

# Estimating the spatial distribution of vocalizing animals from ambient sound spectra using widely spaced recorder arrays and inverse modelling

Sebastian Menze, Daniel Zitterbart, Martin Biuw, and Olaf Boebel

Citation: *The Journal of the Acoustical Society of America* **146**, 4699 (2019); doi: 10.1121/1.5139406

View online: <https://doi.org/10.1121/1.5139406>

View Table of Contents: <https://asa.scitation.org/toc/jas/146/6>

Published by the [Acoustical Society of America](#)

---

## ARTICLES YOU MAY BE INTERESTED IN

### [Automated localization of whales in coastal fjords](#)

*The Journal of the Acoustical Society of America* **146**, 4672 (2019); <https://doi.org/10.1121/1.5138125>

### [Introduction to the Special Issue on Acoustic Source Localization](#)

*The Journal of the Acoustical Society of America* **146**, 4647 (2019); <https://doi.org/10.1121/1.5140997>

### [Acoustic localization of crows in pre-roost aggregations](#)

*The Journal of the Acoustical Society of America* **146**, 4664 (2019); <https://doi.org/10.1121/1.5138133>

### [Noise robust bird call localisation using the generalised cross-correlation with phase transform in the wavelet domain](#)

*The Journal of the Acoustical Society of America* **146**, 4650 (2019); <https://doi.org/10.1121/1.5138593>

### [Convolutional neural network for single-sensor acoustic localization of a transiting broadband source in very shallow water](#)

*The Journal of the Acoustical Society of America* **146**, 4687 (2019); <https://doi.org/10.1121/1.5138594>

### [Long-range frequency-difference source localization in the Philippine Sea](#)

*The Journal of the Acoustical Society of America* **146**, 4727 (2019); <https://doi.org/10.1121/1.5138124>

---

SUBMIT TODAY!

**JASA**  
THE JOURNAL OF THE  
ACOUSTICAL SOCIETY OF AMERICA

**Special Issue: Theory and  
Applications of Acoustofluidics**

# Estimating the spatial distribution of vocalizing animals from ambient sound spectra using widely spaced recorder arrays and inverse modelling

Sebastian Menze,<sup>1,a)</sup> Daniel Zitterbart,<sup>2,b)</sup> Martin Biuw,<sup>3</sup> and Olaf Boebel<sup>4</sup>

<sup>1</sup>*Oceanography and Climate Group, Institute of Marine Research, P. O. Box 1870, Nordnes, Bergen, Norway*

<sup>2</sup>*Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, Massachusetts 02543-1050, USA*

<sup>3</sup>*Marine Mammal Group, Institute of Marine Research, Tromsø, Norway*

<sup>4</sup>*Ocean Acoustics Group, Alfred-Wegener-Institute, Helmholtz-Center for Polar and Marine Research, Klussmannstrasse 3d, 27570 Bremerhaven, Germany*

(Received 15 March 2019; revised 18 November 2019; accepted 18 November 2019; published online 31 December 2019)

The sound energy from marine mammal populations vocalizing over extended periods of time adds up to quasi-continuous “choruses,” which create characteristic peaks in marine sound spectra. An approach to estimate animal distribution is presented, which uses chorus recordings from very sparse unsynchronized arrays in ocean areas that are too large or remote to survey with traditional methods. To solve this under-determined inverse problem, simulated annealing is used to estimate the distribution of vocalizing animals on a geodesic grid. This includes calculating a transmission loss (TL) matrix, which connects all grid nodes and recorders. Geometrical spreading and the ray trace model BELLHOP [Porter (1987). *J. Acoust. Soc. Am.* **82**(4), 1349–1359] were implemented. The robustness of the proposed method was tested with simulated marine mammal distributions in the Atlantic sector of the Southern Ocean using both drifting acoustic recorders [Argo (2018). SEANO] and a moored array as acoustic receivers. The results show that inversion accuracy mainly depends on the number and location of the recorders, and can be predicted using the entropy and range of the estimated source distributions. Tests with different TL models indicated that inversion accuracy is affected only slightly by inevitable inaccuracies in TL models. The presented method could also be applied to bird, crustacean, and insect choruses. © 2019 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/1.5139406>

[JAC]

Pages: 4699–4717

## I. INTRODUCTION

Passive acoustic monitoring (PAM) is increasingly used to study the distribution and migration of vocalizing animals that are otherwise difficult to observe, such as marine mammals (Rogers *et al.*, 2013), birds (Dawson and Efford, 2009), fish (Wall *et al.*, 2013), insects, and amphibians (Pijanowski *et al.*, 2011). Most methods estimate population density and/or spatial distribution based on the detection of transient vocalizations (Marques *et al.*, 2013) recorded by single hydrophones or small-scale arrays. Here, we present an approach to estimate the distribution of vocalizing animals that utilizes ambient sound spectra from widely spaced recorder arrays (>100 km distance) and the cumulative sound energy emitted by a population, rather than signals from individual vocalizations. We developed this method to interpret recordings of low-frequency and far-ranging marine mammal vocalizations in the Southern Ocean, but it could also be applied to other situations involving a large number of signal sources, such as bird, crustacean, and insect choruses, which create a

quasi-continuous chorus that is observed with a sparse array of receivers.

In the ocean, ambient sound (also often termed “ambient noise” or “soundscape”) stems from sea surface motion, precipitation, sea ice motion, glacier calving, shipping, seismic surveys, marine mammals, fish, and crustaceans (Carey and Evans, 2011; McDonald *et al.*, 2008; Nieu Kirk *et al.*, 2012). The cumulative sound energy of a marine mammal population vocalizing during extended periods adds up to a “chorus-like” quasi-continuous signal, which can dominate ambient sound over certain frequency bands (Curtis *et al.*, 1999; Leroy *et al.*, 2018b; Seger *et al.*, 2016). Throughout the remainder of this paper, these parts of the ambient sound are referred to as marine mammal choruses (MMCs), though strictly speaking, they also contain energy from single, discernible calls. Hence, MMCs more accurately represent the acoustic power contributed in specific frequency bands by the target species. A recording containing Antarctic minke whale calls and the Antarctic minke and blue whale MMCs is shown in Fig. 1.

The contribution of the various sources to ambient sound can be determined by analyzing characteristic peaks and slopes in ambient sound spectra. The temporal variability of these spectra can be visualized with long-term spectral

<sup>a)</sup>Electronic mail: [sebastian.menze@imr.no](mailto:sebastian.menze@imr.no)

<sup>b)</sup>Also at: Biophysics Group, Department of Physics, University of Erlangen-Nürnberg, Staudtstrasse 7/B2, 91058 Erlangen, Germany.

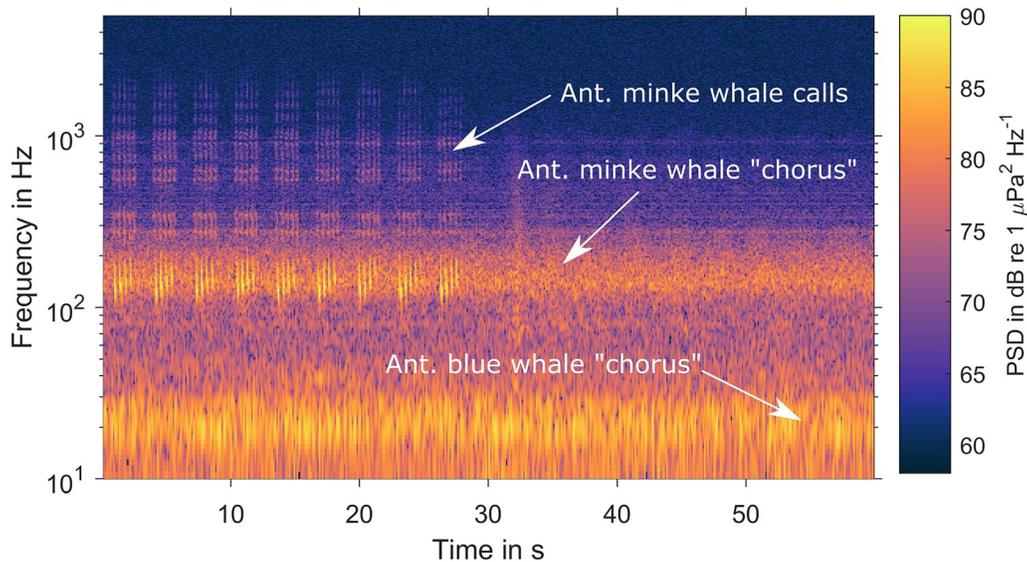


FIG. 1. (Color online) Spectrogram showing Antarctic minke whale calls and “chorus” (MMC) and Antarctic blue whale chorus (MMC). Recorded by a moored Aural M2 recorder (Multi-Electronique, Quebec, CA) on 5 September 2008 at 66°01.13’S and 0°04.77’E (sample rate, 32768 Hz; spectrogram settings: fast Fourier transform (FFT), 8000 points; Hanning window, 50% overlap). The recordings are described in [Menze et al. \(2017\)](#).

averages (LTSA), which display the average power spectral density (PSD) of each recording over the recorder’s deployment period. An example LTSA from the Southern Ocean is displayed in Fig. 2(a) ([Menze et al., 2017](#)). The contribution of the air–sea–ice interaction to ambient sound can be seen as vertical lines, and the contribution of Antarctic blue whales (*Balaenoptera musculus intermedia*), fin whales (*Balaenoptera physalus*), Antarctic minke whales (*Balaenoptera bonaerensis*), and leopard seals (*Hydrurga leptonyx*) can be seen as horizontal lines in the LTSA ([Menze et al., 2017](#)). The spectral peaks in Southern Ocean ambient sound related to Antarctic blue whales, fin whales, and Antarctic minke whales are displayed in Figs. 2(b)–2(d). The MMC sound energy can be calculated by subtracting fitted functions from the measured spectra, resulting in time series of MMC received levels ( $RL_{MMC}$ ). Figure 2(e) compares Antarctic minke whale  $RL_{MMC}$  recorded at 66°S and 69°S; the time series show distinct north-south differences and co-varying patterns. MMCs have also been observed from fin whales in the Mid and North Atlantic ([Nieukirk et al., 2012](#)), fin and blue whales in the North Pacific ([Burtenshaw et al., 2004](#); [Curtis et al., 1999](#)) and Indian Ocean ([Leroy et al., 2018a](#)), Pygmy blue and Antarctic blue whales around Australia ([McCauley et al., 2018](#)), and fin and possibly Bowhead whales in the Arctic ([Ahonen et al., 2017](#)), and exhibit extensive spatial as well as inter- and intra-annual variation. In this study, we explore how the information in such MMC patterns can be used to estimate the spatial distribution of a population of vocalizing animals.

Most approaches to estimate animal distribution or density from acoustic recordings focus on the detection of transient vocalizations, which can also be used to localize individual animals. The spacing, geometry, and clock accuracy of a recorder array, as well as the nature of the sound source, sound speed profile, and bathymetry, determine if and how accurately individual sound sources can be localized. If only a single hydrophone is present, it is often only possible to detect the number of calls per unit time (often

termed call rate or acoustic activity) and  $RL_{MMC}$  at the hydrophone’s location ([Haver et al., 2017](#); [Van Opzeeland et al., 2013](#); [Van Parijs et al., 2009](#)). In shallow water with a dispersive waveguide and impulsive calls, range estimation is possible on a single hydrophone ([Bonnel et al., 2014](#); [Marques et al., 2011](#)). In cases where the vocalizations propagate in a way that allows the identification of multipath arrival patterns or modes, it is also possible to estimate the call source level (SL), the distance from the recorder and source depth, in addition to the number of calls per unit time ([Mouy et al., 2012](#); [Newhall et al., 2012](#); [Valtierra et al., 2013](#)). When arrays with small to medium spacing are used, it is possible to calculate the distance, bearing, and SL of transient sounds via time-difference-of-arrival (TDOA) or beamforming methods ([Harris et al., 2018](#); [Širović et al., 2007](#); [Urazghildiiev and Clark, 2013](#); [Urazghildiiev and Hannay, 2018](#); [Wang et al., 2016](#)). However, when the array spacing becomes so large that a signal is no longer recorded by at least three hydrophones, or individual calls cannot be associated, tracking individual sound sources becomes challenging, and analysis is often limited to comparing the number of calls per unit time and  $RL_{MMC}$  at the different locations ([Risch et al., 2014](#); [Thomisch et al., 2016](#)).

