HAL open science

Characterization with dense array data of seismic sources in the shallow part of the San Jacinto fault zone

Chloé Gradon, Philippe Roux, Ludovic Moreau, Albanne Lecointre, Yehuda Ben Zion

▶ **To cite this version:**

Chloé Gradon, Philippe Roux, Ludovic Moreau, Albanne Lecointre, Yehuda Ben Zion. Characterization with dense array data of seismic sources in the shallow part of the San Jacinto fault zone. Geophysical Journal International, 2021, 224, pp.1133-1140. 10.1093/gji/ggaa411 . insu-03594416

HAL Id: insu-03594416

https://insu.hal.science/insu-03594416

Submitted on 17 Mar 2022

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

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

Distributed under a Creative Commons Attribution 4.0 International License

Geophysical Journal International

Geophys. J. Int. (2021) **224,** 1133–1140
Advance Access publication 2020 August 28
GJI Seismology

doi: 10.1093/gji/ggaa411

Characterization with dense array data of seismic sources in the shallow part of the San Jacinto fault zone

Chloé Gradon,[1] Philippe Roux,[1] Ludovic Moreau,[1] Albanne Lecointre[1] and Yehuda Ben Zion[2]

[1]*Institut des Sciences de la Terre, Université Grenoble Alpes, CNRS UMR* 5275*, Maison des Géosciences,* 1381 *Rue de la Piscine,* 38400 *Saint Martin d'Hères, France. E-mail: chloe.gradon@univ-grenoble-alpes.fr*
[2]*Department of Earth Sciences, University of Southern California, Los Angeles, CA* 90089–0740*, USA*

Accepted 2020 August 21. Received 2020 August 20; in original form 2020 February 5

SUMMARY

We analyse dominant sources identified in a catalogue of more than 156 000 localizations performed using a 26-d data set recorded by a dense array set on the San Jacinto fault near Anza, in California. Events were localized using an array processing technique called Match Field Processing. As for all array processing techniques, the quality of the event position decrease when the events are outside of the array. We thus separate localizations in and outside the array using simple geometrical conditions. We compare the time distribution of the localization to additional data such as meteorological data, day of human activity as well as existing catalogues to determine the nature of the dominant events located using our method. We find that most of the events located outside of the array could be attributed to a surface structure excited by wind. On the other hand, part of the localizations under the array occur during regional earthquakes and could correspond to diffraction on the fault's heterogeneities. The rest of the localizations inside the array could be generated by the fault itself.

Key words: Computational seismology; Earthquake source observations; Wave propagation.

1 INTRODUCTION

Studying the top 500 m of the crust in general, and fault zones in particular, is highly challenging since the shallow material has extreme properties—such as very low *P*- and *S*-velocities (*Vp* and *Vs*), low attenuation *Q* values (Liu *et al.* 2015), and very high *Vp*/*Vs* ratios (Qin *et al.* 2020)—that make it highly susceptible to non-linear behaviour and temporal changes. Local and regional earthquakes, along with rain and temperature changes, produce rock damage (reduction of *Vp*, *Vs*, *Q*) in broad shallow regions around the faults with very large coseismic changes in the top 100 m. To provide detailed understanding of the shallow structure and dynamics of fault zones, temporary dense arrays have been installed in the last few years to enable high-resolution imaging and monitoring studies (Ben-Zion *et al.* 2015, Fig 1). The improved spatial coherency at high frequencies has allowed the use of noise-based tomography to image finer details of the Newport-Inglewood Fault (Lin *et al.* 2013) and the San Jacinto Fault Zone (SJFZ, Hillers *et al.* 2016; Roux *et al.* 2016; Mordret *et al.* 2019; Zigone *et al.* 2019), with scales ranging from hundreds to tens of metres.

Another aspect made possible by continuous recordings at dense arrays is the ability to detect and locate small surface and subsurface events. Taking advantage of spatial coherence on dense seismic arrays, we recently showed that combining Match Field Processing (MFP) with Markov-Chain Monte Carlo (MCMC) sampling provided a large number of new detections in the vicinity of the SJFZ (Gradon *et al.* 2019). The spatial coherence allows us to keep an analytic homogeneous model depending on only three parameters (horizontal positions *x*, *y* and apparent velocity *v*) ensuring high efficiency of the algorithm. The use of MCMC further reduces the computational time compared to grid search approaches and provides statistical information that can be useful when interpreting the data.

Continuous seismic waveforms consist mostly of wind-related signals, air/train/vehicle-traffic events and other natural and anthropogenic sources of ground motion (Riahi & Gerstoft 2016; Inbal *et al.* 2018; Meng & Ben-Zion 2018a,b). These signals are similar to earthquake and tremor waveforms, so they can produce false detection of shallow weak seismic events. To address the challenge of separating subsurface from surface events, the MFP-MCMC localization method used the frequency of the analysis as a 'filter' for separating sources at the surface and sources at depth: surface sources are detected at lower frequencies than deep sources, because the apparent velocity measured at the array increases with depth. In practice, surface sources such as shots and moving vehicles were successfully detected and located by extracting the epicentral position (*x*,*y*) and the apparent velocity *v* of the propagating waves (Gradon *et al.* 2019). In those cases, we obtained a good resolution

© The Author(s) 2020. Published by Oxford University Press on behalf of The Royal Astronomical Society. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

