

# Orienting and locating ocean-bottom seismometers from ship noise analysis

A. Trabattoni<sup>1</sup>, G. Barruol<sup>1</sup>, R. Dreö<sup>2</sup>, A.O. Boudraa<sup>2</sup> and F.R. Fontaine<sup>1,3</sup>

<sup>1</sup> Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France. E-mail: [alister.trabattoni@gmail.com](mailto:alister.trabattoni@gmail.com)

<sup>2</sup> Ecole Navale IRENav/Arts & Métiers ParisTech, BCRM Brest, CC 600, 29240 Brest, France

<sup>3</sup> Université de La Réunion, Laboratoire GéoSciences Réunion, F-97744 Saint Denis, France

Accepted 2019 November 18. Received 2019 November 13; in original form 2019 July 10

## SUMMARY

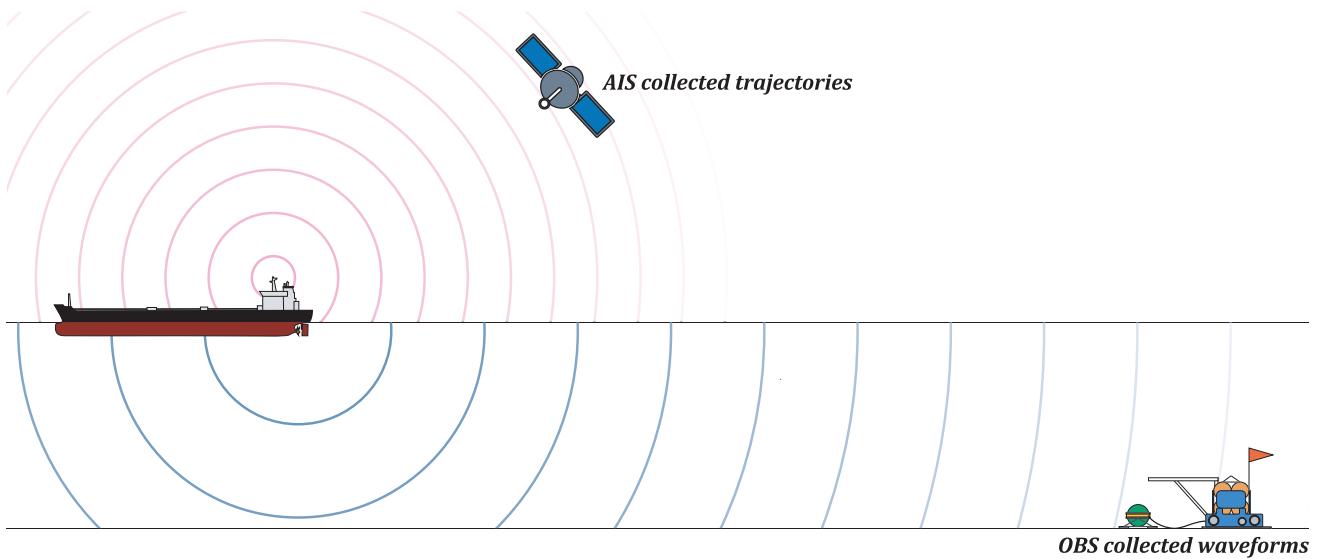
Breakthroughs in understanding the structure and dynamics of our planet will strongly depend upon instrumenting deep oceans. Progress has been made these last decades in ocean-bottom seismic observations, but ocean-bottom seismometer (OBS) temporary deployments are still challenging and face set-up limitations. Launched from oceanographic vessels, OBSs fall freely and may slightly drift laterally, dragged by currents. Therefore, their actual orientation and location on the landing sites are hard to assess precisely. Numerous techniques have been developed to retrieve this key information, but most of them are costly, time-consuming or inaccurate. In this work, we show how ship noise can be used as an acoustic source of opportunity to retrieve both the orientation and the location of OBSs on the ocean floor. To retrieve the OBS orientation, we developed a first method based on a combination of seismic and pressure data through the use of the acoustic intensity. This latter can be used to quantify the OBS orientation from the ship noise direction of arrival (DOA), which can then be compared with known ship trajectories obtained from the automatic identification system (AIS). To accurately relocate OBSs, we also developed a second method based on the hydrophone data which computes distances of acoustical sources by measuring time differences of arrival (TDOA) between direct and reverberated phases. The OBS location is then retrieved by fitting measured ship distances with known ship trajectories. In this study, a full network of OBSs deployed in the SW Indian Ocean was reoriented and a test station was relocated. We demonstrate that our new methods may quantify the OBS orientation with an accuracy of about one degree, and its location with an accuracy of a few tens of metres, depending on the number of ships used in the analysis.

**Key words:** Statistical methods; Time-series analysis; Body waves; Seismic instruments; Seismic noise; Wave propagation.

## INTRODUCTION

The Earth's interior imaging and monitoring is largely based on observations from seismological stations. Unfortunately, most of them are deployed on land and observations are strongly limited by the poor instrumental coverage in ocean basins. Ocean-bottom seismometers (OBSs) have been considerably improved to fill this gap: (i) their autonomy can exceed 1 yr; (ii) the use of three-component, broad-band sensors is becoming standard; (iii) they are now easy to deploy and (iv) their performance are catching up with those of terrestrial stations, enabling them to measure faint signals such as Earth's free oscillations (Deen *et al.* 2017). Despite those impressive advances, the way OBSs are deployed has not changed much since their invention. In most cases, OBSs are

simply dropped in the ocean, straight above the selected landing sites. This implies two major technical issues: (i) Once landed, the OBS's seismometer activates its levelling mechanism to align its vertical component with the gravitational field. However, the orientation of its horizontal components remains unknown. (ii) During its descent in the water column, due to the combined effects of ocean currents and instrument hydrodynamics, the OBS undergoes a lateral shift, which may reach several hundreds of metres depending on the water depth and on the strength of the local currents. Until now, no simple, robust and cheap embedded device has been invented to deal with these issues. This paper proposes new and inexpensive methods to retrieve an OBS orientation and location with unprecedented accuracy, using ship-radiated noise (Fig. 1).



**Figure 1.** Overall sketch of this paper study. Ship noise emitted by merchant's vessels is analysed from seismic and hydroacoustic records to retrieve ocean bottom seismometer (OBS) orientation and localization on the ocean floor. To that end, two data sets are compared: the OBS recordings and the AIS data set of ship localizations.

### Existing sensor orientation methods

To take advantage of the three components of a seismometer, a precise knowledge of the orientation of the sensor is mandatory for many applications in Earth imagery, environmental seismology or bio-seismology. Unfortunately, the use of internal devices (compass, GPS), is hardly practicable or unaffordable, due to the presence of nearby electronics, magnets and/or magnetic rocks, as well as water attenuation of radio waves. Promising solutions such as MEMS-based gyrocompass (D'Alessandro & D'Anna 2014) are emerging but are still at prototype stages. Several methods have been therefore proposed to retrieve the horizontal orientation, by analysing *a posteriori* the recorded data. Most of the developed techniques compare the apparent incoming direction of arrival (DOA) of an event measured by the seismometer, with its theoretical azimuth, the difference indicating the OBS misorientation.

A variety of passive seismological techniques have been developed to orient the horizontal seismometer components, largely based on polarization attributes. Those techniques use a range of seismic observables including surface waves (Laske 1995; Larson 2002; Stachnik *et al.* 2012; Doran & Laske 2017; Scholz *et al.* 2017), body waves (Schulte-Pelkum *et al.* 2001; Fontaine *et al.* 2009; Scholz *et al.* 2017), receiver functions (Janiszewski & Abers 2015; Lim *et al.* 2017) and recordings of ambient noise (Grigoli *et al.* 2012; Zha *et al.* 2013). A crucial downside of passive methods exploiting natural seismic sources is that seismic waves travel through the Earth, which is generally heterogeneous and anisotropic, inducing seismic ray bending and modification of the polarization attributes. This makes it complicated to decipher the actual origin of the polarization deviation induced either by the sensor misorientation, or by the Earth's structure beneath the station (Fontaine *et al.* 2009). Moreover, several months are often required to record a large enough number of good quality teleseismic events, or to ensure that coherent information emerges from ambient noise correlation methods.

Active acoustics and seismic approaches are also widely used to successfully retrieve OBS orientations (Anderson *et al.* 1987; Duennebier *et al.* 1987). They consist in using an artificial source

(acoustic ping, airgun and explosion) to generate acoustic waves. Those approaches benefit from a high signal-to-noise ratio, from a perfect knowledge of the source locations and origin times and from a simple pattern of propagation. Indeed, water waves travel through the ocean which is an isotropic medium with small lateral heterogeneity. These waves do not deviate from their original direction in the horizontal plane, making it easy to pinpoint acoustic sources. Unfortunately, those approaches are time-consuming and require appropriate devices and on-board expertise.

### Existing sensor localization methods

Poor precision on sensors' locations can induce important localization errors when investigating nearby sources; it is also prejudicial for arrays of sensors when the exact geometry of the deployment matters. OBSs' locations are generally inferred from their deployment and recovery points, which relate, respectively, to the location where the OBS is dropped from the ship and to the location where the OBS is recovered at the surface after its deployment period. Unfortunately, as OBSs' descent and ascent in the ocean occur months apart and not necessarily at the same speed nor in the same direction, it is virtually impossible to retrieve the true path followed by the OBS and therefore to assess the accurate location of the sensor on the seafloor.

OBSs may also be located by using the acoustic transponder used to receive the release order (Creager & Dorman 1982), which can also be called upon to assess the distance between the OBS and the ship. By triangulation procedures or using an ultrashort baseline acoustic positioning system, it is possible to retrieve the OBS location. This operation is however time-consuming since one must at least wait for the OBS to reach the ocean floor before performing the localization. If no ultrashort baseline acoustic positioning system is available, it is furthermore necessary to send acoustic pings from a large range of azimuths around the OBS, which requires the ship to make either a large circle around the site or several crossing lines. For active seismic experiments, triangulating shots recorded by the OBS have been used instead of acoustical pings to save time

(Nakamura *et al.* 1987). For passive experiments, using active sources as a calibration step have mainly been applied on dense arrays of hydrophones (e.g. Dosso *et al.* 1998) since a single sequence of emissions could be used to calibrate all instruments at once, making the operation worthwhile. Using the above mentioned active sources methods usually allow to retrieve the OBS location within a few metres but all those field operations can significantly slow down an OBS network deployment. More recently, ambient noise correlation methods have been used on arrays of hydrophones of small spatial extent (tens of metres) to retrieve the relative location from one instrument to another (Sabra *et al.* 2005). In the case of networks with large interstation distances, none of these solutions is suited: (i) active sources cannot be recorded by all sensors at the same time requiring one triangulation step per instrument and (ii) noise correlation, even measurable for seismometers as far as 300 km (Hable *et al.* 2019), loses its accuracy.

### Ship noise as a source of opportunity

An alternative solution to these methods—which is the focus of this paper—is to use the ship-radiated noise as an acoustic source of opportunity to help locating OBSs and orienting their two horizontal components in the geographic reference frame (Fig. 1). Many merchant ships criss-cross the surface of the oceans and their positions are accurately known, thanks to their Automatic Identification System (AIS) designed to avoid collisions. Moreover, ship noise dominates the ocean acoustic landscape, from 10 to 200 Hz (Hildebrand 2009; Kinda *et al.* 2018) and is efficiently recorded by most OBSs.

Ship noise has already been used to localize hydrophone arrays (Ebbeson *et al.* 2008; Morley *et al.* 2009). In single sensor setups, measuring the distance of a ship knowing the instrument location has also been successful (Gao *et al.* 2008; Ferguson *et al.* 2017) along with geoacoustic inversion of the ocean bottom, on which the instrument was resting on (Gervaise *et al.* 2012). To our knowledge, however, no study has yet published an end-to-end solution to orient and/or localize an isolated OBS on the ocean floor using ship noise and its corresponding ship AIS locations.

For such purpose, we propose two new workflows to orient and localize OBSs, based on seismic and hydroacoustic recordings and the AIS trajectories of ships passing by during the OBSs deployment. The orientation workflow was applied to the whole RHUM-RUM network (Réunion Hotspot and Upper Mantle—Réunions Unterer Mantel) and a new reference table of orientations is provided. Obtained values are compared to previous results issued from the polarization analysis of  $P$  and Rayleigh waves from teleseismic events (Scholz *et al.* 2017). Requiring high quality AIS data, the localization workflow could not be applied to the whole network and was tested on a unique test station benefiting from good terrestrial AIS coverage.

