence. *Geophysics*, 76 (6), SA97–SA106,
542 doi:10.1190/geo2010-0188.1.

543 Olivier, G., F. Brenguier, M. Campillo, R. Lynch, and P. Roux, 2015: Body-wave
544 reconstruction from ambient seismic noise correlations in an underground mine.
545 *Geophysics*, 80 (3), KS11–KS25, doi:10.1190/geo2014-0299.1.

546 Poli, P., M. Campillo, H. Pedersen, L. W. Group, and Coauthors, 2012: Body-wave imaging of earth’s
547 mantle discontinuities from ambient seismic noise. *Science*, 338 (6110), 1063–1065.

548 Quiros, D. A., L. D. Brown, and D. Kim, 2016: Seismic interferometry of railroad induced ground
549 motions: body and surface wave imaging. *Geophysical Journal International*, 205 (1), 301–
550 313, doi:10.1093/gji/ggw033.

551 Riahi, N., and P. Gerstoft, 2015: The seismic traffic footprint: Tracking trains, aircraft, and cars
552 seismically. *Geophysical Research Letters*, 42 (8), 2674–2681, doi:10.1002/
553 2015GL063558.

554 Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux, 2005: Ambient noise cross correlation in free
555 space: Theoretical approach. *The Journal of the Acoustical Society of America*, 117 (1), 79–
556 84.

557 Sánchez-Sesma, F. J., and M. Campillo, 2006: Retrieval of the green’s function from cross
558 correlation: the canonical elastic problem. *Bulletin of the Seismological Society of America*,
559 96 (3), 1182–1191.

560 Schuster, G., J. Yu, J. Sheng, and J. Rickett, 2004: Interferometric/daylight seismic imaging.
561 *Geophysical Journal International*, 157 (2), 838–852.

562 Seydoux, L., N. M. Shapiro, J. de Rosny, F. Brenguier, and M. Landès, 2016: Detecting seismic
563 activity with a covariance matrix analysis of data recorded on seismic arrays. *Geophysical*
564 *Journal International*, 204 (3), 1430–1442.

565 Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller, 2005: High-resolution surface- wave
566 tomography from ambient seismic noise. *Science*, 307 (5715), 1615–1618.

567 Silver, P. G., T. M. Daley, F. Niu, and E. L. Majer, 2007: Active source monitoring of cross- well
568 seismic travel time for stress-induced changes. *Bulletin of the Seismological Society of*
569 *America*, 97 (1B), 281–293.

570 Snieder, R., 2004: Extracting the green’s function from the correlation of coda waves: A based on
571 stationary phase. *Phys. Rev. E*, 69, 046 610, doi:10.1103/PhysRevE. 69.046610.

572 Snieder, R., K. Wapenaar, and K. Larner, 2006: Spurious multiples in seismic interferometry
573 primaries. *Geophysics*, 71 (4), SI111–SI124, doi:10.1190/1.2211507.

574 Takano, T., F. Brenguier, M. Campillo, A. Peltier, and T. Nishimura, 2020: Noise- based passive

575 ballistic wave seismic monitoring on an active volcano. *Geophysical Journal* , 220 (1), 501–
576 507.

577 Tsuji, S., K. Yamaoka, R. Ikuta, T. Kunitomo, T. Watanabe, Y. Yoshida, and A. Katsumata, 2018:
578 Secular and coseismic changes in s-wave velocity detected using across in the tokai . *Earth,*
579 *Planets and Space*, 70 (1), 1–10.

580 Wapenaar, K., 2004: Retrieving the elastodynamic green’s function of an arbitrary
581 inhomogeneous medium by cross correlation. *Physical review letters*, 93 (25), 254 301.

582 Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010a: Tutorial on seismic
583 interferometry: Part 1—basic principles and applications. *Geophysics*, 75 (5), 75A195–
584 75A209.

585 Wapenaar, K., and J. Fokkema, 2006: Green’s function representations for seismic .
586 *Geophysics*, 71 (4), SI33–SI46, doi:10.1190/1.2213955.

587 Wapenaar, K., E. Slob, R. Snieder, and A. Curtis, 2010b: Tutorial on seismic interferometry: 2—
588 underlying theory and new advances. *Geophysics*, 75 (5), 75A211–75A227.

589 Zhan, Z., 2020: Distributed acoustic sensing turns fiber-optic cables into sensitive seismic.
590 *Seismological Research Letters*, 91 (1), 1–15.

591 Zhou, W., and H. Paulssen, 2020: Compaction of the Groningen gas reservoir investigated with
592 train noise. *Geophysical Journal International*, doi:10.1093/gji/ggaa364.

593 **List of Figures**

594 Figure 1: Cartoon showing examples of studies related to train seismic tremors.

595 Figure 2: Top: A train tremor recorded 3 km away from a seismic station in Marathon, Canada.

596 Bottom: Spectrogram showing clear spectral lines oscillating as train speed varies.

597 Figure 3: Regions of potential long-range train tremor detection from the main railway route and

598 annual tonnage information in North America. Colors represent annual freight tonnage, which is an

599 indication of the number of trains travelling on the rail sections (sections with annual tonnage < 10

600 MT/year are ignored). Colored lines are 100-km-thick, which is an indication for the distance from

601 which we may detect individual train tremors (50 km from the railway, see details in the text).

