

# Blind source separation of temporally independent microseisms

M. Meschede,<sup>1</sup> E. Stutzmann<sup>1</sup> and M. Schimmel<sup>1,2</sup>

<sup>1</sup>*Institut de Physique du Globe de Paris, CNRS-UMR 7154, Paris 75005, France. E-mail: stutz@ipgp.fr*

<sup>2</sup>*Institut of Earth Sciences Jaume Almera - CSIC, 08028 Barcelona, Spain*

Received 2018 September 28; in original form 2017 December 21

## SUMMARY

Secondary microseismic sources emit seismic waves over long time spans. Reoccurring signals with similar slowness and frequency therefore arrive at seismic arrays. Blind source separation techniques can be used to identify and isolate such reoccurring signals from other signals and from diffuse seismic noise. Along these lines, we use non-negative matrix factorization as blind source separation technique to decompose continuous seismic array records. We model the recorded energy as a mixture of a few components with static slowness–frequency and time-dependent amplitudes. Components and amplitudes are fitted to optimally explain the recorded seismic energy over time. These components represent secondary microseismic signals with quasi-static slowness–frequency vector and fluctuating amplitude. Each fitted component reveals the geographical origin (through the slowness–frequency vector) and time evolution of an active secondary microseism with high precision because it is separated from other signals and diffuse seismic noise. Furthermore, relative traveltimes can be automatically extracted for the signals that correspond to a specific component that can potentially be used in tomographic studies. We show two examples of seismic signals that were extracted with this technique, one focusing on *P* waves from the typhoons Goni and Atsani, and the other showing secondary microseism PKP signals from typhoon Glenda.

**Key words:** Body waves; Seismic noise; Wave propagation.

## 1 INTRODUCTION

Secondary microseisms are generated by ocean wave interactions that continuously emit seismic waves over a time span of many hours—unlike impulsive earthquakes. The origin and frequency of the emitted seismic waves vary only slowly with time, following the evolution of the ocean wavefield (Hasselmann 1963; Kedar *et al.* 2008; Stutzmann *et al.* 2009; Zhang *et al.* 2010; Arduin *et al.* 2011; Stutzmann *et al.* 2012; Farra *et al.* 2016; Liu *et al.* 2016). Their amplitude, however, fluctuates quickly, as a consequence of the random ocean wave interactions.

A seismic array often records a mixture of microseismic signals from different, simultaneously active source regions. Because of the amplitude fluctuations of these signals the mixture is constantly changing: sometimes the signal from one region is stronger, sometimes the signal from another, sometimes the signals have similar amplitude. In this paper, we show that this property allows to separate the signals belonging to each region from the mixtures using blind source separation.

Blind source separation (from now on BSS; e.g. Friedman *et al.* 2001; Cichocki *et al.* 2009; Comon & Jutten 2010) is a technique that is often used not only in acoustic studies (e.g. Parra & Spence 2000; Vincent *et al.* 2006) but also in geophysics (e.g. Ikelle 2007; Moni *et al.* 2012; Takahata *et al.* 2012; Liu & Dragoset 2013). In

our context, the purpose of BSS is to unmix mixtures of signals, retrieving the isolated signals (mixing *components*), and the amplitude with which they contribute to the mixture (mixing *amplitudes*). BSS cannot extract the components and amplitudes from a single mixture without extra information because this is a non-unique problem. However, it becomes applicable when many independent mixtures are used in the unmixing procedure. In this case a component can be identified as a distinct variation in the mixtures with independent amplitude.

Signal mixture are recorded by isolated seismic stations: the power spectrum of the recorded ground displacement, for example, can be seen as the mixture of spectra of different seismic signals from independent underlying physical processes: for example, a mixture of wind generated high-frequency signals and low-frequency volcanic tremor signals is observed over a continuous broad-band spectrum (Cabras *et al.* 2012). Over time, the relative amplitudes of these signals vary because the underlying physical processes are independent and change the mixture continuously. A spectrogram—the power spectrum as a function of time—therefore provides many independent mixtures of the spectra of the underlying signals. Sometimes the volcanic tremor signals are stronger, sometimes the wind generated signals are stronger. BSS can potentially retrieve the spectral shape (components) and amplitudes of the underlying independent signals that correspond to the independent

physical processes from these mixtures (Cabras *et al.* 2008, 2012, 2014).

Signal mixtures appear also in seismic array recordings: Each station of the array records the sum—a mixture—of simultaneously arriving signals. Each station in the array records a different mixture because the signals arrive at slightly different times, depending on the station position and the signal’s horizontal slowness (throughout the paper slowness refers to horizontal slowness). BSS can be used to separate signals with different slowness from these mixtures. Based on this idea, a BSS method called *DUET* has been used in different settings to unmix signals from the volcano Mt Etna (Moni *et al.* 2012), and also from secondary microseisms (Moni *et al.* 2013).

In this paper, we also examine signal mixtures that are recorded by a seismic array. Unlike *DUET*, that exploits the signal mixture variation amongst stations, we *exploit the signal mixture variation with time*. Conceptually, our technique is comparable to the time-varying mixture of signals at a single station that can be decomposed with BSS as outlined above and described in more detail in Cabras *et al.* (2008, 2012, 2014), except that we use seismic energy in slowness–frequency space provided by a seismic array instead of the power spectrum at an isolated station.

In particular, we take advantage of the almost static slowness and frequency of secondary microseisms that continuously generate signals with varying amplitude from the same source region. The time-dependent signal mixture that is recorded at an array can therefore be described as the sum of a few components with static frequency and slowness. BSS can identify such static components, and separate them from other seismic signals. The unmixed components can then be associated with microseisms and examined individually with higher precision than traditional methods allow. This can be seen as an example of unsupervised learning, where we extract signals that reappear over time at a seismic array without explicitly telling the algorithm which signals to extract.

We finally show that it is possible to extract relative traveltimes between stations from each of the unmixed components. Extracting relative traveltimes from secondary microseismic signals is interesting for potential tomographic applications. Zhang *et al.* (2010), Liu *et al.* (2016) and Nishida & Takagi (2016) have shown that traveltimes can be extracted from secondary microseismic signals. Our method allows to retrieve such measurements in a fully automatic way for multiple simultaneously active sources.

After a methodological introduction we demonstrate the main purposes of the method: (i) the extraction of the energy distribution of secondary microseismic sources in frequency–slowness space; (ii) the extraction of the temporal evolution of the amplitude of these sources; (iii) the fully automatic extraction of relative traveltimes between stations of the array from a day record, based on the automatically identified microseismic signals.

## 2 METHOD

### 2.1 Beam computation

In this paper, we look at seismic signals from the perspective of a seismic array: we examine the power spectrum of vertical ground displacement as a function of time  $t$ , frequency  $f$  and horizontal slowness vector  $\mathbf{s}$ , called the *beam*  $B(t, f, \mathbf{s})$ , that is recorded by the array. The beam  $B$  contains the cumulated energy of all signals that arrive at the array, in general from many seismic sources with different slowness and frequency.

To obtain the beam as a function of frequency, we use the complex, analytic bandpass filters that are shown in Fig. 1. The one-sided filter spectra span the secondary microseismic frequency band from 0.10–0.24 Hz. Each filter corresponds to a real and a complex wavelet in time domain. Convolution with these real and complex wavelets of the  $j$ th filter transforms each seismic data trace, sampled at discrete times  $t_i$ , into a complex trace  $a_j(t_i)e^{i\phi_j(t_i)}$ , with instantaneous phase  $\phi_j(t_i)$  and amplitude  $a_j(t_i)$ .

To obtain the beam  $B$ , we shift the instantaneous phase  $\phi_{jl}$  and amplitude  $a_{jl}$  of each trace  $l$  in time by the scalar product of a given horizontal slowness vector  $\mathbf{s}_k$  with a station position  $\mathbf{r}_l$  relative to the array centre. The corresponding time shift  $\delta t_{kl} = \mathbf{s}_k \cdot \mathbf{r}_l$  represents the time shift that a plane wave with slowness  $\mathbf{s}_k$  experiences when it propagates through the array.  $\delta t_{kl}$  can be separated in a sample shift  $\Delta t_{kl}$ , that corresponds to a multiple of the sampling interval of the seismic trace, and a subsample shift  $\epsilon_{kl}$ , that is smaller than the sampling interval of the seismic trace:

$$\delta t_{kl} = \Delta t_{kl} + \epsilon_{kl}. \quad (1)$$

The shifted instantaneous phase of filter  $j$  for slowness  $k$  and station  $l$  is then

$$\phi_{jkl}(t_i) = \phi_{jl}(t_i + \Delta t_{kl}) + 2\pi f_j \epsilon_{kl}. \quad (2)$$

Ignoring subsample time shifts, the shifted amplitude is

$$a_{jkl}(t_i) = a_{jl}(t_i + \Delta t_{kl}). \quad (3)$$

Instantaneous phase and amplitude of the traces of all  $N$  stations of the array are finally averaged to compute the beam:

$$B(t_i, f_j, \mathbf{s}_k) = \text{mean}_{l=1..N} \{ e^{i2\pi\phi_{jkl}(t_i)} \}^2 \text{median}_{l=1..N} \{ a_{jkl}(t_i) \}^2. \quad (4)$$

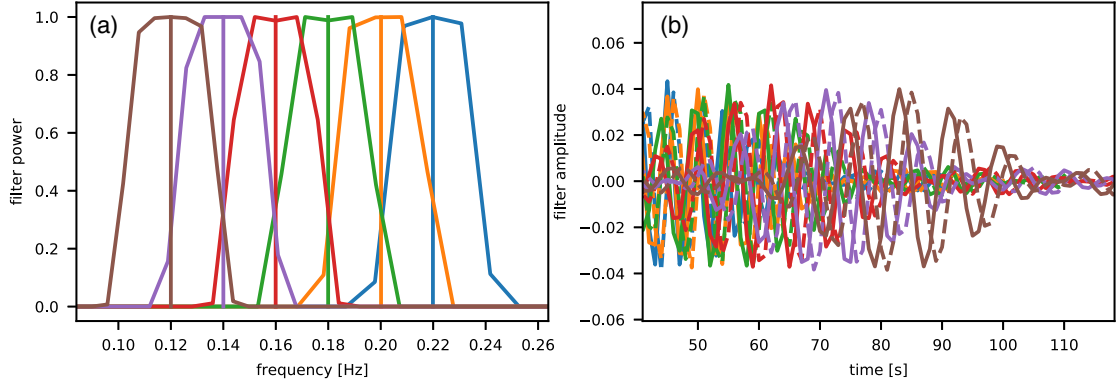
Averaging instantaneous phase and amplitude independently yields a more robust beam, similar to the phase-weighted stack method (Schimmel & Paulssen 1997). Because such a beam depends less on variations in amplitude over the array, it often provides better slowness resolution. We use the median instead of the mean for amplitude averaging because it further reduces the influence of amplitude outliers at isolated stations.

Amplitude variations over the array can, for example, be caused by local interference with other signals but the amplitudes of some seismic signals can show strong local amplitude amplification, such as in the case of PKIKP signals that travel through the Earth’s core. The major effect of such local amplitude variations is to reduce the slowness resolution of the array but not the time resolution.

The slowness resolution of the beam is naturally limited by its size and inter-station distance (aliasing). It can be described in terms of the array response function that indicates how the energy from a monochromatic plane wave spreads in slowness. Different beam-forming techniques exist that can reduce this spread in slowness and improve the beam resolution (Gal *et al.* 2016). The BSS technique that is presented in this paper is complementary and can potentially be combined with these techniques. It exploits the time dependence of the beam to separate sources instead of improving the slowness resolution of the beam.

## 3 TIME AVERAGING VERSUS BLIND SOURCE SEPARATION

The beam  $B$  as a function of time is usually noisy and difficult to examine directly. In the case of microseisms, averaging the beam energy  $B$  over many time steps, comprising a total time length  $T$ , is



**Figure 1.** (a) Analytic filters that cover the secondary microseismic band. The filters are analytic, with zero amplitude at negative frequencies. (b) The imaginary part (dashed lines) and real part (solid) of the wavelet that correspond to a filter shown only for a short time window. The wavelet time length is proportional to the lowest frequency of each filter (finite impulse response).

one possibility to extract relevant signals:

$$\bar{B}(f_j, \mathbf{s}_k) = \text{mean}_{i=1..T} \{B(t_i, f_j, \mathbf{s}_k)\}. \quad (5)$$

Averaging  $B$  over time enhances long-duration signals with consistent slowness and frequency, such as microseisms. Short-duration signals, such as earthquakes, or signals with diffuse slowness and frequency (‘noise’) are attenuated. Simultaneously active sources that are clearly separated in slowness or frequency can also be examined individually.