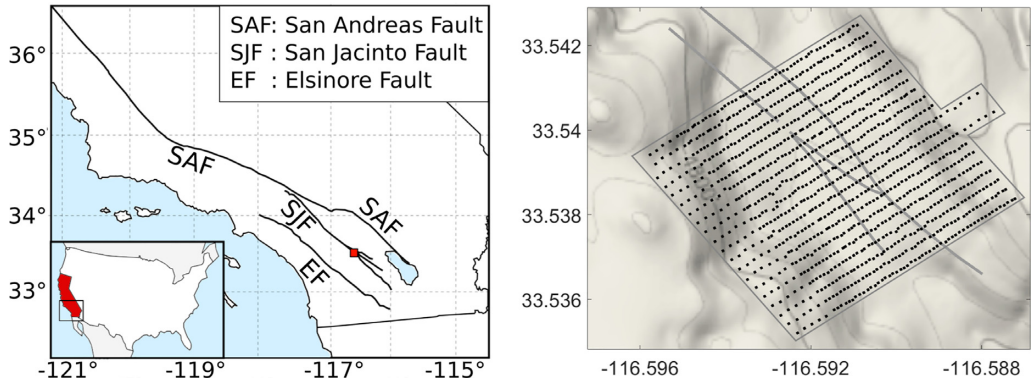

Figure 1. (a) Position of the array on the San Jacinto Fault Zone in Southern California. (b) Geometry of the 1108 sensor array and situation with respect to fault traces (black lines) and local topography (Map data, Google 2017). Each black dot corresponds to a station.

for epicentral coordinates for events inside the array, and velocities that are consistent with previous studies (Mordret *et al.* 2019).

Shallow sources located outside the array were also successfully detected based on the same approach (Gradon *et al.* 2019). However, the use of a homogeneous velocity model to compute the replicas for source localization has limitations, since the trade-off between depth and velocity results in a strong ambiguity on the two parameters. It is therefore essential to identify the dominant sources outside the array to further our analysis. In the last decade, machine learning algorithms trained with labeled data sets of common signals have speed up considerably classification of different parts of continuous waveforms and improve detection of small events (Diersen *et al.* 2011; Reynen & Audet 2017; Chen 2018). However, an *a priori* knowledge on the events characteristics is required to train and/or use the algorithms efficiently.

The goal of this paper is to provide a more comprehensive interpretation of the entire 26-d data set for dominant shallow sources either detected inside or outside the dense array located in the damage structure of the SJFZ. The paper is structured as follows. In Section 2, we detail our event classification algorithm based on the output of the MFP-MCMC algorithm. Sections 3 and 4 give a seismic interpretation to shallow events located either outside or below the array, respectively. Section 5 provide a discussion of the preformed research and useful future studies.

2 EVENT CLASSIFICATION

In our previous study (Gradon *et al.* 2019), we performed a scan of a 26-d long data set recorded on the San Jacinto Fault by a dense array of sensors (Fig. 1). The problem consisted of (i) detecting very shallow events (less than 500 m depth) in poor SNR conditions, (ii) separating between surface-generated events due to anthropogenic activities from microseismic events at shallow depths and (iii) estimating their location. We applied a Match Field Processing technique to achieve this estimation by comparing the data in successive time windows to a simple homogeneous acoustic forward model. This comparison was performed in the frequency domain using the Bartlett operator that consists in the projection of the synthetic field on the cross-spectral density matrix, which contains the auto-correlation and intercorrelation between sensors (Capon 1969; Cros *et al.* 2011; Vandemeulebrouck *et al.* 2013). In our case, we only work with phase information because it is less sensitive to medium heterogeneity.

The MFP technique was developed originally for simpler applications of acoustic waves propagating in the ocean (e.g. Kuperman and Turek 1997; Baggeroer et al. 1993). The synthetic wavefield is usually computed using a fixed medium velocity for a grid of point sources in three dimensions. Applying the method for localization of near-surface sources recorded by seismic arrays around a fault zone is considerably more challenging because the velocity structure of the medium is more complex and a wide variety of noise sources contribute to the recorded ground motion (e.g. Riahi & Gerstoft 2016; Meng & Ben-Zion 2018a,b). The 26-d continuous recording of the data also required an efficient localization tool in terms of computation time. To address these difficulties, we first augmented the method for source localization by using the frequency of study as a 'filter' separating sources at the surface and at depth. Given the dimensions of our array, we determined that using a centre frequency of 4 and 16 Hz would focus the sensitivity of our method to surface sources and deeper sources, respectively. Regarding the heterogeneity of the medium, we computed the synthetic wavefield for a range of apparent velocities as well as source positions. Taking advantage of our 'frequency filter' and of the fact that apparent velocity can be used as a proxy for depth, we kept an analytic homogeneous model depending on only three parameters (X, Y and V_{ap}). We took this cylindrical wave approach instead traditional 3-D back-projection to insure the efficiency of the algorithm. The use of MCMC further reduced the computational time compared to grid search approaches, allowing us to retrieve global maxima and the Bartlett output shape (Gradon *et al.* 2019).

Using the Metropolis–Hastings algorithm and supposing a uniform probability distribution as an *a priori*, we calculated the probability density function (PDF) of the parameters for each time windows. This PDF is represented by a cloud of points providing the most probable epicentral position of the source and corresponding apparent velocity (Fig. 2, Gradon *et al.* 2019). In addition to representing the PDF using the density of points in the cloud, we also include the value of the Bartlett operator output, normalized between 0 and 1, using a colourscale. In the remainder of this paper, the term MFP-MCMC output will refer to both the PDF and the Bartlett operator value.

The previous study demonstrated that the choice of frequency for the MFP-MCMC operator (4 and 16 Hz) constrains the range of depths of the localized sources at the surface and at depth, respectively (Gradon *et al.* 2019). As the present paper focuses on sources at shallow depth, we will only consider the results of the 16 Hz scan. Around 156 000 'shallow depth' localizations are obtained. We take advantage of this large number to perform a statistical approach that helps in identifying the dominant sources present during the 1-month experiment.