It is important to note the difference between density and distribution. In this study, we define density as the average number of vocalizing animals per  $km^2$  within the entire study area, and we define spatial distribution as the number of animals per grid cell for a grid that tessellates the study area. The two most promising methods for estimating animal density from the detection of vocalizations are distance sampling and spatially explicit capture recapture methods ([Harris et al., 2013](#); [Harris et al., 2018](#); [Kusel et al., 2011](#); [Kyhn et al., 2012](#); [Marques et al., 2013](#); [Martin et al., 2013](#); [Thomas and Marques, 2012](#); [Ward et al., 2012](#)). However, due to their reliance on individual call detections, they work best on spatial scales smaller than ocean basins, and require an extensive recorder array ([Carlén et al., 2018](#); [Harris et al.,](#)

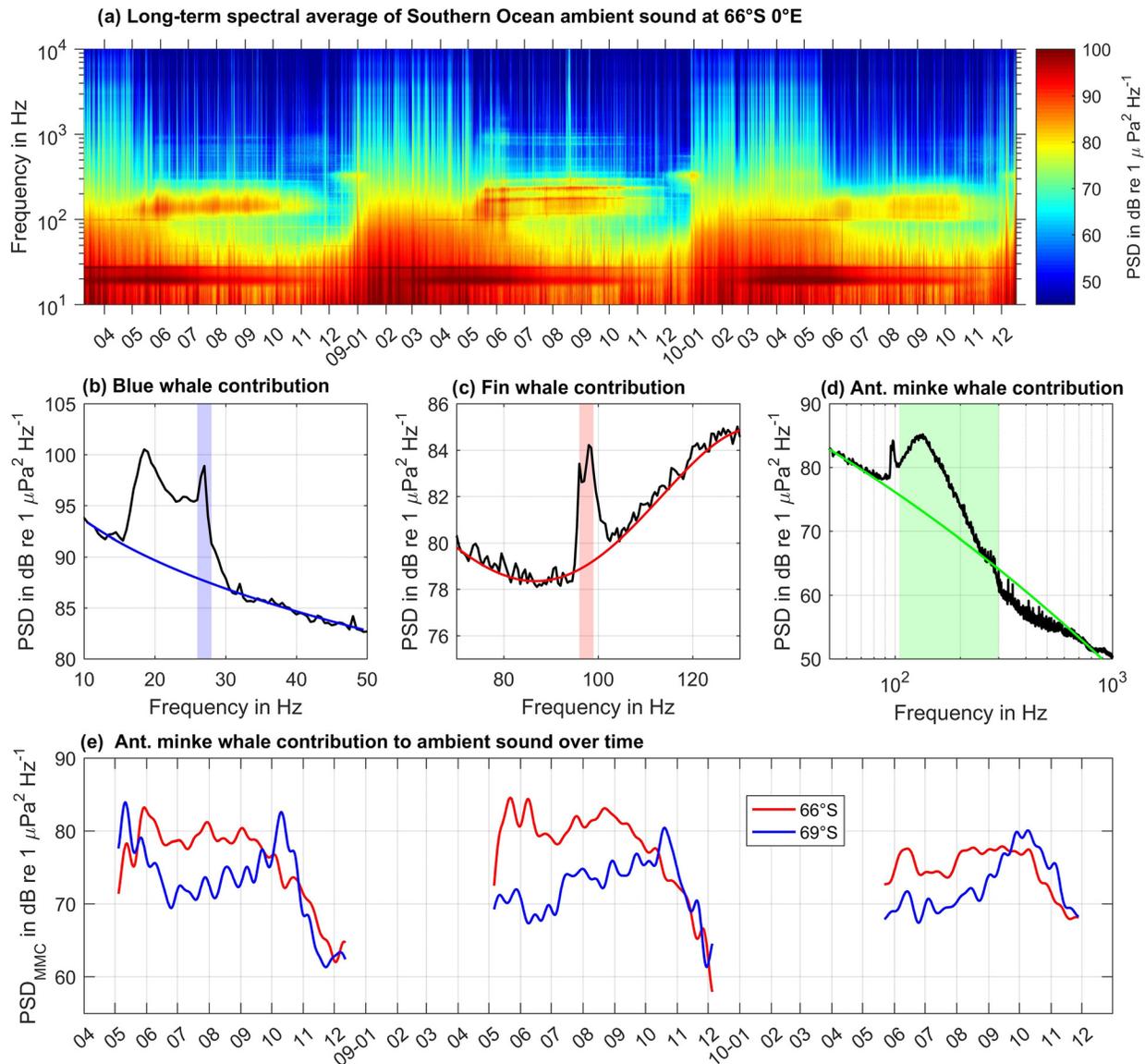


FIG. 2. (Color online) Ambient sound recordings from the Southern Ocean containing marine mammal contributions as described in Menze *et al.* (2017). (a) LTSA of recordings from 66°S, marine mammal contributions visible as horizontal streaks. (b),(c),(d) Marine mammal contribution peaks in example ambient sound spectrum. The black lines represent the measured PSD (at 66°S on 25 May 2008 12.00 h), the colored lines are three fitted interpolation functions and the colored areas are the frequency bands used to calculate the  $PSD_{MMC}$ . (e) Comparison of the Antarctic minke whale chorus ( $PSD_{MMC}$ ) time series (low-pass filtered with a seven-day window Butterworth filter) between 66°S (red) and 69°S (blue).

2018). In this paper, we estimate the spatial distribution of acoustic sources instead of the density of acoustic sources in the study area.

Due to the complex and cumulative nature of the MMC to ambient sound,  $RL_{MMC}$  data have been rarely used to estimate animal distribution. Seger *et al.* (2016) combined MMC recordings and line transect surveys to investigate the spacing among singing Humpback whales. Mellinger *et al.* (2014) discussed an approach to estimate the density of vocalizing fin whales in a reference area around a single hydrophone using acoustic propagation modelling. The difficulty with interpreting the spatial and temporal patterns in  $RL_{MMC}$  is that a higher  $RL_{MMC}$  does not necessarily imply a higher density of animals due to the nonlinearity of underwater sound propagation and the large and unknown number and location of sources involved. For a given location, increased  $RL_{MMC}$  can be caused by a combination of

processes: an increase in the number of vocalizing animals, an increase in SL, an increase in call rate, a decreasing distance to the vocalizing animals, or a decreasing TL between the vocalizing animals and the recorder. We address these issues by using a set of  $RL_{MMC}$  recordings in combination with acoustic propagation models and a parameter estimation algorithm to estimate the distribution of sound sources, which would generate the observed set of  $RL_{MMC}$  recordings. With additional information about the animals' SLs and call rates, it should then be possible to extend the presented approach further and provide an estimate of the number of animals per grid cell. As with any PAM method, we only estimate the distribution of vocalizing animals, while non-vocalizing animals present in the area cannot be detected.

This paper is structured into six sections. Section I is the Introduction and Sec. II describes the inversion method.

Section III describes simulated scenarios to test the robustness of the inversion method and how we quantified inversion accuracy. Section IV presents the results of the simulated test scenarios and relations between inversion accuracy and several metrics. Section V discusses these results and the feasibility of the inversion method. Conclusions are summarized in Sec. VI.

## II. THE INVERSION METHOD

Estimating the spatial distribution (location and amplitude) of sound sources from a finite set of  $RL_{MMC}$  observations is an under-determined, non-linear inverse problem. Similar  $RL_{MMC}$  values could be caused by different source numbers, locations, and amplitudes, and the number of unknown parameters (location and amplitude) is much larger than the number of observations. Following the notation of the Bayesian geophysical inverse problem theory (Mosegaard and Sambridge, 2002; Tarantola, 2005), the  $RL_{MMC}$  observations form the data set  $\mathbf{d}$ , which is connected to the parameter set  $\mathbf{m}$  through the forward model  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ . Here, the forward model  $\mathbf{g}(\mathbf{m})$  simulates the ambient sound created by a set of acoustic sources for which spatial distribution is described by the parameter set  $\mathbf{m}$ . To solve the inverse problem, we are sampling the joint posterior distribution that combines flat prior distributions over the parameters  $\mathbf{m}$  and the least squares misfit between  $\mathbf{d}$  and  $\mathbf{g}(\mathbf{m})$ .

For inverse problems with a small number of parameters, the misfit function can be sampled using a grid search, i.e., calculating the misfit of all possible parameter combinations (also termed the search space). In our case, this is impossible since the number of parameters is in the hundreds to thousands, rendering the search space too large for a grid search. We therefore developed a parameter estimation algorithm that uses a Markov chain Monte Carlo (MCMC) algorithm to sample the misfit function and find the parameters with least misfit between observed and modelled RLs. This is a first exploration of the inverse theory approach toward estimating marine mammal distribution from chorus recordings. Sections II A–II C describe the different parts of the inversion method: the architecture of the forward model, prior estimates, assumptions, and the parameter estimation algorithm.

### A. The forward model and *a priori* assumptions

Estimating  $RL_{MMC}$  requires knowledge about the number or sources (vocalizing marine mammals), their SL, location, and the TL between the source and recorder locations. The TL is not only influenced by the distance between the source and receiver, but also by the sound speed field, sea floor shape and properties, sea surface roughness, sea ice, and bubble clouds. Since it is computationally very costly to include all these parameters in a forward model, we make several assumptions to expedite the calculations.

Our first simplification of the forward model is neglecting time; since we are modelling the contribution of marine mammals to ambient sound, which is quasi-continuous on the scales of minutes to hours, we can simulate the transient vocalizations by a set of continuous sources of identical

frequency. The continuous nature of ambient sound, and the marine mammal contribution to it, arises due to the many sources involved, the multipath propagation that spreads impulsive signals over time, and the repetitive and monotonous nature of many marine mammal vocalizations. Multipath propagation of underwater sound renders initially impulsive signals (such as the Antarctic minke whale calls in Fig. 1) into a quasi-continuous signal (such as the Antarctic minke whale “chorus” in Fig. 1) due to sea floor, internal, and surface reflections. We simulated this process for fin whale vocalizations and found that the pulse train can become a quasi-continuous signal at distances around 100 km away from the source (supplemental Figs. 1 and 2).<sup>1</sup> Since we assume a steady-state situation in our forward model, we observe the time scale so that our model is valid. We aim to estimate source distribution on a basin scale (thousands of km), where the signal travel times between source and recorder are on the scale of minutes to tens of minutes (an underwater sound signal needs approximately 11 min to travel 1000 km). Thus we assume that the SL, call rate, and location of the vocalizing marine mammals and TL are approximately constant on the time scale of 10–30 min. This implies that  $RL_{MMC}$  should be measured on the scale of minutes, ideally between 10 and 30 min, and the time steps between estimates of distribution need to be on the scale of hours. It is unlikely that the large-scale marine mammal distribution and TL change significantly on smaller time scales.

The second assumption is to neglect source depth in the forward model. This is deemed appropriate since the source depth mainly affects TL in the first tens of km (Weirathmueller *et al.*, 2013). Tagging of vocalizing blue whales indicated that calling occurs mainly at depths below 30 m (Lewis *et al.*, 2018).

The third assumption is to discretize and reduce the search space that is sampled by the parameter estimation algorithm. Since we neglect depth and time, the parameter set  $\mathbf{m}$  only needs to describe the source locations and levels. Allowing arbitrary locations, SLs, and number of sources would require an overwhelming computational effort. We reduce the possible source locations to grid nodes. This grid is termed the simulated source grid. Using a rectangular latitude-longitude grid will result in an uneven distribution of nodes across ocean basin scales. Therefore, we calculated node positions with a geodesic algorithm that approximates the shape of a sphere using an icosahedron (Teanby, 2006). It is available as MATLAB code (MathWorks, Natick, MA), and was implemented into the forward model. The estimated received level,  $\widehat{RL}_i$ , at each recorder,  $i$ , is calculated as the (incoherent) sum of the acoustic power from all source grid nodes,

$$g(\mathbf{m}) = \widehat{RL}_i = 20 \log_{10} \left( \sum_{j=1}^{n_{\text{nodes}}} 10^{(SL_j - TL_{ij})/20} \right),$$

where  $j$  is the source grid node index,  $n_{\text{nodes}}$  is the number of grid nodes,  $SL_j = 20 \log_{10}(SP_j)$  is the SL at each node, and  $TL_{ij}$  is the TL between a recorder  $i$  and the source at grid node  $j$ . For efficient computation, the TL between all grid nodes and recorders is calculated into a lookup TL matrix

using a sound propagation model. The two acoustic propagation models implemented for this study are presented in Sec. II B. The parameter estimation algorithm then needs to determine the source pressure at each node  $SP_j$  that produces the least misfit between model and observations. For this it is necessary to reduce the degrees of freedom of the inverse problem to allow the parameter estimation algorithm to find the best  $SP_j$  quickly. Instead of performing a grid search for the best  $SP_j$  (calculating the misfit of all possible parameter combinations), a fixed number of equally loud simulated sources is moved across the grid nodes. The number of simulated sources is set the same as the number of grid nodes  $n_{\text{sources}} = n_{\text{nodes}}$ . This allows all source location combinations ranging from one simulated source at each node to all simulated sources being at one node. The parameter set  $\mathbf{m}$  is then defined as a vector containing the node index that describes where each simulated source is located. The source pressure at a given node is then defined as the sum of all simulated sources (animals) assigned to that node. The sound pressure of each source (animal) is defined as a fraction of the unknown cumulative source pressure (CSP, the total sound energy emitted by all vocalizing animals of the population) and  $n_{\text{sources}}$ . The source pressure  $SP_j$  at a given node  $j$  is thus calculated as the product of the number of simulated sources located at that node and the fraction of the CSP

$$SP_j = \left( \sum_{j \in \mathbf{m}} \right) \frac{\text{CSP}}{n_{\text{nodes}}}.$$

The true value of the CSP is unknown, thus, it needs to be estimated on the basis of typical SLs, population sizes, and call rates as given in the literature. We assume that this could take any value (uniform distribution) between the extreme cases of  $\text{CSP}_{\text{min}}$  (only one animal vocalizing sporadically) and  $\text{CSP}_{\text{max}}$  (all possible existing animals vocalizing constantly).

We then solve the inverse problem (searching the minimum of the misfit function) for a predefined number  $n_{\text{SA chains}}$  of CSP values between  $\text{CSP}_{\text{min}}$  and  $\text{CSP}_{\text{max}}$  independently. The  $SP_j$  estimate is then calculated as the median of the three best (smallest misfit) SA chains to smooth out potential artifacts of a single solution. For small sample sizes (small  $n_{\text{SA chains}}$ ), taking the posterior median is a robust estimator of the parameters (Cronin *et al.*, 2009). The result of the inversion is the estimated source pressure grid, a map that shows where and how much sound pressure is emitted to create the recorded  $\text{RL}_{\text{MMC}}$ . We did not attempt to calculate animal densities from the estimated source pressure grid, but in cases where reliable estimates of animal call rate and SL are available, it should be possible to formulate multipliers that convert source pressure per area to number of animals per area. Conversely, for regions and species where population size is known with reasonable certainty, the migration of the entire vocalizing population could possibly be tracked. Figure 3 shows a flowchart of the inversion method, divided into knowns, prior, and posterior (after inversion) estimates. The inversion method was developed and tested with MATLAB2016a (MathWorks, Natick, MA) and Python2.7 (Python, Fredericksburg, VA).

## B. Sound propagation models

The TL between the recorders, source grid nodes, and test scenario sources was calculated using two methods, geometrical spreading (Lurton, 2010) and raytracing, using the BELLHOP (Porter, 1987) model, although any other underwater sound propagation model may be used as well. Geometrical TL was calculated using a critical radius of 4000 m, where a transition from spherical to cylindrical propagation is assumed, as this value roughly represents the average ocean depth of the study area. For distances shorter than the critical radius, TL was calculated using spherical spreading and absorption only,

$$\text{TL}(r) = 20 \log(r) + \alpha r,$$

where  $r$  is the distance from the source, and  $\alpha$  is the absorption coefficient from the empirical equations of Francois and Garrison (1982). For distances larger than the critical radius, the equation

$$\text{TL}(r) = 20 \log(r_{\text{critical}}) + 10 \log\left(\frac{r}{r_{\text{critical}}}\right) + \alpha r$$

was employed, with  $r_{\text{critical}}$  being the critical radius. The distances between the source and receiver pairs were calculated using great circle lines to account for the curvature of the Earth.

