Investigating the influence of earthquake source complexity on back-projection images using convolutional neural networks

Marina Corradini, I Mcbrearty, D Trugman, Claudio Satriano, P Johnson, Pascal Bernard

Geophysical Journal International

Geophys. J. Int. (2022) **229,** 1824–1839
Advance Access publication 2022 January 27
GJI Seismology

https://doi.org/10.1093/gji/ggac026

Investigating the influence of earthquake source complexity on back-projection images using convolutional neural networks

M. Corradini,[1,4] I.W. McBrearty,[2] D.T. Trugman,[3] C. Satriano,[4] P.A. Johnson[5] and P. Bernard[4]

[1]*European-Mediterranean Seismological Centre, Bruyères le Châtel 91297 Arpajon Cedex, France*
[2]*Department of Geophysics, Stanford University, Stanford, CA 94305, USA*
[3]*Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, 2305 Speedway Stop C1160 Austin, TX 78712-1692, USA*
[4]*Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005, France. E-mail: marina.corradini@emsc-csem.org*
[5]*Geophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA .*

Received 2022 January 14; in original form 2021 July 9

SUMMARY

The retrieval of earthquake finite-fault kinematic parameters after the occurrence of an earthquake is a crucial task in observational seismology. Routinely used source inversion techniques are challenged by limited data coverage and computational effort, and are subject to a variety of assumptions and constraints that restrict the range of possible solutions. Back-projection (BP) imaging techniques do not need prior knowledge of the rupture extent and propagation, and can track the high-frequency (HF) radiation emitted during the rupture process. While classic source inversion methods work at lower frequencies and return an image of the slip over the fault, the BP method highlights fault areas radiating HF seismic energy. Patterns in the HF radiation are attributable to the spatial and temporal complexity of the rupture process (e.g. slip heterogeneities, changes in rupture speed). However, the quantitative link between the BP image of an earthquake and its rupture kinematics remains unclear. Our work aims at reducing the gap between the theoretical studies on the generation of HF radiation due to earthquake complexity and the observation of HF emissions in BP images. To do so, we proceed in two stages, in each case analysing synthetic rupture scenarios where the rupture process is fully known. We first investigate the influence that spatial heterogeneities in slip and rupture velocity have on the rupture process and its radiated wave field using the BP technique. We simulate two different rupture processes using a 1-D line source model: a homogeneous process, where the kinematic parameters are constant along the line, and a heterogeneous process, where we introduce a central segment along the line that has a step change in kinematics. For each rupture model, we calculate synthetic seismograms at three teleseismic arrays and apply the BP technique to reveal how HF emissions are influenced by the three kinematic parameters controlling the synthetic model: the rise time, final slip and rupture velocity. Our results show that the HF peaks retrieved from BP analysis are better associated with space–time heterogeneities of slip acceleration. We then build on these findings by testing whether one can retrieve the kinematic rupture parameters along the fault using information from the BP image alone. We apply a machine learning, convolutional neural network (CNN) approach to the BP images of a large set of simulated 1-D rupture processes to assess the ability of the network to retrieve, from the progression of HF emissions in space and time, the kinematic parameters of the rupture. These rupture simulations include along-strike heterogeneities whose size is variable and within which the parameters of rise-time, final slip and rupture velocity change from the surrounding rupture. We show that the CNN trained on 40 000 pairs of BP images and kinematic parameters returns excellent predictions of the rise time and the rupture velocity along the fault, as well as good predictions of the central location and length of the heterogeneous segment. Our results also show that the network is insensitive towards the final slip value, as expected from theoretical results.

Key words: Neural networks, fuzzy logic; Numerical modelling; Computational seismology; Earthquake source observations.

1 INTRODUCTION

Characterizing earthquake finite-fault parameters, such as the rupture extent, rupture velocity and the spatio-temporal distribution of the slip along the fault, are all fundamental to achieving a better understanding of earthquake dynamics. Earthquake parameters are often estimated through slip inversion techniques that rely on *a priori* assumptions on the fault geometry and the rupture mechanism. Such techniques have succeeded to image the finite-fault parameters in a relatively low-frequency range ($f < 2.0$–3.0 Hz, e.g. Zeng *et al.* 1993; Mai *et al.* 2016). However, if we are interested in uncovering fine-scale, detailed structure in the rupture process, high-frequency (HF) waveform data must be taken into account (e.g. Mendoza & Hartzell 1988; Ide 1999).

The complexity of the source manifests itself in terms of short-period seismic waves, at frequencies higher than the corner frequency of far field waveform spectra f_c (e.g. Madariaga 1977; Spudich & Frazer 1984; Ruiz *et al.* 2011) which is controlled by the source duration. At these higher frequencies, the classic inversion techniques are no longer adequate, both because of computational limitations and by our lack of knowledge of the Earth's structure at those frequencies. Hence, high-resolution imaging techniques, such as the back-projection (BP, Ishii *et al.* 2005) method, have become prominent means to unravel aspects of the rupture complementary to the ones supplied by classic inversion. When applied to teleseismic body waves, the BP method takes advantage of the coherence of HF waveforms among ground motion signals recorded at nearby stations. By time reversing each seismogram by the theoretical traveltimes calculated at each station of a given seismic array, the BP method reconstructs a proxy to the image of seismic emissions radiated during the rupture process. This technique requires minor *a priori* constraints and bypasses the procedure of inverting for Earth structure (e.g. Kiser & Ishii 2017). Numerous observational studies in recent years have demonstrated the ability of BP to illuminate the concentration of HF emission sources excited during earthquake rupturing (e.g. Walker & Shearer 2009; Xu *et al.* 2009; Zhang & Ge 2010; Koper *et al.* 2011; Meng *et al.* 2011; Lay *et al.* 2012; Satriano *et al.* 2014; Vallée & Satriano 2014; Grandin *et al.* 2015). From a theoretical perspective, studies have long attributed the generation of HF radiation during earthquake faulting to small-scale variations or roughness in final slip, slip velocity, or rupture velocity—that is sources of local acceleration and deceleration—over the fault plane (Madariaga 1977; Andrews 1981; Herrero & Bernard 1994; Somerville *et al.* 1999; Ruiz *et al.* 2011). HF radiation enlightened by BP images of large earthquakes could, therefore, help constrain the variability of parameters controlling the rupture process. However, unanimous consent on how the BP image relates to earthquake parameters has not yet been achieved. Specifically, Ishii *et al.* (2005) suggested that the BP image of the earthquake is related to the radiated seismic energy. Both works from Yao *et al.* (2012) and Fukahata *et al.* (2014) point out that it is ambiguous whether the BP image is more likely representing the slip velocity or the slip acceleration. In addition, research from Yin & Denolle (2019) enhances the relationship between BP images and earthquake kinematics by showing that the linear formulation of the BP method better correlates with the true kinematic source properties.

Our work has, therefore, a dual purpose. In the first part of our study, we aim to investigate the link between coherent images of the rupture process and the mechanism of HF generation and, therefore, improve our understanding of what the BP image truly represents. To accomplish this first objective, we simulate two rupture processes using a 1-D line source model (Lancieri & Zollo 2009). This choice allows us to ease the computational effort usually required for a 2-D source model and describe the complex earthquake rupture process in a simplified manner that seeks to represent its key elements by three kinematic parameters: the rise time, the final slip, and the rupture velocity. The two rupture processes we study are:

(1) A homogeneous rupture model, where the kinematic parameters (rise time, final slip and rupture velocity) are constant along the line.

(2) A heterogeneous rupture model, where a segment of 30 km length is introduced in the middle of the fault; within this segment, the rise-time value assumes a different value with respect to the surrounding rupture, whereas the final slip and the rupture velocity remain uniform.

We simulate these two ruptures, generate the synthetic ground motion signals, apply the BP technique, and finally compare the resulting images with the originating rupture model. Because the moment rate and slip rate functions are proportional, using a simplified source enables us to understand the link between the generation of HF radiation and the complexity of the moment rate function. Using a 1-D source allows us to represent the BP image of the rupture conveniently as a 2-D map, where we can easily follow the HF progression in space and time, drawing the attention to the mechanism behind the generation of HF radiation from spatial heterogeneities in the kinematic rupture parameters. Reconstructing the rupture process at three complementary arrays also allows us to better investigate how different arrays can enlighten different aspects of the same rupture process.

After developing a conceptual understanding of the relation between the BP image and the variability of the parameters describing the rupture process, in the second part of our study, we work toward extracting information on the earthquake kinematic properties from the HF radiation enhanced by the BP analysis of earthquakes. To accomplish this second objective, we take a data-driven approach: we again parametrize the system as a 1-D line source model and we simulate 40 000 rupture scenarios with different characterizing traits determined by different choices in the values of the kinematic parameters. We then adopt a convolutional neural network (CNN) approach to look for the statistical link between the BP images and the rupture kinematic parameters, exploring the role of input and target parameters on the accuracy of the CNN predictions. In particular, we aim to assess which kinematic characteristics of the rupture process, as well as the location and the extent of a spatial heterogeneity—when present—can be reliably determined from the BP image of the seismic event.

2 METHODOLOGY I: THE LINE SOURCE AND ITS SYNTHETIC GROUND MOTION SIGNALS

2.1 The line source model

The line source is an intermediate approach between the point source and the extended source model (Lancieri & Zollo 2009). Its advantage lies in the ability to reproduce the typical 2-D source directivity effects on the signal while requiring a lower computational effort, and limiting the degrees of freedom. Use of this model is most appropriate for earthquakes with aspect ratios (fault length divided by width) well in excess of one, such as for large strike slip earthquakes and great subduction zone earthquakes. Our source model is built by placing a series of equally spaced point sources along a line (with total length equal to the fault length). These points are set at the hypocentral depth of the event and are distributed along the strike direction of the fault. The point sources begin to slip at different subsequent activation times, related to the rupture propagation velocity along the line. Once crossed by the rupture front, each point source slips following a ramp function whose duration is the rise time. We describe the line source model with the following parameters:

(i) The hypocentral coordinates of the seismic event.
(ii) The strike direction of the fault.
(iii) The rupture length L.
(iv) The rupture velocity V_r along the line.
(v) The discretization along the line ΔL.
(vi) The rise time t_r, or the duration of the dislocation, for each point source.
(vii) The final slip value s_f, reached by each point source at the end of the dislocation.

Under our application of the line source approximation, each point has the same focal mechanism (slip vector and fault orientation) and the total seismic moment of the full rupture is given by the sum of the seismic moments of each point source. An estimation of the fault length (L) and width (W) is derived from the event magnitude using the Wells & Coppersmith (1994) empirical relationships. To avoid space–time aliasing, the contributions emitted by each elementary source must overlap in time; i.e. the duration of each elementary source (the rise time) must be greater than the time the rupture takes to propagate to the next source (here called τ), as discussed in Lancieri & Zollo (2009). To properly set the point source spacing along the line, different simulations were performed, leading to our preferred sampling given in Table 1.