## EXPERIMENT AND OBSERVABLES

### Seismic and hydro-acoustic data set

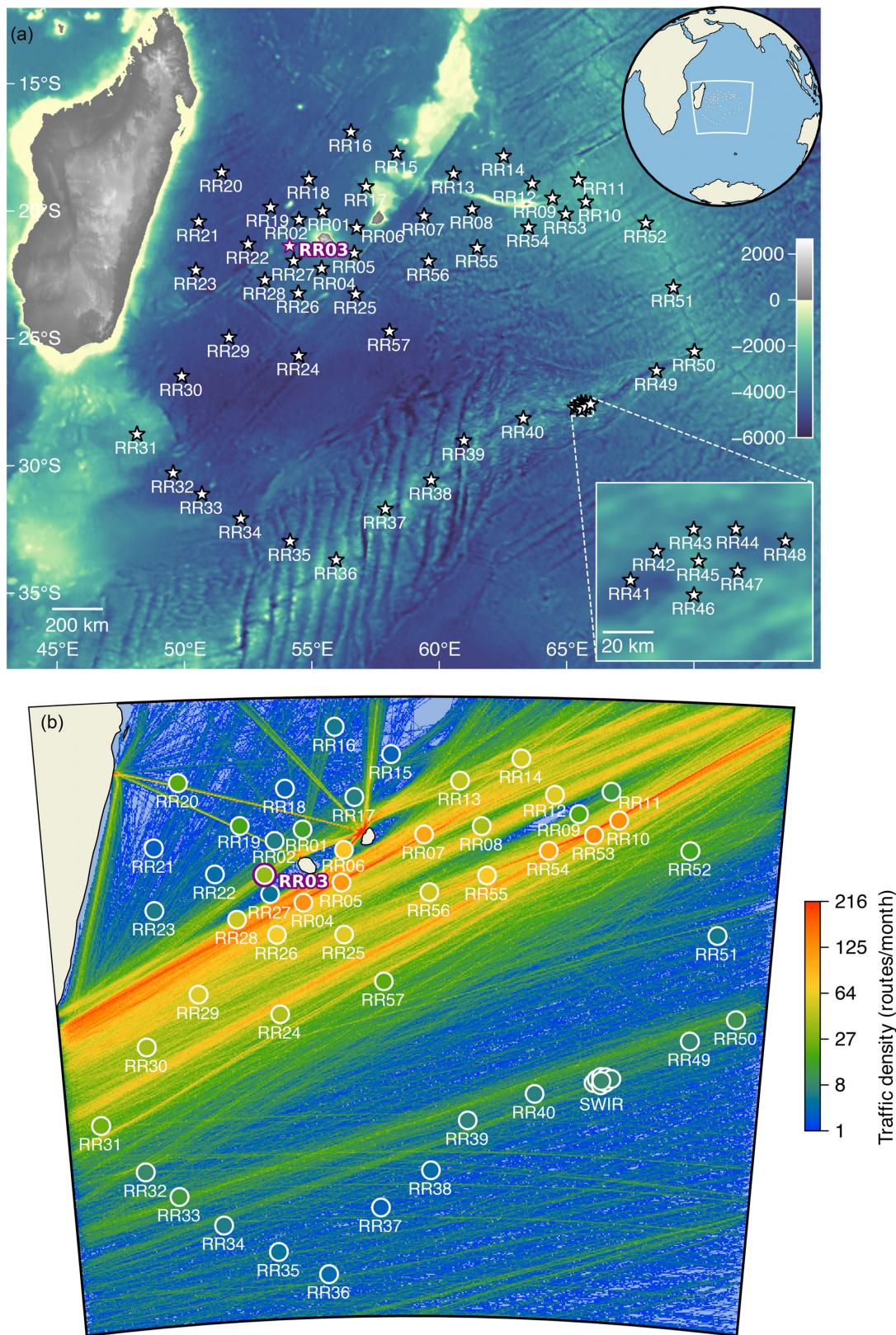
Seismic and hydro-acoustic data used in this study were recorded during the RHUM-RUM experiment (Barruol & Sigloch 2013). This French–German experiment in the Western Indian Ocean, was dedicated to image the whole mantle structure beneath the Réunion volcanic hotspot (Mazzullo *et al.* 2017; Scholz *et al.* 2018; Barruol *et al.* 2019) and covered an area spanning roughly 2000 km  $\times$

2000 km. RHUM-RUM allowed to deploy 57 OBSs from the German DEPAS (Deutscher Geräte-Pool für Amphibische Seismologie) and French INSU (Institut national des sciences de l’Univers) pools, between October 2012 and December 2013 (Fig. 2a). OBSs were not surveyed or otherwise located and their locations on their landing sites were approximated by their deployment locations. The distances at which OBSs were recovered from their deployment illustrate the uncertainty on their localizations (Fig. 3). This plot shows that half of the OBSs were recovered more than 500 m away from their deployment point, for depths ranging from 2200 to 5400 m. Each OBS was equipped by a three-component wide-band or broad-band seismometer, together with a broad-band pressure sensor. While OBS landing areas were chosen to be locally flat, the surrounding bathymetry varied a lot from an OBS to another (Fig. 2a). Seismic and hydroacoustic data was recorded continuously at 50, 62.5 or 100 samples  $s^{-1}$  depending on the stations. This limited our investigations to the Nyquist frequencies of, respectively, 25, 31.25 and 50 Hz. In the lower bound, the microseismic noise limited our investigation to frequencies above 1 Hz. Technical information and performance reports can be found in Stähler *et al.* (2016) and clock corrections in Hable *et al.* (2018). Interestingly, the OBSs orientations were previously determined from the polarization of  $P$  and Rayleigh waves from teleseismic events (Scholz *et al.* 2017). The RHUM-RUM data set is freely available at the RESIF seismological archive centre (see Acknowledgments).

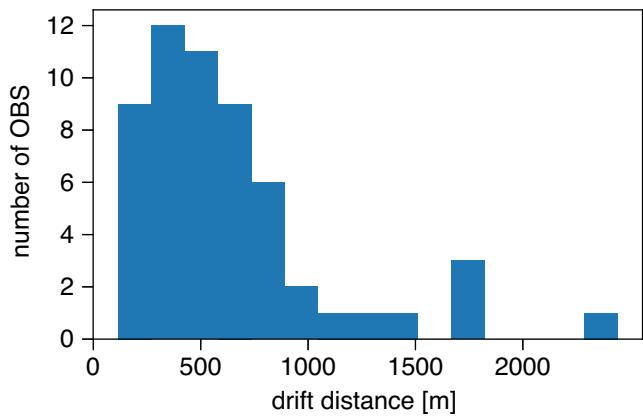
### AIS data set

Our method relies on the accurate knowledge of the noise source position, that is the ship position, which is provided by AIS archives. For maritime monitoring purposes, AIS radio signals are recorded using satellites or from antennas on the coast (at tens of kilometres range). Satellites allow a complete spatial coverage but implies intermittent temporal coverage cadenced by satellite passages which depend on the orbits of the AIS satellite fleets. The used AIS database covering the whole spatial and temporal extent of the RHUM-RUM experiment was purchased from CLS (Collecte Localisation Satellites). This data set provides ship positions on a 1-hr basis. As most ships usually travel with constant heading and constant speed in high seas, positions between two GPS locations were computed by linear interpolation. This may cause positional errors of a few hundreds of metres. It generally does not imply large azimuthal errors at typical OBS–ship ranges and therefore does not limit the reorientation process. Hourly AIS data limits, however, the localization process which accuracy directly depends on the accuracy of inputted ship positions. In order to test the proposed method, an OBS close to La Réunion island (station RR03) was therefore used. Terrestrial AIS data collected in November 2012 by antennas located on land was used, providing ship positions every three minutes.

From visual inspection of spectrograms at expected times of ship passages, we limit ourselves to big vessels like cargos and tankers, which generate high noise levels. Smaller ships can also be detected but are considerably less noisy and usually do not compete with measurements issued from large vessels. The AIS data set was therefore filtered out by type to only retain GPS tracks of cargos and tankers. Unless otherwise specified, ship will refer in the following of this paper to cargo or tanker. To ensure good signal quality, we focused on ships noise produced within 30 km of OBSs, even if they can be detected at more than 100 km. From that perspective, it was decided to only retain ship trajectories with a closest point of



**Figure 2.** RHUM-RUM experiment overview. (a) Bathymetric map of the RHUM-RUM experiment area. A total of 57 OBS were deployed in areas of different topographies. The test station used to illustrate this study (RR03) is highlighted in purple. (b) Map of the marine traffic density obtained from the AIS archives during the whole RHUM-RUM experiment. Circular areas of 30 km are plotted as white disks and represent the distance at which ship noise usually has a good signal quality. Colours inside the circles indicate the monthly average number of ships passing closer than 15 km to an OBS. On the rest of the map, marine traffic density is plotted with the same colour scale. Large differences in the number of ships passing close to an OBS can be observed whether the OBS is located beneath a marine commercial route or not.



**Figure 3.** Histogram of distances between OBS dropping and recovery points of the 57 instruments deployed during the RHUM-RUM experiment. The median drift distance is about 500 m for station deployed at depths ranging from 2200 to 5400 m.

approach (CPA) less than 15 km away from OBSs, so that the ship stayed in a 30 km radius area during a satisfactory time duration.

Fig. 2(b) presents the ship density traffic in the RHUM-RUM area as the number of ship passages in a radius of 15 km per month. The dominant ship traffic is mainly channelled along NE-SW trending commercial routes linking SE-Asia and the south Atlantic. The ships' density in the neighbourhood of an OBS varied a lot according to its location, ranging between a few to hundreds of ships per month. For each station, white circles represent the 30 km radius area used in this study, and the filling colour indicates the traffic density within the circle. For each ship route fulfilling the 15 km condition, we extracted the relevant part of the trajectory where the ship noise should be visible and therefore analysed (Fig. 4a).

### Ship noise signature

As a ship passes in the neighbourhood of an OBS, it radiates noise that is recorded by the seismic and hydroacoustic sensors on the ocean floor. Careful examination of the spectrogram of the hydroacoustic or seismic channels (Fig. 4b), shows that a ship generates a clear increase in the noise level that starts and ends several hours before and after the CPA, depending on the ship velocity and size and on the CPA distance. Ship-radiated noise is mainly characterized by monochromatic bands of energy, spreading from a few Hertz up to the maximum frequency that could be investigated, suggesting that the ship noise spectrum contains even higher frequency signals. Each ship has a signature composed of several monochromatic bands, related to the rotational speed of its engines, the number of propellers and the number of blades of each propeller. Each monochromatic band of energy is affected by a Doppler effect at the CPA, of the order of 0.1 Hz, depending on the ship speed. Interestingly, when several ships are transiting nearby, their frequency bands may overlap, but their Doppler effect helps to discriminate them.

## METHODOLOGY

To retrieve the orientation or the location of an OBS, efficient methods measuring the direction and the distance of a source of known location using the seismometer and the hydrophone of the OBS are necessary. The three-component seismometer enables to measure local ground motion. As a wave passes by, it locally generates a 3-D oscillation whose motion relative to the direction

of propagation defines its polarization. Each OBS also embeds a hydrophone that measures the pressure changes produced by the passage of compressional waves in the water. These two types of sensors allow to use methodologies developed for seismology, for hydroacoustics but also allows using so-called acoustic vector sensors (AVS) methods, involving a combination of both sensors (Cao *et al.* 2017). For a more in-depth understanding of the following proposed methods, the developed codes used in this article can be found in an open-source python package named *Obsea* (see Acknowledgments).

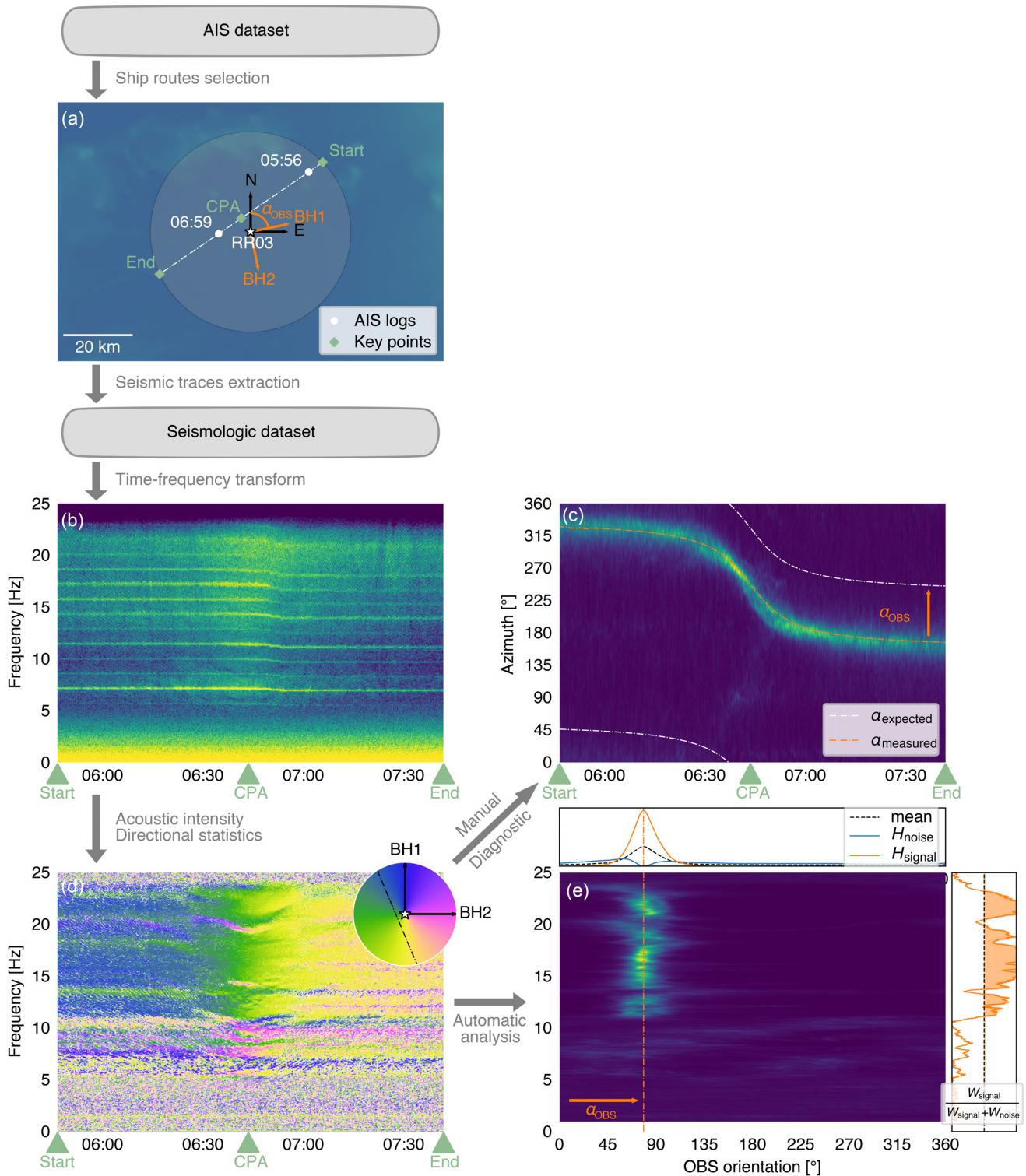
### Orienting OBSs

The proposed method to quantify the OBS orientation relies on a local measurement of the direction of propagation of waves emitted by ships of known trajectories. Ship noise propagates from the source to the OBS as *P* waves, since compressional waves are the only type of wave travelling through fluids. A minor part of *P* waves is also converted to *SV* waves at the ocean floor interface and will be omitted in this approach. The particle motion produced by an incoming *P* wave is collinear to the wave propagation. Its polarization is rectilinear as motion occurs along a straight line. When recording a rectilinear motion and making the hypothesis of an incoming *P* wave, the collinear property is not sufficient to fully constrain the direction of the source: it might as well come from a direction as its opposite. This issue is sometimes named *the 180° ambiguity*.