Raytracing TL was calculated using the two-dimensional (2D) range dependent sound propagation model BELLHOP. Instead of calculating the TL between all source and receiver pairs, we simulated the three-dimensional (3D) sound field using a  $2 \times N$ -dimensional ( $2 \times N$ -D) approach, rotating a set of 2D slices (range and depth) in  $5^\circ$  steps,  $360^\circ$  degree around each source location. The bathymetry for each slice was obtained from the ETOPO-1 topography dataset (Amante and Eakins, 2009). The sound speed over range and depth for each slice was interpolated from the world ocean atlas mean annual climatology dataset (Dushaw *et al.*, 2013). The sea floor was assumed to be an elasto-acoustic half-space with a pressure wave sound speed of  $1800 \text{ m s}^{-1}$  and a density of  $2.0 \text{ g cm}^{-3}$ . Each acoustic source (i.e., whale) was assumed to be at 10 m depth, and all recorders were assumed to be at 100 m depth. Raytracing TL was only calculated for 150 Hz, and sea ice was not accounted for. The implications of these constraints will be discussed in Sec. V. We interpolated a latitude-longitude grid containing the TL at 100 m depth from the 72 range and depth slices for each source. The TL between each source and recorder was then retrieved from this grid. Example slices and interpolated TL values are shown in supplemental Fig. 3.<sup>1</sup>

The two TL models are compared to each other using source and recorder locations in the Weddell Sea (maps of the locations can be found in Secs. IV A–IV D) in Fig. 4. They show a robust correlation, but for close ranges and TL values less than 100 dB the geometrical spreading model overestimates TL in relation to the raytracing model, while it underestimates TL at far ranges and TL values higher than 100 dB. This is also illustrated in supplemental Fig. 4.<sup>1</sup>

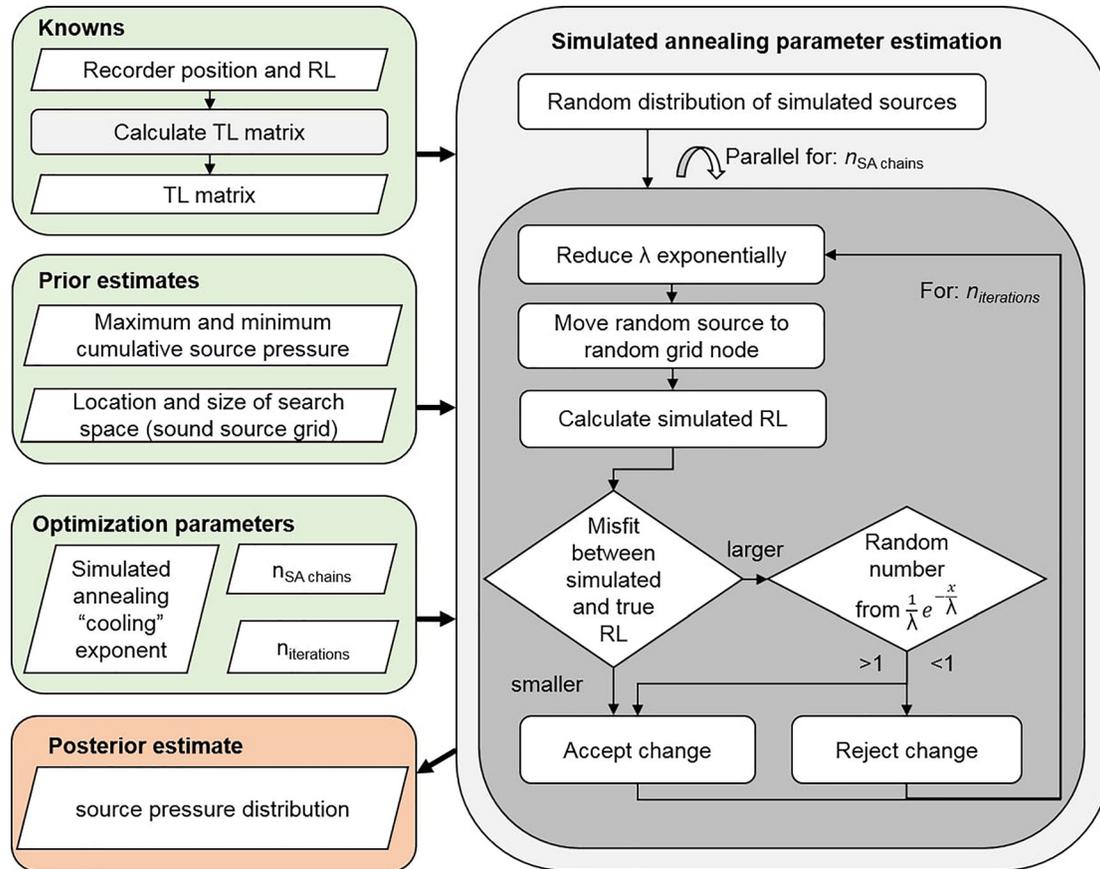


FIG. 3. (Color online) Flowchart of the inversion method and simulated annealing (SA) parameter estimation algorithm. Input variables are marked green, the output variables are marked red. Grey round boxes represent the computational parts of the method.  $\lambda$  is the SA “temperature” parameter that steers how much of an increase in misfit is tolerated at each iteration, and  $x$  is a random variable from the distribution  $f(\lambda, x) = (1/\lambda)e^{-x/\lambda}$ .

which compares the two models over a 1300 km section. It shows that geometrical spreading overestimates the TL in the first 500 km compared to BELLHOP. However, both the geometrical spreading and raytracing models provide a very similar logarithmic TL dependency. The performance and

shortcomings of the two models are evaluated in Sec. V (Discussion).

### C. Parameter estimation

The parameter set  $\mathbf{m}$  is defined as a vector containing the node indices that describe where each simulated source is located. Depending on the size and resolution of the grid, there are hundreds to thousands of grid nodes (=parameters). We sample the misfit function with a MCMC algorithm to find the global minimum within the search space. The movement of the algorithm through the search space is defined in the following manner: initially the simulated sources are distributed randomly (uniform distribution) over the grid nodes. Then, for each iteration, a simulated source is chosen randomly and moved to a random new grid node. Whether a move is accepted or rejected is governed by an acceptance rule. After the decision has been made, a new random move is generated. In this fashion, the algorithm moves through the search space for a fixed number of iterations.

Compared to the large number of parameters, the number of  $RL_{MCMC}$  observations is very small (on the order of tens to hundreds). This implies that the inverse problem is highly under-determined, and the misfit function has many local minima. The local minima and the size of the search space render it challenging for the minimization algorithm to reach the global minimum. An algorithm that only follows

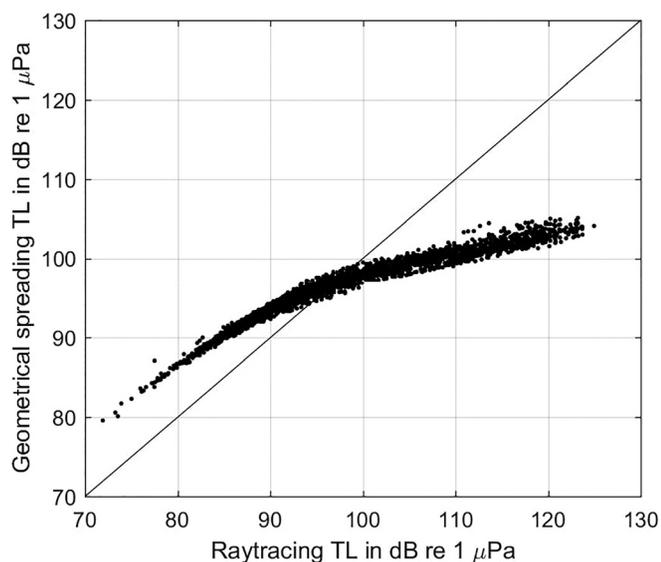


FIG. 4. Comparison of geometrical and raytracing TL models using source and receiver pairs in the Weddell Sea at 150 Hz. Detailed information on the receiver and source locations is given in Sec. III.

decreases in misfit may get trapped at a local minimum, while an algorithm that equally follows decreases and increases in misfit may not converge (get lost in the search space). Therefore, a suitable acceptance rule is essential for finding the global minimum.

We choose the simulated annealing (SA) acceptance rule (Kirkpatrick *et al.*, 1983). The SA algorithm always accepts decreases and increases in misfit with an exponential probability, which is reduced as the number of iterations increases. The probability to accept an increase in misfit is determined by the exponential probability density distribution  $f(x; \lambda)$ . An increase in misfit is accepted when a random number  $x$ , drawn from  $f(x; \lambda)$  is larger than one. For each iteration, a new random number  $x$  is drawn from  $f(x; \lambda)$ ,

$$f(x; \lambda) = \frac{1}{\lambda} e^{-x/\lambda},$$

where  $\lambda$  (termed the SA “temperature”) is the mean of  $f(x; \lambda)$ , and the random variable  $x$  can range from 0 to  $\infty$ . With each iteration,  $\lambda$  is reduced following an exponential function:

$$\lambda = \left(1 - \frac{i_{\text{iteration}}}{n_{\text{iterations}}}\right)^\varepsilon,$$

where  $\varepsilon$  is the SA “cooling” exponent ( $\varepsilon > 0$ ),  $i_{\text{iteration}}$  is the number of the current iteration, and  $n_{\text{iterations}}$  is the total number of iterations. The cooling exponent  $\varepsilon$  determines how fast  $\lambda$  decreases with increasing numbers of iterations, i.e., it controls the speed of the transition from randomly accepting increases in misfit to always rejecting increases in misfit. We found that a cooling exponent between two and six works well, and use  $\varepsilon = 2$  for all inversions in this study. A flowchart of the SA parameter estimations algorithm is displayed in Fig. 3. Given a sufficient number of iterations, the SA algorithm will converge toward the global minimum of the misfit function (Granville *et al.*, 1994).

To illustrate the parameter estimation process, Fig. 5 shows how  $\lambda$  is reduced over the iterations and how the misfit of the different SA chains is reduced over time. Each black line represents a solution (SA chain) moving through the search space. Each solution has a different CSP, which is the reason for the misfit offset between the different solutions already at the start of the iterations. Solutions with very large or very small CSP show large misfits between simulated and true RL over all iterations, whereas solutions with a fitting CSP converge toward lower misfit values after a few thousand iterations.

Figure 6 shows snapshots of the source pressure grid at different iteration stages for the same example scenario with 20 000 iterations and 13 recorders with an average distance of 300 km between the recorders. Initially, the simulated sources are distributed randomly (upper left panel). With increasing iterations the simulated sources are moved across the grid nodes, rendering the simulated source pressure grid increasingly similar to the true source pressure grid (lower right panel). The final estimate resolves the source distribution pattern well considering the small number of recorders

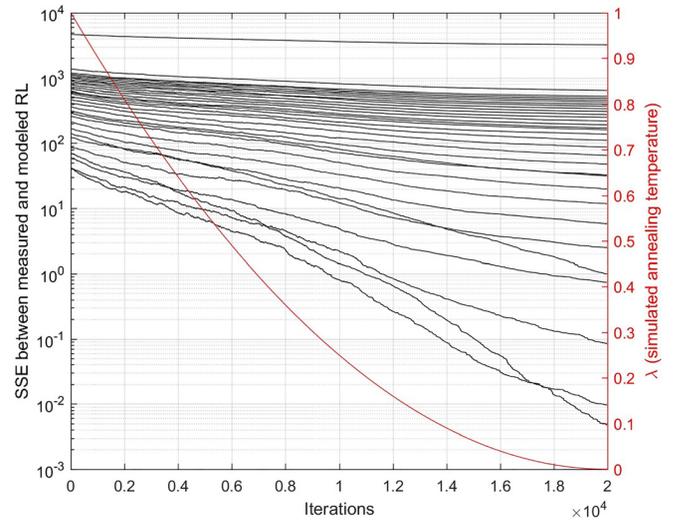


FIG. 5. (Color online) Misfit (SSE = sum of squared errors) between measured and modelled RLs over the iterations of the parameter estimation algorithm (black lines). The different black lines represent different SA chains ( $n_{\text{SA chains}} = 32$ , CSP between  $10^{11}$  and  $10^{13}$   $\mu\text{Pa}$ ). The red line represents  $\lambda$ , the SA temperature [mean of  $P(i, \lambda)$ ] over the iterations (SA cooling exponent  $\varepsilon = 2$ ).

used. The gradient of incorrect sources in the upper left corner of the estimated source distribution represents excess sound energy in the forward model that is moved toward the boundaries of the search space to reduce the misfit between received and modelled RL, and will be discussed in Sec. V.

### III. TEST SCENARIOS

The reliability and sensitivity of the inversion method was investigated using a set of test scenarios. All scenarios were created and analyzed using MATLAB2016a (MathWorks, Natick, MA) on a standard laptop, whereas the SA parameter estimation algorithm was executed on a high-performance computing cluster using 32 central processing units (CPUs) per scenario, computing each SA chain in parallel. The inversion and test scenario codes are available in the supplemental materials and a github repository.<sup>1</sup> The test scenarios were positioned in the Atlantic sector of the Southern Ocean between  $45^\circ\text{S}$  and  $80^\circ\text{S}$  and  $65^\circ\text{W}$  and  $25^\circ\text{E}$ . In all but the last test scenario, the simulated recorder array was a widely spaced mooring array identical to the HAFOS array of the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI; Van Opzeeland *et al.*, 2014; triangles in Fig. 6). The average array spacing, i.e., the distance between neighboring recorders, was 300 km. We plan to apply the inversion method to recordings from this array once they become available in the coming years. Detailed information on the inversion parameters and TL models used for each scenario are given in Table I.