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

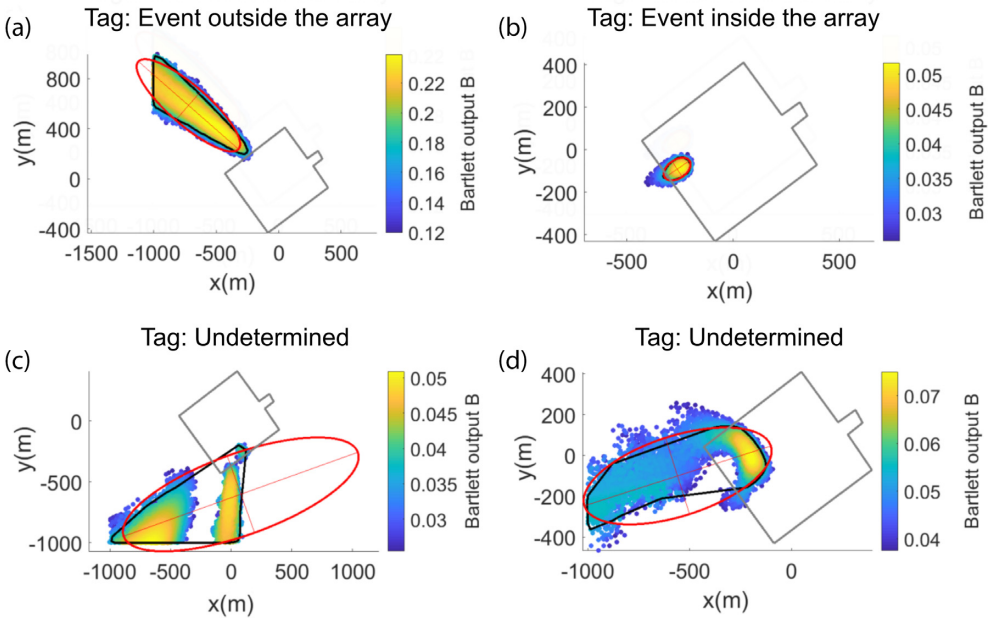

Figure 2. Example of different PDF outputs associated with the MFP-MCMC localization algorithm with their envelopes (black) and fitted ellipses (red). On each panel, the gray line corresponds to the spatial limits of the geophone array. (a) MCMC-MFP output associated with a localization categorized as outside of the array. (b) MCMC-MFP output associated with a localization categorized as inside the array. (c) MCMC-MFP output associated with a localization categorized as undetermined. (d) Other MCMC-MFP output associated with a localization categorized as undetermined.

One main motivation behind this approach is to compensate for the bias introduced by the simple forward model used for localization. This strategy was implemented in an earlier study by Chmiel *et al.* (2019), where MFP was implemented with a minimization algorithm based on the downhill simplex search method. In this case, a localization is manifested by a single point associated with a Bartlett operator value that quantifies the likelihood of a source at this position. A source is localized if its associated Bartlett operator value is higher than a chosen threshold.

In our processing, we sample from the source probability distribution. Thus, we can go one step further and use the shape of our MCMC-MFP output as an additional criterion to select localizations of higher quality. Diffraction laws give us the expected size of the source distribution for a given source position and array geometry (Figs 2a and b). In the case of the Bartlett operator, they state that the 'focal spot' associated to a point source inside the array is of the same order as the half wavelength. In addition, if we consider that only point source are detected, we can expect circular focal spots inside the array and radially elongated focal spots outside of it.

To evaluate the focal spot size automatically, we fit an ellipse to the envelope of the MCMC output at −3 dB, that is taking only into account the points that have an output value for the Bartlett operator greater than half of the maximum. In practice, an event is categorized as outside the array if:

(1) $l_{\text{shortaxis}} < 1/2\ l_{\text{longaxis}}$
(2) $l_{\text{longaxis}} > 200\ \text{m} = 1/4\ l_{\text{ap}}$
(3) $l_{\text{shortaxis}} < 600\ \text{m} = 3/4\ l_{\text{ap}}$
(4) Most of the points of the output are located outside the array,

with $l_{\text{shortaxis}}$ the length of the short axis, l_{longaxis} the length of the long axis and l_{ap} the maximum aperture of the array.

While condition 4 is technically self-sufficient, the other conditions eliminate distorted outputs. Condition 1 and 2 insures that we have an elongated output while condition 3 discards outputs with multiple optima that we cannot interpret.

Table 1. Number of classified localizations.

Shallow depth localizations at 16 Hz	
Nb of localizations inside array	6316 (3 per cent)
Nb of localizations outside array	150 555 (91 per cent)
Nb of rejected localization	10 003 (6 per cent)

Similarly, an event is categorized as inside the array if:

(1) $l_{\text{shortaxis}} > 1/2\ l_{\text{longaxis}}$
(2) $l_{\text{shortaxis}} < 400\ \text{m} = 1/2\ l_{\text{ap}}$
(3) $l_{\text{longaxis}} < 400\ \text{m} = 1/2\ l_{\text{ap}}$
(4) Most of the points of the output are located inside the array.

Fig. 2 illustrates these conditions on four examples of MCMC-MFP outputs, with their envelopes and fitted ellipses. We remove all localizations with outputs that do not meet the conditions presented above (Figs 2c and d). These localizations have misshapen or ambiguous outputs that could correspond to the sampling of side lobes by the MCMC-MFP algorithm or aliasing effects and are therefore considered poor-quality localizations.

The above-mentioned criteria are based on the hypothesis that one point source per time window can be localized. This means that localizations of multiple or distributed sources will be discarded. The conditions of classification are deliberately restrictive because this scheme is very simple and relies on a low number of conditions. Given the fact that we have hundreds of thousands of detections, we decided to use a fast and easily implemented strategy. We choose to keep restrictive conditions to insure the robustness of our classification because the main goal is to find strong evidence of events in the shallow crust. As the remaining localizations are separated into two categories, inside and outside the array, we will study these categories separately.