In seismology, when locating teleseismic events from a single station, two properties are used to retrieve correct DOA from incoming *P* and Rayleigh waves: (i) *P* waves from teleseismic sources come from the Earth's depths and the DOA must be chosen so that it points toward the ground. (ii) Rayleigh waves from teleseismic sources have a retrograde vertical elliptic motion which direction of rotation indicates the direction of the source (Schimmel & Gallart 2004; Schimmel *et al.* 2011). Using those properties, teleseismic events DOA can be measured, and in return used to orient OBSs (Scholz *et al.* 2017). Unfortunately, those properties, cannot be used to retrieve ship noise DOA from an OBS: (i) At long ranges, *P* waves produced by ships arrive with large incidence angles, and because of the impedance increase between the water and the ground, total reflection occurs and the energy travels horizontally. This prevents any reliable measurement of the vertical incidence angle (Harris *et al.* 2013). Although incidence angles may decrease at short distances, the poor reflectivity of the water/sediment interface may generate complex wave behaviour locally, for instance from *P*-wave reflected at deeper earth interfaces and therefore propagating towards the surface. (ii) Rayleigh waves do not propagate into the water as they are interface waves. Another methodology is therefore required to measure full 360° DOA.

### Acoustic intensity based DOA measurement

To overcome this 180° ambiguity and to get the true azimuth of the incoming ship noise, we took advantage of the fourth OBS component provided by the hydrophone. Compressional waves (also called pressure waves) propagate so that the pressure  $p(\mathbf{r}, t)$  at position  $\mathbf{r}$  and time  $t$  evolves in phase with the particle motion  $\mathbf{v}(\mathbf{r}, t)$  in the direction of propagation which is the opposite of the DOA  $\mathbf{u}$ . Euler's equation summarized this property for a plane wave



**Figure 4.** Reorientation workflow for a single ship passage. (a) The AIS data set is processed and a route of a ship passing 15 km or less away from the OBS is extracted (white circles connected by white dash-dotted lines). Start time and end time of the seismological signal of interest are defined by the ship entry and exit time of the 30 km radius area (grey shaded circle). The OBS's horizontal components BH1 and BH2 (orange arrows) are misoriented by a yet unknown angle  $\alpha_{OBS}$  with the geographical reference frame (black arrows). (b) Spectrogram in dB scale computed from the hydrophone channel (high levels in yellow). Ship noise mainly consists of a series of monochromatic bands and harmonics. A slight decrease in pitch can be noted before 7:00 because the ship slightly slows down. (c) Azigram: DOAs in the OBS reference frame is colour-coded and MRL (*Mean Resultant Length*, see text) is transparency-coded. The colour to angle correspondence is plotted in the inset along with a probable guess of the ship path (black dash-dotted line). (d) Time-azimuth representation (high probability in yellow). The theoretical azimuthal curve of the ship (calculated from its AIS trajectory, white dashed line) needs to be translated by an amount equal to the OBS orientation to match the measurements. (e) Orientation-frequency representation (high probability in yellow) and its NMF (*non-negative matrix factorization*, see text) decomposition into signal and noise. The  $W_{signal}$  part exhibits a clear peak which maximum is used as the estimate of the OBS representation. The  $W_{ratio}$  indicates in which frequencies ship noise predominates, that is when  $W_{signal}$  exceeds  $W_{noise}$  (orange fill).

in an infinite homogeneous medium:

$$\mathbf{v}(\mathbf{r}, t) = -\frac{p(\mathbf{r}, t)}{\rho c} \mathbf{u}. \quad (1)$$

Here  $\rho$  is the medium density and  $c$  is the medium sound velocity. In the case of an OBS, the position is stationary, so we will omit the argument  $\mathbf{r}$  from the pressure and the particle velocity.

In the AVS research field, different methods have been developed to retrieve accurate DOA. Among those, intensity-based methods are the most straightforward and use this phase relation between pressure and velocity (Nehorai & Paldi 1994) and have already been successfully used for analysing data recorded in the ocean (e.g. D'Spain *et al.* 1991; Thode *et al.* 2019). The instantaneous acoustic intensity  $\mathbf{I}(t)$  is a vector which points into the direction of propagation of the energy while its amplitude quantifies the flux of energy. It is defined as the product of the pressure and the particle velocity:

$$\mathbf{I}(t) \equiv p(t) \mathbf{v}(t) = -\frac{p(t)^2}{\rho c} \mathbf{u}. \quad (2)$$

For monochromatic acoustic fields (or for each component obtained by decomposing the signal with the Fourier transform), the complex representation can be used. The active acoustic intensity  $\mathbf{I}_a$  is defined as half the real part of the conjugate of the pressure  $p^*$  multiplied by the velocity  $\mathbf{v}^*$  (Mann *et al.* 1987):

$$\mathbf{I}_a \equiv \frac{1}{2} \operatorname{Re}(\underline{\mathbf{v}} \underline{p}^*) = -\frac{|p|^2}{\rho c} \mathbf{u}. \quad (3)$$

Both the instantaneous and the active acoustic intensities point in the opposite direction of  $\mathbf{u}$  as shown in eq. (2) and eq. (3). Therefore, to estimate the incoming DOA unit vector, the instantaneous or the active acoustic intensity can be used as follows:

$$\mathbf{u} = -\frac{\mathbf{I}(t)}{\|\mathbf{I}(t)\|} = -\frac{\mathbf{I}_a}{\|\mathbf{I}_a\|}. \quad (4)$$

All previous equations hold for a unique plane wave propagating in an infinite homogeneous medium. For an OBS lying on the ocean floor, this is not the case: (i) the OBS is located at the boundary between the ocean and the solid earth, that is at an interface of high impedance contrast. The OBS hence records the combination of incident, reflected, refracted and converted waves. (ii) For a given source, several paths are usually possible (waves reflections and/or refraction, complex bathymetry), hence the arrival of several waves from different apparent directions and different arrival times. (iii) The ocean is far from being silent. Numerous sources can be simultaneously recorded at any time and the background noise usually has preferred directions due to the inhomogeneous repartitions of sources in the ocean (in the frequency band of investigation it is primarily due to the concentration of ships along commercial routes).

Assuming that the medium has no or little lateral variations (range-independent model), the effect of the two first issues can be neglected. In such case, we can restrict the analysis to the radial plane, defined as the vertical plane joining the source and the OBS. As the propagation is contained within the radial plane, no transverse particle motion to that plane can occur ensuring that the projection  $P_h(\mathbf{u})$  in the horizontal plane of the acoustic intensity raises the correct azimuth whereas the measured elevation can point in any direction. Thereafter, we will use this normalized projection on the horizontal plane (here  $\Pi_h$  is the projector on the horizontal

plane):

$$\mathbf{u}_h = \frac{\Pi_h(\mathbf{u})}{\|\Pi_h(\mathbf{u})\|}. \quad (5)$$

For sake of clarity, we will omit the  $h$  index and if not explicitly stated, DOA will stand for horizontal DOA.

### Sparsity-based source separation

To compute the direction of a source in a multisource environment, a sparsity-based source separation workflow was used. This kind of method tackles the separation with geometrical considerations (Theis *et al.* 2003) which can be directly translated in terms of the desired source DOA. Sparsity-based source separation workflows consist of three main steps (Gribonval & Zibulevsky 2010). First, a joint sparse transformation is applied to make the supports of the source coefficients in the new representation as disjoint as possible. Secondly, source DOAs are estimated from a direction (or mixture) representation. Finally, if needed, sources can be reconstructed from those estimated DOA.

The short-term Fourier transform (STFT) was used as a joint sparse representation. Noise from different sources is expected to be localized in different time–frequency areas and not to overlap in both time and frequency. Ship noise energy, for example, is spread across time but is mainly concentrated into some small frequency bands (see example Fig. 4b), whereas earthquake signal can spread over many frequencies but is only present during small periods. On real data, perfect separation never occurs but areas where a source significantly predominates over the others can generally be found. STFT might not be the most efficient sparse representation but has the great advantage to be fast to compute and has simple meaning in physics. Additionally, it allows to remove the instrument responses of each channel by water-level deconvolution, which consists in a spectrum division applied to the overall STFT of the signal. In this study, horizontal components of deployed OBS had balanced responses and correcting the small phase mismatch between the seismometer horizontal components and the hydrophone components on the frequency range of interest (above 1 Hz) did not raise any improvement. For simplicity, no instrument correction was performed. STFT was applied to all four channels with a sliding window length of 1024 samples and an overlap of 87.5 per cent (Fig. 4b).

### Azigram representation

Applying eq. (3) to eq. (5) elementwise on the four-component STFT provides a time–frequency representation of DOA. For each time–frequency pixel, directions are colour-coded as azimuths; therefore, we called this representation an azigram, as proposed by Thode *et al.* (2019). To filter relevant DOA from noise, DOA stability is analysed through time. We drew inspiration from the definition of the degree of polarization (DOP) developed by Schimmel & Gallart (2004) and from the analysis of the mixture ratio variance developed by Abrard & Deville (2005). Both approaches assume that, when a mixture of two or more sources are recorded at the same time, the DOA (or similarly the mixture ratio) cannot be stable through time. To measure this stability and estimate the source direction, the mean direction  $\bar{\mathbf{u}}[i, j]$  and the mean resultant length (MRL)  $R[i, j]$  (where  $i$  is the index of time and  $j$  the index of frequency) are computed on small time windows of several azigram values  $N$  (it was chosen  $N = 8$  which gives a time window of

about 10 s according to the chosen STFT parameters and raw signal sampling rate):

$$\mathbf{m}[i, j] = \frac{1}{N} \sum_{k=i-N/2}^{i+N/2} \mathbf{u}[k, j] \quad (6)$$

$$\bar{\mathbf{u}}[i, j] = \frac{\mathbf{m}[i, j]}{\|\mathbf{m}[i, j]\|} \quad (7)$$

$$R[i, j] = \|\mathbf{m}[i, j]\|. \quad (8)$$

The resultant MRL can vary from zero, in case of completely random directions, to one, in case all directions sum up constructively. This MRL can then be used to compute a transparency mask to remove the unstable portions of the azigram formed from the mean directions. An azigram of a passing ship is shown in Fig. 4(c). The DOA of the ship can be efficiently computed mainly for frequencies above 10 Hz. Colours sweep from blue to yellow indicating a variation of the direction of roughly 180° which corresponds to a rectilinear ship passage close to the OBS.

#### Time-azimuth representation

If the azigram representation is useful to explore the signal directivity through time and frequency and to separate indistinguishable mixtures of noise from prominent signals, it can hardly be used to retrieve quantitative values of DOA. A better representation to extract directions can be constructed under the following assumptions: (i) the DOA of any source is assumed to be frequency invariant; (ii) the DOA of each source is assumed to remain stable over small time windows of a few minutes. Based on this, values contained in windows spanning all frequencies over small time durations (about 80 s was used) can be treated regardless of their position in the time–frequency plane. A weighted kernel density estimation (KDE) was used. Mean directions were first binned and weighted according to their MRL to compute a weighted histogram (a resolution of a tenth of a degree was chosen for the binning) and then smoothed with a Gaussian kernel of a few degrees of standard deviation (2° was chosen). Finally, normalization was performed so that the integral over all azimuths of the obtained density is equal to one when all weights (i.e. all MRL) are equal to one. This results in a pseudo probability density function (PPDF) of measured orientation. It is not a PDF function since its integral is smaller than one, but it allows to take into account the possibility of measuring nothing, which would result in a null flat curve. The process was repeated over the above-defined sliding windows resulting in a new time-azimuth representation, shown in Fig. 4(d), where each column is a PPDF indicating the repartition of measured azimuths at a given time in the associated time–frequency window. Each mode (i.e. local maximum) of those PPDFs hopefully corresponds to DOA of a noise source.