We quantified the accuracy of each test scenario inversion using a metric similar to the simple matching coefficient (SMC; Sepkoski, 1974), which divides the sum of true positives and true negatives (number of matches) by the total set size. A SMC of zero means no overlap between two sets, and a SMC of one means a perfect match. We compared the true and estimated source pressure grids using two metrics: normalized accuracy ( $A_n$ ) and binary accuracy ( $A_b$ ). The

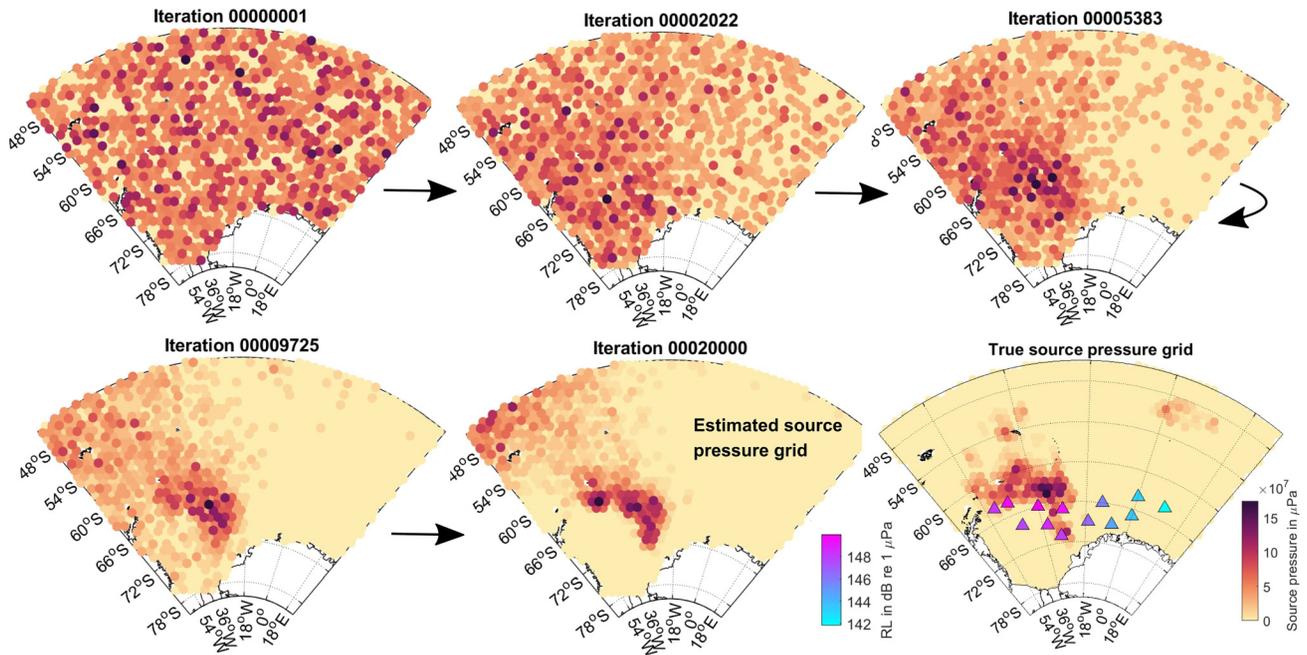


FIG. 6. (Color online) Schematic view of source pressure grid change during SA over 20 000 iterations. Colored dots represent the source pressure at each node in  $\mu\text{Pa}$  (yellow to red hues, normalized for each panel). Lower right plot shows the true source pressure grid and recorder locations where the RLs in dB re 1  $\mu\text{Pa}$  are marked by blue and pink triangles. The color scales are only valid for the panels that show the estimated and true source pressure grid.

binary accuracy compares only presence-absence information, comparing two binary sets (truth and estimate) that are zero where the source pressure is zero, and one where the source pressure is greater than zero. The accuracy is then

$$\text{Accuracy} = \frac{\text{true positives} + \text{true negatives}}{\text{Number of nodes}} = \frac{\text{Number of matching nodes}}{\text{Number of nodes}}.$$

To compare not only presence/absence but also scalar patterns (ratio scale data), the normalized accuracy compares the true and estimated source pressures after normalizing the source pressure at each node into 50 different bin values between 0 and maximum true source pressure. Identical to the binary accuracy, the normalized accuracy is then defined as the number of matching nodes divided by the number of nodes but with 50 instead of 2 classes.

### A. Random source distributions

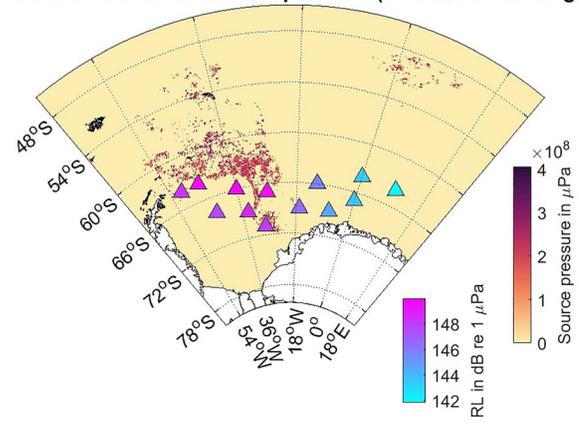
The first test scenario's objective was to investigate the reliability and feasibility of the inversion method and find a metric that correlates with inversion accuracy and can be used when the method is applied to real data. Therefore, we applied the inversion method to 250 random source distributions, an array of 13 recorders with a spacing of approximately 300 km between adjacent recorders and a source pressure grid with 1328 nodes and 111 km distance between adjacent nodes. To simulate the patchy nature of marine mammal distributions, we created a latitude-longitude grid with a resolution of 0.1 arclength (11 km) and randomly assigned SLs to the grid bins of this fine scale source grid.

This was realized in a three-step process: first, random noise with an  $f^{-5}$  spectrum was created (normalized between zero and one) and bins (output of the random number generator) with values below 0.75 set to zero, and bins with values above 0.75 were randomly assigned a value between 0 and 1 with an  $f^{-1}$  noise spectrum. The distribution was then thinned by setting bins back to zero where random noise (normalized between zero and one) with an  $f^{-2}$  spectrum was below 0.6. The resulting random distributions (an example distribution is shown in Fig. 7) show combinations of patchy and filamentous patterns not unlike the modelled habitat suitability distributions for Antarctic minke whales (Bombosch *et al.*, 2014; Herr *et al.*, 2019). We chose the spectral slope of the random distributions manually, yet other exponents or ways of simulating random source distribution to test the inversion method could be used equally well. To simulate source pressure distributions somewhat realistically, the normalized grid was then multiplied with a call rate of 0.5 (animals vocalizing 50% of the time) and source pressures of  $10^9 \mu\text{Pa}$  (180 dB re 1  $\mu\text{Pa}$ ; Širović *et al.*, 2007). For each scenario, the respective "true" source pressure  $\text{SP}_j$  at each source pressure grid node (best possible inversion results) was calculated from the fine scale source grid by smoothing the fine scale grid with a 2D circular averaging filter, the radius of which is the average distance between the nodes (111 km), and then extracting the pressure value at each node's location from the smoothed grid. The node locations and true  $\text{SP}_j$  values of the source pressure grid are shown in Fig. 7(b), which also shows the location and RL of the recorder array. RLs at the recorder array were calculated using geometric spreading TL (the true TL), which was also implemented as the TL model for the inversion.

TABLE I. Overview of test scenarios.

Objective	Recorders	Number and type of replicate source distributions	“True” TL	TL model	Frequency (Hz)	$n_{\text{iterations}}$	$n_{\text{SA chains}}$	CSP ( $\mu\text{Pa}$ )	CSP <sub>min</sub> and CSP <sub>max</sub> ( $\mu\text{Pa}$ )
Effect of different source distributions	Mooring array	250 random distributions	Geometrical spreading	Geometrical spreading	270	20 000	32	$2^{12} \pm 1^{12}$	$1^{11}$ and $1^{13}$
Effect of flawed TL model	Mooring array	100 random distributions on small grid	Geometrical spreading and BELLHOP raytracing	Geometrical spreading	150	15 000	30	$5^{11} \pm 3^{11}$	$1^{11}$ and $1^{13}$
Effect of number of SA chains	Mooring array	1 random distribution	Geometrical spreading	Geometrical spreading	150	15 000	5, 10, 15, 20, 25, 30, 35, 40	$2.3^{12}$	$1^{11}$ and $1^{13}$
Effect of number of iterations	Mooring array	1 random distribution	Geometrical spreading	Geometrical spreading	150	1 000, 5 000, 10 000, 15 000, 20 000, 25 000	32	$2.3^{12}$	$1^{11}$ and $1^{13}$
Effect of frequency	Mooring array	1 random distribution	Geometrical spreading	Geometrical spreading	27, 98, 150, 270	15 000	32	$2.3^{12}$	$1^{11}$ and $1^{13}$
Effect of recorder number and location/feasibility of drifting recorders	Argo floats	1 random distribution	Geometrical spreading	Geometrical spreading	150	20 000	32	$1.6^{12}$	$1^{11}$ and $1^{13}$

(a) True source distribution and pressure (Fine scale source grid)



(b) True source pressure grid

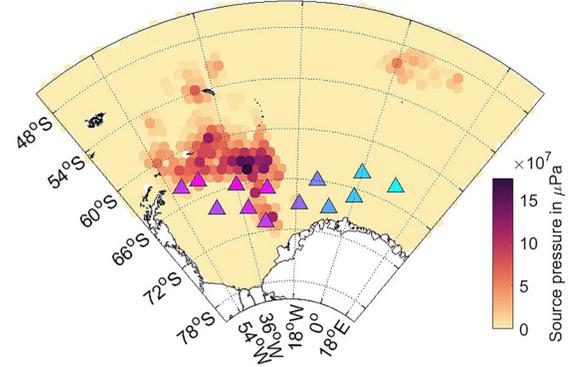


FIG. 7. (Color online) Example test scenario in the Weddell Sea. (a) shows the fine scale source grid where color indicates the amplitude of the acoustic sources (virtual whales) used to calculate the recorded RL (cyan-magenta triangles mark recorder location and respective RL). (b) shows the corresponding true amplitude of the source pressure grid nodes (best possible inversion result).

## B. Inaccurate TL model

The second test scenario’s objective was to investigate the effect of an incorrect and uncertain TL model on inversion success. We created 100 random source distributions using the same recorder array and random fine scale source grid generation as for the previous scenario (Sec. III A), but limited the source distribution to grid nodes between 62.5°S and 72°S and  $-49^\circ\text{E}$  and  $14^\circ\text{E}$  to reduce the computational effort of the raytracing modelling. For each of the 100 distributions, 2 inversions were calculated. The first inversion was calculated with perfect TL knowledge, where both the true and forward model TLs were calculated using geometrical spreading. The second inversion was calculated with a flawed forward model TL, where the true TL was calculated using raytracing (as described in Sec. II A), but the forward model TL was calculated using geometrical spreading.

## C. Robustness of inversion

We tested the robustness of the inversion method toward the number of SA chains and iterations, and the effect acoustic frequency has on inversion accuracy. We used geometrical spreading as the true and forward model TL and the same recorder array and source grid as for the previous scenarios. For a random distribution created by the method

described in Sec. III A, we ran 8 inversions using between 5 and 40 SA chains covering a CSP range between  $10^{11}$  and  $10^{13}$   $\mu\text{Pa}$  (with the true CSP being  $2.3 \times 10^{12}$   $\mu\text{Pa}$ ). For the same distribution and CSP values, we ran 6 inversions with 32 SA chains and between 1000 and 25 000 iterations. The effect of acoustic frequency on inversion success was tested using the same distribution and CSP values and the acoustic frequencies 27, 98, 150, and 270 Hz, since they are the characteristic contributions of marine mammals to the Southern Ocean acoustic environment (Menze *et al.*, 2017).

#### D. Simulation of drifting recorders using Argo float tracks

The last scenario tested the feasibility of using drifting platforms, such as Argo floats (Argo, 2018), as a receiver array. We extracted the tracks of all Argo floats within the study area (between  $45^\circ\text{S}$  and  $80^\circ\text{S}$  and  $-65^\circ\text{E}$  and  $25^\circ\text{E}$ ) between the 1.1.2013 and 29.5.2013 from the Coriolis Global Data Assembly Center.<sup>2</sup> The Argo tracks are displayed in supplemental Fig. 5.<sup>1</sup> We created a random distribution using the method described for the previous scenarios. For each day between 1 January 2013 and 29 May 2013, we used the positions of the available Argo float profiles as recorder locations and ran an inversion using 20 000 iterations and 32 SA chains.

### IV. RESULTS

#### A. Random source distributions and inversion accuracy

We estimated the area in which the inversion method produced reliable results by correlating the true and estimated source pressure at each node over the 250 random source distributions. The resulting map of correlation coefficients is displayed in Fig. 8. Correlation coefficients are high ( $>0.5$ ) within an oval area centered around the recorder location. This area roughly corresponds to the area we termed the “trust zone,” which we defined as the area where more than one recorder is present within a 1000 km radius. The trust zone could be defined equally well using other definitions, but we choose our approach as a first conservative approximation of the area in which we expect the recorder

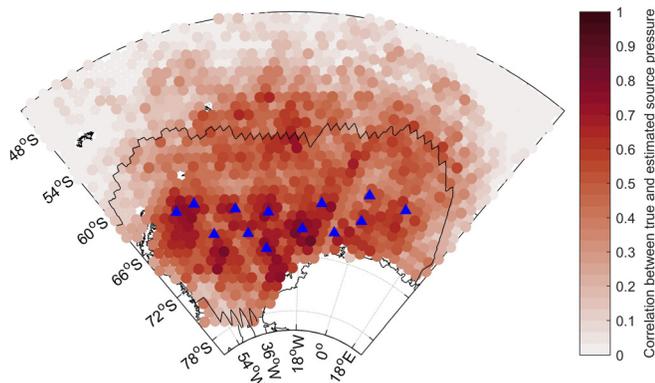


FIG. 8. (Color online) Map showing the correlation (red hues) between true and estimated source pressure over the 250 random distributions. Blue triangles mark recorder location. The black contour encircles the trust zone (more than one recorder present within a 1000 km radius).

setup and inversion algorithm to produce reliable results. This choice is discussed in more detail in Sec. V. The heterogeneous patterns in correlation are likely artifacts caused by the small number of test scenarios.

We calculated the normalized and binary accuracy of the 250 inversions. The inversions proved remarkably successful given the small number (13) of recorders in the array, and accuracy values ranged between 0.2 and 1 with a median  $A_n$  of 0.7 and median  $A_b$  of 0.8 for nodes within the trust zone. Simulations confirmed that the  $A_n$  expected by chance is 0.3, and the  $A_b$  expected by chance is 0.5. Both the binary and normalized accuracies show a decreased inversion success when  $\text{SP}_j$  is calculated from only the best SA chain (solution) instead of the median of the three best SA chains. This is shown in Fig. 9(a), which compares the cumulative density function (CDF) of the normalized and binary accuracies for the 250 random source distributions. As indicated by the correlation map in Fig. 8, inversions were most successful within the trust zone. The CDF of accuracy within the trust zone and entire grid are compared in Fig. 9(b), confirming that the inversion was more accurate within the trust zone than across the entire grid. Hereinafter, all  $A_n$  and  $A_b$  values in the paper are calculated using only nodes within the trust zone if not stated otherwise.