Time averaging has a significant problem: weaker signals can easily be masked by stronger ones. This even happens when both signals are well separated in slowness because leaked energy from one source can still be stronger than energy from significantly weaker secondary sources at different slowness (this leakage is related to the array response function). Because secondary microseism amplitudes vary over orders of magnitude, this happens often.

Nevertheless, the amplitude of microseisms fluctuates also rapidly in time. Sometimes, when the stronger source has very low amplitude, weak sources can be studied with higher precision. Time averaging dumps this information, whereas BSS exploits these amplitude fluctuations in time to extract the strong and the weak sources. It explains the beam as the mixture of several components in contrast to a single average.

### 3.1 Blind source separation with non-negative matrix factorization

We now formalize and illustrate BSS using non-negative matrix factorization (NNMF) for the case of a synthetic beam  $B$  that could have been observed with a seismic array.

We define a microseismic *source*  $S$  to be the product of a *normalized energy distribution*  $E$  in frequency  $f$  and slowness  $\mathbf{s}$  space and an *amplitude*  $A$  that varies in time:

$$S(t, f, \mathbf{s}) = A(t)E(f, \mathbf{s}), \quad (6)$$

$A$  and  $E$  are non-negative, because they represent energies. Additionally,  $E$  is normalized such that its maximum is one.

Now consider a beam  $B$  that is composed of  $N$  sources  $S_m$  ( $m = 1..N$ ) with amplitudes  $A_m(t)$  and energy distributions  $E_m(f, \mathbf{s})$ . Assuming that the source signals that are recorded at the array are sufficiently uncorrelated, that is, that there is no particular phase relation between the signals, the beam  $B$  can be written as the sum

of the individual sources:

$$B(t, f, \mathbf{s}) \approx \sum_{m=1}^N A_m(t)E_m(f, \mathbf{s}). \quad (7)$$

The beam  $B$  at time  $t$  is a mixture (superposition) of the *components*  $E_m(f, \mathbf{s})$  with *coefficients*  $A_m(t)$ . We are going to show that it is possible to find components  $E_m(f, \mathbf{s})$  and amplitudes  $A_m(t)$  that explain the beam  $B$  in an optimal way with a least-squares inversion. The inversion is particularly stable if the beam is known over a long-enough time span and if the amplitudes vary sufficiently, that is, if many independent mixtures of the components are known.

Consider the example shown in Fig. 2. In this synthetic example, three monochromatic sine waves constitute a hypothetical seismic wavefield that is recorded by a seismic array in Alaska.

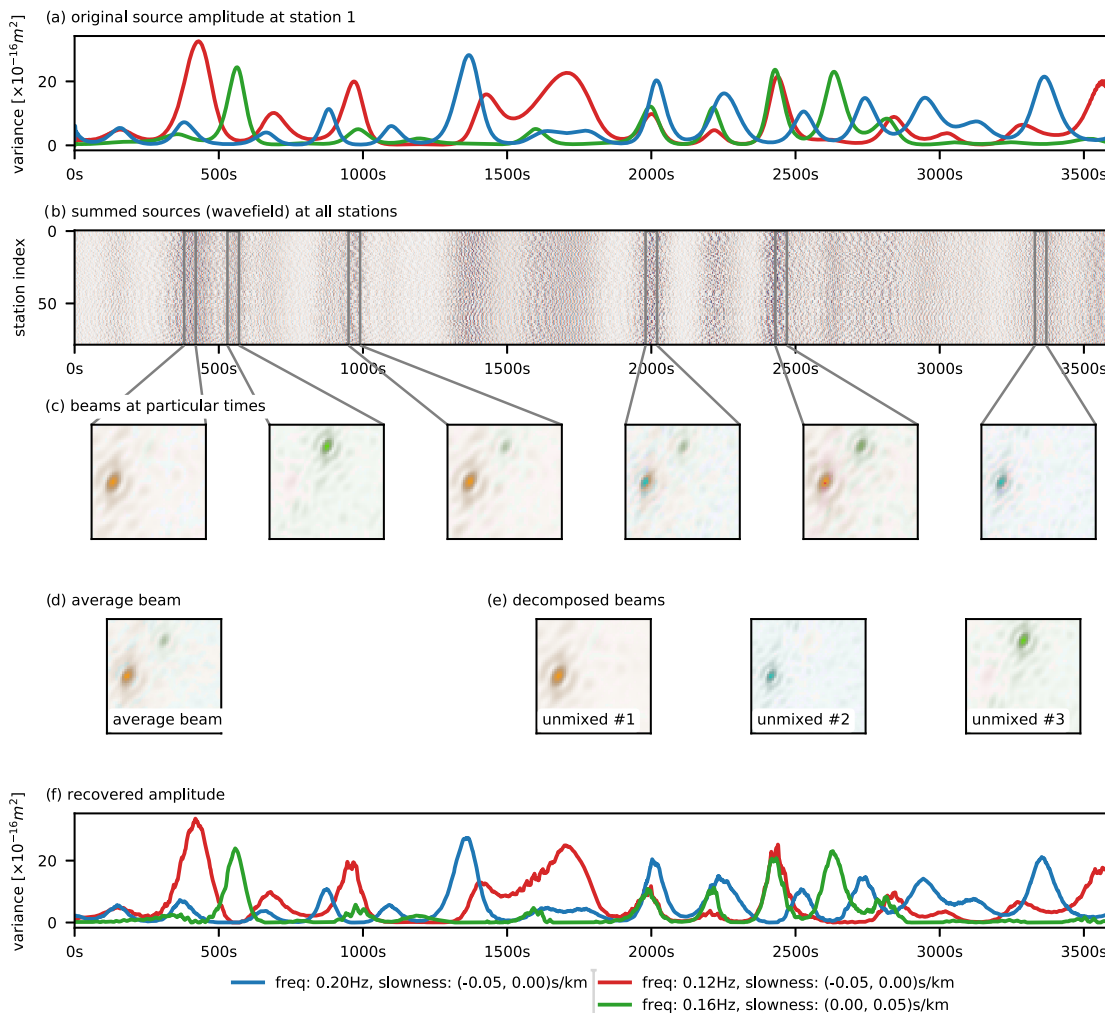
The three signals  $s_m(t)$  have (signal #0) slowness  $\mathbf{s}_0 = (-0.05, 0.00) \text{ s km}^{-1}$  and frequency  $f_0 = 0.12 \text{ Hz}$ , (signal #1) slowness  $\mathbf{s}_1 = (-0.00, 0.05) \text{ s km}^{-1}$  and frequency  $f_1 = 0.16 \text{ Hz}$ , and (signal #2) with slowness  $\mathbf{s}_2 = (-0.053, 0.00) \text{ s km}^{-1}$  and frequency  $f_2 = 0.20 \text{ Hz}$ . Such slownesses and frequencies correspond to typical  $P$  waves from secondary microseisms. The signal amplitudes  $a_m(t)$  are drawn every 100 s from a log-uniform distribution between 1 and  $10 \times 10^{-8} \text{ m}$  and smoothly interpolated in between. In the same way, random phases  $\phi_m(t)$  are drawn every 100 s from a uniform distribution between 0 and  $2\pi$  and interpolated smoothly in between.  $a_m(t)$  and  $\phi_m(t)$  represent the rapidly fluctuating amplitudes and phases of secondary microseismic signals:

$$s_m(t) = a_m(t) \sin(2\pi f_m t + \phi_m(t)). \quad (8)$$

The time-dependent variance, or power (energy per time), of the sine wave signals as measured by the beam  $B$  is  $A_m(t) = (a_m(t)/2)^2$  and is shown in Fig. 2(a).

The synthetic wavefield, as observed by the seismic array, is the sum of the three individual signals that are shifted by the time necessary to reach each station. For each station  $l$  with position  $\mathbf{r}_l$  relative to the array centre, the signals are therefore shifted by  $\delta t_{ml} = \mathbf{s}_m \cdot \mathbf{r}_l$ . A synthetic 1 hr record  $u_l(t)$  is then generated for each station of the array by summing the shifted individual signals  $s_m(t)$ :

$$u_l(t) = \sum_{m=1}^N s_m(t - \delta t_{ml}). \quad (9)$$



**Figure 2.** Illustration of the separation of temporally independent sources. Panel (a) shows the amplitudes of three monochromatic synthetic sources. The synthetic sources have distinct slowness and/or frequencies and slowly varying random phase. Panel (b) shows the synthetic wavefield that has been constructed for seismic array stations in Alaska. Different sources appear with different patterns in the wavefield because they have different slowness and frequency. Panel (c) shows some time samples of the beam that corresponds to the wavefield as recorded by the array. The beam clearly recovers the signal amplitude and slowness but it is a mixture of all sources. Panel (d) shows the time average of the beam. The source mixture is retrieved but the original sources are lost in the averaging process. Panel (e) shows the decomposed beam, that is, three components that optimally explain the source mixture that is observed in the full time-dependent beams shown in panel (c). Finally, panel (f) shows the recovered amplitudes of each signal over time that resemble to a high degree the original ones shown in panel (a).

The wavefield  $u_l(t)$  is shown in Fig. 2(b), as recorded by the whole array. The distinct patterns that correspond to the different signals can be visually identified.

Fig. 2(c) shows time samples of the beam  $B$  as a function of  $\mathbf{s}$ . The beam maximum over all frequencies at a particular slowness (from now on just *beam amplitude*) is shown with colour intensity. The frequency of the beam maximum (from now on *dominant frequency*) is shown with colour hue. For example, the red colour of the first beam that is shown corresponds to the frequency  $f = 0.12$  Hz of source #1, the green colour of the second beam to frequency  $f = 0.16$  Hz of source #2, and the blue colour of the last beam to frequency  $f = 0.20$  Hz of source #3. The beam is sometimes dominated by a single source and sometimes by a mixture of evenly strong sources, corresponding to the amplitudes in Fig. 2(a).

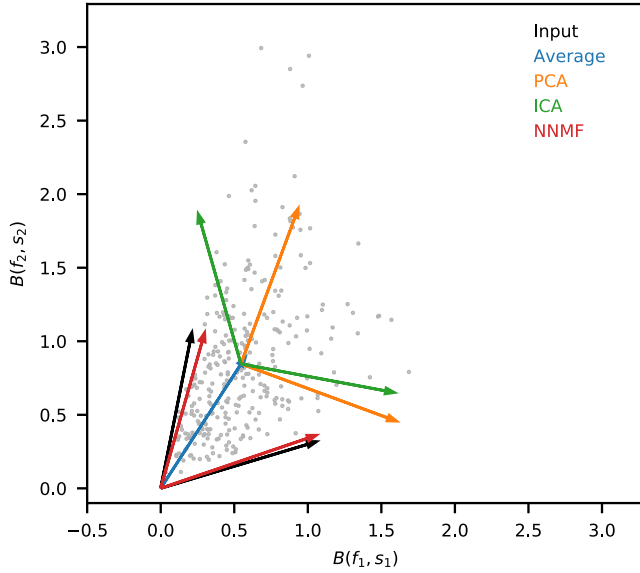
The time average of the 1 hr beam is shown in Fig. 2(d). Although the low-frequency source #1 is well resolved in this beam, source #2 is more difficult to identify and source #3 vanishes completely behind the stronger source #1. This illustrates that the time average

does not resolve secondary sources very well. Using BSS, it is possible to retrieve the original sources from the time-dependent beam.

We now transform the continuous eq. (7) into its discretized form. The continuous beam  $B(t, f, \mathbf{s})$  becomes the *beam matrix*  $B_{ik}$  where index  $i$  corresponds to a time sample  $t_i$ , and index  $k$  to a slowness–frequency pair  $(f_k, \mathbf{s}_k)$ . The beam matrix can be seen as the set of images as plotted in Fig. 2(c). BSS extracts common structures in this set of images. In other words, the whole set of beam time samples is constructed as the linear combination of a few components. In discrete form, eq. (7) can be written as

$$B_{ik} \approx \sum_{m=1}^N A_{im} E_{mk}. \quad (10)$$

To find the time-dependent amplitudes  $A$ , and the static energy distributions  $E$  for a given number  $N$  of components, we minimize the following least-squares error under the condition that  $A_{im}$  and



**Figure 3.** Illustration of PCA (orange), ICA (green) and NNMF (red). Each grey point is a random superposition (mixture) of two input components (black). NNMF is able to recover the input components because it takes into account two main properties, non-negativity and non-orthogonality, of the point cloud that are known *a priori*.