An examination of the distribution of the localizations (Table 1) reveals that the majority of the shallow depth localizations at 16 Hz corresponds to sources outside of the array.

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

3 LOCALIZATIONS OUTSIDE THE ARRAY

Because localizations of events outside of the array have low radial resolution, we can reduce the outputs from a (X, Y, v) point could to a direction and a velocity. This will result in a simpler representation and facilitate our analysis. We determine a backazimuth from the MCMC-MFP outputs of localizations outside of the array. This is achieved by fitting a line to the PDF output and computing the angle between this line and the north. The localizations can then be represented in polar plots with the apparent velocity as the radial coordinate (Fig. 3a). We see dominant localizations around 309° with apparent velocities ranging from 3500 to 5000 m s^{-1}. This direction coincides with the strike direction of the SJFZ (blue cones in Fig. 3a).

Because apparent velocity can be used as a proxy for depth, we can hypothesize that the dominant sources outside the array, with backazimuths centred on 309°, are good candidates for subsurface sources. Their apparent velocities, ranging from 3500 to 5000 m s^{-1} indicate waves coming from depth, but still traveling in the subsurface. However, the position of the sources outside the array could also mean that these waves are diving body waves emitted from sources at the surface. The use of the 16 Hz frequency does filter surface sources inside the array, as they emit waves at lower apparent velocities, but diving P waves emitted by sources far from the array would have higher velocity and could be measured at such frequency.

Fig. 3(b) shows the analytical incidence angle, or tilt angle, associated with apparent velocities ranging from 1000 to 6500 m s^{-1}. We derive this incidence angle using Snell's law and analytical ray tracing for propagation in a homogeneous gradient model (Meng & Ben-Zion 2018a). In practice, we use an average 1-D gradient velocity model (shown as black line in Fig. 3c) obtained from a shallow velocity model inverted by Mordret *et al.* (2019) using ambient noise recorded at the array (background colours, Fig. 3c).

Apparent velocities associated with the dominant source correspond to incidence angles θ_i ranging from 10° to 13° (yellow area in Fig. 3b). Using this information and our model, we compute the wave paths associated with the dominant source. For apparent velocities ranging from 3500 to 5000 m s^{-1}, the rays reach the surface 2.2–3.9 km northwest of the array. We represented the corresponding area in Fig. 4. It is located on a residential sector that is a potential source of surface noise. Road 371 is also contained in the area of interest. The signal could also be produced by wind interacting with tall structures present along this direction (Johnson *et al.* 2019). We note the presence of a monopole cell tower 3 km away from the centre of the array that could be a good candidate for such a source. Indeed, a 1-D simulation shows that the 3rd flexural mode of this tower corresponds to a resonant frequency at 16 Hz.

We examine the variations of the cumulative localizations with time to find information on the nature of the source. If a source is anthropogenic or due to traffic, we expect a variation of the localizations during the week, with a decrease in activity on Sundays, for example. Fig. 5(a) shows the localizations stacked in 15-min bins for the duration of the experiment (black curve). An average was computed using a moving window of 4.5-hr length (blue). Sundays are highlighted in grey. In Fig. 5(b), we represent the spectrogram of the activity. It exhibits two main components at 12-hr and 24-hr periods. However, there are no clear correlations between the activity of the dominant source and anthropogenic activities.

We also compare the average daily variations of localizations to weather data measured at a station 3 km northeast of the array. To remove potential long-term trends prior to averaging, the number of hourly localizations is normalized by the number of daily localizations. Thus, we obtain an hourly percentage of localization for each day. These hourly percentages are averaged and represented in green in Fig. 6. For the weather data, we average the variation relative to the daily mean. As expected, the temperature and wind gust velocity show a daily cycle, which does not directly correlates with the 12-hr periodicity of the MCMC-MFP localizations (Fig. 6). However, the maxima of MCMC-MFP detections, at sunset and sunrise respectively, correspond to strong variations of temperature and of the wind gusts velocity.

4 LOCALIZATIONS INSIDE THE ARRAY

We also obtain a significant number (∼6000) of localizations under the array. Figs 7(a) and (b) show the stack of the localizations in 10 m × 10 m bins for two apparent velocity ranges. For the sake of clarity, we only stack the maximum of the MFP output instead of the entire PDF distribution.

For lower apparent velocities, which correspond to shallower depth (Gradon *et al.* 2019), the dominant cluster of source is 200-m distant from the geological fault trace (Fig. 7a). Using the 1-D gradient model introduced in the previous section (Fig. 3), these sources are located at depth between 150 and 390 m. This is consistent with the velocity model from Mordret *et al.* (2019), which reveals a complex structure ranging from 50 to 400 m deep, at this location, with alternating low and high velocities anomalies (Figs 7c–e). We hypothesize that this complex structure yields localized seismic activity. At higher apparent velocities (Fig. 7b), sources are clustering closer to the geological fault trace where microseismic activity is also expected.

Another interpretation would be that we localize scatterers instead of sources. In this case, far-field waves diffracted by Fault-related heterogeneities under the array would trigger a localization by the algorithm. To investigate this further, we compare the time distribution of our localization to expected first arrival times computed from a regional catalogue from Ross *et al.* (2019), and to picks made on the data from a borehole located under the array.

We first computed theoretical arrival times at the array for the regional catalogue from Ross *et al.* (2019) using a mean velocity of 6000 m s^{-1} and compared them to our catalogue. We also took into account the magnitude and distance of each earthquake to remove manually the events that were too far away to be detected. Only 50 of the localizations occur less than 2 s after the expected first arrival time at the array (Fig. 8a). For longer time delays, we verified that our detections were contained within the coda of the earthquake. To do so, we used data from a high quality 3C borehole station located under the array to extract the time signals corresponding to the earthquakes in the regional catalogue. If one of our detection: (1) occur less than 2 s after a regional earthquake first arrival or (2) is contained in the coda of an earthquake visible on the borehole data, we consider it to be correlated with an earthquake. Using the borehole data, we were also able to find correlation between some of our detection and additional earthquakes that were not in the catalogue. Some of the additional earthquakes were more than 30 s long and could be global earthquakes that were not included in the regional catalogue.