The true and estimated CSPs within the trust zone are compared in Fig. 10. They agree well with a correlation coefficient of 0.9. To show example source distributions, Fig. 11 compares example true and estimated source distributions from the test scenario, sorted from best to worst normalized accuracy. The inversion method managed to estimate the presence and absence of sources well in most cases, even when no source was present in the trust zone or sources were distributed across multiple clusters. In some of the estimated source pressure grids, a gradient of sound sources is present at the boundary of the search space in the general direction of the true source distribution. As will be shown later, this represents excess sound energy in the forward model that is moved toward the boundaries of the search space to reduce the misfit between received and modelled RLs.

To investigate why some of the random source distributions could be estimated successfully while others could not, we compared the effect of several metrics on inversion accuracy and found that information entropy is one of the most useful metrics to predict inversion accuracy. Information entropy (Shannon, 1948) is a measure of information content (Borda, 2011), which reaches its maximum when the elements of the set are uniformly distributed. Using only nodes within the trust zone, the entropy  $H(P(\text{SP}_j))$  of the estimated and true  $\text{SP}_j$  was calculated from the sample distribution  $P(\text{SP}_j)$  of the  $\text{SP}_j$  values in the following manner:

$$H(P(\text{SP}_j)) = - \sum P(\text{SP}_j) \log(P(\text{SP}_j)).$$

Source pressure distributions with high entropy contain a large variety of different  $\text{SP}_j$  values, whereas distributions with low entropy contained many similar  $\text{SP}_j$  values, mostly a high number of empty nodes with  $\text{SP}_j = 0$ .

Figure 12(a) shows how inversion accuracy varies with the RL range of each source distribution and the misfit

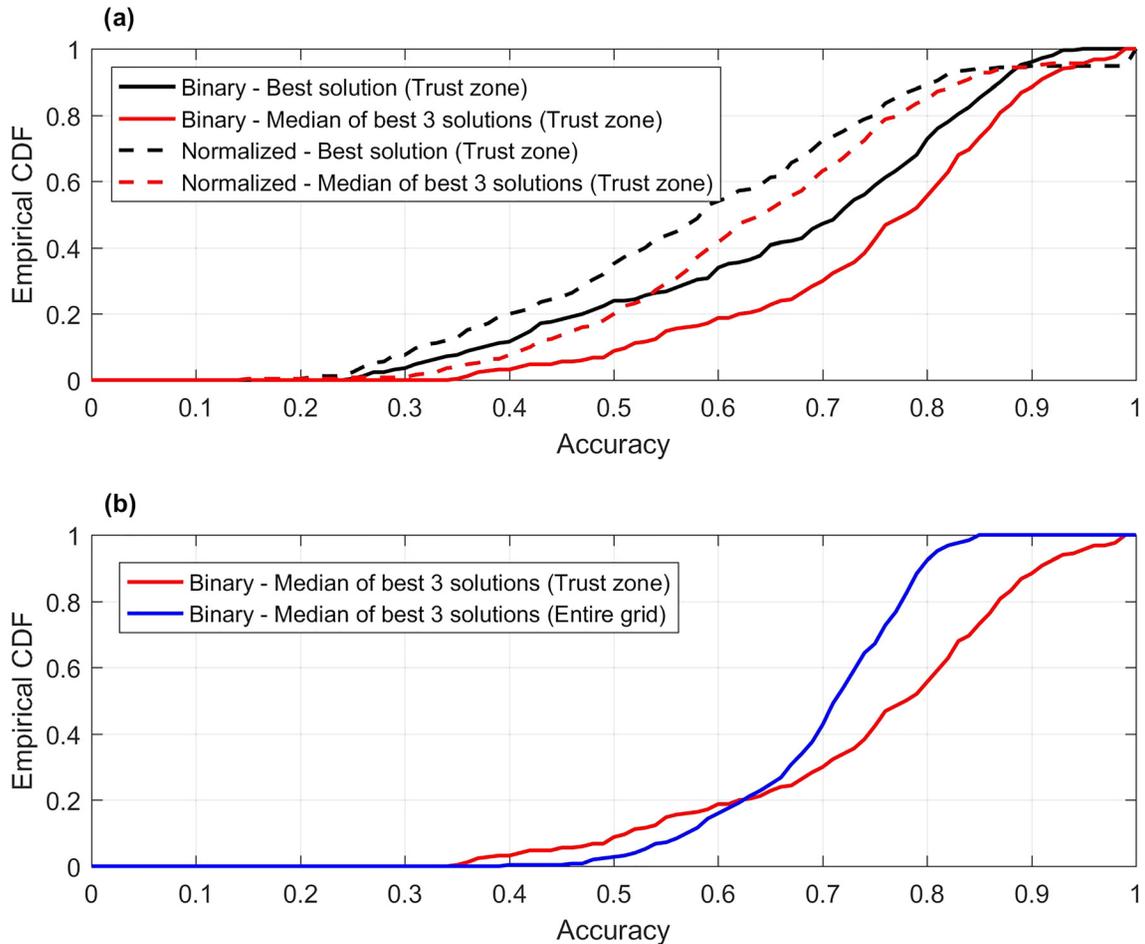


FIG. 9. (Color online) Comparison of the cumulative density functions (CDFs) of the accuracy values of the 250 random source distributions. (a) compares the accuracy of the best (black) and median of the three best SA chains (red). (b) compares the accuracy of the entire source grid and the source grid nodes within the trust zone.

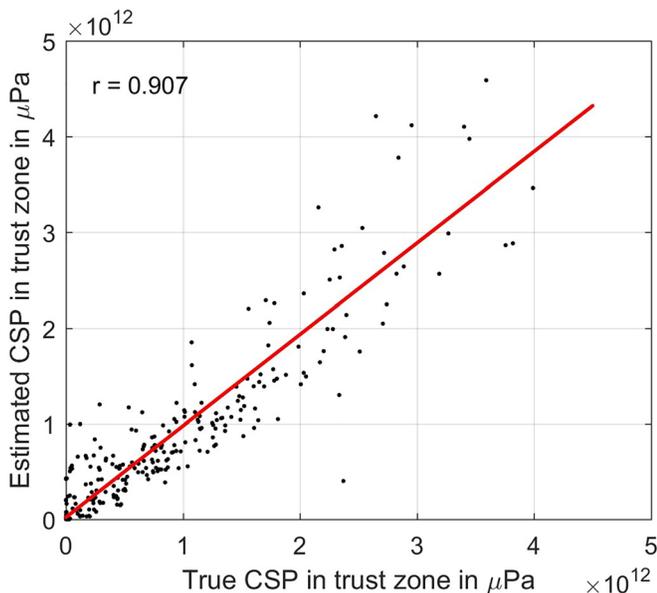


FIG. 10. (Color online) Comparison of the true and estimated CSPs in  $\mu\text{Pa}$ . Each dot represents 1 of the 250 random source distributions. The red line is a linear fit to the data.

between true and simulated RL. Two clusters can be identified: a group of distributions with a RL range below 20 dB, which contains both low and high  $A_n$  values (0.4–1), and a cluster with RL ranges above 20 dB, which contains mainly high  $A_n$  values (0.6–0.9). Figure 12(c) shows that the cluster with RL ranges above 20 dB represents distributions closer to the recording array, which create a correspondingly larger RL range. The blue hues in the right cluster indicate a smaller average distance between sources and receivers ( $<2000$  km). It is also separated from the other distributions through higher misfit values [yellow hues in Fig. 12(a)]. The cluster with RL ranges below 20 dB shows a large gradient of  $A_n$  values that corresponds to the gradient of true source pressure entropy (supplemental Fig. 6).<sup>1</sup> Accuracy shows an inverse relation to the entropy of the true source pressure [Fig. 12(b)], which also corresponds to an increase in CSP. This means that the inversion works best for distributions with low variance (such as many empty nodes) and less well for distributions with high variance. Accuracy is increasing with increasing misfit between true and estimated RLs [Fig. 12(e)] for misfit values between  $-100$  and  $-50$  dB, which also exhibit high estimated source pressure entropy values (yellow hues) and shows no clear relationship for misfit values above  $-50$  dB. The entropy of the estimated source

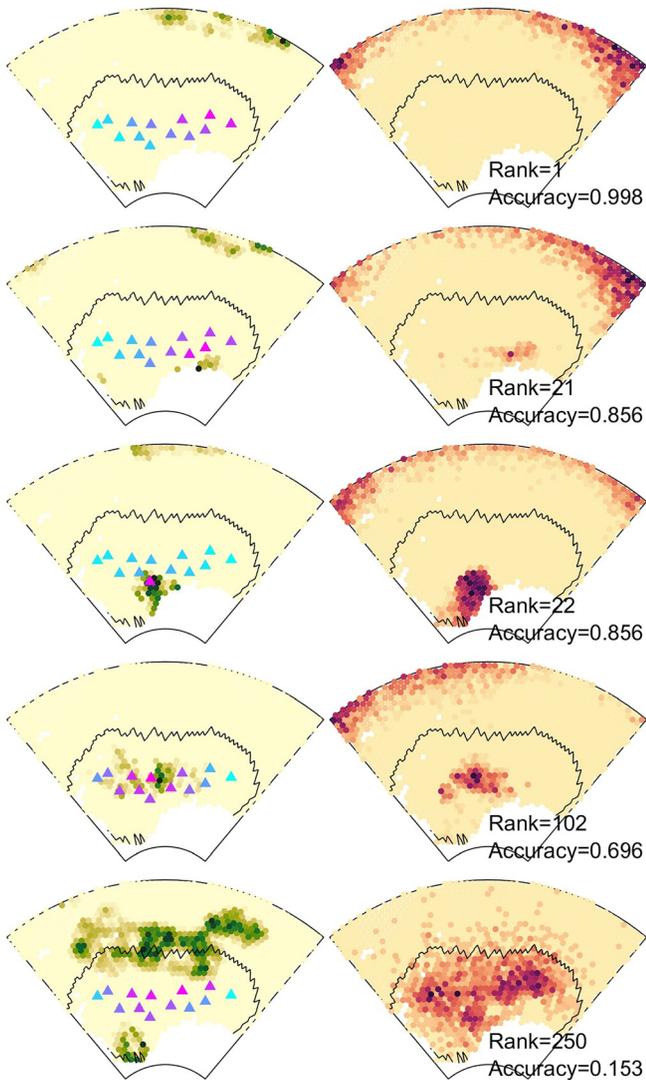


FIG. 11. (Color online) Various examples of true and estimated distributions sorted from the most accurate inversion (rank 1,  $A_n = 0.996$ ) to the least accurate inversion (rank 250,  $A_n = 0.162$ ). Maps in the left columns show the modelled source pressure grid and recorder locations (triangles, color indicates RL), and maps in the right columns show the estimated source pressure grid. Node color indicates the source pressure in  $\mu\text{Pa}$ , normalized for each scenario (row). The black contours encircle the trust zone (more than one recorder present within a 1000 km radius).

pressure shows a marked relationship to accuracy [Fig. 12(f)]: accuracy decreases with increasing entropy of the estimated source pressure following a linear function ( $r = 0.87$ ). However, the relationship between the entropy of the true and estimated source pressure is not linear and shows only limited correlation [Fig. 12(d)]. Another metric of the quality of the estimate is the width (variance) of the estimated source pressure distribution for each node. We compared the mean (averaged over all nodes) range of the best three source pressure estimates to normalized accuracy in Fig. 12(g). We used the range as an indicator of the variance due to the small number of solutions. Estimates with a small range (below  $10^9 \mu\text{Pa}$ ) between the three best solutions show the highest accuracy, whereas estimates with a range larger than  $10^9 \mu\text{Pa}$  show a large spread in accuracy. This spread corresponds to a gradient in the entropy of the estimated source pressure

(color). When only estimates with low entropy are considered (blue dots), a robust relation between the range of the estimates and accuracy exists.

The scatterplots in Fig. 12 show that the entropy and spread (variance) of the estimated source pressure can be used as a metric for inversion accuracy when no other information is available. The best inversion accuracy was achieved for estimated source distributions with low entropy, meaning that many nodes have similar values (are empty) and distributions were patchy; however, inversion was also successful for distributions with high entropy when the RL gradient/range was sufficiently high.

We also analyzed the true and false positive rates of the estimated source pressure grids, considering only source presence/absence information (supplemental Fig. 7<sup>1</sup>), and found that an increasing true positive rate corresponds to decreasing misfit and RL range, whereas an increasing false positive rate corresponds to increasing entropy of the estimated source pressure.

## B. Effect of inaccurate TL model

The effect of a flawed TL model on inversion accuracy was tested by using 100 random distributions with the ray-tracing and geometrical spreading TL models. Figure 13 compares the inversion accuracy for inversion with a perfect and inaccurate TL model. Both the binary and normalized accuracies show a clear but small negative offset in the CDF (mean offset is 0.06) when the TL model is inaccurate compared to the perfect TL model. The inversion method still produced reliable source pressure grid estimates when complex multipath propagation of sound was approximated with a simple geometrical spreading model, at least for the deep ocean with upward refracting sound speed profile in the study area.