$E_{mk}$  are non-negative:

$$\underset{A_{im} > 0, E_{mk} > 0}{\text{minimize}} \quad \sum_{ik} |B_{ik} - \sum_m A_{im} E_{mk}|^2. \quad (11)$$

Eq. (11) with the imposed constraints is identical to an NNMF. As a non-negative equivalent of principal component analysis (PCA), NNMF is routinely used in many different domains, and many algorithms that solve the constraint minimization problem of eq. (11) exist. In this paper, we use an algorithm implemented in the open software library scikit-learn (Cichocki & Phan 2009; Févotte & Idier 2011).

Fig. 2(e) shows the result of the unmixing procedure for the synthetic example. The algorithm correctly extracts the position of origin and the frequency of the three input signals. The power of the unmixed sources, shown in Fig. 2(f), also closely resembles the input power of the synthetic signals.

### 3.2 Comparison of different blind source separation techniques

NNMF is particularly well suited for the case of non-negative beam values. We illustrate here its advantages compared to other techniques for the case of a continuously recorded beam on a conceptual synthetic example shown in Fig. 3.

Consider the beam  $B = [B_1, B_2]$ , measured over time at two distinct slowness and frequency values ( $B_1 = B(t_i, f_1, s_1)$  and  $B_2 = B(t_i, f_2, s_2)$ ) over 300 time steps. Each grey scattered point in Fig. 3 corresponds to the beam at a particular time. The  $x$ -component of point  $i$  corresponds to the beam at time  $t_i$ , frequency  $f_1$  and slowness  $s_1$ . The  $y$ -component corresponds to the beam at time  $t_i$ , frequency  $f_2$  and slowness  $s_2$ . The particular values of the beam were generated as mixture of two components  $E_1 = [1, 0.3]$  and  $E_2 = [0.2, 1]$  (shown with black arrows, called input in the legend). The component amplitudes  $A_1(t_i)$  and  $A_2(t_i)$  are drawn for every time step from a  $\chi^2$  distribution that has twice the variance for source  $E_1$  than for source  $E_2$ . Component 1 can be seen as having dominant

frequency and slowness  $f_1, s_1$  and leaks some energy (value 0.3) into component 2 with frequency and slowness  $f_2, s_2$ . Component 2 on the other hand leaks energy (value 0.2) into source  $E_1$ . The components are non-orthogonal because of the energy leakage. The synthetic beam  $B$  that generates the point cloud can be written as

$$B(t_i, f, s) = A_1(t_i)E_1(f, s) + A_2(t_i)E_2(f, s). \quad (12)$$

The distribution of this point cloud can now be described using different methods. We illustrate in Fig. 3 time averaging, PCA, independent component analysis (ICA) and NNMF [a general introduction to these methods can be found in Friedman *et al.* (2001)].

Time averaging (blue arrow) extracts a mixture of both sources, but is dominated by the stronger source  $E_1$ . If the input sources were orthogonal, the mean amplitude of source  $E_1$  could have been approximated by the value  $f_1, s_1$  coordinate of the mean, and the amplitude of source  $E_2$  by the  $f_2, s_2$  coordinate of the mean. When source energy overlaps, as in real observations, this is not possible anymore.

PCA (implementation details in Halko *et al.* 2011) describes the point cloud with two orthogonal components that are the directions of the strongest variation of the data in a least-squares sense, ordered from the strongest to weakest. PCA implicitly assumes that the data is multivariate Gaussian distributed around the mean. This is the reason why we draw the PCA component arrows from the mean and not from the origin in Fig. 3. The PCA does not use the main properties about the data set that are known *a priori*: first, beam components can be non-orthogonal, for example, due to leaked energy or noise. Second, beam components are positive definite. Because it lacks this crucial information, the PCA interprets the point cloud distribution inaccurately, and extracts, for example, one negative component (downward pointing arrow).

ICA (implementation details in Hyvärinen & Oja 2000) does not constrain components to be orthogonal, but also allows for negative components. The corresponding ICA components (green arrows) are even less close to the input components than the components that were found by the PCA.

Finally, we use NNMF (implementation details in Cichocki & Phan 2009; Févotte & Idier 2011) to extract non-orthogonal, positive definite components (red arrows). NNMF extracts components that are very close to the input components because it exploits the *a priori* known properties, non-negativity and non-orthogonality, of the point cloud in contrast to PCA and ICA.

In contrast to the time average, BSS methods, in particular NNMF, allow to extract the relevant components from a beam more accurately.

### 3.3 Source number and regularization

The number of independent components  $N$  determines the capacity of our model to represent the beam  $B$ . As in other inversions, there is no simple rule for how it should be chosen. In contrast to the synthetic example in Fig. 2, the number of active microseisms is unknown and often not well defined in reality: Real signals from secondary microseisms do have some variations in frequency and slowness, and could be represented by one component that averages the signal energy distribution, or by several components that describe the slowness–frequency variations more accurately. The beam  $B$  also contains ‘noise’, often from energy that leaks from surface waves to other seismic arrivals. Seismic noise often has consistent slowness and frequency and will therefore usually be represented by one or several components in our model. We call

a component a *noise component* if its energy is distributed evenly over slowness.

Roughly the component number balances the two common extremes of many inverse problems: a low component number is able to fit the beam with high bias and low variance with respect to noise or source variations (underfitting). In other words, it provides stable components that are little influenced by noise or source variations but only recovers the most important structures of the beam. On the other hand, a high component number fits the beam with low bias but high variance (overfitting): weak sources are recovered but the components are strongly influenced by noise and source variations. In practice, we search for a middle ground to fit the dominant active sources in the beam with few stable components.

Another possibility to tune the model is the introduction of L1 or L2 regularization. L1/L2 regularization adds an adjustable costs to the elements of the amplitude, the component, or both matrices in eq. (11). Both regularization types favour less extreme components and amplitudes in exchange for a less accurate fit of the beam. This is particularly useful to avoid fitting noise. L1 regularization has a tendency to favour sparse components and amplitude that have zero values, whereas L2 regularization has a tendency to equalize component and amplitude values (see Appendix A). In practice, there is no ideal damping parameter as in other inverse problems. The regularization parameter has to be tuned for each problem. However, we have found that very little L1/L2 regularization is required if we fit a lot of data with few components. In the examples that are shown below, regularization has therefore only a minor influence, as long as no extreme values are used.

### 3.4 Suppression of noise using sample weights

In many cases, two distinct types of components can be distinguished: unmixed source components, whose beam energy is concentrated around a specific slowness value, and noise components, whose beam energy is not concentrated. Typically, we are mainly interested in the source components, and not particularly in the noise components.

One possibility to remove noise components is to weight the beam time samples with a measure that indicates how strongly the energy concentrates in slowness–frequency space. A good measure of the time-dependent noise level of the beam is the median of the beam energy over all slownesses and frequencies. If the beam  $B$  is divided at each time step by the noise level at this time step, it is transformed from units of energy into signal to noise ratio units.

### 3.5 Relative traveltimes

The complex filters provide the instantaneous phase of each trace at every time step, which permits to extract phase shifts between stations with respect to the expected phase of a plane wave with given slowness. Such phase shifts are directly related to traveltime shifts that can be caused by heterogeneous structure in the Earth, which is why they are highly interesting for seismological studies (Zhang *et al.* 2010).

If a component corresponds to a plane wave, it can be represented by the frequency  $f_m$  and slowness  $\mathbf{s}_m$  with maximum component energy. Whenever the component amplitude is high, we can expect that such a plane wave appears in the waveforms. To extract phase and traveltime shifts of the plane wave that correspond to the component  $m$ , we start with the phase traces  $\phi_{lm}(t_i)$  for stations  $l$  and

component frequency  $f_m$  that are aligned to the component slowness  $\mathbf{s}_m$ , as described in eq. (2).

The shifted phases are distributed around the mean shifted phase of the stations  $l$  at times when such plane waves arrive with the component slowness  $\mathbf{s}_m$ . This mean shifted phase can be computed for each time step as

$$\bar{\phi}_m(t_i) = \arg \left\{ \sum_{l=1}^N e^{i\phi_{lm}(t_i)} \right\}. \quad (13)$$

When the mean shifted phase of all stations is subtracted from the shifted phases of each station, we obtain relative phase shifts for each time step. These relative phase shifts are now concentrated around zero whenever a coherent plane wave is observed with the frequency and slowness of the component.

When the plane wave arrives with a slightly different slowness  $\Delta\mathbf{s}(t_i)$ , a linear trend  $2\pi f_m \Delta\mathbf{s}(t_i) \cdot \mathbf{r}_l$  remains in the phases. We therefore compute a slowness correction at each time step using a linear least-squares regression:

$$\Delta\mathbf{s}(t_i) = \text{minimize} \left\{ \sum_l |2\pi f_m \Delta\mathbf{s}(t_i) \cdot \mathbf{r}_l - (\phi_{lm}(t_i) - \bar{\phi}_m(t_i))|^2 \right\}. \quad (14)$$

The phase that corresponds to this slowness correction is subtracted as well from the relative phase shifts. We finally obtain the slowness corrected relative phases  $\Delta\phi_{lm}(t_i)$  for component  $m$ , each station  $l$  and time  $i$ :

$$\Delta\phi_{lm}(t_i) = \phi_{lm}(t_i) - \bar{\phi}_m(t_i) - 2\pi f_m \Delta\mathbf{s}(t_i) \cdot \mathbf{r}_l. \quad (15)$$

These phases are distributed around zero when the component amplitude is high. The deviations from zero that are consistent over time for a given station correspond to a time shift. We average the vector that corresponds to a phase in the complex plane over the whole day, weighted such that times where the component has high amplitude compared to the other signals contribute more to the average than times where the relative component amplitude is weak. In particular we use  $w_m(t_i) = (A_m(t_i)/\sum A_m(t_i))^\alpha$  as weights, with  $\alpha = 2$ , which provides us with a complex coefficient  $c$ :

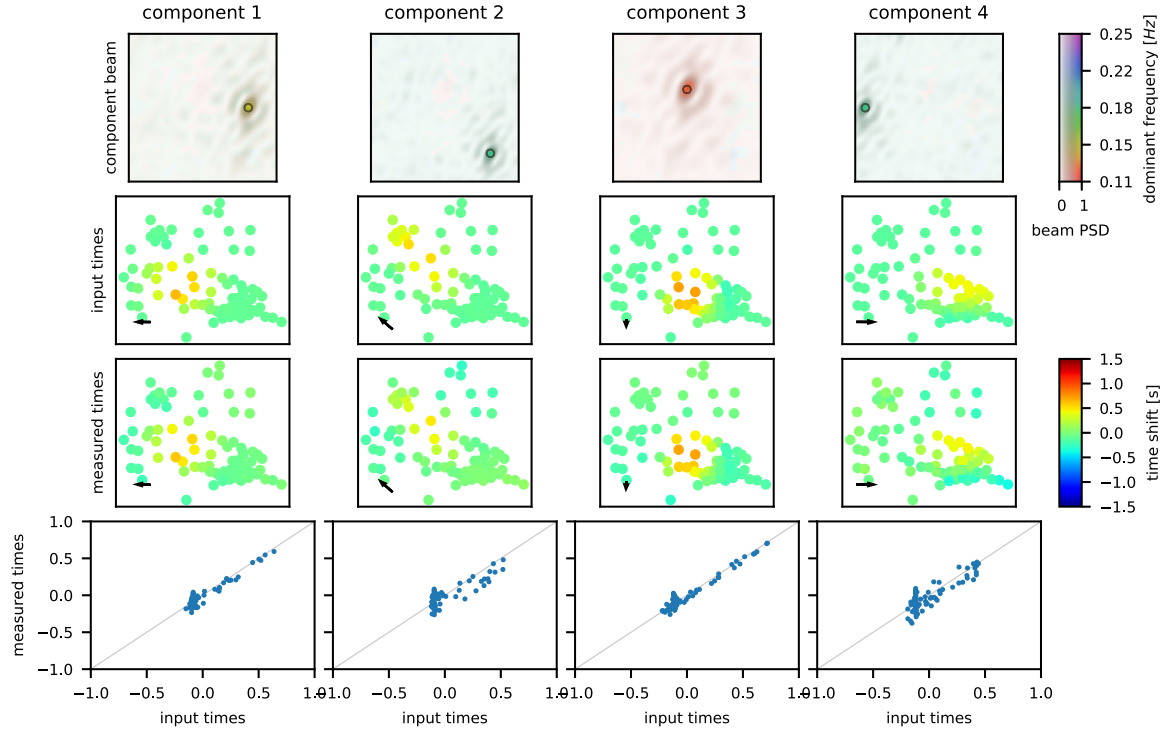
$$c_{lm} = \frac{1}{\sum_i w_m(t_i)} \sum_{i=1}^T (w_m(t_i) e^{i\Delta\phi_{lm}(t_i)}). \quad (16)$$

The phase of  $c_{lm}$  can be converted to traveltimes:

$$\Delta t_{lm} = \arg\{c_{lm}\} / (2\pi f_m). \quad (17)$$

The amplitude  $|c_{lm}|$  is one, if the relative phases are perfectly consistent over the whole day. This is usually not the case because of other signals that interfere and degrade the observations. The value  $|c_{lm}|$  provides a measure of the consistency of our measurement, and is therefore from now on called *consistency* of the measured traveltime shift at station  $l$  for component  $m$ . On high quality days with strong signals, we obtain consistencies of about  $|c_{lm}| = 0.5$ , on lower quality days, consistencies of  $|c_{lm}| = 0.1$ – $0.2$  are more common.