Overlaying the theoretical curve (white dashed line Fig. 4d) of the ship azimuth variation estimated from its AIS data on top of this time-azimuth representation exhibits the misorientation bias. The translation ( $\alpha_{\text{OBS}}$  on Fig. 4d) needed to shift the curve on top of highly probable areas indicates the OBS misorientation. This shifting allows a good visual check of the consistency between the measured azimuth and the expected one and allows to manually assess a first estimate of the OBS orientation.

#### Orientation-frequency representation

The time-azimuth representation is relevant to track moving sources. The OBS orientation azimuth  $\alpha_{\text{OBS}}$  is time-invariant, so at this stage, integration over the whole ship passage duration was performed. The azimuthal curve  $\alpha_{\text{expected}}(t)$  in the geographical coordinate reference system (CRS) is computed using the AIS positions of the ship. To convert the DOA from the geographical CRS into the OBS CRS, adding the OBS orientation is sufficient:

$$\alpha_{\text{expected}}(t) = \alpha_{\text{measured}}(t, f) + \alpha_{\text{OBS}}, \quad \forall(t, f) \in D_{\text{ship}}, \quad (9)$$

where  $D_{\text{ship}}$  is the time–frequency plane domain where ship noise predominates. This holds in theory, but in practice the measured DOA is impacted by noise and slightly deviates through time and frequency. Subtracting the measured DOA in the geographical CRS from the expected DOA in the OBS CRS is then a time (and frequency) invariant estimate of the OBS orientation:

$$\alpha_{\text{OBS}} = \alpha_{\text{expected}}(t) - \alpha_{\text{measured}}(t, f), \quad \forall(t, f) \in D_{\text{ship}}. \quad (10)$$

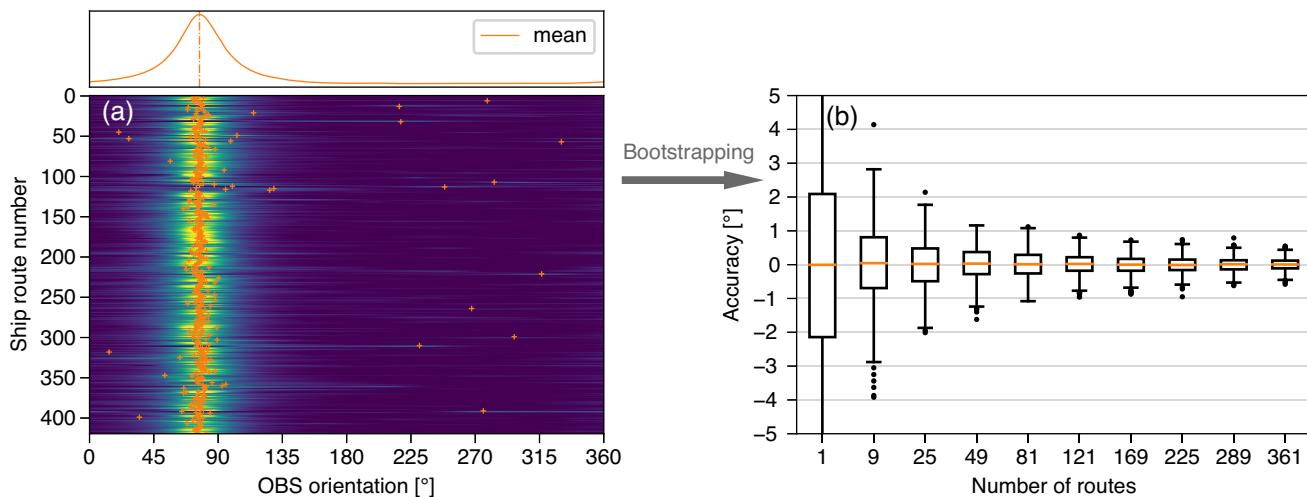
This OBS orientation estimate was performed for each value of the azigram and allowed to introduce a new orientation-frequency presented in Fig. 4(e). The methodology used is the same than the one used to compute the time-azimuth representation except that: (i) values of the azigram were grouped per frequency and not by small vertical time windows, (ii) the estimates of OBS orientation obtained using eq. (10) are used instead of raw azigram values. For each frequency, the obtained group of OBS orientation estimates are used to compute PPDFs which are then gathered in the new orientation-frequency representation (Fig. 4e). Each row represents a PPDF of measured OBS orientation for a given frequency. Frequency bands where the ship signal predominates are easily identifiable visually as they show a clear peak at the most probable OBS orientation.

#### Extraction of the OBS orientation

The OBS orientation is also frequency-invariant, but due to its harmonic structure, ship noise generally dominates on some frequency bands. While not totally the case, it was supposed that PPDFs can be satisfactorily approximated as a linear mixture of two components: (i) a component representing the PPDF of measured OBS orientation when the ship signal predominates, (ii) a component representing the background anisotropic noise. To decompose orientation-frequency representations into those two components, non-negative matrix factorization (NMF) was used (Lee & Seung 1999). NMF combines two properties: (i) Dimensionality reduction, that is the property of summarizing all the information in a few features. (ii) Automatic clustering, that is the property of grouping similar objects. In this particular case, frequencies where ship noise predominates are expected to be separated from frequencies where the background noise predominates. NMF decomposes a 2-D matrix  $V$  into the product of two lower rank matrices  $W$  and  $H$  (in our case of rank two):

$$V \approx WH \approx [W_{\text{signal}} \quad W_{\text{noise}}] \begin{bmatrix} H_{\text{signal}} \\ H_{\text{noise}} \end{bmatrix}. \quad (11)$$

NMF requires all entries in  $V$ ,  $W$  and  $H$  to be positive. In our case, the two rows of  $H$  are a base of PPDF of OBS orientation and  $W$  indicates at which frequency each of those rows are present. As the factorization is not unique, it was chosen to normalize the columns of  $W$  so that their maximum is one.  $H$  was corrected



**Figure 5.** OBS orientation statistics from multiple ship passages. (a)  $H_{\text{signal}}$  for the entire 420 ships routes used for reorienting station RR03 (high probabilities in yellow). Orange crosses mark each maximum (outliers can be noted). To aggregate all measurements, stacking is performed (top curve) and the maximum of the resulting curve is used as the OBS orientation final estimate. (b) Uncertainties as a function of the number of ships used for determining the actual orientation of the RR03 OBS, determined by bootstrapping. Standard boxplot with median, interquartile range (IQR) and whiskers with maximum 1.5 IQRs are plotted for several numbers of ships passages. Great improvements are observed until about 50 ships which are probably due to a good outlier rejection. An uncertainty of less than one degree is achieved at about 80 ships (with 99 per cent confidence).

accordingly.

$$\|W_{\text{signal}}\|_\infty = \|W_{\text{noise}}\|_\infty = 1. \quad (12)$$

To assign each extracted component by the NMF algorithm to the signal part or the noise part, the  $H$  component with lower circular standard deviation was interpreted as the signal part. Fig. 4(e) illustrates the decomposition, via NMF, of the orientation-frequency representation. The dimensionality reduction property can be observed on the two components of  $H$ : the  $H_{\text{signal}}$  part summarizes the PPDFs of measured OBS orientation while the  $H_{\text{noise}}$  part summarizes the noise. The automatic clustering property can be observed on the two  $W$  components. When  $W_{\text{signal}} > W_{\text{noise}}$ , it can be considered that ship noise predominates. NMF succeeds in extracting the desired information with two components whereas in low signal-to-noise scenarios, simply averaging all the per frequency PPDF results in a meaningless flat curve.

#### Statistics on ship passages

PPDFs computed for each ship passage were simply stacked (i.e. averaged) to blindly aggregate information from multiple ship passages. The mode (i.e. the maximum) of the resulting mean PPDF was used as the overall estimate of the OBS orientation. Stacking PPDFs instead of taking the average of the individual modes allows a more robust estimation. Outlier peaks do not contribute to the main peak and therefore the result is not altered. The analysis of a single ship passage as reported in Fig. 4, was repeated for all ship passages recorded at station RR03 and the PPDF were stacked as shown in Fig. 5(a). A clear peak indicates the overall estimate of the OBS orientation.

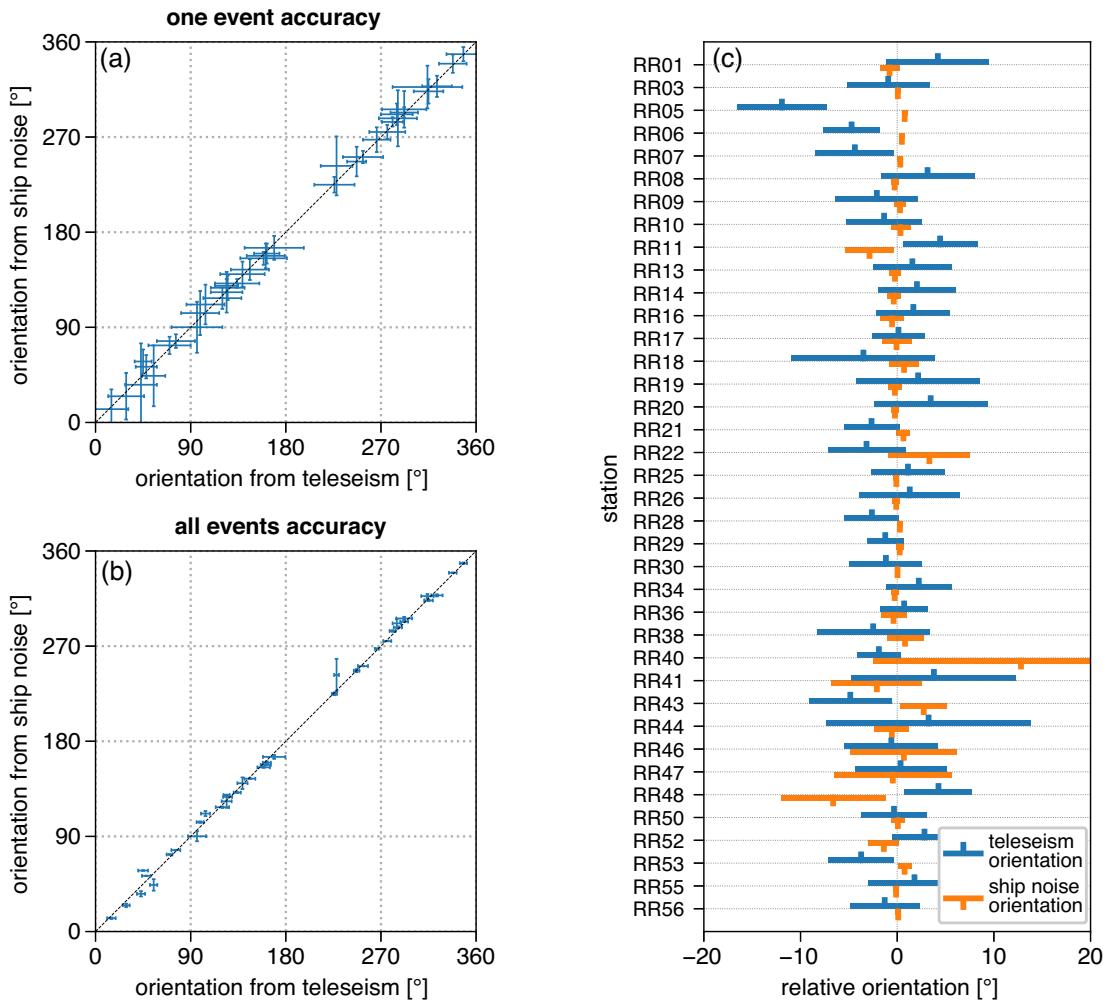
We estimated uncertainties relying on one ship passage by computing standard metrics on all the OBS orientation estimates provided by the whole ship traffic. Bootstrapping techniques were used to account for the gain in accuracy when aggregating information from several passages. Virtual collections of ship passages are created by randomly sampling a given number of ship passages from the actual collection (each ship passage can be potentially drawn

multiple times). Each virtual collection is considered as a potential traffic scenario and an associated OBS orientation is computed. Comparing numerous results from the set of virtual scenarios can then be used to estimate uncertainties with the desired metrics. We chose to build 1000 virtual collections to evaluate the uncertainties and to obtain robust statistical estimates. The process was then repeated with collections featuring different numbers of ship passages, to simulate sets of scenarios with different traffic density.

For each set of virtual collection, the estimated set of orientation and the orientation obtained using all ship passages were compared. This allows performing standard boxplot of errors considering that reference as the best assessment of the real OBS orientation. Fig. 5(b) shows the result for station RR03. From zero to around 50 ships, we observe a great improvements in the knowledge of the final OBS orientation, probably because outlier effects are efficiently absorbed by stacking. Beyond 50 ships, the progression is slower and roughly follows a square root law dictated by the law of large numbers. In that case, for fewer than 50 ships, better results could be performed by manually removing outliers but this was considered as a useless task as more ships passages were available. It is important to point out that while uncertainties smaller than one degree seem to be largely attained, it cannot exclude the presence of bias. Not perfectly balanced instrument response, complex topography, strongly anisotropic background noise, etc. could potentially produce biased results which still would lead to a small spreading of obtained OBS orientations.