Over the 26 d, around 50 per cent of the MCMC-MFP localizations below the array can be correlated to regional earthquakes outside of the array. The amount of MCMC-MFP localizations

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

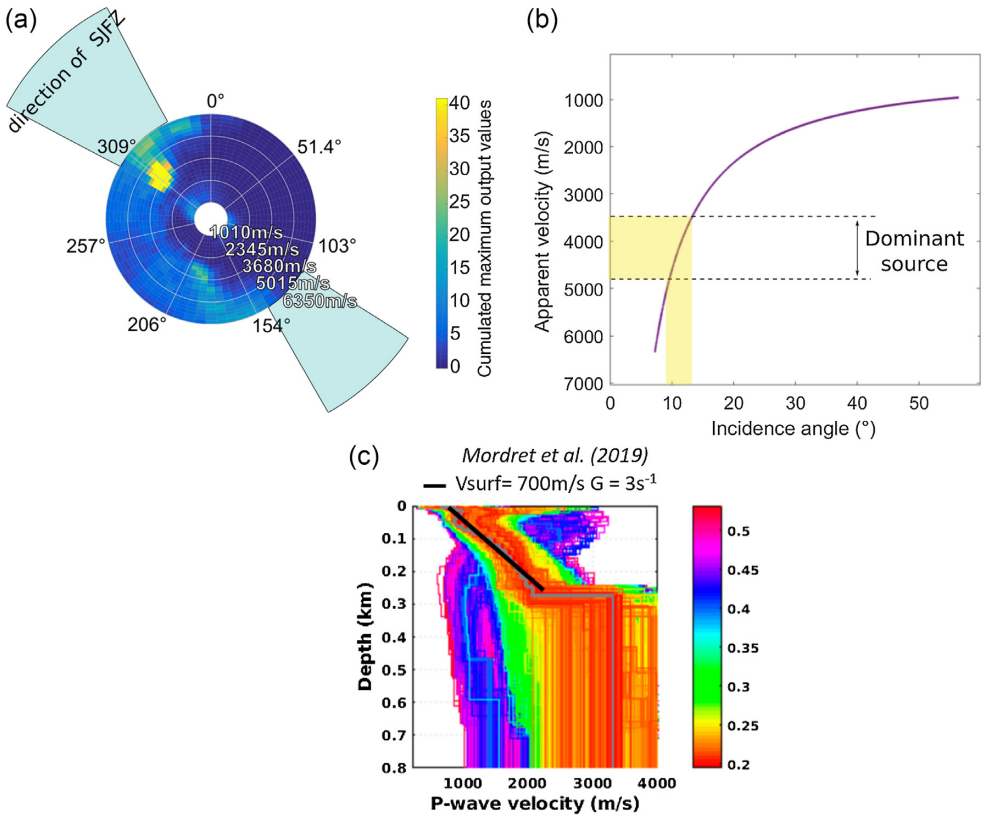

Figure 3. (a) Stacked localizations according to the apparent velocity and backazimuth at 16 Hz, weighted by the value of the maximum output for each localization. The blue cones indicate the direction of the fault. (b) Apparent velocity according to incidence angle computed for an average 1-D gradient velocity model. The yellow area corresponds to the incidence angle and apparent velocities of the dominant source. (c) Background: 1-D velocity model inverted by Mordret *et al.* (2019). Black line: average 1-D gradient velocity model used for this study.

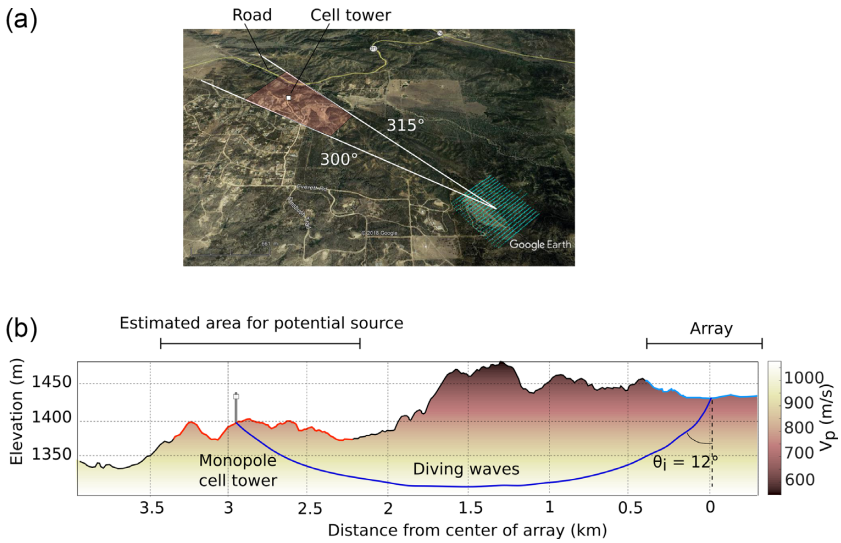

Figure 4. (a) Superposition of the source direction on the Google earth view. The red area correspond to the point of origin for *P* waves estimated for apparent velocities ranging from 3500 to 5000 m s^{-1} using the 1-D gradient velocity profile shown Fig. 3(c). (b) Bathymetry along the source direction. The blue area corresponds to the array position. The red area correspond to the point of origin for *P* waves. The *P*-wave velocity profile is shown in the background.

during regional earthquakes varies visibly from day to day ranging from 90 to 20 per cent of the total daily localizations (Fig. 8b). The days with low percentage (136, 143 and 148) have the highest number of localizations. This could be due to the presence of additional sources. Only half of our localizations can be linked to seismic activity outside of the array. The other half could either be related to undetected regional events or to the fault activity under the array.