A synthetic test that shows this procedure and its quality is presented in Fig. 4. We have generated four synthetic monochromatic signals with varying slowness and frequency using the same approach as in Fig. 2 to determine their amplitude and phase. We generate 2 hr of synthetic ground displacement waveforms for each station of a seismic array in Alaska, which is composed of the networks TA and AK.



**Figure 4.** Synthetic benchmark for traveltimes extraction. A 2 hr synthetic wavefield (recorded at the Alaska array) consisting of a mixture of four monochromatic signals with different frequencies and slownesses is unmixed. The four unmixed components are shown in Row 1. The input slowness of the synthetic signals is indicated with black circles and very well recovered through the unmixing procedure. In addition to the plane wave time shifts expressed by the signal slowness and station position  $r_i \cdot s_i$ , time shift perturbations were added to each signal that resemble a slow velocity heterogeneity at the centre of the array. The time shift perturbations that have been added to each signal are shown in Row 2. Row 3 shows how well these time shifts are recovered automatically for each component. Row 4 shows histograms of the input and the measured time shifts.

In addition to the normal time shift that a plane wave experiences when propagating with a certain slowness, we add a station-dependent time shift to each signal that corresponds to a slow velocity zone in the centre of the array. The exact setup of the slow velocity zone is not of importance, but the heterogeneity slows the wave front, leading to realistic time shifts of  $< 1$  s at stations that lie behind the heterogeneity, from the perspective of a signal’s direction of propagation.

Fig. 4, Row 1 shows the four components (from column 1–4) that have been extracted using NMF from the time-dependent beam. They correspond very well to the input slowness (indicated with a black circle) and frequency of the four input signals. Row 2 shows the input traveltimes, that are positive (red colour) behind the heterogeneity at the centre of the array from the perspective of the components direction of propagation that is indicated with a black arrow. Note that the signals arrive from below, which is the reason why only stations that are close to the centre are affected. Row 3 shows the measured traveltimes (2 hr average) at each stations for each component. The measured traveltimes resemble very closely the input ones, which is confirmed by the scatter plots in Row 4.

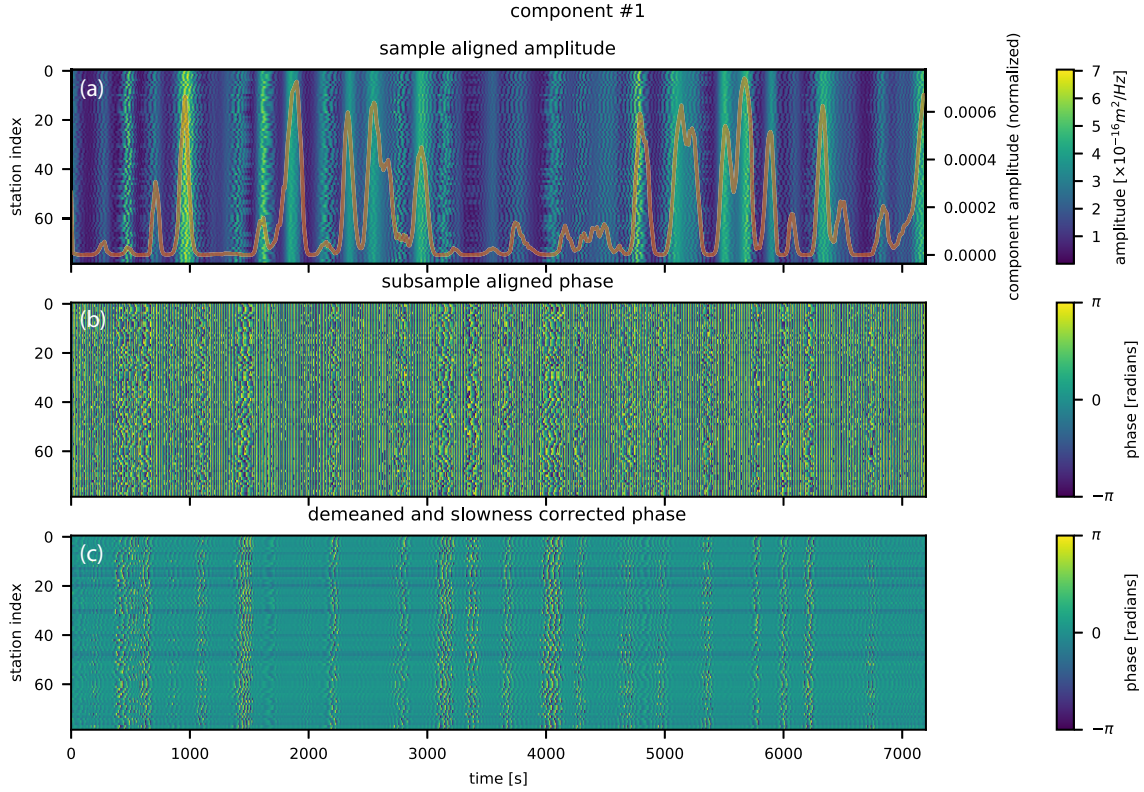
The steps of traveltimes extraction are illustrated in detail for component #1 in Fig. 5: Figs 5(a) and (b) show the amplitudes  $a_{lm}(t_i)$  and instantaneous phases  $\phi_{lm}(t_i)$  of all stations for the frequency  $f_m$ , and align to the slowness  $s_m$  with maximum component energy, respectively. The component amplitude that was obtained with NMF is shown in Fig. 5(a) (orange line). Whenever the component amplitude is high, the trace amplitudes and phases in Figs 5(a) and (b) align over stations (vertically). After the mean of the aligned phases has been subtracted (Fig. 5c), phase shifts that are consistent

over time (horizontal patterns) appear and can be converted into traveltimes shifts.

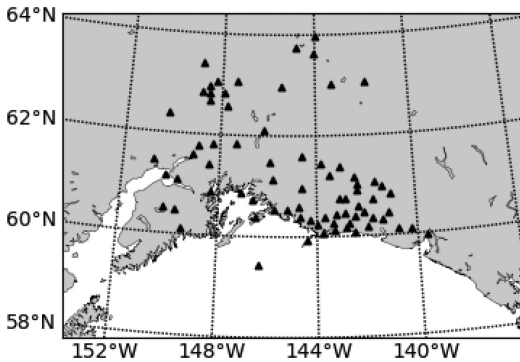
The consistencies  $|c_{lm}|$  of all stations and components are  $> 0.9$  in this synthetic example. This means that the measured traveltimes shifts are consistent at times when the component amplitude is high and therefore robust. The consistency value is very important because it allows to objectively estimate the quality of the measurement. Subjective parameters that determine the quality of this measurement are the component number and the filters that were chosen initially. The traveltimes shifts can only be as precise as the frequency resolution of the filter: a filter bandwidth of 0.02 Hz implies that the phase is transformed to traveltimes with a relative error of  $< 0.01/0.12 = 8$  per cent, which is larger for the low frequency wavelet at 0.12 Hz (see eq. 17). The measurement is based on the assumption that an unmixed component corresponds to a plane wave. Traveltimes will be inaccurate for unmixed components that are mixtures of plane waves with different slowness and frequency or that contain noise only. In such a case, the consistency values will go down. The exact component number of the unmixing procedure is not important as long as it isolates plane waves with specific slowness and frequency.

## 4 EXAMPLES

We illustrate the procedure with two real data examples; each of them decomposes the beam energy of a full day. As in the synthetic traveltimes test shown in Fig. 4, we use vertical ground motion data that were recorded by a seismic array in Alaska consisting of 64 stations (AK and TA seismic networks), shown in Fig. 6



**Figure 5.** Stepwise extraction of relative traveltimes from the unmixing component #1 of the 2 hr synthetic data in Fig. 4. Panel (a) shows the amplitude of the traces, corresponding to the frequency  $f_m$  with maximum component energy. The traces are sample aligned according to the slowness  $s_m$  with the main component energy. Superimposed, with an orange line, is the inferred component amplitude that was found with non-negative matrix factorization. Whenever the component amplitude is high, the amplitudes align vertically, as expected. The three other signals can also be seen when the trace amplitude is high but does not align as well. Panel (b) shows the equivalent plot for the subsample aligned instantaneous phases of the traces, again for the filter with frequency  $f_j$ . It is less obvious to see, but again, whenever the component amplitude (orange line in a) is high, the phases align. This becomes more obvious in panel (c) where we subtracted the mean of the phase and corrected for any remaining time shift that can be explained by adjusting the slowness. Now, phase shifts appear that are constant for particular stations (horizontal features in the plot). These phase shifts can be observed in particular when the component amplitude is low, and they cannot be observed when the component amplitude is high. Therefore, we compute the time averaged phase shift from panel (c) as the weighted mean over time of the phases with weights that reflect the component amplitude.



**Figure 6.** Map of the stations of the ‘Alaska’ virtual array that is used in this paper.

(see Acknowledgements section for detailed references). The raw recordings of each station are transformed from instrument response to vertical ground displacement with flat frequency response in the secondary microseismic frequency band from 0.05–0.30 Hz and a sampling rate of 1 Hz.

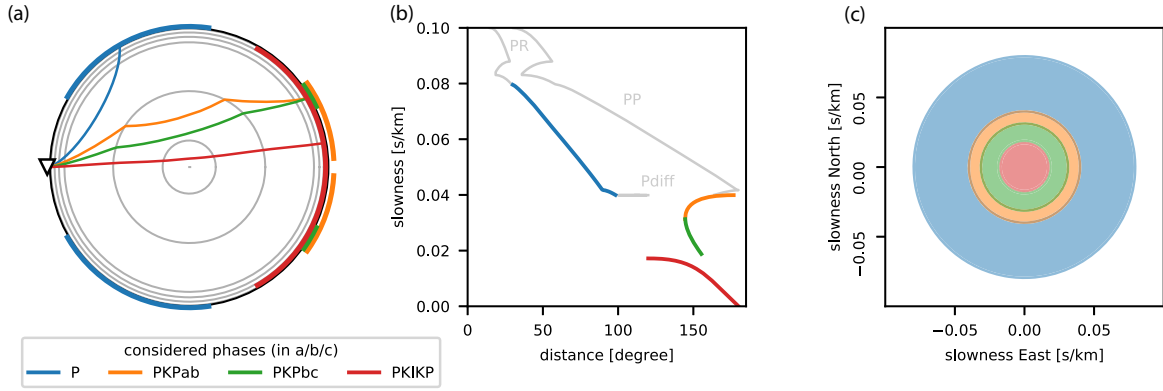
The beam  $B$  as a function of time, frequency and slowness is computed in 20 s intervals (4321 samples for one day), using the filters with bandwidth 0.02 Hz and central frequencies 0.12, 0.14,

..., 0.22 Hz shown in Fig. 1, on a Cartesian slowness grid that spans the west–east and south–north components from  $-0.08$  to  $0.08$  s  $\text{km}^{-1}$ . The response function of the array is shown in Appendix B.

Beam energy at a particular slowness can be associated with different seismic phases: the strongest amplitude phases on the vertical component in the considered slowness–frequency range are the  $P$  waves that are shown in Fig. 7(a). Phases that are reflected multiple times, such as PP, are not shown. Each ray path is associated with a particular horizontal slowness, and to a distance between source and receiver. The relation of horizontal slowness with distance is shown in Fig. 7(b) for each phase. Fig. 7(c) shows the slowness ranges that are associated with each seismic phase in the beam  $B$ . Energy with a particular slowness and azimuth can be retraced to its origin using the slowness–distance relation of Fig. 7(b), under the assumption that it corresponds to the phases that are listed in Fig. 7(c).