#### Application to the whole network

For each station, the full workflow of orientation retrieval was performed in order to estimate three values: (i) the global estimated OBS orientation using all ship routes  $\alpha_{\text{OBS}}$ ; (ii) the 95 per cent uncertainties interval on this orientation estimated from the bootstrap method using all ship routes  $2\sigma_{\text{all}}$  and (iii) a robust estimation of the 95 per cent uncertainties interval on the orientation using isolated ship passages  $2\sigma_{\text{one}}$ . For this last value, interquartile range (IQR) was used assuming a Gaussian distribution. This robust estimator



**Figure 6.** Comparison of OBS orientations between our results and those obtained from a previous study analysing teleseismic wave polarizations (Scholz *et al.* 2017). (a) Comparisons of accuracies of teleseismic versus ship noise techniques for a single event shows good matches within their uncertainties. (b) Using all events the uncertainties shrink and in some case, values obtained from ship noise and teleseismic analysis do not perfectly agree. (c) Relative orientation to the average issued from the two studies allows to zoom on the error bars and better appreciate agreements between results.

rejects outliers which, in case ship passages are studied separately, would have been manually removed. This accuracy without outliers is a good indicator of the capacity of a station to measure a source direction as it does not depend on the number of recorded ships. Processing more than 15 000 ship passages (each roughly representing 2 hr of real-time data) took, on average, a few hours on laptop computer.

All results are summed up in Table 1. Orientations were computed for 38 stations over the 57 deployed during the RHUM-RUM experiment. For the 19 remaining stations, technical issues prevented the retrieval of the OBS orientation (Stähler *et al.* 2016; Scholz *et al.* 2017). Uncertainties using all ship passages vary from such small angles as 0.3° up to 6.1° (except for RR40 where stacking ships passages did not allow to increase the accuracy, likely due to strong topographic effects in the southwest Indian ridge area).

#### Comparison with a previous seismologic study

To validate our research, results of a previous work using teleseismic events to retrieve OBS orientation (Scholz *et al.* 2017) were used. This work proposed two independent methods to retrieve

OBS orientation based on teleseismic *P* and Rayleigh waves polarizations. From the associated results reported in Table 1, values using the *culled R-pol* method with circular mean estimator were used. Whereas it was stated that the *P-pol* method raised slightly better results (Scholz *et al.* 2017), the *R-pol* measurements benefit from being available for all working stations, which is not the case for the other method. From those values, the standard error of the mean (SEM) is extrapolated by dividing the given standard deviation by the square root number of used earthquakes to take into account the gain in accuracy obtained by gathering information from different events. These standard uncertainties are equivalent to the  $\sigma_{\text{all}}$  of this paper.

Retrieved orientations from both studies were compared along with their related uncertainties (Fig. 6). In most cases, both orientations agree and fall within the uncertainty of each other. Some mismatches can be observed but can be explained by several reasons: (i) In some cases, the small number of events used by the seismic method does not ensure a reliable estimate of the orientation and its related uncertainty (e.g. station RR05). (ii) Simplistic assumptions were made by both studies: neither the local structure beneath the sensor nor the topography was taken into account which can result in biased results for some stations. The mean

**Table 1.** Retrieved orientations for the RHUM-RUM network by ship noise analysis and by teleseisms analysis (Scholz *et al.* 2017). For each method, a set of metrics are compared:  $N$  indicates the number of events used (ship passage or seismic);  $\alpha_{\text{OBS}}$  is the retrieved orientation;  $2\sigma_{\text{all}}$  is the 95 per cent confidence interval aggregating all events and;  $2\sigma_{\text{one}}$  is the 95 per cent confidence interval of separately observed events. Red highlighted lines are stations for which measurements could not be performed because at least one of the sensor of the OBS failed or was too noisy. For two stations (RR12, RR31), only the hydrophone failed which prevented the use of our method, whereas the teleseismic based method worked. Grey highlighted lines are INSU stations equipped with differential pressure gauges and broadband sensor decoupled from the OBS structure. More information related to the OBS are available through the RESIF web services (see Acknowledgements).

| Station | Deployment |           |       | Ships |                       |                        |                        | Teleseisms |                       |                        |                        |
|---------|------------|-----------|-------|-------|-----------------------|------------------------|------------------------|------------|-----------------------|------------------------|------------------------|
|         | Latitude   | Longitude | Depth | $N$   | $\alpha_{\text{OBS}}$ | $2\sigma_{\text{all}}$ | $2\sigma_{\text{one}}$ | $N$        | $\alpha_{\text{OBS}}$ | $2\sigma_{\text{all}}$ | $2\sigma_{\text{one}}$ |
| RR01    | -20,0069°  | 55,4230°  | 4298  | 190   | 318,0°                | (1,0°)                 | (10,0°)                | 7          | 323°                  | (5,3)°                 | (14)°                  |
| RR02    | -20,3392°  | 54,4984°  | 4436  |       |                       |                        |                        |            |                       |                        |                        |
| RR03    | -21,3732°  | 54,1294°  | 4340  | 420   | 77,0°                 | (0,3°)                 | (6,4°)                 | 18         | 76°                   | (4,2)°                 | (18)°                  |
| RR04    | -22,2553°  | 55,3846°  | 4168  |       |                       |                        |                        |            |                       |                        |                        |
| RR05    | -21,6626°  | 56,6676°  | 4092  | 1388  | 57,7°                 | (0,3°)                 | (11,4°)                | 3          | 45°                   | (4,6)°                 | (8)°                   |
| RR06    | -20,6550°  | 56,7639°  | 4216  | 901   | 129,2°                | (0,3°)                 | (6,5°)                 | 14         | 124°                  | (2,9)°                 | (11)°                  |
| RR07    | -20,1945°  | 59,4058°  | 4370  | 1213  | 52,7°                 | (0,3°)                 | (11,0°)                | 6          | 48°                   | (4,1)°                 | (10)°                  |
| RR08    | -19,9259°  | 61,2907°  | 4190  | 453   | 157,6°                | (0,4°)                 | (11,7°)                | 14         | 161°                  | (4,8)°                 | (18)°                  |
| RR09    | -19,4924°  | 64,4485°  | 2976  | 225   | 127,4°                | (0,6°)                 | (11,6°)                | 14         | 125°                  | (4,3)°                 | (16)°                  |
| RR10    | -19,6437°  | 65,7558°  | 2310  | 1472  | 287,7°                | (1,0°)                 | (26,5°)                | 21         | 286°                  | (3,9)°                 | (18)°                  |
| RR11    | -18,7784°  | 65,4629°  | 3941  | 156   | 35,7°                 | (2,5°)                 | (39,1°)                | 15         | 43°                   | (3,9)°                 | (15)°                  |
| RR12    | -18,9255°  | 63,6474°  | 3185  |       |                       |                        |                        | 10         | 27°                   | (2,8)°                 | (9)°                   |
| RR13    | -18,5427°  | 60,5635°  | 4130  | 516   | 313,2°                | (0,6°)                 | (11,5°)                | 12         | 315°                  | (4,0)°                 | (14)°                  |
| RR14    | -17,8448°  | 62,5299°  | 3420  | 618   | 12,6°                 | (0,7°)                 | (18,7°)                | 16         | 15°                   | (4,0)°                 | (16)°                  |
| RR15    | -17,7402°  | 58,3330°  | 3959  |       |                       |                        |                        |            |                       |                        |                        |
| RR16    | -16,8976°  | 56,5335°  | 4426  | 62    | 159,8°                | (1,2°)                 | (9,5°)                 | 10         | 162°                  | (3,8)°                 | (12)°                  |
| RR17    | -19,0427°  | 57,1322°  | 2205  | 68    | 246,8°                | (1,5°)                 | (14,0°)                | 11         | 247°                  | (2,7)°                 | (9)°                   |
| RR18    | -18,7504°  | 54,8878°  | 4743  | 35    | 296,2°                | (1,5°)                 | (17,3°)                | 8          | 292°                  | (7,4)°                 | (21)°                  |
| RR19    | -19,8500°  | 53,3805°  | 4901  | 186   | 117,6°                | (0,7°)                 | (10,1°)                | 8          | 120°                  | (6,4)°                 | (18)°                  |
| RR20    | -18,4774°  | 51,4600°  | 4820  | 187   | 155,3°                | (0,4°)                 | (6,1°)                 | 14         | 159°                  | (5,9)°                 | (22)°                  |
| RR21    | -20,4217°  | 50,5599°  | 4782  | 31    | 284,3°                | (0,7°)                 | (4,4°)                 | 12         | 281°                  | (2,9)°                 | (10)°                  |
| RR22    | -21,3007°  | 52,4994°  | 4920  | 43    | 291,5°                | (4,2°)                 | (10,2°)                | 14         | 285°                  | (4,0)°                 | (15)°                  |
| RR23    | -22,3290°  | 50,4487°  | 4893  |       |                       |                        |                        |            |                       |                        |                        |
| RR24    | -25,6805°  | 54,4881°  | 5074  |       |                       |                        |                        |            |                       |                        |                        |
| RR25    | -23,2662°  | 56,7249°  | 4759  | 691   | 274,8°                | (0,3°)                 | (6,6°)                 | 20         | 276°                  | (3,8)°                 | (17)°                  |
| RR26    | -23,2293°  | 54,4698°  | 4259  | 828   | 144,6°                | (0,4°)                 | (9,8°)                 | 12         | 146°                  | (5,2)°                 | (18)°                  |
| RR27    | -21,9657°  | 54,2889°  | 4277  |       |                       |                        |                        |            |                       |                        |                        |
| RR28    | -22,7152°  | 53,1595°  | 4540  | 616   | 72,9°                 | (0,3°)                 | (8,2°)                 | 51         | 70°                   | (2,8)°                 | (20)°                  |
| RR29    | -24,9657°  | 51,7488°  | 4825  | 727   | 267,5°                | (0,4°)                 | (11,7°)                | 48         | 266°                  | (1,9)°                 | (13)°                  |
| RR30    | -26,4861°  | 49,8917°  | 5140  | 430   | 293,2°                | (0,1°)                 | (5,0°)                 | 12         | 292°                  | (3,8)°                 | (13)°                  |
| RR31    | -28,7648°  | 48,1394°  | 2710  |       |                       |                        |                        | 22         | 76°                   | (3,6)°                 | (17)°                  |
| RR32    | -30,2903°  | 49,5555°  | 4670  |       |                       |                        |                        |            |                       |                        |                        |
| RR33    | -31,1170°  | 50,6835°  | 4904  |       |                       |                        |                        |            |                       |                        |                        |
| RR34    | -32,0783°  | 52,2113°  | 4260  | 39    | 131,5°                | (0,4°)                 | (4,1°)                 | 39         | 134°                  | (3,4)°                 | (21)°                  |
| RR35    | -32,9694°  | 54,1473°  | 4214  |       |                       |                        |                        |            |                       |                        |                        |
| RR36    | -33,7018°  | 55,9578°  | 3560  | 47    | 224,9°                | (1,3°)                 | (7,4°)                 | 60         | 226°                  | (2,5)°                 | (19)°                  |
| RR37    | -31,7010°  | 57,8876°  | 4036  |       |                       |                        |                        |            |                       |                        |                        |
| RR38    | -30,5650°  | 59,6858°  | 4540  | 47    | 317,3°                | (1,9°)                 | (20,2°)                | 32         | 314°                  | (5,8)°                 | (33)°                  |
| RR39    | -29,0165°  | 60,9755°  | 4700  |       |                       |                        |                        |            |                       |                        |                        |
| RR40    | -28,1461°  | 63,3020°  | 4750  | 85    | 242,7°                | (15,3°)                | (27,9°)                | 44         | 228°                  | (2,3)°                 | (15)°                  |
| RR41    | -27,7330°  | 65,3344°  | 5430  | 60    | 90,1°                 | (4,7°)                 | (24,1°)                | 8          | 96°                   | (8,5)°                 | (24)°                  |
| RR42    | -27,6192°  | 65,4376°  | 4776  |       |                       |                        |                        |            |                       |                        |                        |
| RR43    | -27,5338°  | 65,5826°  | 4264  | 43    | 111,6°                | (2,4°)                 | (18,8°)                | 18         | 104°                  | (4,2)°                 | (18)°                  |
| RR44    | -27,5324°  | 65,7480°  | 4548  | 50    | 165,2°                | (1,8°)                 | (11,3°)                | 7          | 169°                  | (10,6)°                | (28)°                  |
| RR45    | -27,6581°  | 65,6019°  | 2822  |       |                       |                        |                        |            |                       |                        |                        |
| RR46    | -27,7909°  | 65,5835°  | 3640  | 53    | 140,3°                | (5,5°)                 | (11,9°)                | 19         | 139°                  | (4,8)°                 | (21)°                  |
| RR47    | -27,6958°  | 65,7553°  | 4582  | 58    | 123,2°                | (6,1°)                 | (19,2°)                | 10         | 124°                  | (4,7)°                 | (15)°                  |
| RR48    | -27,5792°  | 65,9430°  | 4830  | 54    | 44,1°                 | (5,4°)                 | (28,8°)                | 10         | 55°                   | (3,5)°                 | (11)°                  |
| RR49    | -26,2742°  | 68,5354°  | 4444  |       |                       |                        |                        |            |                       |                        |                        |
| RR50    | -25,5181°  | 70,0222°  | 4100  | 135   | 348,4°                | (0,7°)                 | (6,7°)                 | 22         | 348°                  | (3,4)°                 | (16)°                  |
| RR51    | -22,9989°  | 69,1911°  | 3463  |       |                       |                        |                        |            |                       |                        |                        |
| RR52    | -20,4722°  | 68,1094°  | 2880  | 212   | 24,8°                 | (1,6°)                 | (22,0°)                | 27         | 29°                   | (3,3)°                 | (17)°                  |
| RR53    | -20,1213°  | 64,9664°  | 2940  | 1548  | 103,5°                | (0,7°)                 | (20,9°)                | 28         | 99°                   | (3,4)°                 | (18)°                  |
| RR54    | -20,6424°  | 63,5082°  | 2499  |       |                       |                        |                        |            |                       |                        |                        |
| RR55    | -21,4417°  | 61,4959°  | 4462  | 886   | 251,1°                | (0,3°)                 | (5,8°)                 | 16         | 253°                  | (4,8)°                 | (19)°                  |
| RR56    | -21,9694°  | 59,5853°  | 4230  | 388   | 339,4°                | (0,3°)                 | (8,6°)                 | 13         | 338°                  | (3,6)°                 | (13)°                  |
| RR57    | -24,7264°  | 58,0496°  | 5200  |       |                       |                        |                        |            |                       |                        |                        |