In this study, we locate the earthquake in Northwestern China, and we use the geometry of the $M_w = 7.5$, 2001 Kunlun earthquake (Klinger *et al.* 2005; Vallée *et al.* 2008). The generated source models do not aim at reproducing the exact rupture process of the Kunlun event, but they rather explore different rupture scenarios. However, our synthetic model does share the same hypocentral coordinates, seismic moment, focal mechanism and rupture propagation direction of the M_w 7.5, Kunlun earthquake, as summarized in Table 1. Hypocentral coordinates and focal mechanism are obtained from the Global CMT Moment Tensor Solution (Dziewonski *et al.* 1981). We simulate 40 000 rupture scenarios of the M_w 7.5 strike-slip event adopting the 1-D line source model. To capture a range of different scenarios, we vary the kinematic parameters (rise time, final slip and rupture velocity) for each simulation. Each kinematic parameter is randomly selected within a range of values shown in Table 2 according to the empirical relationships proposed by Wells & Coppersmith (1994) and Geller (1976). In addition, we introduce a heterogeneous segment whose length L_H ranges from one space sample, that is 400 m to 40 km. Within the heterogeneous segment, the rise time, the final slip and the rupture velocity assume a different value from the surrounding length. Although simple in its setting, the range of models we simulate in the study highlights the salient features of the rupture process we aim at investigating.

Table 1. Hypocentral coordinates and focal mechanism of the $M_w = 7.8$, 2001 Kunlun earthquake, used as a frame for our synthetic study. Rupture length L and width W derived from Wells and Coppersmith empirical relations. Space ΔL and time Δt sampling chosen in order to avoid space–time aliasing, as discussed by Lancieri & Zollo (2009). P- and S-wave velocities and quality factor used for the Green's functions.

Hypocentral coordinates (°)	Focal mechanism (°)
Lon = 90.59, Lat = 35.93	Strike = 78, Dip = 61, Rake = −12
Hypocentral depth (km)	Rupture Geometry (km)
$Z = 13$	$L = 100$, $W = 15$
P-wave velocity (km s^{-1})	Space sampling (km)
$a = 5.8$	$\Delta L = 0.4$ km
S-wave velocity (km s^{-1})	Time sampling (s)
$\beta = 3.46$	$\Delta t = 0.02$ s
Attenuation	
$Q = 730$	

Table 2. Ranges of kinematic parameters for rupture process simulation. We call t_r the rise time, s_f the final slip, V_r the rupture velocity, L_H the length of the heterogeneous segment and x_c its central position. Kinematic parameters are uniformly distributed within the range listed below.

t_r (s)	s_f (m)	V_r (m s^{-1})	L_H (km)	x_c (km)
1–8	1–10	2500–3460	0.4–40	20–80

2.1.1 *Two case-studies: the homogeneous and the heterogeneous rupture processes*

In the first part of our study, we aim at reproducing two examples of rupture processes: the homogeneous and the heterogeneous ones. In the first case, the kinematic parameters of rise time, final slip and rupture velocity are set constant along the 1-D line, whereas in the second example a step change in the rise time is set for the points lying within the central segment. In Fig. 1(a), we show the slip-rate functions for each point of the line in space and time in the case of a homogeneous rupture model, where the kinematic parameters are constant along the line. In the marginal plot along the y-axis we show the sum of the slip-rate functions in time, whereas in the marginal plot along the x-axis we show the sum of the slip-rate functions in space, which returns the moment-rate function. In this simple case, the marginal plot along the x-axis shows the well-known trapezoid function described by Haskell. It is worth noting that in the simple case of a 1-D source model, the moment rate function is proportional to the slip rate function. In Fig. 1(b), we show the slip-rate functions of each point of the line for a heterogeneous case. Here, we place a 30-km-long segment in the middle of the line fault and we change the rise time value for the source points within it. The rupture propagates with constant rupture velocity ($V_r = 3$ km s^{-1}) and the source points reach the same final slip value ($s_f = 8$ m). The rise time, on the contrary, has an abrupt change from 6 to 3 s, determining a faster activation of the 'heterogeneous' points. Here, the abrupt change in the rise time value affects the rupture process and makes the moment rate function more complex. In particular, the effect of the decrease in the rise time value within the heterogeneous segment

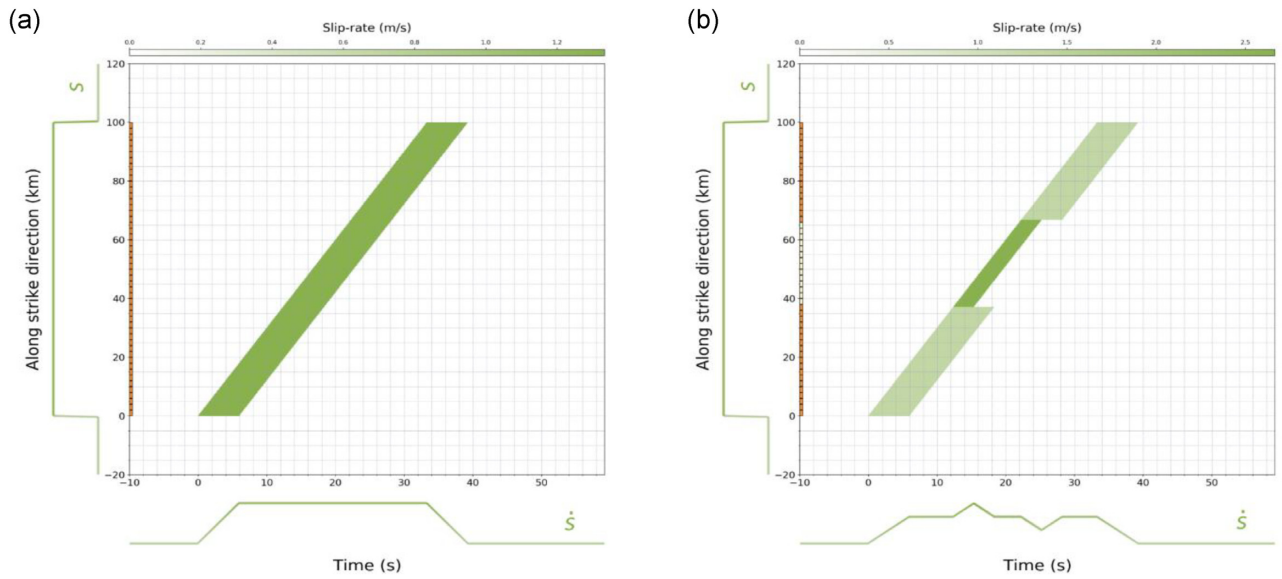

Figure 1. Slip-rate functions in space and time for a homogeneous (a) and a heterogeneous (b) rupture processes. The along-strike direction is shown on the *y*-axis (in orange, the segment that ruptured during the simulation). Time is shown on the *x*-axis. Each source point dislocates following a ramp function (a box-car in terms of slip-rate). The marginal plot along the *y*-axis shows the sum of the slip-rate functions in time, whereas the marginal plot along the *x*-axis shows the sum of the slip-rate functions in space.

determines a shorter duration of the box-car and a higher value of its amplitude. Such a complexity can be seen in the marginal plot along the *x*-axis, where the trapezoid function is perturbed by the presence of the heterogeneity.

2.2 Modelling body waves at teleseismic distance

For each rupture scenario generated in the previous section, we compute synthetic seismograms at three arrays of existing seismic stations at teleseismic distances (between 50° and 90°): Alaska (AK), Europe (EU) and Australia (AU). Each array is comparable to an antenna that tracks in space and time the strongest coherent sources of HF seismic energy. The arrays are located at three complementary locations with respect to the epicentre of the event, allowing us to obtain three different viewpoints of the same rupture process. Each array is composed of 55 stations (black triangles in Fig. 2). The interest of working at teleseismic distances lies in the possibility of investigating HF emissions using only the far-field term of the Green's function in the representation theorem (Aki & Richards 2002). This approximation leads to many simplifications of the calculations and simple physical models of the generation of HF waves. In the current study, Green's functions are calculated in a spherically averaged Earth's velocity model AK135 (Kennett *et al.* 1995). The ground motion associated with the displacement on the fault plane is computed via the representation integral in the frequency domain. Teleseismic body waves are computed by taking into account only the direct *P* arrival and associated depth phases, pP and sP. We compute the displacement associated with a teleseismic *P* wave in a geometrical ray solution following the formula of Okal (1992). However, in our study, we do not model absolute ground motion amplitudes: we neglect the geometrical spreading and the response of both the receiving site and the recording instrument, and we normalize the individual ground motion signals with respect to the maximum amplitude amongst all ground motion signals. To compute the reflection coefficients for pP and sP phases into the synthetic seismograms, we use the calculations made by Aki & Richards (2002). Values for *P*- and *S*-wave velocity are shown in Table 1.

2.3 Description of the synthetic data set

We composed our data set of 55 *Z*-component synthetic displacement traces for each array in Fig. 2. The AK and AU arrays are located almost in the directive position with respect to the rupture propagation, whereas the EU array is located almost in the antidirective position. To identify HF pulses, the synthetic displacement traces are then differentiated in time and bandpass filtered between 0.5 and 4 Hz. The filter is a zero-phase 4-pole Butterworth, to obtain an acausal filtered signal where the arrival time of the signal's peak is respected (Fig. 3).

3 METHODOLOGY II: THE BP IMAGING TECHNIQUE

3.1 Back projection

To compute the BP image, we follow the conventional method for source imaging first proposed by Ishii *et al.* (2005), where the body waves arriving at an array of recording stations are projected back to a reference source grid under the ray theory asymptotic approximation. The steps to follow when applying the BP method are described in the following summarizing points:

1. We define the coordinates of the BP gridpoints and assume a velocity model to calculate theoretical traveltimes. We use a 1-D grid of potential source locations that is 300 km long, and with a step size of 1 km (Fig. 2), to represent the line source. All gridpoints are placed at the hypocentral depth of *x* km;
2. We define the coordinates of the BP gridpoints and assume a velocity model to calculate theoretical traveltimes. We use a 1-D grid of potential source locations placed over the line source. In particular, the grid is 300 km long with a grid-step of 1 km (Fig. 2) and all gridpoints are placed at the hypocentral depth. This is a

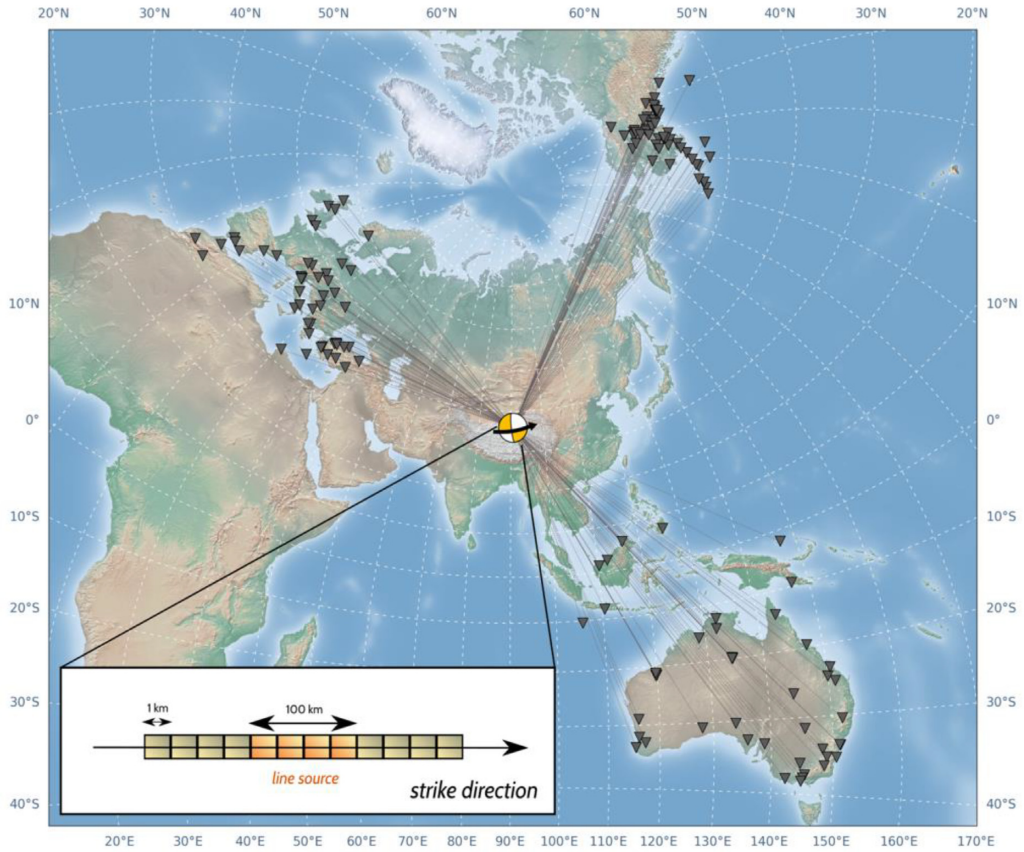

Figure 2. Geographical setting of this study. Focal mechanism and epicentral location of the synthetic test are placed in Northwestern China. Alaskan (AK), Australian (AU) and European (EU) arrays are composed each of 55 seismic stations (black triangles). The bottom window shows a zoom over the 300-km-long 1-D BP grid that surrounds the line source.