Prior to unmixing, the beam is transformed to signal to noise ratio units, as explained in Section 3.4. We define the time-dependent noise level as the median beam energy over all slownesses and frequencies in the considered ranges from  $-0.08$  to  $0.08$  s  $\text{km}^{-1}$  and 0.12 to 0.22 Hz at the respective time. This emphasizes components that have concentrated energy in slowness–frequency space. We then unmix the beam using NMF with five components. Because we use few components and have many time samples of the beam,





**Figure 7.** Illustration of the phases that we examine in this paper. Panel (a) shows ray paths and distance ranges of the phases, (b) their slowness–distance relation, (c) the slowness regions that correspond to each phase in the beam. The phases that are shown are: the P phase (blue), that is, a  $P$  wave that travels mostly through the lower mantle; the PKP phases that travel through the outer core and split into PKPab (orange) and PKPbc (green) with slightly different ray paths; the PKIKP phase that passes through the inner core. Note that the distance ranges of PKP and PKIKP overlap, which is the reason why a secondary microseism in these regions can be observed with both phases. Some other phases in the examined slowness ranges with significant amplitude on the vertical component are indicated in grey in panel (b): PR (blue colour) are  $P$  waves that travel short distances and are heavily influenced by upper-mantle heterogeneities; The PP phase has similar slowness than P but comes from twice as far, fortunately it is usually weaker in amplitude than the direct P phase; Pdiff is diffracted around the core and can be observed at a slowness that corresponds to a large distance P phase, or an almost antipodal PKPab phase.

no additional regularization is required. The resulting unmixed components are normalized such that their maximum is one (division by the unnormalized component’s maximum) and the mixing amplitudes are adapted accordingly (multiplication with unnormalized component’s maximum). We use the average mixing amplitude of each component to define its average strength in signal to noise ratio units. Component amplitudes are transformed back from signal to noise ratio units to displacement energy by multiplication with the time-dependent noise level.

The day-averaged wave interaction pressure, computed from the wind-driven ocean wave model WAVEWATCH III (e.g. Tolman 2008; Ardhuin *et al.* 2011) is shown for reference for each example. WAVEWATCH III provides maps of the surface pressure that is generated by ocean wave interactions. This surface pressure is considered to be the dominant source of the seismic signals in the secondary microseismic frequency band (e.g. Kedar *et al.* 2008; Stutzmann *et al.* 2012). The source pressure maps serve merely as an independent reference with which our observations can be compared. It is only approximate because the pressure model does not take into account the source site effect and resonance in the ocean water column which alter the frequency and amplitude of the seismic signals (Gualtieri *et al.* 2014; Meschede *et al.* 2017). As for the beam energy, the pressure amplitude is shown with colour intensity and the dominant frequency of the pressure with colour hue. Additionally, we show the positions of tropical cyclones that are active during a day with black dots because cyclones are known sources of secondary microseisms.

## 4.1 Example 1: typhoons Atsani and Goni

### 4.1.1 Unmixed sources

The setting of the first example are secondary microseisms from the twin typhoons Atsani and Goni on 2015 August 21.

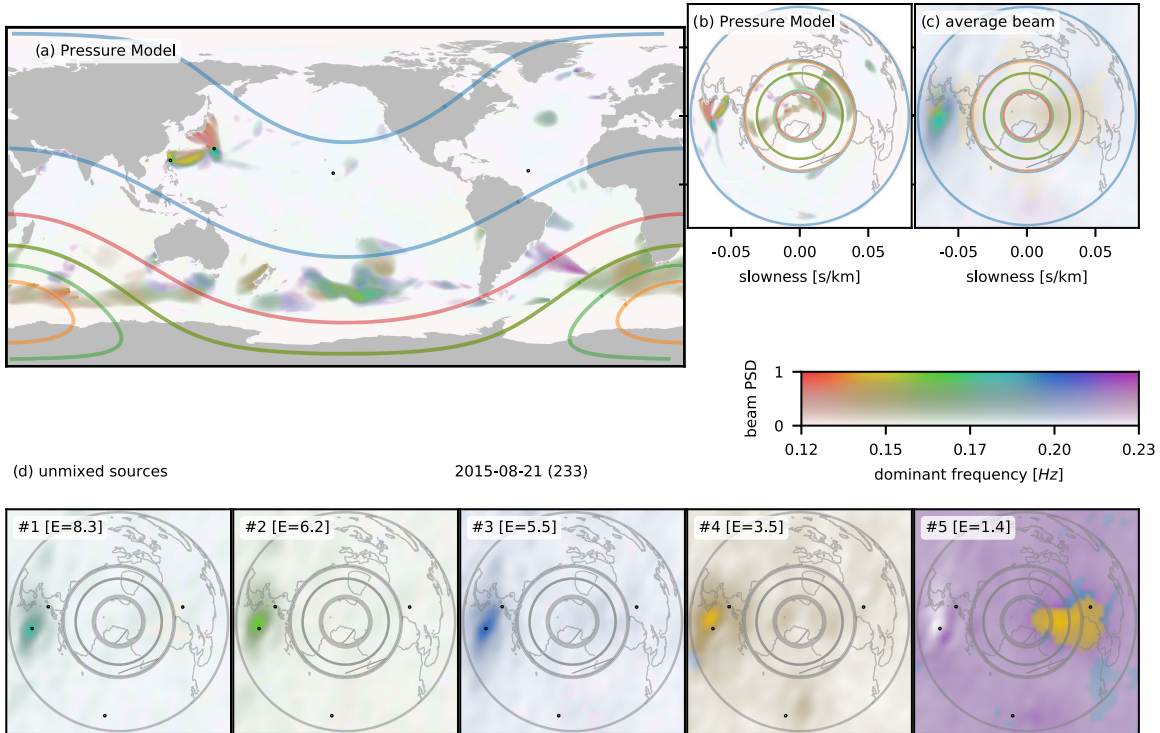
The day-averaged wave interaction pressure from the model WAVEWATCH III is shown in Fig. 8(a). The regions that are covered by the different seismic phases are indicated with lines that are coloured according to Fig. 7 (the short distance orange and green line overlap and therefore appear with a different colour). Black

dots show the positions of the tropical cyclones Goni at the coast of Taiwan, Atsani southeast of Japan, Kilo in the central Pacific and Dani in the central Atlantic. Strong wave interactions can be seen at the position of Goni and Atsani in the  $P$ -wave distance range. The wave interaction pressure has higher dominant frequency at the position of Atsani (blue colour) and lower dominant frequency towards the Japanese coast (red colour). Other regions with strong wave interaction pressure are the South-Atlantic and South-Indian ocean that lie in the PKPab/PKPbc and PKIKP distance ranges. The South-Pacific is in the shadow zone of the Earth’s core from the Alaskan perspective, and cannot be seen by any of the seismic phases that we consider. It could however generate Pdiff phases that arrive with the slowness of a large distance  $P$  wave or of a nearly antipodal PKPab wave (see grey line in Fig. 7b).

The wave interaction pressure projected to slowness space is shown in Fig. 8(b). To project the map shown in Fig. 8(a) to slowness domain, we transform each point, defined by its distance and azimuth from the Alaska array centre, to slowness and azimuth using the relation shown in Fig. 7(b). On the left side (westward direction) in the  $P$ -wave slowness range, the wave interaction pressure corresponding to the cyclones Goni and Atsani can be seen. From the Alaska array, the South-Indian ocean lies in westward (left) direction and the South-Atlantic in eastward (right) direction. In the PKPab and PKPbc slowness regions, strong wave-wave interaction pressure can be seen in these directions, as well as in the PKIKP slowness range that covers the antipode.

The observed beam energy, averaged over the whole day, is shown in Fig. 8(c). It shows very clearly signal energy that corresponds to  $P$  waves from the cyclones Goni and Atsani. The colour indicates that its dominant frequency is 0.18 Hz (turquoise/blue) at the position of cyclone Atsani and 0.14 Hz (faint yellow) towards the Japanese coast. This frequency dependence corresponds to the one that we expect from the ocean wave hindcast shown in Figs 8(a) and (b). Physically, long period signals from the coast are generated by the interaction of incident ocean waves with waves reflected at the coast that are generally stronger at longer periods. Signal energy at other slownesses is not clearly identifiable in the day averaged beam.

The lower panels of Fig. 8 show the decomposed beam for this day. We have used NNMF to find five sources that optimally explain



**Figure 8.** Observation of secondary microseism signals on the 2015 August 21 by the seismic stations in Alaska. Panel (a) shows a model of ocean wave interaction pressure that has been computed from the wind-driven ocean wave hindcast WAVEWATCH III. Coloured lines indicate the regions of the seismic phases P (orange), PKPab (green), PKPbc (red) and PKIKP (purple). Black dots correspond to cyclones Goni (Taiwan), Atsani (South-East of Japan), Kilo (Central Pacific) and Dani (Central Atlantic). (b) Ocean wave interaction pressure projected to slowness domain. The slowness regions of the different seismic phases are indicated with coloured lines. (c) Day averaged beam energy. (d) Unmixed beam components, ordered from the strongest (#1) to the weakest (#5) signal to noise ratio E.

the 4321 time samples of the beam. The unmixed sources are ordered from the strongest (#1) to the weakest source (#5) in signal-to-noise ratio units. Sources #1 to #4 correspond to the Goni–Atsani source region. Source #5 corresponds to a source region in the South-Atlantic.

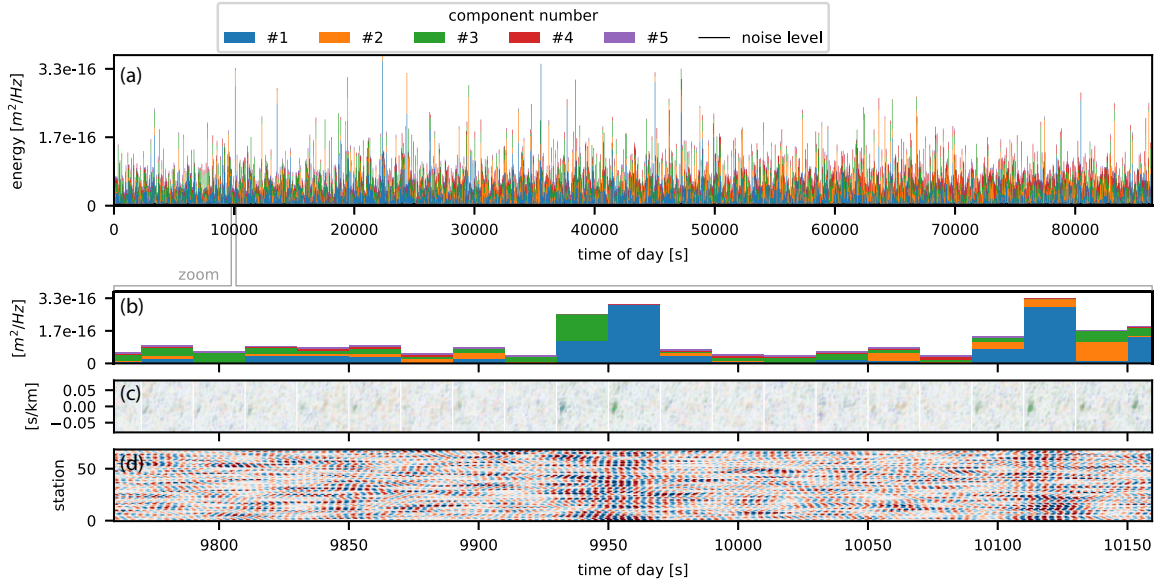
Source #1 is centred at the location of cyclone Atsani with dominant frequencies of about 0.18 Hz. Source #2 is close to source #1 but slightly shifted towards the Japanese coastline with lower frequencies of about 0.16 Hz. Source #3 is a high-frequency source (0.20 Hz) at the position of Atsani. Source #4 is close to the Japanese coastline and has the lowest frequency (0.14 Hz). Sources #1 to #4 belong to the Goni–Atsani source region, whose frequency composition is shown in detail. The frequency composition of the Goni–Atsani source region corresponds to our expectations from the wave–wave interaction pressure model: sources at the coastline have significantly lower frequency than sources at the cyclone.

Source #5 is weaker and corresponds to PKPab/PKPbc and maybe PKIKP sources from the South-Atlantic source region. It even seems that there is some additional energy at the source location of cyclone Dani. It is mixed with a noise component. A closer look on sources #4 and #5 reveals that there is some beam energy at the position of cyclone Kilo. Indeed, we find that increasing the source number to 10 enables us to separate sources from this cyclone as well.