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

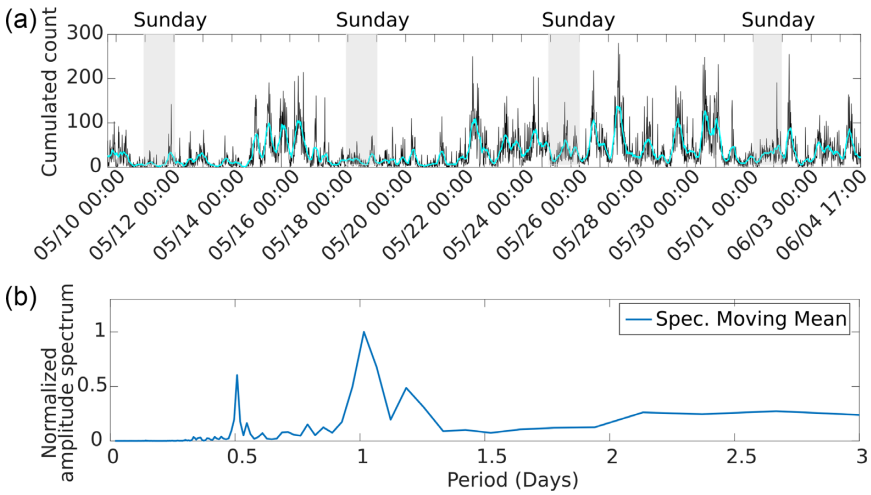

Figure 5. (a) Time variations of localization counts for backazimuths around 309° and apparent velocities around 3750 m s^{-1}, according to local time for 26 successive days. Black line: Time variations of the localization counts normalized by the daily number of localization. Blue line: moving mean over 4.5-hr windows (b) Spectrum according to period (in days) of the moving mean of normalized localization counts.

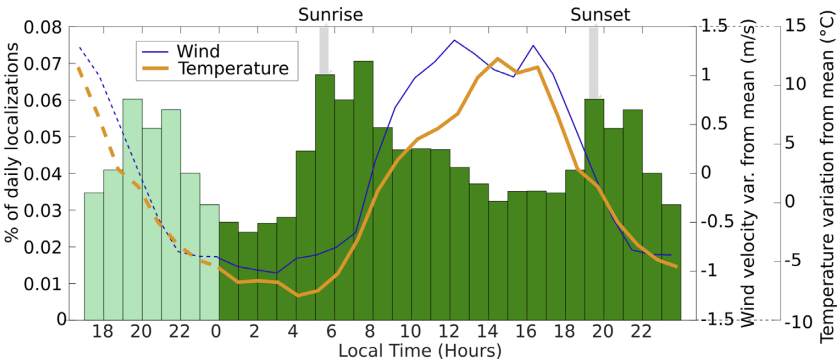

Figure 6. Dark green bars: Average distribution of localizations with backazimuths around 309° and apparent velocities around 3750 m s^{-1}, stacked over 24 hr according to local time. Light green bars: duplicate of the stack between 17:00 and 00:00, used to highlight the periodicity. Orange line: Temperature variations relative to the daily mean averaged over 26 d. Dark blue line: Wind gust velocity variations relative to the daily mean averaged over 26 d.

5 DISCUSSION

The results of this work highlight the high number and diversity of sources that produce packets of seismic energy that are included in continuous recorded waveform. Localizing sources at shallow depth in a fault zone environment raises several challenges. Shallow sources are expected to have weak energy, high frequency content and short duration (Kwiatek & Ben-Zion 2014). These characteristics result in signals with poor SNR. In the SJFZ, weak seismic events overlap with remote ongoing seismicity, which is dominated by surface waves (Roux *et al.* 2016).

The large number of shallow detections and localizations (more than 156 000 in 26 d) provided by the dense array deployment at the SJFZ has significant implications on the rheology and dynamics of the top 500 m of the crust, because at least a fraction of the localizations are likely to be shallow depths events in the vicinity of faults generating phase-coherent wave fronts with high apparent velocities. On the other hand, we observe that more than 90 per cent of the events were located outside of the array and may find their origin in wind gust excitation of a tower locater 3 km away from the centre of the array and other obstacles above the surface. This underlines the need for caution when interpreting localizations obtained from continuous records processing (Inbal *et al.* 2018; Meng & Ben-Zion 2018b; Johnson *et al.* 2019). The ∼6300 microseismic events that were located beneath the array originate from identified weakness zones in the SJFZ. They seem to have two origins that are difficult to distinguish from each other. Half of them occur close in time to regional events and may simply correspond to waves scattered by various structures in the SJFZ below the array. The other half do not seem connected to earthquakes in any catalogue and may be due to genuine failure process occurring at shallow depth along the SJFZ.

For these last set of events, it would be interesting to estimate the magnitude of the sources and examine if the frequency-size distribution of events in the top crust follows the Gutenberg–Richter statistics. Unfortunately, most of these microseismic events are buried within ambient seismic noise and can only be extracted through phase-matching techniques as performed with the MFP-MCMC algorithm. The magnitude of these events is then very difficult to establish with satisfactory confidence interval. Taking advantage of the array gain provided by the dense array can help provide relative magnitude estimates. This type of analysis may be the subject of a future study.