## C. Sensitivity tests

Inversion accuracy was not impacted by changes in frequency. No significant change was detected among 27, 98, 150, and 270 Hz, and the binary accuracies were 0.83, 0.85, 0.85, and 0.87, respectively. However, inversion accuracy showed a marked relationship with the number of SA chains (solutions), which determines the resolution with which the CSP range is sampled. Figure 14 shows how inversion accuracy increases with an increasing number of SA chains. Within the trust zone, accuracy increases until around 20 solutions, whereas the accuracy of the entire grid increases continuously up to 40 solutions. This can also be seen when visually comparing the true [Fig. 14(b)] and estimated source pressure grids from inversions with an increasing number of SA chains [Figs. 14(c)–14(j)]. Five SA chains proved way too little to approximate the source distribution adequately, whereas inversions using 10–25 SA chains resolved the central cluster of sources but showed excess sound sources at the northern search space boundary. Inversion using more than 30 SA chains resolved the central cluster of sources and did not show excess sound sources at the search space boundaries. These results indicate that the inversion algorithm stores excess sound energy at the search space

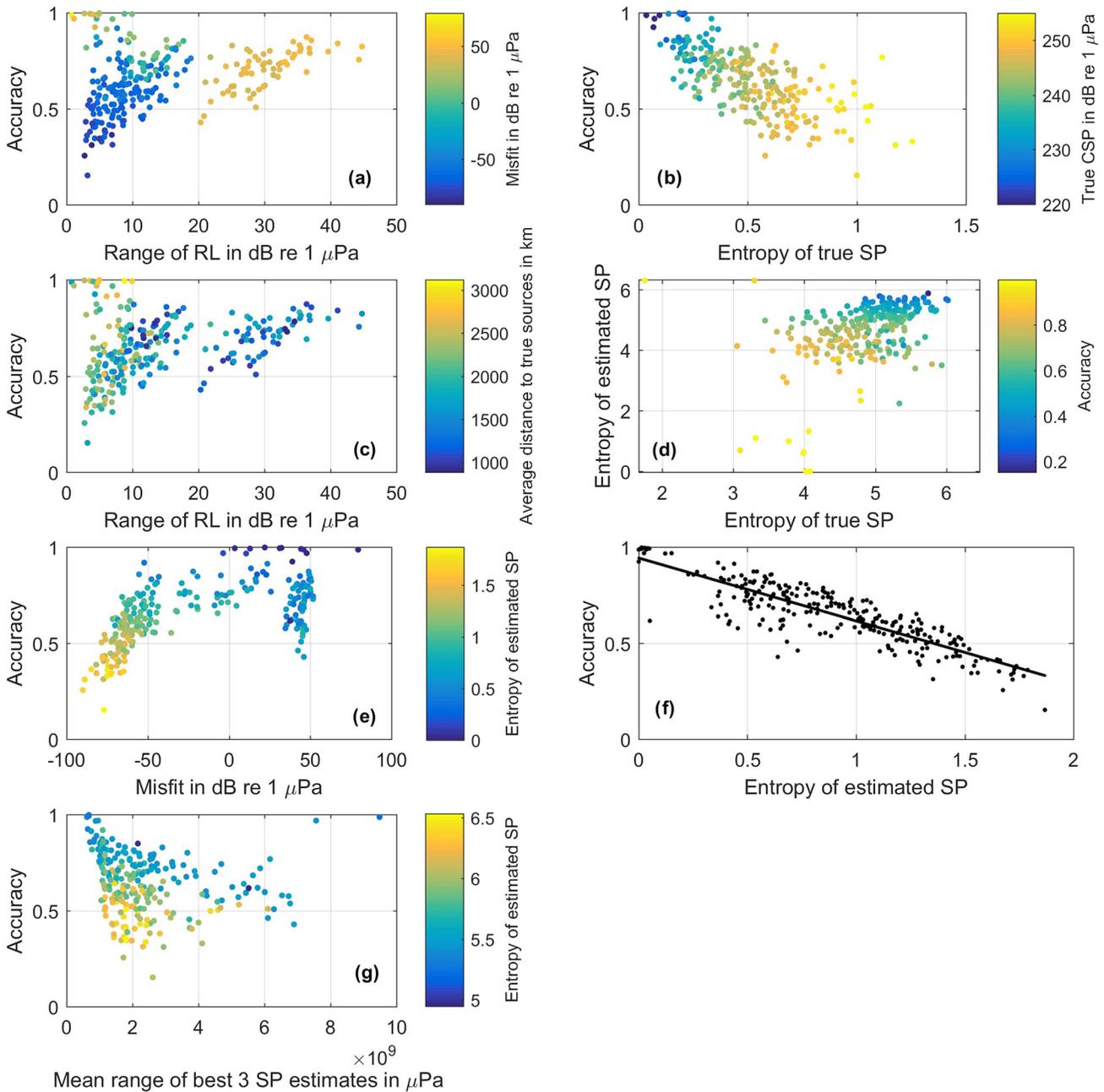


FIG. 12. (Color online) Scatterplots comparing the normalized inversion accuracy ( $A_n$  within the trust zone) and several metrics. Each dot represents 1 of the 250 random source distributions. (a) The range of RLs and the misfit between true and simulated RL (colors), (b) entropy of the true source pressure (SP) and true CSP (colors), (c) the range of RLs and the average distance between the true sources and recorders (colors), (d) entropy of the true and estimated SP (colors indicate accuracy), (e) misfit between true and estimated RL and entropy of the estimated source pressure, (f) entropy of the estimated source pressure, where the black line represents fitted linear function, and (g) the mean range of the three best estimated source pressure estimates and entropy of the estimated SP (color).

boundaries when the CSP distribution is too coarsely sampled (too few SA chains). The number of iterations to achieve successful inversion proved to be remarkably low (Fig. 15) in the test scenario. For the scenario described in Sec. III C, the increase in inversion accuracy flattened out after approximately 5000 iterations.

#### D. Simulation of drifting recorders using Argo float tracks

The suitability of Argo floats as drifting ambient sound recorders was tested using a random source distribution and

the location of Argo float profiles over 71 days. The true source distribution, estimated source distribution, and recorder locations for six example days (sorted after inversion accuracy) are displayed in Fig. 16. When a sufficient number of Argo profiles (recorders) were present and their locations were spread evenly over the grid, inversion was successful with normalized accuracies up to 0.7 for the entire source grid. But, on days with very few or less evenly distributed floats, the inversion was unsuccessful. To investigate the necessary conditions for successful inversion, the scatterplots in Fig. 17 compare normalized accuracy over the number and location of recorders and the node entropy of

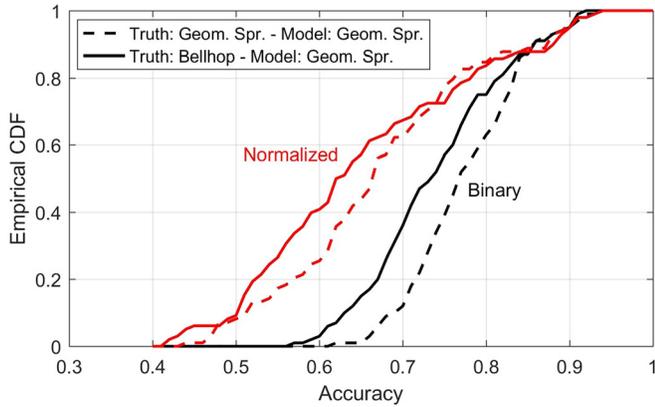


FIG. 13. (Color online) Comparison of the CDFs of accuracy (red, normalized accuracy; black, binary accuracy) of the 100 random source distribution estimates within the trust zone. The dashed lines show the accuracy for inversions with a perfect TL model, where the true and forward model TLs were both calculated using geometrical spreading. The solid lines show the accuracies for inversions with a flawed TL forward model, where the true TL was calculated using raytracing with BELLHOP, whereas the forward model TL was calculated using geometrical spreading.

the estimated source pressure. Whereas the entropy  $H(P(SP_j))$  quantifies the flatness of the source pressure *sample distribution*, the node entropy  $H(SP_j)$  determines the flatness of the source pressure grid directly by summing over the nodes

$$H(SP_j) = - \sum SP_j \log(SP_j).$$

Both entropy metrics are low when the source pressure distribution has a low variance (many similar values, mainly empty nodes) and high when the source pressure distribution has a high variance.

We found that when less than 15 recorders were present, inversion accuracy ranged between 0.4 and 0.75, whereas

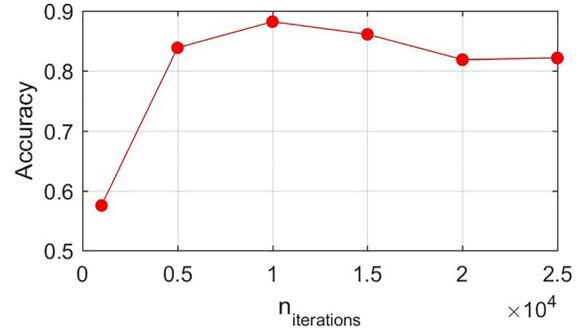


FIG. 15. (Color online) Binary inversion accuracy over the number of iterations for the test scenario.

accuracy was between 0.6 and 0.75 when more than 15 recorders were present. We found a close and almost linear relationship between the RL range and inversion accuracy, independent of the number or recorders. This indicates that inversion accuracy depends on both the number of recorders and the RL gradient (range). As for the 250 random source distributions in the first test scenario, we found a close relationship between inversion accuracy and the entropy of the estimated source pressure. In this scenario, the relationship between node entropy and accuracy was linear. Normalized accuracies were above 0.7 on 45% of the simulated days.

## V. DISCUSSION

The test scenarios showed that it is possible to estimate the distribution of sound sources from ambient sound using widely spaced recorder arrays, but also demonstrated the limitations of the method and explored the prerequisites for successful inversion. Sections VA–VD interpret the results of the test scenarios and discuss the feasibility to apply this inversion method to real ambient sound data.

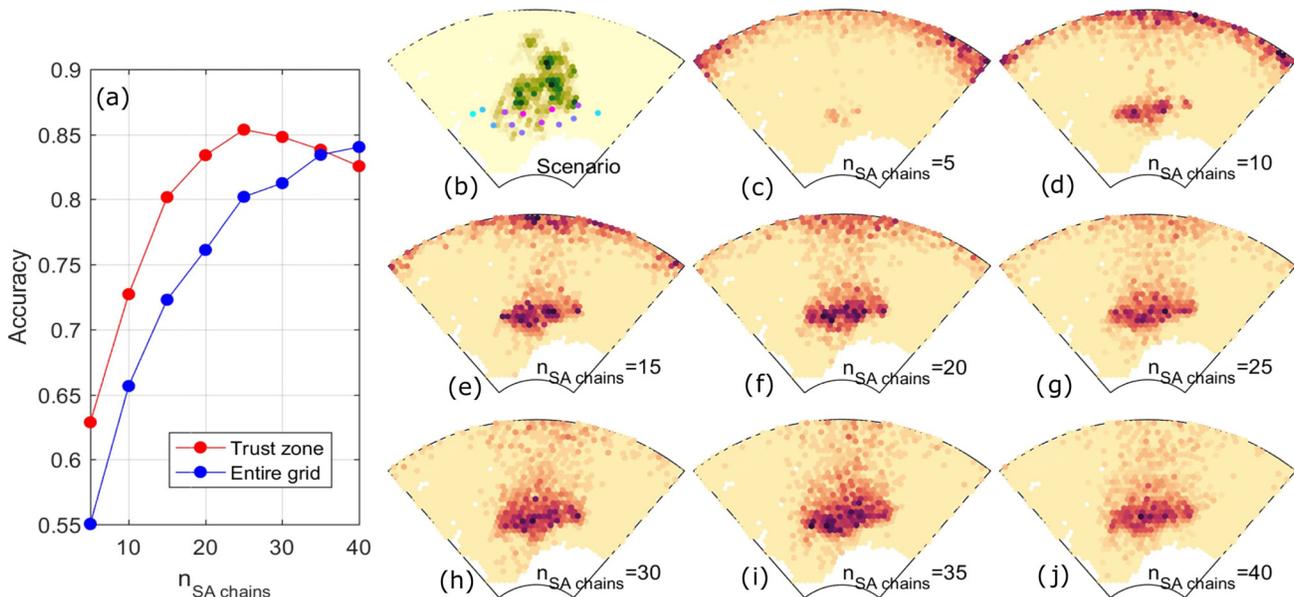


FIG. 14. (Color online) The effect of the number of SA chains on inversion accuracy. (a) shows how binary accuracy in the trust zone and entire grid increases with increasing number of SA chains. (b) shows the test scenario source distribution (green hues) and RL (blue-pink hues). (c)–(j) show the estimated source pressure grid for inversion using 5–40 SA chains.

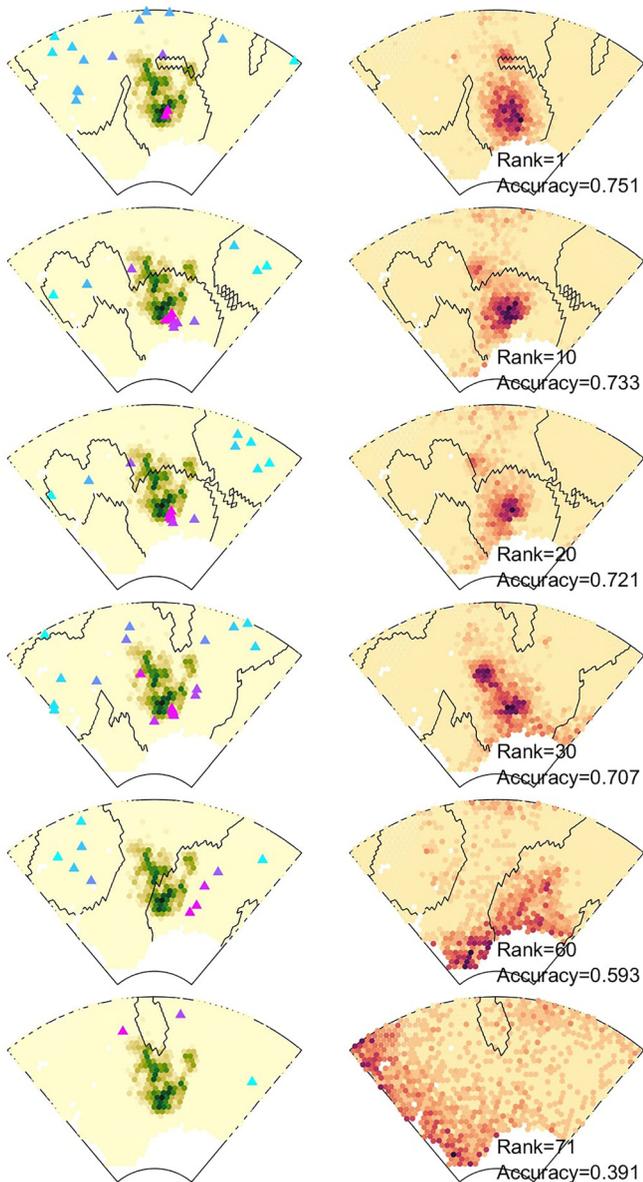


FIG. 16. (Color online) Example true and estimated distributions, sorted from best inversion to worst (rank 71) using normalized accuracy over the entire grid. Maps on the left show the true source pressure grid and recorder (Argo float) locations (triangles, color indicates RL), and maps of the right show the estimated source pressure grid. Node color indicates the source pressure in  $\mu\text{Pa}$ , normalized for each scenario (row). The black contours encircle the trust zone (more than one recorder present within a 1000 km radius).