#### 4.1.2 Time development and mixing components

Each unmixing component in Fig. 8 comes with an amplitude for each of the 4321 time samples of the beam. These amplitudes describe how strong a component contributes to the beam at a given time. The amplitudes for each component are shown in Fig. 9(a), stacked on top of each other. The noise level of the beam is indicated with a black colour but is not distinguishable from the time axis in this example because it is very low. Fig. 9(a) demonstrates that the component amplitudes fluctuate quickly, not only from one component to another, but also for all amplitudes. Some longer duration developments of the amplitudes can also be seen. Source #1 (blue) is stronger at the beginning of the day, whereas sources #2 (orange) and #4 (red) become stronger towards the end, indicating that source region moved towards the Japanese coast towards the end of the day and became lower frequency.

A zoomed extract of this plot is shown in Fig. 9(b) for times 9760–10 150 [s]. Fig. 9(c) shows the corresponding beams, and Fig. 9(d) the associated wavefield. More precisely, Fig. 9(d) shows the recorded seismic displacement at each station of the array, time shifted according to the slowness of source #1 and the positions of each station, and filtered around the dominant frequency of source #1. The wavefield is therefore aligned vertically in Fig. 9, whenever a plane wave arrives with the slowness of source #1. This happens, in particular, at time sample 9960 s which is almost exclusively composed of source #1. At this instant, the beam shows source #1, and the wavefield is well aligned, showing the actual



**Figure 9.** (a) Temporal evolution of the amplitudes of the unmixed sources over the whole day. (b–d) Zoom at times 22 180–22 580 of (b) the source amplitudes, (c) the beam time samples and (d) the ground displacement at each station of the array (wavefield), that has been aligned according to the slowness of source #1.

seismic signal that has been identified and detected from the decomposition. The seismic signal again demonstrates the fluctuating amplitude of the signal.

Signals from secondary microseisms are not always so clearly visible as in this example. Most of the time, they are somewhat hidden by other signals and just appear gradually when many time samples are averaged. However, because the amplitude of microseismic signals fluctuates strongly at a seismic array, often some signals arrive with stronger amplitude than the noise level and can be examined individually.

#### 4.1.3 Extracting relative traveltimes

The quality with which secondary microseism signals, as the one shown in Fig. 9(d), can be observed, suggests that we can extract traveltimes from them. To this end, we use the procedure explained in Section 3.5. The traveltimes shifts that correspond to the five components of this example are shown in Fig. 10 for each station over a topographic map of Alaska. The size of the scatter points is proportional to the consistency value  $|c_l|$  with which the traveltimes was observed over the day, smaller points are less consistent than larger points. The maximum consistency of  $|c_l| = 0.5$  is obtained for component #1 that is the strongest. A black arrow shows the direction of the slowness vector with maximum energy for each component.

Few consistent trends are visible in the traveltimes: slower velocity along the southeastern coastline, faster velocity in the southwest, slower velocities around the array centre (Denali gap), and faster velocities in the north and north of the southeastern coastline. This corresponds to our expectations from the geology of Alaska and tomographic studies of this region (Martin-Short *et al.* 2015).

## 4.2 Example 2: multiple sources including PKP waves from Cyclone Glenda

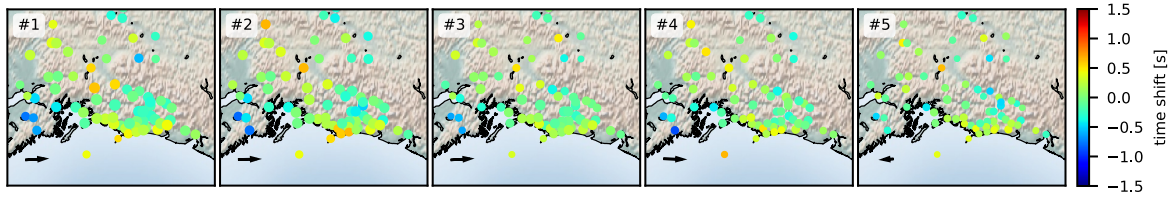
The context for the second example is represented by several secondary microseism source regions that were active on 2015 March 1.

The major source regions with strong wave interaction pressure in the model (shown in Figs 11 a and b) are in the North-Pacific in the regional  $P$ -wave slowness range, in the North Pacific around Kamchatka and in the North Atlantic around Greenland in the  $P$ -wave slowness range, in the South Atlantic and South Indian Ocean off the coast of Madagascar in the PKP slowness range (cyclone Glenda, indicated with a black dot), and at the Antarctic coast in the PKIKP slowness range.

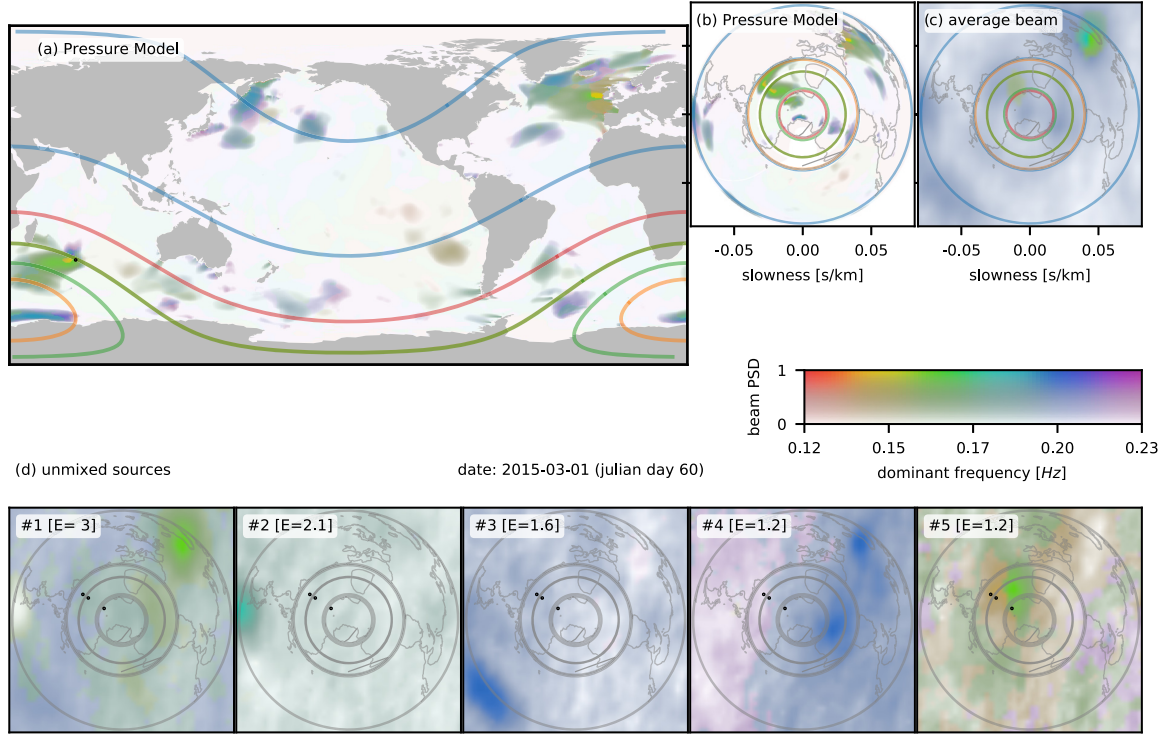
In the time averaged observed beam, shown in Fig. 11(c), the dominant source in the North-Atlantic is the only source region that can be clearly identified. Other source regions can be seen only vaguely. On the other hand, the unmixed sources, shown in Fig. 11(d) separate the individual sources regions: component #1 corresponds to the North Atlantic, component #2 to the Kamchatka, component #3 to the North Pacific, component #4 to the South Atlantic, and #5 to the South Indian source regions. All components are mixed with noise and also component #4 mixes high frequency energy from South Atlantic PKP waves with high frequency energy from North Atlantic  $P$  waves. Increasing the number of components (not shown), improves the unmixing and separates the two wave types in component #4, as well as noise from source components. However, increasing the number of components splits also some of the source regions into several components, and it is a matter of choice whether one wants to have a larger number of more precise components or a smaller number of undivided ones.

Compared to example 1, the components in this example have a lower signal to noise ratio as becomes clear from the black line that shows the noise level in Fig. 12(a). Also many components are active simultaneously as can be seen in Fig. 12(b) that shows a zoomed time window of Fig. 12(a). Correspondingly, the beam energy is noisy (Fig. 12c) and signals are more difficult to identify. Nevertheless, the waveforms shown in Fig. 12(d) can be seen to align at 34860s, when the component #1 amplitude (Fig. 12b) is high and the amplitude of other components is low.

Traveltimes can also be extracted from such a noisier record, using our automated approach. In example 2, the consistency value  $|c_l| < 0.2$  of the average phases that are extracted is much lower than in example 1. Nevertheless, the traveltimes patterns that are



**Figure 10.** Traveltime differences that were extracted for the unmixed components #1–#5. The source direction from where maximum beam energy arrives at the array is indicated with a black arrow.



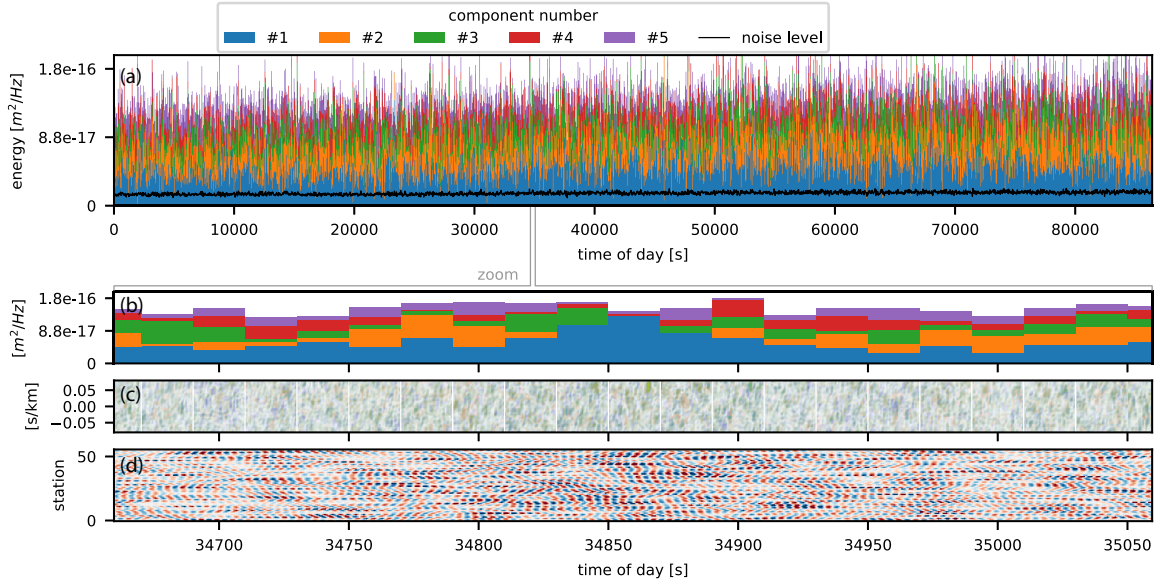
**Figure 11.** Observation of secondary microseism signals on 2015 March 1 by the Alaska array. Panel (a) shows a model of ocean wave interaction pressure that has been computed from the wind-driven ocean wave hindcast WAVEWATCH III. Coloured lines indicate the regions of the seismic phases P (orange), PKPab (green), PKPbc (red) and PKIKP (purple). The black dot corresponds to cyclone Glenda (South Indian Ocean). (b) Ocean wave interaction pressure projected to slowness domain. The slowness regions of the different seismic phases are indicated with coloured lines. (c) Day averaged beam energy. (d) Unmixed beam components, ordered from the strongest (#1) to the weakest (#5) signal to noise ratio  $E$ .

extracted, and shown in Fig. 13, are similar to the ones extracted from example 1. Comparing the traveltimes extracted in example 2 from component #2 that is located at Japan with the traveltimes extracted in example 1 from any of the components #1 to #4, one can identify common patterns such as: the low velocity zone at the southeastern coast of Alaska, the fast velocity zone in the southwest, the faster velocities away from the coast in the east, and the slower velocities in the centre to northwest part of the array. In contrast to example 1, the components come from various directions to the array. Component #1 and #2 show slightly different patterns of traveltimes that might be exploitable in a tomographic study. Traveltimes from component #4 are somewhat inconsistent with the ones from component #1 although both seem to examine signals with similar slowness. However, remember that source #4 is only partly unmixed. This difference can be explained considering the higher noise level that particularly degrades the weaker components. To fully automatize traveltime extraction, we should therefore impose a strict threshold on the consistency value  $|c_l|$  that is a measure of the quality of the obtained traveltimes.