Another continuation of this work could be the inclusion of machine learning to our possessing. In order to process the high number of localizations in our catalogue, we only used a very simple classification method based on the shape of the MCMC-MFP outputs. Using machine learning would allow us to use conditions that are

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

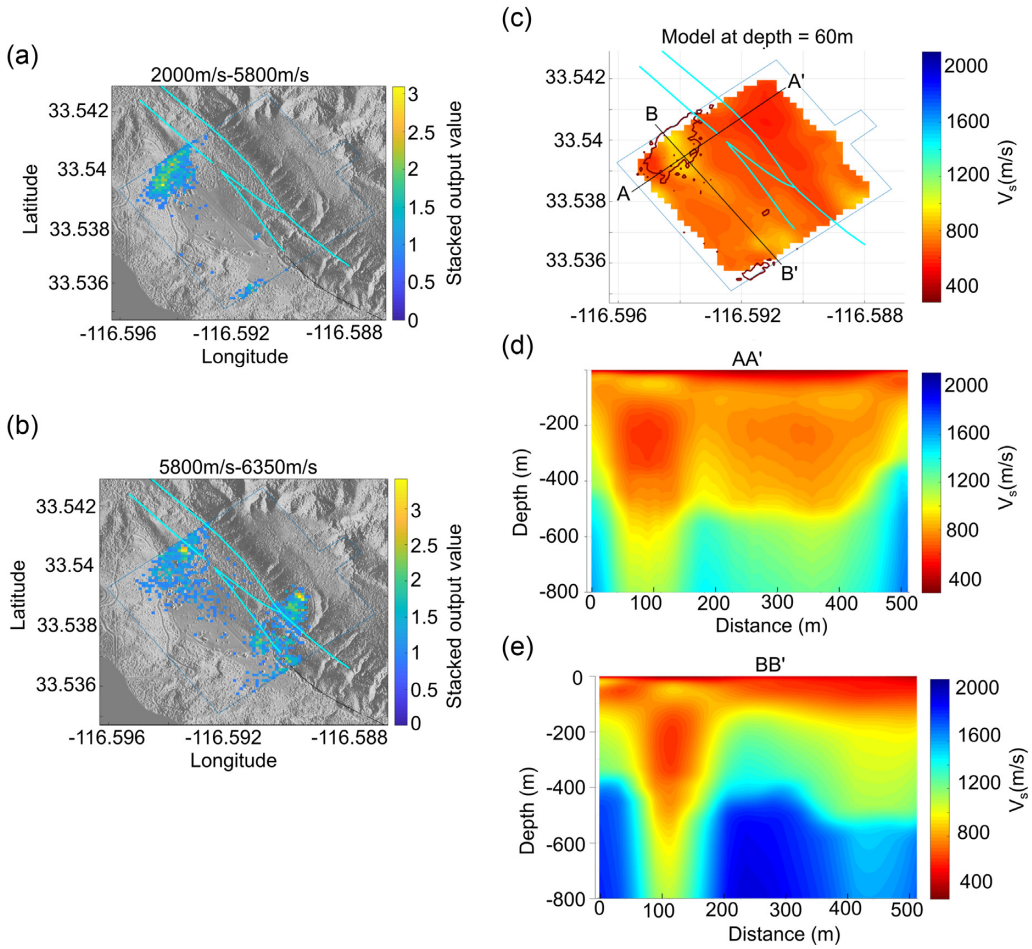

Figure 7. Cumulated outputs of localizations under the array at 16 Hz. The maximum of the outputs are stacked into 10 m × 10 m bins and summed over different velocity ranges (a) Results at 16Hz summed for apparent velocities between 2000 and 5800 m s^{-1}. (b) Results at 16 Hz summed for apparent velocities between 5800 and 6350 m s^{-1}. (c) Slice of the topography model from Mordret *et al.* (2019) at depth 60 m. The red line corresponds to the outline of cumulated localization between 2000 and 5800 m s^{-1}. (e) Vertical slice of the model along plane AA′ (c). (d) Vertical slice of the model along plane BB′ (c).

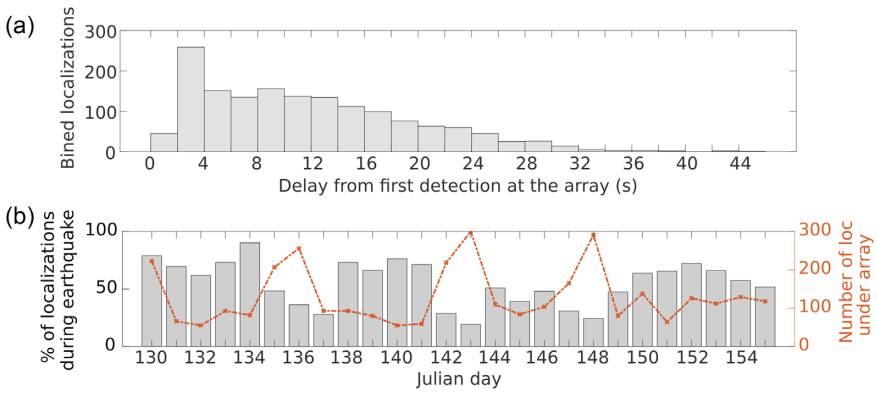

Figure 8. (a) Time delay between observed or estimated first arrival of regional earthquakes and time of MCMC-MFP localization under the array. The first arrival are either picked on borehole data or estimated from the catalogue from Ross *et al.* (2019). (b) Daily percentage of MCMC-MFP localization during regional earthquakes (grey) and total number of MCMC-MFP localizations under the array (orange).

more complex and increase the number of classes. Another issue with our processing is the fact that we reduce our signal to short time windows (0.25 s), with several windows often covering a single event. Using our catalogue, we could reprocess the data to extract shorter or longer time windows containing the signal associated to each event. These adjustable time windows could then be used as *a priori* information to detect new events using machine learning techniques or template matching.

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022

ACKNOWLEDGEMENTS

ISTerre is part of Labex OSUG@2020. YBZ acknowledges support from the National Science Foundation (grant EAR-1818589) and the Department of Energy (award DE-SC0016520).