difference between the two methods is  $3.8^\circ$  and in half of the cases is smaller than  $2.7^\circ$ . The mean accuracy using ship noise analysis is  $1.8^\circ$  against  $4.4^\circ$  using teleseisms analysis. For half of the stations, an accuracy of  $0.7^\circ$  is obtained using the ship noise analysis while a median accuracy of  $4.0^\circ$  is attained using teleseism analysis. It is worth noticing that the mean number of events per station recorded during the whole experiment is about 400 for ships against about 20 for teleseisms. This explains why ship noise analysis significantly outperforms teleseism analysis. This new technique clearly benefits from much more events of longer duration which allow increasing the accuracy by stacking information.

### Locating OBSs

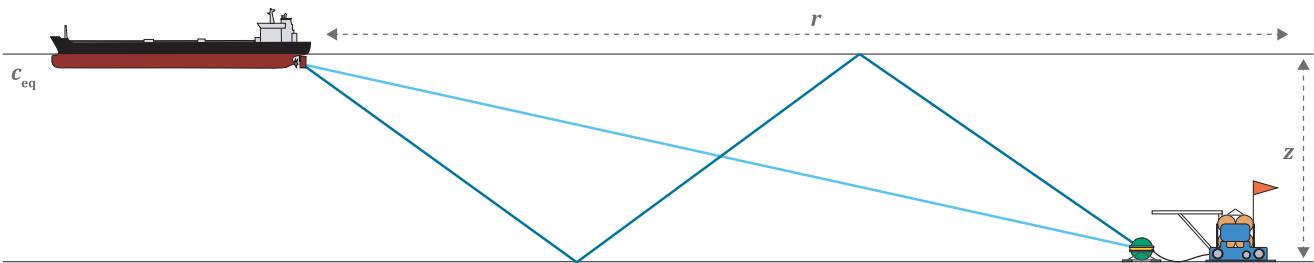
The following method proposes to localize OBSs by ranging acoustic sources in the ocean by analysing the time differences of arrival (TDOA) between direct and reflected phases. This only requires the use of the single hydrophone component of the OBS. Waves emitted by sources at the ocean surface can follow several paths to reach the ocean floor, mainly by bouncing back and forth between the ocean surface and floor (Fig. 7). For impulsive sources (e.g. airgun shots and whales), echoes generally follow the direct arrival and can be temporarily separated on the hydrophone recordings. In that case, the measurement of the vertical angle of incidence of the direct arrival can be used to range acoustic sources (Harris *et al.* 2013). However, this cannot be applied to continuous sources as different phases merge and interfere. In the case of a moving source (like a ship), spectrograms exhibit fringes that evolve with time, resulting in oblique striations (visible in Fig. 8b). This phenomenon is known as the Lloyd Mirror effect (e.g. Wilmut *et al.* 2007).

Two main frameworks exist to link those striations to the environment property and to the source's and receiver's locations: (i) a modal approach which models the spatial (hence temporal as the ship moves) variations of the interferences through the waveguide invariant theory (e.g. Gervaise *et al.* 2012); (ii) an approach which links the frequency variation of the interferences to the TDOA between echoes, for instance thanks to cepstrum analysis (e.g. Gao *et al.* 2008; Gao 2013). As the modal approach needs the relative distance from the source to the receiver as input to spatially resolve modes, we focused on the cepstrum analysis, that we detail below.

### Cepstrum analysis

In a single source scenario, a signal  $x(t)$  received by the hydrophone can be written as the convolution of a source term  $s(t)$  and a propagation term  $g(t)$  (i.e. the Green's function or impulse response):

$$x(t) = s(t) * g(t). \quad (13)$$



**Figure 7.** OBS relocation scheme: noise may propagate along several acoustic paths from the ship to the OBS. Here, the direct and the two times reflected paths are schematized. The measurement of the time difference of arrival (TDOA, see text) between these two phases can be linked to the OBS–ship distance (and therefore to the OBS location) by an appropriate modelling of the wave propagation from the ship to the OBS.

The Fourier spectral density (spectrum) of this signal is:

$$|X(f)|^2 = |S(f)|^2 |G(f)|^2. \quad (14)$$

The log spectrum of this signal is:

$$\log |X(f)|^2 = \log |S(f)|^2 + \log |G(f)|^2. \quad (15)$$

In this study, we modelled the propagation term by considering only two phases: (*d*) the direct path and (*r*) the first reverberated path that bounces once on the ocean floor and once on the ocean surface (Fig. 7). Each phase is associated with an attenuation coefficient and a time of arrival ( $\delta(t)$  is the Dirac function):

$$g(t) = \alpha_d \delta(t - t_d) + \alpha_r \delta(t - t_r) \quad (16)$$

$$G(f) = \alpha_d \exp(-j2\pi t_d f) + \alpha_r \exp(-j2\pi t_r f) \quad (17)$$

$$|G(f)|^2 = \alpha_d^2 + \alpha_r^2 + 2\alpha_d \alpha_r \cos(2\pi(t_r - t_d)f) \quad (18)$$

$$\log |G(f)|^2 \approx \log |\alpha_d^2 + \alpha_r^2| + \frac{2\alpha_d \alpha_r}{\alpha_d^2 + \alpha_r^2} \cos(2\pi \tau_{dr} f). \quad (19)$$

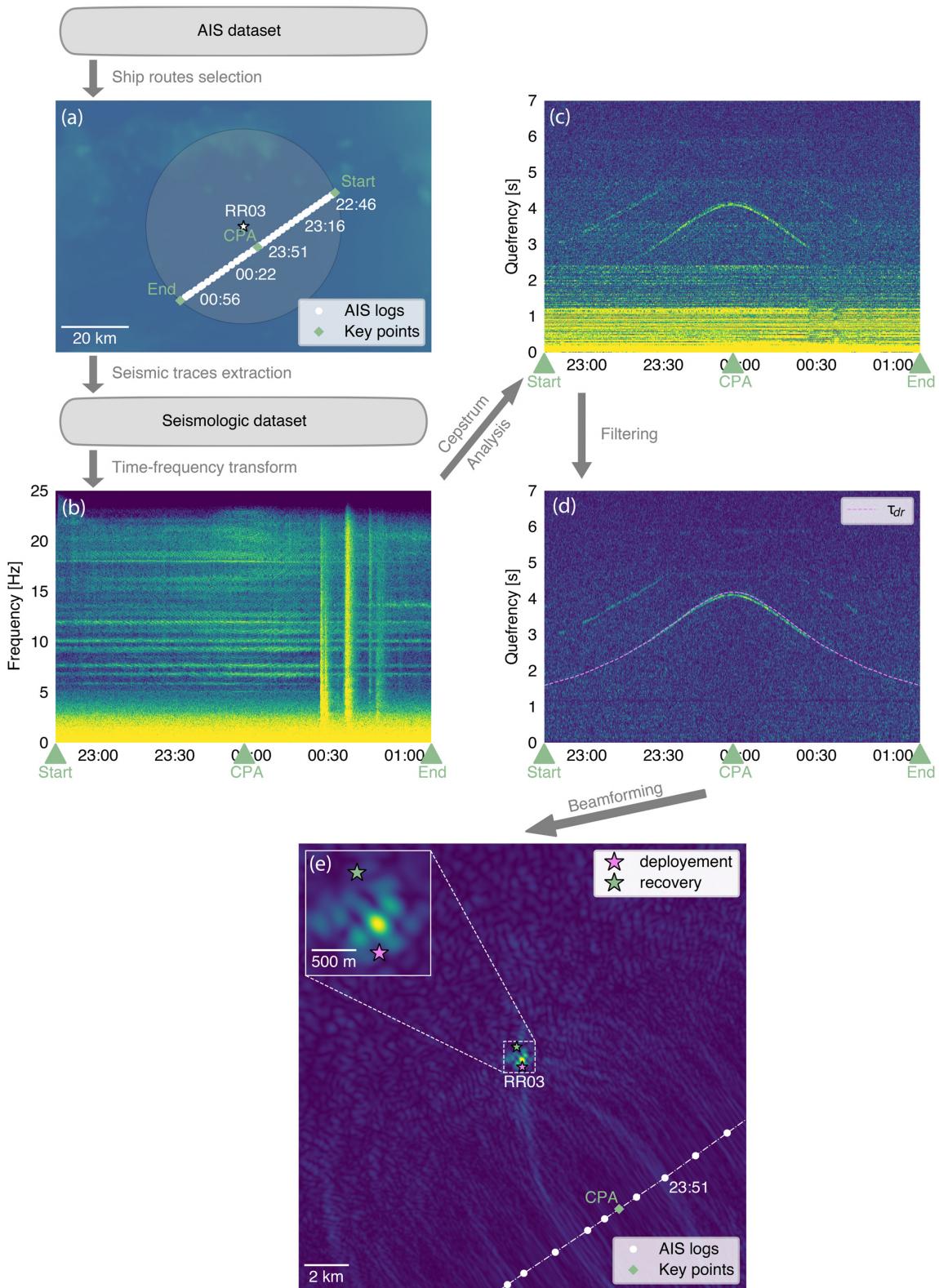
We note from eq. (14) and eq. (18) that the spectrum of the recorded signal is made of an envelope (the spectrum of the source term) that modulates an oscillating function (the spectrum of the propagation term). Combining eqs (15) and (19), we note that the log spectrum of the recorded signal has an additive component with 'fundamental frequency'  $\tau_{dr} = t_r - t_d$ . Therefore, the 'spectrum' of the log spectrum would likewise show a peak at the desired TDOA. It is from that observation that cepstrum analysis arose (Oppenheim & Schafer 2004). The cepstrum  $C_X(\tau)$  is defined as the inverse Fourier transform of the logarithmic spectrum. It is defined in a domain which, while being the reciprocal of frequency, is not really the time domain and which has been called the quefrency domain:

$$C_X(\tau) \equiv \mathcal{F}^{-1} [\log |F[x(t)]|^2]. \quad (20)$$

It can be noted that, due to the linearity of the Fourier transform, the cepstrum can also be decomposed into the sum of a source term and a propagation term:

$$C_X(\tau) = C_S(\tau) + C_G(\tau). \quad (21)$$

In this work, the cepstrum  $C_X(\tau)$ , which is real-valued was transformed into analytic signal. This analytical representation is useful tool for providing instantaneous values for parameters such as amplitude and phase of  $C_X(\tau)$ . This extra step allows to easily extract the envelope of the cepstrum (by taking its modulus, which is useful



**Figure 8.** Localization workflow for a single ship passage. (a and b) Same as in Fig. 4. Some seismic activity can be observed at the end of the spectrogram as vertical bands. (c) Raw cepstrrogram (high levels in yellow). (d) The time-invariant features are filtered out and denoising is applied. This isolates the features related to phases paths interferences. The theoretical TDOA between the direct and two times reflected path  $\tau_{dr}$  is overlaid (purple dashed line) assuming that the OBS did not drift from its deployment point. (e) Beamforming in the quefrency domain over a square of  $20 \times 20 \text{ km}^2$  centred on the deployment point. The ship path (white dotted line) can be considered as the antenna used for beamforming. The background colours are the beamforming scores of OBS localization (high values in yellow). In the left upper corner, a zoom of a  $1.6 \text{ km}^2$  around the OBS deployment (red star) and recovery (green star) localizations is plotted. Great focusing indicates the most probable OBS location on the ocean floor and validates the simple propagative model used.

for representation) and also, later on, to choose between coherent summation (complex values are summed) or incoherent summation (modulus are summed and the phase is discarded). The analytic cepstrum can be performed before the inverse Fourier transform  $\mathcal{F}^{-1}$  stage by discarding the negative frequency components and doubling the positive frequencies (sgn is the signum function):

$$\hat{C}_X(\tau) = \mathcal{F}^{-1} [(1 + \text{sgn}(f)) \log|\mathcal{F}[x(t)]|^2]. \quad (22)$$

Applying eq. (22) for each column (each spectrum) of a spectrogram gives a cepstrogram. In the case of a passing ship (Fig. 8c) the cepstrogram is composed of: (i) time-invariant features, that are identified as the source term and; (ii) a clear time-varying line which is identified as the propagation term. Unfortunately, the harmonic structure of ship radiated noise translates into cepstral features spanning over a large part of the quefrency domain and overlapping with the propagation term.