simplification of the standard approach in BP analysis where a 2-D grid is used;

3. For each gridpoint, theoretical *P*-wave traveltimes are calculated at each station of the array in the AK135 velocity model (Kennett 2005);

4. Filtered seismograms are time shifted according to the values of traveltimes;

5. Shifted signals are stacked, returning a signal related to the gridpoint. If the shifted signals sum constructively, the stack is high, meaning that the gridpoint is a plausible source of coherent HF radiation (from the array point of view). If the stack is low, in contrast, the gridpoint likely did not contribute coherent HF energy during the rupture process.

Several methods exist for aligning and stacking seismic traces, for instance, the *N*th root stacking (Muirhead 1968) and the *F*-ratio (Melton & Bailey 1957). In our study, seismic traces are combined via the shift-and-sum approach, because this linear approach is the simplest technique for stacking the station traces, it requires the least a priori assumptions, and it does not deform the amplitude of the signals. Mathematically, the linear stack $\boldsymbol{s_i}(t)$ at the *i*th potential gridpoint can be expressed as follows:

$$\boldsymbol{s_i}\ (\boldsymbol{t}) = \sum_j \dot{\boldsymbol{u}}_j \left(\boldsymbol{t} - \boldsymbol{t}_{ij}^{\boldsymbol{P}}\right), \tag{1}$$

where $\dot{\boldsymbol{u}}_j$ is the velocity trace at the *j*th station and $\boldsymbol{t}_{ij}^{\boldsymbol{P}}$ is the theoretical *P*-wave traveltime from the *i*th source to the *j*th station. In the calculation of the BP image, we replace the stacked signal by its square value, the beam power. The beam can be highly noisy especially for HF filtered signals, where further peaks appear because of the filter. To further reduce the noise, the beam is smoothed using a zero-phase Gaussian filter, parametrized by its standard deviation σ. We tested different values of Gaussian smoothing windows (σ), and chose the value of $\sigma = 0.4$ s, which allows us to only focus on the main energy bursts neglecting the smaller peaks coming from the filtering and stacking procedures (see Fig. A2).

3.2 Space–time resolution of BP images: reducing the artefacts

The space–time resolution of BP images is quantified by the array response function (ARF), that is the space–time BP image of an instantaneous point source (e.g. Xu *et al.* 2009; Meng *et al.* 2012). The ARF is generally a 3-D function, with two spatial dimensions (the fault grid) and one time dimension. It's worth noting that in our case the ARF is a 2-D function, since we use a 1-D spatial grid of potential source locations. The size and shape of ARF depend on two factors: the frequency content of the data and the geometry of the array, respectively. In principle, the higher the frequency we are looking at, the more detailed the BP image will be (Schweitzer *et al.* 2012). Likewise, both the aperture and the position of the array with respect to the rupture direction control the resolution of the BP image. Specifically, good resolution in BP images is achieved when using large-aperture arrays (Xu *et al.* 2009). Low-frequency ($f < 1$ Hz) P-waves are less affected by scattering and smaller-scale

Figure 3. Example of a teleseismic synthetic signal generated at AU, AK and EU array for a homogeneous (a) and a heterogeneous (b) rupture process. In each plot, we show the displacement trace on the top, the unfiltered velocity trace in the middle, and the velocity trace bandpass filtered between 0.5 and 4 Hz at the bottom.

heterogeneity, thus their stack generally returns a high coherency. However, they do not provide a good degree of detail on BP images. On the contrary, HF ($f > 1$ Hz) *P* waves provide a high resolution on BP images, but at the same time their waveforms are more easily distorted by small-scale heterogeneity and scattering (e.g. Frankel & Clayton 1986, Takemura & Furumura 2013), often leading to a less coherent stack. In addition, similarity among ground motion signals breaks down as interstation distance increases (Xu *et al.*

2009). The frequency range 0.5–4 Hz is typical in BP analysis at teleseismic distance, because it offers a good balance between the resolution of the BP image and the waveform coherence (Xu *et al.* 2009).

The size and shape of the ARF (which depend on the relative position between the array and the rupture direction) are at the origin of the two typical artefacts of BP imaging, namely spatial 'smearing' and time 'walking' (e.g. Walker & Shearer 2009; Xu *et al.* 2009; Meng *et al.* 2012). The smearing effect makes the beam power signal of the true HF emission source to be blurred in space, whereas the walking effect leads to having a false energy radiation appearing before and after the true emission time.

An additional problem is the presence of HF emissions carried by depth phases, which can obscure some of the first-order features of the rupture process (Yagi *et al.* 2012). Okuwaki *et al*. (2019) show that BP has a depth-dependent bias proportional to the amplitude of the depth phases in the Green's function. At shallow depths (e.g. less than 10 km) these artefacts are caused by non-physical interactions of the back-projected *P* wave and depth phases. For deep earthquakes, on the contrary, depth phases are separated enough from the direct *P* wave and they can be used to constrain the source depth (Kiser *et al.* 2011). However, in this case they might still complicate the interpretation of the BP image, as further HF peaks appear in the image, but they are not generated by the complexities of the source.

A key issue in BP analysis is therefore to remove the 'smearing' and 'walking' artefacts and the spurious emissions carried by depth phases to obtain only the true location of the HF radiation sources in space and time. In Fig. 4, we show by way of example the BP images of a homogeneous rupture process calculated at the AK, AU and EU arrays using the conventional approach described in Ishii *et al.* (2005). The figure shows that a few pulses of HF emission retrieved by the BP analysis are found in correspondence with the initiation and the stopping phases of the rupture. This happens because in a homogeneous rupture process all kinematic parameters are constant along the line fault, and only the initiation and the stopping phases of the rupture generate abrupt changes in the slip-rate function (see the slip-rate function in the bottom subplot of Fig. 4), that in turn produce HF seismic waves. Hence, the two bursts of HF radiation both at the initiation and at the end of the simulation are an effect of the finite duration of the rise-time. The presence of several pulses of HF emissions both at the initiation and at the stopping of the rupture is due to the contamination of depth-phases (see in particular the BP image obtained at the EU array). In addition, the HF emissions also show a 'walking' effect under the form of a time tilting that depends on the relative position between the array and the rupture direction (Fig. 4). Several approaches have been proposed in the attempt to improve resolution and reduce the artefacts in BP images (e.g. Lay *et al.* 2010; Wang *et al.* 2012; Haney 2014; He *et al.* 2015; Nakahara & Haney 2015; Wang & Mori 2016). In the first part of our study, we use the method described in Wang *et al.* (2016, hereinafter W2016).

In the following lines, we describe the general procedure used by W2016, by detailing the two steps used in their method. We will then move to describing the procedure we have used, which was inspired by the W2016 method. In their study, the authors propose a two-step procedure, first to correct the time tilt of HF emission patches and second to reduce the smear around the true HF energy peaks. The first step of the approach proposed by W2016 is performed by selecting a reference station lying in a central position within the array. In the conventional BP technique, the signal is shifted by the theoretical traveltime t_{ij}^{P} between the gridpoint i and the station *j*. In the approach suggested by W2016, the signal recorded at a station *j* is shifted by the difference between the traveltime at the station *j* and the traveltime at the reference station:

$$\boldsymbol{s}_i\ (\boldsymbol{t}) = \sum_j \dot{\boldsymbol{u}}_j\ \left(\boldsymbol{t} - \boldsymbol{\tau}_{ij}^{P}\right). \tag{2}$$

and

$$\boldsymbol{\tau}_{ij}^{P} = \boldsymbol{t}_{ij}^{P}\ - \boldsymbol{t}_{iJ}^{P}. \tag{3}$$

where $\dot{\boldsymbol{u}}_j$ is the velocity trace at the *j*th station, $\boldsymbol{\tau}_{ij}^{P}$ is the *i*th source traveltime difference between the *j*th station and the reference station J. This procedure returns a BP image of the rupture as if it were seen by the reference station, where the HF emissions are no longer tilted in time. However, by following the 'time correction' of the approach proposed by W2016, we obtain a BP image where the time axis is no longer the absolute time at the source, but it rather corresponds to the apparent time at the reference station. It is instructive to note that, in this new reference system, the directivity effect of the source appears in the BP images in terms of time stretching or compression of the HF emissions (see the Appendix).

In their second step, W2016 suggest looking at the BP image of an earthquake as a convolution between the BP image of the true HF peaks and the ARF. The ARF can be evaluated in different ways: theoretically, with synthetic Green's functions (e.g. Rost & Thomas 2002), or with empirical Green's functions. The latter is usually performed by applying the BP analysis to a smaller earthquake that can be assumed to be a point source. The procedure proposed by W2016 can also help remove the contamination of depth phases into the BP images. In their study, W2016 use an aftershock as an effective point source and a non-negative least squares (NNLS) algorithm to perform the deconvolution of a BP image of the array response from the BP image of the true HF peaks. In our study we use a synthetic point source activating at the hypocentre as array response and perform an image deconvolution using the Richardson and Lucy restoration algorithm (Richardson 1972; Lucy 1974), hereinafter R&L, available in the *scikit-image* image processing library in Python. The R&L algorithm is an iterative procedure for recovering from a blurred, contaminated image an underlying, original image through the deconvolution with a known point spread function (the array response function, in our case). In Fig. 5, we show an example of the R&L restoration algorithm applied to the BP image of a homogeneous rupture process calculated at the EU array, where we can see the removal of depth phases and a highly reduced smearing around the true HF peaks. To more easily identify the position of the HF peaks in a space–time plot, we then search for the HF peak maxima in space and time in the matrix resulting from the R&L algorithm. Once the HF maxima have been extracted, we restore the time axis in the BP images from the apparent time of the reference station to the absolute time at the source by inversion of the W2016 time correction equation.

3.3 Relation between BP and fault slip

To better understand the relation between the BP image and the fault slip, we select two representative cases among the 40 000 rupture scenarios:

1) A homogeneous rupture model, where the kinematic parameters are constant along the line.