### A. Accuracy and reliability of the inversion method

The random distribution and Argo float test scenarios showed that inversion accuracy can be predicted using the entropy and spread of the estimated source pressure grid [Figs. 12(f) and 17(b)] and range or gradient of the RLs [Figs. 12(c) and 17(c)]. Both the test scenarios with fixed recorders and random source distributions, and Argo float scenario with variable recorders and a fixed source distribution, indicated that an inversion is likely inaccurate when the estimated sources distribution has a high entropy and accurate when the source distribution has low entropy (many empty nodes and a patchy distribution). The reason for this is likely that the misfit function does not have a pronounced

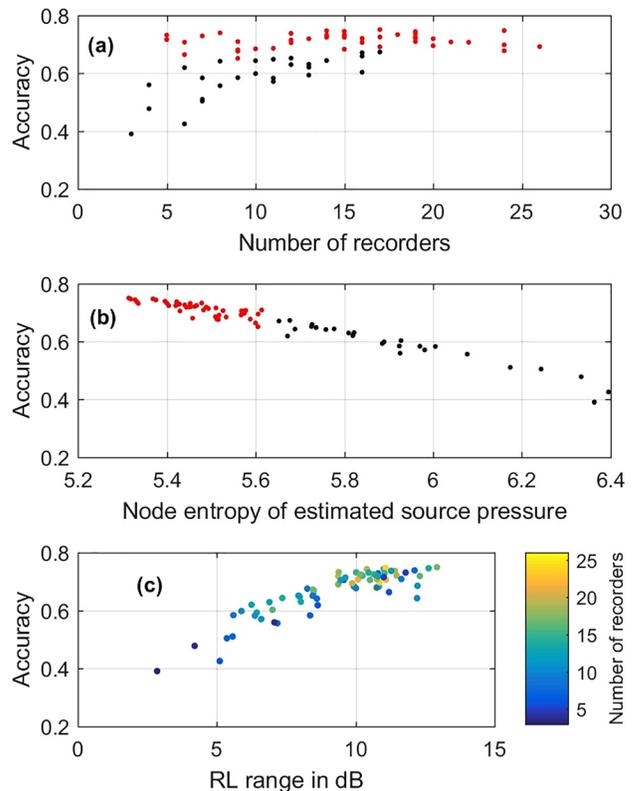


FIG. 17. (Color online) Scatterplots comparing normalized inversion accuracy over the entire grid for several metrics. (a) Number of recorders (Argo floats), (b) node entropy of the estimated source pressure, (c) range of RL values and number of recorders (color). Red dots indicate inversions with a node entropy of the estimated source pressure below 5.65.

global minimum when the inversion algorithm does not have sufficient information (too few recorders or too small RL gradient), forcing the parameter estimation algorithm to spread the sources over the search space. The comparisons between accuracy and RL range [Fig. 12(c) and 17(c)] demonstrated that an increased RL gradient, and resulting increase in RL information, benefits inversion accuracy, but inversion can also be successful with small RL gradients when the true source distribution contains no sources in the trust zone or has a low entropy (many empty nodes).

The test scenario with fixed recorders and random source distributions showed an inverse relationship between inversion accuracy and true source pressure entropy [Fig. 12(b)], indicating that the recorder array used in this scenario is most suitable to locate clustered distribution and regions with no sources. This could be related to a lack of gradients in the RL dataset for more uniform source distributions. Adding more recorders to the array and adjusting the spacing of the array would increase the information present in the RL dataset and, thus, improve the inversion accuracy. The effect of recorder array geometry on inversion accuracy will be studied with further simulations that would extend the scope of this paper.

It was crucial to test the effect of an inaccurate TL model on inversion accuracy, since TL models are only, more or less, a rough approximation of the true TL as it is challenging to model underwater sound propagation correctly. Most available models are only 2D, do not include

sea ice, and are computationally expensive. In high latitude oceans, such as the Weddell Sea, the effect of sea ice on TL can be profound, but few operational TL models that include TL from sea ice exist. We therefore compared the inversion accuracy of 100 random distributions with a perfect and a flawed TL model (Fig. 13). When using a flawed TL model, by approximating the true raytracing TL with a geometrical spreading model, the accuracy of the CDF shifted, on average, 0.06 toward smaller values. This means that inversion accuracy is only slightly affected by the flawed TL model in our study area (open upward refracting ocean), and thus simple TL models (such as geometrical spreading) could be used for inversions based on real data. This is likely the case due to the long distances and many source-receiver pathways of the inverse problem. Since the recorder array is widely spaced, small-scale variations in TL are not resolved, and the many pathways likely average out TL errors. As long as the TL model resolves the non-linear gradient of the TL on the scale of hundreds to thousands of km (supplemental Fig. 4<sup>1</sup>), inversion accuracy is only slightly decreased when approximating true TL with the geometrical spreading model. If this holds for ocean areas with more complex propagation characteristics than the deep offshore Southern Ocean remains to be studied with further simulations. Ocean areas with waveguides or complex topography will likely need more sophisticated TL forward models for successful inversion.

The area in which the inversion produces reliable estimates, the trust zone, was approximated by studying the correlation between the true and estimated source pressures of hundreds of random source distributions (Fig. 8). Such an approach could also be applied to estimate the trust zone when real data are used. The size and shape of the trust zone depends on the number and location of recorders; placing a large number of recorders uniformly over the study area is likely the best way to record a suitable dataset for inversion. This is supported by the results of the ARGO float simulations (Fig. 17). Within the trust zone, the inversion algorithm successfully estimated the CSP for most of the 250 random source distributions ( $r = 0.9$ ; Fig. 10). It is important that the pressure values are not biased since the source pressure at each node is the basis of eventually estimating the number of animals per area by multiplying source pressure per area with (yet unknown) species specific coefficients. Increasing the number of SA chains, which determines how many different CSP samples are calculated, can likely increase this correlation even more.

The sensitivity tests (Figs. 14 and 15) showed that successful estimates of source distribution can be computed with reasonable effort ( $\sim 10\,000$  iterations and 30 SA chains for the Weddell Sea test scenario). As expected, the more SA chains are used for the inversions, the better the estimate becomes. Using too few SA chains under-samples the CSP distribution, resulting in estimates with either a too low or too high CSP. The parameter estimation algorithm stores this excess sound energy (which cannot be located sufficiently) at the boundaries of the search space to match the general gradient of RL in the recorder array (Fig. 14). It was expected that accuracy increases with an increasing number

of iterations until a certain value is reached; however, the comparatively small number of iterations needed to calculate accurate inversions was smaller than expected. This means that the source grid size and resolution and the number of recorders can be increased with realizable computational effort.

Another important aspect is that the recorders need to be calibrated sufficiently because biases in the RL data could affect inversion accuracy, and the inversion method relies on absolute RL values and small gradients. However, the inaccurate TL test scenario (Fig. 13) showed that small errors in the forward model are tolerated by the inversion method, thus, small errors in RL should be tolerated similarly by the inversion method. Ideally each recording device should be calibrated before deployment. If this is not possible, the gain should be chosen so that part of the recorded spectra hit the noise floor of the recording device. This noise floor can then be compared to the factory calibration values of the hydrophone and recording device, and eventual offsets detected. An example of this post-deployment calibration check can be found in Menze *et al.* (2017). It is also a suitable way of quantifying the recorders self-noise. If it is too high, faint MMC peaks in the ambient sound might not be detected.

## B. Requirements for successful inversion

To apply the inversion method to real MMC data and get reliable source pressure distribution estimates, several prerequisites need to be fulfilled. First, the number of recorders needs to be large enough, and they need to record a large enough  $RL_{MMC}$  gradient. For the Weddell Sea scenario, already up to ten recorders can be sufficient, but more are preferred (Figs. 12 and 17). The recorders are best spread evenly over the study area to record as much RL gradient (large range of RL values) as possible to maximize the information content of the RL dataset. One of the most important requirements is that the MMC to ambient sound should be detectable in the first place. This depends not only on the number of vocalizing animals in the area but also on noise from shipping, seismic surveys, and sea surface motion. In regions with high marine traffic, the MMC peaks are likely masked by shipping noise, leading to a lack of  $RL_{MMC}$  measurements and low inversion accuracy. The inversion method is thus most suitable in remote regions far away from anthropogenic activity, which are also difficult to survey with traditional methods due to their remoteness. The inversion method is based on minimizing the misfit between recorded and modelled  $RL_{MMC}$ , thus, offsets and biases in the recorded  $RL_{MMC}$  can lead to erroneous inversion results. The recorders need to be properly calibrated to provide reliable  $RL_{MMC}$  data. Second, the number of iterations and solutions needs to be sufficiently high. Third, the source pressure grid should be large enough to cover all possibly expected source locations and have an adequate resolution (distance between grid nodes). Fourth, the TL matrix between the recorders and grid nodes should be calculated as accurately as possible.

### C. Argo floats as ambient sound recorders

The Argo float test scenario showed that ambient sound data from drifting recorders could successfully be used for inversion (Fig. 17). However, the test also showed that successful inversion in the study area was only possible on approximately 45% of the simulated days, and on the other days there were too few profiles or profiles at unfavorable locations. This is related to the sparse number of Argo floats in the Southern Ocean (Reeve *et al.*, 2016); mid latitude areas have much better Argo float coverage than high latitude areas, thus, Argo floats are likely suitable for inversions in most of the world's oceans with the current Argo float array.

To obtain ambient sound data suitable for MMC inversion from Argo floats, several specifications need to be fulfilled. The location of the float needs to be known [the float needs to surface to get a global positioning system (GPS) fix or be localized acoustically]. The float needs to be able to record 5–30 min of sound in the right frequency band and sufficient dynamic range with a calibrated hydrophone. Furthermore, the floats would need to be able to calculate and transmit a power spectrum of the recording. Finally, all floats need to record at approximately the same time and date. The timing does not need to be accurate on the scale of seconds but should agree on the scale of minutes to ensure that only spatial variability of the MMC is recorded. To ensure consistency, the float should also stay at a fixed depth. It is likely most practical to record ambient sound for 10 min at the floats drifting depth (approximately 1000 m) before the float surfaces to measure a temperature and salinity profile. The technology to record ambient sound and transmit spectra with Argo floats has already been developed and successfully tested (Matsumoto *et al.*, 2013; Nystuen *et al.*, 2011), but a large transnational effort is necessary to create and deploy an Argo float array sufficient for MMC inversion. We propose that the scope of future Argo float deployments not only contain oceanographic and bio-geo-chemical sensors but also a calibrated hydrophone and necessary data processing capabilities, which cannot only be used to study marine mammal distribution but also rain fall rate and air–sea–ice interaction (Cazau *et al.*, 2018; Ma *et al.*, 2005).

### D. Application of the inversion method

We could demonstrate that successful inversion of MMCs is possible with the HAFOS mooring array (Van Opzeeland *et al.*, 2014). Inversion should be possible with all four MMCs (Blue, fin, and Antarctic minke whales, and leopard seals) and could allow year-round monitoring of the distribution of vocalizing marine mammals in the Weddell Sea. To obtain values of animal distribution and density (average number of animals in study area), the source pressure (SP) per area values needs to be multiplied with the population specific call rate (CR) and SL values

$$\frac{n_{\text{animals}}}{\text{area}} = \frac{\text{SP}}{\text{area}} \text{CR} 10^{\text{SL}/20}.$$

Reliable values for CR and SL are very difficult to obtain and, therefore, we did not investigate such density estimation

yet. These multipliers are likely not constant with time and region and similar to the multipliers used in call detection estimation methods (Thomas and Marques, 2012). In addition to MMCs in the Southern Ocean, the inversion method could be applied to the MMC of fin whales in the Mid and North Atlantic (Nieukirk *et al.*, 2012), fin and blue whales in the North Pacific (Curtis *et al.*, 1999), fin and possibly Bowhead whales in the Arctic (Ahonen *et al.*, 2017), and fin and blue whales in the Indian Ocean (Leroy *et al.*, 2018a). Data from the widely spaced recorder arrays used in these studies show temporal and spatial patterns in  $RL_{\text{MMC}}$  suggestive of seasonal migration.

## VI. CONCLUSION

We presented and tested an approach to estimate the distribution of vocalizing marine mammals based on inverse modelling and the spatial variation in ambient sound spectra instead of the detection of individual, transient vocalizations. Despite the under-determinedness of this inverse problem, the parameter estimation algorithm successfully estimated the spatial distribution of sound sources in a set of test scenarios, which showed that inversion accuracy depends on the number (and gradient) of RL observations, number of SA chains, and sound source distribution entropy. The accuracy of the estimates is only slightly affected by inevitable inaccuracies in the TL model. Test simulations indicated that drifting platforms, such as Argo floats, can be suitable to gather MMC data. Applying the method to ambient sound recordings from the Southern Ocean renders it possible to study the distribution and migration of vocalizing marine mammals on unprecedented spatial scales and temporal resolution, and compliments existing visual and acoustic estimation methods. The approach we explored in this paper could also be applied to recordings of other species that generate chorus-like sounds, e.g., insects, amphibians, and birds, provided that the sounds propagate far enough and are generated often enough to form a chorus. Calibrated recorders are used, and the TL between the recorders and sound sources can be sufficiently modelled.

## ACKNOWLEDGMENTS

We would like to thank Randi Ingvaldsen and the Institute of Marine Research for providing supervision, funding (Research council of Norway Grant No. 228896), and scientific freedom to S.M., and the Office of Naval Research for funding D.Z. with Grant No. N00014-18-1-2811. We also thank the reviewers and Len Thomas for their feedback. The authors declare no competing interests.

<sup>1</sup>See supplementary material at <https://doi.org/10.1121/1.5139406> for a pdf with the supplementary figures and a zip folder containing the code used for this study. The supplementary material can also be found in the github repository <https://github.com/sebastianmenze/ambient-sound-inversion>

<sup>2</sup>See <http://www.argodatamgt.org/Access-to-data/Argo-data-selection> (Last viewed 2/23/2018).

Ahonen, H., Stafford, K. M., de Steur, L., Lydersen, C., Wiig, Ø., and Kovacs, K. M. (2017). “The underwater soundscape in western Fram Strait: Breeding ground of Spitsbergen’s endangered bowhead whales.” *Mar. Pollut. Bull.* **123**(1-2), 97–112.