#### Propagation term extraction

Removing the source term allows to improve the TDOA retrieval (Ferguson *et al.* 2019). Separating time-invariant signals from moving ones is particularly suited for singular value decomposition (SVD) clutter filtering (Demene *et al.* 2015). The mean cepstrum is subtracted from the whole cepstrogram and a SVD is applied. In our case, a SVD just like principal component analysis (PCA) summarizes the cepstrogram in a sum of unitary time-invariant cepstrums (the singular vectors) associated to the quantity of signal they represent (the singular values). The way the energy is spread among the time-invariant singular vectors is related to the time-invariant aspect of the signal. The source term can ideally be represented by a unique singular vector of high singular value but, in practice, few singular vectors are necessary. Indeed, ship radiated noise slightly fluctuates, partially due to the Doppler effect and to small variations of the engine rate. In contrast, the propagation term needs to be represented by numerous time-invariant singular vectors of low singular values, due to its highly time-variant nature. Singular vectors were sorted by their singular value strength; the most energetic singular vectors, mainly containing the source term, were removed. The remaining singular vectors containing the propagation term and noise were kept to form the filtered cepstrogram as presented in Fig. 8(d). The number of singular values removed was automatically chosen using the kneedle algorithm (Satopaa *et al.* 2011).

To further improve TDOA estimation, an additional filtering was applied. More precisely, a high-pass filter was applied along quefrencies. This has a similar effect than filtering in the time domain (indeed, quefrencies are in seconds and Fourier transform in both time or quefrency domain leads to the frequency domain). The frequency content of the reverberated signal in the cepstrogram is directly linked to the frequency content of ship noise and cut-off frequencies are chosen accordingly. In practice, a simple high-pass filter with a cut-off frequency of 5 Hz was sufficient to remove the main part of the undesired signal. The result is a tremendously accurate measurement of the TDOA (Fig. 8d).

#### Propagative model

To retrieve useful information on the OBS location, the computed TDOA must be compared to a physical model of propagation. This latter was assessed by a direct ray tracing approach that easily allows computing theoretical phases arrival. In deep water, wavelengths of interest (roughly 150 m at 10 Hz) are usually at least one order

of magnitude smaller than other typical lengths. The ray tracing approach gives, in that case, accurate results. As a simplification, we assumed that the OBS is laying on a flat surface. This assumption is fulfilled in flat sedimentary basins, where the topography does not change over tens of kilometres or more (this is typically the case for the RR03 station). To further simplify the propagative model, the concept of equivalent celerity was used. To compute the traveltime  $t_{ray}$  from the ocean surface to the ocean floor at depth  $z$  and distance  $d$  in a range-independent model, computing the distance along a straight line raises the correct traveltime if divided by an equivalent celerity  $c_{eq}$  computed as:

$$\frac{1}{c_{eq}} = \frac{1}{z} \int_0^z \frac{dz}{c(z)}. \quad (23)$$

In this work, the equivalent celerity ( $c_{eq} = 1502 \text{ m/s}$ ) was computed from CTD (conductivity temperature depth) measurements in the area and the depth ( $z = 4340 \text{ m}$ ) was determined by multi-beam bathymetric measurements performed during the deployment at each OBS location (available through the RESIF data centre, see Acknowledgments). The traveltimes for the direct and first reverberated path could be assessed along with the TDOA:

$$\tau_{dr}(t) = \frac{\sqrt{d(t)^2 + (3z)^2} - \sqrt{d(t)^2 + z^2}}{c_{eq}}. \quad (24)$$

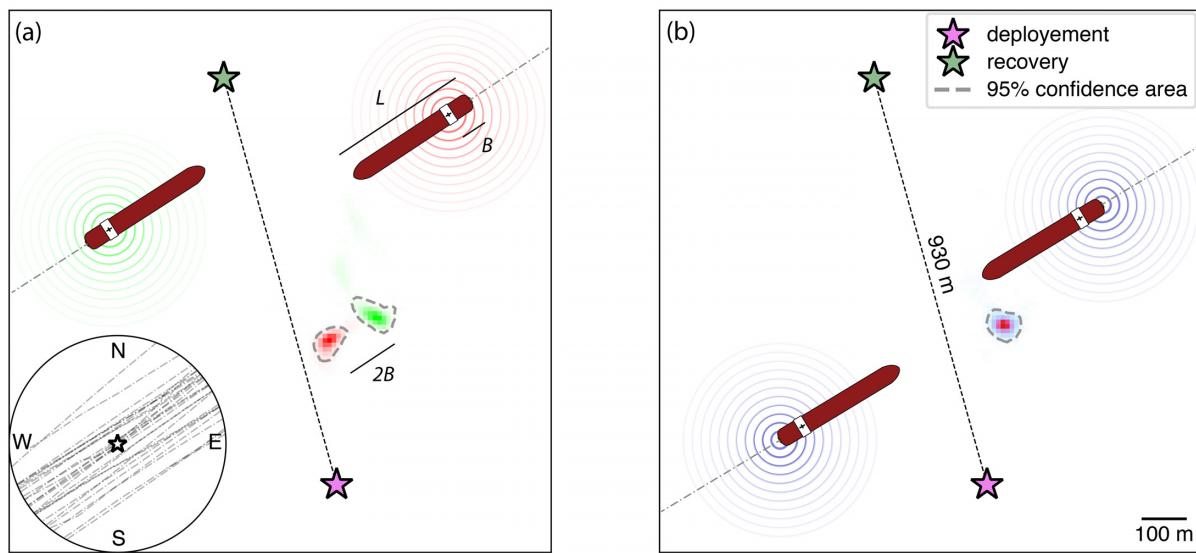
On Fig. 8(d),  $\tau_{dr}(t)$  is shown as a dashed purple line overlaid on the filtered cepstrogram assuming that the OBS did not drift and was located exactly beneath the deployment point. The theoretical curve almost matches the measurement, yet some mismatch can be corrected by a better OBS location hypothesis, what we are looking for.

#### Beamforming in the quefrency domain

To localize the OBS from isolated distance measurement, a nearfield beamforming approach was developed. Because of the wave propagation reciprocity, source and receiver can be virtually exchanged. The ship path can thereby be considered as a linear array of receivers. A 2-D grid search approach was performed and for each hypothetical OBS position, the signal of the filtered cepstrum was picked and summed coherently according to the time delay law defined by eq. (24). This summation is expected to be constructive at the actual OBS location. The envelope of this beamforming image for a single ship was computed and allows to create an image of the OBS location (Fig. 8e). The image exhibits a clear focal point located between the deployment and recovery points. The quality of the focusing testifies the validity of the used simple propagative model.

#### Localization accuracy

To determine an accurate OBS localization on the seafloor and to estimate the uncertainties, different measurements from different ship passages were compared. For our test station (RR03), 28 relevant ship passages extracted from terrestrial AIS logs of November 2012 were used (shown in map view in the insert Fig. 9a), each allowing to compute an image of the OBS location. All the images were then summed. To estimate the uncertainties, the process was repeated 10 000 times for random sets of 28 images sampled with replacement from the original 28 images (bootstrap approach, each set can contain a same image several times). A 2-D histogram of the OBS localization was computed and a 95 per cent confidence area



**Figure 9.** Probability maps of the OBS localization by bootstrapping 28 ship passages, according to two hypotheses on the location of the acoustical centre of emission location of ships. (a) the raw AIS position is used as the acoustical centre. Generally, the AIS antenna (black cross) is located close to the bridge (white rectangle). The length of the ship is noted  $L$  while  $B$  denotes the distance from the AIS antenna to the ship's stern. The inset represents the tracks used to retrieve the OBS location in a radius of 30 km. Ship routes are split in two: those crossing from east to west (red), and those crossing in the opposite direction (green). Each set of ships generate a clear individual patch indicating two different OBS locations. The two patches are aligned along the major ship route orientation and are at a distance comparable at two times the typical length  $B$ . (b) By correcting the AIS data, the stern of the ship is used as the acoustical centre which approximates the propeller location. A clear and unique patch (in pink) indicates the OBS most probable location between the deployment point and the recovery point. The 95 per cent confidence area (delimited by a grey dotted line) describes a zone of a few tens of metres wide.

was extracted by setting the threshold which allows rejecting the cells of the histogram which account for 5 per cent of the computed OBS localizations.

#### Acoustical centre location

Computing the 2-D histogram of OBS localization without further processing produced a bimodal distribution that results from a discrepancy between the AIS and the noise source actual locations. Splitting ship routes into two sets of eastward and westward travelling directions and computing the OBS localization for each set produced two distinct unimodal patches visible as shown in Fig. 9(a, in red and green). AIS archives of ships, which antenna position was provided, showed that AIS antennas are generally located close to the bridge at a stern distance  $B$  which can be approximated as 20 per cent of the overall ship distance  $L$ . Assuming that ship noise is mainly emitted by the propeller, hence at the ship stern, explains why the two patches are observed at a distance of  $\sim 2B$  aligned along the ship commercial route. Correcting the AIS data of 20 per cent  $L$  backward along each ship track before processing the OBS localization effectively merged the two patches into a single OBS localization (Fig. 9b) with a radius of confidence of a few tens of metres. This simple correction of the ship AIS location therefore demonstrates that the dominant ship noise is emitted by the propeller at the rear of the vessel. As seen in Fig. 9(b), the retrieved location is slightly closer to the deployment point, which agrees with the fact that OBS speed is higher during its descent than during its ascent.

Despite requiring good AIS coverage and a flat topography, the method showed promising capabilities to relocate OBS with a simple propagative model. It significantly improved the knowledge on the OBS location compared to the sole information available from the deployment and recovery points. It suggests that, in the 1–25 Hz frequency band, ship noise is mainly radiated by the propeller. It

also proves the ability of an OBS to range acoustical sources with high accuracy.

#### CONCLUSION

In this work, we demonstrated that ship noise analysis can be successfully used to orient and localize OBSs on the ocean floor. Two new fully automated workflows were developed for those purposes and were successfully applied to data recorded during the RHUM-RUM experiment in the SW Indian Ocean. Benefiting from the abundance of ships passing in this area and from the accurate knowledge of their trajectories from AIS data, this approach allowed to retrieve the OBS orientation with an accuracy better than  $1^\circ$  for most instruments. A result that performs far better than orientation retrieval from teleseismic event analysis. Passively retrieving the localization of an OBS within a few tens of metres was also successfully performed. It required more frequent ship positions than provided by the satellite AIS data and was therefore restricted to one station covered by high rate terrestrial AIS data. This study thus enables *a posteriori* orientation and localization retrieval of OBSs with no time cost during the cruise; the only cost being purchasing AIS archives of transiting ships. In addition of the three-component seismometer, the embedded hydrophone was extensively used, which is rather uncommon in seismology. It extended the classical polarization analysis into an acoustic intensity analysis, allowing for a more complete description of  $P$  waves. This study also confirms the potential of ships as acoustical sources for geophysical purposes. Just like earthquakes, ships can be used advantageously to set light on unknown parameters of interest. They benefit from great localization accuracy in time and space, are abundant in most seas, but require dedicated methods to handle continuous acoustic sources. The localization method suggests that the propeller is the main source of ship radiated noise within the [1–50] Hz frequency

range. The methods we developed also testify the ability of OBSs to fully localize ships and more generally all kind of moving acoustical sources in the oceans. These acoustical sources include marine mammals and sea ice activity, which bolsters OBSs potential as watchful environmental observers of the oceans.

## ACKNOWLEDGEMENTS

The RHUM-RUM project (<http://www.rhum-rum.net>) was funded by ANR (Agence Nationale de la Recherche) in France (project ANR-11-BS56-0013), and by DFG (Deutsche Forschungsgemeinschaft) in Germany (grants SI1538/2-1 and SI1538/4-1). Additional support was provided by CNRS (Centre National de la Recherche Scientifique, France), TAAF (Terres Australes et Antarctiques Françaises, France), IPEV (Institut Polaire Français Paul Emile Victor, France), AWI (Alfred Wegener Institut, Germany) and the Université de La Réunion, France. We thank DEPAS (Deutsche Geräte-Pool für Amphibische Seismologie, Germany), GEOMAR (GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Germany) and INSU-IPGP (Institut National des Sciences de l'Univers—Institut de Physique du Globe de Paris, France) for providing the ocean-bottom seismometers. The RHUM-RUM data set (<http://dx.doi.org/10.15778/RESIF.YV2011>) has been assigned the FDSN network code YV and is hosted and served by the French RESIF data centre (<http://seismology.resif.fr>). We warmly thank cruise participants and crew members on *R/V Marion Dufresne*, cruise MD192 and on *R/V Meteor*, cruise M101. This work benefited from the work of Alexandre Bouillon and Bruno Geay during their Ecole Navale internship at the Université de La Réunion in 2016. Codes used to explore OBS data sets and to retrieve orientations and localizations are available in a package named *obsea*, available at <https://github.com/atrabattoni/obsea> and planned to grow with future advances on acoustic source detection, localization and classification from OBS recordings. The author's PhD was funded by the STEP'UP doctoral school of *Université de Paris* and by the *Direction Générale de l'Armement*.