2) A heterogeneous rupture model, where a segment of 30 km length is introduced in the middle of the fault; within this segment,

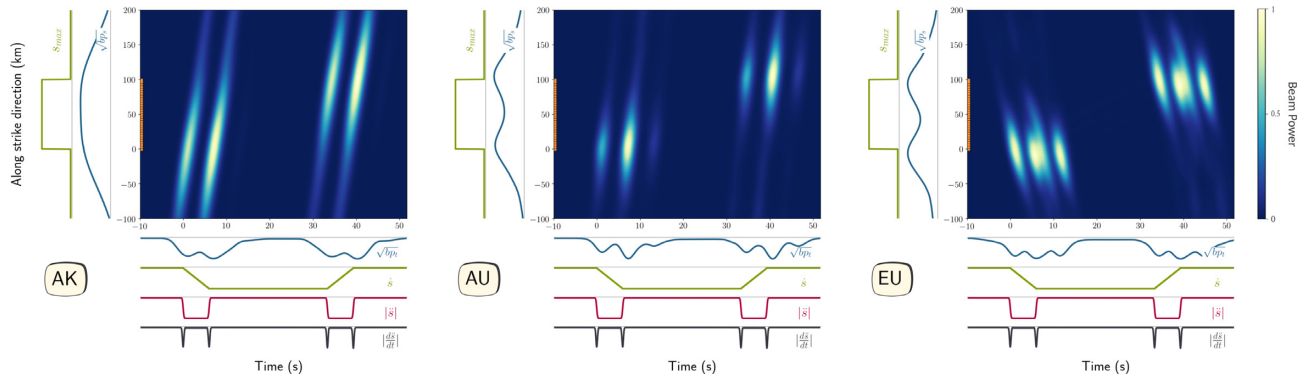

Figure 4. BP images of a homogeneous rupture process (a) and a heterogeneous rupture process (b) calculated at the three arrays. Time is plotted along the *x*-axis and the along-strike direction is plotted along the *y*-axis. In the left-hand subplots we show the square root of the beam power integrated in time (in blue) and the maximum value of the slip-rate function (in green). In the bottom subplots we show the square root of the beam power integrated in space (in blue), the slip-rate function (in green) and the absolute values of the slip-acceleration function (in red) and its time derivative (in black). The blurred and tilted patches correspond to the BP reconstruction of the HF emissions related to the source complexity and the contamination of the depth phases.

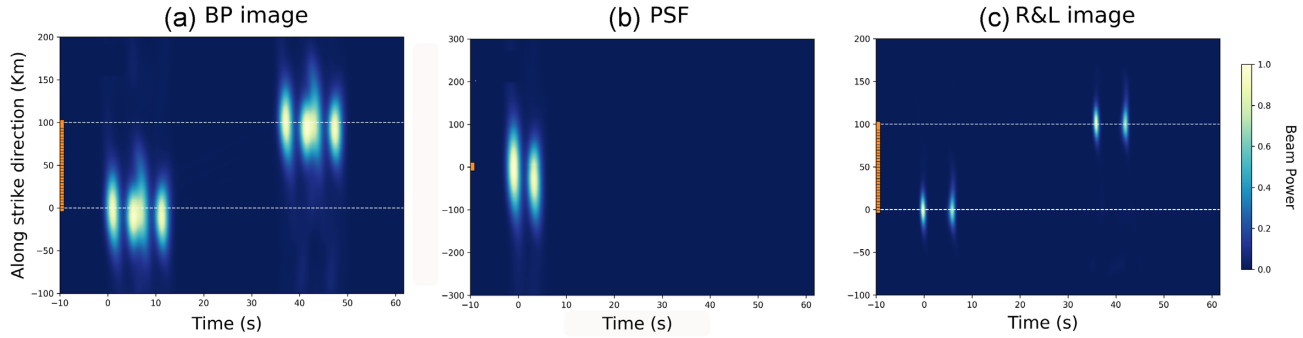

Figure 5. Example of the Richardson and Lucy restoration algorithm applied to the BP image of a homogeneous rupture process calculated at the EU array. (a) Original BP image; (b) point spread function (PSF, also known as array response function) obtained as the BP image of a point source activating at the hypocentre and (c) restoration result after the R&L algorithm.

the rise-time value assumes a different value with respect to the surrounding rupture.

3.3.1 Homogeneous rupture model

In a homogeneous rupture model, the kinematic parameters are constant along the line: the rupture propagates along the line fault with constant velocity and the source points have the same value of rise time and final slip. Specifically, the case we selected presents the following values: rupture velocity $V_r = 3$ km s^{-1}; rise time $t_r = 6$ s; final slip $s_f = 8$ m. In a 1-D line source model, the moment rate and slip rate functions are proportional. In addition, in a homogeneous rupture process, the shape of the slip-rate function is very simple, and the unique abrupt changes are the slopes associated with the initiation and stopping of the rupture. An example of teleseismic synthetic traces generated at the three arrays is shown in Fig. 3(a), where the signal is characterized by two abrupt changes, corresponding to the initiation and the stopping of the rupture, whereas the remaining portion of the displacement signal is flat, indicating a constant rupture process. In Fig. 3(a) we note that the HF bandpass filter behaves like a time-derivative on the velocity trace as, more specifically, the pulses on the HF-filtered signal highlight the discontinuities of the velocity trace. In particular, the finite duration rise-time produces a two-pulse HF radiation pattern, both for the initiation and the stopping of the rupture. The effect of the depth phases is almost imperceptible on the AK array, because of the radiation pattern. Here, four HF pulses are visible in association with the discontinuities in the slip-rate function due to the initiation and the stopping phases of the rupture. In contrast, the waveforms at the AU and EU arrays are measurably perturbed by the depth phases. Here, synthetics show at least three pulses for the AU array and four pulses for the EU array.

In Fig. 6(a), we show the location in space and time of the HF peaks obtained via the R&L restoration algorithm to eliminate the contamination of depth phases in the BP images. The main plot in the background (the diagonal red and blue lines) shows the normalized slip acceleration functions for each point of the line: since we are describing the dislocation at each point of the source as a ramp function, the slip velocity function of each point will be a box-car, and the slip acceleration function will be a two-delta signal (red and blue spikes) whose distance is the chosen value of rise time. In this framework, a given place on the fault experiences sudden onset, constant acceleration, then steady sliding, and then deceleration as the rupture front passes. We normalize the slip-acceleration functions because we are mainly interested in highlighting the relation between its discontinuities and the HF peaks. At the bottom of Fig. 6(a) we show the total slip-acceleration function a which is calculated as the sum over space of the slip acceleration at a given time. The second plot in the marginal at the bottom shows the time derivative of the total slip acceleration function in absolute value. The third plot in the marginal at the bottom shows the square root

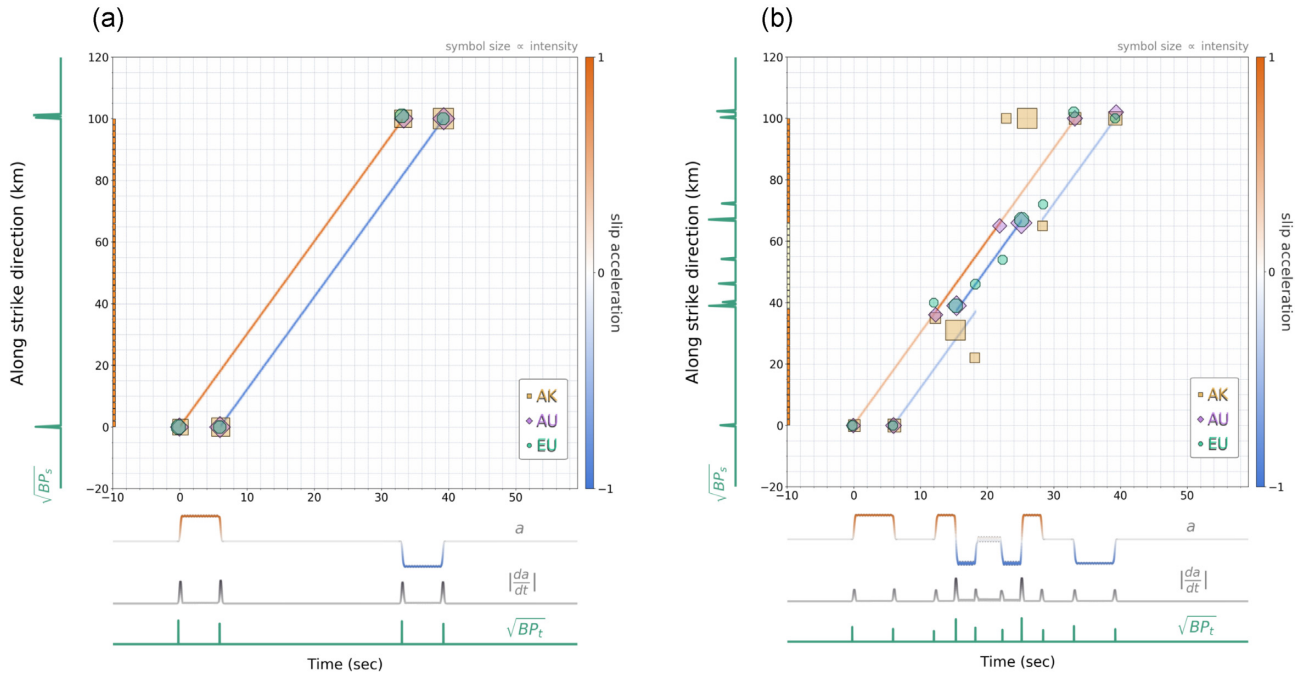

Figure 6. HF peaks (coloured symbols) obtained for a homogeneous (a) and a heterogeneous (b) rupture process retrieved at AK, AU and EU array. HF peak size is proportional to the value of the beam power. In the main plot we show the normalized slip-acceleration function for each point of the line. The along-strike direction is on the *y*-axis (in orange we plot the line source) and the time is on the *x*-axis. The marginal on the left is the square root of the beam power signal for the EU array. The marginal on the bottom are: the total slip-acceleration function a; its time derivative in absolute value; the square root of the beam power signal for the EU array.

of the beam power signal for the EU array. The main plot shows that the three arrays are in agreement in terms of the location of the peaks in space and time and their beam power intensity. Four HF peaks are observed: the first two, generated at the hypocentre, corresponding to the rupture initiation; the other two, generated at the end of the ruptured segment, corresponding to the rupture stopping. In particular, if we compare the HF peaks with the subplots in the time axis, we note that the HF peaks appear at the discontinuities of the slip acceleration function, indicating a strong association between the R&L deconvolved BP images and the absolute value of the time derivative of the slip-acceleration function.

3.3.2 *Heterogeneous rupture model*

In the heterogeneous case, a 30-km-long segment is placed in the middle of the line fault and we change the rise time value for the source points within it. The rupture propagates with constant rupture velocity ($V_r = 3$ km s^{-1}) and the source points reach the same final slip value ($s_f = 8$ m). The rise time, on the contrary, has an abrupt change from 6 to 3 s, determining a faster activation of the 'heterogeneous' points. An example of teleseismic synthetic traces for this heterogeneous rupture process generated at the three arrays is shown in Fig. 3(b), where the signals show a higher complexity beyond the initiation and the stopping of the rupture. In particular, the decrease in the rise time value within the heterogeneous segment produces additional elastic waves, since the moment rate function experiences a transient acceleration when encountering the heterogeneity. Multiple pulses in fact appear on the HF-filtered traces Fig. 6(b): they mark not only the initiation and the stopping phases, but also the major discontinuities in the slip-rate function due to the location of the heterogeneity.