- Amante, C., and Eakins, B. W. (2009). "ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis, in *NOAA Technical Memorandum NESDIS NGDC-24*, National Geophysical Data Center, NOAA, available at <https://doi.org/10.7289/V5C8276M> (Last viewed 5/1/2016).
- Argo (2018). "Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)" (SEANOE), available at <http://doi.org/10.17882/42182>.
- Bombosch, A., Zitterbart, D. P., Van Opzeeland, I., Frickenhaus, S., Burkhardt, E., Wisz, M. S., and Boebel, O. (2014). "Predictive habitat modelling of humpback (*Megaptera novaeangliae*) and Antarctic minke (*Balaenoptera bonaerensis*) whales in the Southern Ocean as a planning tool for seismic surveys," *Deep Sea Res., Part I* **91**, 101–114.
- Bonnell, J., Thode, A. M., Blackwell, S. B., Kim, K., and Macrander, A. M. (2014). "Range estimation of bowhead whale (*Balaena mysticetus*) calls in the Arctic using a single hydrophone," *J. Acoust. Soc. Am.* **136**(1), 145–156.
- Borda, M. (2011). *Fundamentals in Information Theory and Coding* (Springer, Berlin).
- Burtenshaw, J. C., Oleson, E. M., Hildebrand, J. A., McDonald, M. A., Andrew, R. K., Howe, B. M., and Mercer, J. A. (2004). "Acoustic and satellite remote sensing of blue whale seasonality and habitat in the Northeast Pacific," *Deep Sea Res., Part II* **51**(10–11), 967–986.
- Carey, W. M., and Evans, R. B. (2011). *Ocean Ambient Noise* (Springer, New York).
- Carlén, I., Thomas, L., Carlström, J., Amundin, M., Teilmann, J., Tregenza, N., Tougaard, J., Koblitz, J. C., Sveegaard, S., Wennerberg, D., Loisa, O., Dähne, M., Brundiers, K., Kosecka, M., Kyhn, L. A., Ljungqvist, C. T., Pawliczka, I., Koza, R., Arciszewski, B., Galatius, A., Jabbusch, M., Laaksonlaita, J., Niemi, J., Lyytinen, S., Gallus, A., Benke, H., Blankett, P., Skóra, K. E., Acevedo-Gutiérrez, A. (2018). "Basin-scale distribution of harbour porpoises in the Baltic Sea provides basis for effective conservation actions," *Biol. Conserv.* **226**, 42–53.
- Cazau, D., Bonnell, J., and Baumgartner, M. (2018). "Wind speed estimation using acoustic underwater glider in a near-shore marine environment," *IEEE Trans. Geosci. Remote Sens.* **57**, 2097–2106.
- Cronin, B., Stevenson, I. H., Sur, M., and Kording, K. P. (2009). "Hierarchical Bayesian modeling and Markov chain Monte Carlo sampling for tuning-curve analysis," *J. Neurophysiol.* **103**, 591–602.
- Curtis, K. R., Howe, B. M., and Mercer, J. A. (1999). "Low-frequency ambient sound in the North Pacific: Long time series observations," *J. Acoust. Soc. Am.* **106**(6), 3189–3200.
- Dawson, D. K., and Efford, M. G. (2009). "Bird population density estimated from acoustic signals," *J. Appl. Ecol.* **46**(6), 1201–1209.
- Dushaw, B. D., Worcester, P. F., Dziaciuch, M. A., and Menemenlis, D. (2013). "On the time-mean state of ocean models and the properties of long range acoustic propagation," *J. Geophys. Res. Oceans* **118**(9), 4346–4362, <https://doi.org/10.1002/jgrc.20325>.
- Francois, R. E., and Garrison, G. R. (1982). "Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Am.* **72**(6), 1879–1890.
- Granville, V., Krivanek, M., and Rasson, J. P. (1994). "Simulated annealing: A proof of convergence," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(6), 652–656.
- Harris, D., Matias, L., Thomas, L., Harwood, J., and 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**(5), 3522–3535.
- Harris, D. V., Miksis-Olds, J. L., Vernon, J. A., and Thomas, L. (2018). "Fin whale density and distribution estimation using acoustic bearings derived from sparse arrays," *J. Acoust. Soc. Am.* **143**(5), 2980–2993.
- Haver, S. M., Klinck, H., Nieuwkirk, S. L., Matsumoto, H., Dziak, R. P., and Miksis-Olds, J. L. (2017). "The not-so-silent world: Measuring Arctic, equatorial, and Antarctic soundscapes in the Atlantic Ocean," *Deep Sea Res., Part I* **122**, 95–104.
- Herr, H., Kelly, N., Dorschel, B., Huntemann, M., Kock, K., Lehnert, L. S., Siebert, U., Viquerat, S., Williams, R., and Scheidat, M. (2019). "Aerial surveys for Antarctic minke whales (*Balaenoptera bonaerensis*) reveal sea ice dependent distribution patterns," *Ecol. Evol.* March, **9**, 5664–5682.
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). "Optimization by simulated annealing," *Science* **220**(4598), 671–680.
- Kusel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., and Ward, J. (2011). "Cetacean population density estimation from single fixed sensors using passive acoustics," *J. Acoust. Soc. Am.* **129**(6), 3610–3622.
- Kyhn, L. A., Tougaard, J., Thomas, L., Duve, L. R., Stenback, J., Amundin, M., Desportes, G., and Teilmann, J. (2012). "From echolocation clicks to animal density—Acoustic sampling of harbor porpoises with static dataloggers," *J. Acoust. Soc. Am.* **131**(1), 550–560.
- Leroy, E., Samaran, F., Stafford, K., Bonnel, J., and Royer, J. (2018a). "Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the Southern Indian Ocean," *Endanger. Species Res.* **37**, 289–300.
- Leroy, E. C., Royer, J.-Y., Bonnel, J., and Samaran, F. (2018b). "Long-term and seasonal changes of large whale call frequency in the Southern Indian Ocean," *J. Geophys. Res. Oceans* **123**, 8568–8580, <https://doi.org/10.1029/2018JC014352>.
- Lewis, L. A., Calambokidis, J., Stimpert, A. K., Fahlbusch, J., Friedlaender, A. S., McKenna, M. F., Mesnick, S. L., Oleson, E. M., Southall, B. L., Szesciorka, A. R., and Širović, A. (2018). "Context-dependent variability in blue whale acoustic behaviour," *R. Soc. Open Sci.* **5**(8), 180241.
- Lurton, X. (2010). *An Introduction to Underwater Acoustics: Principles and Applications* (Springer, Berlin).
- Ma, B. B., Nystuen, J. A., and Lien, R.-C. (2005). "Prediction of underwater sound levels from rain and wind," *J. Acoust. Soc. Am.* **117**(6), 3555.
- Marques, T. A., Munger, L. M., Thomas, L., Wiggins, S. M., and Hildebrand, J. A. (2011). "Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting," *Endanger. Species Res.* **13**(3), 163–172.
- Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). "Estimating animal population density using passive acoustics," *Biol. Rev. Cambridge Philos. Soc.* **88**(2), 287–309.
- Martin, S. W., Marques, T. A., Thomas, L., Morrissey, R. P., Jarvis, S., Dimarzio, N., Moretti, D., and Mellinger, D. K. (2013). "Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors," *Mar. Mammal Sci.* **29**, 142–158.
- Matsumoto, H., Jones, C., Klinck, H., Mellinger, D. K., Dziak, R. P., and Meinig, C. (2013). "Tracking beaked whales with a passive acoustic profiler float," *J. Acoust. Soc. Am.* **133**(2), 731–740.
- McCauley, R. D., Gavrilov, A. N., Jolliffe, C. D., Ward, R., and Gill, P. C. (2018). "Pygmy blue and Antarctic blue whale presence, distribution and population parameters in southern Australia based on passive acoustics," *Deep Sea Res., Part II* **157–158**, 154–168.
- McDonald, M. A., Hildebrand, J. A., Wiggins, S. M., and Ross, D. (2008). "A 50 year comparison of ambient ocean noise near San Clemente Island: A bathymetrically complex coastal region off Southern California," *J. Acoust. Soc. Am.* **124**(4), 1985–1992.
- Mellinger, D. K., Küsel, E. T., Harris, D., Thomas, L., and Matias, L. (2014). "Estimating singing fin whale population density using frequency band energy," *J. Acoust. Soc. Am.* **136**(4), 2275.
- Menze, S., Zitterbart, D. P., van Opzeeland, I., and Boebel, O. (2017). "The influence of sea ice, wind speed and marine mammals on Southern Ocean ambient sound," *R. Soc. Open Sci.* **4**(1), 160370.
- Mosegaard, K., and Sambridge, M. (2002). "Monte Carlo analysis of inverse problems," *Inverse Probl.* **18**, R29–R54.
- Mouy, X., Hannay, D., Zykov, M., and Martin, B. (2012). "Tracking of Pacific walrus in the Chukchi Sea using a single hydrophone," *J. Acoust. Soc. Am.* **131**(2), 1349.
- Newhall, A. E., Lin, Y.-T., Lynch, J. F., Baumgartner, M. F., and Gawarkiewicz, G. G. (2012). "Long distance passive localization of vocalizing sei whales using an acoustic normal mode approach," *J. Acoust. Soc. Am.* **131**(2), 1814–1825.
- Nieuwkirk, S. L., Mellinger, D. K., Moore, S. E., Klinck, K., Dziak, R. P., and Goslin, J. (2012). "Sounds from airguns and fin whales recorded in the mid-Atlantic Ocean, 1999–2009," *J. Acoust. Soc. Am.* **131**(2), 1102–1112.
- Nystuen, J., Riser, S., Wen, T., and Swift, D. (2011). "Interpreted acoustic ocean observations from Argo floats," *J. Acoust. Soc. Am.* **129**, 2400.
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H., and Pieretti, N. (2011). "Soundscape ecology: The science of sound in the landscape," *BioScience* **61**(3), 203–216.
- Porter, M. B. (1987). "Gaussian beam tracing for computing ocean acoustic fields," *J. Acoust. Soc. Am.* **82**(4), 1349–1359.
- Reeve, K. A., Boebel, O., Kanzow, T., Strass, V., Rohardt, G., and Fahrbach, E. (2016). "A gridded data set of upper-ocean hydrographic properties in the Weddell Gyre obtained by objective mapping of Argo float measurements," *Earth Syst. Sci. Data* **8**(1), 15–40.
- Risch, D., Castellote, M., Clark, C. W., Davis, G. E., Dugan, P. J., Hodge, L. E., Kumar, A., Lucke, K., Mellinger, D. K., Nieuwkirk, S. L., Popescu,

- C. M., Ramp, C., Read, A. J., Rice, A. N., Silva, M. A., Siebert, U., Stafford, K. M., Verdaat, H., and Van Parijs, S. M. (2014). "Seasonal migrations of North Atlantic minke whales: Novel insights from large-scale passive acoustic monitoring networks," *Mov. Ecol.* **2**(1), 24.
- Rogers, T. L., Ciaglia, M. B., Klinck, H., and Southwell, C. (2013). "Density can be misleading for low-density species: Benefits of passive acoustic monitoring," *PLoS One* **8**(1), e52542.
- Seger, K. D., Thode, A. M., Urbán-R., J., Martínez-Loustalot, P., Jiménez-López, M. E., and López-Arzate, D. (2016). "Humpback whale-generated ambient noise levels provide insight into singers' spatial densities," *J. Acoust. Soc. Am.* **140**(3), 1581–1597.
- Sepkoski, J. J. (1974). "Quantified coefficients of association and measurement of similarity," *J. Int. Assoc. Math. Geol.* **6**(2), 135–152.
- Shannon, C. E. (1948). "A mathematical theory of communication," *Bell Syst. Tech. J.* **27**, 379–423.
- Širović, A., Hildebrand, J. A., and Wiggins, S. M. (2007). "Blue and fin whale call source levels and propagation range in the Southern Ocean," *J. Acoust. Soc. Am.* **122**(2), 1208–1215.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, Philadelphia, PA).
- Teanby, N. A. (2006). "An icosahedron-based method for even binning of globally distributed remote sensing data," *Comput. Geosci.* **32**, 1442–1450.
- Thomas, L., and Marques, T. A. (2012). "Passive acoustic monitoring for estimating animal density," *Acoust. Today* **8**(3), 35.
- Thomisch, K., Boebel, O., Clark, C., Hagen, W., Spiesecke, S., Zitterbart, D., and Van Opzeeland, I. (2016). "Spatio-temporal patterns in acoustic presence and distribution of Antarctic blue whales *Balaenoptera musculus* intermedia in the Weddell Sea," *Endanger. Species Res.* **30**, 239–253.
- Urazghildiev, I. R., and Clark, C. W. (2013). "Comparative analysis of localization algorithms with application to passive acoustic monitoring," *J. Acoust. Soc. Am.* **134**(6), 4418–4426.
- Urazghildiev, I. R., and Hannay, D. E. (2018). "Passive acoustic detection and estimation of the number of sources using compact arrays," *J. Acoust. Soc. Am.* **143**(5), 2825–2833.
- Valtierra, R. D., Glynn Holt, R., Cholewiak, D., and Van Parijs, S. M. (2013). "Calling depths of baleen whales from single sensor data: Development of an autocorrelation method using multipath localization," *J. Acoust. Soc. Am.* **134**(3), 2571–2581.
- Van Opzeeland, I., Samaran, F., Stafford, K. M., Findlay, K., Gedamke, J., Harris, D., and Miller, B. S. (2014). "Towards collective circum-Antarctic passive acoustic monitoring: The Southern Ocean hydrophone network (SOHN)," *Polarforschung* **83**(2), 47–61.
- Van Opzeeland, I., Van Parijs, S., Kindermann, L., Burkhardt, E., and Boebel, O. (2013). "Calling in the cold: Pervasive acoustic presence of humpback whales (*Megaptera novaeangliae*) in Antarctic coastal waters," *PLoS One* **8**(9), 1–7.
- Van Parijs, S., Clark, C., Sousa-Lima, R., Parks, S., Rankin, S., Risch, D., and Van Opzeeland, I. (2009). "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," *Mar. Ecol.: Prog. Ser.* **395**, 21–36.
- Wall, C. C., Simard, P., Lembke, C., and Mann, D. A. (2013). "Large-scale passive acoustic monitoring of fish sound production on the West Florida Shelf," *Mar. Ecol.: Prog. Ser.* **484**, 173–188.
- Wang, D., Garcia, H., Huang, W., Tran, D. D., Jain, A. D., Yi, D. H., Gong, Z., Jech, J. M., Godø, O. R., Makris, N. C., and Ratilal, P. (2016). "Vast assembly of vocal marine mammals from diverse species on fish spawning ground," *Nature* **531**, 366–370.
- Ward, J. A., Thomas, L., Jarvis, S., Dimarzio, N., Moretti, D., Marques, T. A., Dunn, C., Claridge, D., Hartvig, E., and Tyack, P. (2012). "Passive acoustic density estimation of sperm whales in the Tongue of the Ocean, Bahamas," *Mar. Mammal Sci.* **28**, 444–455.
- Weirathmueller, M. J., Wilcock, W. S. D., and Soule, D. C. (2013). "Source levels of fin whale 20 Hz pulses measured in the Northeast Pacific Ocean," *J. Acoust. Soc. Am.* **133**(2), 741–749.