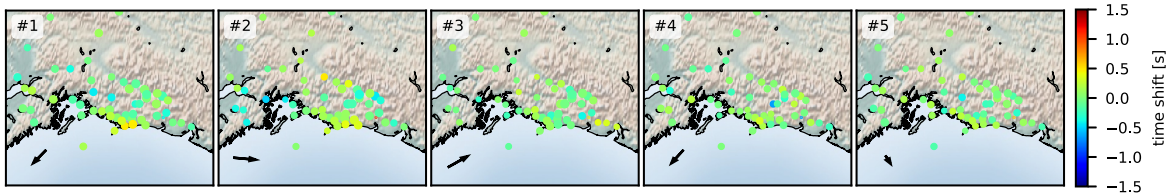
## 5 CONCLUSIONS

We have demonstrated that recurrent seismic signals from secondary microseisms can be identified and decomposed fully automatically using BSS. We have used NMF as BSS technique to decompose the non-negative beam energy in slowness–frequency space. This method can be used to examine microseisms with high precision and find and extract weak signals from the mixture of signals that are observed over the course of the day by a seismic array. It provides knowledge of the short duration time development of each signal, which can, for example, be used to automatically extract traveltimes from secondary microseisms in a fully automatic way. It has been shown by other authors that such traveltime measurements from secondary microseisms can resolve subsurface structure.

We have used the proposed methodology on data from a seismic array in Alaska to extract  $P$  waves from secondary microseisms associated with the cyclones Goni and Atsani, as well as  $P$  waves from other regions with strong wave interactions, and  $PKP$  waves from cyclone Glenda. We have confirmed the assumption that the amplitude of secondary microseismic signals fluctuates rapidly. Such



**Figure 12.** (a) Temporal evolution of the amplitudes of the unmixed sources over the whole day. (b–d) Zoom at times 34660–35060 s of the day. Panel (b) shows the source amplitudes, (c) the beam time samples, and (d) the ground displacement at each station of the array (wavefield), which has been aligned according to the slowness of source #1.



**Figure 13.** Traveltime differences that were extracted for the unmixed components #1 to #5. The source direction from where maximum beam energy arrives at the array is indicated with a black arrow.

signals can be strong enough to surpass the background noise level such that they can be examined without any time averaging or stacking directly in the wavefield. Finally, we have extracted relative traveltimes with a fully automatic procedure from automatically detected and characterized microseismic signals for the seismic array in Alaska. The quality of these measurements is automatically quantified and leads to observations that correspond to our geological knowledge of the region and that can potentially be used in future tomographic studies.

In this paper, we observed and exploited strong fluctuations in amplitude of signals from secondary microseisms. These fluctuations allow us to examine individual sources with high precision because they reduce the amount of interference between different signals (signals are disjoint in time). The origin of these fluctuations can be the random character of the ocean wavefield that leads to fluctuations of the secondary microseism source amplitude. It would be interesting to analyse in detail if the amplitude distribution of the microseismic signals that are recorded at a seismic array corresponds to the one that is expected for the interaction pressure of ocean waves. Ocean waves themselves are typically assumed to behave as a Gaussian random field and are modelled accordingly. These statistics are less well examined for the nonlinear interactions of the ocean waves that generate secondary microseismic sources. Such questions are beyond the scope of this paper but can certainly be addressed using techniques as the proposed one to access the amplitude of individual microseismic sources with high time resolution.

## ACKNOWLEDGEMENTS

This paper was funded through the ANR (Agence Nationale de Recherche) project number ANR-14-CE01-0012.

The following open scientific software packages were used: obspy (Krischer *et al.* 2015), numpy & scipy (van der Walt *et al.* 2011), scikit-learn (Pedregosa *et al.* 2011) and matplotlib (Hunter *et al.* 2007). Numerical computations were performed on the S-CAPAD platform, IPGP, France.

The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used in this study for access to waveforms and related metadata from the following seismic networks: AK (Alaska Earthquake Center & Univ. of Alaska Fairbanks 1987), TA (IRIS Transportable Array 2003). IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681.

The secondary microseism pressure model with and without coastal reflections was developed as part of the IOWAGA project (ERC grant 240009). It is freely available from Ifremer (<http://ftp.ifremer.fr/ifremer/ww3/HINDCAST/SISMO/>).

This is IPGP Contribution Number: 3996

## REFERENCES

Alaska Earthquake Center, Univ. of Alaska Fairbanks, 1987. Alaska regional network, *Seismic Network*.

- Ardhuin, F., Stutzmann, E., Schimmel, M. & Mangeney, A., 2011. Ocean wave sources of seismic noise, *J. geophys. Res.*, **116**(C9), 1-21.
- Cabras, G., Carniel, R. & Wassermann, J., 2008. Blind source separation: an application to the Mt. Merapi volcano, Indonesia, *Fluctuation Noise Lett.*, **8**(03n04), L249–L260.
- Cabras, G., Carniel, R. & Jones, J., 2012. Non-negative matrix factorization: an application to Erta 'Aale volcano, Ethiopia, *Boll. Geofis. Teor. Appl.*, **53**(2), 231–242.
- Cabras, G., Carniel, R., Jones, J.P. & Takeo, M., 2014. Reducing wind noise in seismic data using non-negative matrix factorization: an application to Villarrica volcano, Chile, *Geofis. Int.*, **53**(1), 77–85.
- Cichocki, A. & Phan, A.-H., 2009. Fast local algorithms for large scale non-negative matrix and tensor factorizations, *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, **92**(3), 708–721.
- Cichocki, A., Zdunek, R., Phan, A.H. & Amari, S.-I., 2009. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley & Sons.
- Comon, P. & Jutten, C., 2010. *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press.
- Farra, V., Stutzmann, E., Gualtieri, L., Schimmel, M. & Ardhuin, F., 2016. Ray-theoretical modeling of secondary microseism P waves, *Geophys. J. Int.*, **206**(3), 1730–1739.
- Févotte, C. & Idier, J., 2011. Algorithms for nonnegative matrix factorization with the  $\beta$ -divergence, *Neural Comput.*, **23**(9), 2421–2456.
- Friedman, J., Hastie, T. & Tibshirani, R., 2001. *The Elements of Statistical Learning, Vol. 1*, Springer Series in Statistics, Springer.
- Gal, M., Reading, A.M., Ellingsen, S.P., Koper, K.D., Burlacu, R. & Gibbons, S.J., 2016. Deconvolution enhanced direction of arrival estimation using one- and three-component seismic arrays applied to ocean induced microseisms, *Geophys. J. Int.*, **206**(1), 345–359.
- Gualtieri, L., Stutzmann, E., Farra, V., Capdeville, Y., Schimmel, M., Ardhuin, F. & Morelli, A., 2014. Modelling the ocean site effect on seismic noise body waves, *Geophys. J. Int.*, **197**, 1096–1106.
- Halko, N., Martinsson, P.-G. & Tropp, J.A., 2011. Finding structure with randomness: stochastic algorithms for constructing approximate matrix decompositions, *SIAM Rev.*, **53**, 217–288.
- Hasselmann, K., 1963. A statistical analysis of the generation of microseisms, *Rev. Geophys.*, **1**(2), 177–210.
- Hunter, J.D. *et al.*, 2007. Matplotlib: a 2D graphics environment, *Comput. Sci. Eng.*, **9**(3), 90–95.
- Hyvärinen, A. & Oja, E., 2000. Independent component analysis: algorithms and applications, *Neural Netw.*, **13**(4), 411–430.
- Ikelle, L., 2007. Coding and decoding: Seismic data modeling, acquisition and processing, in *SEG Technical Program Expanded Abstracts 2007*, Society of Exploration Geophysicists, pp. 66.
- IRIS Transportable Array, 2003. Usarray transportable array, International Federation of Digital Seismograph Networks, *Seismic Network*.
- Kedar, S., Longuet-Higgins, M., Webb, F., Graham, N., Clayton, R. & Jones, C., 2008. The origin of deep ocean microseisms in the north atlantic ocean, *Proc. R. Soc. A*, **464**, 777–793.
- Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C. & Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discovery*, **8**(1), 1–17.
- Liu, K.-H. & Dragoset, W.H., 2013. Blind-source separation of seismic signals based on information maximization, *Geophysics*, **78**(4), V119–V130.
- Liu, Q. *et al.*, 2016. Source locations of teleseismic P, SV, and SH waves observed in microseisms recorded by a large aperture seismic array in China, *Earth planet. Sci. Lett.*, **449**, 39–47.
- Martin-Short, R., Allen, R.M., Bastow, I.D., Totten, E. & Richards, M.A., 2015. Mantle flow geometry from ridge to trench beneath the Gorda-Juan de Fuca plate system, *Nat. Geosci.*, **8**(12), 965–968.
- Meschede, M., Stutzmann, E., Farra, V., Schimmel, M. & Ardhuin, F., 2017. The effect of water column resonance on the spectra of secondary microseism P waves, *J. geophys. Res.*, **122**(10), 8121–8142.
- Moni, A., Bean, C.J., Lokmer, I. & Rickard, S., 2012. Source separation on seismic data: application in a geophysical setting, *IEEE Signal Process. Mag.*, **29**(3), 16–28.
- Moni, A., Craig, D. & Bean, C.J., 2013. Separation and location of microseism sources, *Geophys. Res. Lett.*, **40**(12), 3118–3122.
- Nishida, K. & Takagi, R., 2016. Teleseismic s wave microseisms, *Science*, **353**(6302), 919–921.
- Parra, L. & Spence, C., 2000. Convolutional blind separation of non-stationary sources, *IEEE Trans. Speech Audio Process.*, **8**(3), 320–327.
- Pedregosa, F. *et al.*, 2011. Scikit-learn: machine learning in python, *J. Mach. Learning Res.*, **12**, 2825–2830.
- Schimmel, M. & Paulssen, H., 1997. Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.*, **130**(2), 497–505.
- Stutzmann, E., Schimmel, M., Patau, G. & Maggi, A., 2009. Global climate imprint on seismic noise, *Geochem. Geophys. Geosyst.*, **10**(11), 1-12.
- Stutzmann, E., Ardhuin, F., Schimmel, M., Mangeney, A. & Patau, G., 2012. Modelling long-term seismic noise in various environments, *Geophys. J. Int.*, **191**(2), 707–722.
- Takahata, A.K., Nadalin, E.Z., Ferrari, R., Duarte, L.T., Suyama, R., Lopes, R.R., Romano, J.M. & Tygel, M., 2012. Unsupervised processing of geophysical signals: a review of some key aspects of blind deconvolution and blind source separation, *IEEE Signal Process. Mag.*, **29**(4), 27–35.
- Tolman, H.L., 2008. A mosaic approach to wind wave modeling, *Ocean Modelling*, **25**(1-2), 35–47.
- van der Walt, S., Colbert, S.C. & Varoquaux, G., 2011. The NumPy array: a structure for efficient numerical computation, *Comput. Sci. Eng.*, **13**(2), 22–30.
- Vincent, E., Gribonval, R. & Févotte, C., 2006. Performance measurement in blind audio source separation, *IEEE Transactions Audio Speech Lang. Process.*, **14**(4), 1462–1469.
- Zhang, J., Gerstoft, P. & Shearer, P.M., 2010. Resolving P-wave travel-time anomalies using seismic array observations of oceanic storms, *Earth planet. Sci. Lett.*, **292**(3),

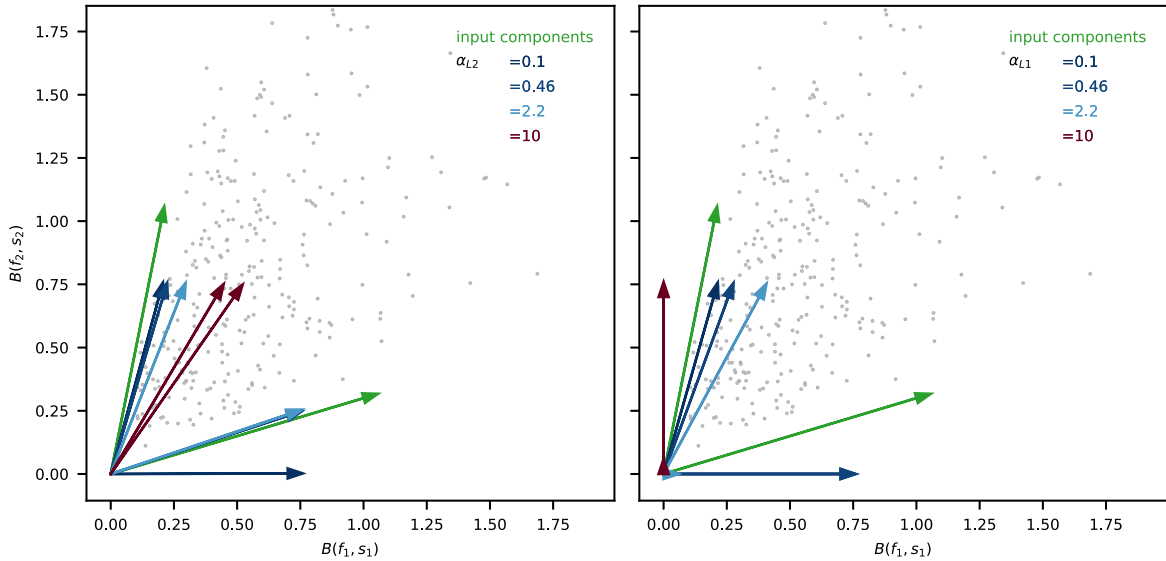
419–427.