In Fig. 6(b), the three arrays are not always in agreement in terms of the location of the peaks in space and time, nor do they retrieve the same number of peaks. As previously seen in the homogeneous rupture process, HF peaks are again observed at the hypocentre, corresponding to the rupture initiation, and at the end of the ruptured segment, corresponding to the rupture stopping. Additional HF peaks appear at the edges of the heterogeneous segment in space, and they match the discontinuities of the slip acceleration function in time. The result of the peak extraction following the R&L restoration algorithm is not optimal for the AK array, whereas it works well for the AU and EU arrays. For the EU array in particular, the beam power closely matches the absolute value of the time derivative of the slip-acceleration function. Along the strike direction, the localization of the peaks accurately identifies the hypocentre and the end of the ruptured segment, whereas the localization of the peaks generated by the presence of the heterogeneous segment is less clear.

4 A CNN APPROACH

The second step of this work is to understand the statistical link between the raw BP image of an earthquake and the kinematic parameters controlling the rupture process. CNNs have recently become a successful tool in earthquake seismology, where they have been used as a powerful alternative to many classical techniques (e.g. Perol *et al.* 2018; Rouet-Leduc *et al.* 2020). Even when complex and non-linear interactions exist among observed data and model parameters of interest, CNNs can find useful functional relationships between the input and outputs, and provide accurate predictions. CNNs are frequently applied to both time-series and image inputs, where the input data can have any number of parallel input channels. The input data are processed through numerous convolutional layers, which extract relevant features from the data that the later

portions of the CNN can use to predict the output (Fig. 7). Here we apply a CNN approach by training on the 40 000 BP images of the rupture scenarios previously generated and attempting to predict rupture kinematic parameters of each realization. To take full advantage of the features contained in the BP image, whether they be `walking' and `smearing' artefacts, or true source emissions, for each rupture process we calculate the BP image using the conventional approach by Ishii *et al.* (2005). In the calculation of the BP image, we replace the stacked signal by its square value, the beam power. However, in this second part of our study, we want to apply a minimal pre-processing and allow the CNN to find the best link between the raw BP image and the kinematic parameters. Because each array introduces distinct artefacts into the BP image, it may be possible to pre-process the BP images to mitigate the geometric effects prior to apply the CNN. However, such pre-processing procedures are inherently non-unique and risk losing potential valuable information from the raw images. Hence, hereafter the beam is not smoothed with a Gaussian filter and, most importantly, no peak extraction or other manipulation is made on the raw BP images.

4.2 CNN architecture

The CNN is designed as a regression model that is trained in a supervised way, to predict a target output, for each given input. In our study, the input is a 2-D matrix containing the pixels of the BP images of the simulated rupture processes, and the target is a 1-D vector containing the kinematic parameters corresponding to those simulations. Specifically, the target vector d is defined as:

$$\boldsymbol{d} = [\boldsymbol{t_r},\ \boldsymbol{s_f},\ \boldsymbol{V_r},\ \boldsymbol{H_{tr}},\ \boldsymbol{H_{sf}},\ \boldsymbol{H_{Vr}},\ \boldsymbol{x_c},\ \boldsymbol{L_H}], \tag{4}$$

where t_r is the rise time, s_f is the final slip, V_r is the rupture velocity; H_{tr}, H_{sf} and H_{Vr} are the heterogeneities, respectively, in the rise time, in the final slip and in the rupture velocity; x_c and L_H are the central position and the length of the heterogeneous segment. The objective of the CNN algorithm is to learn from the training data a functional mapping between the input matrix and target vector (LeCun *et al*. 1998, Goodfellow *et al.* 2016).

We tested various network architectures in our study. Our preferred one has the structure sketched in Fig. 7, because it required a small number of learnable parameters (improving its generalization performance) and was computationally efficient in training. In this architecture, the BP image is first downsampled with a max-pooling operation, with equal pooling lengths of [3] w × [3] h along the image dimensions. Once downsampled, the BP image is passed through two convolutional layers, each with a convolution, a max-pooling, and an activation function. The convolution is performed with 5 and 10 filters, in the first and second layers, respectively. The kernel size for both layers is [5] w × [5] h. The max-pooling operation is performed with pooling lengths of [4] w × [4] h in the first layer and of [3] w × [3] h in the second layer of the loop. A rectified linear unit (ReLu, Ramachandran *et al.* 2017) is used as an activation function. Two fully connected layers link these convolutional layers with the output target (Fig. 7). We refrain from using a large number of convolution and pooling layers, as is common in CNN applications, because we want to preserve a substantial portion of the 'bright spots' in BP images, whose shape depends on the relative position between the source and the array (as discussed in Section 3.2) and thus carry information on the geometry of the problem. Network weights are updated during the learning to minimize the loss function, in our case, we use the mean square error (MSE):

$$\mathrm{MSE} = \frac{1}{M}\sum_{n=1}^{M}(d_n - y_n)^2, \tag{5}$$

where M is the number of samples in the data set, d_n is the normalized target vector containing the true kinematic parameters for the sample n, and y_n is the normalized output vector containing the prediction of the kinematic parameters for the sample n. The kinematic parameters used in the simulations are uniformly distributed, hence we adopt a min-max normalization on target data to map them to the range 0 to 1:

$$\tilde{d}_n = \frac{d_n - \min(d)}{\max(x) - \min(x)}, \tag{6}$$

where $\tilde{\boldsymbol{d}}_{\boldsymbol{n}}$ is the normalized target. During the training, the updates to the model weights are controlled by the learning rate, which quantifies how fast the model adapts to the problem. Specifically, the learning rate will constrain the step by which the new weights differ from the old weights as we try to minimize the loss function. In our study, we initialize the weights using the Glorot initialization scheme and set the learning rate equal to 7.5×10^{-5}. The Adam algorithm (Kingma & Ba 2014), applied in the PyTorch framework (Paszke *et al*. 2017), is used to train the networks. The original data set composed of 40 000 simulations of rupture processes is divided into three subsets: training (70 per cent), validation (20 per cent) and testing (10 per cent). The network is trained for up to 500 iterations over the training data set using a batch size of 135. The model's parameters for which the MSE reaches its minimum in validation are selected as the best model's parameters, which are then used for testing.

5 STATISTICAL LINK BETWEEN THE BP IMAGE AND THE TARGET PARAMETERS

5.1 Effect of target on CNN predictions

In Fig. 8, we present the results obtained with the design of the CNN sketched in Fig. 7 and the target vector defined in eq. (4). We include outputs of both the stacked version of the method (sum of the BP images obtained from the three arrays), and the predictions for using each array separately. The plots show the output of the CNN versus the target values, which are the predictions of the CNN versus the true values of kinematic parameters, for the testing data set. To summarize the prediction accuracies, we calculate the regression score function, R^2, for each of the different components of the target vector. This score is defined as:

$$\boldsymbol{R}^2 = 1 - \frac{\sum_{i=1}^{n}(\boldsymbol{d_i} - \boldsymbol{y_i})^2}{\sum_{i=1}^{n}\left(\boldsymbol{d_i} - \hat{\boldsymbol{d}}\right)^2}, \tag{7}$$

where d is the target vector, y is the CNN prediction and $\hat{\boldsymbol{d}}$ is the mean value of the target. This score captures how well the model is able to return a prediction close to the target parameter in a linear regression setting. Its range of possible values is (−1, 1], where the best possible score is 1.0 and a score of 0.0 implies the model predictions are no better than a simple guess based on the average value. A negative value for R^2 implies the model predicts worse than the simple guess. The highest values of R^2 are found for the predictions of the rise time t_r ($R^2 = 0.912$) the rupture velocity V_r ($R^2 = 0.914$) and the central position of the heterogeneous segment

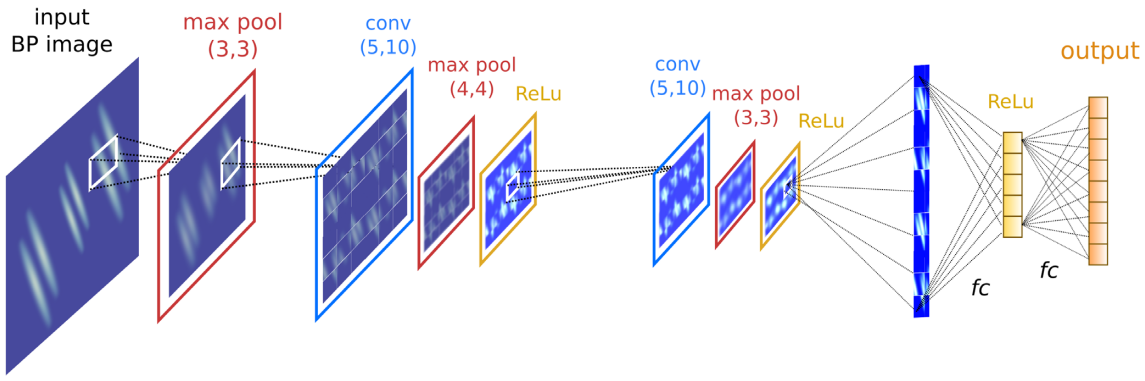

Figure 7. Sketch illustrating the design of the CNN used in this study. The input is the BP image which undergoes a max pooling operation first. Then, once downsampled, the BP image is passed through two convolutional layers, each composed of a convolution (conv), pooling and activation (ReLu) operation. The network ends with two fully connected layers (fc). The output of the network is the vector containing the kinematic parameters of the rupture process.

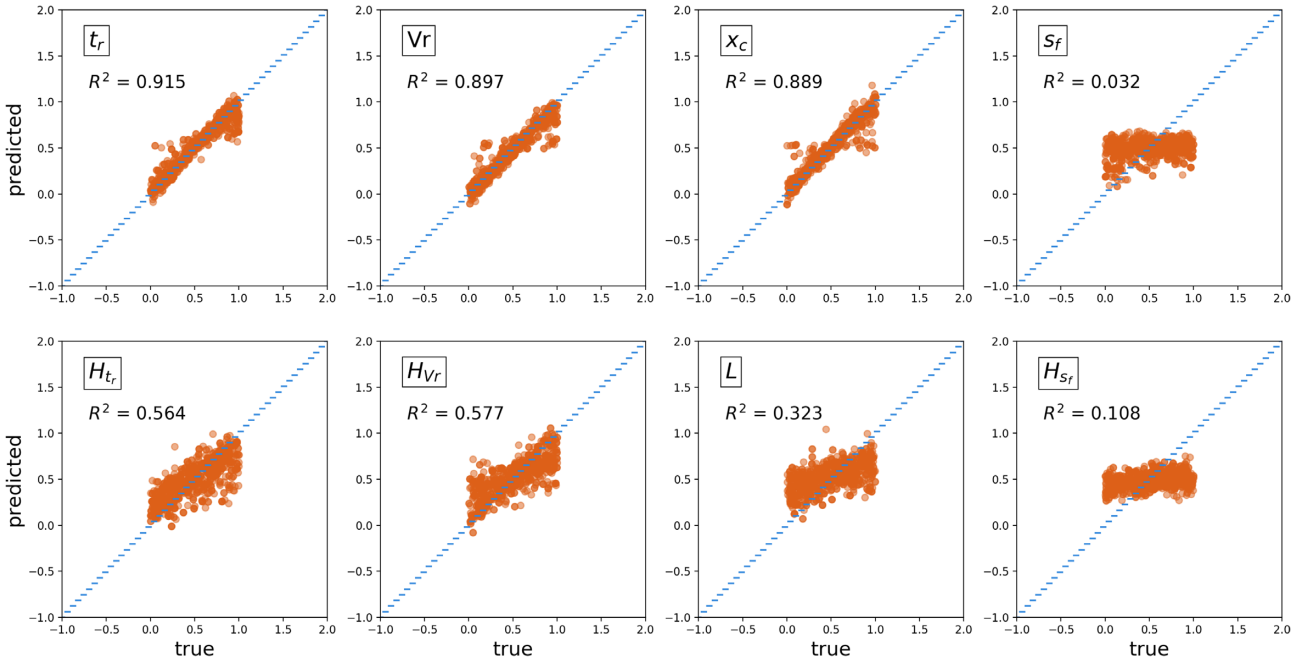

Figure 8. Predictions of the CNN versus the true values of kinematic parameters, for the testing data set. The regression score function R^2 is shown in each subplot. The input BP image is obtained as the sum of the BP images calculated at the three arrays AK, AU and EU. We adopt a min–max normalization on target data to map them to the range 0–1.

x_c ($R^2 = 0.885$). Accurate predictions are also obtained for the heterogeneous values in rise time and rupture velocity, H_{tr} ($R^2 = 0.576$) and H_{vr} ($R^2 = 0.499$). However, the length of the heterogeneous segment L_H is not well predicted ($R^2 = 0.355$) and poor predictions are obtained for the final slip s_f ($R^2 = 0.123$) and its heterogeneous values H_{sf} ($R^2 = 0.017$). We attribute this shortcoming to the inherent insensitivity of the BP approach to image the low-frequency aspects of the rupture process.