## REFERENCES

- Abrard, F. & Deville, Y., 2005. A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources, *Signal Process.*, **85**, 1389–1403, doi:10.1016/j.sigpro.2005.02.010.
- Anderson, P.N., Duennbier, F.K. & Cessaro, R.K., 1987. Ocean borehole horizontal seismic sensor orientation determined from explosive charges, *J. geophys. Res.*, **92**, 3573, doi:10.1029/jb092ib05p03573.
- Barruol, G. & Sigloch, K., 2013. Investigating La Réunion hot spot from crust to core, *EOS, Trans. Am. Geophys. Un.*, **94**, 205–207, doi:10.1002/2013eo230002.
- Barruol, G. *et al.*, 2019. Large-scale flow of Indian Ocean asthenosphere driven by Réunion plume, *Nat. Geosci.*, **12**, 1043–1049, doi:10.1038/s41561-019-0479-3.
- Cao, J., Liu, J., Wang, J. & Lai, X., 2017. Acoustic vector sensor: reviews and future perspectives, *IET Signal Process.*, **11**, 1–9, doi:10.1049/iet-spr.2016.0111.
- Creager, K.C. & Dorman, L.R.M., 1982. Location of instruments on the seafloor by joint adjustment of instrument and ship positions, *J. geophys. Res.*, **87**, 8379, doi:10.1029/jb087ib10p08379.
- D'Alessandro, A. & D'Anna, G., 2014. Retrieval of ocean bottom and downhole seismic sensors orientation using integrated MEMS gyroscope and direct rotation measurements, *Adv. Geosci.*, **40**, 11–17, doi:10.5194/adgeo-40-11-2014.
- D'Spain, G.L., Hodgkiss, W.S. & Edmonds, G.L., 1991. Energetics of the deep ocean's infrasonic sound field, *J. acoust. Soc. Am.*, **89**, 1134–1158, doi:10.1121/1.400677.
- Deen, M., Wielandt, E., Stutzmann, E., Crawford, W., Barruol, G. & Sigloch, K., 2017. First observation of the Earth's permanent free oscillations on ocean bottom seismometers, *Geophys. Res. Lett.*, **44**, 10 988–10 996, doi:10.1002/2017gl074892.
- Demene, C. *et al.*, 2015. Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases Doppler and ultrasound sensitivity, *IEEE T. Med. Imaging*, **34**, 2271–2285, 10.1109/tmi.2015.2428634.
- Doran, A.K. & Laske, G., 2017. Ocean-bottom seismometer instrument orientations via automated Rayleigh-wave arrival-angle measurements, *Bull. seism. Soc. Am.*, **107**, 691–708, doi:10.1785/0120160165.
- Dosso, S.E., Fallat, M.R., Sotirin, B.J. & Newton, J.L., 1998. Array element localization for horizontal arrays via Occam's inversion, *J. acoust. Soc. Am.*, **104**, 846–859, doi:10.1121/1.423359.
- Duennbier, F.K., Anderson, P.N. & Fryer, G.J., 1987. Azimuth determination of and from horizontal ocean bottom seismic sensors, *J. geophys. Res.*, **92**, 3567, doi:10.1029/jb092ib05p03567.
- Ebbeson, G.R., Thierion, L. & Heard, G.J., 2008. Array element localization of a bottom-mounted hydrophone array using ship noise, *Can. Acoust.*, **36**, 170–171.
- Ferguson, E.L., Ramakrishnan, R., Williams, S.B. & Jin, C.T., 2017. Convolutional neural networks for passive monitoring of a shallow water environment using a single sensor, in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5–9 March 2017, New Orleans, LA, USA, pp. 2657–2661, doi:10.1109/icassp.2017.7952638.
- Ferguson, E.L., Williams, S.B. & Jin, C.T., 2019. Improved multi-path time delay estimation using Cepstrum subtraction *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019)*, pp. 2657–2661, in Brighton, UK, May 2019, doi:10.1109/icassp.2019.8683176.
- Fontaine, F.R., Barruol, G., Kennett, B.L.N., Bokelmann, G.H.R. & Reynaud, D., 2009. Upper mantle anisotropy beneath Australia and Tahiti from P wave polarization: implications for real-time earthquake location, *J. geophys. Res.*, **114**, doi:10.1029/2008jb005709.
- Gao, Y., Clark, M. & Cooper, P., 2008. Time delay estimate using cepstrum analysis in a shallow littoral environment, in *Proceedings of the 3B.3 UDT Europe*, Glasgow UK, pp. 9.
- Gao, Y., 2013. Power cepstra measured in shallow water environments, in *Proceedings of the MTS/IEEE OCEANS – Bergen*, pp. 1–7, doi:10.1109/OCEANS-Bergen.2013.6608044.
- Gervaise, C., Kinda, B.G., Bonnel, J., Stéphan, Y. & Vallez, S., 2012. Passive geoacoustic inversion with a single hydrophone using broadband ship noise, *J. acoust. Soc. Am.*, **131**, 1999–2010, doi:10.1121/1.3672688.
- Gribonval, R. & Zibulevsky, M., 2010. Sparse component analysis, in *Handbook of Blind Source Separation*, eds Comon, P. & Jutten, C., pp. 367–420, Elsevier, doi:10.1016/B978-0-12-374726-6.00015-1.
- Grigoli, F., Cesca, S., Dahm, T. & Krieger, L., 2012. A complex linear least-squares method to derive relative and absolute orientations of seismic sensors, *Geophys. J. Int.*, **188**, 1243–1254, doi:10.1111/j.1365-246x.2011.05316.x.
- Hable, S., Sigloch, K., Barruol, G., Stähler, S.C. & Hadzioannou, C., 2018. Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations, *Geophys. J. Int.*, **214**, 2014–2034, doi:10.1093/gji/ggy236.
- Hable, S., Sigloch, K., Stutzmann, E., Kiselev, S. & Barruol, G., 2019. Tomography of crust and lithosphere in the western Indian Ocean from noise cross-correlations of land and ocean bottom seismometers, *Geophys. J. Int.*, **219**, 924–944, doi:10.1093/gji/ggz333.
- Harris, D., Matias, L., Thomas, L., Harwood, J. & Geissler, W.H., 2013. Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic, *J. acoust. Soc. Am.*, **134**, 3522–3535, doi:10.1121/1.4821207.
- Hildebrand, J.A., 2009. Anthropogenic and natural sources of ambient noise in the ocean, *Mar. Ecol. Prog. Ser.*, **395**, 5–20, doi:10.3354/meps08353.
- Janiszewski, H.A. & Abers, G.A., 2015. Imaging the Plate Interface in the Cascadia Seismogenic Zone: new constraints from offshore receiver functions, *Seismol. Res. Lett.*, **86**, 1261–1269, doi:10.1785/0220150104.

- Kinda, G.B., Courtois, F.L., Stéphan, Y., Boutonnier, J.-M., Royer, J.-Y. & Barruol, G., 2018. Underwater ambient noise spatial and temporal coherence at basin scale, *J. acoust. Soc. Am.*, **144**, 1732–1732, doi:10.1121/1.5067681.
- Larson, E.W.F., 2002. Determining surface wave arrival angle anomalies, *J. geophys. Res.*, **107**, ESE 7-1–ESE 7-15, doi:10.1029/2000jb000048.
- Laske, G., 1995. Global observation of off-great-circle propagation of Long-Period surface waves, *Geophys. J. Int.*, **123**, 245–259, doi:10.1111/j.1365-246x.1995.tb06673.x.
- Lee, D.D. & Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization, *Nature*, **401**, 788–791, doi:10.1038/44565.
- Lim, H., Kim, Y.H., Song, T.-R.A. & Shen, X., 2017. Measurement of seismometer orientation using the tangential P-wave receiver function based on harmonic decomposition, *Geophys. J. Int.*, **212**, 1747–1765, doi:10.1093/gji/ggx515.
- Mann, J.A., Tichy, J. & Romano, A.J., 1987. Instantaneous and time-averaged energy transfer in acoustic fields, *J. acoust. Soc. Am.*, **82**, 17–30, doi:10.1121/1.395562.
- Mazzullo, A., Stutzmann, E., Montagner, J.-P., Kiselev, S., Maurya, S., Barruol, G. & Sigloch, K., 2017. Anisotropic tomography around La Réunion Island from Rayleigh waves, *J. geophys. Res.*, **122**, 9132–9148, doi:10.1002/2017jb014354.
- Morley, M.G., Dosso, S.E. & Chapman, N.R., 2009. Array element localization using ship noise, *J. acoust. Soc. Am.*, **125**, 1403–1409, doi:10.1121/1.3076201.
- Nakamura, Y., Donoho, P.L., Roper, P.H. & McPherson, P.M., 1987. Large-offset seismic surveying using ocean-bottom seismographs and air guns: instrumentation and field technique, *Geophysics*, **52**, 1601–1611, doi:10.1190/1.1442277.
- Nehorai, A. & Paldi, E., 1994. Acoustic vector-sensor array processing, *IEEE T. Signal Proces.*, **42**, 2481–2491, doi:10.1109/78.317869.
- Oppenheim, A.V. & Schafer, R.W., 2004. DSP history - from frequency to quefrency: a history of the cepstrum, *IEEE Signal Proc. Mag.*, **21**, 95–106, doi:10.1109/msp.2004.1328092.
- Sabra, K.G., Roux, P., Thode, A.M., DSpain, G.L., Hodgkiss, W.S. & Kuperman, W.A., 2005. Using ocean ambient noise for array self-localization and self-synchronization, *IEEE J. Ocean. Eng.*, **30**, 338–347, doi:10.1109/joe.2005.850908.
- Satopaa, V., Albrecht, J., Irwin, D. & Raghavan, B., 2011. Finding a "kneeble" in a haystack: detecting knee points in system behavior, *2011 31st International Conference on Distributed Computing Systems Workshops*, 20–24 June 2011, doi:10.1109/icdcsw.2011.20.
- Schimmel, M. & Gallart, J., 2004. Degree of polarization filter for frequency-dependent signal enhancement through noise suppression, *Bull. seism. Soc. Am.*, **94**, 1016–1035, doi:10.1785/0120030178.
- Schimmel, M., Stutzmann, E., Arduhin, F. & Gallart, J., 2011. Polarized Earths ambient microseismic noise, *Geochem. Geophys. Geosyst.*, **12**, 1525–2027, doi:10.1029/2011gc003661.
- Scholz, J.-R., Barruol, G., Fontaine, F.R., Mazzullo, A., Montagner, J.-P., Stutzmann, E., Michon, L. & Sigloch, K., 2018. SKS splitting in the Western Indian Ocean from land and seafloor seismometers: plume plate and ridge signatures, *Earth planet. Sci. Lett.*, **498**, 169–184, doi:10.1016/j.epsl.2018.06.033.
- Scholz, J.-R., Barruol, G., Fontaine, F.R., Sigloch, K., Crawford, W.C. & Deen, M., 2017. Orienting ocean-bottom seismometers from P-wave and Rayleigh wave polarizations, *Geophys. J. Int.*, **208**, 1277–1289, doi:10.1093/gji/ggw426.
- Schulte-Pelkum, V., Masters, G. & Shearer, P.M., 2001. Upper mantle anisotropy from long-period polarization, *J. geophys. Res.*, **106**, 21 917–21 934, doi:10.1029/2001jb000346.
- Stachnik, J.C., Sheehan, A.F., Zietlow, D.W., Yang, Z., Collins, J. & Ferris, A., 2012. Determination of New Zealand ocean bottom seismometer orientation via Rayleigh-wave polarization, *Seismol. Res. Lett.*, **83**, 704–713, doi:10.1785/0220110128.
- Stähler, S.C. et al., 2016. Performance report of the RHUM-RUM ocean bottom seismometer network around La Réunion western Indian Ocean, *Adv. Geosci.*, **41**, 43–63, doi:10.5194/adgeo-41-43-2016.
- Theis, F.J., Jung, A., Puntonet, C.G. & Lang, E.W., 2003. Linear geometric ICA: fundamentals and algorithms, *Neural Comput.*, **15**, 419–439, doi:10.1162/089976603762552979.
- Thode, A.M., Sakai, T., Michalec, J., Rankin, S., Soldevilla, M.S., Martin, B. & Kim, K.H., 2019. Displaying bioacoustic directional information from sonobuoys using azigrams, *J. acoust. Soc. Am.*, **146**, 95–102, doi:10.1121/1.5114810.
- Wilmut, M.J., Chapman, N.R., Heard, G.J. & Ebbeson, G.R., 2007. Inversion of Lloyd mirror field for determining a source's track, *IEEE J. Ocean. Eng.*, **32**, 940–947, doi:10.1109/JOE.2007.904315.
- Zha, Y., Webb, S.C. & Menke, W., 2013. Determining the orientations of ocean bottom seismometers using ambient noise correlation, *Geophys. Res. Lett.*, **40**, 3585–3590, doi:10.1002/grl.50698.