602 Figure 4: Schematic representation of seismic interferometry for opportune sources. (a) A railway

603 surrounding a dense geophone array; an example from the Marathon (Canada) deployment. Different

604 train locations (stars) allow the illumination of the array with different azimuths. Yellow kernels are

605 schematic views of the propagation of diving P waves. (b-d) 3 different scenarios of wave

606 interference: (b and c) leading to a proper measurement of a diving P wave and (d) leading to a

607 spurious or virtual refraction measurement.

608 Figure 5: Chart illustrating the processing steps for opportune sources (in blue) compared to the

609 standard ambient noise correlation workflow (in orange).

610 Figure 6: Case studies of train tremor correlations over a dense array at Marathon/Canada. a) One-

611 minute cross-correlation for a quiet time period. b) Map of the study zone in the north of Marathon,

612 Ontario, Canada. Grey dots are the 1020 seismic stations. The black dashed line is the railroad (CPRS).

613 The red and blue cross are the position studied. c) Train seismic record. d) - e) left: stacked section

614 over 6 trains. d) -e) right: 1-minute beamforming panels for 6 train passages.

615 Figure 7: Detection of train passages in the Coachella Valley. a): Layout of the dense nodal arrays used

616 for the monitoring of the San Jacinto Fault. b) Map of the study area showing the Fresnel zone (yellow

617 ellipse) where train signals contribute coherently to the P-waves in the correlations, travelling
618 between PFO and CIR. The orange areas show the zones of sensitivity to the structure for a wave
619 reconstructed by interferometry between the two arrays. The railway and the main highway are
620 shown in blue. The active tectonic faults are shown in black. The three broadband stations used for
621 building the train catalogue are shown with red and black circles. c) The train catalog is built by slant-
622 stacking the envelopes of the train tremors. The colored rectangles show the location of the Fresnel
623 zone along the railway, green for a time-window with trains, red for a time-window without train and
624 rejected from further analysis.

625 Figure 8: Workflow for monitoring applications. a) i: Cross-correlation computation and correlation
626 gathers construction for every 30 min-long segments of selected continuous data from the dens
627 arrays. a) ii: Vespagrams of the correlation gathers used for the 30-min window selection. The black
628 rectangles in the leftmost panel are used to measure the different selection criteria. a) iii: The three
629 selection criteria associated to each 30-min window. The red boxes are rejected, the green boxes are
630 kept for the next step. a) iv: Stack of the selected 30-min gathers into daily gathers, then the daily
631 gathers into the Reference gather. b) Monitoring results. The bottom histogram shows the number of
632 hours of continuous noise records stacked to obtain the daily correlation gathers. c) i: Stack of every
633 30-min windows without train selection, similar to the section shown in [Brenquier et al. \(2019\)](#). c)ii:
634 Histogram of the signal-to-noise ratio improvement between the Reference correlation gathers
635 without and with train signal selection

636

637

638

639

640

641 Authors:

642 Laura Pinzon-Rincon: laura-alejandra.pinzon-rincon@univ-grenoble-alpes.fr, Université602Grenoble
643 Alpes, ISTerre, CS 40700, 38058 GRENOBLE Cedex 9, France.

644 François Lavoué: francois.lavoue@univ-grenoble-alpes.fr, Université Grenoble Alpes, ISTerre,604CS
645 40700, 38058 GRENOBLE Cedex 9, France

646 Aurélien Mordret: aurelien.mordret@univ-grenoble-alpes.fr, Université Grenoble Alpes,606ISTerre,
647 CS 40700, 38058 GRENOBLE Cedex 9, France

648 Pierre Boué: pierre.boue@univ-grenoble-alpes.fr, Université Grenoble Alpes, ISTerre, CS60840700,
649 38058 GRENOBLE Cedex 9, France

650 Florent Brenguier: florent.brenguier@univ-grenoble-alpes.fr, Université Grenoble Alpes,610ISTerre,
651 CS 40700, 38058 GRENOBLE Cedex 9, France

652 Philippe Dales: pdales@gmail.com, Institute of Mine Seismology 19 Jeanine Street
653 Sudbury612Ontario, P3B 0E5 Canada

654 Yehuda Ben-Zion: benzion@usc.edu, Department of Earth Sciences and Southern California
655 Earthquake Center, University of Southern California, Zumberge Hall, Room 109, USC Earth Sciences,
656 Los Angeles, CA 90089-0740

657 Frank Vernon: flvernon@ucsd.edu, IGPP, Scripps Institution of Oceanography, UC San617Diego, 9500
658 Gilman Dr #0225, La Jolla, CA 92093-0225,CA, USA

659 Christopher J. Bean: chris.bean@dias.ie, Geophysics Section, Dublin Institute for Advanced
660 Studies, 5 Merrion Square North, Dublin 2, D02 Y006, Ireland

661 Daniel Hollis: dan.hollis@sisprobe.com, Sisprobe, IMAG building, office #230, 700 avenue Centrale,
662 38401 Saint Martin d'Hères, FRANCE

663