Difficult-to-predict parameters in the target vector slow down the loss function's convergence towards its minimum. Therefore, we define a different target vector, where we remove the final slip value and its heterogeneity and keep the following kinematic parameters:

$$d = [t_r,\ V_r,\ H_{tr},\ H_{Vr},\ x_c, L_H]\ . \tag{8}$$

In Fig. 9, we show the prediction versus the true values of the reduced target vector in eq. (8). As expected, the performance on the remaining parameters does not change significantly with a different choice of target vector, however it does show a modest improvement and increased training convergence rate.

5.2 Effect of input on CNN predictions

As complementary information, we test whether a different choice of input parameter could affect the CNN results. In the previous cases, the input of the CNN was a matrix containing the BP images of the simulated rupture processes, obtained as the sum of the BP images calculated at the three arrays AK, AU and EU. Here, we test whether the combination of the three arrays or the employment of only one array at a time in the CNN approach could help achieve better results.

In Fig. 10, we present the comparison between the results of the CNN approach applied to the sum of the BP images and to the single-array BP image using the reduced target vector. It is instructive to

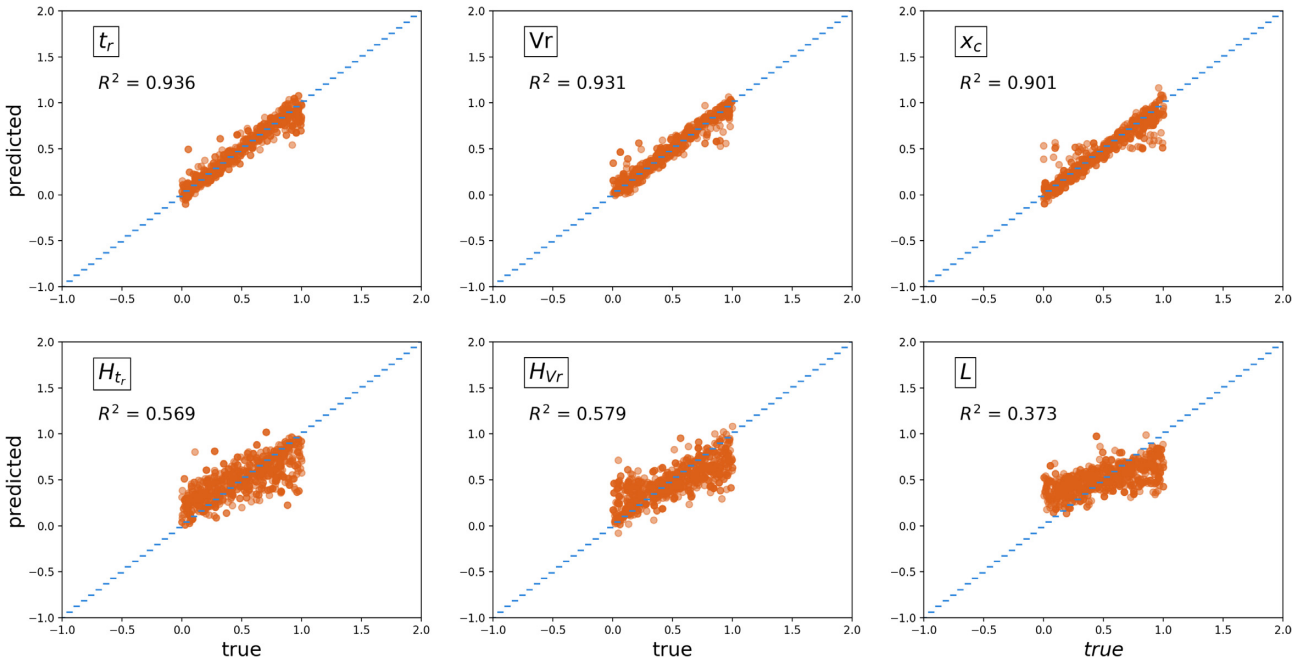

Figure 9. Predictions of the CNN versus the true values of kinematic parameters of the reduced target vector for the testing data set. The input BP image is obtained as the sum of the BP images calculated at the three arrays AK, AU and EU. Target data have been scaled to 0–1 via min–max normalization.

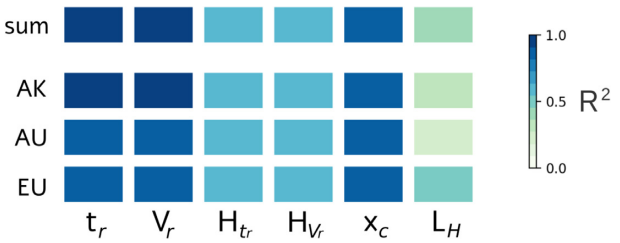

Figure 10. Regression score function (R^2) for the CNN predictions of the kinematic parameters in the reduced target vector when using the sum of the three BP images (on the top) or the single-array BP image as input for the network. The score represents the quality of the model indicating how close the predictions are to the target parameters in a linear regression setting.

note that not all arrays perform the same. In particular, slightly better predictions of the rise time and its heterogeneous value can be found when using the BP image of the AK array. The comparison also shows an improvement in the predictions of the length of the heterogeneous segment when we use an input BP image calculated at the EU array. On the contrary, when using the AK or the AU array in isolation, the predictions of the reduced target vector do not improve in comparison to the initial test where the input is the sum of the three BP images. In these cases, the smearing or the walking effect on BP images may hamper the accuracy of the CNN predictions. In our case, the EU and AU array response functions are quite focused along the space direction, whereas the AK array shows higher smearing. The EU array shows the strongest walking effect, because of its large backazimuth, whereas the HF emissions on AK and AU BP images are not very tilted in time (see as an example the Fig. 4). However, only the EU array allows the CNN to return the best predictions of the length of the heterogeneous segment, even better than the outputs obtained with the summed BP image. On the contrary, both AK and AU arrays perform worse than the summed BP image for the length of the heterogeneous segment. Even though the smearing effect could encumber the ability of the network to extract information on the BP images, we cannot attribute to it the shortcomings of the CNN on AK and AU arrays.

We question, therefore, whether the different performances of the CNN on individual arrays could be attributable to the backazimuth of the array relative to the rupture. In the conventional BP analysis, the delay-and-sum approach does not carry directivity effects in the duration of the rupture, nor in the time separations between HF pulses in the BP image. Hence, the time axis of the BP image is not affected by the relative position between the rupture direction and the array position. Nevertheless, different tilt angles for the smeared HF radiations are seen in the BP images depending on the backazimuth. Because of the array configuration and position respect to the rupture direction, coherence among ground motion signals can vary from one array to another, determining a different pattern of HF peaks in the BP image among arrays at complementary azimuths (e.g. Xu *et al.* 2009). As a further analysis, we investigate whether the backazimuth and thus the different HF emission pattern and tilt in the BP image, can favour or restrain the quality of the CNN predictions. We focus the analysis on the EU array because of its strong walking effect. We rotate the previously used line source at nine different azimuths ranging from 0°, being the azimuth of the EU array reference station, to 320°. For each azimuth, we generate 40 000 rupture scenarios and apply the CNN method on them. In Fig. 11, we show the prediction of the CNN for the parameters in eq. (6) and compare the quality of the CNN prediction obtained at nine different azimuths. For each kinematic rupture parameter, we highlight in orange the variability of R^2 as a function of azimuth in the corresponding regression score colour-bar. From this analysis, we can note that the central position of the heterogeneous segment is the best retrieved parameter for all azimuths, whereas the length of the heterogeneous segment shows the widest range of regression scores. But even this range is rather modest [0.35–0.50], suggesting that azimuthal changes have a minimal effect on the quality of the CNN predictions.

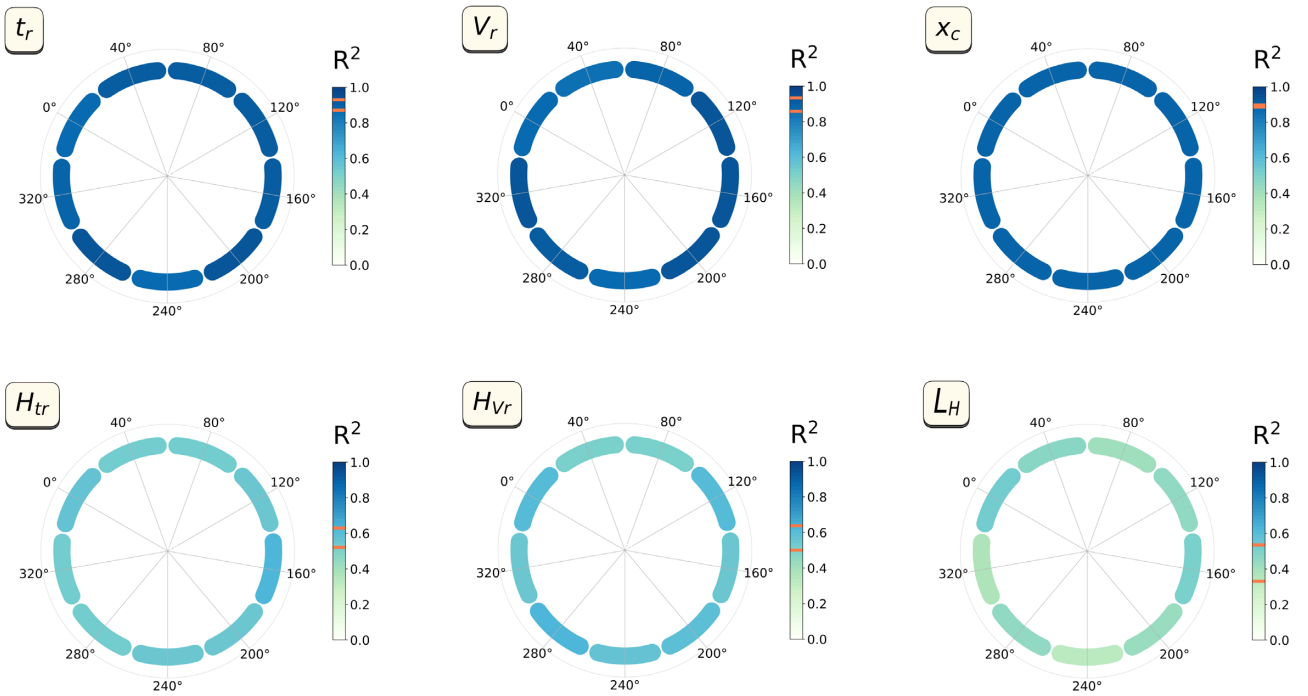

Figure 11. Azimuthal dependence of CNN predictions for the kinematic parameters in the reduced target vector. The variability of the regression score is represented with an orange interval. For this test the EU array is used and 0° is the backazimuth of the EU array with respect to the line fault strike.