## APPENDIX A: L1 VERSUS L2 REGULARIZATION

Regularization is introduced to NNMF as an additional cost to eq. (11):

$$\underset{A_{im}>0; E_{mk}>0}{\text{minimize}} \left[ \sum_{ik} |B_{ik} - \sum_m A_{im} E_{mk}|^2 + \alpha_{L2} \sum_{mk} (E_{mk}^2 + A_{mk}^2) + \alpha_{L1} \sum_{mk} (|E_{mk}| + |A_{mk}|) \right]. \quad (\text{A1})$$

Fig. A1 illustrates its effect on a synthetic beam, represented as a point cloud as in Fig. 3. L2 regularization with  $\alpha_{L1} = 0$  (left) equalizes component values (and amplitudes that are not shown) at high damping parameters  $\alpha_{L2}$ . L1 regularization with  $\alpha_{L2} = 0$  (right) favours sparse solutions at high damping parameters  $\alpha_{L1}$  because it corresponds to a norm with corners (for details see e.g. Friedman *et al.* 2001).

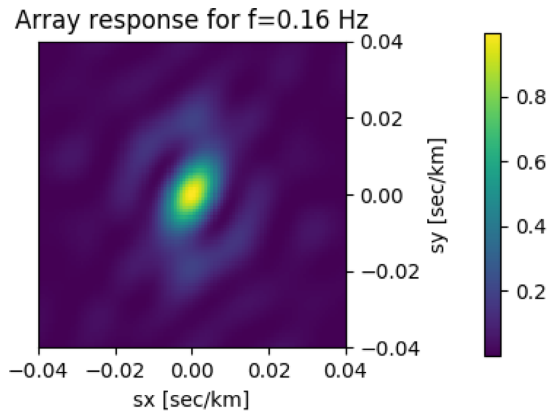


**Figure A1.** L2 (left) and L1 (right) regularized NNMF for different damping parameters  $\alpha$ . L1 regularization favours sparse components that have zero  $x$  or  $y$  values. L2 regularization has a tendency to equalize the components.

## APPENDIX B: ARRAY RESPONSE AND PLANE WAVE ASSUMPTION

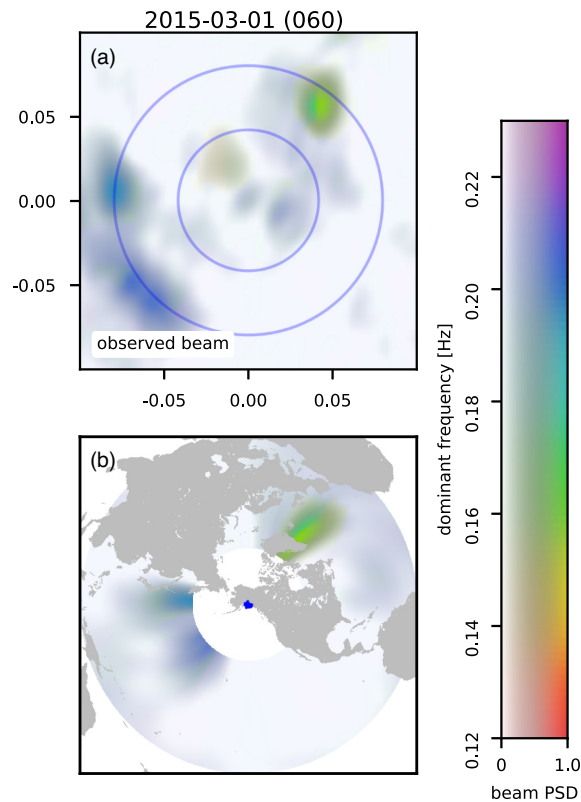
Fig. B1 shows the array response of the Alaska array that is used in this paper for a frequency of 0.16 Hz. The aperture of this virtual array is 600 km. The width of the main lobe at 90 per cent of the maximum is  $0.004 \text{ s km}^{-1}$  along the  $s_x$  axis and  $0.006 \text{ s km}^{-1}$  along the  $s_y$  axis. It corresponds to a surface of about  $400 \text{ km} \times 800 \text{ km}$  for a source at a distance of  $50^\circ$  from the network.

Beams are computed using eq. (4) using the plane wave approximation. In Fig. 11(c), sources close to Greenland are detected with a



**Figure B1.** Array response of the Alaska array that is used in this paper for a period of 0.16 Hz.

dominant frequency between 0.15 and 0.17 Hz. The distance between the sources and the Alaska array is about  $50^\circ$ . In that case, the time difference between the arrival of a wave at the centre of the network and at the further station is 22.480 s in spherical Earth and 20.528 s



**Figure B2.** Day average beam energy computed using a subarray with 400 km aperture. Colours represent the frequency of the beam maximum and colour hue gives the normalized beam amplitude. The maximum of the beam close to Greenland is similar to that observed with larger seismic array as in Fig. 11(c).

using plane wave approximation. The time error is 1.95 s, that is about 1/3 of the wave period, and therefore the further station will still sum up constructively to the total beam energy.

In order to further investigate the beam resolution as a function of array aperture, we selected a subset of the Alaska stations corresponding to a virtual array aperture of 400 km. The beam computed for the same day as in Fig. 11(c) is plotted in Fig. B2. Beams in Figs B2 and 11 c are very similar: two sources at two close frequencies (0.15 and 0.16 Hz) are clearly detected near Greenland. Therefore, even though considering the curvature may improve the results, we conclude that the distortion of the wave front can be neglected in this study.



## Key words

Authors are requested to choose key words from the list below to describe their work. The key words will be printed underneath the summary and are useful for readers and researchers. Key words should be separated by a semi-colon and listed in the order that they appear in this list. An article should contain no more than six key words.

COMPOSITION and PHYSICAL PROPERTIES	Seismic cycle	Instability analysis
Composition and structure of the continental crust	Space geodetic surveys	Interferometry
Composition and structure of the core	Tides and planetary waves	Inverse theory
Composition and structure of the mantle	Time variable gravity	Joint inversion
Composition and structure of the oceanic crust	Transient deformation	Neural networks, fuzzy logic
Composition of the planets		Non-linear differential equations
Creep and deformation		Numerical approximations and analysis
Defects	GEOGRAPHIC LOCATION	Numerical modelling
Elasticity and anelasticity	Africa	Numerical solutions
Electrical properties	Antarctica	Persistence, memory, correlations, clustering
Equations of state	Arctic region	Probabilistic forecasting
Fault zone rheology	Asia	Probability distributions
Fracture and flow	Atlantic Ocean	Self-organization
Friction	Australia	Spatial analysis
High-pressure behaviour	Europe	Statistical methods
Magnetic properties	Indian Ocean	Thermobarometry
Microstructure	Japan	Time-series analysis
Permeability and porosity	New Zealand	Tomography
Phase transitions	North America	Waveform inversion
Plasticity, diffusion, and creep	Pacific Ocean	Wavelet transform
	South America	
GENERAL SUBJECTS	GEOMAGNETISM and ELECTROMAGNETISM	PLANETS
Core	Archaeomagnetism	Planetary interiors
Gas and hydrate systems	Biogenic magnetic minerals	Planetary volcanism
Geomechanics	Controlled source electromagnetics (CSEM)	
Geomorphology	Dynamo: theories and simulations	SEISMOLOGY
Glaciology	Electrical anisotropy	Acoustic properties
Heat flow	Electrical resistivity tomography (ERT)	Body waves
Hydrogeophysics	Electromagnetic theory	Coda waves
Hydrology	Environmental magnetism	Computational seismology
Hydrothermal systems	Geomagnetic excursions	Controlled source seismology
Infrasound	Geomagnetic induction	Crustal imaging
Instrumental noise	Ground penetrating radar	Earthquake dynamics
Ionosphere/atmosphere interactions	Magnetic anomalies: modelling and interpretation	Earthquake early warning
Ionosphere/magnetosphere interactions	Magnetic fabrics and anisotropy	Earthquake ground motions
Mantle processes	Magnetic field variations through time	Earthquake hazards
Ocean drilling	Magnetic mineralogy and petrology	Earthquake interaction, forecasting, and prediction
Structure of the Earth	Magnetostratigraphy	Earthquake monitoring and test-ban treaty verification
Thermochronology	Magnetotellurics	Earthquake source observations
Tsunamis	Marine electromagnetics	Guided waves
Ultra-high pressure metamorphism	Marine magnetics and palaeomagnetism	Induced seismicity
Ultra-high temperature metamorphism	Non-linear electromagnetics	Interface waves
GEODESY and GRAVITY	Palaeointensity	Palaeoseismology
Acoustic-gravity waves	Palaeomagnetic secular variation	Rheology and friction of fault zones
Earth rotation variations	Palaeomagnetism	Rotational seismology
Geodetic instrumentation	Rapid time variations	Seismic anisotropy
Geopotential theory	Remagnetization	Seismic attenuation
Global change from geodesy	Reversals: process, time scale, magnetostratigraphy	Seismic instruments
Gravity anomalies and Earth structure	Rock and mineral magnetism	Seismic interferometry
Loading of the Earth	Satellite magnetism	Seismicity and tectonics
Lunar and planetary geodesy and gravity		Seismic noise
Plate motions	GEOPHYSICAL METHODS	Seismic tomography
Radar interferometry	Downhole methods	Site effects
Reference systems	Fourier analysis	Statistical seismology
Satellite geodesy	Fractals and multifractals	Surface waves and free oscillations
Satellite gravity	Image processing	Theoretical seismology
Sea level change		

Tsunami warning  
 Volcano seismology  
 Wave propagation  
 Wave scattering and diffraction

#### TECTONOPHYSICS

Backarc basin processes  
 Continental margins: convergent  
 Continental margins: divergent  
 Continental margins: transform  
 Continental neotectonics  
 Continental tectonics: compressional  
 Continental tectonics: extensional  
 Continental tectonics: strike-slip and transform  
 Cratons  
 Crustal structure  
 Diapirism  
 Dynamics: convection currents, and mantle plumes  
 Dynamics: gravity and tectonics  
 Dynamics: seismotectonics  
 Dynamics and mechanics of faulting  
 Dynamics of lithosphere and mantle  
 Folds and folding  
 Fractures, faults, and high strain deformation zones  
 Heat generation and transport

Hotspots  
 Impact phenomena  
 Intra-plate processes  
 Kinematics of crustal and mantle deformation  
 Large igneous provinces  
 Lithospheric flexure  
 Mechanics, theory, and modelling  
 Microstructures  
 Mid-ocean ridge processes  
 Neotectonics  
 Obduction tectonics  
 Oceanic hotspots and intraplate volcanism  
 Oceanic plateaus and microcontinents  
 Oceanic transform and fracture zone processes  
 Paleoseismology  
 Planetary tectonics  
 Rheology: crust and lithosphere  
 Rheology: mantle  
 Rheology and friction of fault zones  
 Sedimentary basin processes  
 Subduction zone processes  
 Submarine landslides  
 Submarine tectonics and volcanism  
 Tectonics and climatic interactions  
 Tectonics and landscape evolution  
 Transform faults  
 Volcanic arc processes

#### VOLCANOLOGY

Atmospheric effects (volcano)  
 Calderas  
 Effusive volcanism  
 Eruption mechanisms and flow emplacement  
 Experimental volcanism  
 Explosive volcanism  
 Lava rheology and morphology  
 Magma chamber processes  
 Magma genesis and partial melting  
 Magma migration and fragmentation  
 Mud volcanism  
 Physics and chemistry of magma bodies  
 Physics of magma and magma bodies  
 Planetary volcanism  
 Pluton emplacement  
 Remote sensing of volcanoes  
 Subaqueous volcanism  
 Tephrochronology  
 Volcanic gases  
 Volcanic hazards and risks  
 Volcaniclastic deposits  
 Volcano/climate interactions  
 Volcano monitoring  
 Volcano seismology