REFERENCES

Ben-Zion, Y. *et al.*, 2015. Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone, *Geophys. J. Int.,* **202**(1), 370–380.

Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proc. IEEE,* **57**(8), 1408–1418.

Chen, Y., 2018. Automatic microseismic event picking via unsupervised machine learning, *Geophys. J. Int.,* **212**(1), 88–102.

Chmiel, M., Roux, P. & Bardainne, T., 2019. High-sensitivity microseismic monitoring: automatic detection and localization of subsurface noise sources using matched-field processing and dense patch arrays, *Geophysics,* **84**(6), doi:10.1190/geo2018-0537.1.

Cros, E., Roux, P., Vandemeulebrouck, J. & Kedar, S., 2011. Locating hydrothermal acoustic sources at old faithful geyser using matched field processing, *Geophys. J. Int.,* **187**(1), 385–393.

Diersen, S., Lee, E.J., Spears, D., Chen, P. & Wang, L., 2011. Classification of seismic windows using artificial neural networks, *Proc. Comp. Sci.,* **4,** 1572–1581.

Gradon, C., Moreau, L., Roux, P. & Ben-Zion, Y., 2019. Analysis of surface and seismic sources in dense array data with match field processing and Markov chain Monte Carlo sampling, *Geophys. J. Int.,* **218**(2), 1044–1056.

Hillers, G., Roux, P., Campillo, M. & Ben-Zion, Y., 2016. Focal spot imaging based on zero lag cross-correlation amplitude fields: application to dense array data at the San Jacinto fault zone, *J. geophys. Res.,* **121**(11), 8048–8067.

Inbal, A., Cristea-Platon, T., Ampuero, J.-.P., Hillers, G., Agnew, D. & Hough, SE., 2018. Sources of long-range anthropogenic noise in Southern California and implications for tectonic tremor detection, *Bull. seism. Soc. Am.,* 1–20, doi:10.1785/0120180130.

Johnson, C.W., Meng, H., Vernon, F. & Ben-Zion, Y., 2019. Characteristics of ground motion generated by wind interaction with trees, structures, and other surface obstacles, *J. geophys. Res.,* **124**(8), 8519–8539.

Kwiatek, G. & Ben-Zion, Y., 2014. Theoretical limits on detection and analysis of small earthquakes, *J. geophys. Res.,* **119,** 8132–8153.

Lin, F.-C., Li, D., Clayton, R.W. & Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array, *Geophysics,* **78**(4), Q45–56.

Liu, X., Ben-Zion, Y. & Zigone, D., 2015. Extracting seismic attenuation coefficients from cross-correlations of ambient noise at linear triplets of stations, *Geophys. J. Int.,* **203**(2), 1149–1163.

Meng, H. & Ben-Zion, Y., 2018a. Detection of small earthquakes with dense array data: example from the San Jacinto fault zone, Southern California, *Geophys. J. Int.,* **212**(1), 442–457.

Meng, H. & Ben-Zion, Y., 2018b. Characteristics of airplanes and helicopters recorded by a dense seismic array near Anza California, *J. geophys. Res.,* **123**(6), 4783–4797.

Mordret, A., Roux, P., Boué, P. & Ben-Zion, Y., 2019. Shallow three-dimensional structure of the San Jacinto fault zone revealed from ambient noise imaging with a dense seismic array, *Geophys. J. Int.,* **216**(2), 896–905.

Qin, L., Ben-Zion, Y., Bonilla, L.F. & Steidl, J.H., 2020. Imaging and monitoring temporal changes of shallow seismic velocities at the Garner Valley near Anza, California, following the M7.2 2010 El Mayor-Cucapah earthquake, *J. geophys. Res.,* **125**(1), doi:10.1029/2019jb018070.

Reynen, A. & Audet, P., 2017. Supervised machine learning on a network scale: application to seismic event classification and detection, *Geophys. J. Int.,* **210**(3), 1394–1409.

Riahi, N. & Gerstoft, P., 2016. Locating sources in a dense array through network-based clustering, in *2016 Information Theory and Applications Workshop (ITA),* 1–8. IEEE. doi:10.1109/ITA.2016.7888149.

Ross, Z.E., Trugman, D.T., Hauksson, E. & Shearer, P.M., 2019. Searching for hidden earthquakes in Southern California, *Science,* **364**(6442), 767–771.

Roux, P., Moreau, L., Lecointre, A., Hillers, G., Campillo, M., Ben-Zion, Y., Zigone, D. & Vernon, F., 2016. A methodological approach towards high-resolution surface wave imaging of the San Jacinto fault zone using ambient-noise recordings at a spatially dense array, *Geophys. J. Int.,* **206**(2), 980–992.

Vandemeulebrouck, J., Roux, P. & Cros, E., 2013. The plumbing of old faithful geyser revealed by hydrothermal tremor, *Geophys. Res. Lett.,* **40**(10), 1989–1993.

Zigone, D., Ben-Zion, Y., Lehujeur, M., Campillo, M., Hillers, G. & Vernon, F.L., 2019. Imaging subsurface structures in the San Jacinto fault zone with high-frequency noise recorded by dense linear arrays, *Geophys. J. Int.,* **217**(2), 879–893.

Kuperman, W. A. & Turek, G., 1997. Matched Field Acoustics, *Journal of Engineering and Applied Science,* **11**(8):141–148, doi:10.1006/mssp.1996.0066.

Baggeroer, A. B., Kuperman, W. A. & Mikhalevsky, P. N., 1993. An Overview of Matched Field Methods in Ocean Acoustics, *IEEE Journal of Oceanic Engineering,* **18,** 2(8):401–424, doi:10.1109/48.262292.

Downloaded from https://academic.oup.com/gji/article/224/2/1133/5898672 by guest on 17 March 2022