6 DISCUSSION AND CONCLUSION

The objective of this study was to investigate the link between BP images and the kinematic parameters of the rupture, and, in particular, slip motion on the fault (velocity or acceleration). For this purpose, we studied a large data set of synthetic line rupture processes, characterized by a heterogeneous segment (in terms of final slip, rise time or slip velocity) with variable length and position. Synthetic traces, filtered between 0.5 and 4 Hz, are back projected following the approach of Wang *et al.* (2016) and the HF peaks are retrieved through deconvolution between the BP image and the array response function (ARF) for the given frequency range. For the particular horizontal line-source configuration chosen here, depth phases strongly contaminate BP images, introducing, for each HF peak, two 'ghost' peaks associated with pP and sP phases. Deconvolution with an ARF that includes depth phases restores the original number of peaks with accurate time resolution, but occasionally the peak is shifted modestly in space. The role played by depth phases is amplified by the simple horizontal line source geometry chosen here, where all the fault points lay at the same depth. In a 1-D line source, the locus of points on the fault from which radiation arrives at the observation point at the same time, also called the *isochrones* (Bernard & Madariaga 1984; Spudich & Frazer 1984), are in fact just a single point. In this setting, the recording stations 'see' at a given time the energy emitted from just a single point, where depth phases emerge from the same depth. In a more realistic 2-D fault, the isochrones on the fault plane are generally incoherent and lay at different depths. In this case, the recording stations will 'see' at a given time the energy emitted from different points along the fault, and depth phases emerge from different depths. Moreover, the spatial extension of the ARF introduces a further averaging scale which should reduce the coherency of depth phases. The resolution of BP becomes very poor at depth, hence improved BP techniques such as the hybrid BP proposed by Yagi *et al.* (2012) could be useful in mitigating the effect of depth phases on BP images of 2-D synthetic fault models.

An important question is whether BP images are associated with slip velocity or acceleration (Fukahata *et al.* 2014). Here we show that filtering plays an important role, since the 0.5–4 Hz bandpass filter, typically used in BP analyses (Xu *et al.* 2009) behaves like a time derivative for the seismograms. Comparison between HF peaks extracted from BP images with the slip rate and slip acceleration function, shows that the beam power is more directly related to the absolute value of the time derivative of the slip acceleration function, when narrow-band filtering is used. In the case of our simple line source models with ramp acceleration functions, the deconvolved BP images (Figs 5 and 6) clearly highlight the abrupt discontinuities in slip acceleration. This results in detecting two HF peaks from each spatial point, during the onset and termination of the rupture.

The ability of BP images to retrieve the rupture kinematic parameters was tested using a CNN approach on BP images. CNN are data-driven predictive models, whose performance depends on the specification of input and target parameters, and the richness of the training data set. We found in our application that the CNN is able to predict the rise time, the rupture velocity, the heterogeneous values in rise time and rupture velocity, the length, and the central position of the heterogeneous segment. However, the CNN fails at predicting the final slip. We attribute this shortcoming to the inherent insensitivity of the BP approach to the low-frequency aspects of the rupture process. Our study has focused on studying the impact on BP images of HF radiation due to basic geometric and kinematic rupture properties, however, future developments could include training the CNN to infer other source characteristics that are typically well resolved by the BP method, such as the source length and orientation, or the earthquake moment.

As additional information, we also tested whether the information coming from one single array could be thorough for the network

or, on the contrary, if the combination of the information coming from the three arrays could provide us with better predictions. We tested whether the backazimuth of the array relative to the rupture could influence the quality of the CNN predictions. Our analysis shows no strong azimuthal dependence in the quality of the CNN predictions depending on the relative position between the source and the array. Thus, from a CNN perspective, stacking multiple arrays may not always provide the best outcome, in contrast to this being a relatively commonplace practice in teleseismic BP analysis. It is worth noting that we trained our CNN on a simplistic case of a line source, with the objective to assess the resolving power of BP in a controlled test. Generalization of our CNN approach will require training the CNN on more realistic 2-D source models, for example using fractal slip distribution (Ruiz *et al.* 2011), dynamic modelling, or real earthquakes. In the future, a promising approach would be to train a CNN that also infers for, or allows, variable input fault geometry, such that the same network could be applied to a wider range of earthquake rupture scenarios. Additionally, rather than predict the small set of descriptive kinematic parameters (eq. 8), the CNN could directly predict the continuous slip rate function at each spatial gridpoint. These additional complexities will enable possible real time monitoring, whereby large earthquakes occur, and a CNN is applied to rapidly estimate the continuous space–time kinematic rupture history from the recorded data at teleseismic arrays.

Our study demonstrates the potential upsides of machine learning approaches in providing reasonably accurate predictions for the other kinematic parameters of the rupture process. Although not fully addressing the question of the generalizability of the CNN method, our analysis does reveal new promise in processing BP images to obtain relevant information on the rupture process. Further analysis on the waveform content of synthetic data, as well as a careful analysis on the similarity between the BP images of real data and synthetic data in a fixed frequency band, would enrich our study, potentially making it a suitable approach for real data too, as long as waveforms are carefully pre-processed. Future work will address these complexities and more, and we hope to see kinematic rupture histories imaged with deep learning methods in the future.

CROSSREF FUNDING DATA REGISTRY

This study was partially supported by the European Research Council under the European Union Horizon 2020–SERA research and innovation program. Numerical computations were performed on the S-CAPAD platform, at Institut de physique du globe de Paris (IPGP), France.

DATA AVAILABILITY

Data available on request.

AUTHOR CONTRIBUTION STATEMENT

MC and CS conceived the study. MC performed the computations and processed the data. MC, CS and PB interpreted and discussed the link between the BP image and the slip function. MC, IM and DTT conceived the computational framework of the CNN analysis. IM designed the CNN architecture. MC, IM and DTT interpreted the data and discussed the results of the CNN analysis. MC drafted the manuscript. All of the authors provided critical feedback on the results and discussion and helped shape the manuscript. PB and PAJ supervised the project.

REFERENCES

Andrews, D. J. 1981. A stochastic fault model: 2. Time-dependent case, *J. geophys. Res.,* **86**(B11), 10 821–10 834.

Bernard, P. & Madariaga, R. 1984. High-frequency seismic radiation from a buried circular fault, *Geophys. J. Int.,* **78**(1), 1–17.

Dziewonski, A. M., Chou, T.-A. & Woodhouse, J. H. 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Frankel, A. & Clayton, R. W. 1986. Finite difference simulations of seismic scattering: implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. geophys. Res.,* **91**(B6), 6465–6489.

Fukahata, Y., Yagi, Y. & Rivera, L. 2014. Theoretical relationship between back-projection imaging and classical linear inverse solutions, *Geophys. J. Int.,* **196**(1), 552–559.

Geller, Robert J., 1976. *Scaling relations for earthquake source parameters and magnitudes,* Bulletin of the Seismological Society of America, **66**(5), 1501–1523, doi.org/10.1785/BSSA0660051501.

Goodfellow, I., Bengio, Y. & Courville, A. 2016. *Deep Learning*. MIT Press.

Grandin, R., Vallée, M., Satriano, C., Lacassin, R., Klinger, Y., Simoes, M. & Bollinger, L.(2015), Rupture process of the Mw = 7.9 2015 Gorkha earthquake (Nepal): insights into Himalayan megathrust segmentation, *Geophys. Res. Lett.,* **42,** 8373–8382.

Haney, M. M. 2014. Backprojection of volcanic tremor, *Geophys. Res. Lett.,* **41**(6), 1923–1928.

He, X., Ni, S., Ye, L., Lay, T., Liu, Q. & Koper, K. D. 2015. Rapid seismological quantification of source parameters of the 25 April 2015 Nepal earthquake, *Seismol. Res. Lett.,* **86**(6), 1568–1577.

Herrero, A. & Bernard, P. 1994. A kinematic self-similar rupture process for earthquakes, *Bull. seism. Soc. Am.,* **84**(4), 1216–1228.

Ide, S. 1999. Source process of the 1997 Yamaguchi, Japan, earthquake analyzed in different frequency bands, *Geophys. Res. Lett.,* **26**(13), 1973–1976.

Ishii, M., Shearer, P. M., Houston, H. & Vidale, J. E. 2005. Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array, *Nature,* **435**(7044), 933–936.

Kennett, B. L. N., 2005. Seismological tables: ak135, *Research School of Earth Sciences Australian National University Canberra, Australia,* 1-289.

Kennett, B. L. N., Engdahl, E. R. & Buland, R. 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.,* **122,** 108–124.

Kingma, D. P. & Ba, J. 2014. Adam: a Method for Stochastic Optimization, CoRR, abs/1412.6980, arXiv:1412.6980.

Kiser, E. & Ishii, M. 2017. Back-projection imaging of earthquakes, *Annu. Rev. Earth planet. Sci.,* **45,** 271–299.

Kiser, E., Ishii, M., Langmuir, C. H., Shearer, P. M. & Hirose, H. 2011. Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases, *J. geophys. Res.,* **116,** B06310.

Klinger, Y., Xu, X., Tapponnier, P., Van der Woerd, J., Lasserre, C. & King, G. 2005. High-resolution satellite imagery mapping of the surface rupture and slip distribution of the M w∼ 7.8, 14 November 2001 Kokoxili earthquake, Kunlun fault, northern Tibet, China, *Bull. seism. Soc. Am.,* **95**(5), 1970–1987.

Koper, K. D., Hutko, A. R. & Lay, T. 2011. Along-dip variation of teleseismic short-period radiation from the 11 March 2011 Tohoku earthquake (Mw 9.0), *Geophys. Res. Lett.,* **38**(21), doi:10.1029/2011GL049689.

Lancieri, M. & Zollo, A. 2009. Simulated shaking maps for the 1980 Irpinia earthquake, Ms 6.9: insights on the observed damage distribution, *Soil Dyn. Earthq. Eng.,* **29**(8), 1208–1219.

Lay, T. *et al.*, 2012. Depth-varying rupture properties of subduction zone megathrust faults, *J. geophys. Res.,* **117**(B4), doi:10.1029/2011JB009133.

Lay, T., Ammon, C. J., Kanamori, H., Koper, K. D., Sufri, O. & Hutko, A. R. 2010. Teleseismic inversion for rupture process of the 27 February 2010 Chile (Mw 8.8) earthquake, *Geophys. Res. Lett.,* **37**(13), doi:10.1029/2010GL043379.

1998. Gradient-based learning applied to document recognition, *Proceedings of the IEEE,* **86**(11), 2278–2324.

Lucy, L. B. 1974. An iterative technique for the rectification of observed distributions, *Astron. J.,* **79,** 745.
Madariaga, R. 1977. High-frequency radiation from crack (stress drop) models of earthquake faulting, *Geophys. J. Int.,* **51**(3), 625–651.
Mai, P. M *et al.* 2016. The earthquake-source inversion validation (SIV) project, *Seismol. Res. Lett.,* **87**(3), 690–708.
Melton, B. S. & Bailey, L. F. 1957. Multiple signal correlators, *Geophysics,* **22**(3), 565–588.
Mendoza, C. & Hartzell, S. H. 1988. Aftershock patterns and main shock faulting, *Bull. seism. Soc. Am.,* **78**(4), 1438–1449.
Meng, L., Ampuero, J. - P., Luo, Y., Wu, W. & Ni, S. 2012. Mitigating artifacts in back-projection source imaging with implications for frequency-dependent properties of the Tohoku-Oki earthquake, *Earth, Planets Space,* **64**(12), 5, doi:10.5047/eps.2012.05.010.
Meng, L., Inbal, A. & Ampuero, J. P. 2011. A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.,* **38**(7), doi:10.1029/2011GL048118.
Muirhead, K. 1968. Eliminating false alarms when detecting seismic events automatically, *Nature,* **217**(5128), 533.
Nakahara, H. & Haney, M M. 2015. Point spread functions for earthquake source imaging: an interpretation based on seismic interferometry, *Geophys. J. Int.,* **202**(1), 54–61.
Okal, E. A. 1992. A student's guide to teleseismic body wave amplitudes, *Seismol. Res. Lett.,* **63**(2), 169–180.
Okuwaki, R., Kasahara, A., Yagi, Y., Hirano, S. & Fukahata, Y. 2019. Back-projection to image slip, *Geophys. J. Int.,* **216**(3), 1529–1537.
Paszke, A. *et al.*, 2017. Automatic differentiation in pytorch.
Perol, T., Gharbi, M. & Denolle, M. 2018. Convolutional neural network for earthquake detection and location, *Sci. Adv.,* **4**(2), e1700578.
Ramachandran, P., Zoph, B. & Le, Q. V. 2017. Searching for activation functions, arXiv preprint arXiv:1710.05941.
Richardson, W. H. 1972. Bayesian-based iterative method of image restoration, *J. Opt. Soc. Am.,* **62**(1), 55–59.
Rost, S. & Thomas, C. 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 1008.
Rouet-Leduc, B., Hulbert, C., McBrearty, I. W. & Johnson, P. A. 2020. Probing slow earthquakes with deep learning, *Geophys. Res. Lett.,* **47**(4), e2019GL085870.
Ruiz, J., Baumont, D., Bernard, P. & Berge, C. 2011. Modelling directivity of strong ground motion with a fractal, k-square, kinematic source model, *Geophys. J. Int.,* **186,** 226–244.
Satriano, C., Dionicio, V., Miyake, H., Uchida, N., Vilotte, J. P. & Bernard, P. 2014. Structural and thermal control of seismic activity and megathrust rupture dynamics in subduction zones: lessons from the Mw 9.0, 2011 Tohoku earthquake, *Earth planet. Sci. Lett.,* **403,** 287–298.
Schweitzer, J., Fyen, J., Mykkeltveit, S., Gibbons, S. J., Pirli, M., Kühn, D. & Kværna, T. 2012. Seismic arrays, in *New manual of seismological observatory practice 2 (NMSOP-2),* pp. 1–80, Deutsches GeoForschungsZentrum GFZ.
Somerville, P. *et al.*, 1999. Characterizing crustal earthquake slip models for the prediction of strong ground motion, *Seismol. Res. Lett.,* **70**(1), 59–80.
Spudich, P. & Frazer, L. N. 1984. Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop, *Bull. seism. Soc. Am.,* **74**(6), 2061–2082.
Takemura, S. & Furumura, T. 2013. Scattering of high-frequency *P* wavefield derived by dense Hi-net array observations in Japan and computer simulations of seismic wave propagations, *Geophys. J. Int.,* **193**(1), 421–436.
Vallée, M., Landès, M., Shapiro, N. M. & Klinger, Y. 2008. The 14 November 2001 Kokoxili (Tibet) earthquake: high-frequency seismic radiation originating from the transitions between sub-Rayleigh and supershear rupture velocity regimes, *J. geophys. Res.,* **113**(B7), doi:10.1029/2007JB005520.
Vallée, M. & Satriano, C. 2014. Ten-year recurrence time between two major earthquakes affecting the same fault segment, *Geophys. Res. Lett.,* **41**(7), 2312–2318.
Walker, K. T. & Shearer, P. M. 2009. Illuminating the near-sonic rupture velocities of the intracontinental Kokoxili Mw 7.8 and Denali fault Mw 7.9 strike-slip earthquakes with global P wave back projection imaging, *J. geophys. Res.,* **114**(B2), doi:10.1029/2008JB005738.
Wang, D. & Mori, J. 2016. Short-period energy of the 25 April 2015 Mw 7.8 Nepal earthquake determined from back projection using four arrays in Europe, China, Japan, and Australia, *Bull. seism. Soc. Am.,* **106**(1), 259–266.
Wang, D., Mori, J. & Uchide, T. 2012. Supershear rupture on multiple faults for the Mw 8.6 Off Northern Sumatra, Indonesia earthquake of April 11, 2012, *Geophys. Res. Lett.,* **39**(21),.
Wang, D., Takeuchi, N., Kawakatsu, H. & Mori, J. 2016. Estimating high frequency energy radiation of large earthquakes by image deconvolution back-projection, *Earth planet. Sci. Lett.,* **449,** 155–163.
Wells, D. L. & Coppersmith, K. J. 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. seism. Soc. Am.,* **84**(4), 974–1002.
Xu, Y., Koper, K. D., Sufri, O., Zhu, L. & Hutko, A. R. 2009. Rupture imaging of the Mw 7.9 12 May 2008 Wenchuan earthquake from back projection of teleseismic P waves, *Geochem., Geophys., Geosyst.,* **10**(4), doi:10.1029/2008GC002335.
Yagi, Y., Nakao, A. & Kasahara, A. 2012. Smooth and rapid slip near the Japan Trench during the 2011 Tohoku-oki earthquake revealed by a hybrid back-projection method, *Earth planet. Sci. Lett.,* **355,** 94–101.
Yao, H., Shearer, P. M. & Gerstoft, P. 2012. Subevent location and rupture imaging using iterative backprojection for the 2011 Tohoku Mw 9.0 earthquake, *Geophys. J. Int.,* **190**(2), 1152–1168.
Yin, J. & Denolle, M. A., 2019. Relating teleseismic backprojection images to earthquake kinematics, *Geophys. J. Int.,* **217**(2), 729–747.
Zeng, Y., Aki, K. & Teng, T. L. 1993. Mapping of the high-frequency source radiation for the Loma Prieta earthquake, California, *J. geophys. Res.,* **98**(B7), 11 981–11 993.
Zhang, H. & Ge, Z. 2010. Tracking the rupture of the 2008 Wenchuan earthquake by using the relative back-projection method, *Bull. seism. Soc. Am.,* **100**(5B), 2551–2560.

APPENDIX

The procedure suggested by W2016 transforms the BP image conventionally calculated in the time at the source into a BP image in the apparent time of the reference station. In this new reference system, the HF emissions no longer suffer from the 'walking effect', however they are stretched or contracted along the time axis because of the directivity effect. In Fig. A1, we show, by way of example, the comparison between the BP image of a homogeneous rupture process calculated at the EU array following the conventional approach of Ishii *et al.* (2005) and the BP image of the same rupture process calculated using the W2016 reference station correction. We note that, in the reference station system, the HF radiation is stretched along the time axis. This happens because in this example the EU array is located in the antidirective position with respect to the direction of propagation of the rupture.

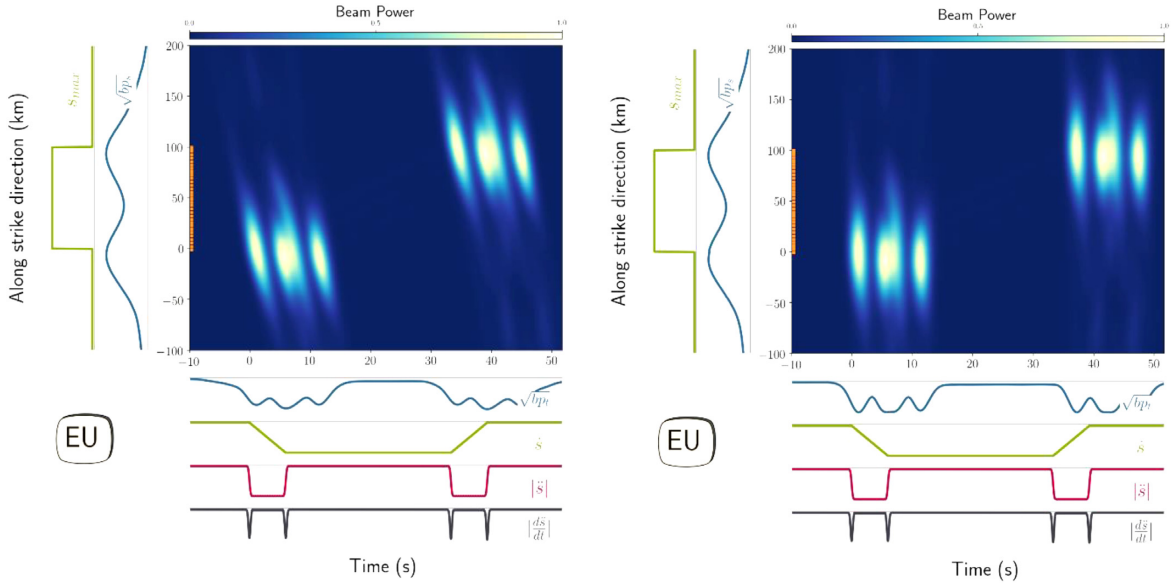

Figure A1. HF emissions radiated by a homogeneous rupture process retrieved by BP analysis using the conventional approach (on the top) and the W2016 approach (at the bottom). Time is on the *x*-axis, the along-strike direction is on the *y*-axis. The square root of the beam power integrated in time (in blue) and the maximum value of the slip-rate function (in green) are plotted in the left subplots. The square root of the beam power integrated in space (in blue), the slip-rate function (in green) and the absolute values of the slip-acceleration function (in red) and its time derivative (in black) are plotted in the bottom subplots. In the conventional BP analysis (BP image on the top), the HF emissions are tilted along the time axis. In the W2016 approach (BP image at the bottom), the HF emissions are stretched in time because of the directivity effect.

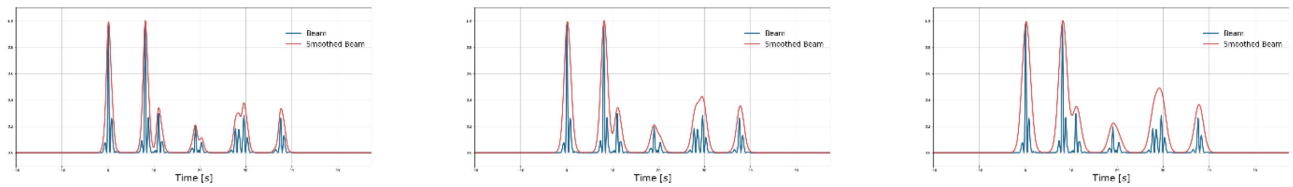

Figure A2. Comparison of beam and smoothed beam signals when using the three different values of σ chosen in the study (from left- to right-hand panels: $\sigma = 0.2$, $\sigma = 0.3$, $\sigma = 0.4$) to reduce BP artefacts.